NO:
DATE:

Exam:

1.) A markov decision process is a markov reward process with decisions. It is an environment in which all states are markov decisions.

A ~~Markov~~ MDP is a ~~turple~~ tuple $\{S, A, P, R, \gamma\}$

S is a finite set of states
A is a finite set of actions
P is a state transition probability matrix, $P^a_{ss'} = P[S_{t+1} = s' \mid S_t = s, A_t = a]$
R is a reward function : $R^a_s = E[R_{t+1} \mid S_t = s, A_t = a]$
γ is a discount factor $\gamma \in [0, 1]$

2.) The process to satisfy the Markov Property is the future is independent of the past given ~~state~~ the present

a state S_t is Markov if and only if

$$P[S_{t+1} \mid S_t] = P[S_{t+1} \mid S_1, \ldots, S_t]$$

3.) Policy is a distribution over action given states meaning it fully defines the behaviour of the agent while value action tell us if in state how much total (discounted) reward can is expect in the future. ~~The~~ It gives the long term value of states.

VALIANT®

NO:
DATE:

4.) The discount is the present value of future rewards. This values immediate reward above delayed reward
γ close to 0 leads to "myopic" evaluation
γ close to 1 leads to far sighted evaluation

Sometimes its possible to use undiscounted Markov Reward processes $\gamma = 1$ if all sequences terminate

5. States {Sunny, Cloudy}
Actions {Go out, Stay inside}
Discount = 0.5

	Go Out	Stay Inside
☀	+2	0
☁	+1	+3

$$P = \begin{bmatrix} 0.0 & 1.0 \\ 1.0 & 0.0 \end{bmatrix}$$

R go out $\begin{bmatrix} +2 \\ +1 \end{bmatrix}$ R stay inside $\begin{bmatrix} 0 \\ +3 \end{bmatrix}$

a.) If Sunny
$r\pi = 0.5 \times (2) + 0.5 \times (0) = 1 + 0 = 1$

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

b.) If cloudy
$r\pi = 0.5 \times (1) + 0.5 \times (3) = 0.5 + 1.5 = 2$

VALIANT®

NO:
DATE:

c. V_1 (Sunny)

$V_1 = 1 + 0.5\,(0V_1 + 1V_2)$

$V_1 = 1 + 0V_1 + 0.5V_2$

$V_1 = -0V_1 - 0.5V_2 = 1$

$$\boxed{1V_1 - 0.5V_2 = 1}$$

d. V_2 (Cloudy)

$V_2 = 2 + 0.5\,(1V_1 + 0V_2)$

$V_2 = 2 + 0.5V_1 + 0$

$V_2 = -0.5V_1 - 0V_2 = 2$

$$\boxed{-0.5V_1 + 1V_2 = 2}$$

$1V_1 = 1 + 0.5V_2$

$$\frac{1V_1}{1} = \frac{1 + 0.5V_2}{1}$$

$$V_1 = \frac{1 + 0.5V_2}{1}$$

$$-0.5\left(\frac{1 + 0.5V_2}{1}\right) + 0.1 = 2$$

$$\left(-0.5 \times \frac{1}{1}\right)\left(-0.5 \times \frac{0.5}{1}V_2\right) + 1 = 2$$

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

$$-0.5 - 0.25 + 1V_2 = 2$$

$$-0.5 + (1 - 0.25)V_2 = 2$$

$$-0.5 + 0.75V_2 = 2$$

$$0.75V_2 = 2 + 0.5$$

$$\boxed{V_2\,(\text{cloudy}) = \frac{2 + 0.5}{0.75} = \frac{2.5}{0.75} = 3.33}$$

$V_1 = \frac{1 + 0.5\, v_2}{1}$

$V_1 = \frac{1 + 0.5\,(3.33)}{1} = \frac{1 + 1.66}{1} = \frac{2.66}{1}$

V_1 (Sum) =	2.66

6.

A	B	C
D	E	F
G	H	I

States = (A - I)

Actions = (UP, DOWN, LEFT, RIGHT)

Policy = 0.25

Reward = -1

Discount = $\gamma = 1$.

