Caelan Garrett
6.S078 Assignment 6.2
12/10/13

I implemented the Tiger Problem with restarts every time a door is open as described in class. I didn't directly hardcode the problem in the POMDP solver. Instead, I allowed the user (by modifying the initialization) to specify states, actions, observations, transition probabilities, observation probabilities, rewards, and the discount factor (set to 1 for the Tiger Problem). Therefore theoretically, the solver implementing Monahan's algorithm should work for arbitrary POMDPs, but I'm too scared/lazy to define another one. I measured the number of generated vectors and runtime for both the no pruning and pruning variants. Additionally, I measure the number of vectors after pruning and the corresponding percentage pruned. It's safe to say that pruning made a big difference, especially on Iteration 3. I also plot the alpha vectors at each iteration. Black horizontal lines are the y-axis markers of expected reward in intervals of 10, and the purple horizontal line is the x-axis at expected reward of 0. I colored the alpha vectors by the first move at the root of the policy tree. Red vectors indicate open left, green vectors indicate listen, and blue vectors indicate open right. I defined the x-axis such that p = 0 indicates the belief that the tiger is not left and p = 1 indicates the belief that the tiger is left.

No Pruning	Iteration 0	Iteration 1	Iteration 2	Iteration 3
Generated Vectors	3	27	2187	14348907
Runtime	0.000537872314453	0.00166988372803	0.0955879688263	213.929979086

| Pruning	Iteration 0	Iteration 1	Iteration 2	Iteration 3
Generated Vectors	3	27	75	147
After Pruning Vectors	3	5	7	5
Percentage Pruned	0.0	0.814814814815	0.906666666667	0.965986394558
Runtime	0.00774097442627	0.0714750289917	0.292954206467	0.803614854813
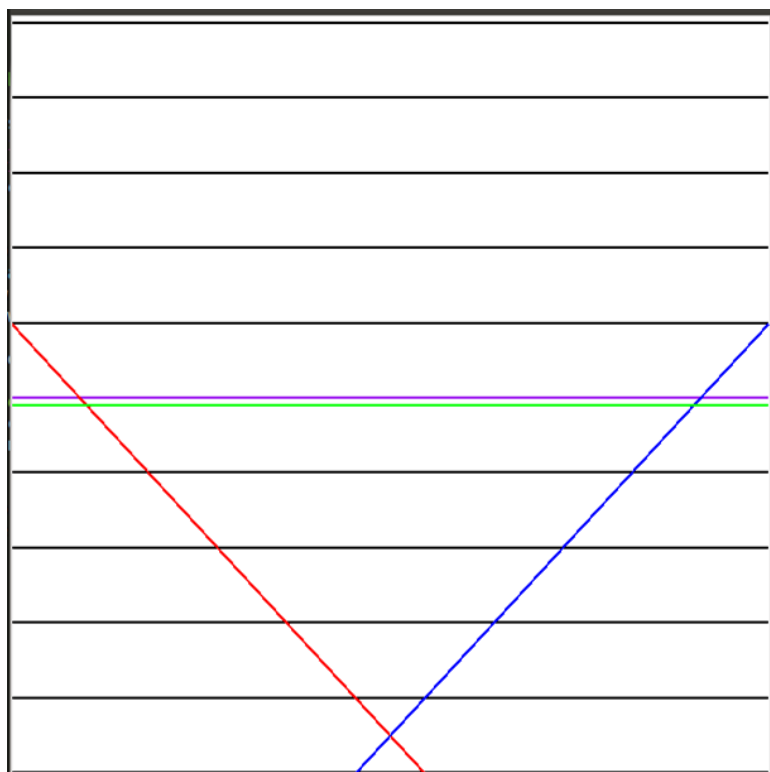
Alpha Vectors for Iteration 0

