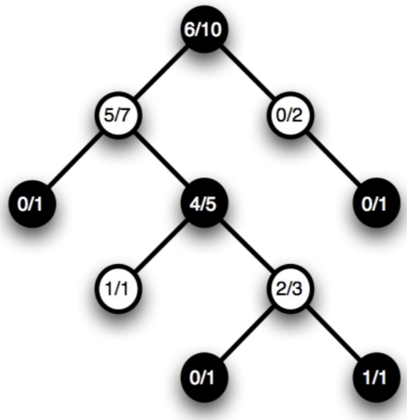
# Theory of Computer Games

## Homework # 5

311605004
劉子齊

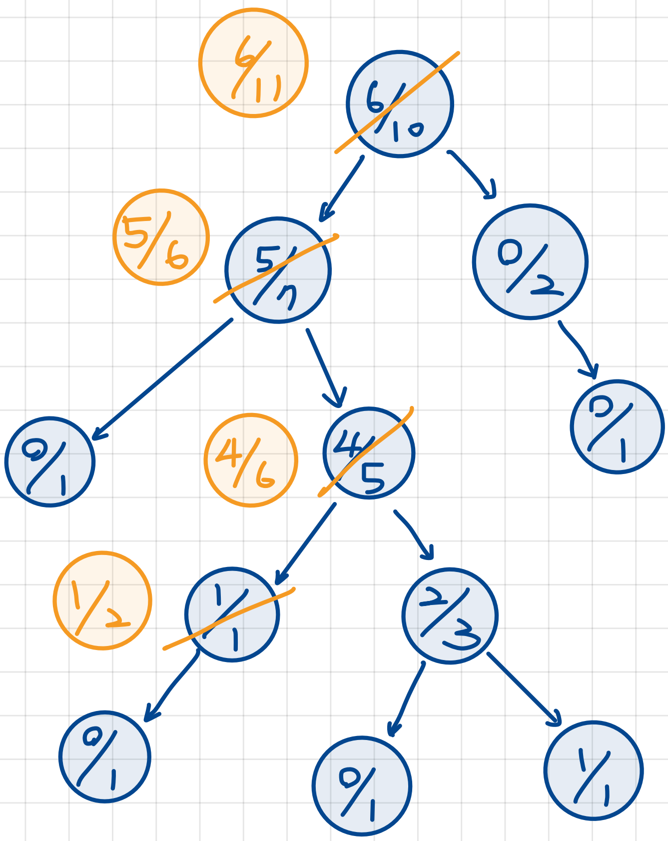



1. The above MCTS tree is built without considering opponents (i.e., all moves are for one player) by following the UCT formula.

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

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.

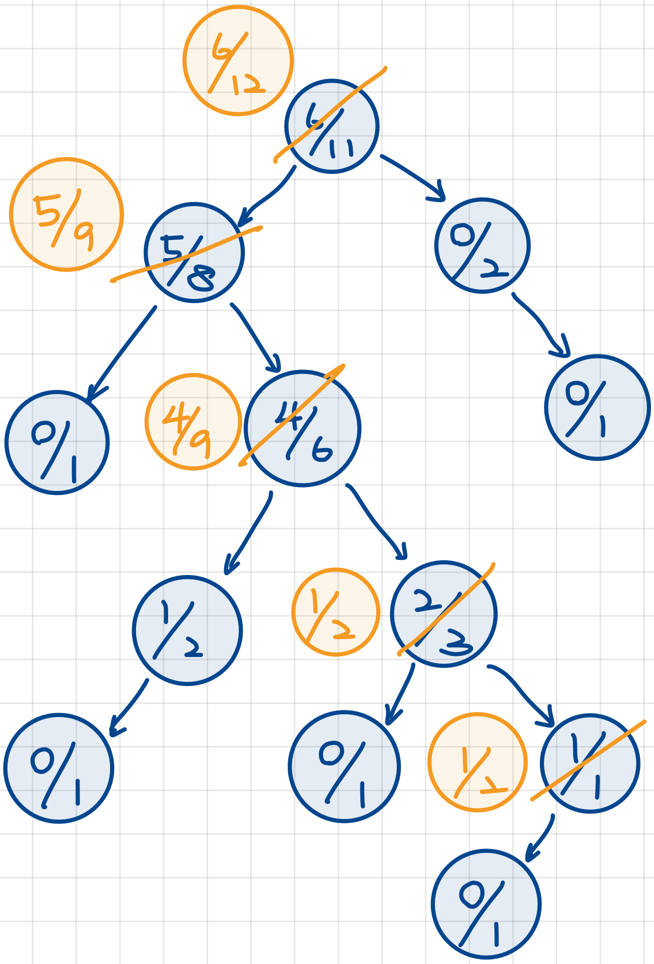



$$\frac{5}{7} + 0.1 \cdot \sqrt{\frac{\log 10}{7}} > 0 + 0.1 \cdot \sqrt{\frac{\log 10}{2}} = 0.07\ldots$$

$$0 + 0.1 \cdot \sqrt{\frac{\log 7}{1}} < \frac{4}{5} + 0.1\sqrt{\frac{\log 7}{5}}$$

$$1 + 0.1 \cdot \sqrt{\frac{\log 5}{1}} > \frac{2}{3} + 0.1 \cdot \sqrt{\frac{\log 5}{3}}$$

