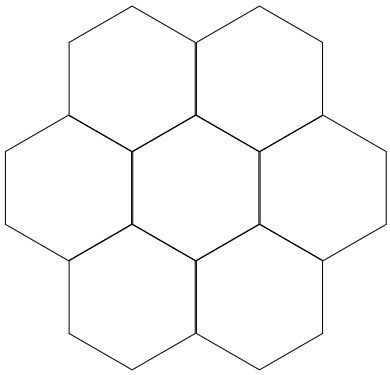
Homework 5

CS 3793/5233 – Fall 2016 — assigned November 10, 2016
Tom Bylander, Instructor — due December 6, 2016 before class

1. The game Hex-Hax-Hox is similar to Tic-Tac-Toe, but is played on the following board:



The game is won either by 3 in a straight line through the middle or 3 in a line on the outside. There are 9 possible ways to get 3 in a line. Let X_n be the number of possible 3 in a lins with exactly n X's and 0 O's. Similarly, Let O_n be the number of possible 3 in a lines with exactly n O's and 0 X's. Solve the following exercises using the following evaluation function:

$$Eval = 100X_3 + 10X_2 + X_1 - (100O_3 + 10O_2 + O_1)$$

   (a) (10 pts.) Approximate the number of possible Hex-Hax-Hox games, not taking symmetry into account.

   (b) (10 pts.) Show the whole game tree starting from an empty board down to depth 2 (X moves first, and O moves second), taking symmetry into account. You should have 2 positions on level 1 and 5 positions on level 2.

   (c) (10 pts.) Mark on your tree the evaluations of all the positions at level 2. Also, mark on your tree the backed-up values for the positions at levels 1 and 0, using the minimax algorithm, and use them to choose the best starting move.

   (d) (10 pts.) Show how alpha-beta pruning would evaluate the game tree starting from an empty board down to level 3 (X moves first). Assume the nodes are generated in *the optimal order for alpha-beta pruning*. Do not show the nodes that are not evaluated.

   (e) (10 pts.) Show that your instructor will not make any money with his invented game by showing that X has an forced win starting from an empty board.

2. (20 pts.) Consider applying k-means using $k = 2$ and sum-of-squares error to the following four points: $(1, 1), (1, 2), (2, 1), (2, 3)$. What are the stable assignments and what are their errors?

3. (30 pts.) In the following game, the agent starts at the square marked with a *. The agent can move left or right deterministicly, with a reward of -1 per move except when the agent reaches the 1 square (getting a reward of 1) or the 9square (getting a reward of 9). The game ends when the agent reaches either the 1 square or the 9 square. A value of a policy is evaluated using total reward.

1	*					9

   (a) (10 pts.) What is the value of the policy that always chooses a left action? What is the value of the policy that always chooses a right action?

   (b) (10 pts.) Suppose a left/right action succeeds with probability 0.9. Otherwise, the agent stays in the same square. What is the expected value of the policy that always chooses a left action? What is the expected value of the policy that always chooses a right action? I'm looking for exact answers. Hint: Look up geometric distributions and determine the average number of actions to move one square.

   (c) (10 pts.) What probability of success would result in identical values for the two strategies? Explain your answer.

   (d) (100 pts., shared extra credit) Suppose a left/right action succeeds with probability 0.9. Otherwise, the agent moves in the other direction. What is the expected value of the policy that always chooses a left action? What is the expected value of the policy that always chooses a right action? I'm looking for an answer correct to three significant digits.