c166f19.midterm- Requires Respondus LockDown Browser

Due No due date **Points** 200 **Questions** 24
Available Oct 24 at 11am - Oct 24 at 12:30pm about 2 hours **Time Limit** 70 Minutes
Requires Respondus LockDown Browser

This quiz was locked Oct 24 at 12:30pm.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	70 minutes	190 out of 200

Score for this quiz: **190** out of 200
Submitted Oct 24 at 12:18pm
This attempt took 70 minutes.

Question 1

5 / 5 pts

Given the Probability Distribution Table Below:

A	B	C	D	P(A, B, C, D)
0	0	0	0	0.4
0	0	0	1	0.1
0	0	1	0	0.2
0	0	1	1	0.02
0	1	0	0	0.02
0	1	0	1	0.02
0	1	1	0	0.02

0	1	1	1	0.02
1	0	0	0	0.04
1	0	0	1	0.03
1	0	1	0	0.03
1	0	1	1	0.02
1	1	0	0	0.03
1	1	0	1	0.02
1	1	1	0	0.02
1	1	1	1	0.01

Calculate the P(D=0|A=1,B=0,C=1)

Correct!

0.6

orrect Answers 0.6 (with margin: 0.05)

Question 2

5 / 5 pts

Given the Probability Distribution Table Below:

A	B	C	D	P(A, B, C, D)
0	0	0	0	0.4
0	0	0	1	0.1
0	0	1	0	0.2

0	0	1	1	0.02
0	1	0	0	0.02
0	1	0	1	0.02
0	1	1	0	0.02
0	1	1	1	0.02
1	0	0	0	0.04
1	0	0	1	0.03
1	0	1	0	0.03
1	0	1	1	0.02
1	1	0	0	0.03
1	1	0	1	0.02
1	1	1	0	0.02
1	1	1	1	0.01

Calculate the P(D=1|A=1,B=0,C=1)

Correct!

0.4

orrect Answers

0.4 (with margin: 0.05)

Question 3

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Calculate the value of states using Value Iteration algorithm for time step K=0, 1, 2, 3

Provide the value for State (1,1) at time step 2:

$V_2(1,1)$ =

Correct!

0.75

orrect Answers

Between 0.73 and 0.77

Question 4

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Calculate the value of states using Value Iteration algorithm for time step K=0, 1, 2, 3

Provide the value for State (2,1) at time step 2:

$V_2(2,1)$ =

Correct!

5

orrect Answers

Between 4.9 and 5.1

Question 5

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Calculate the value of states using Value Iteration algorithm for time step K=0, 1, 2, 3

Provide the value for State (3,1) at time step 2:

$V_2(3,1)$ =

Correct!

10

orrect Answers 10 (with margin: 0)

Question 6

5 / 5 pts

Given the MDP as described in question 1, after two iterations of the Value Iteration algorithm, what action would the Policy Extraction Algorithm assign to state (1,1)?

- Can't determine.
- Down

Correct!

- Right

Question 7

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Calculate the value of states using Value Iteration algorithm for time step K=0, 1, 2, 3

Provide the value for State (1,1) at time step 3:

$V_3(1,1)$ =

Correct!

4

orrect Answers

Between 3.9 and 4.1

Question 8

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Calculate the value of states using Value Iteration algorithm for time step K=0, 1, 2, 3

Provide the value for State (2,1) at time step 3:

$V_3(2,1)$ =

Correct!

5

orrect Answers

Between 4.9 and 5.1

Question 9

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Calculate the value of states using Value Iteration algorithm for time step K=0, 1, 2, 3

Provide the value for State (2,1) at time step 3:

$V_3(3,1)$ =

Correct!

10

orrect Answers

10 (with margin: 0)

Question 10

5 / 5 pts

Given the MDP as described in question 1, after three iterations of the Value Iteration algorithm, what action would the Policy Extraction Algorithm assign to state (1,1)?

- Right

Correct!

