



درس هوش مصنوعی و سیستم های خبره

---

تمرین دوازدهم

---

مدرس : دکتر محمدی

دانشجو : سارا سادات یونسی / ۹۸۵۳۳۰۵۳

## بخش عملی

کد در گیت لب گذاشته شده.

[sara\\_younesi / ReinforcementLearning · GitLab](#)

چرا در حالت manual هدف به سمتی که شما حرکت می‌دهید نمی‌رود؟

چون حرکت ما دارای نویز است و فقط ما تعیین کننده اکشن هستیم ولی امکان دارد ترنزیشنی متفاوت از آنچه که ما می‌خواهیم به ما بدهد.

بستگی به احتمال  $(s, a, s')$  دارد

دارای نویز ۰,۲ هستیم .

```
Select Anaconda Prompt [anaconda3]
Took action: east
Ended in state: (2, 1)
Got reward: 0.0

Started in state: (2, 1)
Took action: north
Ended in state: (2, 2)
Got reward: 0.0

Started in state: (2, 2)
Took action: east
Ended in state: (3, 2)
Got reward: 0.0

Started in state: (3, 2)
Took action: exit
Ended in state: TERMINAL_STATE
Got reward: 1

EPISODE 19 COMPLETE: RETURN WAS 0.34867844010000015

BEGINNING EPISODE: 20

Started in state: (0, 0)
Took action: north
Ended in state: (0, 1)
Got reward: 0.0

Started in state: (0, 1)
Took action: north
Ended in state: (0, 2)
Got reward: 0.0

Started in state: (0, 2)
Took action: east
Ended in state: (1, 2)
Got reward: 0.0

Started in state: (1, 2)
Took action: east
Ended in state: (1, 2)
Got reward: 0.0

Started in state: (1, 2)
Took action: east
Ended in state: (2, 2)
Got reward: 0.0

Started in state: (2, 2)
Took action: east
Ended in state: (3, 2)
Got reward: 0.0

Started in state: (3, 2)
Took action: exit
Ended in state: TERMINAL_STATE
Got reward: 1

EPISODE 20 COMPLETE: RETURN WAS 0.5314410000000002

AVERAGE RETURNS FROM START STATE: 0.30090242981497733
```

توضیحی بر کارکرد توابع :













