Sprawozdanie

Zadanie 1.1

1. wygenerowany labirynt 15, 15 A\*

Obraz zawierający zrzut ekranu, diagram, kwadrat, Prostokąt

Opis wygenerowany automatycznie

1. wygenerowany labirynt 25, 25 A\*

Obraz zawierający zrzut ekranu, kwadrat, diagram, Prostokąt

Opis wygenerowany automatycznie

1. wygenerowany labirynt 35, 35 A\*

Obraz zawierający zrzut ekranu, diagram, kwadrat, Prostokąt

Opis wygenerowany automatycznie

1. wygenerowany labirynt 15, 15 BFS

Obraz zawierający zrzut ekranu, Oprogramowanie graficzne, Oprogramowanie multimedialne, diagram

Opis wygenerowany automatycznie

1. wygenerowany labirynt 25, 25 BFS

Obraz zawierający zrzut ekranu, diagram, kwadrat, piksel

Opis wygenerowany automatycznie

1. wygenerowany labirynt 35, 35 BFS

Obraz zawierający zrzut ekranu, diagram, kwadrat, sztuka

Opis wygenerowany automatycznie

A\* jest często szybszy niż BFS dzięki wykorzystaniu heurystyki i bardziej skoncentrowanemu podejściu, ale wymaga on zdefiniowania odpowiedniej heurystyki dla danego problemu. Jeśli jednak nie ma dostępnej dobrej heurystyki, BFS może być bardziej niezawodny, chociaż kosztem potencjalnej straty efektywności.

1. Strategia przeszukiwania:

A\*: Wykorzystuje heurystykę w celu oszacowania kosztu od aktualnego węzła do celu.

Obejmuje zarówno rzeczywiste koszty przejścia, jak i heurystyczną ocenę od danego węzła do celu.

Podejście bardziej zorientowane celowo, skupia się na obszarach, które wydają się być bardziej obiecujące.

BFS: Przechodzi przez węzły w kolejności wszerz, sprawdzając najpierw wszystkie węzły na jednym poziomie przed przejściem do następnego poziomu.

Nie korzysta z heurystyki ani rzeczywistych kosztów przejścia. Rozważa wszystkie węzły na danym poziomie przed przechodzeniem do kolejnego.

1. Koszty:

A\*: Minimalizuje koszty, sumując rzeczywiste koszty przejścia od początku do danego węzła i heurystyczną ocenę od tego węzła do celu.

Skupia się na znalezieniu najbardziej efektywnej ścieżki, biorąc pod uwagę zarówno rzeczywiste koszty, jak i przewidywane koszty do celu.

BFS: Przeszukuje graf przy stałym koszcie kroku, bez uwzględniania heurystyki ani rzeczywistych kosztów przejścia.

Skupia się na znalezieniu najkrótszej ścieżki, ale nie uwzględnia ewentualnych różnic w kosztach między poszczególnymi ścieżkami.

1. Celność i optymalność:

A\*: Dzięki heurystyce często znajduje optymalne ścieżki.

Jest bardziej zorientowany celowo, co pozwala na skupienie się na obszarach prowadzących do celu.

BFS: Znajduje najkrótszą ścieżkę, ale nie zawsze jest optymalny w sensie minimalizacji kosztów rzeczywistego przejścia.

Może prowadzić do rozważenia wielu niepotrzebnych węzłów w poszukiwaniu najkrótszej ścieżki.

