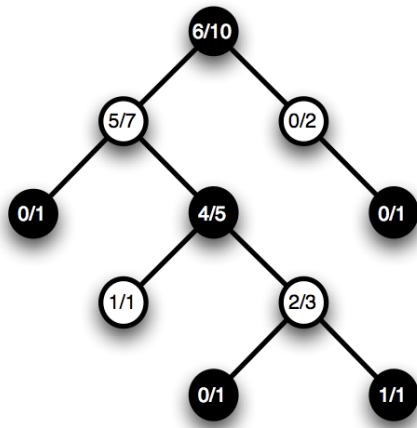Homework/Pop Quiz #5 of the course: Theory of Computer Games.


1. The above MCTS tree is built without considering opponents (i.e., all moves are for one player) by following the UCT formula.

$$a^* = \text{argmax}_{a \in legal}\left(Q\left(s, a\right) + c\sqrt{\frac{\log N(s)}{N(s,a)}}\right)$$

where Q(s,a) is the winning rate of the move a from state s, c is 0.1, N(s) is the number of samples on state s, and N(s,a) is the total number of samples on action a for state s.

   (a) Indicate which leaf to choose for the next UCT iteration, and depict how the tree will be changed if the expanded node is a loss.
   (b) Repeat (a), after (a) is done.
   (c) Repeat (a), after (b) is done.
   (d) Repeat (a) but for a win leaf, after (c) is done
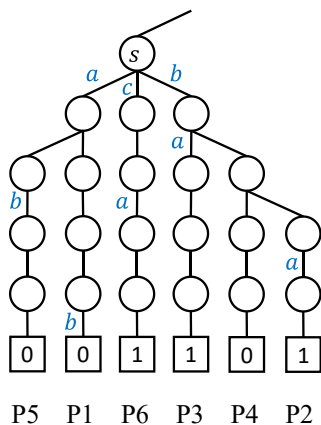
2. For the above UCT, assume that the playout sequence is P1, P2, P3, P4, P5, P6. Calculate all the values of Q(s,a), ~Q(s,a), N(s,a), ~N(s,a), after each playout. Note: ~Q(s,a) and ~N(s,a) are the RAVE version of Q(s,a) and N(s,a).

3. Calculate *Q(s,a), ~Q(s,a), N(s,a), ~N(s,a)*, again, assuming the following prior knowledge:

   H(s,a) = 0.6, *H(s,b)* = 0.55, *H(s,c)* = 0.5

   C(s,a) = 5, *C(s,b)* = 5, *C(s,c)* = 4

   ~C(s,a) = 8, *~C(s,b)* = 6, *~C(s,c)* = 6

   Note: *H(s,a)* is the initial value of *Q(s,a)* and *~Q(s,a)*, while *C(s,a)* and *~C(s,a)* are the initial values of *N(s,a)* and *~N(s,a),*.