

ارائه پروژه درس هوش مصنوعی

عنوان پروژه : Pendulum

اعضای گروه :

محمد حسین حسینی - ۹۹۵۲۱۱۹۹

زهرا سادات طباطبائی - ۹۹۵۲۱۴۱۵

فهرست

4

خروجی و نتیجه

- نمودار
- گیف

3

توضیحات پروژه

- Action Space
- Observation Space

2

نصب پروژه

- نصب gymnasium
- نصب Requirement.txt

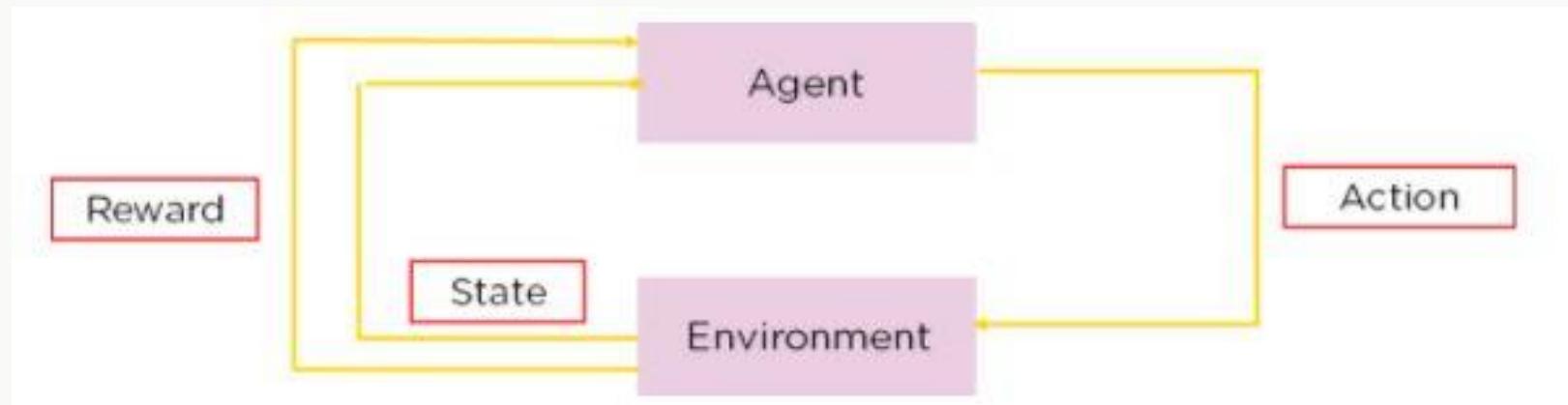
1

الگوریتم Q-learning

- Q-learning چیست؟
- سودوکد

Q-learning چیست؟

الگوریتم Q-Learning، یکی از الگوریتم‌های بسیار معروف از نوع off-policy در حوزه یادگیری تقویتی است. عامل یادگیرنده با الگوریتم Q-Learning، بعد از مشاهده اکشنی را انتخاب می‌کند. سپس محیط به Agent، حالت بعدی سیستم و Reward ناشی از اکشن را بر می‌گرداند. Agent با مشاهده اطلاعات دریافتی از محیط، اکشن بعدی را انتخاب می‌کند و این فرآیند تا زمان رسیدن به سیاست بهینه ادامه پیدا می‌کند.



در شکل زیر جزئیات الگوریتم Q-Learning نمایش داده شده است :

Algorithm 14: Sarsamax (Q-Learning)

Input: policy π , positive integer $num_episodes$, small positive fraction α , GLIE $\{\epsilon_i\}$

Output: value function Q ($\approx q_\pi$ if $num_episodes$ is large enough)

Initialize Q arbitrarily (e.g., $Q(s, a) = 0$ for all $s \in \mathcal{S}$ and $a \in \mathcal{A}(s)$, and $Q(\text{terminal-state}, \cdot) = 0$)

for $i \leftarrow 1$ **to** $num_episodes$ **do**

$\epsilon \leftarrow \epsilon_i$

 Observe S_0

$t \leftarrow 0$

repeat

 Choose action A_t using policy derived from Q (e.g., ϵ -greedy)

