GRIDWORLD MDP

GRIDWORLD is 3×4 with 1 wall

∴ Total States = 12 − 1 = 11

Action: North, South, East, West.

	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
S_0	0.1	0.1			0.8						
S_1											
S_2		0.1		0.1		0.8					
S_3											
S_4					0.2			0.8			
S_5						0.1	0.1			0.9	
S_6								1.0			
S_7								0.9	0.1		
S_8								0.1	0.8	0.1	
S_9									0.1	0.8	0.1
S_{10}										0.1	0.9

```
0      1      2      3
•      •      •      +

4             5      6
•             •      -

7      8      9      10
A      •      •      •
```

CORRECT : 0.8
WRONG : 0.2 evenly
+ : +1.0
− : −1.0

ACTION : SOUTH

0 → 4

$S0\, S_4 = 0.8$
we can move to S_1 & S_0 with 0.2
even Probability

∴ $S_0\, S_0 = 0.1$
$S0\, 1 = 0.1$

2 → 5

$S_2\, S_5 = 0.8$
we can move to S_1 & S_3
∴ $S_2\, S_1 = 0.1 = 0.1$

$$3 \rightarrow 6$$

Now 6 in kind of end state with bad reward. So the moment that happens we will move to start i.e. **7**

$$\therefore \quad S_3 \, S_7 = 1.0$$

$$4 \rightarrow 7$$

$$S_4 \, S_7 = 0.8$$

Now since on right we have wall so only wrong move could be S_4 ↓

$$\therefore \quad S_{4} \, S_4 = 0.2$$

$$5 \rightarrow 9$$

$$S_5 \, S_9 = 0.8$$

We can be in same state (due to wall) or move to S_6

$$\therefore \quad S_5 \, S_5 = 0.1 = S_5 \, S_6$$

($\because S_5 \, S_6$ won't be like $3 \rightarrow 6$ as here we are just analysing the chance whereas in $3 \rightarrow 6$ we were actually making a move.)

$$6 \rightarrow 10$$

$\because$ 6 in end state so restart to 7

$$\therefore \quad S_6 \, S_7 = 1.0$$

($\because$ Nothing in Down)
↓

$$7 \rightarrow 7$$

$$\therefore \quad S_7 \, S_7 = 0.8$$

Now can move to S_7 or S_8

$$\therefore \quad S_7 \, S_7 = 0.8 + 0.1 = 0.9$$
$$S_7 \, S_8 = 0.1 \quad\quad = 0.1$$

$$8 \rightarrow 8$$

$$S_8 \, S_8 = 0.8$$
$$S_8 \, S_7 = 0.1$$
$$S_8 \, S_9 = 0.1$$

$$9 \rightarrow 9$$
$$S_9 \, S_9 = 0.8$$
$$S_9 \, S_{10} = 0.1$$
$$S_9 \, S_8 = 0.1$$

$$10 \rightarrow 10$$
$$S_{10} \, S_{10} = 0.8$$
$$S_{10} \, S_9 = 0.1$$
$$S_{10} \, S_{10} = S_{10} S_{10} + 0.1 = 0.9$$