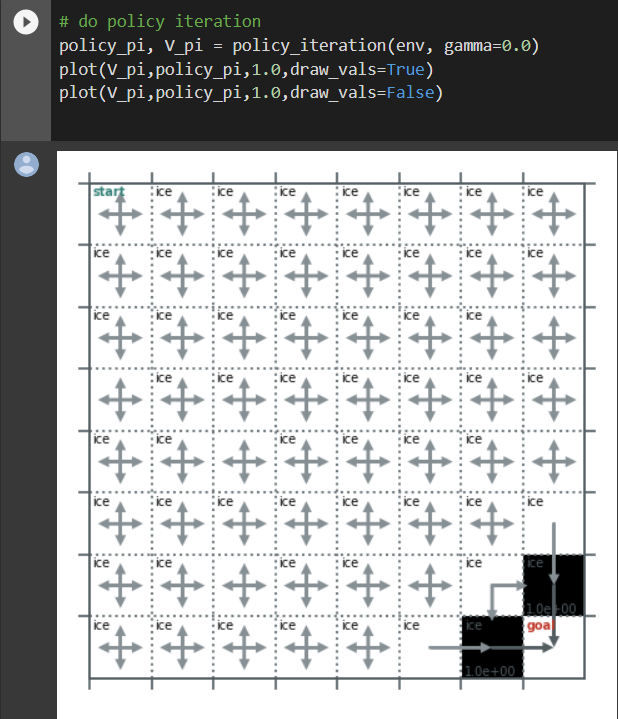
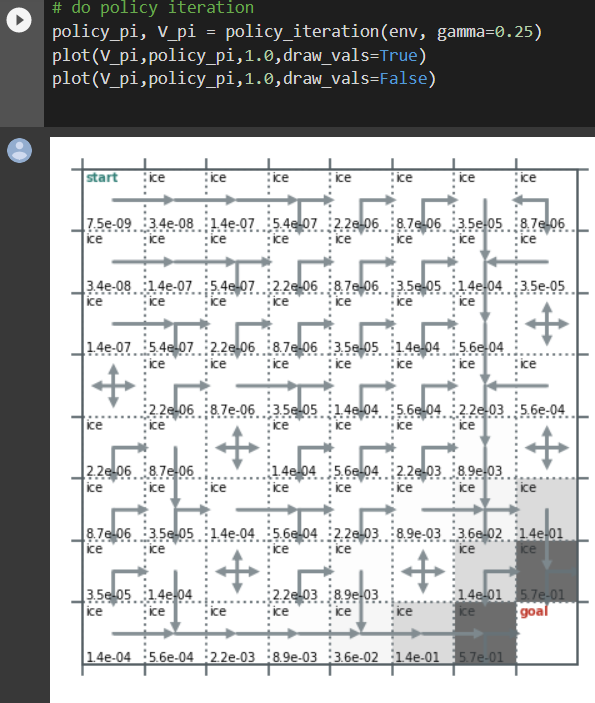
Zadanie 2/2 SWD

Mateusz Mika

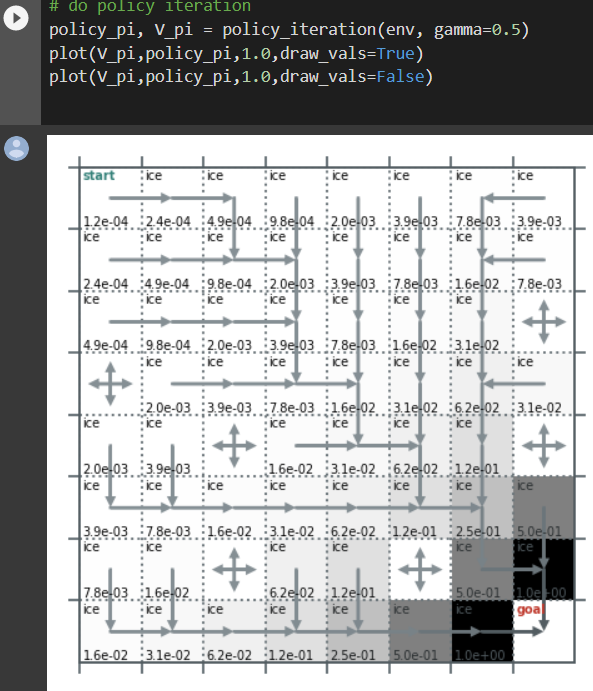
Test Frozen Lake dla Gamma 0 (policy iteration)



