CSCI4150: Artificial Intelligence - Exam2

Your name: Samad Farooqui

RIN: 661540118

Please read instructions carefully. Do not worry if you cannot finish everything. Do not write down disorganized answers in the hope of getting partial credit; it's better to do a few questions completely right. Please write your answers down clearly (think before you write). You can use extra pages.
Good luck!
–Lirong

69

Problem 1	Problem 2	Problem 3	Problem 4	Problem 5	Problem 6	Total
10 /10	19 /20	22 /35	13 /15	0 /15	5 /5	/100

Problem 1: Game Theory (10pt). Consider the following non-zero sum game in matrix-normal form

Possible NE: DR? or UL?

R dominates C
U dominates M

	L	C	R
U	(4,6)	(3,1)	(1,5)
~~M~~	~~(2,2)~~	~~(1,4)~~	~~(1,4)~~
D	(3,3)	(0,0)	(2,6)

1. **5pt**. What are the (pure) Nash Equilibria of this game? You don't need to show your calculation.

(D,R) and (U,L)

2. **5pt**. What is the 2×2 matrix that results from the iterative elimination of dominated strategies? Show the dominance relationships in your calculation.

R dominates C since $5 \geq 1$, $4 \geq 4$, and $6 \geq 0$ for the column player. U dominates M since $4 \geq 2$, $3 \geq 1$, and $1 \geq 1$ for the row player. The matrix is

	L	R
U	(4,6)	(1,5)
D	(3,3)	(2,6)

a≻b 5 → ← 3 b≻a

a≻c → 5 ← 3 c≻a

a≻d 2+1+3 d≻a 2

b≻c 2+3 5 c≻b 2+1 3

b≻d 2+1 3 d≻b 2+3 5

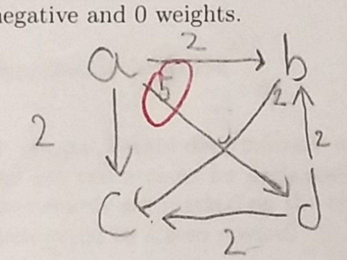
c≻d 2 +1 d≻c 2+3

Problem 2: Voting (20pt). Let P denote the following profile:

$$P = 2@[\underbrace{a \succ b \succ c \succ d}_{P_1}] + 2@[\underbrace{d \succ c \succ b \succ a}_{P_2}] + 1@[\underbrace{c \succ b \succ a \succ d}_{P_3}] + 3@[\underbrace{a \succ d \succ b \succ c}_{P_4}]$$

For these problems, $I = \{a, b, c, d\}$, $i \in I$

1. **5pt**. Draw the weighted majority graph. You don't need to draw edges with negative and 0 weights.



-1

2. **5pt**. Who is the Plurality winner? You must show your calculations.

P_1: 2 for a, 0 for everyone else
P_2: 2 for d, 0 $\forall i$ where $i \neq d$
P_3: 1 for c, 0 for $\forall i \neq c$
P_4: 3 for a, 0 for $\forall i \neq a$

a total: 2+3 = 5
b total: 0
c total: 1
d total: 2

$\boxed{\text{a wins}}$

3. **5pt**. Who is the Borda winner? You must show your calculations.

Borda: Place x vote	P_1	P_2	P_3	P_4	Total
a:	3x2=6	0x2=0	1x1=1	3x3=9	16
b:	2x2=4	1x2=2	2x1=2	1x3=3	11
c:	1x2=2	2x2=4	3x1=3	0	9
d:	0x2=0	3x2=6	1x0=0	3x2=6	12

$\boxed{\text{a wins}}$

4. **5pt**. Who is the STV winner? You must show your calculations.

Round 1 w/ P: B had 0 first-place wins, so out.
Round 2 P': 2@[a≻c≻d] + 2@[d≻c≻a] + 1@[c≻a≻d] + 3@[a≻d≻c]
C had 1 first place win, a had 5, d had 2. C is out.
Round 3 P'': 2@[a≻d] + 2@[d≻a] + 1@[a≻d] + 3@[a≻d]
6 first place for a, 2 first place for d.

