DIANA NICOLE DANGA

COM221

1.) Define a Markov Decision Process (MDP). List its key components (5 pts.)

↳ A Markov Decision Process is a mathematical tool that is used for decision-making. It describes an environment for reinforcement learning where the "environment" is fully observable. It revolves around the idea that the future is independent of the past given the present.

Key components:

- States
- actions
- policies
- Rewards
- Discount factors

2.) What does it mean for a process to satisfy the Markov Property?

↳ The process satisfies the Markov property if the future state depends on the present state, and not the past actions that led to it, because past is irrelevant to the future, all that matters is the present that holds all relevant information.

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

↳ The policy defines the behavior of the agent. It says the actions that agent can do. Meanwhile, the policies is associated value functions, which is all about the value returned upon following a policy.

/ /

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

↳ to determine the agent's preference for immediate reward over future ones. When $\gamma = 0$, the agent ignores the future ones and only focus to have the immediates ones, while $\gamma = 1$ it will equally give importance to all future rewards together with the immediate rewards.

5.) Two-State weather MDP

	Go out	stay inside
sunny	2	0
cloudy	1	3

a.) Compute the average expected reward for sunny

$r_\pi = 0.5\,(2) + 0.5\,(0) = 0.1 + 0 = 1$

b.) Compute the average expected reward for Rainy

$r_\pi = 0.5\,(1) + 0.5\,(3) = 0.5 + 1.5 = 2$

$r_\pi = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

c.) using the bellman expectation equation, solve for v_π (Sunny)

$v_1 = 1 + 0.5\,(0.0v_1 + 1.0v_2)$

$v_1 = 1 + 0.0v_1 + 0.5v_2$

$v_1 = -0.0v_1 - 0.5v_2 = 1$

$\boxed{1v_1 - 0.5v_2 = 1}$

d.) using the bellman expectation equation, solve for v_π (cloudy)

$v_2 = 2 + 0.5\,(1.0v_1 + 0.0v_2)$

$v_2 = 2 + 0.5v_1 + 0.0v_2$

$v_2 = -0.5v_1 - 0.0v_2 = 2$

$\boxed{-0.5v_1 - 1v_2 = 2}$

$$1v_1 - 0.5v_2 = 1$$
$$v_1 = 1 + 0.5v_2$$

$$-0.5(1 + 0.5v_2) + 1v_2 = 2$$
$$-(0.5 \times 1)(-0.5 \times 0.5v_2) + 1v_2 = 2$$
$$-0.5 - 0.25v_2 + 1v_2 = 2$$
$$-0.5 + (1v_2 - 0.25v_2) = 2$$
$$-0.5 + 0.75v_2 = 2$$

$$\boxed{V_\pi(\text{Rainy}) = \frac{0.75\cancel{v_2}}{0.\cancel{7}5v_2} = \frac{2 + 0.5}{0.75} = \frac{2.5}{0.75} = 3.33}$$

$$v_1 = 1 + 0.5(3.33)$$
$$= 1 + 1.665$$
$$\boxed{V_\pi(\text{sunny}) = 2.665}$$

6.) Consider the following gridworld (MDP):

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

A	B	C
D	wall	F
G	H	TS

0	0	0
0	x	0
0	0	

a.) Step 1:

i.) $V_{k+1}(A) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+0)} + \overset{\rightarrow}{(-1+0)} + \overset{\uparrow}{(-1+0)} + \overset{\downarrow}{(-1+0)}\right]$

$= \frac{1}{4}[-1 -1 -1 -1]$

$= -1$

2) $U_{k+1}(B) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+0)} + \overset{\rightarrow}{(-1+0)} + \overset{\uparrow}{(-1+0)} + \overset{\downarrow}{(-1+0)}\right]$

$= \frac{1}{4}[-1-1-1-1]$

$= -1$

3) $U_{k+1}(C) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+0)} + \overset{\rightarrow}{(-1+0)} + \overset{\uparrow}{(-1+0)} + \overset{\downarrow}{(-1+0)}\right]$

$= \frac{1}{4}[-1-1-1-1]$

$= -1$

4) $U_{k+1}(D) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+0)} + \overset{\rightarrow}{(-1+0)} + \overset{\uparrow}{(-1+0)} + \overset{\downarrow}{(-1+0)}\right]$