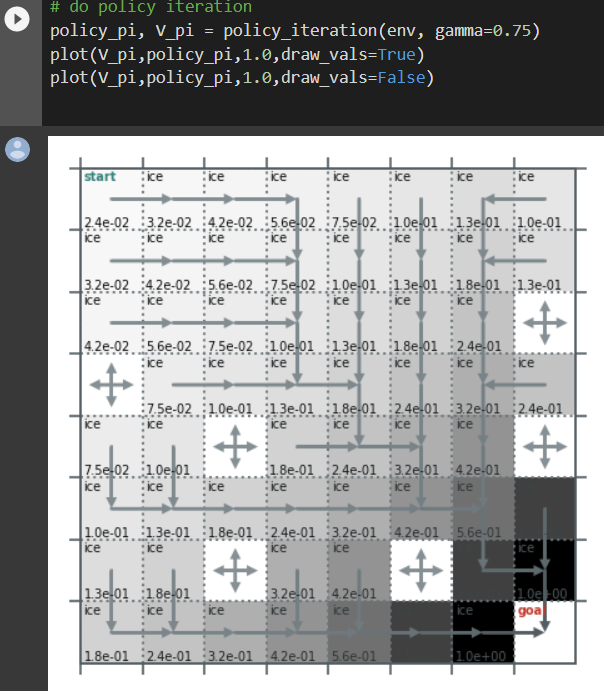
Test dla Gamma 0.25 (policy iteration)



Test dla Gamma 0.5 (policy iteration)



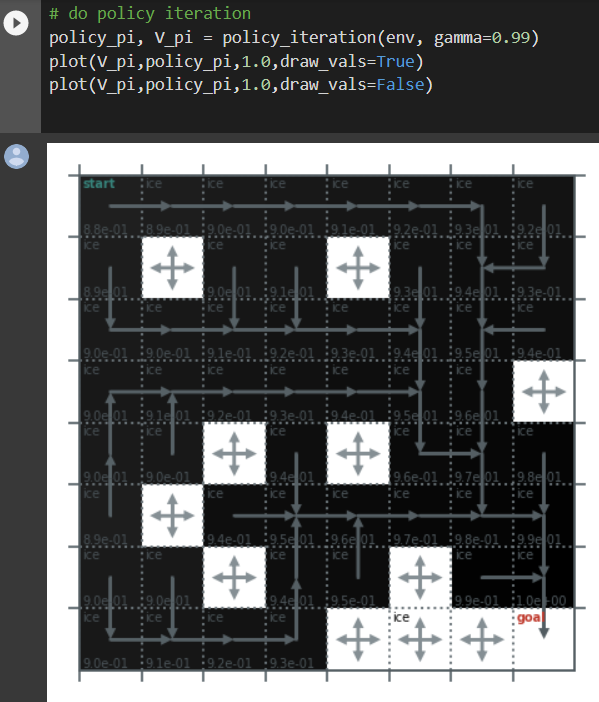
Test dla Gamma 0.75 (policy iteration)



