

## Wprowadzenie do RL 10

### Zadanie 1

Zaimplementuj algorytm **Q-learning** w środowisku **FrozenLake** wykorzystując **sieć neuronową jako aproksymator funkcji Q**.

Sprawdź czy po nauczaniu **strategia zachłanna** polegająca na wybieraniu akcji o największej przewidywanej wartości zwrotu (o największej wartości Q) pozwala dotrzeć do stanu 15.

### Zadanie 2

Zaimplementuj algorytm **Q-learning** w środowisku **CartPole** wykorzystując **sieć neuronową jako aproksymator funkcji Q**.

- Przetestuj działanie algorytmu dla **różnych architektur sieci neuronowej** (możesz **dodawać/odejmować warstwy** sieci oraz **zmieniać ilość neuronów w warstwach**)
- Jak na proces uczenia wpływa zmiana wartości  $\epsilon$  oraz ilości epizodów treningowych?
- Porównaj wyniki uzyskane dla stałej wartości  $\epsilon$  oraz dla  $\epsilon$  zmniejszanego po każdym epizodzie według formuły:

$$\epsilon = \epsilon - (1/\text{train\_episodes})$$

gdzie `train_episodes` to ilość epizodów w czasie których agent się uczy. Przyjmij wartość początkową  $\epsilon = 1$ .

**UWAGA:** w przypadku obu zadań wygeneruj **wykres zmiany błędu w kolejnych epizodach** oraz **wykres sumy nagród zdobytych w kolejnych epizodach**.