(b) Repeat (a), after (a) is done.



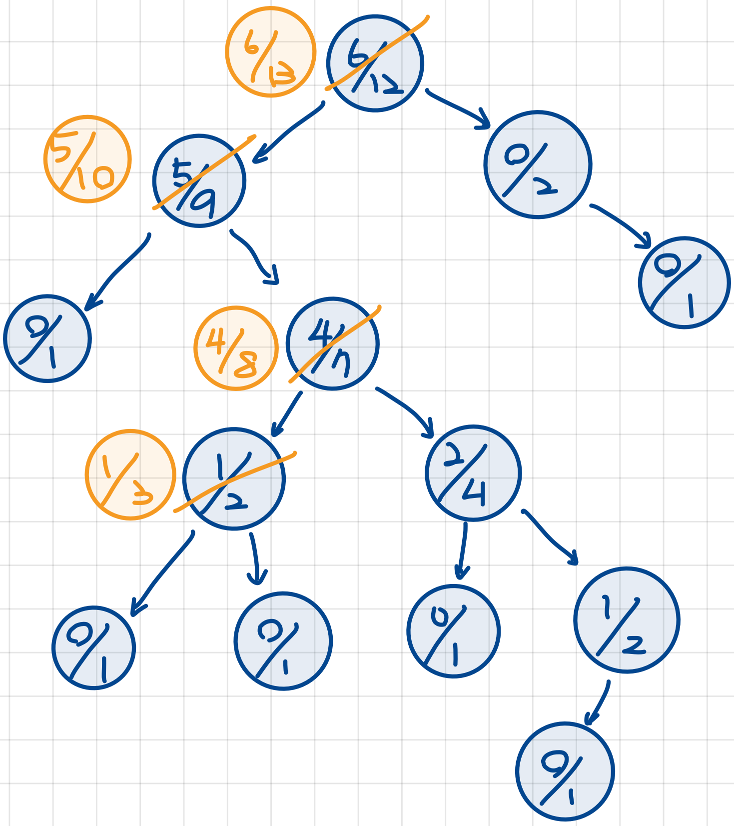

$$\frac{5}{8}+0.1\cdot\sqrt{\frac{\log 11}{8}} > 0+0.1\cdot\sqrt{\frac{\log 11}{2}}$$

$$0+0.1\cdot\sqrt{\frac{\log 8}{1}} < \frac{4}{6}+0.1\cdot\sqrt{\frac{\log 8}{6}}$$

$$\frac{1}{2}+0.1\cdot\sqrt{\frac{\log 6}{2}} < \frac{2}{3}+0.1\cdot\sqrt{\frac{\log 6}{3}}$$

$$0+0.1\cdot\sqrt{\frac{\log 3}{1}} < 1+0.1\cdot\sqrt{\frac{\log 3}{1}}$$

(c) Repeat (a), after (b) is done.