- Down
- Can't determine.

Question 11

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Using the policy of always going right in states (1,1) and (2,1), Calculate the value of states using this policy for time step K=0, 1, 2, 3

Provide the value for State (1,1) at time step 2:

$V_2(1,1)$ =

Correct!

0.75

orrect Answers

Between 0.73 and 0.77

Question 12

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Using the policy of always going right in states (1,1) and (2,1), Calculate the value of states using this policy for time step K=0, 1, 2, 3

Provide the value for State (2,1) at time step 2:

$V_2(2,1)$ =

Correct!

-5

orrect Answers

Between -4.9 and -5.1

Question 13

0 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Using the policy of always going right in states (1,1) and (2,1), Calculate the value of states using this policy for time step K=0, 1, 2, 3

Provide the value for State (1,1) at time step 3:

$V_3(1,1)$ =

'ou Answered

0.75

orrect Answers

Between -0.45 and -0.55

Question 14

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit = 1
(2, 1) Actions: down, right	(2, 2) Action: Exit = -10
(3, 1) Action: Exit = 10	

When taking action down, it is successful with probability 0.75, otherwise you go right.

When taking action right, it is successful with probability 0.75, otherwise you go down.

When taking action Exit, it is successful with probability 1.0.

The only reward is when taking action Exit, and there is no discounting.

Using the policy of always going right in states (1,1) and (2,1), Calculate the value of states using this policy for time step K=0, 1, 2, 3

Provide the value for State (2,1) at time step 3:

$V_3(2,1)$ =

'ou Answered

-5

orrect Answers

-4.9 (with margin: -5.1)

Question 15

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit
(2, 1) Actions: down, right	(2, 2) Action: Exit
(3, 1) Action: Exit	

This time you do not know the transition model or reward function. You observe the following episodes under policy π:

Episode 1: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 2: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 3: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (2,2)), ((2,2), Exit, -10, x))]

Episode 4: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 5: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 5: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 7: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 8: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Transitions are defined as (State, Action, Reward, New State).

Using our model based learning approach, what is P((1,2) | (1,1), down)?

Correct!

0.5

orrect Answers

0.5 (with margin: 0)

Question 16

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit
(2, 1) Actions: down, right	(2, 2) Action: Exit
(3, 1) Action: Exit	

This time you do not know the transition model or reward function. You observe the following episodes under policy π:

Episode 1: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 2: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 3: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (2,2)), ((2,2), Exit, -10, x))]

Episode 4: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 5: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 5: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 7: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 8: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Transitions are defined as (State, Action, Reward, New State).

Using our model based learning approach, what is P((2,2) | (2,1), down)?

Correct!

0.25

orrect Answers

0.25 (with margin: 0)

Question 17

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit
(2, 1) Actions: down, right	(2, 2) Action: Exit
(3, 1) Action: Exit	

This time you do not know the transition model or reward function. You observe the following episodes under policy π:

Episode 1: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 2: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 3: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (2,2)), ((2,2), Exit, -10, x))]

Episode 4: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 5: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 5: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 7: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 8: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Transitions are defined as (State, Action, Reward, New State).

Assume α=1.0, and γ=1.0, and initially all state are assumed to have a value of 0.

Using Temporal-Difference Learning, after seeing all of the episodes, what would the value estimate be for state (1,1).

Correct!

1

orrect Answers

1 (with margin: 0)

Question 18

10 / 10 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit

(2, 1) Actions: down, right	(2, 2) Action: Exit
(3, 1) Action: Exit	

This time you do not know the transition model or reward function. You observe the following episodes under policy π:

Episode 1: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 2: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 3: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (2,2)), ((2,2), Exit, -10, x))]

Episode 4: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 5: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 5: [(1,1), down, 0, (2,1)), ((2,1), down, 0, (3,1)), ((3,1), Exit, 10, x))]

Episode 7: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Episode 8: [(1,1), down, 0, (1,2)), ((1,2), Exit, 1, x)]

Transitions are defined as (State, Action, Reward, New State).

