ME 538
Homework 2: Agent Coordination
Due Nov 3, 2014 at 11:59 PM

Arthur's El Farol bar problem [2] is a perfect example of a congestion game. In this problem each agent $i$ decides whether to attend a bar by predicting, based on its previous experience, whether the bar will be too crowded to be "rewarding" at that time, as quantified by a system reward $G$. The congestion game structure means that if most agents think the attendance will be low (and therefore choose to attend), the attendance will actually be high, and vice-versa.

For this assignment, we use a modified version of the bar problem where the $N$ agents pick one out of $K$ nights to attend the bar every week [1, 3]. The system reward in any particular week is:

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

where $x_k(z)$ is the total attendance on night $k$; and $b$ is a real-valued parameter representing the optimal number of people in the bar. The system dynamics are as follows:

Initialize: week $\leftarrow 0$
Repeat till week $>$ Time Limit

- Each agent chooses an action.
- Those actions lead to a system state.
- The system state leads to a system reward.
- Each agent receives a reward (i.e., agent reward) .
- week $\leftarrow$ week + 1

**Learning Agents:** Each agent $i$ keeps a K-dimensional vector of its estimates of the reward it would receive for taking each of those K actions (action-value learning). Each week, each agent $i$ picks the night to attend based on sampling this vector.

**Problem 1:** Derive a simple "local" reward for each agent in this problem (e.g., the agent reward should depend on information easily available to the agent). Discuss the factoredness and learnability of the reward you derived.

**Problem 2:** Derive a difference reward for each agent. What is a good counterfactual "action" $c_i$ for this case? For at least two values of $c_i$, discuss the locality of the information and the factoredness and learnability of the resulting difference rewards.

**Problem 3:** Perform a simulation for this problem with the following parameters:

a) capacity of each night is 4 (b=4), k=6, and there are 30 agents in the system.
b) capacity of each night is 4 (b=4), k=5, and there are 50 agents in the system.

Plot the performance of three agent rewards (G, difference, and local) and a histogram of sample attendance profiles. Discuss the simulation results.

# References


[1] A. K. Agogino and K. Tumer, Analyzing and visualizing multiagent rewards in dynamic and stochastic environments, *Journal of Autonomous Agents and Multi Agent Systems*, 17:320–338, 2008.

[2] W. B. Arthur. Complexity in economic theory: Inductive reasoning and bounded rationality. *The American Economic Review*, 84(2):406–411, May 1994.

[3] D. H. Wolpert and K. Tumer. Optimal payoff functions for members of collectives. *Advances in Complex Systems*, 4(2/3):265–279, 2001.