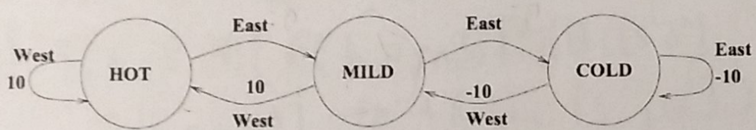
$\boxed{\text{a wins}}$

Problem 3: MDPs and Reinforcement Learning (35pt) This problem has two parts, one focused on Q-learning in a deterministic world and one focused on Q-learning in a nondeterministic world. If you get stuck, be sure to scan the remaining parts, because some latter parts can be answered readily without answering previous parts.

Part A. Deterministic world

$a: \{W, E\}$

Consider the simple 3-state deterministic weather world sketched in the figure below. There are two actions for each state, namely moving **West** and **East**. The non-zero rewards are marked on the transition edges, hence edges without a number correspond to a zero reward.



1. **2pt**. How many possible policies are there in this deterministic world? 4

2. **8pt**. A robot starts in the state **Mild**. It is actively learning its Q-table and moves in the world for 4 steps choosing actions according to a toss of a coin, namely **West, East, East, West**. The initial values of its Q-table are 0 and the discount factor is $\gamma = 0.5$. Fill in the Q-values after each action below.

4

	Initial State: MILD		Action: West New State: HOT		Action: East New State: MILD		Action: East New State: COLD		Action: West New State: MILD	
	East	West	East	West	East	West	East	West	East	West
HOT	0	0	0	0	5	0	5	9	5	0
MILD	0	0	0	10	0	10	2.5	10	2.5	10
COLD	0	0	0	0	0	0	0	0	0	-5

$Q_{t+1}($

$0 + 0.5(10)$

$10 + 0.5(10)$

$R(C, W, M)$

$-10 + 0.5(10)$

hot : h   west : w
mild : m   east : e
cold : c

10

3. **10pt.** Prove that the *always-west* policy: [for all state s, $\pi(s)$ =West] is better than the *always-east* policy: [for all state $s, \pi(s)$ =East]. Hint: you can prove it by showing that for each state, its reward under *always-west* is higher than its reward under *always-east*. Proof by cases.

$\gamma : 0.2$

State : Hot
at iteration i

$Q_{i+1}(\text{hot, west}) = \underbrace{T(\text{hot, west, hot})}[R(\text{hot, west, hot}) + \gamma Q(s', a')]$ Only one possible s' in going west.

$= T(h, w, \text{hot})[10 + \gamma Q(s', a')]$

$Q_{i+1}(\text{hot, east}) = T(h, e, \text{mild})[R(h, e, \text{mild}) + \gamma Q_i(s', a')]$

$= T(h, e, \text{mild})[-10 + \gamma Q_i(s', a')]$ $Q_i(h, w) > Q_i(h, e)$ since r larger since max values are constant.

State : Mild

$Q_{i+1}(m, w) = T(m, w, h)[R(m, w, h) + \gamma Q_i(s', \max a')] = T(m, w, h)[10 + \gamma Q_i(s', \max a')]$

$Q_{i+1}(m, e) : T(m, e, c)[R(m, e, c) + \gamma Q_i(s', \max a')]$ $T(m, e, c)[0 + \gamma Q_i(s', \max a')]$

$Q_{i+1}(m, w) > Q_{i+1}(m, e)$ since reward is higher $10 > 0$.

State : cold.