Assume α=1.0, and γ=1.0, and initially all state are assumed to have a value of 0.

Using Temporal-Difference Learning, after seeing all of the episodes, what would the value estimate be for state (2,1).

Correct!

10

orrect Answers 10 (with margin: 0)

Question 19

5 / 5 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit
(2, 1) Actions: down, right	(2, 2) Action: Exit
(3, 1) Action: Exit	

This time you do not know the transition model or reward function. In addition you are modeling the world using an approximate state represented by two features (f1, f2). Now it just so happens that f1=row, and f2 = col. Our weights for these features are initialized as 0. Q-state values are calculated as:

$$Q(s, a) = w_{1,a} * f1(s) + w_{2,a} * f2(s)$$

So there are different weights $w_{i,a}$ for each action a. For example:

$$Q((1,1), down) = w_{1,down} * f1(1,1) + w_{2,down} * f2(1,1)$$

This allows for different Q-values for each action in state.

Assume the following weights:

	Feature 1	Feature 2
right	-1	-1
down	1	1
exit	0	0

What is the approximate Q-value for q-state (1,1), right

- 0
- 2

None of these

Correct!

-2

Question 20

5 / 5 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit
(2, 1) Actions: down, right	(2, 2) Action: Exit
(3, 1) Action: Exit	

This time you do not know the transition model or reward function. In addition you are modeling the world using an approximate state represented by two features (f1, f2). Now it just so happens that f1=row, and f2 = col. Our weights for these features are initialized as 0. Q-state values are calculated as:

$$Q(s, a) = w_{1,a} * f1(s) + w_{2,a} * f2(s)$$

So there are different weights $w_{i,a}$ for each action a. For example:

$$Q((1,1), down) = w_{1,down} * f1(1,1) + w_{2,down} * f2(1,1)$$

This allows for different Q-values for each action in state.

Assume the following weights:

	Feature 1	Feature 2
right	-1	-1
down	1	1

exit	0	0

What is the approximate Q-value for q-state: (1,1), down

- None of these
- Correct! 2
- -2
- 0

Question 21

5 / 5 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit
(2, 1) Actions: down, right	(2, 2) Action: Exit
(3, 1) Action: Exit	

This time you do not know the transition model or reward function. In addition you are modeling the world using an approximate state represented by two features (f1, f2). Now it just so happens that f1=row, and f2 = col. Our weights for these features are initialized as 0. Q-state values are calculated as:

$$Q(s, a) = w_{1,a} * f1(s) + w_{2,a} * f2(s)$$

So there are different weights $w_{i,a}$ for each action a. For example:

$Q((1,1),\ down) = w_{1,down} * f1(1,1) + w_{2,down} * f2(1,1)$

This allows for different Q-values for each action in state.

Assume the following weights:

	Feature 1	Feature 2
right	-1	-1
down	1	1
exit	0	0

What is the approximate Q-value for q-state: (1,1), exit

Correct!

- (•) 0
- ( ) None of these
- ( ) 2
- ( ) -2

Question 22

5 / 5 pts

The following MDP world consists of 5 states and 3 actions:

(1, 1) Actions: down, right	(1, 2) Action: Exit
(2, 1) Actions: down, right	(2, 2) Action: Exit
(3, 1) Action: Exit	

This time you do not know the transition model or reward function. In addition you are modeling the world using an approximate state represented by two features (f1, f2). Now it just so happens that f1=row, and f2 = col. Our weights for these features are initialized as 0. Q-state values are calculated as:

$$Q(s, a) = w_{1,a} * f1(s) + w_{2,a} * f2(s)$$

So there are different weights $w_{i,a}$ for each action a. For example:

$$Q((1,1), down) = w_{1,down} * f1(1,1) + w_{2,down} * f2(1,1)$$

This allows for different Q-values for each action in state.