Zadanie 1.2

1. A\* x=5 y=10 BFS x=10 y=5

Obraz zawierający zrzut ekranu, diagram, kwadrat, piksel

Opis wygenerowany automatycznie

1. A\* x=18 y=24 BFS x=24 y=18

Obraz zawierający zrzut ekranu, diagram

Opis wygenerowany automatycznie

1. A\* x=31 y=25 BFS x=25 y=31

Obraz zawierający zrzut ekranu, diagram, kwadrat, piksel

Opis wygenerowany automatycznie

1. Przeszukiwanie rekurencyjne labiryntu za pomocą algorytmu DFS 15, 15

Obraz zawierający zrzut ekranu, kwadrat, diagram, Prostokąt

Opis wygenerowany automatycznie

1. Przeszukiwanie rekurencyjne labiryntu za pomocą algorytmu DFS 25, 25

Obraz zawierający zrzut ekranu, kwadrat, Prostokąt, wzór

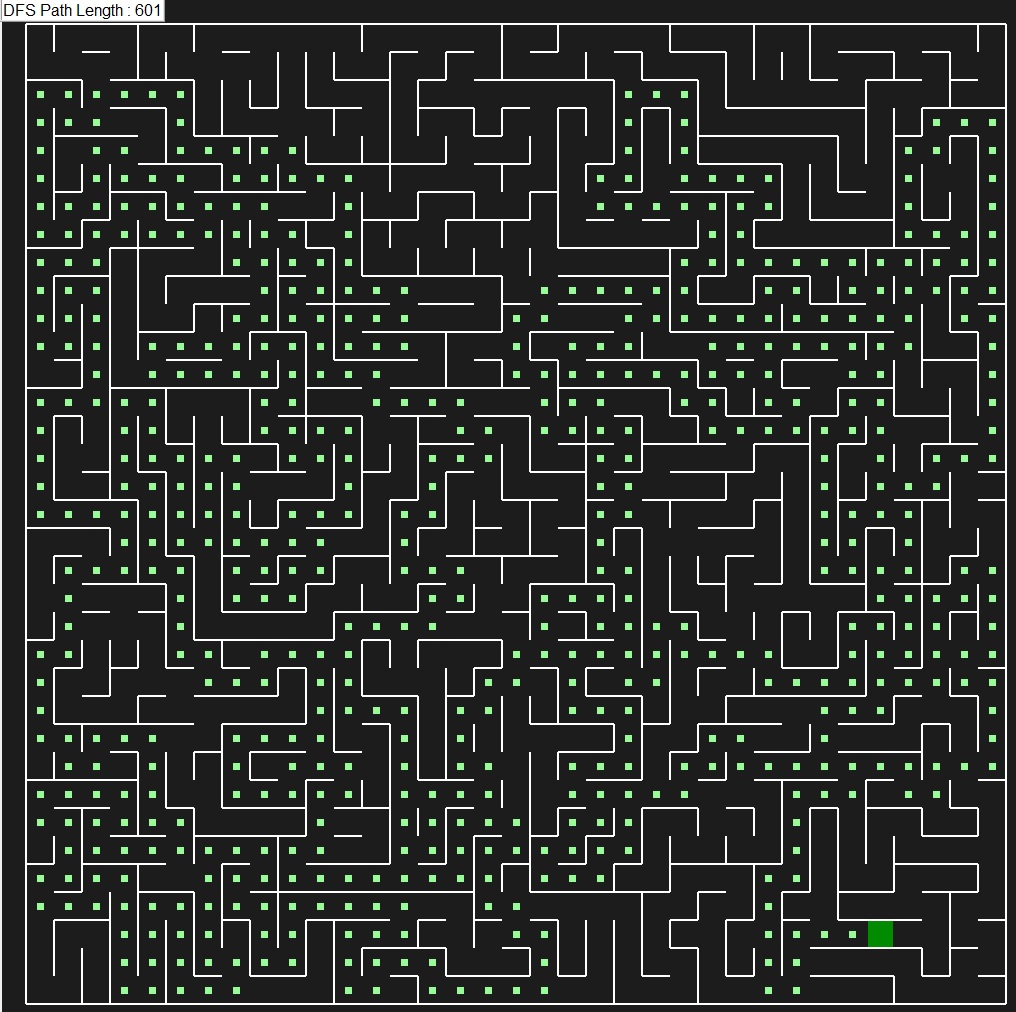
Opis wygenerowany automatycznie

1. Przeszukiwanie rekurencyjne labiryntu za pomocą algorytmu DFS 25, 25

Obraz zawierający zrzut ekranu, wzór, tekst, diagram

Opis wygenerowany automatycznie

1. Przeszukiwanie rekurencyjne labiryntu za pomocą algorytmu DFS 35, 35



Zadanie 2.1

Niestety na gammie 1 zadanie nie chciało się rozwiązać więc zmieniłem 1 na 0.99

Gamma 0 policy\_iteration

Obraz zawierający tekst, kwadrat, zrzut ekranu, numer

Opis wygenerowany automatycznie

Gamma 0 value\_iteration

Obraz zawierający tekst, kwadrat, zrzut ekranu, numer

Opis wygenerowany automatycznie

Gamma 0.25 policy\_iteration

Obraz zawierający tekst, diagram, numer, Plan

Opis wygenerowany automatycznie

Gamma 0.25 value\_iteration

Obraz zawierający tekst, diagram, Plan, numer

Opis wygenerowany automatycznie

Gamma 0.5 policy\_iteration

Obraz zawierający tekst, diagram, krzyżówka, Plan

Opis wygenerowany automatycznie

Gamma 0.5 value\_iteration

Obraz zawierający tekst, diagram, Plan, kwadrat

Opis wygenerowany automatycznie

Gamma 0.75 policy\_iteration

Obraz zawierający tekst, diagram, Plan, wykres

Opis wygenerowany automatycznie

Gamma 0.75 value\_iteration

Obraz zawierający tekst, diagram, Plan, kwadrat

Opis wygenerowany automatycznie

Gamma 0.99 policy\_iteration

Obraz zawierający Prostokąt, kwadrat, zrzut ekranu, diagram

Opis wygenerowany automatycznie

Gamma 0.99 value\_iteration

Obraz zawierający kwadrat, Prostokąt, wzór, zrzut ekranu

Opis wygenerowany automatycznie

1. Jaki wpływ na jakość znalezienia drogi do frisbee ma wartość gamma - Wartość gamma wpływa na jakość znalezienia drogi do frisbee, regulując uwzględnianie przyszłych nagród w decyzjach algorytmów.