$Q_{i+1}(c, w) = T(c, w, m)[R(c, w, m) + \gamma Q_i(s', \max a')]$

$= T(c, w, m)[-10 + \gamma Q_i(s', \max a')]$

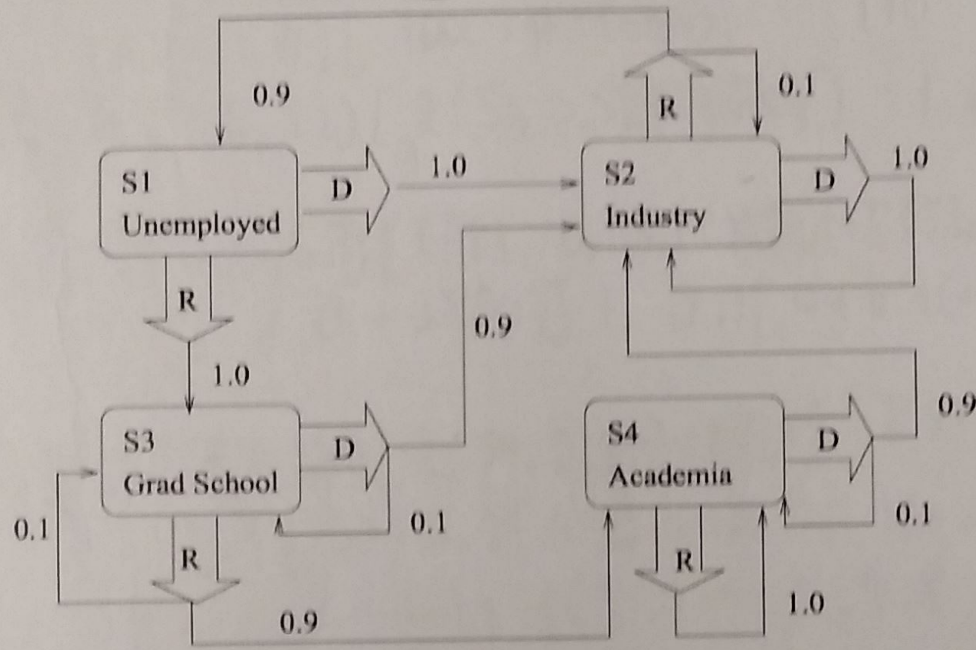
$Q_{i+1}(c, e) = T(c, e, c)[R(c, e, c) + \gamma Q_i(s', \max a')]$

$= T(c, e, c)[-10 + \gamma Q_i(s', \max a')]$

So $Q_{i+1}(c, w) > Q_i(c, e)$ since reward is the same and all the other actions have it so $w > e$. Therefore, for every state, the Q value of going west will be better than going east since reward is always higher for going west: even in cold state, reward will be higher as going east will have you stay in cold, suffering more.

Part B. Nondeterministic world

Consider the Markov Decision Process below. Actions have nondeterministic effects, i.e., taking an action in a state always leads to one next state, but which state is the next is determined by transition probabilities. The transition probabilities are shown in the figure attached below. There are two actions out of each state: **D** (going for development) and **R** (going for research).



Consider the following deterministic *ultimately-care-only-about-money* reward for any transition starting at a state:

State: S	S1	S2	S3	S4
Reward: $R(S)$	0	100	0	10

That is, the reward $R(S, a, S') = R(S)$ in the above table. Suppose $\gamma = 0.9$, and further suppose that the optimal policy is $\pi^*(s) = D$, for any $s \in \{S1, S2, S3, S4\}$.

$R(S, a, S') - R(S)$

$\max_a$

1. **12pt**. Compute the optimal expected values from each state, namely $V^*(S1), V^*(S2), V^*(S3), V^*(S4)$ according to this policy. (*Hints: Start by calculating $V^*(S2)$*)

$$V^*(S2) = \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V_i^*(s')]$$

6

$$= T(S2,D,S2)[R(S2,D,S2) + \gamma(100)]$$

$$= (1.0)[100 + 0.9(100)] = 190$$

$$V^*(S1) = T(S1,D,S2)[R(S1,D,S2) + \gamma V_i(S2)] = 1.0[0 + 171] = 171$$

$$V^*(S3) = T(S3,D,S3)[R(S3,D,S3) + \gamma V_i(S3)] + T(S3,D,S2)[R(S3,D,S2) + \gamma V(S2)]$$

$$= 0.1[0 + 0.9(0)] + 0.9[0 + 0.9(190)] = 154$$

$$V^*(S4) = T(S4,D,S4)[R(S4,D,S4) + \gamma V^*(S4)] + T(S3,D,S2)[R(S4,D,S2) + \gamma V^*(S2)]$$

$$= 0.1[10 + 0.9(10)] + 0.9[10 + 0.9(190)] =$$

$$1.9 + 162.9 = 164.8$$

2. **3pt**. Compute the Q-value $Q(S2, R)$.