Step 1:

1.) $V_{k+1}(A) = 1/4\,[(-1 + V(A)) + (-1 + V(B)) + (-1 + V(A)) + (-1 + V(A))]$

$V_{k+1}(A) = 1/4\,[(-1+0) + (-1+0) + (-1+0) + (-1+0)$

$V_{k+1}(A) = [(-1 + (-1)), (-1 + (-1)), (-1 + (-1)), (-1 + (-1))] = -1$

2.) $V_{k+1}(B) = 1/4\,[(-1 + V(A)) + (-1 + V(C)) + (-1 + V(B)) + (-1 + V(B))]$

$V_{k+1}(B) = 1/4\,[(-1+0) + (-1+0) + (-1+0)$

$V_{k+1}(B) = 1/4\,[(-1 + (-1)), (-1 + (-1)), (-1 + (-1)), (-1 + (-1)) = -1$

VALIANT

NO:
DATE:

3.) $V_{k+1}(C) = 1/4\,[(-1 + V(B)) + (-1 + V(C)) + (-1 + V(E)) + (-1 + V(C))$
$V_{k+1}(C) = 1/4\,[(-1+0) + (-1+0) + (-1+0) + (-1+0)$
$V_{k+1}(C) = 1/4\,[(-1+(-1)) + (-1+(-1)) + (-1+(-1)) + (-1+(-1))] = -1$

4.) $V_{k+1}(D) = 1/4\,[(-1 + V(D)) + (-1 + V(D)) + (-1 + V(G)) + (-1 + V(A))$
$V_{k+1}(D) = 1/4\,[(-1+0), (-1+0), (-1+0), (-1+0)$
$V_{k+1}(D) = 1/4\,[(-1+(-1)) + (-1+(-1)) + (-1+(-1)) + (-1+(-1)) = -1$

5. $V_{k+1}(F) = 1/4\,[(-1 + V(F)) + (-1 + V(F)) + (-1 + V(\blacksquare)) + (-1 + V(G))$
$V_{k+1}(F) = 1/4\,[(-1+0), (-1+0), (-1+0), (-1+0)$
$V_{k+1}(F) = 1/4\,[(-1+(-1)) + (-1+(-1)) + (-1+(-1)) + (-1+(-1)] = -1$

6. $V_{k+1}(G) = 1/4\,[(-1 + V(G)) + (-1 + V(H)) + (-1 + V(G)) + (-1 + V(D))$
$V_{k+1}(G) = 1/4\,[(-1+0) + (-1+0) + (-1+0) + (-1+0)$
$V_{k+1}(G) = 1/4\,[(-1+(-1)) + (-1+(-1)) + (-1+(-1)) + (-1+(-1)] = -1$

7. $V_{k+1}(H) = 1/4\,[(-1 + V(G)) + (-1 + V(1)) + (-1 + V(H)) + (-1 + V(H))$
$V_{k+1}(H) = 1/4\,[(-1+0) + (-1+0) + (-1+0) + (-1+0)$
$V_{k+1}(H) = 1/4\,[(-1+(-1)) + (-1+(-1)) + (-1+(-1)) + (-1+(-1)] = -1$

8. $q_{k+1}(A, \text{LEFT})$
$= -1 + V(A)$
$= -1 + (-1)$
$q_{k+1}(A, \text{LEFT}) = -2$

NO:
DATE:

9.) $q_{k+1}(A, \text{RIGHT})$
$= -1 + V(B)$
$= -1 + (-1)$
$q_{k+1}(A, \text{RIGHT}) = -2$

10.) $q_{k+1}(A, \text{UP})$
$= -1 + (V(A)$
$= -1 + (-1)$
$q_{k+1}(A, \text{UP}) = -2$

11.) $q_{k+1}(A, \text{Down})$
$= -1 + V(D)$
$= -1 + (-1)$
$q_{k+1}(A, \text{Down}) = -2$

12.) $q_{k+1}(B, \text{LEFT})$
$= -1 + V(A)$
$= -1 + (-1)$
$q_{k+1}(B, \text{LEFT}) = -2$

13.) $q_{k+1}(B, \text{RIGHT})$
$= -1 + (V(C)$
$= -1 + (-1)$
$q_{k+1}(B, \text{RIGHT}) = -2$

14.) $q_{k+1}(B, \text{UP})$
$= -1 + V(B)$
$= -1 + (-1)$
$q_{k+1}(B, \text{Down}) = -2$

15. $q_{k+1}(B, \text{Down})$
$= -1 + V(B)$
$= -1 + (-1)$
$q_{k+1}(B, \text{Down}) = -2$

16.) $q_{k+1}(C, \text{LEFT})$
$= -1 + V(B)$
$= -1 + (-1)$
$q_{k+1}(C, \text{LEFT}) = -2$

17.) $q_{k+1}(C, \text{RIGHT})$
$= -1 + V(C)$
$= -1 + (-1)$
$q_{k+1}(C, \text{RIGHT}) = -2$

18. $q_{k+1}(C, \text{UP})$
$= -1 + V(C)$
$= -1 + (-1)$
$= -2$

19. $q_{k+1}(C, \text{DOWN})$
$= 1 + V(F)$
$= -1 + (-1)$
$= -2$

20. $q_{k+1}(\quad, \text{LEFT})$
$= 1 + V(D)$
$= -1 + (-1)$
~~$q_{k+1}(\text{DICE}$~~ $= -2$

