

Quiz 1

ASEN 6519-001: Advanced Survey of Sequential Decision Making

October 6, 2021

Question 1. Answer the following questions about ACAS-Xu:

- a) (0.5pt) In addition to reducing the probability of near mid-air collisions, what other objective is the ACAS-Xu system optimized for?
- b) (0.5pt) In general, modeling the horizontal and vertical positions and velocities of two aircraft requires a 12-dimensional state (3 position and 3 velocity variables for each aircraft), which would result in an intractable MDP. ACAS-Xu avoids this problem by splitting the problem into vertical and horizontal subproblems. However these subproblems have only 4-dimensional and 6-dimensional state spaces respectively, which does not add up to 12. What additional simplification is made to avoid keeping track of 12 variables between the two problems?
- c) (1pt) Using the concept of computational complexity classes, explain why using the QMDP approximation is beneficial in ACAS-Xu.

Question 2. (1pt) In actor-critic methods, how important is it for the absolute value returned by the Q-network to match the true average cumulative reward from a simulator started at that state? Justify your answer based on evidence from one of the papers read in class.

Question 3. (1pt) Describe the main way that soft actor critic differs from traditional actor-critic methods.

Question 4. (2pt) Would most computer scientists expect that it is possible to reduce QSAT to the Markov Decision Process¹? Why or why not?

Question 5. (2pt) Suppose you observe the value of 100 identically distributed independent random variables, X_i for $i \in \{1, \dots, 100\}$, and observe the sample standard deviation to be 9.7. In a separate experiment, you sample 10,000 identically distributed independent random variables, Y_i for $i \in \{1, \dots, 10,000\}$ and observe the sample standard deviation to be 430.2. If you estimate the mean of X_i and Y_i using this data, which estimate do you expect to be more accurate. Justify your answer.

Question 6. (2pt) Consider an undiscounted MDP with a state space of size $|S|$ and a finite horizon H . Based on the relative size of H and $|S|$, describe how you would decide between using sparse sampling and value iteration to solve the problem, and justify your answer with an argument based on the computational complexity of the algorithms.

¹"The Markov Decision Process refers to the abstract problem called the Markov Decision Process, i.e the set of all MDP instances. Reduction of QSAT to the MDP means that every instance of QSAT can be transformed in polynomial time to an MDP