Modified El Farol: Simulating Repeated Games

Mike Mangione
Boston University
mjbm@bu.edu

Abstract

The El Farol game has been previously modeled as a simulation using update functions for adjusting player strategy. Similar papers have suggested that increased variability in player methodologies would enhance model realism and increase true convergence on maximal payoff across players. This paper proposes a simulation for the El Farol game within a Q-Learning framework, along with player communication and multiple sub-games. With increased versatility and modularity, this framework provides a robust solution for exploring repeated, collaborative games with a high-degree of variability in input. This model produces a realistic deliverable for limited resource allocation across semi-collaborative parties, and its structure lends itself to adaptation for other problem spaces.

I. Introduction:

Game theory has been responsible for countless developments across a number of domains. Political scientists use it to predict the motives and actions of governments, economist deploy it to analyze markets, and neuroscientists have modeled brain activity patterns with it. [Sanfey] While a number of game theory models can be used to describe social interactions, there are a subset of problems that are directly aimed at understanding social dynamics and semi-collaborative behavior. The Prisoner's Dilemma is perhaps the quintessential example, as it is a two player general game which explicitly rewards collaborative choices. While two player collaborative games are frequently explored, they capture only a small subsection of collaborative games. For this reason, the N-player El Farol game presents an interesting alternative, as it allows greater flexibility in the evolution of player strategy and can incorporate advanced communication elements between player nodes.

II. The El Farol Game:

i. Structure:

The El Farol game assumes a fixed population of players, and gives them two choices:

- Go to the El Farol Bar
- Stay home

If less than sixty percent of the total player population chooses to attend the bar, then they get a positive payoff for that game. Conversely, an attendance of more than sixty percent results in negative payoff for the bar-goers. There is no payoff value for the players that choose to stay home. As such, a sample payoff matrix for one player is provided below:

$$\begin{bmatrix} > 60\% : \text{at bar} & \text{at home} \\ < 60\% : \text{at bar} & \text{at home} \end{bmatrix} = \begin{bmatrix} a & 0 \\ -a & 0 \end{bmatrix}$$

This matrix, while a helpful aid for considering the game, is a simplification of the problem space. This assumes that the population of the bar is a singular adversary, when in reality it is a collection of multiple players facing the same potential payoff. Therefore, the true payoff matrix for the game would be an N-dimensional matrix with the choices represented for all N

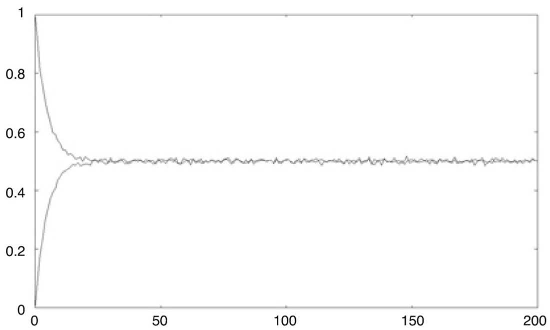


Figure 1: *A 10,000 player simulation with capacity of 50% in the Szilagyi Model*

players in the fixed population.

As such, a deterministic pure strategy for all N-players is not desirable, as it would produce one of two sub-optimal outcomes:

- All players attend the bar, and all receive negative payoff
- All players attend the bar, and all receive no payoff

A deterministic mixed-strategy, while allowing for occasional positive payoffs, will also be insufficient. For example, if each player attends the bar with sixty percent probability, then more than half of the time all bar players will see negative payoff.

ii. Simulations:

Simulations for El Farol have been modeled that allow for individual players to create heuristic mixed strategies for attending the bar. Szilagyi's Pavlovian Model, a time-series model that adjusts relative to previous successes and failures, presents a convergence on the capacity of the bar with fluctuation amplitudes roughly equal to twenty percent of the population size. With machine learning frameworks, the iterations required for convergence are reduced. The sample below illustrates the speed and accuracy in convergence for this model.

iii. Limitations:

As the El Farol game exists, it is a simple game with little similarity to the real-world problem it seeks to address. Although the simple framework has been modeled with great repetition, few papers have implemented the problem for extended frameworks. A number of papers reference possible extensions to increase simulation realism, a number of which introduce the capacity for machine learning frameworks.

