

# Payoff function for El Farol problem when using fictitious Markov play with incomplete information

Carlos E. Álvarez

August 29, 2024

El Farol bar problem consists in finding an agent-based strategy where  $N$  agents try to maximize the benefit they obtain by attending to El Farol Bar. The bar has a threshold of attendance  $T_A$ , and  $A$  is the number of agents that attended to the bar.

To attend to the bar has a cost of  $(-1)$  units for an agent. If at the end of a round  $A \leq T_A N$ , all attendees get a reward of  $(+2)$  units for that round, while if  $A > T_A N$  there is no reward for the attendees. An agent that does not attend simply gets no reward nor pays any cost.

The state  $S^{(t)}$  of the system of  $N$  agents at round  $t$  can be represented by a string of  $N$  binary digits

$$S^{(t)} = \underbrace{01101011 \cdots 10010}_N,$$

where 1 means that the agent in that position attended to the bar, while 0 means that it did not. Let's call  $s_i^{(t)}$  to the state of the  $i$ th agent at round  $t$ , that is, the  $i$ th position on the string  $S^{(t)}$ .

The payoff function  $Po$  measures the amount of units an agent  $i$  gets on a round, and it is a function of the state taken by the agent on the next round  $s_i^t$ , the assumed state of the system  $S^{(t)}$ , from which the attendance  $A$

is calculated, and the maximum number of allowed attendees

$$A_{mx} = T_A N. \quad (1)$$

Assuming that we have complete information about the attendance, the payoff function is

$$Po(s_i^{(t)}, S^{(t)}, A_{mx}) = \begin{cases} 1 & , \text{ if } s_i^{(t)} = 1, A^{(t)} \leq A_{mx} \\ -1 & , \text{ if } s_i^{(t)} = 1, A^{(t)} > A_{mx} \\ 0 & , \text{ if } s_i^{(t)} = 0 \end{cases} \quad (2)$$

## 1 Incomplete information

Equation (2) can be computed by an agent if it knows the complete state of the system  $S$ , needed to obtain the assistance  $A$ . What happens if the agent only has access to the state of a number of bits  $b < N$ ? In this case, the agent is only certain about its own state and the state of  $b - 1$  of its partners. Regarding the possible state of the unknown partners, an agent can only speculate. Following the principle of maximum entropy (references) the agents will make the most unbiased assumption about the state of the unknown bits, that is, take the distribution with the maximum entropy which, under no further constraints, is the uniform distribution.

The payoff function will depend then on the state that agent  $i$  plans to be in the next round  $s_i^{(t)}$ , the assumed state of the set of bits known to the agent  $i$  ( $K_i^{(t)}$ ), from which only a partial attendance can be obtained, and  $A_{mx}$ . As we assume a uniform distribution for the state of the remaining  $N - b$  bits, where  $b = |K_i|$ , the amount of units gained or lost by attending to the bar is replaced by an average gain over the unknown bits.

Let  $0 \leq k_1 \leq b$  represent the number of known bits in state 1, and  $0 \leq u_1 \leq N - b$  represent the number of unknown bits in state 1. Let us define first a pre-payoff function

$$Po_{pre}(s_i^{(t)}, K_i^{(t)}, A_{mx}) = \begin{cases} 1 & , \text{ if } s_i^{(t)} = 1, k_1 + u_1 \leq A_{mx} \\ -1 & , \text{ if } s_i^{(t)} = 1, k_1 + u_1 > A_{mx} \\ 0 & , \text{ if } s_i^{(t)} = 0 \end{cases} \quad (3)$$

and note that we must let the value of  $u_1$  to vary, due to the agent's lack of information. Instead of 1 and -1, we can then use an average value for the scores, over the distribution of the unknown bits. To do this we can count the number of possible states that respect the inequality in the upper row of equation (3). The number of states that respect  $k_1 + u_1 \leq A_{mx}$  is

$$n_r = \sum_{u_1=0}^{A_{mx}-k_1} \binom{N-b}{u_1}, \quad (4)$$

while the number of states that do not respect this inequality is

$$n_n = 2^{N-b} - n_r. \quad (5)$$

The probability of obtaining a payoff of 1, given that the agent attended, in (3) is given by

$$P[k_1 + u_1 \leq A_{mx} | s_i^{(t)} = 1] = \frac{n_r}{2^{N-b}}, \quad (6)$$

while the probability of obtaining a payoff of -1, given that the agent attended is

$$P[k_1 + u_1 > A_{mx} | s_i^{(t)} = 1] = \frac{n_n}{2^{N-b}}. \quad (7)$$

Then, using (6) and (7), it is possible to define a payoff function for particular values of  $A_{mx}$ ,  $N$ , and  $b$ .

As an example, assume  $N = 4$ ,  $b = 2$ , and  $T_A = 0.5$ , which means that  $A_{mx} = T_A N = 2$ , and say that the two bits the agent knows about are the 1st and 2nd bits, with values  $K_i^{(t)} = \text{01}$ , that is  $k_1 = 1$ . Then

$$\begin{aligned} P[k_1 + u_1 \leq A_{mx} | s_i^{(t)} = 1] &= \frac{1}{2^2} \sum_{u_1=0}^1 \binom{2}{u_1} \\ &= \frac{1}{4} \left( \underbrace{\frac{2!}{(2-0)!0!}}_{\text{0100}} + \underbrace{\frac{2!}{(2-1)!1!}}_{\text{0101, 0110}} \right) \\ &= \frac{3}{4}, \end{aligned}$$

while

$$P[k_1 + u_1 > A_{mx}|s_i^{(t)} = 1] = 1 - P[k_1 + u_1 \leq A_{mx}|s_i^{(t)} = 1] = \frac{1}{4}(\underbrace{1}_{0111}) = \frac{1}{4}.$$

Therefore, the average payoff expected for attending to the bar under these conditions is

$$\begin{aligned}\overline{Po} &= 1 \times P[k_1 + u_1 \leq A_{mx}|s_i^{(t)} = 1] - 1 \times P[k_1 + u_1 > A_{mx}|s_i^{(t)} = 1] \\ &= 2P[k_1 + u_1 \leq A_{mx}|s_i^{(t)} = 1] - 1 \\ &= \frac{1}{2}.\end{aligned}$$