Game Theory - Nash Equilibrium in El Farol Bar Problem

Hyuntaek Oh
ohhyun@oregonstate.edu
Due. Nov. 12, 2024

I. Introduction

In the field of game theory and multiagent systems, Nash Equilibrium stands as crucial concept, clarifying the conditions under which agents in a game can no longer unilaterally improve their outcomes through individual strategy changes. This concept not only makes our theoretical understanding better but also improves the design and analysis of algorithms for distributed systems where multiple decision-makers interact. Bar problem, which is a classical problem of complex interaction among agents in multiagent systems, where the dynamics of 50 agents choosing among 6 different nights to maximize their respective rewards were explored. This scenario implies an ideal testbed for investigating complex interactions and strategic decision-making in a controlled environment.

This paper addresses the computational exploration of Nash Equilibrium within the framework of three distinct reward systems: local, global, and difference rewards. Each system offers assessment of agents' behavior, with the setting involving 50 agents, a scaling factor b=4, and k=6 possible nights for attendance. The core challenge lies in identifying how these different reward systems influence the establishment of Nash Equilibrium, potentially guiding system-wide outcomes and individual agent strategies.

Our aim is to observe how various reward mechanisms affect the equilibrium states in a multiagent system and to evaluate the efficacy and implications of each approach. By simulating, the investigation provides practical implications for designing more effective multiagent systems.

II. Method

In this section, we investigate Nash Equilibrium within a multiagent bar attendance problem involving 50 agents, each choosing among 6 nights to attend, with the objective to maximize their respective rewards. We compare three reward systems: local, global, and difference rewards. Each reward structure is modeled using exponential functions that reflect reducing returns as more agents attend same days, with scaling factor b=4, which influence the reward based on the attendance. Each of the 50 agents must choose one night out of six to attend, with the goal of maximizing their reward under varying structures.

In the question (a), the local reward each agent uses to compute Nash Equilibrium is:

$$L(z) = x_k(z)e^{\frac{-x_k(z)}{b}},$$

where z is the system state and $x_k(z)$ is the number of agents attending the same night k. With the local reward, each agent focuses on maximizing their own attendance night's outcome regardless of the overall system state. This can potentially lead to suboptimal global outcomes if agents assemble in the bar certain nights, causing overcrowding and reducing returns due to the exponential penalty on higher attendance.

Similarly, in the question (b), the global reward each agent utilizes to calculate Nash Equilibrium is:

$$G(z) = \sum_{k=1}^{K} x_k(z)e^{\frac{-x_k(z)}{b}},$$

The global reward aggregates the outcome across all nights, providing a single reward value that all agents aim to maximize collectively. This setup encourages coordination among agents to distribute themselves more evenly across nights. It can result in a more balanced system where no single night is overly preferred, as agents' rewards depend on the total system performance.

Lastly, in the question (c), the difference reward employed by agent is:

$$D_i(z) = x_i(z)e^{\frac{-x_i(z)}{b}} - (x_i(z)-1)e^{\frac{-(x_i(z)-1)}{b}}$$

This reward structure is designed to align the individual agent's incentive with the global optimum by making them aware of the impact of their action on the system. It can prevent scenarios where agents might otherwise crowd a single night based on incomplete information about overall system states.

For setting simulation parameters, agents are initially assigned to nights randomly. Each simulation runs for up to 5,000 iterations or until no agent changes their night choice, indicating a potential Nash Equilibrium. Agents consider switching nights if the expected reward for a different night is higher, based on current attendance.

III. Experiments and Results

We conducted a series of simulations using a discrete-event simulation model built in Python. The environment consisted of 50 agents, each independently deciding which of six nights to attend based on three different reward structures. Each simulation was initialized with agents randomly distributed across the nights. The main goal was to observe how each reward structure influenced the agents' final distribution and the speed at which the system reached Nash Equilibrium. Each simulation was allowed to run until no changes were observed in the agents' night choices. Given randomly distributed attendance, the algorithm calculates current states and three different rewards respectively at first step and checks the expected rewards which is better than other nights by changing one agent's choice for any other nights iteratively.

A. Local Reward

We conducted an experiment finding Nash Equilibrium by using local reward equation with randomly distributed attendance, which is $[7, 4, 9, 10, 11, 9]$. The result is below.