First, players might be given unique models or initializations that enable them to make decisions with distinct processes. Previous frameworks relied on the players utilizing the same frameworks from limited strategy instantiations, so we would expect the speed with which they converge to be due in some part to their similarity in choice. We would expect a different convergence guarantee if players utilize unique models.

Second, the basic El Farol framework allows for no communication between player nodes. For a model of social interaction, it should be essential that some form of communication exist. A better tailored framework should allow subsets of the population to communicate, as they would in reality.

Third, the supposition that all players choose whether to go to the bar at the same point is improbable, and it does not translate well to the modeling of any real system. Even live two person games that require simultaneous decisions do not occur in the same instance, and thus introduce the possibility for time-based advantages.

Fourth, a single payoff matrix of N degree for N players will not encapsulate the entire set of decisions a potential bar patron would consider when choosing to leave for a bar. It will be essential, for a realistic model, to add additional sub-games that more appropriately model how and why players choose to go to the bar.

III. Modified El Farol:

i. Motivation:

It would be essential, to properly model the social games and motivations involved in attending a bar, to introduce additional factors for consideration to the El Farol problem. These additions may also help model more complex problems in more complex domains, like resource allocation and the modeling of financial markets.

ii. Structure:

When considering the structure of a physical bar, in many ways it can be considered a series of lines and chokepoints (or, for our purposes, sub-games.). As such, a realistic model should be interested in the flow controls imposed by the bar structure. The Q-Learning machine learning technique, in which nodes flow through a series of weighted links between locations to determine optimal pathways, is well suited to this problem. The vertices of the Q-Learning locations can inherently be modeled as games, where nodes use individual strategies to follow links of their choosing to distinct locational nodes. The meta-structures of the nodes (Locations and Players), iteration structures and player strategies are explained in the following sub-sections.

ii.1 Players:

As previously stated, the modified model uses Player object nodes with a number of shared properties between other Player nodes and the location nodes.

Definition 1 *All players are instantiated with the following characteristic properties, in the form of percentages relative to the total number of players:*

- *Optimism: The threshold that will prevent the player from entering the queue to the bar (e.g. an optimism value of 60% for N = 100 will prevent the player from entering a line with 60 people in it)*
- *Patience: The threshold for line length that the player will wait through to enter the bar (e.g. a patience level of 60% for N = 100 will cause the player to leave a 60 person line)*
- *Extroversion: The threshold for bar attendance that the player will tolerate (e.g. an extroversion level of 60% for N = 100 will cause the player to leave a bar with 60 patrons)*

We define variables p_o, p_p, p_e that map to optimism, patience, and extroversion traits respectively for all players p in superset P.

Definition 2 *The payoff functions from each sub-game for each player p are mapped to two distinct variables:*

- *Happiness: The summary payoff accrued through iterations in a single day*
- *Cumulative Happiness: The summary payoff accrued through iterations across all days*

For all players p in superset P, happiness maps to p_h and cumulative happiness maps to p_c

ii.2 Time Scale:

Definition 3 *Players update strategies and engage in sub-games with respect to two iterative variables:*

- *Iterations: Determines the location of players, their ability to participate in sub-games, and strategy updates*
- *Days: Comprised of a fixed number of Iterations*

For a single simulation, each day d in superset of days D will iterate through each iteration i in superset of iterations I.

ii.3 Layout:

Definition 4 *The layout of the modified El Farol problem is comprised of locations nodes S, Q, and B which map to Street, Queue, and Bar locations.*
We define function γ for location nodes which returns the number of players currently in that location.
We define value B_c as the capacity of the Bar node, which is a fixed number corresponding to the number of players the Bar can accommodate.

