CS 7642: Reinforcement Learning and Decision Making **Homework #5**

Bar Brawl

1 Problem

1.1 Description

You are the proprietor of an establishment that sells beverages of an unspecified, but delicious, nature. The establishment is frequented by a set P of patrons. One of the patrons is the instigator and another is the peacemaker.

On a given evening, a subset S $\subseteq$ P is present at the establishment. If the instigator is in S but the peacemaker is not in S, then a fight will break out. If the instigator is not in S or if the peacemaker is in S, then no fight will occur.

Your goal is to predict whether a fight will break out among a subset of the patrons without initially knowing the identity of the instigator or the peacemaker.

1.2 Procedure

Develop a KWIK learner for this problem. When presented with S, the patrons at the establishment, your learner will respond with "FIGHT", "NO FIGHT", or "I DON'T KNOW". Additionally, it should be capable of learning from the tuple of S and the true outcome for the evening. For each problem, the following input will be given:

- `atEstablishment : boolean[0...numberOfEpisodes][0...numberOfPatrons]`
- `fightOccured : boolean[0...numberOfEpisodes]`

Specifically:

- For each episode, you should present your learner with the next row of atEstablishment and the corresponding row of fightOccured.
- If your learner returns "FIGHT" or "NO FIGHT", you may continue on to next episode.
- If your learner returns "I DON'T KNOW", then you should present the pair (atEstablishment, fightOccured) to you learner to learn from.

You will submit a list of integers corresponding to the returned values of each episode.

1.3 Examples

The test case will be considered successful if no wrong answers are returned and the number of "I DON'T KNOW"s does not exceed the max allowed.

- `atEstablishment = [[True,True], [True,False], [False,True], [True,True], [False,False], [True,False], [True,True]]`
- `fightOccurred = [False, True, False, False, False, True, False]`
- Output: `0, -1, 0, 0, 0, 1, 0` (No incorrect answers and only one "I DON'T KNOW".)

KEY for output:
0 = NO FIGHT
1 = FIGHT
-1 = I DON'T KNOW

1.4 Resources

1.4.1 Readings

- Li-Littman-Walsh-2008.pdf Li et al. 2011

1.5 Submission Details

The due date is indicated on the Canvas page for this assignment.

Make sure you have set your timezone in Canvas to ensure the deadline is accurate. To complete the assignment calculate answers to the specific problems given and submit results to Canvas.

References

[Li+11] Lihong Li et al. "Knows what it knows: a framework for self-aware learning". In: *Machine learning* 82.3 (2011), pp. 399–443.