2. Czym różnią się metody policy\_iteration oraz value\_iteration? - Metoda policy\_iteration polega na cyklicznym procesie oceny i poprawy polityki, gdzie początkowo analizuje się aktualną politykę, a następnie ją doskonali. Z kolei metoda value\_iteration bezpośrednio wylicza optymalną wartość stanu i następnie tworzy politykę opartą na tych wartościach.
3. Dlaczego value\_iteration znajduje lepszą drogę niż policy\_iteration - Value\_iteration jest skuteczniejsza niż policy\_iteration, ponieważ bezpośrednio oblicza optymalne wartości stanów, omijając analizę początkowej polityki.

Zadanie 2.2

Gamma 0 policy\_iteration

Obraz zawierający tekst, zrzut ekranu, kwadrat, Prostokąt

Opis wygenerowany automatycznie

Gamma 0 value\_iteration

Obraz zawierający tekst, kwadrat, numer, zrzut ekranu

Opis wygenerowany automatycznie

Gamma 0.2 policy\_iteration

Obraz zawierający tekst, zrzut ekranu, numer, Czcionka

Opis wygenerowany automatycznie

Gamma 0.2 value\_iteration

Obraz zawierający tekst, zrzut ekranu, krzyżówka, kwadrat

Opis wygenerowany automatycznie

Gamma 0.4 policy\_iteration

Obraz zawierający tekst, zrzut ekranu, numer, Czcionka

Opis wygenerowany automatycznie

Gamma 0.4 value\_iteration

Obraz zawierający tekst, zrzut ekranu, numer, Czcionka

Opis wygenerowany automatycznie

Gamma 0.6 policy\_iteration

Obraz zawierający tekst, zrzut ekranu, numer, Czcionka

Opis wygenerowany automatycznie

Gamma 0.6 value\_iteration

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Gamma 0.8 policy\_iteration

Obraz zawierający tekst, zrzut ekranu, numer, Czcionka

Opis wygenerowany automatycznie

Gamma 0.8 value\_iteration

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

Gamma 0.99 policy\_iteration

Obraz zawierający tekst, zrzut ekranu, krzyżówka, kwadrat

Opis wygenerowany automatycznie

Gamma 0.99 value\_iteration

Obraz zawierający tekst, zrzut ekranu, Czcionka, numer

Opis wygenerowany automatycznie

1. Jak poślizg na lodzie ma wpływ na skomplikowanie trasy - Poślizg na lodzie istotnie zwiększa trudność trasy, wprowadzając element losowości w działaniu agenta. Agent może napotykać trudności w precyzyjnym przewidywaniu, jakie konkretne akcje doprowadzą go do celu, co skutkuje chaotyczniejszą trasą.
2. Czy wartość gamma ma wpływ na skomplikowanie trasy, a jeżeli tak, to jakie. -Tak, wartość gamma wpływa na skomplikowanie trasy. Wyższa wartość gamma może prowadzić do bardziej długotrwałych, lecz bardziej przemyślanych ścieżek, podczas gdy niższa wartość gamma może skłaniać do krótkoterminowych i bardziej ryzykownych decyzji.

