CSE 573 Written HW 2

Alexander Van Roijen

March 15, 2020

I Value Iteration

I.I compute the utility of each action in terms of γ

U(up) = $-50 + \Sigma_{i=1}^{100}\gamma^i$
U(down) = $50 - \Sigma_{i=1}^{100}\gamma^i$

I.II Assuming a discounted reward function, for what values of γ should the agent choose up instead of down initially?

want U(up)>U(down). $-50+\Sigma_{i=1}^{100}\gamma^i = 50-\Sigma_{i=1}^{100}\gamma^i \rightarrow 100 < 2*\Sigma_{i=1}^{100}\gamma^i \rightarrow \Sigma_{i=1}^{100}\gamma^i > 50 \rightarrow \frac{\gamma^{101}-1}{\gamma-1} - 1 > 50 \rightarrow \gamma \geq 0.985$

II reinforcement learning

II.I a

$sample = reward + \gamma \max(Q(s'))$
$Q(s,a) = (1-\alpha)Q(s,a) + \alpha sample$
$\alpha = 0.25, \gamma = 1$
Note from the first sample, max Q(A) = 4
Q(J,Pass)= $0.75*0+0.25*(0+1*3) = 3/4$ We get this from the max q value for the sample prior on $s\prime = A$
Q(J,Bet)= $0.75*4+0.25*(0+1*3) = 4+3/4 We get this again from the same max q value as before.$

II.II b

$Q(s,a) = w_1 f_1(s,a) + w_2 f_2(s,a)$
After the first sample
$\text{diff} = r + \gamma \max_a'(Q(s',a')) - Q(s,a) = 8 + 0 - 0 = 8$
$w_1 = w_1 + \alpha * (\text{diff}) f_1(s,a) = 0 + 0.25 * 8 * 1 = 2$
$w_2 = w_2 + \alpha * (\text{diff}) f_2(s,a) = 0 + 0.25 * 8 = 2$
Further note that our Q value has updated to now be $Q(A, Bet) = 2$
After the second sample
$\text{diff} = r + \gamma \max_a'(Q(s',a')) - Q(s,a) = 0 + 1 * 2 - (2 + 0) = 0$
$w_1 = w_1 + \alpha * (\text{diff}) f_1(s,a) = 2 + 0.25 * 0 * 1 = 2$
$w_2 = w_2 + \alpha * (\text{diff}) f_2(s,a) = 2 + 0.25 * 0 * 0 = 2$

III reinforcement learning 2

action	Probability
Left	$\epsilon/2$
Right	$1-\epsilon$
Fire	$\epsilon/2$

The results are exactly because we act according to our optimal policy $1-\epsilon$ percent of the time, and uniformly randomly with no particular priority the remaining ϵ

IV Probability and uncertainty

IV.I A

Solving the initial conditions like a system of equations, we get that $P(A = T) = 0.5, P(B = T) = 0.8, P(B = F) = 0.2$ which means
$P(A = F, B = T) = 0.5 * 0.8 = 0.4$
$P(A = F, B = F) = 0.5 * 0.2 = 0.1$

IV.II B

IV.III B1

$$
\begin{aligned}
P(a,b|c) &= \frac{P(a,b,c)}{P(c)} && (1)\\
&= \frac{P(a,b,c)}{P(c)}\frac{P(b,c)}{P(b,c)} && (2)\\
&= \frac{P(a,b,c)}{P(b,c)}\frac{P(b,c)}{P(c)} && (3)\\
&= P(a|b,c)P(b|c)\square && (4)
\end{aligned}
$$

IV.IV B2

$$\frac{P(a,b,c)}{P(c)} = \frac{0.01}{P(c)} = P(a|b,c)P(b|c) = 0.2 * 0.1 \rightarrow P(c) = \frac{0.01}{0.02} = 0.5\square \tag{5}$$

IV.V B3

$$P(a,b|c) = P(a|c)P(b|c) = \frac{P(a,b,c)}{P(c)} = \frac{P(a,c)}{P(c)}\frac{P(b,c)}{P(c)} \tag{6}$$

$$= P(a,b,c) = \frac{P(a,c)}{P(c)} * P(b,c) = \frac{P(a,b,c)}{P(b,c)} = \frac{P(a,c)}{P(c)} = P(a|b,c) = P(a|c)\square \tag{7}$$

IV.VI B4

$$P(a,b,c) = P(a,b,c)\frac{P(c)}{P(c)} = P(a,b|c)P(c) = P(a|c)P(b|c)P(c) = P(a)P(b|c)P(c) \tag{8}$$

V HMM

V.I A

$P(W_1|E_1) = \frac{P(E_1|W_1)P(W_1)}{P(E_1)} \rightarrow a^1 = P(W_1 = 1|E_1 = A) = 0.3 * 0.8/(0.3 * 0.8 + 0.6 * 0.2) = 2/3$ and $a^0 = P(W_1 = 0|E_1 = A) = 0.6 * 0.2/(0.3 * 0.8 + 0.6 * 0.2) = 1/3$

V.II B

$P(W_2|E_1) = \Sigma_{w_1} P(W_2|W_1) * P(W_1|E_1) \rightarrow \Sigma_{w_1} P(W_2 = 0|W_1) * P(W_1|E_1 = A) = P(W_2 = 0|W_1 = 0) * P(W_1 = 0|E_1 = A) + P(W_2 = 0|W_1 = 1) * P(W_1 = 1|E_1 = A) = 0.9 * a^0 + 0.7 * a^1 = b^0$
$b^1 = \Sigma_{w_1} P(W_2 = 1|W_1) * P(W_1|E_1 = A) = P(W_2 = 1|W_1 = 0) * P(W_1 = 0|E_1 = A) + P(W_2 = 1|W_1 = 1) * P(W_1 = 1|E_1 = A) = 0.1 * a^0 + 0.3 * a^1 = b^1$

