- **Your Name:** *Qihang Dai*

- **Your PennKey:** *ahgdyycc*

- **Your PennID:** *78803164*

1   Multiple Choice & Written Questions

1. (a) bob loves cookie

   (b) $lnP(w1 = loves, w2 = cookie|w0 = bob) = lnP(w1 = loves|w0 = bob) + lnP(w2 = cookie|w1 = loves, w0 = bob) = ln0.5 + ln0.4 = ln0.2 =$
   $lnP(w1 = hates, w2 = cookie|w0 = bob) = lnP(w1 = hates|w0 = bob) + lnP(w2 = cookie|w1 = hates, w0 = bob) = ln0.4 + ln0.2 = ln0.08$

   (c) no. greedy sampling may be the best for the first word, but it is not the best for the second word. And the combination of this two gram may not be the highest probability among the 2-gram model.

   (d) $lnP(w1 = loves, w2 = Bob|w0 = Bob) = lnP(w1 = loves|w0 = Bob) + lnP(w2 = Bob|w1 = loves, w0 = Bob) = ln0.5 + ln0.25 = ln0.125$
   $lnP(w1 = hates, w2 = cherry|w0 = Bob) = lnP(w1 = hates|w0 = Bob) + lnP(w2 = cherry|w1 = hates, w0 = Bob) = ln0.4 + ln0.7 = ln0.28$

   Compared with answer calculated in (a), we should keep the two sentence with higher probability. So the answer is "Bob loves cherry" and "Bob loves cookie".

2. (a)
   $$[0.52 0.37 0.31 0.16]$$

   (b)
   $$[0.297 0.255 0.241 0.207]$$

   (c)
   $$[0.2796 0.4032 0.262 0.3138]$$

3. (a)

State s	Condition	i	Vi(s)
P	V(P)>0	2	75
Q	V(Q)>0	2	30
P	V(P) = V*(P)	2	75
Q	V(Q) = V*(Q)	5	75

   (b)

State s	Condition	i	Vi(s)
P	V(P) = V*(P)	2	48

Figure 1 grids (value iteration process):

Grid V0:
P0	0	0	Q0
0	V0		0
+75			+30

Grid V1:
P			Q
75	V1		30
+75			+30

Grid V2:
P 75			Q 30
75	V2		30
+75			+30

Grid V3:
P 75	75	30	Q 30
75	V3		30
+75			+30

Grid V4:
P 75	75	75	Q 30
75	V4		30
+75			+30

Grid V5:
P 75	75	75	Q 75
75	V5		30
+75			+30
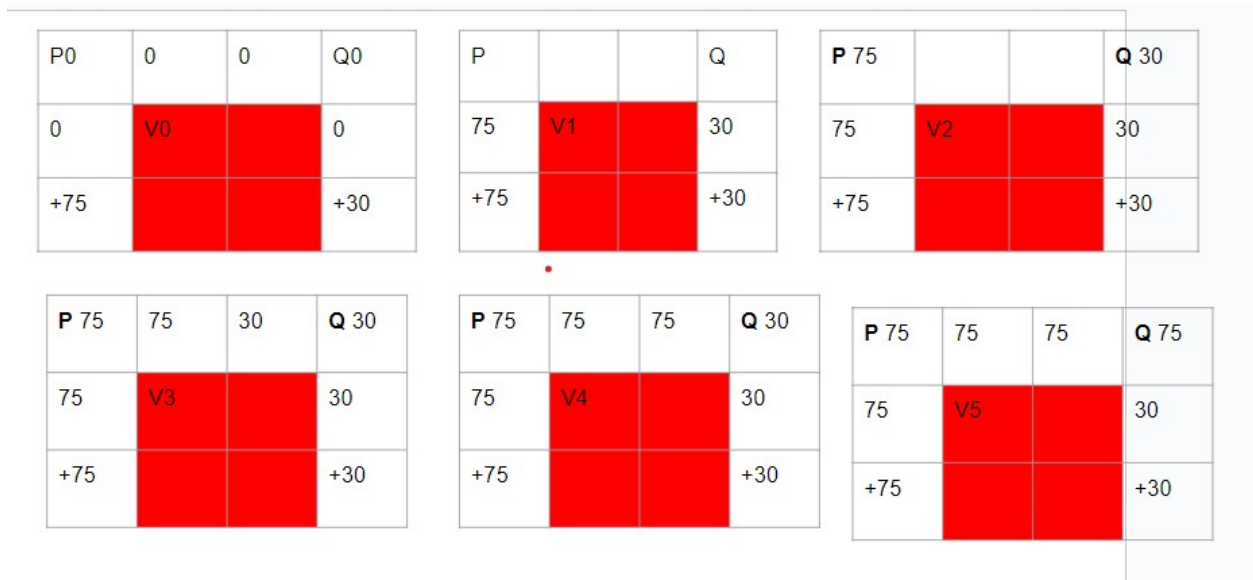
Figure 1: (a) The value iteration process