Zadanie 2.3

Gamma 0 policy\_iteration

Obraz zawierający tekst, krzyżówka, zrzut ekranu, kwadrat

Opis wygenerowany automatycznie

Gamma 0 value\_iteration

Obraz zawierający tekst, numer, zrzut ekranu, kwadrat

Opis wygenerowany automatycznie

Gamma 0 value\_iteration\_2

Obraz zawierający tekst, krzyżówka, numer, kwadrat

Opis wygenerowany automatycznie

Gamma 0.2 policy\_iteration

Obraz zawierający tekst, diagram, krzyżówka, numer

Opis wygenerowany automatycznie

Gamma 0.2 value\_iteration

Obraz zawierający tekst, diagram, Plan, numer

Opis wygenerowany automatycznie

Gamma 0.2 value\_iteration\_2

Obraz zawierający tekst, diagram, Plan, numer

Opis wygenerowany automatycznie

Gamma 0.4 policy\_iteration

Obraz zawierający tekst, diagram, numer, Plan

Opis wygenerowany automatycznie

Gamma 0.4 value\_iteration

Obraz zawierający tekst, diagram, Plan, numer

Opis wygenerowany automatycznie

Gamma 0.4 value\_iteration\_2

Obraz zawierający tekst, diagram, numer, Czcionka

Opis wygenerowany automatycznie

Gamma 0.6 policy\_iteration

Obraz zawierający tekst, diagram, numer, Plan

Opis wygenerowany automatycznie

Gamma 0.6 value\_iteration

Obraz zawierający tekst, diagram, numer, Plan

Opis wygenerowany automatycznie

Gamma 0.6 value\_iteration\_2

Obraz zawierający tekst, diagram, numer, Czcionka

Opis wygenerowany automatycznie

Gamma 0.8 policy\_iteration

Obraz zawierający tekst, diagram, krzyżówka, numer

Opis wygenerowany automatycznie

Gamma 0.8 value\_iteration

Obraz zawierający tekst, diagram, Plan, wykres

Opis wygenerowany automatycznie

Gamma 0.8 value\_iteration\_2

Obraz zawierający tekst, diagram, Plan, numer

Opis wygenerowany automatycznie

Gamma 0.99 policy\_iteration

Obraz zawierający tekst, krzyżówka, diagram, kwadrat

Opis wygenerowany automatycznie

Gamma 0.99 value\_iteration

Obraz zawierający tekst, krzyżówka, diagram, kwadrat

Opis wygenerowany automatycznie

Gamma 0.99 value\_iteration\_2

Obraz zawierający tekst, krzyżówka, diagram, kwadrat

Opis wygenerowany automatycznie

1. Czym się różni metoda value\_iteration od value\_iteration\_2? - Wartościowanie różniące się między value\_iteration a value\_iteration\_2 polega na tym, że value\_iteration jest bardziej zoptymalizowane pod względem wydajności, obliczając jednocześnie wartości stanów i funkcję akcji. Z kolei value\_iteration\_2 to bardziej rozbudowana wersja, która oddzielnie oblicza Q-function, wykorzystując jedno kroki do przodu. Oba algorytmy prowadzą do podobnych wyników, ale value\_iteration\_2 jest bardziej rozbudowaną i zrozumiałą implementacją.

Zadanie 3

Jak dobrze zachowuje się system po wykonaniu różnej ilości epizodów. Pokazać konkretne zachowania, które się zmieniły pomiędzy małą ilością epizodów, a dużą ilością epizodów.

W większości przypadków widzieliśmy zdecydowany wzrost running reword względem epizodów aczkolwiek przy niższej gammie (0.5 lub mniej) zauważaliśmy znaczny spadek wydajności i większą ilość epizodów która w niektórych przypadkach oscylowała w granicach 10 pkt w running reword.

Jak gamma ma wpływ na szybkość uczenia i zachowanie się systemu. Czy widać znaczący wpływ tego parametru?

Gamma istotnie wpływa na tempo uczenia i zachowanie systemu. Niska wartość gamma skupia system na krótkoterminowych nagrodach, co może przyspieszyć uczenie, ale prowadzić do impulsywnego zachowania. Wysoka wartość gamma sprzyja bardziej długoterminowemu planowaniu, co może spowolnić proces uczenia, ale prowadzić do bardziej zrównoważonego zachowania. W przypadku tego zadania na niższej gammie nie dało się rozwiązać zadania w sensownym czasie (odnoszę się do max\_steps\_per\_episode=5000 i gamma = 0.20).

