Group Formation by Group Joining and Opinion Updates via Multi-Agent Online Gradient Ascent

Chuang-Chieh Lin, *Tamkang University, Taiwan*,
Chih-Chieh Hung, *National Chung Hsing University, Taiwan*[*]
Chi-Jen Lu, *Academia Sinica, Taiwan*
Po-An Chen, *National Yang Ming Chiao Tung University, Taiwan*

***Abstract*—This article aims to exemplify best-response dynamics and multi-agent online learning by group formation. This extended abstract provides a summary of the full paper in IEEE Computational Intelligence Magazine on the special issue *AI-eXplained* (AI-X). The full paper includes interactive components to facilitate interested readers to grasp the idea of pure-strategy Nash equilibria and how the system of strategic agents converges to a stable state by the decentralized online gradient ascent with and without regularization.**

I. Introduction

Game theory has been applied in a variety of situations due to its predictability of outcomes in real world. It can also be used in solving problems, such as saddle-point optimization which has been used extensively in generative adversarial network models [1]. A *game* consists of strategic agents, each of which acts rationally to maximize its own reward (or utility) or minimize its cost. A *Nash equilibrium* is a stable state composed of the strategies of all agents such that none of the agents wants to change its own strategy unilaterally. Therefore, such a stable state is possibly achievable or even predictable. However, how to achieve a Nash equilibrium in a game may not be quite straightforward, especially when agents behave in a "decentralized" way. Indeed, when an agent's reward function depends on the strategies of the other agents, the maximizer of one agent's reward function is not necessarily a maximizer for any other agent, and it may change whenever any other agent changes its strategy.

In this article, we consider the group formation of strategic agents as an example to demonstrate the strategic behaviors of the agents. A strategic agent can either join a group or change its opinion to maximize its reward. The eventual equilibrium of the game hopefully suggests predictable outcomes for the whole society. For the case in which agents apply group-joining strategies, we consider *pure-strategy Nash equilibria* (PNE) as the solution concept, where a pure strategy means a strategy played with a probability of 1. For the case in which agents change their opinions, we assume that each agent plays an *online gradient ascent algorithm*, which guarantees the time-average convergence to a hindsight optimum for a single agent (see [2] for the cost-minimization case), in a decentralized way, and then we examine the possibly convergent state of the system.

[*]Corresponding Author: Chih-Chieh Hung (smalloshin@nchu.edu.tw).

II. Group and Opinion Formation

Given a set V of n agents $v_1, v_2, \ldots, v_n$. Each agent v_i is represented as a *public preference vector* z_i and a *private preference vector* s_i, such that the former (we call it an *opinion*) corresponds to the preference revealed to all the agents while the latter corresponds to its *belief*, which is unchangeable. We consider $s_i, z_i \in \mathcal{K}$ such that $\mathcal{K} := \{x \in [-1,1]^k : \|x\|_2 \leq 1\} \subset \mathbb{R}^k$ is the feasible set. One can realize that each dimension of the domain stands for a certain social issue, such that -1 maps to the far-left politics, while 1 maps to far-right politics. The bounded 2-norm constraint is in line with the bounded rationality of a person, or the bounded budget for a group. We use $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ and $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ to denote two profiles that include each agent's opinion and belief, respectively. Each agent is initially regarded as a group. The *opinion of a group* is the average of the opinions of its members. Similar to the *monotone* setting in [3], a group wins with higher odds if its opinion brings more utility to all the agents. The *reward* (i.e., payoff) of an agent is the expected utility that it can get from all the groups. Specifically, assume that we currently have $m \leq n$ groups $G_1, G_2, \ldots, G_m$, and denote by $|G_i| = n_i$ the number of members in group G_i. Let $\mathcal{G} = (G_1, G_2, \ldots, G_m)$ denote the profile of the groups. To ease the notation, we denote by $\tau = (\mathbf{z}, \mathbf{s}, \mathcal{G})$ the *state* of the game. The reward function of agent i is $r_i(\tau) = \sum_{j=1}^{m} p_j(\tau)\langle s_i, \bar{g}_j\rangle$, where $\bar{g}_j = \sum_{v \in G_j} z_i/|n_j|$ represents the average opinion of group G_j and the winning probability $p_j(\tau)$ of group G_j is

$$p_j(\tau) = \frac{e^{n_j \langle \bar{g}_j, \sum_{v_r \in V} s_r\rangle}}{\sum_{i \in [m]; n_i > 0} e^{n_i \langle \bar{g}_i, \sum_{v_r \in V} s_r\rangle}},$$

where $[m]$ denotes $\{1, 2, \ldots, m\}$. We consider the following strategic behaviors of an agent in such a game:

- Group Joining:
  - Seeking a specific group that hopefully maximizes the agent's reward and joins the group.
- Opinion Updating without Regularization:
  - Each agent in a certain group tries to maximize its reward by changing its own opinion.
- Opinion Updating with Regularization:
  - Each agent in a certain group tries to maximize its reward by changing its own opinion, while the reward includes the regularization $-\|s_i - z_i\|_2^2$, which

hopefully limits how strategic an agent can be by preventing it from moving too far from its own belief.

III. Group Joining and Pure-Strategy Nash Equilibria

A. 1D Representation

When the opinions and beliefs are assumed to be in $[-1, 1] \subset \mathbb{R}$, we can illustrate these vectors as well as the dynamics of changes on a real line. For example, in Fig. 1 we have five agents v_1, v_2, v_3, v_4, v_5. By assuming v_1, v_2, v_3, and v_4 to have their opinions z_1, z_2, z_3, z_4 fixed, we can observe the changes in the winning probabilities and the rewards of all the agents by moving z_5 from -1 to 1.

