Learning Payoff Functions in Infinite Games

Yevgeniy Vorobeychik, Michael P. Wellman, and **Satinder Singh**
University of Michigan
Artificial Intelligence Laboratory
1101 Beal Avenue
Ann Arbor, MI 48109-2110 USA
{ yvorobey, wellman, baveja }@umich.edu

Reviewer: Gyuseok Lee
Date: 09/19/2022

- Overview of paper content (0.5 points)

Motivation of this paper is how to deal with the situation with incomplete information and real-valued action. To solve this, the authors propose the pay-off function learning as a standard regression problem to low-degree polynomials. To measure the proposed method, they focused on relative utility of prescribed strategies not just the accuracy of payoff functions. For experimental setup, authors tested two games: one is a two-player version of the first price sealed-bid auction with well-known solution, and the other is a five-player market-based scheduling game with no known solution.

The problem definition of this paper is how to make the payoff function approximation by selecting a function $\hat{u}$. Their goal in approximating payoff functions is typically not predicting payoffs themselves, but rather in assessing strategic behavior. Specifically, the authors propose ϵ-Nash equilibrium of game, whose ϵ is used for approximation error of $\hat{u}$. The more detailed explanation is like below.

Definition 4 *A strategy profile* $\sigma = (\sigma_1, \dots, \sigma_m)$ *constitutes an* ϵ-Nash equilibrium *of game* $[I, \{\Delta(S_i)\}, \{u_i(s)\}]$ *if for every* $i \in I$, $\sigma_i' \in \Delta(S_i)$, $u_i(\sigma_i, \sigma_{-i}) + \epsilon \geq u_i(\sigma_i', \sigma_{-i})$. (1)

$$\epsilon = \max_{i \in I} \left[u_i(s_i^*(\hat{\sigma}_{-i}), \hat{\sigma}_{-i}) - u_i(\hat{\sigma}_i, \hat{\sigma}_{-i}) \right] \quad (2)$$

Where s_i^* denotes i's best correspondence, defined by $s_i^*(\sigma_{-i}) = \{x : x \in armax_{s_i} u_i(s_i, \sigma_{-i})\}$

- Strengths (0.5 points)

The first thing proposed as the quadratic form has an advantage to analytically solve for Nash equilibrium.
Plus, their result in both experiments show that when data is sparse, such methods can provide better approximations of their underlying game at least in terms of ϵ-Nash equilibrium of game than discrete approximation using the same data set.

- Weaknesses (0.5 points)

There is a limitation that the space of strategy becomes complicated.
I think that this method would be specific by domain or the setting of game, so I wonder that the approximated payoff function can be sub-optimal function we cannot search.

- Questions or Discussion (0.5 points)

I think that approximation for the payoff function would be related to AutoML (i.e., find best (hyper) parameter, model structure and so on), so I am curious between AutoML and this research.