Night	Number of Agents Attending	Rewards per night
Night 1	9	0.948593
Night 2	8	1.082682
Night 3	8	1.082682
Night 4	9	0.948593
Night 5	8	1.082682
Night 6	8	1.082682

TABLE I: Rewards based on Local Rewards and Agent Attendance by Night

From the table above, the simulation with the local reward concluded with agents fairly evenly distributed across most nights, except for minor imbalances. In particular, nights had attendance from 8 to 9 agents, with corresponding rewards between 0.94853 and 1.082682, meaning that local reward leads to individually rational decisions that stabilize around a balanced distribution.

In the local reward setting, agents aim to maximize their own night's reward regardless of overall system performance. Nash Equilibrium is reached when no single agent can increase their reward by unilaterally switching nights. Given the similar local rewards obtained across nights, it implies that agents found no benefit in switching nights beyond minor fluctuations. This behavior suggests that a state where agents' decisions are stabilized, thus meeting the Nash Equilibrium condition where the expected reward for moving is not greater than the reward for staying.

Equilibrium is reached since each agent's decision to switch nights would not yield a significantly higher reward. In other words, due to the balanced attendance across nights, each agent decided to stay current state after many times of switching nights and computing expected rewards. Thus, each agent's current choice is as good as any other in terms of individual reward.

B. Global Reward

We also conducted a similar experiment using Global reward to find Nash Equilibrium with the randomly initialized array $[9, 8, 12, 10, 8, 3]$. The result is below.

As can be seen, the Table 2 displays that Global Reward structure resulted in an extreme preference for one night (Night 3 with 30 agents), while other nights had optimal attendance. In the table, the sum of each reward by night is

Night	Number of Agents Attending	Rewards per night
Night 1	4	1.471517
Night 2	4	1.471517
Night 3	30	0.016592
Night 4	4	1.471517
Night 5	4	1.471517
Night 6	4	1.471517

TABLE II: Rewards based on Global Reward and Agent Attendance by Night

7.374181, which is greater than the sum of each reward using local rewards, 6.227914, meaning that this Nash Equilibrium is different from one using local rewards. In other words, Nash Equilibrium with Global Reward structure minimizes the reward by one night (Night 3) and maximizes rewards by other nights.

Despite the imbalance, the global reward calculation method promotes a collective maximization of rewards, which is reflected in a consistent global reward across all nights. In this setup, agents collectively benefit more by heavily attending one night, thereby maximizing their exponential reward collectively, even if it means overcrowding one night.

In this case, Nash Equilibrium is reached since the distributed reward from attending any night becomes equivalent. Even though the system state is imbalanced, no agent gains more by changing nights.

C. Difference Reward

Similarly, an experiment is conducted with the randomly initialized array $[4, 9, 10, 9, 11, 7]$. The result is below.

Night	Number of Agents Attending	Rewards per night
Night 1	4	0.054418
Night 2	4	0.054418
Night 3	30	-0.004003
Night 4	4	0.054418
Night 5	4	0.054418
Night 6	4	0.054418

TABLE III: Rewards based on Difference Rewards and Agent Attendance by Night

As shown in Table 3, the distribution was nearly uniform with 4 agents per night, except for one night where 30 agents attended on Night 3. This Nash Equilibrium is similar to the case using global reward, not local reward, with different rewards per night.

The difference reward structure is designed to minimize the negative impact on an agent might have by joining an already crowded night. Despite one outlier, the almost uniform distribution among other nights suggests that agents were effectively discouraged from overcrowding any single night.

Nash Equilibrium in the Difference Reward structure reflects that agents are individually optimizing not only their own reward but also considering the cost of their presence to the group's overall reward. Agents find no better alternative that would provide a higher reward than the current distribution, except one outlier.

IV. Conclusion

We investigated the behavior of 50 agents deciding among 6 nights to maximize their rewards under three distinct reward systems-local, global, and difference rewards-within a Nash Equilibrium framework. Each structure's design profoundly influenced agents' distribution strategies and the system's dynamics, causing to different results and efficiencies.

These results from three different reward systems underscore the impact of reward structure design on agent behavior and system outcomes in multiagent environments. The local reward structure, while simple and promoting individual rationality, does not ensure optimal global outcomes. The global reward structure effectively aligns agent actions towards collective goals but may lead to inefficiencies such as overcrowding. Meanwhile, the difference reward structure provides a promising approach by closely aligning individual actions with global system efficiency.