

NAMA : SAVIRA AGUSTIN

NIM : 1301173406

KELAS : IF 41-12

1/	kuning		END	- Tujuan bergerak secepat mungkin untuk menemukan kotak END
		merah		- Menginjak kotak kuning kesehatan terganggu
	END			- Menginjak kotak merah mati
		kuning		

Tentukan nilai reward setiap kotak :

- > kotak merah : -3
- > kotak kuning : -2
- > kotak kosong : -1
- > kotak END : 8

2/	1	2	3	Finish	4	- Nilai reward start = 0, finish = 8, kotak merah = -3, kotak kosong = -1
	5	6	7		8	
	Start	9	10	11	12	- alfa (learning rate) = 1, gamma (discount rate) = 1, aksi atas, bawah, kiri, kanan
	13	14	15	16		- Nilai initial tabel $Q(s,a) = 0$

- a. Tentukan tabel Q pada saat initial
- b. Update tabel Q untuk episode 9-5-1-2-3-4
- c. Update tabel Q hasil b untuk episode 9-10-11-12

Jawab !

a. $Q(s,a) = 0$

States

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
a																
c	atas	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
t	bawah	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
i	kiri	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
n	kanan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
s																

b) Episode : 9-5-1-2-3-4

-3	11	-1	12	-1	13	8	14
-1	15	-1	16	-1	17	-1	18
0	9	-1	10	-1	11	-1	12
-1	13	-1	14	-3	15	-3	16

diketahui :

$$\alpha = 1$$

$$\gamma = 1$$

$$Q(s,a) \leftarrow Q(s,a) + \alpha [r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$$

1. Step 1 : $Q(s_9, atas) \leftarrow 0 + 1 * [-1 + 1 * \max \{0, 0, 0, 0\} - 0]$

$$Q(s_9, atas) \leftarrow -1$$

2. Step 2 : $Q(s_5, atas) \leftarrow 0 + 1 * [-3 + 1 * \max \{0, 0, 0, 0\} - 0]$

$$Q(s_5, atas) \leftarrow -3$$

3. Step 3 : $Q(s_1, kanan) \leftarrow 0 + 1 * [-1 + 1 * \max \{0, 0, 0, 0\} - 0]$

$$Q(s_1, kanan) \leftarrow -1$$

4. Step 4 : $Q(s_2, kanan) \leftarrow 0 + 1 * [-1 + 1 * \max \{0, 0, 0, 0\} - 0]$

$$Q(s_2, kanan) \leftarrow -1$$

5. Step 5 : $Q(s_3, kanan) \leftarrow 0 + 1 * [8 + 1 * \max \{0, 0, 0, 0\} - 0]$

$$Q(s_3, kanan) \leftarrow 8$$

States

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
a	atas	0	0	0	0	-3	0	0	0	-1	0	0	0	0	0	0	0
c	bawah	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
t	kiri	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
i	kanan	-1	-1	8	0	0	0	0	0	0	0	0	0	0	0	0	0
o																	
n																	
s																	

c) Episode : 9-10-11-12

-3	11	-1	12	-1	13	8	14
-1	15	-1	16	-1	17	-1	18
0	9	-1	10	-1	11	-1	12
-1	13	-1	14	-3	15	-3	16

$$Q(s,a) \leftarrow Q(s,a) + \alpha [r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$$

1. Step 1 : $Q(s_9, \text{kanan}) \leftarrow 0 + 1 * [-1 + 1 * \max \{0, 0, 0, 0\} - 0]$
 $Q(s_9, \text{kanan}) \leftarrow -1$

2. Step 2 : $Q(s_{10}, \text{kanan}) \leftarrow 0 + 1 * [-1 + 1 * \max \{0, 0, 0, 0\} - 0]$
 $Q(s_{10}, \text{kanan}) \leftarrow -1$

3. Step 3 : $Q(s_{11}, \text{kanan}) \leftarrow 0 + 1 * [-1 + 1 * \max \{0, 0, 0, 0\} - 0]$
 $Q(s_{11}, \text{kanan}) \leftarrow -1$

		States															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
actions	atas	0	0	0	0	-3	0	0	0	-1	0	0	0	0	0	0	0
	bawah	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	kiri	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	kanan	-1	-1	0	0	0	0	0	0	-1	-1	-1	0	0	0	0	0

