

Sieci Samouczące Się

Laboratorium 2 – Sprawozdanie

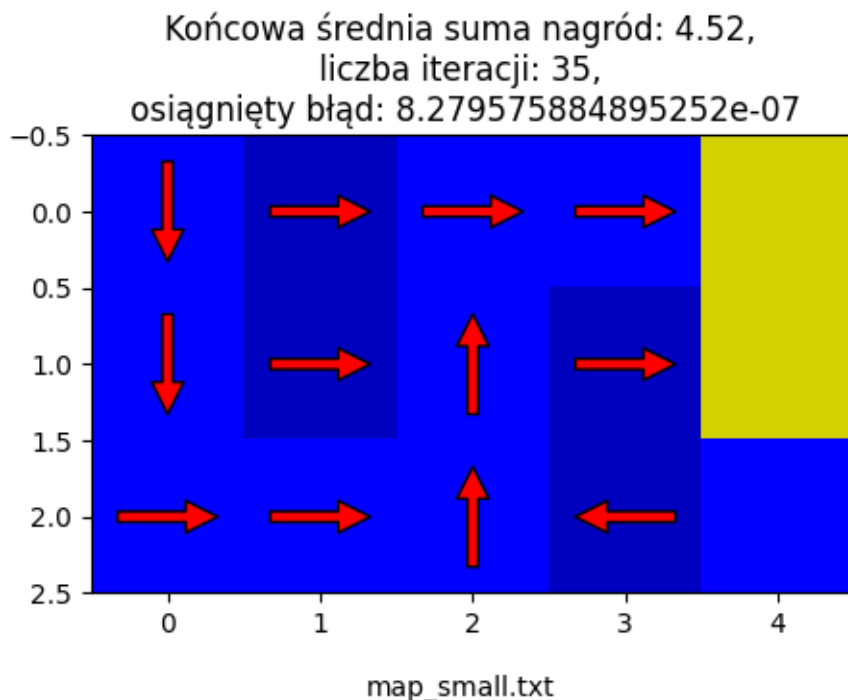
Programowanie dynamiczne, przyjęty wariant – iteracja wartości, implementacja inspirowana slajdami wykładowymi.

Działanie:

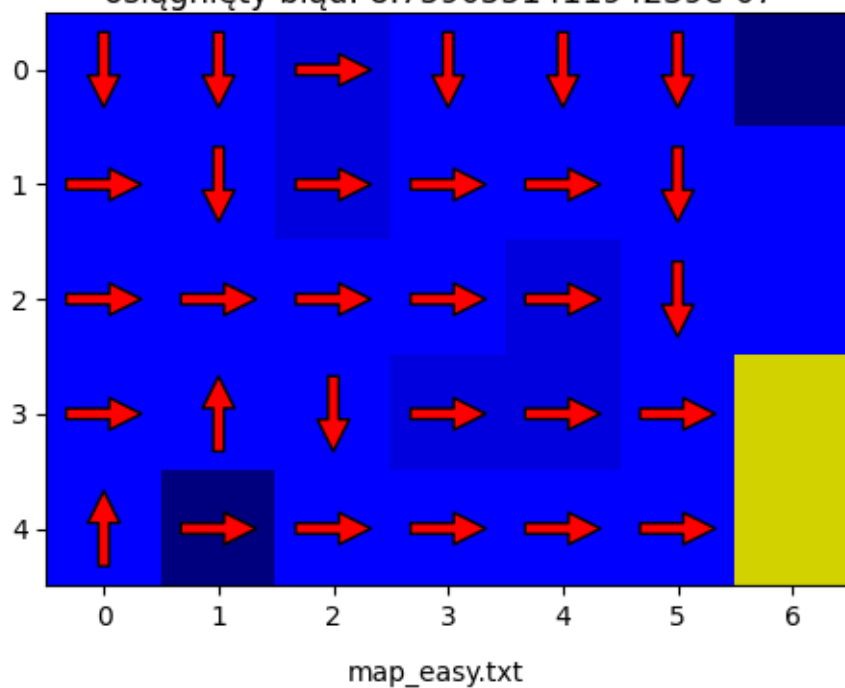
Zaczynamy od ustalenia zmiennych i inicjalizacji wartości stanów V wypełnionej zerami. Następnie wchodzimy do pętli trwającej do poprawienia różnicy błędów lub osiągnięcia limitu iteracji. W pętli tworzymy zmienną v_pom jako kopię V i ustalamy błąd na 0 dla obecnej iteracji. Potem przechodzimy po wszystkich stanach, dla każdego stanu sprawdzamy możliwe ruchy i akcje; dla nich wyznaczane są nagrody (w tym ujemne za wyjście poza mapę) oraz prawdopodobieństwa ustalone dla każdej akcji w tym stanie. Wyliczamy średnie nagrody, średnie v , ostatecznie dodajemy do tablicy V maksymalny uzyskany wynik i obliczamy błąd pomiędzy v_pom a V od danego stanu i kontynuujemy pętlę.

