

نمونه آزمایشی حل مسئله SHM با استفاده از Markov Decision Process

در این نمونه یک مدل ساده به شرح زیر تصور شده است:

ساختمانی در نظر بگیرید که تنها دو سنسور برای تشخیص رطوبت دارد. و دو نوع واقعه در هر لحظه میتواند رخ بدهد. یا بران میبارد یا بران نمیبارد. بنابراین در کل ۸ حالت خواهیم داشت مطابق جدول زیر:

وضعیت هوا	سنسور دوم	سنسور اول	حالت ها
صاف	خاموش	خاموش	S0
صاف	روشن	خاموش	S1
صاف	خاموش	روشن	S2
صاف	روشن	روشن	S3
برانی	خاموش	خاموش	S4
برانی	روشن	خاموش	S5
برانی	خاموش	روشن	S6
برانی	روشن	روشن	S7

در چنین سناریویی Reward های زیر در نظر گرفته شده است:

امتیاز	حالت
-0.5	S0
0	S1
0	S2
-1	S3
-1	S4
1	S5
1	S6
-0.5	S7

با کد نوشته شده جدول گذار زیر تولید شده است:

احتمال وقوع	حالت پس از اتفاق	اتفاق	حالت
0.2	S2	on1	S0
0.2	S0	off1	S0
0.2	S1	on2	S0
0.2	S0	off2	S0

S0	on1	S6	0.2
S0	off1	S4	0.4
S0	on2	S5	0.2
S0	off2	S4	0.4
S1	on1	S3	0.2
S1	off1	S1	0.2
S1	on2	S1	0.2
S1	off2	S0	0.2
S1	on1	S7	0.2
S1	off1	S5	0.4
S1	on2	S5	0.4
S1	off2	S4	0.2
S2	on1	S2	0.2
S2	off1	S0	0.2
S2	on2	S3	0.2
S2	off2	S2	0.2
S2	on1	S6	0.4
S2	off1	S4	0.2
S2	on2	S7	0.2
S2	off2	S6	0.4
S3	on1	S3	0.2
S3	off1	S1	0.2
S3	on2	S3	0.2
S3	off2	S2	0.2
S3	on1	S7	0.4
S3	off1	S5	0.2
S3	on2	S7	0.4
S3	off2	S6	0.2
S4	on1	S2	0.3
S4	off1	S0	0.1
S4	on2	S1	0.3
S4	off2	S0	0.1
S4	on1	S6	0.3
S4	off1	S4	0.3
S4	on2	S5	0.3
S4	off2	S4	0.3
S5	on1	S3	0.3
S5	off1	S1	0.1
S5	on2	S1	0.1
S5	off2	S0	0.3
S5	on1	S7	0.3
S5	off1	S5	0.3
S5	on2	S5	0.3
S5	off2	S4	0.3
S6	on1	S2	0.1
S6	off1	S0	0.3
S6	on2	S3	0.3
S6	off2	S2	0.1
S6	on1	S6	0.3

S6	off1	S4	0.3
S6	on2	S7	0.3
S6	off2	S6	0.3
S7	on1	S3	0.1
S7	off1	S1	0.3
S7	on2	S3	0.1
S7	off2	S2	0.3
S7	on1	S7	0.3
S7	off1	S5	0.3
S7	on2	S7	0.3
S7	off2	S6	0.3

پس از آن کد اصلی در خصوص شبیه سازی MDP اجرا میشود که نتیجه زیر را در بر خواهد داشت.

```

eitd@eitd-RnD: ~/Desktop/MDP_Lab/test1_begining
File Edit View Search Terminal Help
eitd@eitd-RnD:~/Desktop/MDP_Lab/test1_begining$ python mdp.py 0.004 0.001
State - Value
S3 - -1.0
S2 - 0.0
S1 - 0.0
S0 - -0.5
S7 - -0.5
S6 - 1.0
S5 - 1.0
S4 - -1.0

Optimal policy is
State - Action
S3 - off1
S2 - off2
S1 - off1
S0 - on2
S7 - off1
S6 - off2
S5 - off1
S4 - on2

```

در واقع Optimal Policy اشاره شده. دستورالعمل کم حجمی است که به سنسور داده خواهد شد تا در هر لحظه تصمیم بگیرد اقدام بهینه چیست.