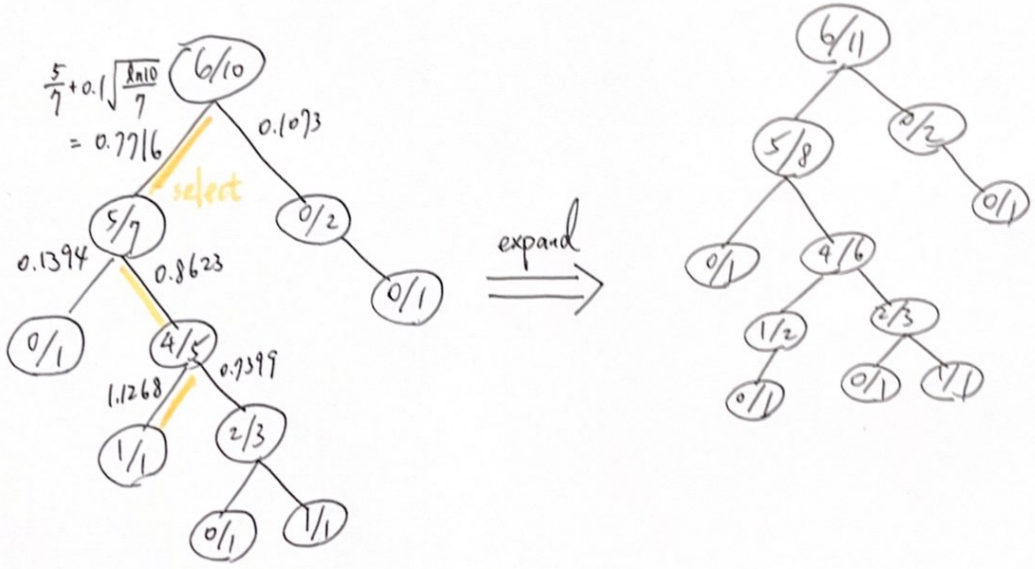
HW5 311551059 陳昱羽

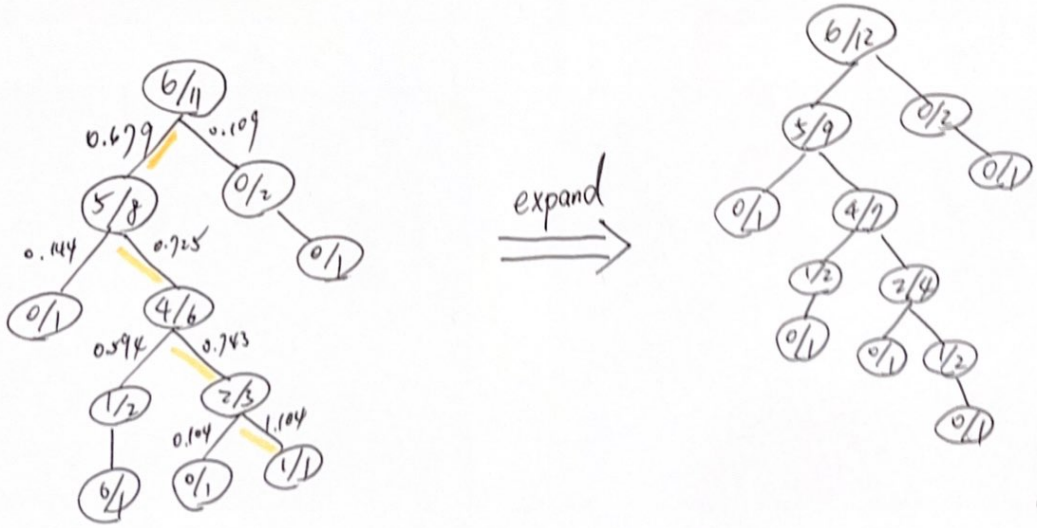
1.

(a) $a^* = \arg\max_a \left(Q(s,a) + c\sqrt{\frac{\log N(s)}{N(s,a)}} \right)$

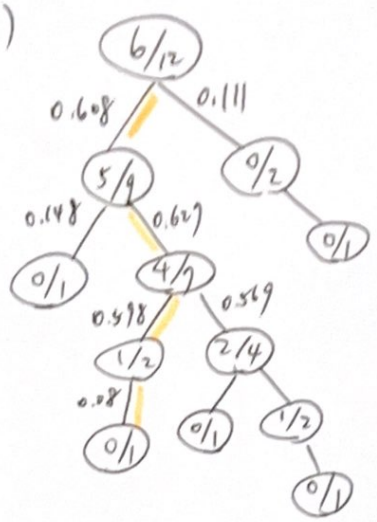
(log in computer → ln)



