Sprawozdanie

Przedmiot: Odkrywanie związków w danych wielowymiarowych

Autorka: Monika Besler

Zadanie 1:

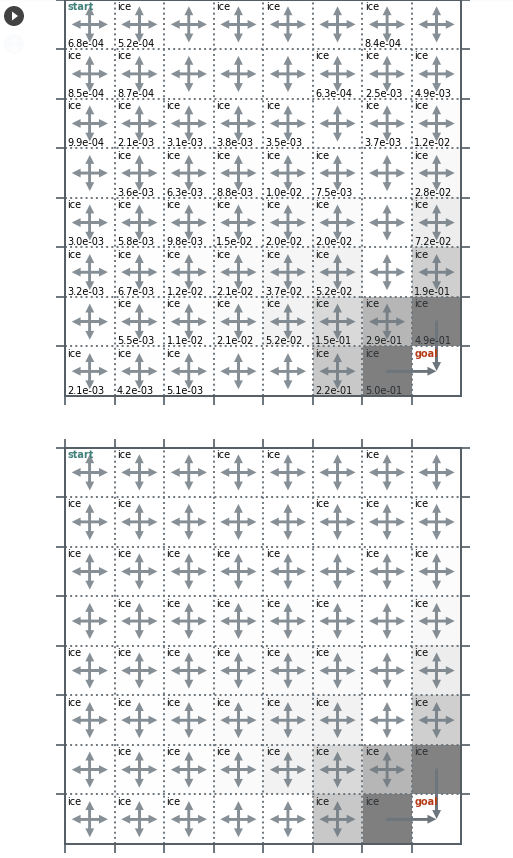
Obraz zawierający tekst

Opis wygenerowany automatycznie

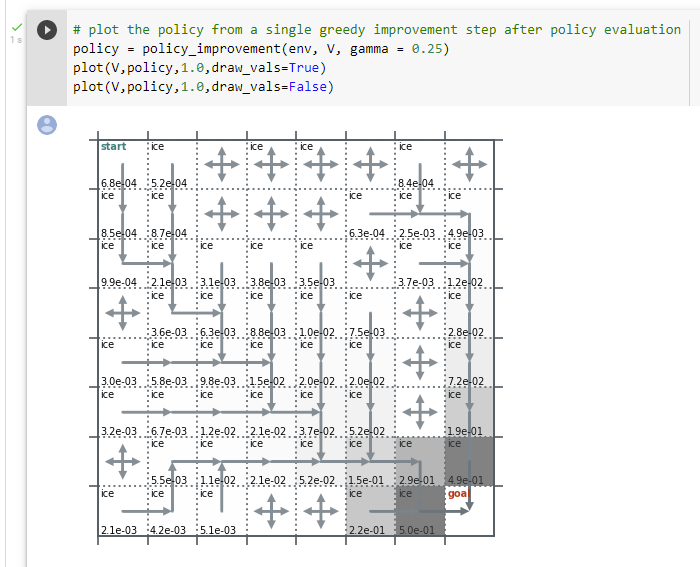
Wartość gamma = 0.0

Obraz zawierający tekst

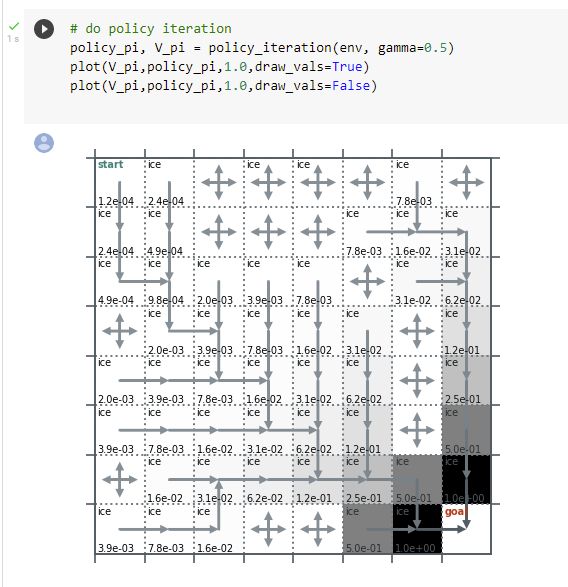
Opis wygenerowany automatycznie



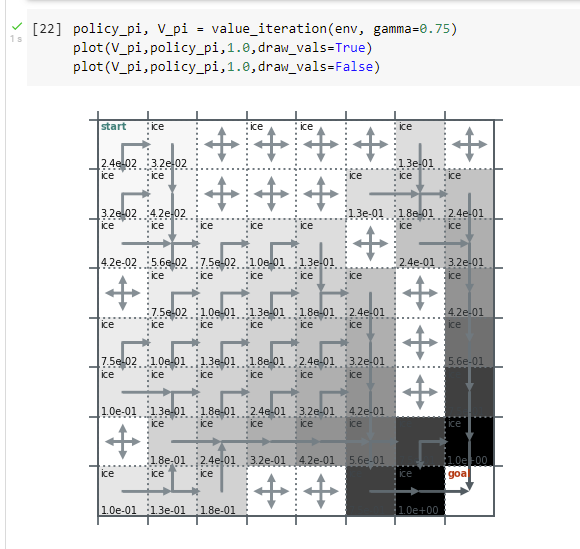
Wartość gamma = 0.25



Wartość gamma = 0.5



Wartość gamma = 0.25



Wartość gamma = 1

Obraz zawierający tekst, krzyżówka

Opis wygenerowany automatycznie

1. Jaki wpływ na jakość znalezienia drogi do frisbee ma wartość gamma,

W tym przypadku optymalną wartością gamma jest 0.5, ponieważ ważne jest, wyważenie krótko i długoterminowych zysków. Zarówno w przypadku wartości 0, jaki i wartości 1 – algorytm zostawia dużą dowolność w szukaniu drogi, a proponowana droga jest mniej efektywna i dalsza od optymalnej.

1. Czym różnią się metody policy\_iteration oraz value\_iteration?

Zarówno policy\_iteration jak i value\_iteration są algorytmami programowania dynamicznego, które znajdują optymalną politykę pi w środowisku uczenia się ze wzmocnieniem.

W policy\_iteration zaczynamy od ustalonej polityki. I odwrotnie, w value\_iteration od wyboru funkcji wartości. Następnie, w obu algorytmach, iteracyjnie wprowadzamy poprawki aż do osiągnięcia zbieżności.

1. Dlaczego value\_iteration znajduje lepszą drogę niż policy\_iteration?

Algorytm value\_iteration jest prosty. Łączy on dwie fazy iteracji polityki w jedną operację aktualizacji. Jednak funkcja iteracji wartości przechodzi przez wszystkie możliwe działania naraz, aby znaleźć maksymalną wartość działania. W związku z tym algorytm iteracji wartości jest cięższy obliczeniowo.

Oba algorytmy mają gwarancję zbieżności do optymalnej polityki na końcu. Jednak algorytm policy\_iteration osiąga zbieżność w mniejszej liczbie iteracji. W związku z tym algorytm policy\_iteration jest uważany za szybszy od algorytmu value\_iteration.