Quiz: Epsilon-Greedy Policies

In the previous concept, you learned about ϵ-greedy policies.

You can think of the agent who follows an ϵ-greedy policy as always having a (potentially unfair) coin at its disposal, with probability ϵ of landing heads. After observing a state, the agent flips the coin.

- If the coin lands tails (so, with probability $1 - \epsilon$), the agent selects the greedy action.
- If the coin lands heads (so, with probability ϵ), the agent selects an action *uniformly* at random from the set of available (non-greedy **AND** greedy) actions.

In order to construct a policy π that is ϵ-greedy with respect to the current action-value function estimate Q, we need only set

$$\pi(a|s) \leftarrow \begin{cases} 1 - \epsilon + \frac{\epsilon}{|\mathcal{A}(s)|} & \text{if } a = \arg\max_{a' \in \mathcal{A}(s)} Q(s, a') \\ \frac{\epsilon}{|\mathcal{A}(s)|} & \text{otherwise} \end{cases}$$

for each $s \in \mathcal{S}$ and $a \in \mathcal{A}(s)$. Note that ϵ must always be a value between 0 and 1, inclusive (that is, $\epsilon \in [0, 1]$).

In this quiz, you will answer a few questions to test your intuition.

QUESTION 1 OF 4

Which of the values for epsilon yields an epsilon-greedy policy that is guaranteed to **always** select the greedy action? Select all that apply.

(1) epsilon = 0

Quiz: Epsilon-Greedy Policies

- [ ] (3) epsilon = 0.5
- [ ] (4) epsilon = 1
- [ ] (5) This is a trick question! The *true answer* is that none of the values for epsilon satisfy this requirement.

SUBMIT

QUESTION 2 OF 4

Which of the values for epsilon yields an epsilon-greedy policy that is guaranteed to **always** select a non-greedy action? Select all that apply.

- [ ] (1) epsilon = 0
- [ ] (2) epsilon = 0.3
- [ ] (3) epsilon = 0.5
- [ ] (4) epsilon = 1
- [x] (5) This is a trick question! The *true answer* is that none of the values for epsilon satisfy this requirement.

SUBMIT

Quiz: Epsilon-Greedy Policies

Which of the values for epsilon yields an epsilon-greedy policy that is equivalent to the equiprobable random policy (where, from each state, each action is equally likely to be selected)?

- [ ] (1) epsilon = 0
- [ ] (2) epsilon = 0.3
- [ ] (3) epsilon = 0.5
- [x] (4) epsilon = 1
- [ ] (5) This is a trick question! The *true answer* is that none of the values for epsilon satisfy this requirement.

SUBMIT

QUESTION 4 OF 4

Which of the values for epsilon yields an epsilon-greedy policy where the agent has the *possibility* of selecting a greedy action, but *might* select a non-greedy action instead? In other words, how might you guarantee that the agent selects each of the available (greedy and non-greedy) actions with nonzero probability?

- [ ] (1) epsilon = 0
- [x] (2) epsilon = 0.3

Quiz: Epsilon-Greedy Policies

(4) epsilon = 1

(5) This is a trick question! The *true answer* is that none of the values for epsilon satisfy this requirement.

SUBMIT

NEXT