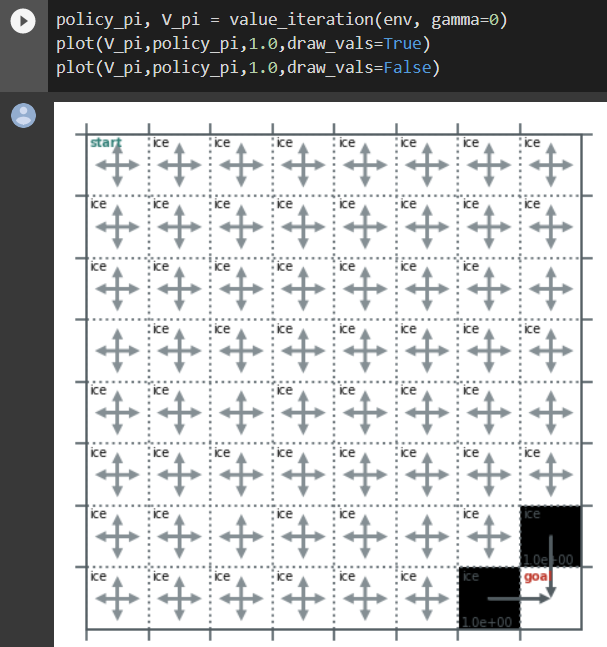
Test dla Gamma 1 (policy iteration)

Dla wartości gamma = 1 funkcja wykonuje się bez końca aż do wyświetlenia błędu

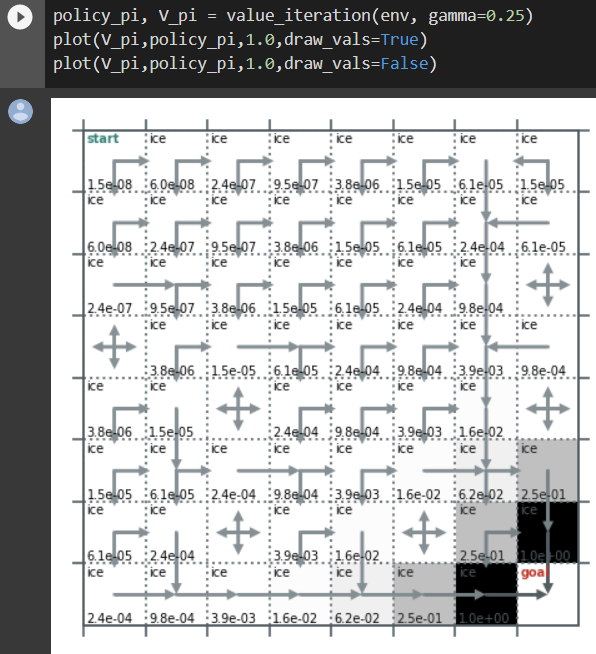
Użycie wartości 0.99 daje za to wynik jak poniżej:



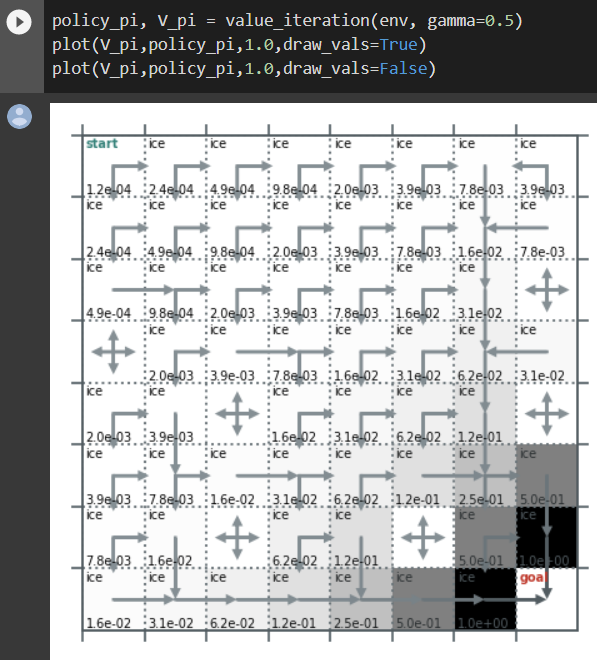
Test dla Gamma 0 (value iteration)