1000 epizodów oraz Gamma 0.9

Obraz zawierający tekst, zrzut ekranu, Czcionka, czarne i białe

Opis wygenerowany automatycznie

running reward: 6.52 at episode 10

running reward: 11.43 at episode 20

running reward: 11.69 at episode 30

running reward: 12.04 at episode 40

running reward: 11.54 at episode 50

running reward: 11.08 at episode 60

running reward: 10.92 at episode 70

running reward: 10.58 at episode 80

running reward: 10.01 at episode 90

running reward: 10.24 at episode 100

running reward: 10.33 at episode 110

running reward: 10.26 at episode 120

running reward: 10.42 at episode 130

running reward: 10.45 at episode 140

running reward: 10.35 at episode 150

running reward: 10.44 at episode 160

running reward: 10.33 at episode 170

running reward: 10.38 at episode 180

running reward: 10.59 at episode 190

running reward: 11.57 at episode 200

running reward: 12.53 at episode 210

running reward: 19.48 at episode 220

running reward: 27.20 at episode 230

running reward: 30.27 at episode 240

running reward: 24.20 at episode 250

running reward: 22.10 at episode 260

running reward: 41.82 at episode 270

running reward: 51.20 at episode 280

running reward: 63.35 at episode 290

running reward: 50.25 at episode 300

running reward: 67.52 at episode 310

Solved at episode 316!

2000 epizodów oraz Gamma 0.9

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie  
running reward: 6.86 at episode 10

running reward: 11.24 at episode 20

running reward: 12.18 at episode 30

running reward: 13.43 at episode 40

running reward: 13.23 at episode 50

running reward: 13.11 at episode 60

running reward: 13.73 at episode 70

running reward: 16.96 at episode 80

running reward: 17.11 at episode 90

running reward: 14.95 at episode 100

running reward: 13.95 at episode 110

running reward: 12.82 at episode 120

running reward: 13.62 at episode 130

running reward: 13.85 at episode 140

running reward: 13.18 at episode 150

running reward: 13.61 at episode 160

running reward: 13.93 at episode 170

running reward: 24.93 at episode 180

running reward: 36.56 at episode 190

running reward: 34.82 at episode 200

running reward: 34.99 at episode 210

running reward: 41.03 at episode 220

running reward: 34.39 at episode 230

running reward: 31.47 at episode 240

running reward: 32.31 at episode 250

running reward: 38.09 at episode 260

running reward: 64.62 at episode 270

running reward: 84.89 at episode 280

running reward: 85.41 at episode 290

Solved at episode 298!

10000 epizodów oraz Gamma 0.9

Obraz zawierający tekst, Czcionka, biały, typografia

Opis wygenerowany automatycznie

Obraz zawierający tekst, Czcionka, zrzut ekranu, typografia

Opis wygenerowany automatycznie

Obraz zawierający tekst, Czcionka, typografia

Opis wygenerowany automatycznie

running reward: 6.97 at episode 10

running reward: 13.14 at episode 20

running reward: 15.83 at episode 30

running reward: 17.22 at episode 40

running reward: 18.80 at episode 50

running reward: 25.55 at episode 60

running reward: 35.66 at episode 70

running reward: 33.33 at episode 80

running reward: 50.56 at episode 90

running reward: 47.75 at episode 100

running reward: 69.80 at episode 110

Solved at episode 119!

5000 epizodów oraz Gamma 0.2

Obraz zawierający tekst, Czcionka, zrzut ekranu, typografia

Opis wygenerowany automatycznie

running reward: 6.06 at episode 10

running reward: 8.29 at episode 20

running reward: 10.55 at episode 30

running reward: 11.37 at episode 40

running reward: 11.82 at episode 50

running reward: 11.55 at episode 60

running reward: 10.98 at episode 70

running reward: 10.98 at episode 80

running reward: 10.81 at episode 90

running reward: 10.38 at episode 100

running reward: 10.46 at episode 110

running reward: 10.05 at episode 120

running reward: 9.95 at episode 130

running reward: 9.96 at episode 140

running reward: 10.24 at episode 150

running reward: 10.03 at episode 160

running reward: 10.08 at episode 170

running reward: 9.59 at episode 180

running reward: 9.48 at episode 190

running reward: 9.36 at episode 200

running reward: 9.42 at episode 210

running reward: 9.59 at episode 220

running reward: 9.51 at episode 230

running reward: 9.92 at episode 240

running reward: 9.91 at episode 250

running reward: 9.63 at episode 260

running reward: 9.76 at episode 270

running reward: 10.09 at episode 280

running reward: 10.20 at episode 290

running reward: 10.10 at episode 300

running reward: 10.01 at episode 310

running reward: 9.80 at episode 320

running reward: 9.97 at episode 330

running reward: 10.66 at episode 340

running reward: 10.58 at episode 350

running reward: 11.23 at episode 360

running reward: 11.45 at episode 370

running reward: 11.43 at episode 380

running reward: 12.75 at episode 390

running reward: 12.68 at episode 400

running reward: 12.97 at episode 410

running reward: 13.22 at episode 420

running reward: 14.05 at episode 430

running reward: 17.13 at episode 440

running reward: 33.14 at episode 450

running reward: 28.77 at episode 460

running reward: 27.19 at episode 470

running reward: 28.90 at episode 480

running reward: 44.55 at episode 490

running reward: 75.59 at episode 500

Solved at episode 505!