Figure 2 grids (stochastic value iteration process):

Grid V0:
P0	0	0	Q0
0	V0		0
+75			+30

Grid V1:
P			Q
60	V1		24
+75			+30

Grid V2:
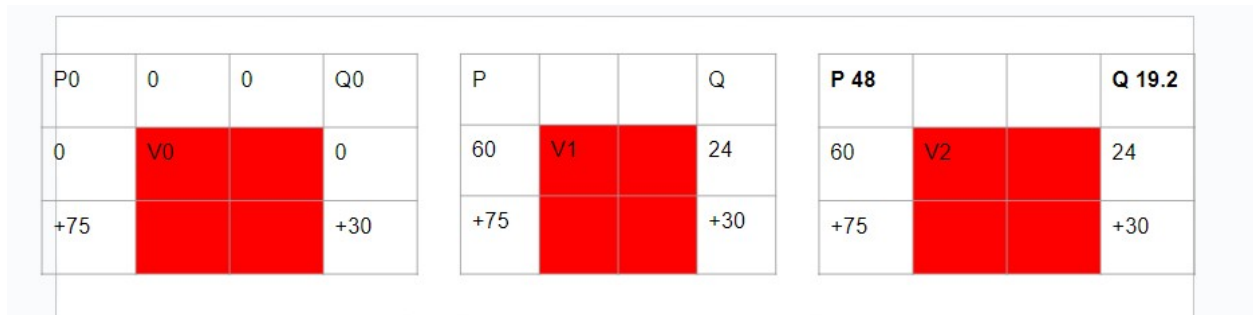| P 48			Q 19.2
60	V2		24
+75			+30

Figure 2: (b) The stochastic value iteration process

4. (a) see figures below

   (b) V* = 6.4
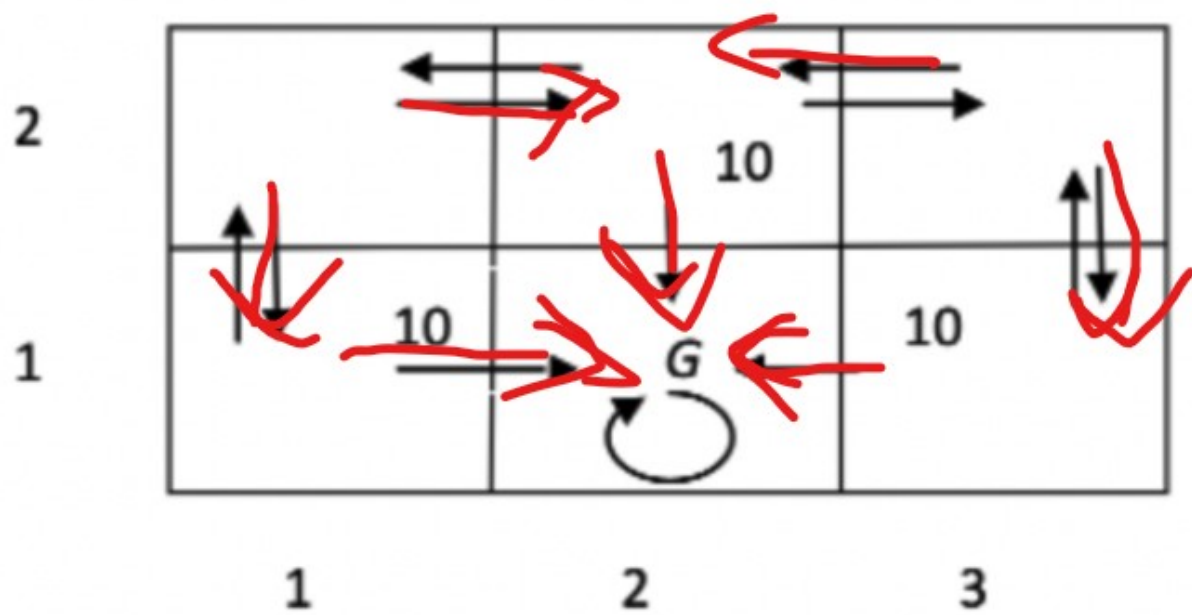
   (c) grid[row2, col1] = 8, grid[row2,col2] = 10, grid[row2, col3] = 8, grid[rpw1, col3] = 10

Figure 3: (a) The optimal action