Opening a does better here than in future horizons because you only get one move. If you are really not confident, it still is safer to listen.
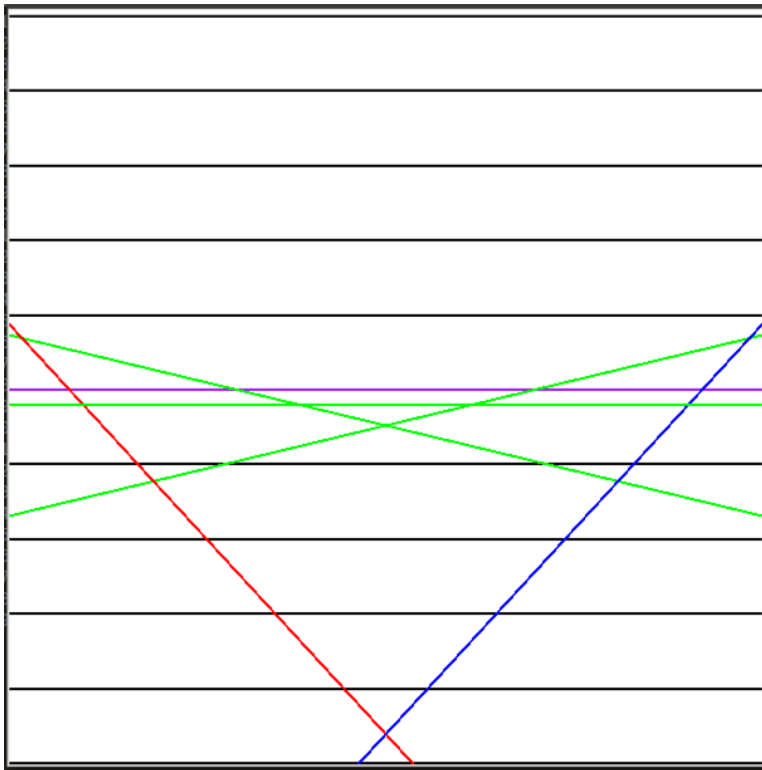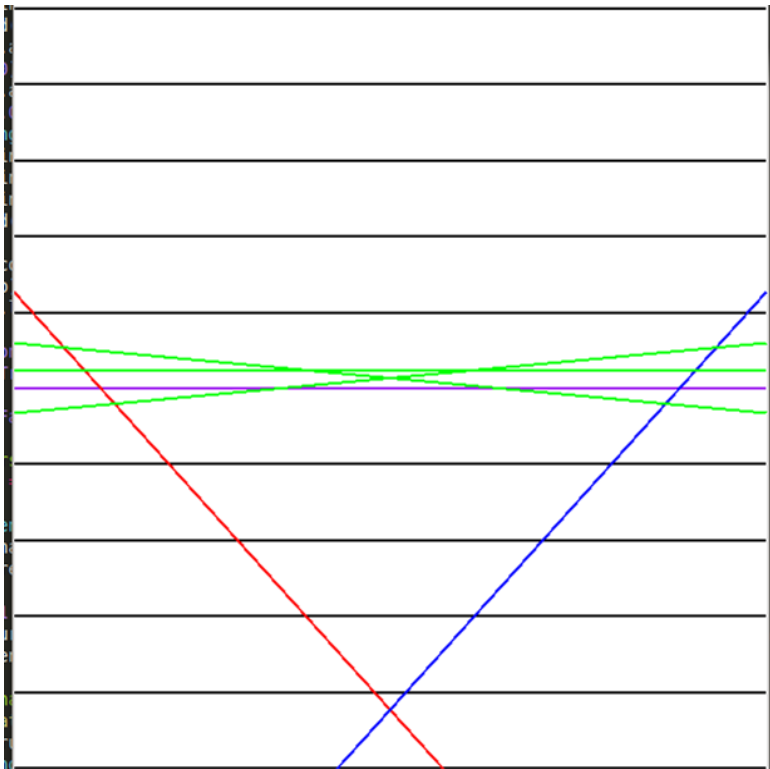

Alpha Vectors for Iteration 1

Unless you are extremely confident in your belief, it is much better to listen than open a door. In the middle, the policy tree tells you to listen twice because one observation won't make you confident enough in your belief.
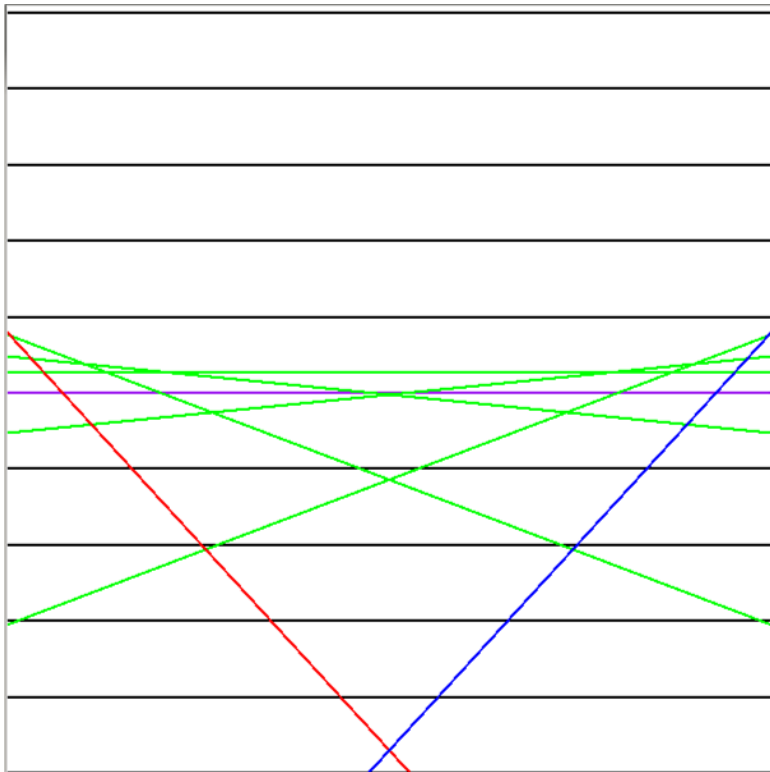
Alpha Vectors for Iteration 2

The main difference from the last iteration is that even if you start out indifferent to which door the tiger is behind, you can update your belief using two observations to become relatively confident and achieve positive expected value.

Alpha Vectors for Iteration 3

Here you have the ability to win the game twice to get a larger maximum expected reward. If you already are confident in your belief, immediately opening the door allows the game to reset so you can play the game again and achieve additional expected reward.