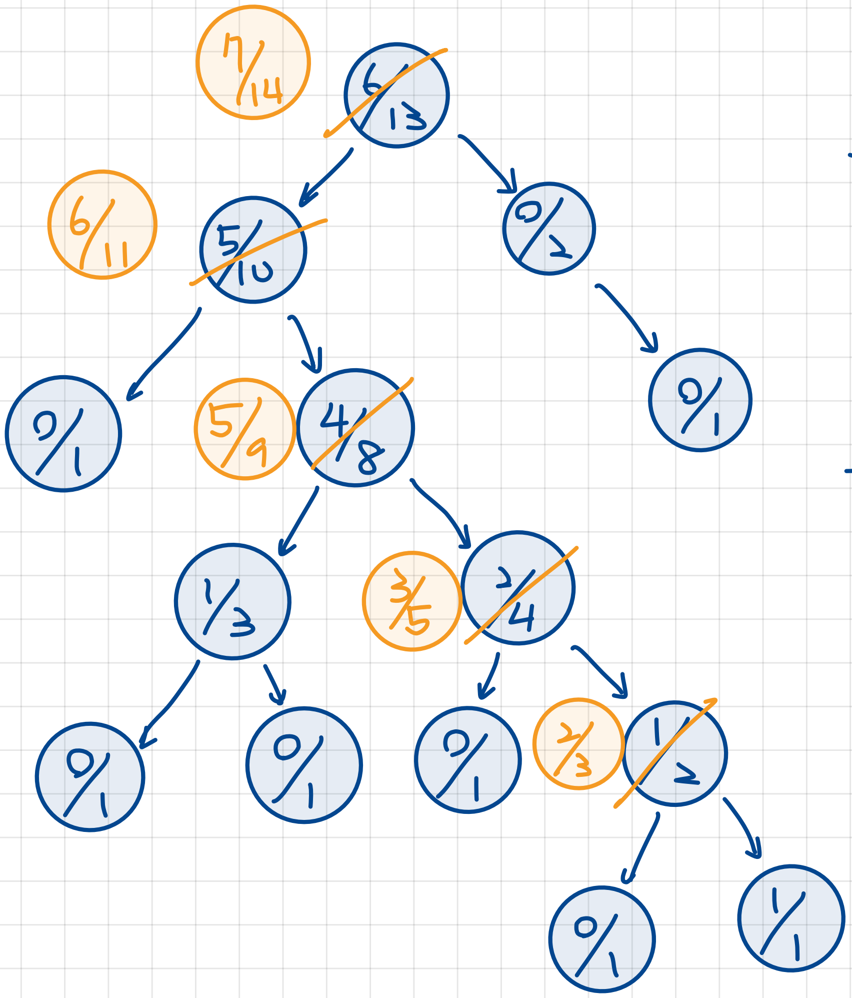



$$\frac{5}{9}+0.1\cdot\sqrt{\frac{\log 12}{9}} > 0+0.1\cdot\sqrt{\frac{\log 12}{2}}$$

$$0+0.1\cdot\sqrt{\frac{\log 9}{1}} < \frac{4}{7}+0.1\cdot\sqrt{\frac{\log 9}{7}}$$

$$\frac{1}{2}+0.1\cdot\sqrt{\frac{\log 7}{2}} > \frac{2}{4}+0.1\cdot\sqrt{\frac{\log 7}{4}}$$

(d) Repeat (a) but for a win leaf, after (c) is done




$$\frac{5}{10}+0.1\cdot\sqrt{\frac{\log 13}{10}} > 0+0.1\cdot\sqrt{\frac{\log 13}{2}}$$

$$0+0.1\cdot\sqrt{\frac{\log 10}{1}} < \frac{4}{8}+0.1\cdot\sqrt{\frac{\log 10}{8}}$$

$$\frac{1}{3}+0.1\cdot\sqrt{\frac{\log 8}{3}} < \frac{2}{4}+0.1\cdot\sqrt{\frac{\log 8}{4}}$$

$$0+0.1\cdot\sqrt{\frac{\log 4}{1}} < \frac{1}{2}+0.1\cdot\sqrt{\frac{\log 4}{2}}$$

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)*.

| | P1 | P2 | P3 | P4 | P5 | P6 |
|---|---|---|---|---|---|---|
| Q(s,a) | 0/1 | 0/1 | 0/1 | 0/1 | 0/2 | 0/2 |
| ~Q(s,a) | 0/1 | 1/2 | 2/3 | 2/3 | 2/4 | 3/5 |
| N(s,a) | 1 | 1 | 1 | 1 | 2 | 2 |
| ~N(s,a) | 1 | 2 | 3 | 3 | 4 | 5 |

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)*,.

| | P1 | P2 | P3 | P4 | P5 | P6 |
|---|---|---|---|---|---|---|
| N(S,a) | 5+1=6 | 6 | 6 | 6 | 6+1=7 | 7 |
| ~N(s,a) | 8+1=9 | 9+1=10 | 10+1=11 | 11 | 11+1=12 | 12+1=13 |
| Q(s,a) | $0.6 - \frac{0.6}{6} = \frac{3}{6}$ | $\frac{3}{6}$ | $\frac{3}{6}$ | 36 | (3+0)/(6+1) $= \frac{3}{7}$ | $\frac{3}{7}$ |
| ~Q(s,a) | $0.6 - \frac{0.6}{9} = \frac{4.8}{9}$ | (4.8+1)/(9+1) $= \frac{5.8}{10}$ | (5.8+1)/(10+1) $= \frac{6.8}{11}$ | $\frac{6.8}{11}$ | (6.8+0)/(11+1) $= \frac{6.8}{12}$ | (6.8+1)/(12+1) $= \frac{7.8}{13}$ |