- *As such, when the Bar is at capacity, $\gamma(B) = B_c$*

The location of player p is given by value $p_L \in [S, Q, B]$

ii.4 Trait Updating:

Definition 5 *The function λ, for any characteristic of all players p in superset P of all players, updates the characteristic traits of the player with an a priori model. This function is currently called when a player loses a sub-game* *and when the final iteration of a day is completed. This function and its implementation is discussed at length in the Analysis section.*

iii. Simulation Pseudocode:

Given the above definitions, the pseudocode in Algorithm 1 represents the modified simulation.

iii.1 Player Network:

The implementation of player communication has been frequently discussed for the El Farol framework, and this modified structure utilizes communication by broadcasting player node values. For each player, at the start of the simulation, each player instantiates a node network of which they are the center. Each player connects to the same number of other players, and the players they connect to are selected randomly. These networks contribute the trait updating function, in that they allow the player to compare their successes in the simulation with a fixed population of nodes. The implemented update function λ is replicated below for some trait x for some player p with player network Z and step value V:

$$n = z \in Z \mid z_h = max(Z_h)$$

$$\lambda(p_x) = p_x + (p_x - n_x)/V$$

As such, the player will reach the character trait of the leader of its network after V updates. This model may be modified to use cumulative happiness by seeking the cumulative happiness leader in the network instead of the happiness leader.

Algorithm 1 Modified El Farol Simulation

```
∀p ∈ P of length N, p_{q_i} ∈ I
and ∃v ∈ Q | p_{q_l} = γ(v)
d, i = 0, 0
for d ∈ D do
    for i ∈ I do
        for p ∈ P do
            if p_L = S then
                if p_o ∗ N > γ(Q) then
                    p_L = Q
                    p_{q_i} = i
                else
                    p_o = λ(p_o)
                end if
            else if p_L = Q then
                p_h = p_h − p_p
                if i − p_{q_i} > p_{q_l} then
                    p_L = B
                    p_{q_i}, p_{q_l} = 0, 0
                else if i − p_{q_i} > p_p ∗ N or γ(B) = B_c then
                    p_L = S
                    p_{q_i}, p_{q_l} = 0, 0
                    p_p = λ(p_p)
                end if
            else if p_L = B then
                p_h = p_h + p_e
                if γ(B) > p_e ∗ N then
                    p_L = S
                    p_e = λ(p_e)
                end if
            end if
        end for
    end for
    for p ∈ P do
        p_o, p_p, p_e = λ(p_o), λ(p_p), λ(p_e)
        p_c = p_c + p_h
        p_h = 0
    end for
end for
```

iv. Experimentation:

iv.1 Implementation:

This framework was coded in Python, along with a simple GUI. The following values are accessible via the GUI:

- Cumulative Happiness Model: Boolean value to selects p_h or p_c for update function
- Population Size: N number of players for the simulation

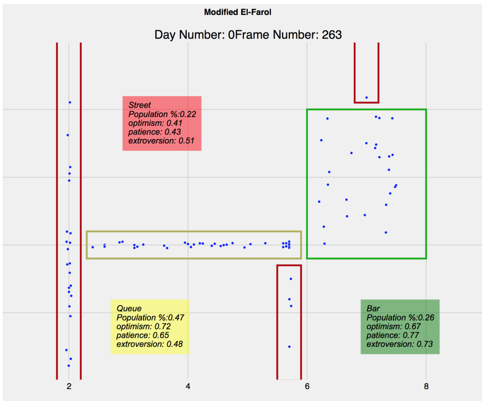


Figure 2: *Animation Frame from 100 player Modified El Farol Simulation*

- Percent capacity of bar: Percentage X such that $B_c = N * X$
- Number of connections: Number of player nodes in each player's network
- Number of Days
- Number of Iterations per Day
- Data collection rate: Number R such that: data collection rate = # of Iterations / R
- Animation: Boolean value for showing animation of simulation (See Fig. 2)

