Midterm Examination

DATE: 9/3/25

Midterm Exam
Neil Adrian B. Baltar COM221

1. Define a Markov Decision Process (MDP). List it's key components. (5pts)

= Markov Decision Process describes an environment in Reinforcement Learning where the environment is fully observable. It aims to get the possible state and action in the environment. Its components are as follows:

- States
- State transition Probability
- Discount
- Actions
- Reward

2. What does it mean for a process to satisfy the Markov property? (5pts)

= When the current state process is enough to predict the future state process

3. Explain the difference between a policy and a value function. (5pts)

= Policy is the strategy that the agent adopts and value function tells the expected cumulative reward while taking action in a state for a given policy.

4. What is the role of the discount factor (γ) in a MDP? (5pts)

= Discount Factor determines the importance of future rewards compared to immediate rewards

• What happens when $\gamma = 0$ and when $\gamma \to 1$

If the discount factor is 0 then the agent will value immediate rewards. While if is 1, it will try to maximize future rewards.

VICTORY

NO.:
DATE:

5. Two State weather MDP. (15 pts)

Reward Function:

	Go out	Stay Inside
Sunny	+2	0
Rainy	+1	+3

State Transition Matrix

	Sunny	Rainy
Sunny	0.0	1.0
Rainy	1.0	0.0

Discount Factor is $\gamma = 0.5$. Agent uses Uniform Random Policy

(a.) Compute the average expected reward for Sunny. (2 pts)

$r\pi = 0.5 \times (2) + 0.5 \times (0) = 1 + 0 = \boxed{1}$

(b.) Compute the average expected reward for Rainy. (2 pts)

$r\pi = 0.5 \times (1) + 0.5 \times (3) = 0.5 + 1.5 = \boxed{2}$

(c.) Using the bellman expectation equation, solve for $v\pi$(Sunny) (5pts)

$V_1 = 1 + 0.5(0.0v_1 + 1.0v_2)$ $\quad V_2 = 2 + 0.5(1.0v_1 + 0.0v_2)$

$V_1 = 1 + 0v_1 + 0.5v_2$ $\quad V_2 = 2 + 0.5v_1 + 0v_2$

$V_1 - 0v_1 - 0.5v_2 = 1$ $\quad V_2 - 0v_2 - 0.5v_1 = 2$

$V_1 - 0.5v_2 = 1$ $\quad V_2 - 0.5v_1 = 2$

$V_1 = 1 + 0.5v_2$

$V_1 = 1 + 0.5(3.33)$ $\quad \boxed{v\pi(\text{Sunny}) = 2.665}$

$V_1 = 1 + 1.665$

(d.) Using the bellman expectation equation, solve for $v\pi$(~~[illegible]~~ Rainy) (6pts)

$V_1 - 0.5v_2 = 1$

$V_1 = 1 + 0.5v_2$

$V_2 - 0.5v_1 = 2$

$V_2 - 0.5(1 + 0.5v_2) = 2$

$V_2 - (0.5 \times 1)(0.5 \times 0.5v_2) = 2$

$V_2 - 0.25v_2 - 0.5 = 2$

$0.75v_2 - 0.5 = 2$

$\frac{0.75v_2}{0.75} = \frac{2 + 0.5}{0.75}$

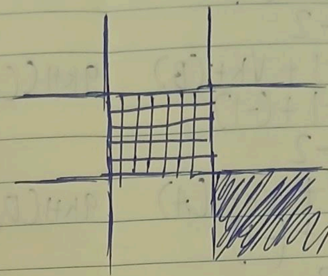
$V_2 = \frac{2 + 0.5}{0.75} = \frac{2.5}{0.75} = 3.33$

$\boxed{v\pi(\text{Rainy}) = 3.33}$

NO.:
DATE:

6 Gridworld MDP (15pts)

A	B	C
D	E	F
G	H	I



Discount Factor $\gamma = 1$

(a) Using dynamic programming, compute the optimal state-value function $V_*(S)$ for all non-terminal states (9pts)

	$V_k(S)$	$V_{k+1}(S)$
A	0	?
B	0	?
C	0	?
D	0	?
F	0	?
G	0	?
H	0	?