$= \frac{1}{4}[-1-1-1-1]$

$= -1$

5) $U_{k+1}(F) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+0)} + \overset{\rightarrow}{(-1+0)} + \overset{\uparrow}{(-1+0)} + \overset{\downarrow}{(-1+0)}\right]$

$= \frac{1}{4}[-1-1-1+0]$

$= -1$

6) $U_{k+1}(G) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+0)} + \overset{\rightarrow}{(-1+0)} + \overset{\uparrow}{(-1+0)} + \overset{\downarrow}{(-1+0)}\right]$

$= \frac{1}{4}[-1-1-1-1]$

$= -1$

7) $U_{k+1}(H) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+0)} + \overset{\rightarrow}{(-1+0)} + \overset{\uparrow}{(-1+0)} + \overset{\downarrow}{(-1+0)}\right]$

$= \frac{1}{4}[-1-1-1-1]$

$= -1$

-1	-1	-1
-1	x	-1
-1	-1	

Step 2:

A) $q(A,L) = -1 + (-1)$

$= -2$

$q(A,U) = -1 + (-1)$

$= -2$

$q(A,R) = -1 + (-1)$

$= -2$

$q(A,D) = -1 + (-1)$

$= -2$

$\pi_{k+1}(A) = L, R, U, D$

B.) $q(B,L) = -1 + (-1) = -2$ $\quad$ $q(B,U) = -1 + (-1) = -2$ $\quad$ $\pi_{k+1}(B) = L, R, U, D$

$q(B,R) = -1 + (-1) = -2$ $\quad$ $q(B,D) = -1 + (-1) = -2$

C.) $q(C,L) = -1 + (-1) = -2$ $\quad$ $q(C,U) = -1 + (-1) = -2$ $\quad$ $\pi_{k+1}(C) = L, R, U, D$

$q(C,R) = -1 + (-1) = -2$ $\quad$ $q(C,D) = -1 + (-1) = -2$

D.) $q(D,L) = -1 + (-1) = -2$ $\quad$ $q(D,U) = -1 + (-1) = -2$ $\quad$ $\pi_{k+1}(D) = L, R, U, D$

$q(D,R) = -1 + (-1) = -2$ $\quad$ $q(D,D) = -1 + (-1) = -2$

F) $q(F,L) = -1 + (-1) = -2$ $\quad$ $q(F,U) = -1 + (-1) = -2$ $\quad$ $\pi_{k+1}(F) = D$

$q(F,R) = -1 + (-1) = -2$ $\quad$ $q(F,D) = -1 + (0) = -1$

G.) $q(G,L) = -1 + (-1) = -2$ $\quad$ $q(G,U) = -1 + (-1) = -2$ $\quad$ $\pi_{k+1}(G) = L, R, U, D$

$q(G,R) = -1 + (-1) = -2$ $\quad$ $q(G,D) = -1 + (-1) = -2$

H) $q(H,L) = -1 + (-1)$
$= -2$

$q(H,U) = -1 + (-1)$
$= -2$

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

$q(H,R) = -1 + (0)$
$= -1$

$q(H,D) = -1 + (-1)$
$= -2$

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

↑←↓→	↑←↓→	↑←↓→
↑←↓→	Wall	↓
↑←↓→	→	0

Step 1:

1.) $V_{k+2}(A) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-1))} + \overset{\rightarrow}{(-1+(-1))} + \overset{\uparrow}{(-1+(-1))} + \overset{\downarrow}{(-1+(-1))}\right]$
$= \frac{1}{4}[-2-2-2-2]$
$= -2$

2.) $V_{k+2}(B) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-1))} + \overset{\rightarrow}{(-1+(-1))} + \overset{\uparrow}{(-1+(-1))} + \overset{\downarrow}{(-1+(-1))}\right]$
$= \frac{1}{4}[-2-2-2-2]$
$= -2$

3.) $V_{k+2}(C) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-1))} + \overset{\rightarrow}{(-1+(-1))} + \overset{\uparrow}{(-1+(-1))} + \overset{\downarrow}{(-1+(-1))}\right]$
$= \frac{1}{4}[-2-2-2-2]$
$= -2$