Once the input parameters are specified, the simulation and data collection are executed in parallel. As stated above, data is collected at a rate relative to the number of iterations. Specifically, at each data collection interval, the following array element is added to a Numpy matrix for each player:

$$[\text{iteration number}, p_o, p_p, p_e, p_h, \text{day number}]$$

After the simulation is completed, the array is parsed and formatted as a .csv file, which is subsequently read and output into a series of matplotlib graphical elements.

iv.2 Data:

The goal of this new framework was to create an environment that allows for high variation in input and a corresponding high variation in output. However, heuristically sampling different inputs has produced some general trends for the model, so long as certain conditions and assumptions are provided for. As such, we have aggregated representative simulations for input variable combinations of interest.

iv.3 Analysis:

As proof of concept, consider a simulation of N = 100, capacity = 10, and network size = 10. Given the parameters as described above, one might be able to predict various simulation outcomes. First, it is likely that player extroversion will be maximized over multi-day models, as it is a strategy that incurs no negative cost to the player and is essential for increasing the likelihood of staying in the Bar. Second, optimism will likely not converge quickly due to the neutral payoff for players not attempting to enter the bar. Third, patience might reach a low convergence point, due to the negative payoff associated with it and the reduced queue length due to increased bar attendance.
Figure 3 illustrates the trait convergence patterns for Non-Cumulative and Cumulative Happiness models after the first and final days in a ten day sequence.
Both happiness models share some commonalities in output. Players who are initialized with high patience and extroversion traits generally see high payoff values in the first day. Furthermore, the lowest payoff values on the first day patience graph follow a negative linear slope, indicating a high correlation between patience value and lack of happiness. This is especially true for the Non-Cumulative model, but the pattern somewhat persists in the Cumulative model as well. Understandably, little convergence is seen in the first day, but it is clear that certain strategies are optimal even in this short span.
By the tenth day, however, the convergence patterns are clear. Extroversion is almost universally at 1, with the only deviations found at the lowest happiness-level players. However, between the Cumulative and Non-Cumulative models, there are substantial differences in how the models converge. The Non-Cumulative Optimism spans the range between 0.6 and 0.8, with a tight patience value of 0.8. The Cumulative patience values fluctuate slightly more, while its optimism values are slightly tighter. It is worth clarifying that for all four graphs, the happiness graphed is non-cumulative, because only the non-cumulative model shows players

with net positive non-cumulative happiness in the final day. This is likely due to the variation allowed by the non-cumulative model, which itself functions more like a deterministic mixed strategy model.

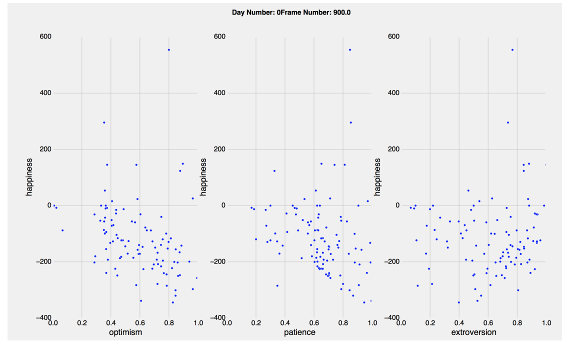
IV. Conclusion:

i. Results:

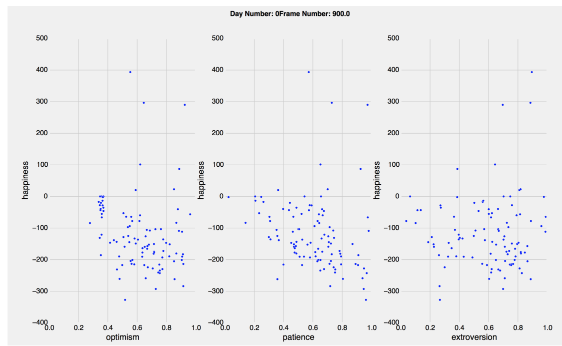
The basic El Farol game works well to illustrate the inherent challenge of collaborative, repeated general games. This modified frameworks takes necessary steps, as outlined in academia, to push the model into the territory of real-time meta-strategy simulation. By introducing chronological elements, communication, and trait updating functions, the player behavior is more dynamic and perhaps more akin to the reality of the bar problem.