$V_{k+1}(A) = \frac{1}{4}[(-1+V_k(A)) + (-1+V_k(B)) + (-1+V_k(D)) + (-1+V_k(A))] = -1$

$V_{k+1}(B) = \frac{1}{4}[(-1+V_k(B)) + (-1+V_k(C)) + (-1+V_k(B)) + (-1+V_k(A))] = -1$

$V_{k+1}(C) = \frac{1}{4}[(-1+V_k(C)) + (-1+V_k(C)) + (-1+V_k(F)) + (-1+V_k(B))] = -1$

$V_{k+1}(D) = \frac{1}{4}[(-1+V_k(A)) + (-1+V_k(D)) + (-1+V_k(G)) + (-1+V_k(D))] = -1$

$V_{k+1}(F) = \frac{1}{4}[(-1+V_k(C)) + (-1+V_k(F)) + (-1+V_k(I)) + (-1+V_k(F))] = -1$

$V_{k+1}(G) = \frac{1}{4}[(-1+V_k(D)) + (-1+V_k(H)) + (-1+V_k(G)) + (-1+V_k(G))] = -1$

$V_{k+1}(H) = \frac{1}{4}[(-1+V_k(H)) + (-1+V_k(I)) + (-1+V_k(H)) + (-1+V_k(G))] = -1$

NO.:
DATE:

$q_{k+1}(A, LEFT) = -1 + V_{k+1}(A)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(A, RIGHT) = -1 + V_{k+1}(B)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(A, UP) = -1 + V_{k+1}(A)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(A, DOWN) = -1 + V_{k+1}(D)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(B, LEFT) = -1 + V_{k+1}(A)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(B, RIGHT) = -1 + V_{k+1}(C)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(B, UP) = -1 + V_{k+1}(B)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(B, DOWN) = -1 + V_{k+1}(B)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(C, LEFT) = -1 + V_{k+1}(B)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(C, RIGHT) = -1 + V_{k+1}(C)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(C, UP) = -1 + V_{k+1}(C)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(C, DOWN) = -1 + V_{k+1}(F)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(D, LEFT) = -1 + V_{k+1}(D)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(D, RIGHT) = -1 + V_{k+1}(D)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(D, UP) = -1 + V_{k+1}(A)$
$= -1 + \cancel{V_{k+1}}(-1)$
$= -2$

$q_{k+1}(D, DOWN) = -1 + V_{k+1}(G)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(F, LEFT) = -1 + V_{k+1}(F)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(F, RIGHT) = -1 + V_{k+1}(F)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(F, UP) = -1 + V_{k+1}(C)$
$= -1 + (-1)$
$= -2$

$q_{k+1}(F, DOWN) = -1 + V_{k+1}(I)$
$= -1 + 0$
$= -1$

$$q_{k+1}(G, LEFT) = -1 + V_{k+1}(G) = -1 + (-1) = -2$$

$$q_{k+1}(G, RIGHT) = -1 + V_{k+1}(H) = -1 + (-1) = -2$$

$$q_{k+1}(G, UP) = -1 + V_{k+1}(D) = -1 + (-1) = -2$$

$$q_{k+1}(G, DOWN) = -1 + V_{k+1}(G) = -1 + (-1) = -2$$

$$q_{k+1}(H, LEFT) = -1 + V_{k+1}(G) = -1 + (-1) = -2$$

$$q_{k+1}(H, RIGHT) = -1 + V_{k+1}(I) = -1 + 0 = -1$$

$$q_{k+1}(H, UP) = -1 + V_{k+1}(H) = -1 + (-1) = -2$$

$$q_{k+1}(H, DOWN) = -1 + V_{k+1}(H) = -1 + (-1) = -2$$

$\pi_{k+1}(A) = \{LEFT, RIGHT, UP, DOWN\}$

$\pi_{k+1}(B) = \{LEFT, RIGHT, UP, DOWN\}$

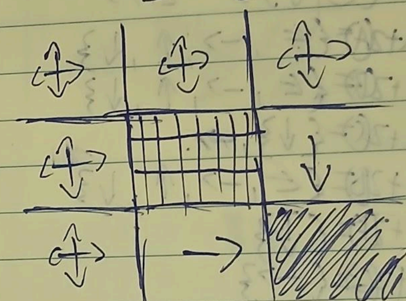
$\pi_{k+1}(C) = (LEFT, RIGHT, UP, DOWN\}$

$\pi_{k+1}(D) = \{LEFT, RIGHT, UP, DOWN\}$

$\pi_{k+1}(F) = \{DOWN\}$

$\pi_{k+1}(G) = \{LEFT, RIGHT, UP, DOW\}$

$\pi_{k+1}(H) = \{RIGHT\}$



	$V_k(S)$	$V_{k+1}(S)$	$V_{k+2}(S)$
A	0	-1	?
B	0	-1	?
C	0	-1	?
D	0	-1	?
F	0	-1	?
G	0	-1	?
H	0	-1	?