 Take action A_t and observe R_{t+1}, S_{t+1}

$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t))$

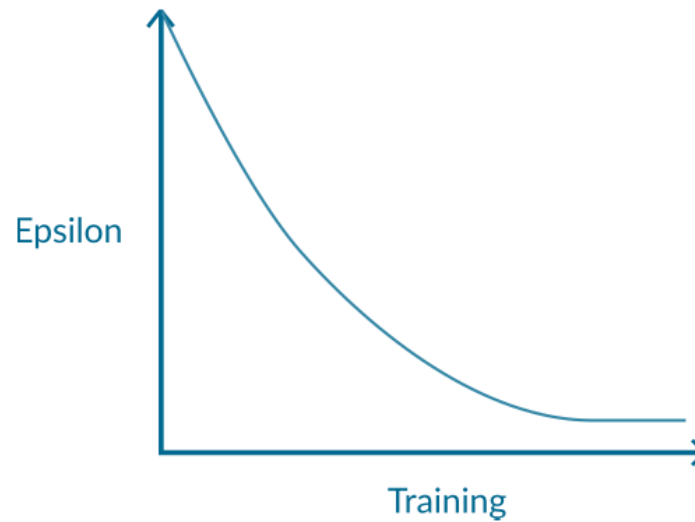
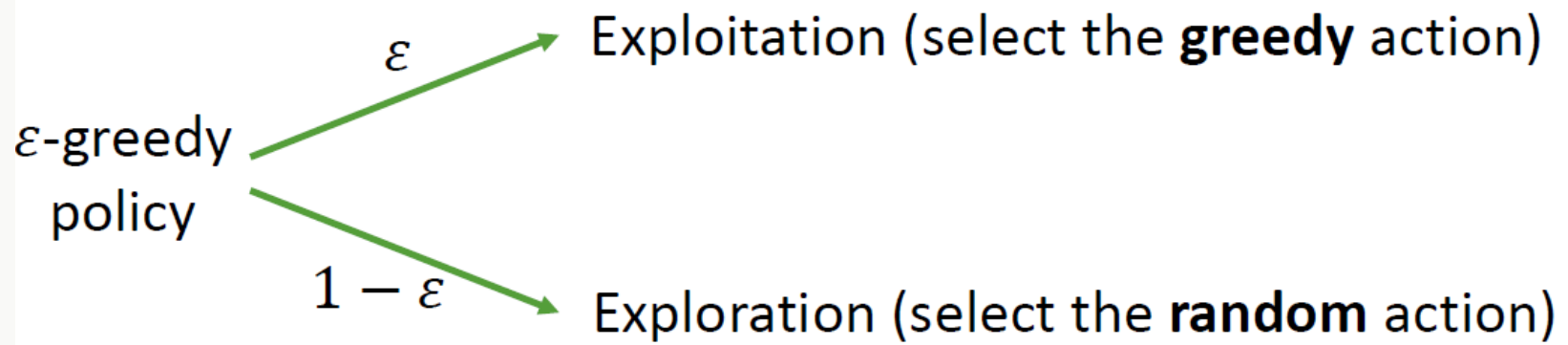
$t \leftarrow t + 1$

until S_t is terminal;

end

return Q

توضیح در اسلاید بعد



نصب و اجرای پروژه



1

`pip install gymnasium[classic_control]`



2

`pip install -r requirement.txt`



3

`python Qlearning.py`



توضیحات پروژه

Action Space

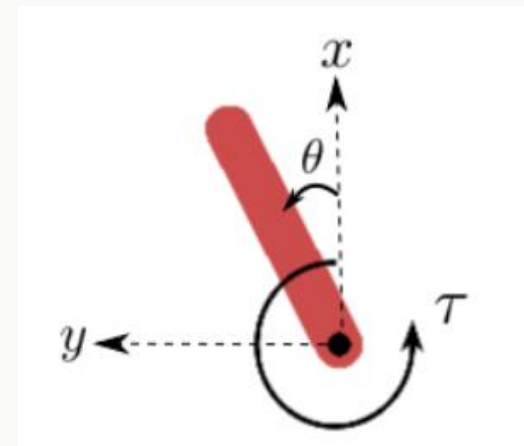
The action is a `ndarray` with shape `(1,)` representing the torque applied to free end of the pendulum.

Num	Action	Min	Max
0	Torque	-2.0	2.0

Observation Space

The observation is a `ndarray` with shape `(3,)` representing the x-y coordinates of the pendulum's free end and its angular velocity.

Num	Observation	Min	Max
0	$x = \cos(\theta)$	-1.0	1.0
1	$y = \sin(\theta)$	-1.0	1.0
2	Angular Velocity	-8.0	8.0



Action Space	<code>Box(-2.0, 2.0, (1,), float32)</code>
Observation Shape	<code>(3,)</code>
Observation High	<code>[1. 1. 8.]</code>
Observation Low	<code>[-1. -1. -8.]</code>
Import	<code>gymnasium.make("Pendulum-v1")</code>



نتیجه