4) $V_{k+2}(D) = \frac{1}{4}\left[(\overset{\leftarrow}{-1}+(-1)) + (\overset{\rightarrow}{-1}+(-1)) + (\overset{\uparrow}{-1}+(-1)) + (\overset{\downarrow}{-1}+(-1))\right]$

$= \frac{1}{4}\left[-2-2-2-2\right]$

$= -2$

5.) $V_{k+2}(F) = \frac{1}{4}\left[(\overset{\leftarrow}{-1}+(-1)) + (\overset{\rightarrow}{-1}+(-1)) + (\overset{\uparrow}{-1}+(-1)) + (\overset{\downarrow}{-1}+(0))\right]$

$= \frac{1}{4}\left[-2-2-2-1\right]$

$= -1.75$

6.) $V_{k+2}(G) = \frac{1}{4}\left[(\overset{\leftarrow}{-1}+(-1)) + (\overset{\rightarrow}{-1}+(-1)) + (\overset{\uparrow}{-1}+(-1)) + (\overset{\downarrow}{-1}+(-))\right]$

$= \frac{1}{4}\left[-2-2-2-2\right]$

$= -2$

7.) $V_{k+2}(H) = \frac{1}{4}\left[(\overset{\leftarrow}{-1}+(-1)) + (\overset{\rightarrow}{-1}+(0)) + (\overset{\uparrow}{-1}+(-1)) + (\overset{\downarrow}{-1}+(-1))\right]$

$= \frac{1}{4}\left[-2-1-2-2\right]$

$= -1.75$

	$V_k(s)$	$V_{k+1}(s)$	$V_{k+2}(s)$
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

/ /

Step 2:

A.) $q(A,L) = -1 + (-2) = -3$ $\quad q(A,U) = -1 + (-2) = -3$

$q(A,R) = -1 + (-2) = -3$ $\quad q(A,D) = -1 + (-2) = -3$

$\pi_{k+1}(A) = L, R, U, D$

B.) $q(B,L) = -1 + (-2) = -3$ $\quad q(B,U) = -1 + (-2) = -3$

$q(B,R) = -1 + (-2) = -3$ $\quad q(B,D) = -1 + (-2) = -3$

$\pi_{k+1}(B) = L, R, U, D$

C.) $q(C,L) = -1 + (-2) = -3$ $\quad q(C,U) = -1 + (-2) = -3$

$q(C,R) = -1 + (-2) = -3$ $\quad q(C,D) = -1 + (-1.75) = -2.75$

$\pi_{k+1}(C) = D$

D.) $q(D,L) = -1 + (-2) = -3$ $\quad q(D,U) = -1 + (-2) = -3$

$q(D,R) = -1 + (-2) = -3$ $\quad q(D,D) = -1 + (-2) = -3$

$\pi_{k+1}(D) = L, R, U, D$

G.) $q(G,L) = -1 + (-2) = -3$ $\quad q(G,U) = -1 + (-2) = -3$

$q(G,R) = -1 + (-1.75) = -2.75$ $\quad q(G,D) = -1 + (-2) = -3$

$\pi_{k+1}(G) = R$

↑ ← ↓ →	↑ ← ↓ →	↓
↑ ← ↓ →	wall	↓
→	→	0

Step 1:

1.) $V_{k+3}(A) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-2))} + \overset{\rightarrow}{(-1+(-2))} + \overset{\uparrow}{(-1+(-2))} + \overset{\downarrow}{(-1+(-2))}\right]$

$= \frac{1}{4}\left[-3-3-3-3\right]$

$= -3$

2.) $V_{k+3}(B) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-2))} + \overset{\rightarrow}{(-1+(-2))} + \overset{\uparrow}{(-1+(-2))} + \overset{\downarrow}{(-1+(-2))}\right]$

$= \frac{1}{4}\left[-3-3-3-3\right]$

$= -3$

3.) $V_{k+3}(C) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-2))} + \overset{\rightarrow}{(-1+(-2))} + \overset{\uparrow}{(-1+(-2))} + \overset{\downarrow}{(-1+(-1.75))}\right]$