$V_{k+2}(A) = -2$

$V_{k+2}(B) = -2$

$V_{k+2}(C) = -2$

$V_{k+2}(D) = -2$

$V_{k+2}(F) = -1.75$

$V_{k+2}(G) = -2$

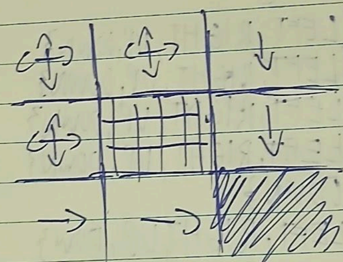
$V_{k+2}(H) = -1.75$

NO.:
DATE:

$q_{k+2}(A, \leftarrow) = -3$
$q_{k+2}(A, \rightarrow) = -3$
$q_{k+2}(A, \uparrow) = -3$
$q_{k+2}(A, \downarrow) = -3$
$q_{k+2}(B, \leftarrow) = -3$
$q_{k+2}(B, \rightarrow) = -3$
$q_{k+2}(B, \uparrow) = -3$
$q_{k+2}(B, \downarrow) = -3$
$q_{k+2}(C, \leftarrow) = -3$
$q_{k+2}(C, \rightarrow) = -3$
$q_{k+2}(C, \uparrow) = -3$
$q_{k+2}(C, \downarrow) = -2.75$

$q_{k+2}(D, \leftarrow) = -3$
$q_{k+2}(D, \rightarrow) = -3$
$q_{k+2}(D, \uparrow) = -3$
$q_{k+2}(D, \downarrow) = -3$
$q_{k+2}(G, \leftarrow) = -3$
$q_{k+2}(G, \rightarrow) = -2.75$
$q_{k+2}(G, \uparrow) = -3$
$q_{k+2}(G, \downarrow) = -3$

$\pi_{k+2}(A) = \{\leftarrow, \rightarrow, \uparrow, \downarrow\}$
$\pi_{k+2}(B) = \{\leftarrow, \rightarrow, \uparrow, \downarrow\}$
$\pi_{k+2}(C) = \{\downarrow\}$
$\pi_{k+2}(D) = \{\leftarrow, \rightarrow, \uparrow, \downarrow\}$
$\pi_{k+2}(F) = \{\downarrow\}$
$\pi_{k+2}(G) = \{\rightarrow\}$
$\pi_{k+2}(H) = \{\rightarrow\}$



	$V_k(S)$	$V_{k+1}(S)$	$V_{k+2}(S)$	$V_{k+3}(S)$
A	0	−1	−2	?
B	0	−1	−2	?
C	0	−1	−2	?
D	0	−1	−2	?
F	0	−1	−1.75	?
G	0	−1	−2	?
H	0	−1	−1.75	?

$V_{k+3}(A) = -3$
$V_{k+3}(B) = -3$
$V_{k+3}(C) = -2.9375$
$V_{k+3}(D) = -3$
$V_{k+3}(F) = -2.75$
$V_{k+3}(G) = -2.9375$
$V_{k+3}(H) = -2.75$

$q_{k+4}(A, \leftarrow) = -4$

$q_{k+4}(A, \rightarrow) = -4$

$q_{k+4}(A, \uparrow) = -4$

$q_{k+4}(A, \downarrow) = -4$

$q_{k+4}(B, \leftarrow) = -4$

$q_{k+4}(B, \rightarrow) = -3.9375$

$q_{k+4}(B, \uparrow) = -4$

$q_{k+4}(B, \downarrow) = -4$

$q_{k+4}(D, \leftarrow) = -4$

$q_{k+4}(D, \rightarrow) = -4$

$q_{k+4}(D, \uparrow) = -4$

$q_{k+4}(D, \downarrow) = -3.9375$

$\pi_{k+4}(A) = \leftarrow, \rightarrow, \uparrow, \downarrow$

$\pi_{k+4}(B) = \rightarrow$

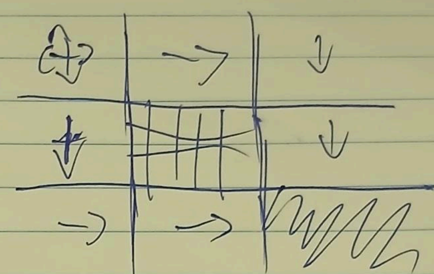
$\pi_{k+4}(C) = \downarrow$

$\pi_{k+4}(D) = \downarrow$

$\pi_{k+4}(F) = \downarrow$

$\pi_{k+4}(G) = \rightarrow$

$\pi_{k+4}(H) = \rightarrow$



	$V_k(S)$	$V_{k+1}(S)$	$V_{k+2}(S)$	$V_{k+3}(S)$	$V_{k+4}(S)$
A	0	-1	-2	-3	?
B	0	-1	-2	-3	?
C	0	-1	-2	-2.9375	?
D	0	-1	-2	-3	?
F	0	-1	-1.75	-2.75	?
G	0	-1	-2	-2.9375	?
H	0	-1	-1.75	-2.75	

$V_{k+5}(A) = -4$

$V_{k+5}(B) =$

$V_{k+5}(C) =$

$V_{k+5}(D) =$

$V_{k+5}(F) =$

$V_{k+5}(G) =$

$V_{k+5}(H) =$