프로그램 결과값만 캡쳐해서 올립니다.

```
Initial Q table is

col=0        1        2        3        4        5
row: 0
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
row: 1
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
row: 2
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
row: 3
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
row: 4
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
row: 5
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
 0.00,    0.00,    0.00,    0.00,    0.00,    0.00
```

```
Learning starts.

episode= 0    epsilon= 0.4
episode= 5000     epsilon= 0.36924653855465434
episode= 10000    epsilon= 0.3408575155864846
episode= 15000    epsilon= 0.3146511444266214
episode= 20000    epsilon= 0.2904596148294764
episode= 25000    epsilon= 0.26812801841425576
episode= 30000    epsilon= 0.24751335672245633
episode= 35000    epsilon= 0.22848362553952595
episode= 40000    epsilon= 0.21091696961721942
episode= 45000    epsilon= 0.19470090238398868
```

Learning is finished. the Q table is:

col=0	1	2	3	4	5
row: 0					
0.00,	0.00,	0.00,	0.00,	0.00,	0.00
0.00,	0.00,	0.00,	0.00,	0.00,	0.00
0.00,	0.00,	0.00,	0.00,	0.00,	0.00
0.00,	0.00,	0.00,	0.00,	0.00,	0.00
row: 1					
0.00,	-9.00,	-9.00,	-9.00,	-9.00,	0.00
0.00,	5.31,	5.90,	5.31,	-9.00,	0.00
0.00,	5.31,	-9.00,	6.56,	-9.00,	0.00
0.00,	-9.00,	4.78,	5.31,	5.90,	0.00
row: 2					
0.00,	4.78,	0.00,	5.90,	0.00,	0.00
0.00,	-9.00,	0.00,	-9.00,	0.00,	0.00
0.00,	5.90,	0.00,	7.29,	0.00,	0.00
0.00,	-9.00,	0.00,	-9.00,	0.00,	0.00
row: 3					
0.00,	5.31,	-9.00,	6.56,	0.00,	0.00
0.00,	6.56,	7.29,	-9.00,	0.00,	0.00
0.00,	-9.00,	7.29,	8.10,	0.00,	0.00
0.00,	-9.00,	5.90,	6.56,	0.00,	0.00
row: 4					
0.00,	0.00,	6.56,	7.29,	0.00,	0.00
0.00,	0.00,	8.10,	9.00,	0.00,	0.00
0.00,	0.00,	-9.00,	-9.00,	0.00,	0.00
0.00,	0.00,	-9.00,	7.29,	0.00,	0.00
row: 5					
0.00,	0.00,	0.00,	0.00,	0.00,	0.00
0.00,	0.00,	0.00,	0.00,	0.00,	0.00
0.00,	0.00,	0.00,	0.00,	0.00,	0.00
0.00,	0.00,	0.00,	0.00,	0.00,	0.00

· 코드   + 넥스트

Test starts.

Episode: 0 start state: (1 , 1)
The move is: right that leads to state (1 , 2)
The move is: right that leads to state (1 , 3)
The move is: down   that leads to state (2 , 3)
The move is: down   that leads to state (3 , 3)
The move is: down   that leads to state (4 , 3)
The move is: right that leads to state (4 , 4)
Episode has ended. Total reward received in episode = 9

Episode: 1 start state: (1 , 1)
The move is: down   that leads to state (2 , 1)
The move is: down   that leads to state (3 , 1)
The move is: right that leads to state (3 , 2)
The move is: down   that leads to state (4 , 2)
The move is: right that leads to state (4 , 3)
The move is: right that leads to state (4 , 4)
Episode has ended. Total reward received in episode = 9

```
Episode: 2  start state: ( 1 ,  1 )
The move is: right   that leads to state ( 1 ,   2 )
The move is: right   that leads to state ( 1 ,   3 )
The move is: down    that leads to state ( 2 ,   3 )
The move is: down    that leads to state ( 3 ,   3 )
The move is: down    that leads to state ( 4 ,   3 )
The move is: right   that leads to state ( 4 ,   4 )
Episode has ended. Total reward received in episode =  9

Episode: 3  start state: ( 1 ,  1 )
The move is: right   that leads to state ( 1 ,   2 )
The move is: right   that leads to state ( 1 ,   3 )
The move is: down    that leads to state ( 2 ,   3 )
The move is: down    that leads to state ( 3 ,   3 )
The move is: down    that leads to state ( 4 ,   3 )
The move is: right   that leads to state ( 4 ,   4 )
Episode has ended. Total reward received in episode =  9

Episode: 4  start state: ( 1 ,  1 )
The move is: down    that leads to state ( 2 ,   1 )
The move is: down    that leads to state ( 3 ,   1 )
The move is: right   that leads to state ( 3 ,   2 )
The move is: right   that leads to state ( 3 ,   3 )
The move is: down    that leads to state ( 4 ,   3 )
The move is: right   that leads to state ( 4 ,   4 )
Episode has ended. Total reward received in episode =  9
Program ends!!!
```