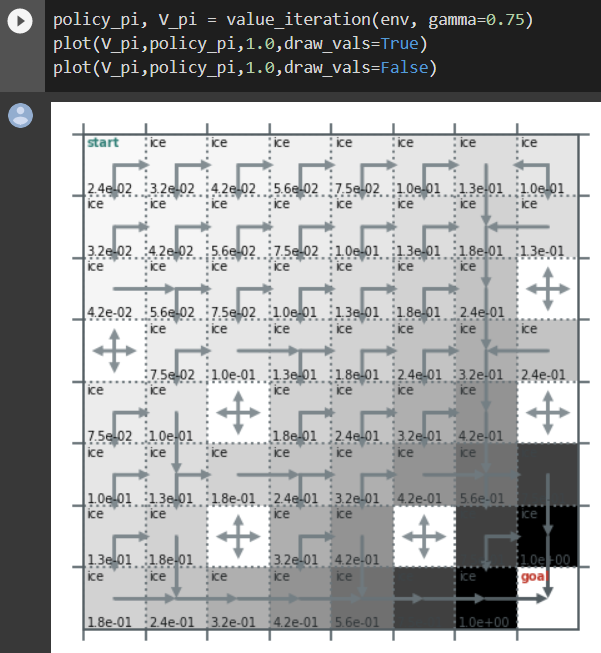
Test dla Gamma 0.25 (value iteration)



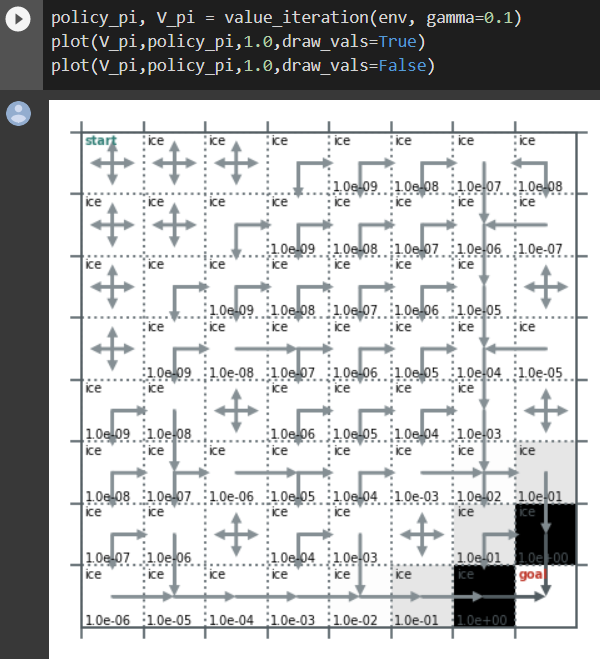
Test dla Gamma 0.5 (value iteration)



Test dla Gamma 0.75 (value iteration)



Test dla Gamma 1 (value iteration)



Zadanie1.

1.

Zwiększenie współczynnika gamma w programi FrozenLake dla policy\_iteration jak również value\_iteration powoduje zwiękrzenie dokładności

estymacji. Zaczynając od wartości gamma 0 zmierzając w strone wartości 1 rośnie ilość prób wykonanych na ścieżce do celu. Przy czym zadanie

wartości gamma równej 1 spowodowało zawieszenie działąnia algorytmu.

2. Metoda

policy\_iteration na każdym kroku wykonywana jest ocena zasad i czynność jest przeprowadzana aż do momentu uzyskania zbieżności,

a następnie zasady są aktualizowane i proces się powtarza.

value\_iteration natomiast wykonuje tylko pojedynczą iterację na każdym kroku. Następnie dla każdego stanu przyjmuje maksymalną

wartość działania jako szacowaną wartość stanu. Gdy te wartości stanu zbiegną się do wartości optymalnych,

wówczas można uzyskać optymalną politykę.

3.

W praktyce value\_iteration działa znacznie lepiej niż policy\_iteration i znajduje optymalną funkcję wartości stanu w znacznie mniejszej

liczbie kroków.