You have until 5pm on November 16th to complete this take-home exam. You may use any resources available to you (e.g., book, notes, internet, software). Although you may discuss the general concepts presented in this class with others, you must not consult with others inside or outside of the class on these specific questions. Please submit your responses to Canvas.

Aircraft Collision Avoidance Problem

Suppose we have two aircraft, A and B, that are approaching each other head-on. We are trying to prevent the aircraft from colliding. At each step, we can make exactly one of the following three decisions:

1. a_{+1}: Tell A to go up and B to go down.
2. a_{-1}: Tell A to go down and B to go up.
3. a_0: Do nothing.

We will assume that if we tell A to do something, it will follow our instructions. However, if we tell B to do something, it may not respond. The state variables include:

1. $r \in \{0, 1\}$, which denotes whether B is responsive.
2. $h \in \{-10, -9, ..., 10\}$, which is the altitude of A relative to B.
3. $t \in \{10, 9, ..., 0\}$, which is the time until the two aircraft pass each other.

We will represent states as a vector $[r, h, t]^\intercal$. Assume h and t are measured exactly. The reward model can be additively decomposed into the sum of $R(s)$ and $R(a)$. For all states s with both $t = 0$ and $h = 0$ true, $R(s) = -1$; otherwise $R(s) = 0$. We also have $R(a_0) = 0$ and $R(a) = \lambda$ for all other actions.

The transition model is as follows. Regardless of action, r remains the same and t is decremented by 1 until reaching 0. Any action taken from a state with $t = 0$ immediately ends the episode. If we take action a_i, then $\dot{h}_A = i$. If responsive, then $\dot{h}_B = -i$ after taking a_i; otherwise $\dot{h}_B = 0$. Let $\hat{h} = \max\{\min\{h + \dot{h}_A - \dot{h}_B, 10\}, -10\}$. At the next time step, $h = \hat{h}$ with probability 0.5, $h = \min\{\hat{h} + 1, 10\}$ with probability 0.25, and $h = \max\{\hat{h} - 1, -10\}$ with probability 0.25.

1. (1pt) What is the size of the state space?
2. (1pt) What is the size of the observation space?
3. (1pt) What is the dimensionality of our belief state?
4. (1pt) Assume our initial belief is uniform over all states with $t = 10$. After the first observation, how many components of the belief vector will be non-zero?
5. (2pt) Suppose we have a belief b that assigns probability 1 to state $[1, 10, 1]^\intercal$; what is $Q^*(b, a_{+1})$ (assume $\lambda = -0.5$)? Provide an exact numerical value and explain.
6. (2pt) Suppose we have a belief b that assigns probability 1 to state $[0, 10, 1]^\intercal$; what is $U^*(b)$ (assume $\lambda \leq 0$)? Provide an exact numerical value and explain.
7. (2pt) Is it possible for $U^*([r, h, t]^\intercal) \neq U^*([r, -h, t]^\intercal)$ for some λ, r, h, and t? If so, provide an example. If not, provide a simple explanation.
8. (2pt) As $\lambda \to -\infty$, what is $\min_s U^*(s)$? Why?
9. (2pt) Suppose we have a belief b that assigns probability 1 to state $[0, 9, 0]^\intercal$. State an action that will maximize $Q^*(b, a)$ when $\lambda = 5$. Is it unique?
10. (3pt) Draw a two-step conditional plan from the state $[0, 1, 10]^\intercal$ where the action associated with the root node is a_0. Only show the observation branches that have a non-zero probability of occurring.

11. (1pt) If we are using the fast informed bound (FIB) to approximate the optimal value function, how many alpha vectors will there be?
12. (2pt) If α_{QMDP} is an alpha vector generated by QMDP and α_{FIB} is an alpha vector generated by FIB, can there exist a b such that $b^\intercal \alpha_{\text{QMDP}} < b^\intercal \alpha_{\text{FIB}}$? Why or why not?
13. (2pt) Suppose we have a belief state b that assigned probability 0.5 to $[0, 0, 1]^\intercal$ and probability 0.5 to $[1, 0, 1]^\intercal$. What is the value for $U^*(b)$ in terms of λ (which may take on *any* negative value)?
14. (1pt) Why would you not use a particle filter to update your belief for this problem?
15. (2pt) Suppose your initial belief is uniform over the state space and then you observe that aircraft A is 3 units above aircraft B after executing a_0. What probability would an exact Bayesian update of your belief state assign to aircraft B being non-responsive? Why?
16. (0 pt) Write a little paragraph about what you learned in this class.