21. $q_{k+1}(D, \text{RIGHT})$
$= -1 + V(D)$
$= 1 + (-1)$
$= -2$

22.) $q_{k+1}(D, \text{UP})$
$= -1 + V(A)$
$= -1 + (-1)$
$= -2$

23. $q_{k+1}(D, \text{Down})$
$= -1 + V(G)$
$= -1 + (-1)$
$= -2$

VALIANT®

NO:

DATE:

24.) $q_{k+1}(F, \text{LEFT})$
$= -1 + v(F)$
$= -1 + (-1)$
$= -2$

25 $q_{k+1}(F, \text{RIGHT})$
$= -1 + v(F)$
$= -1 + (-1)$
$= -2$

26. $q_{k+1}(F, \text{UP})$
$= -1 + v(C)$
$= -1 + (-1)$
$= -2$

27. $q_{k+1}(F, \text{DOWN})$
$= -1 + v(1)$
$= -1 + (0)$
$= -1$

28. $q_{k+1}(G, \text{LEFT})$
$= -1 + v(G)$
$= -1 + (-1)$
$= -2$

29. $q_{k+1}(G, \text{RIGHT})$
$= -1 + v(H)$
$= -1 + (-1)$
$= -2$

30.) $q_{k+1}(G, \text{UP})$
$= -1 + v(D)$
$= -1 + (-1)$
$= -2$

31. $q_{k+1}(G, \text{DOWN})$
$= -1 + v(G)$
$= -1 + (-1)$
$= -2$

32. $q_{k+1}(H, \text{LEFT})$
$= -1 + v(G)$
$= -1 + (-1)$
$= -2$

33. $q_{k+1}(H, \text{RIGHT})$
$= -1 + v(1)$
$= -1 + (0)$
$= -1$

34. $q_{k+1}(H, \text{UP})$
$= -1 + v(H)$
$= -1 + (-1)$
$= -2$

35. $q_{k+1}(H, \text{DOWN})$
$= -1 + v(H)$
$= -1 + (-1)$
$= -2$

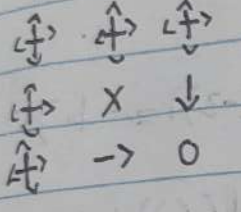
π_{k+1} A {LEFT, RIGHT, UP, DOWN}
π_{k+1} B {LEFT, RIGHT, UP, DOWN}
π_{k+1} C {LEFT, RIGHT, UP, DOWN}
π_{k+1} D {LEFT, RIGHT, UP, DOWN}
π_{k+1} F {DOWN}

π_{k+1} G {L, R, UP, DOWN}
π_{k+1} H {RIGHT}

A
B
C
D
F
G
H

NO:
DATE:

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



$V(A) = 1/4\,[-1 + V_{K+1}(A)) + (-1 + V_{K+1}(B)) + (-1 + V_{K+1}(A)) + (-1 + V_{K+1}(A))$

$V_{K+2}(A) = 1/4\,[(-1,(-1)) + (-1+(-1)) + (-1+(-1)) + (-1+(-1))$

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

$V_{K+2}(B) = 1/4\,[-1 + V_{K+1}(A)) + (-1 + V_{K+1}(C)) + (-1 + V_{K+1}(B)) + (-1 + V_{K+1}(B))$

$V_{K+2}(B)\ 1/4\,[(-1,(-1)) + (-1+(-1)) + (-1,(-1)) + (-1+(-1))$

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

$V_{K+2}(C) = 1/4\,[(-1 + V_{K+1}(B)) + (-1 + V_{K+1}(C)) + (-1 + V_{K+1}(F)) + (-1 + V_{K+1}(H)$

$V_{K+2}(C) = 1/4\,[(-1 + (-1)) + (-1 + (-1)) + (-1 + (-1)) + (-1 + -1))$

$V_{K+2}(C) \cdot (-2 + (-2)) + (-2 + (-2)) + (-2 + (-2)) + (-2 + (-2)) = -2$

VALIANT®

NO:

DATE:

$V_{K+2}(D) = 1/4 [(-1 + V_{K+1}(D)) + (-1 + V_{K+1}(D)) + (-1 + V_{K+1}(G)) + (-1 + V_{K+1}(A))$

$= 1/4 [(-1 + (-1)) + (-1 + (-1)) + (-1 + (-1)) + (-1 + (-1))]$

$V_{K+2}(D) = [(-2 + (-2)) + (-2 + (-2)) + (-2 + (-2)) + (-2 + (-2))] = -2$

$V_{K+2}(F) = 1/4 [(-1 + V_{K+1}(F)) + (-1 + V_{K+1}(F)) + (-1 + V_{K+1}(C)) + (-1 + V_{K+1}(C))]$

$= 1/4 [(-1 + (-1)) + (-1 + (-1)) + (-1 + (0)) + (-1 + (-1))$

$V_{K+2}(F) = (-2 + (-2)) + (-2 + (-2)) + (-2 + (-1)) + (-2 + (-2))] = -1.75$

$V_{K+2}(G) = 1/4 [(-1 + V_{K+1}(G)) + (-1 + V_{K+1}(H)) + (-1 + V_{K+1}(G)) + (-1 + V_{K+1}(D))]$

$V_{K+2}(G) = 1/4 [(-1 + (-1)) + (-1 + (-1)) + (-1 + (-1)) + (-1 + (-1))]$

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

$V_{K+2}(H) = 1/4 (-1 + V_{K+1}(G)) + (-1 + V_{K+1}(1)) + (-1 + V_{K+1}(1)) + (-1 + V_{K+1}(H))]$

$V_{K+2}(H) = 1/4 (-1 + (-1)) + (-1 + (-1)) + (-1 + (-1)) + (-1 + (-1))$

$V_{K+2}(H) = (-2 + (-2)) + (-2 + (-1)) + (-2 + (-2)) + (-2 + (-2)) = 1.7$

NO:
DATE:

$9K+2(A,\text{LEFT}) = -1 + (A) = -1 + (-2) = -3$

$9K+2(A,\text{RIGHT}) = -1 + v(B) = -1 + (-2) = -3$

$9K+2(A,\text{UP}) = -1 + v(A) = -1 + (-2) = -3$

$9K+2(A,\text{DOWN}) = -1 + v(D) = -1 + (-2) = -3$

$9K+2(B,\text{LEFT}) = -1 + (A) = -1 + (-2) = -3$

$9K+2(B,\text{RIGHT}) = 1 + v(C) = -1 + (-2) = -3$

$9K+2(B,\text{UP}) = -1 + v(B) = -1 + (-2) = -3$

$9K+2(B,\text{DOWN}) = -1 + v(B) = -1 + (-2) = -3$

$9K+2(C,\text{LEFT}) = -1 + (B) = -1 + (-2) = -3$

$9K+2(C,\text{RIGHT}) = -1 + v(C) = -1 + (-2) = -3$

$9K+2(C,\text{UP}) = -1 + v(C) + (-2) = -3$

$9K+2(C,\text{DOWN}) = -1 + v(F) = -1 + (1.75) = -2.75$

$9K+2(D,\text{LEFT}) = -1 + (D) = -1 + (-2) = -3$

$9K+9(D,\text{RIGHT}) = -1 + v(D) = -1 + (-2) = -3$

$9K+9(D,\text{UP}) = -1 + v(A) = -1 + (-2) = -3$

$9K+2(D,\text{DOWN}) = -1 + v(G) = -1 (-2) = -3$

$9K+2(E,\text{LEFT}) = -1 + (E) = -1 + (-2) = -3$

$9K+2(E,\text{RIGHT}) = -1 + (E) = -1 + (-2) = -3$

$9K+2(\text{UP}) = -1 + (C) = -1 + (-2) = -3$

$9K+2(\text{DOWN})$

VALIANT

NO:
DATE:

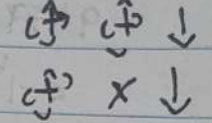
$q_{K+2}(G, \text{LEFT})$
$= -1 + V(G)$
$= -1 + (-2)$
$= -3$

$q_{K+2}(G, \text{RIGHT})$
$= -1 + V(H)$
$= -1 + (-1.75)$
$= -2.75$

$q_{K+2}(G, \text{UP})$
$= -1 + V(D)$
$= -1 + (-2)$
$= -3$

$q_{K+2}(G, \text{DOWN})$
$= -1 + (-2)$
$= -3$

$\pi_{K+2}(A) \{L, R, U, D\}$
$\pi_{K+2}(B) \{L, R, U, D\}$
$\pi_{K+2}(C) \{D\}$
$\pi_{K+2}(D) \{L, R, U, P\}$
~~$\pi_{K+2}(E) \{D\}$~~
$\pi_{K+2}(G) \{R\}$



$V_{K+3}(A) = 1/4 [(-1 + V_{K+2}(A)) + (-1 + V_{K+2}(B)) + (-1 + V_{K+2}(C)) + (-1 + V_{K+2}($
$V_{K+3}(A) = 1/4 [(-1 + (-2)) + (-1 + (-2)) + (-1 + (-2)) + (-1 + (-2))$
$V_{K+3}(A) = 1/4 = [$

VALIANT®