ii. Future Extensions:

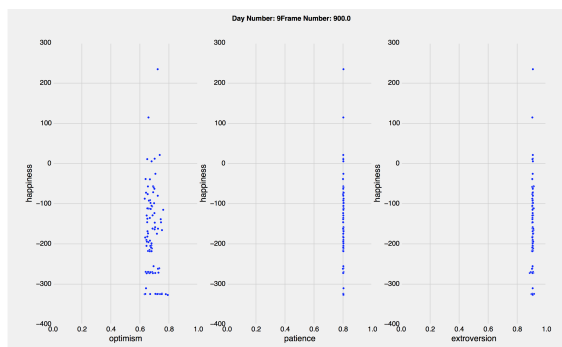
Obviously, the simple update function implemented is not the only machine learning function possible, and one of the benefits of this system's modularity is that it encourages experimentation with these functions. For future developments, it would be worthwhile to explore better tailored function to the framework, including ones that incorporate the historical actions taken by players and their sequential happiness over rounds. In terms of increasing the problem complexity, both the Kolkata Bar problem and increased variability in player strategies should be considered.
Furthermore, the data analytic framework can be better improved an automated. While simple scatter plots are effective for making visual approximations of strategy densities, there are likely better functions for mapping these outputs.
Lastly, the large number of parameters might allow for interesting meta-simulations. For example, given some fixed inputs, we could allow a machine learning model to determine optimal values for the other variables (e.g. number of days/iterations, network size). At the very least, greater heuristic experimentation should produce interesting insights for this framework.



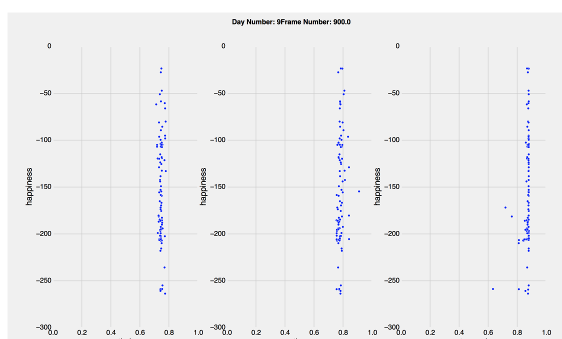
(a) Non-Cumulative Happiness, Day 0



(b) Cumulative Happiness, Day 0



(c) Non-Cumulative Happiness, Day 9



(d) Cumulative Happiness, Day 9

Figure 3: *N = 100, 900 iterations*

If you are interested in working with this simulation, simply navigate to the modified-el-farol/code folder and execute the following:
$ chmod +x sample_animation.py
$ ./sample_animation.py
Simply answer the GUI's prompts, and view the output simulation for yourself!

V. Citations:

- Chen, Shu-Heng, and Umberto Gostoli. "Coordination in the El Farol Bar problem: The role of social preferences and social networks." 2012 IEEE Congress on Evolutionary Computation (2012): n. pag. Web.
- "On the Complexity of the El Farol Bar Game: A Sensitivity Analysis." Agent-Directed Simulation Symposium (ADS 2016) (2016): n. pag. Web.
- Sanfey, A. G. "Social Decision-Making: Insights from Game Theory and Neuroscience." Science 318.5850 (2007): 598-602. Web.
- Sheikhha, Fatemeh, Hadi Tabatabaee Malazi, and Roya Amjadifard. "Adaptive Parasitized El Farol Bar Problem." 2009 WRI World Congress on Computer Science and Information Engineering (2009): n. pag. Web.
- "Talk:El Farol Bar problem." Wikiwand. N.p., n.d. Web. 08 May 2017.
- 5.3.1 The El Farol Bar Model. N.p., n.d. Web. 08 May 2017.