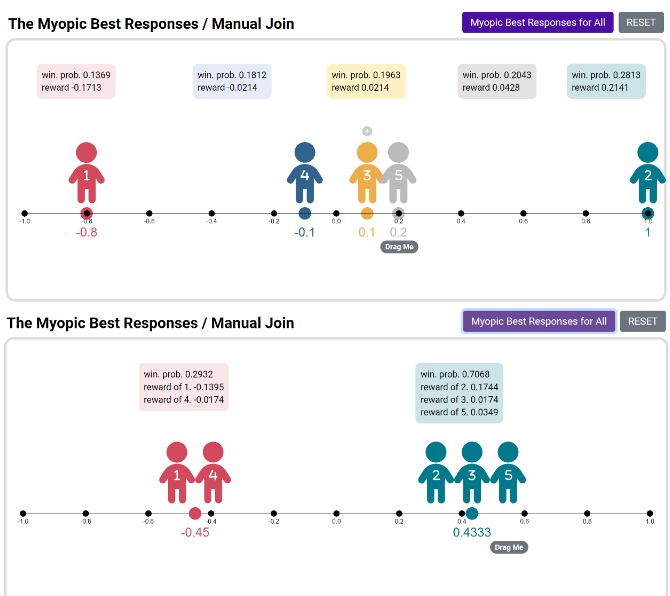

Fig. 1. 1D Representation: myopic best-responses and a PNE.

B. Group Joining: Myopic Best Responses

We assume that agent i decides to join G_j, for which $j = \arg\max_\ell p_\ell(\tau) \cdot \langle \bar{g}_\ell, s_i \rangle$. We call this strategy a *myopic best response*. An agent joins a group by considering not only its winning probability but also the utility that the agent can get from the group before joining. The state in the bottom of Fig. 1 is an example of a PNE.

IV. Opinion Updates by Online Learning

A. 2D Representation

We illustrate the opinions and beliefs as well as the dynamics of opinion changes in $\mathcal{K} := \{x \in [-1, 1]^2 : \|x\|_2 \leq 1\} \subset \mathbb{R}^2$. The 2-norm constraint that $\|z_i\|_2, \|s_i\|_2 \leq 1$ correlates the dimensions. A projection of the opinion is required if the constraint is not satisfied.

B. Online Gradient Ascent

We consider the setting that each agent tries to maximize its own reward by "changing its opinion" without deviating from the group to which it belongs. Each agent runs the online gradient ascent algorithm to iteratively update its opinion so as to maximize its reward. The update is done by adding a certain quantity (tuned by the learning rate η) toward the direction of the gradient. A "projection" $\Pi_{\mathcal{K}}(x)$ which projects x onto the feasible set $\mathcal{K}$ by dividing its 2-norm is performed if necessary.

Algorithm: Multi-Agent Online Gradient Ascent

Input: feasible set $\mathcal{K}$, T, learning rate η.

for $t \leftarrow 1$ to T **do**
  for each agent i **do**
    observe reward $r_i(\tau)$, where state $\tau = (\mathbf{z}, \mathbf{s}, \mathcal{G})$
    $z_{i,t+1} \leftarrow \Pi_{\mathcal{K}}(z_{i,t} + \eta \nabla_{z_i} r_i(\tau))$
  end for
end for

C. Online Gradient Ascent with Regularization

The reward function for agent i including the regularizer, is defined as $r_i(\tau) = \sum_{j=1}^{m} p_j(\tau)\langle s_i, \bar{g}_j \rangle - \|z_i - s_i\|_2^2$. Since $-\|z_i - s_i\|_2^2$ is always non-positive, an agent will be constrained to consider "not being too far from its belief." Our experimental illustrations show that such a regularization helps the game converge to a state where agents' opinions will not be too far from their beliefs (see Fig. 2).

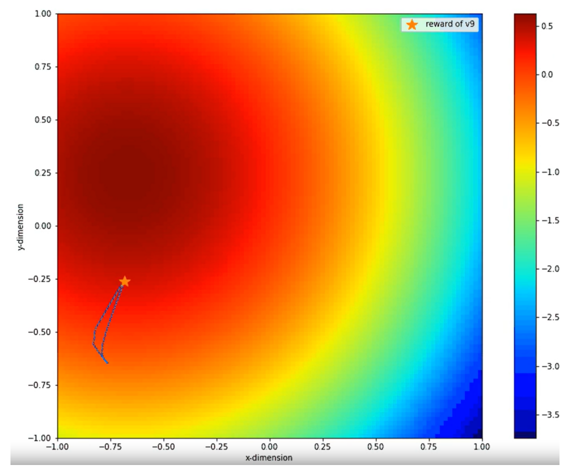

Fig. 2. Opinion updates via the online gradient ascent with regularization.

V. Conclusion

This article presents a preliminary study on the dynamics of group formation. From the illustrations, readers can realize what a pure-strategy Nash equilibrium in a system of multi-agents is and also learn how an online gradient ascent algorithm as one of the dynamics can reach a stable state.

Acknowledgments

This work is supported by the National Science and Technology Council under grant nos. NSTC 110-2222-E-032-002-MY2, NSTC 111-2221-E-005-047- and NSTC 111-2410-H-A49-022-MY2. We also thank Victorien Yen for helping implement the JavaScript codes in the immersive article.

References

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, pp. 139–144, 2020.

[2] F. Orabona, "A modern introduction to online learning," *arXiv:1912.13213*, 2022.

[3] C.-C. Lin, C.-J. Lu, and P.-A. Chen, "On the efficiency of an election game of two or more parties: How bad can it be?" *arXiv:2303.14405*, 2023.