$= \frac{1}{4}\left[-3-3-3-2.75\right.$

$= -2.94$

4.) $V_{k+3}(D) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-2))} + \overset{\rightarrow}{(-1+(-2))} + \overset{\uparrow}{(-1+(-2))} + \overset{\downarrow}{(-1+(-2))}\right]$

$= \frac{1}{4}\left[-3-3-3-3\right]$

$= -3$

5.) $V_{k+3}(F) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-1.75))} + \overset{\rightarrow}{(-1+(-1.75))} + \overset{\uparrow}{(-1+(-2))} + \overset{\downarrow}{(-1+(0))}\right]$

$= \frac{1}{4}\left[-2.75-2.75-3-1\right.$

$= -2.38$

6.) $V_{k+3}(b) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-2))} + \overset{\rightarrow}{(-1+(-1.75))} + \overset{\uparrow}{(-1+(-2))} + \overset{\downarrow}{(-1+(-2))}\right]$

$= \frac{1}{4}[-3 - 2.75 - 3 - 3]$

$= -2.94$

7. $V_{k+3}(H) = \frac{1}{4}\left[(-1+(-2)) + (-1+0) + (-1+(-1.75)) + (-1+(-1.75))\right]$

$= \frac{1}{4}(-3 + -1 - 2.75 - 2.75$

$= -2.38$

	$V_k(s)$	$V_{k+1}(s)$	$V_{k+2}(s)$	$V_{k+3}(s)$
A	0	-1	-2	-3
B	0	-1	-2	-3
C	0	-1	-2	-2.94
D	0	-1	-2	-3
F	0	-1	-1.75	-2.38
b	0	-1	-2	-2.94
H	0	-1	-1.75	-2.38

A B C
D F
b H

Step 2:

A) $q(A,L) = -1 + (-3)$
$= -4$

$q(A,U) = -1 + (-3)$
$= -4$

$q(A,R) = -1 + (-3)$
$= -4$

$q(A,D) = -1 + (-2)$
$= -4$

$\pi_{k+1}(A) = L, R, U, D$

A B C
D F
G H

B.) $q(B,L) = -1 + (-3)$
$= -4$

$q(B,U) = -1 + (-3)$
$= -4$

$\pi_{k+1}(B) = R$

$q(B,R) = -1 + (-2.94)$
$= -3.94$

$q(B,D) = -1 + (-3)$
$= -4$

D.) $q(D,L) = -1 + (-3)$
$= -4$

$q(D,U) = -1 + (-3)$
$= -4$

$\pi_{k+1}(D) = D$

$q(D,R) = -1 + (-3)$
$= -4$

$q(D,D) = -1 + (-2.94)$
$= -3.94$

A.)

	$V_k(s)$	$V_{k+1}(s)$	$V_{k+2}(s)$	$V_{k+3}(s)$	$V_{k+4}(s)$
A	0	-1	-2	-3	
B	0	-1	-2	-3	
C	0	-1	-2	-2.94	
D	0	-1	-2	-3	
F	0	-1	-1.75	-2.38	
G	0	-1	-2	-2.94	
H	0	-1	-1.75	-2.38	

B.)

↑←↓→	→	↓
↓	Wall	↓
→	→	0

Step 1:

1.) $V_{k+2}(A) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-3))} + \overset{\rightarrow}{(-1+(-3))} + \overset{\uparrow}{(-1+(-3))} + \overset{\downarrow}{(-1+(-3))}\right]$
$= \frac{1}{4}[-4-4-4-4]$
$= -4$

2.) $V_{k+2}(B) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-3))} + \overset{\rightarrow}{(-1+(-2.94))} + \overset{\uparrow}{(-1+(-3))} + \overset{\downarrow}{(-1+(-3))}\right]$
$= \frac{1}{4}[-4-3.94-4-4]$
$= -3.99$

3.) $V_{k+1}(C) = \frac{1}{4}\left[\overset{\leftarrow}{(-1+(-3))} + \overset{\rightarrow}{(-1+(-2.94))} + \overset{\uparrow}{(-1+(-2.94))} + (-1+(-2.38))\right]$

$= \frac{1}{4}\left[-4 - 3.94 - 3.94 - 3.38\right]$

$= -3.82$