V.III C

$P(W_2|E_1, E_2) = P(W_2|E_1)P(E_2|W_2) \rightarrow c^0 = P(W_2 = 0|E_1 = A)P(E_2 = B|W_2 = 0) = b^0 * 0.4$
$\rightarrow c^1 = P(W_2 = 1|E_1 = A)P(E_2 = B|W_2 = 1) = b^1 * 0.7$

V.IV D

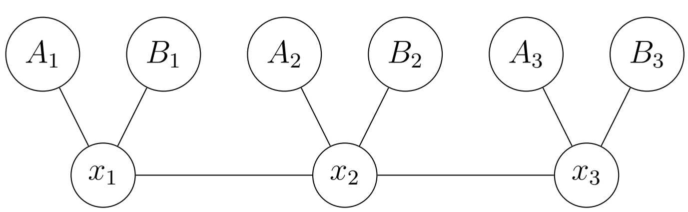
$P(W_3|E_1) = \Sigma_{w_2} P(W_3, w_2|E_1) = \Sigma_{w_2} P(W_3|w_2) * P(w_2|E_1) \rightarrow d^0 = \Sigma_{w_2} P(W_3 = 0|w_2) * P(w_2|E_1 = A) = b^0 * 0.9 + b^1 * 0.7$
$\rightarrow d^1 = \Sigma_{w_2} P(W_3 = 1|w_2) * P(w_2|E_1 = A) = b^0 * 0.1 + b^1 * 0.3$

V.V E

$P(W_3|E_1, E_3) = P(W_3|E_1)P(E_3|W_3) \rightarrow e^0 = P(W_3 = 0|E_1 = A, E_3 = B) = P(W_3 = 0|E_1 = A)P(E_3 = B|W_3 = 0) = d^0 * 0.4$
$\rightarrow e^1 = P(W_3 = 1|E_1 = A, E_3 = B) = P(W_3 = 1|E_1 = A)P(E_3 = B|W_3 = 1) = d^1 * 0.7$

VI Modified HMM

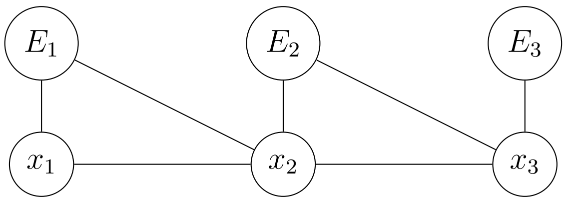
VI.I A



VI.II B

$P(X_t|a_{1:t}, b_{1:t}) \propto P(X_t|a_{1:t-1}, b_{1:t-1})P(a_t, b_t|X_t)$

VI.III C



VI.IV D

$P(X_t|e_{1:t-1}) = \Sigma_{x_{t-1}} P(X_t|x_{t-1}, e_{t-1})P(x_{t-1}|e_{1:t-1})$

VII Bayes nets representation

VII.I a

$P(+g, +a, +b, +s) = P(S|A, B, G)P(A, B, G) = P(S|A, B, G)P(A, B, G)\frac{P(G)}{P(G)} = P(S = +|A = +, B = +)P(A = +|G = +)P(B = +|G = +)P(G = +) = 0.1*1.0*0.4*1.0 = .04$

VII.II b

$P(+a) = \Sigma_g P(+a|G = +)P(G = +) = 1 * 0.1 + 0.1 * 0.9 = .1 + .09 = .19$

VII.III c

$P(A|B) = \frac{\Sigma_g P(A,B|g)P(g)}{P(B)} \rightarrow \frac{(P(A+|g+)P(B+|g+)P(g+)+P(A+|g-)P(B+|g-)P(g-))}{P(B+)} = \frac{1*0.4*0.1+0.1*0.4*0.9}{0.4} = \frac{0.04+0.036}{0.4} = 0.19$ which makes sense as A is independent of B as long as we dont know s $= 0.19$

VII.IV d

$P(A|B, S)$ lets do this in two parts, first $P(A, B, S) = P(A+)P(B+|A+)P(S|B+, A+) = 0.19 * 0.4 * 1 = 0.076$ by the chain rule.

$P(B,S) = \Sigma_a P(S|B,a)P(a)P(B) = P(s{+}|b{+},a{+})P(a{+})P(b{+}) + P(s{+}|b{+},a{-})P(a{-})P(b{+}) = 1 * 0.19 * 0.4 + 0.8 * 0.81 * 0.4 = 0.335$
By dividing the two we get $0.076/0.335 = 0.227$

VIII Bayesian Inference

We initially have $P(B|w = \text{true}) = \Sigma_{a,e,r} P(B,a,e,r|W = \text{True})$

VIII.I A

Using properties of this $P(A|B,E)P(B)P(E)P(R|E)P(w = \text{True}|A) = P(A|B,E)P(B)P(R,E)P(w = \text{Tr}$

VIII.II B

Eliminating A gives us $\Sigma_a P(w = \text{true}, a|B,E)P(B)P(E)P(R|E) = P(w = \text{true}|B,E)P(B)P(E)P(R|E)$
Eliminating E gives us $\Sigma_e P(w = \text{true}|B,E)P(B)P(R,E) = P(w = \text{true}|B,e)P(B)P(R)$
Finally when eliminating R we are left with $\Sigma_r P(w = \text{true}|B,e)P(B)P(r) = P(w = \text{true}|B,e)P(B)$ which is simply a function of our query B. With B=b we can solve for our exact query after normalization.

VIII.III C

We can avoid having the table of w, B, E if we sum out A last, so rather, eliminate the variables in order R, E, A