1

$$T(S2,R,S2)[R(S2,R,S2) + Q_i(S2,D)]$$

$$+ T(S2,R,S1)[R(S2,R,S1) + Q_i(S2,D)]$$

$$= 0.1[100 + 2] + 0.9[100 + 2]$$

$$= 102$$

Problem 4: Naïve Bayes Classifier Learning (15pt). You are a robot in an animal shelter, and must learn to discriminate Dogs from Cats. You choose to learn a Naïve Bayes classifier. You are given the following (noisy) examples:

Sound	**Fur**	**Color**	**Class**
Meow	Coarse	Brown	Dog
Bark	Fine	Brown	Dog
Bark	Coarse	Black	Dog
Bark	Coarse	Black	Dog
Meow	Fine	Brown	Cat
Meow	Coarse	Black	Cat
Bark	Fine	Black	Cat
Meow	Fine	Brown	Cat

S=M: 4/8
S=B: 4/8
f:C: 4/8
f:f: 4/8
C:Br

Recall that Baye's rule allows you to rewrite the conditional probability of the class given the attributes as the conditional probability of the attributes given the class. As usual, Z is a normalizing constant that makes the probabilities sum to one.

$$\text{P(Class|Sound, Fur, Color)} = Z * \text{P(Sound, Fur, Color|Class)P(Class)}$$

1. **3pt**. Now assume that the attributes (Sound, Fur, and Color) are conditionally independent given the Class. Rewrite the expression of the Naïve Bayes classifier with the conditional independence assumption.

3

P(Class|Sound, Fur, Color) = P(class) P(sound|class) P(fur|class) P(color|class) · Z

7

2. **7pt**. Fill in numerical values for the following expressions (no Laplace Smoothing). Leave your answers as common fractions (e.g., 1/4, 3/5).

P(Dog)		P(Cat)	
P(Sound=Meow | Class=Dog)	1/4	P(Sound=Meow | Class=Cat)	3/4
P(Sound=Bark | Class=Dog)	3/4	P(Sound=Bark | Class=Cat)	1/4
P(Fur=Coarse | Class=Dog)	3/4	P(Fur=Coarse | Class=Cat)	1/4
P(Fur=Fine | Class=Dog)	1/4	P(Fur=Fine | Class=Cat)	3/4
P(Color=Brown | Class=Dog)	1/2	P(Color=Brown | Class=Cat)	1/2
P(Color=Black | Class=Dog)	1/2	P(Color=Black | Class=Cat)	1/2

3

3. **5pt**. Consider a new example: Sound=Bark, Fur=Coarse, Color=Brown. Calculate the class probabilities using the common fractions from above. **Show your calculation to receive full scores.**

P(Dog|Sound = Bark, Fur = Coarse, Color = Brown) =
P(Cat|Sound = Bark, Fur = Coarse, Color = Brown) =

$$\frac{(0.5)(3/4)(3/4)(1/2)}{(0.5)(3/4)(3/4)(1/2)(0.5)(1/4)(1/4)(1/2)}$$

$$\frac{(0.5)(1/4)(1/4)(0.5)}{(0.5)(3/4)(3/4)(1/2)(0.5)(1/4)(1/4)(1/2)}$$

0

Problem 5: Perceptron (15pt). Suppose you are given the following feature vectors: $x_1 = (1,0), x_2 = (1,2), x_3 = (0,-1), x_4 = (-1,-1), x_5 = (-2,1)$. Their corresponding labels are $y_1 = 1, y_2 = 1, y_3 = -1, y_4 = -1, y_5 = -1$. **Note: there is no bias term in this problem.** Suppose we run perceptron on this dataset starting with $w_0 = (0,0)$. Write down the values of w_1, w_2, w_3, w_4 and w_5 after each training instance, that is, w_i is the updated vector after the i-th training instance x_i.

No feature . . . .

$W_1 = (1,0)(1) + (0,-1)\cdot 1 + (0,-1)\cdot -1 + (-1,-1)\cdot -1 + (-2,1)\cdot -1$

$W_2 = (1,0)\cdot(1) + (0,-1)\cdot 1 + (0,1)\cdot -1 + (1,1)\cdot -1 + (2,-1)\cdot -1$

$W_3 = (1,0)(1) + (0,-1)\cdot 1 + (0,-1)\cdot -1 + (-1,-1)\cdot -1 + (-2,1)\cdot -1$

$W_4 = (1,0)\cdot 1 + (0,-1)\cdot 1 + (0,1)+1 + (1,1)\cdot -1 + (2,-1)\cdot -1$

$W_5 = (1,0)\cdot 1 + (0,-1)\cdot 1 \quad (0,-1)+1 \quad (-1,-1)+1 + (-2,1)\cdot -1$