Wyniki:

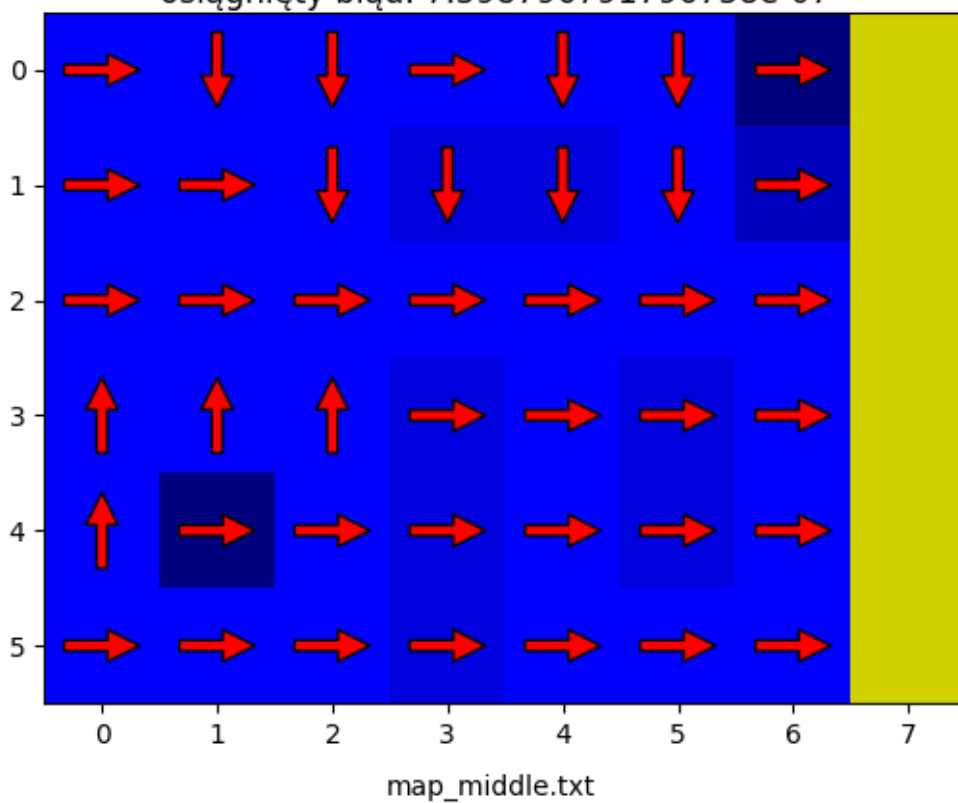
Ustalony błąd do osiągnięcia: 0.000001; Gamma: 0.9; Limit iteracji: 1000



Końcowa średnia suma nagród: 6.65,
liczba iteracji: 36,
osiągnięty błąd: 8.759055141194239e-07



Końcowa średnia suma nagród: 7.703,
liczba iteracji: 38,
osiągnięty błąd: 7.598796791796758e-07



Końcowa średnia suma nagród: 183.324,
liczba iteracji: 71,
osiągnięty błąd: 8.056439972925489e-07

