

Wprowadzenie do RL 7

UWAGA: Nie korzystamy z notatników Colab, ale z czystego Pythona. Przed przystąpieniem do zadań proszę zapisać na dysku skrypty z folderu **Lab_7**. Skrypt **cartpole.py** zawiera środowisko **CartPole**.

Zadanie 1

Przetestuj działanie środowiska **CartPole** (skrypt: [CartPole_random_action.py](#)) dla losowo generowanych akcji. Jaką długość miał **najkrótszy/najdłuższy epizod**, który zaobserwowałeś?

Zadanie 2

Przetestuj działanie środowiska **CartPole** w wersji z **dyskretyzacją**. Upewnij się, że rozumiesz **sposób dyskretyzacji**. Odpowiedz na pytania:

- A. Co jest wypisywane na ekranie?
- B. Jaka jest interpretacja wypisywanych liczb?

(skrypt: [CartPole_random_action_discretization.py](#))

Zadanie 3

Napisz program implementujący poniższy **algorytm** dla środowiska **CartPole** wyliczenia polityki π . Przyjmij że π to **polityka stochastyczna**, która obu akcjom (w lewo i w prawo) przypisuje **prawdopodobieństwo 0.5**.

Tabular TD(0) for estimating v_π

```
Input: the policy  $\pi$  to be evaluated
Algorithm parameter: step size  $\alpha \in (0, 1]$ 
Initialize  $V(s)$ , for all  $s \in \mathcal{S}^+$ , arbitrarily except that  $V(\text{terminal}) = 0$ 
Loop for each episode:
  Initialize  $S$ 
  Loop for each step of episode:
     $A \leftarrow$  action given by  $\pi$  for  $S$ 
    Take action  $A$ , observe  $R, S'$ 
     $V(S) \leftarrow V(S) + \alpha [R + \gamma V(S') - V(S)]$ 
     $S \leftarrow S'$ 
  until  $S$  is terminal
```

(uzupełnij skrypt: [cartpole_TD\(0\)_start.py](#))