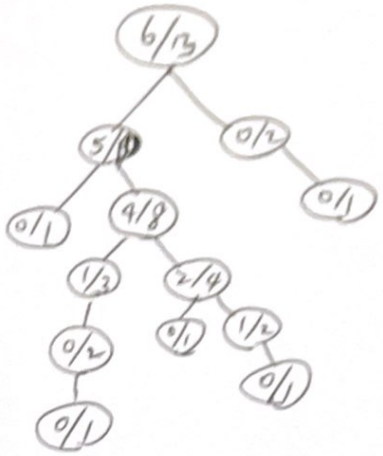
(b)



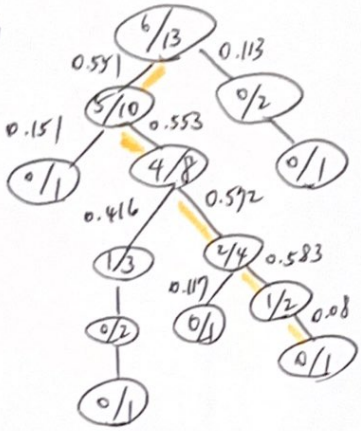
(c)



expand

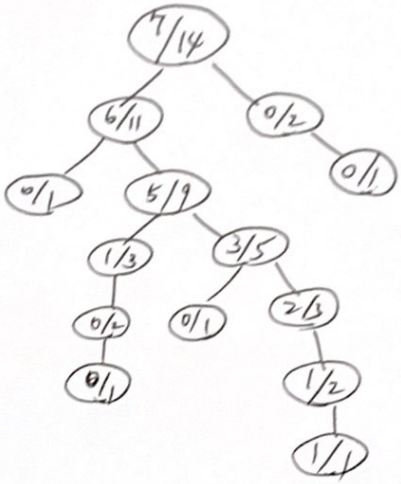


(d)



expand

(win)



2.

P_1: $Q(s,a) = \frac{0}{1}$, $N(s,a) = 1$, $\sim Q(s,a) = \frac{0}{1}$, $\sim N(s,a) = 1$ (loss)

P_2: $Q(s,a) = \frac{0}{1}$, $N(s,a) = 1$, $\sim Q(s,a) = \frac{1}{2}$, $\sim N(s,a) = 2$ (win)

P_3: $Q(s,a) = \frac{0}{1}$, $N(s,a) = 1$, $\sim Q(s,a) = \frac{2}{3}$, $\sim N(s,a) = 3$ (win)

P_4: $Q(s,a) = \frac{0}{1}$, $N(s,a) = 1$, $\sim Q(s,a) = \frac{2}{3}$, $\sim N(s,a) = 3$ ~~(loss)~~ (not related to a)

P_5: $Q(s,a) = \frac{0}{2}$, $N(s,a) = 2$, $\sim Q(s,a) = \frac{2}{4}$, $\sim N(s,a) = 4$ (loss)

P_6: $Q(s,a) = \frac{0}{2}$, $N(s,a) = 2$, $\sim Q(s,a) = \frac{3}{5}$, $\sim N(s,a) = 5$ (win)

3.

P_1: $Q(s,a) = \frac{0.6 \times 5 + 0}{5+1} = \frac{3}{6}$

$\sim Q(s,a) = \frac{0.6 \times 8 + 0}{8+1} = \frac{4.8}{9}$

$N(s,a) = 5+1 = 6$

$\sim N(s,a) = 8+1 = 9$

P_2: $Q(s,a) = \frac{3+0}{6+0} = \frac{3}{6}$

$\sim Q(s,a) = \frac{4.8+1}{9+1} = \frac{5.8}{10}$

$N(s,a) = 6$

$\sim N(s,a) = 9+1 = 10$

P_3: $Q(s,a) = \frac{3+0}{6+0} = \frac{3}{6}$

$\sim Q(s,a) = \frac{5.8+1}{10+1} = \frac{6.8}{11}$

$N(s,a) = 6$

$\sim N(s,a) = 10+1 = 11$

P_4: $Q(s,a) = \frac{3+0}{6+0} = \frac{3}{6}$

$\sim Q(s,a) = \frac{6.8+0}{11+0} = \frac{6.8}{11}$

$N(s,a) = 6$

$\sim N(s,a) = 11$

P_5: $Q(s,a) = \frac{3+0}{6+1} = \frac{3}{7}$

$\sim Q(s,a) = \frac{6.8+0}{11+1} = \frac{6.8}{12}$

$N(s,a) = 6+1 = 7$

$\sim N(s,a) = 11+1 = 12$

P_6: $Q(s,a) = \frac{3+0}{7+0} = \frac{3}{7}$

$\sim Q(s,a) = \frac{6.8+1}{12+1} = \frac{7.8}{13}$

$N(s,a) = 7$

$\sim N(s,a) = 12+1 = 13$