Assume the following weights:

	Feature 1	Feature 2
right	-1	-1
down	1	1
exit	0	0

What action would policy extraction return for state (1,1)

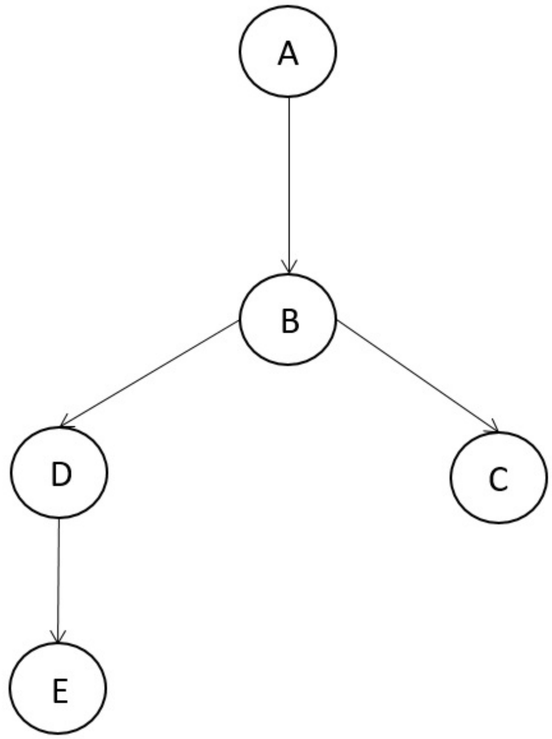
Correct!

- (selected) down
- right
- None of these
- exit

Question 23

10 / 10 pts

Assume all of the variables in the Bayesian Network below are Boolean (with values in the set {0,1}). Determine the new factor produced by Variable Elimination after the first variable is eliminated when ordering the variables alphabetically (A before B) to calculate the probability query: P(D|C=1, E=1)



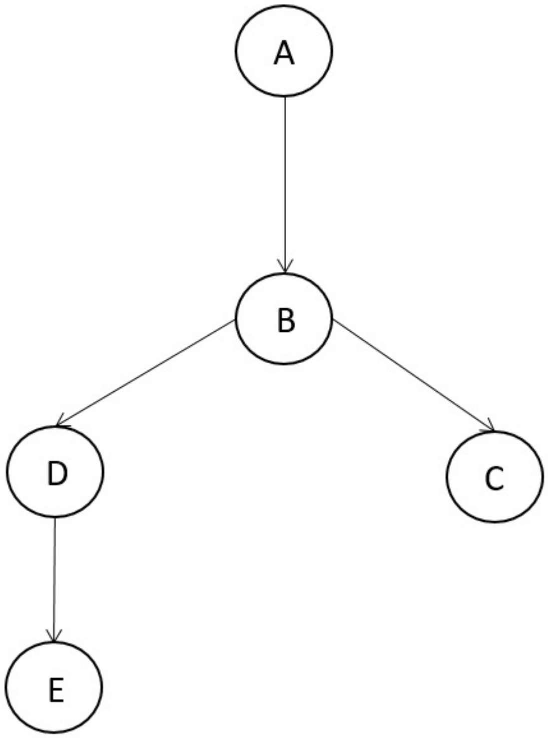
Select all the variables contained in the new factor after the first variable is eliminated:

- [ ] E
- [x] B — Correct!
- [ ] D
- [ ] A
- [ ] C

Question 24

10 / 10 pts

Assume all of the variables in the Bayesian Network below are Boolean (with values in the set {0,1}). Determine the new factor produced by Variable Elimination after the second variable is eliminated when ordering the variables alphabetically (A before B) to calculate the probability query: P(D|C=1, E=1)



Select all the variables contained in the new factor after the second variable is eliminated:

Correct!

- [x] D
- [ ] E

Correct!

- [x] C
- [ ] B
- [ ] A

Quiz Score: **190** out of 200