NO.:
DATE:

Carl Arvin C. Hipolito

COM221

* States S: (A, B, C, D, E, F, G, H, I)
* Actions A: (UP, DOWN, LEFT, RIGHT)
* Policy P: From every state, choose each action w/ probability 0.25
* Reward (R: -1) per step
* Discount Factor ($\gamma = 1$)
* Undiscounted MDP: ($\gamma = 1$)
* Non-terminal states (A, B, D, E, F, H)
* Terminal state (C, G, I)
* Agent follows a uniform random policy

RULES:

* From each state, actions move in that direction if possible, otherwise you stay in the same square.
* Reward is -1 until the terminal state is reached.

GOAL:

* The goal is to reach state C, G, I which gives 0 reward and ends the episode.
* To reach the goal, we need to find the optimal policy π.

STEP 1:

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

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

3. $V_{k+1}(D) = \frac{1}{4}[(-1 + V(A)) + (-1 + V(E)) + (-1 + V(G)) + (-1 + V(D))]$
$= \frac{1}{4}[(-1) + (-1) + (-1) + (-1)]$
$V_{k+1}(D) = -1$

4. $V_{k+1}(E) = \frac{1}{4}[(-1 + V(B)) + (-1 + V(F)) + (-1 + V(H)) + (-1 + V(D))]$
~~$\frac{1}{4}[(-1) + (-1)]$~~
$= \frac{1}{4}[(-1) + (-1) + (-1) + (-1)]$
$V_{k+1}(E) = -1$

5. $V_{k+1}(F) = \frac{1}{4}[(-1 + V(C)) + (-1 + V(F)) + (-1 + V(I)) + (-1 + V(E))]$
$= \frac{1}{4}[(-1) + (-1) + (-1) + (-1)]$
$V_{k+1}(F) = -1$

6. $V_{k+1}(H) = \frac{1}{4}[(-1 + V(E)) + (-1 + V(I)) + (-1 + V(H)) + (-1 + V(G))]$
$= \frac{1}{4}[(-1) + (-1) + (-1) + (-1)]$
$V_{k+1}(H) = -1$

	$V_k(S)$	$V_{k+1}(S)$	$V_{k+2}(S)$
A	0	-1	-2
B	0	-1	-1.75
D	0	-1	-1.75
E	0	-1	-2
F	0	-1	-1.5
H	0	-1	-1.5

7.

-1	-1	0
-1	-1	-1
0	-1	0

VICTORY

NO.:
DATE:

STEP 2:

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

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

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

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

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

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

14. $q_{k+1}(B, RIGHT) = -1 + V(C)$
$= -1 + 0$
$q_{k+1}(B, RIGHT) = -1$

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

16. $q_{k+1}(B, DOWN) = -1 + V(E)$
$= -1 + (-1)$
$q_{k+1}(B, DOWN) = -2$

17. $\pi_{k+1}(B) = \{RIGHT\}$

18. $q_{k+1}(D, LEFT) = -1 + V(D)$
$= -1 + (-1)$
$q_{k+1}(D, LEFT) = -2$

19. $q_{k+1}(D, RIGHT) = -1 + V(E)$
$= -1 + (-1)$
$q_{k+1}(D, RIGHT) = -2$

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

21. $q_{k+1}(D, DOWN) = -1 + V(G)$
$= -1 + 0$
$q_{k+1}(D, DOWN) = -1$

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

23. $q_{k+1}(E, LEFT) = -1 + V(C)$
$= -1 + (-1)$
$q_{k+1}(E, LEFT) = -2$

24. $q_{k+1}(E, RIGHT) = -1 + V(E)$
$= -1 + (-1)$
$q_{k+1}(E, RIGHT) = -2$

25. $q_{k+1}(E, UP) = -1 + V(B)$
$= -1 + (-1)$
$q_{k+1}(E, UP) = -2$

26. $q_{k+1}(E, DOWN) = -1 + V(H)$
$= -1 + (-1)$
$q_{k+1}(E, DOWN) = -2$

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

28. $q_{k+1}(F, LEFT) = -1 + V(E)$
$= -1 + (-1)$
$q_{k+1}(F, LEFT) = -2$

29. $q_{k+1}(F, RIGHT) = -1 + V(F)$
$= -1 + (-1)$
$q_{k+1}(F, RIGHT) = -2$

NO.:
DATE:

7.

30. $q_{k+1}(F, UP) = -1 + V(C)$

$= -1 + (0)$

$\boxed{q_{k+1}(F, UP) = -1}$

31. $q_{k+1}(F, DOWN) = -1 + V(J)$

$= -1 + 0$

$\boxed{q_{k+1}(F, DOWN) = -1}$

32. $\boxed{\pi_{k+1}(F) = \{UP, DOWN\}}$

33. $q_{k+1}(H, LEFT) = -1 + V(G)$

$= -1 + 0$

$\boxed{q_{k+1}(H, LEFT) = -1}$

34. $q_{k+1}(H, RIGHT) = -1 + V(I)$

$= -1 + 0$

$\boxed{q_{k+1}(H, RIGHT) = -1}$

35. $q_{k+1}(H, UP) = -1 + V(E)$

$= -1 + (-1)$

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

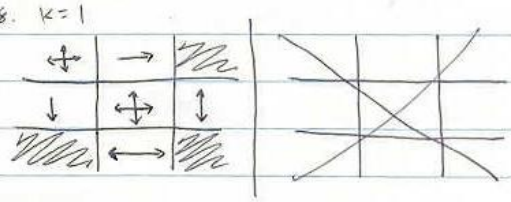
36. $q_{k+1}(H, DOWN) = -1 + V(H)$

$= -1 + (-1)$

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

37. $\pi_{k+1}(H) = \{LEFT, RIGHT\}$

38. $k = 1$



39. $V_k(A) =$ —

40. $V_k(B) = -1$

41. $V_k(D) =$ —

42. $V_k(E) =$ —

43. $V_k(F) =$ —

44. $V_k(H) =$ —

($k = 2$)

$V_{k+2}(A) = \frac{1}{4}[(-1 + V^{k+1}(A)) + (-1 + V^{k+1}(B)) + (-1 + V^{k+1}(D)) + (-1 + V^{k+1}(A))]$

$= \frac{1}{4}[(-2) + (-2) + (-2) + (-2)]$

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

$V_{k+2}(B) = \frac{1}{4}[(-1 + V_{k+1}(B)) + (-1 + V_{k+1}(C)) + (-1 + V_{k+1}(E)) + (-1 + V_{k+1}(A))]$

$= \frac{1}{4}[(-2) + (-1) + (-2) + (-2)]$

$\boxed{V_{k+2}(B) = -1.75}$

$V_{k+2}(D) = \frac{1}{4}[(-1 + V_{k+1}(A)) + (-1 + V_{k+1}(E)) + (-1 + V_{k+1}(G)) + (-1 + V_{k+1}(D))]$

$= \frac{1}{4}[(-2) + (-2) + (-2) + (-2)]$

$\boxed{V_{k+2}(D) = -1.75}$

$V_{k+2}(E) = \frac{1}{4}[(-1 + V_{k+1}(B)) + (-1 + V_{k+1}(F)) + (-1 + V_{k+1}(H)) + (-1 + V_{k+1}(D))]$

$= \frac{1}{4}[(-2) + (-2) + (-2) + (-2)]$

$\boxed{V_{k+2}(E) = -2}$

$V_{k+2}(F) = \frac{1}{4}[(-1 + V_{k+1}(C)) + (-1 + V_{k+1}(F)) + (-1 + V_{k+1}(I)) + (-1 + V_{k+1}(E))]$

$= \frac{1}{4}[(-1) + (-2) + (-1) + (-2)]$

$\boxed{V_{k+2}(F) = -1.5}$

$V_{k+2}(H) = \frac{1}{4}[(-1 + V_{k+1}(E)) + (-1 + V_{k+1}(I)) + (-1 + V_{k+1}(H)) + (-1 + V_{k+1}(G))]$

$= \frac{1}{4}[(-2) + (-1) + (-2) + (-1)]$

$\boxed{V_{k+2}(H) = -1.5}$

-2	-1.75	0
-1.75	-2	-1.5
0	-1.5	0

NO.:
DATE:

45. $q_{k+2}(A, \leftarrow) = -1 + (-2)$
$= -3$
$q_{k+2}(A, \rightarrow) = -1 + (-1.75)$
$= -2.75$
$q_{k+2}(A, \uparrow) = -1 + (-2)$
$= -3$
$q_{k+2}(A, \downarrow) = -1 + (-1.75)$
$= -2.75$
$\pi_{k+2}(A) = \{RIGHT, DOWN\}$

46. $q_{k+2}(B, \leftarrow) = -1 + (-2) = -3$
$q_{k+2}(B, \rightarrow) = -1 + (0) = -1$
$q_{k+2}(B, \uparrow) = -1 + (-1.75) = -2.75$
$q_{k+2}(B, \downarrow) = -1 + (-2) = -3$
$\pi_{k+2}(B) = \{RIGHT\}$

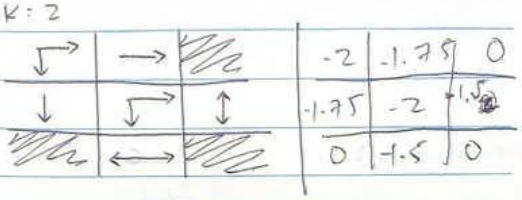
47. $q_{k+2}(D, \leftarrow) = -1 + (-1.75) = -2.75$
$q_{k+2}(D, \rightarrow) = -1 + (-2) = -3$
$q_{k+2}(D, \uparrow) = -1 + (-2) = -3$
$q_{k+2}(D, \downarrow) = -1 + (0) = -1$
$\pi_{k+2}(D) = \{DOWN\}$

48. $q_{k+2}(E, \leftarrow) = -1 + (-1.75) = -2.75$
$q_{k+2}(E, \rightarrow) = -1 + (-1.5) = -2.5$
$q_{k+2}(E, \uparrow) = -1 + (-1.75) = -2.75$
$q_{k+2}(E, \downarrow) = -1 + (-1.5) = -2.5$
$\pi_{k+2}(E) = \{RIGHT, DOWN\}$

49. $q_{k+2}(F, \leftarrow) = -1 + (-2) = -3$
$q_{k+2}(F, \rightarrow) = -1 + (-1.5) = -2.5$
$q_{k+2}(F, \uparrow) = -1 + (0) = -1$
$q_{k+2}(F, \downarrow) = -1 + (0) = -1$
$\pi_{k+2}(F) = \{UP, DOWN\}$

50. $q_{k+2}(H, \leftarrow) = -1 + (0) = -1$
$q_{k+2}(H, \rightarrow) = -1 + (0) = -1$
$q_{k+2}(H, \uparrow) = -1 + (-2) = -3$
$q_{k+2}(H, \downarrow) = -1 + (-1.5) = -2.5$
$\pi_{k+2}(H) = \{LEFT, RIGHT\}$

$K = 2$



51. {RIGHT, DOWN}
52. {RIGHT}
53. {DOWN}
54. {RIGHT, DOWN}
55. {UP, DOWN}
56. {LEFT, RIGHT}

49. -2
40. -1.75
41. -1.75
42. -2
43. -1.5
44. -1.5