5000 epizodów oraz Gamma 0.5

Obraz zawierający tekst, ubrania, materiał, wzór

Opis wygenerowany automatycznie

running reward: 10.63 at episode 10

running reward: 25.74 at episode 20

running reward: 44.65 at episode 30

running reward: 64.22 at episode 40

running reward: 72.51 at episode 50

running reward: 74.62 at episode 60

running reward: 100.61 at episode 70

Solved at episode 70!

5000 epizodów oraz Gamma 0.9

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie



running reward: 7.97 at episode 10

running reward: 10.39 at episode 20

running reward: 11.94 at episode 30

running reward: 12.89 at episode 40

running reward: 12.08 at episode 50

running reward: 12.92 at episode 60

running reward: 13.68 at episode 70

running reward: 12.94 at episode 80

running reward: 13.20 at episode 90

running reward: 21.33 at episode 100

running reward: 33.90 at episode 110

running reward: 74.22 at episode 120

Solved at episode 125!

5000 epizodów oraz Gamma 0.99

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie

running reward: 8.00 at episode 10

running reward: 15.42 at episode 20

running reward: 14.56 at episode 30

running reward: 13.40 at episode 40

running reward: 12.67 at episode 50

running reward: 12.77 at episode 60

running reward: 13.05 at episode 70

running reward: 13.13 at episode 80

running reward: 13.51 at episode 90

running reward: 15.75 at episode 100

running reward: 31.99 at episode 110

running reward: 60.97 at episode 120

running reward: 82.10 at episode 130

running reward: 77.87 at episode 140

running reward: 73.49 at episode 150

running reward: 77.65 at episode 160

Solved at episode 167!

5000 epizodów oraz Gamma 0.1 (Niestety ale wykonałem to 3 razy i ilość epizodów przekroczyła 10 000 ale i tak nie udało mi się zakończyć tego z sukcesem)

running reward: 6.42 at episode 10

running reward: 13.05 at episode 20

running reward: 18.53 at episode 30

running reward: 19.65 at episode 40

running reward: 19.36 at episode 50

running reward: 16.33 at episode 60

running reward: 15.84 at episode 70

running reward: 14.67 at episode 80

running reward: 13.01 at episode 90

running reward: 12.19 at episode 100

running reward: 11.48 at episode 110

running reward: 11.14 at episode 120

running reward: 12.03 at episode 130

running reward: 12.04 at episode 140

running reward: 13.79 at episode 150

running reward: 14.26 at episode 160

running reward: 18.63 at episode 170

running reward: 18.89 at episode 180

running reward: 17.41 at episode 190

running reward: 14.81 at episode 200

running reward: 12.82 at episode 210

running reward: 11.79 at episode 220

running reward: 12.29 at episode 230

running reward: 11.36 at episode 240

running reward: 11.28 at episode 250

running reward: 11.26 at episode 260

running reward: 10.90 at episode 270

running reward: 10.71 at episode 280

running reward: 11.32 at episode 290

running reward: 11.79 at episode 300

running reward: 12.34 at episode 310

running reward: 12.87 at episode 320

running reward: 13.03 at episode 330

running reward: 12.66 at episode 340

running reward: 12.15 at episode 350

running reward: 12.45 at episode 360

running reward: 11.93 at episode 370

running reward: 11.15 at episode 380

running reward: 11.37 at episode 390

running reward: 10.82 at episode 400

running reward: 10.60 at episode 410

running reward: 10.41 at episode 420

running reward: 10.19 at episode 430

running reward: 11.08 at episode 440

running reward: 10.43 at episode 450

running reward: 10.27 at episode 460

running reward: 10.73 at episode 470

running reward: 11.15 at episode 480

running reward: 10.85 at episode 490

running reward: 10.66 at episode 500

running reward: 10.85 at episode 510

running reward: 10.85 at episode 520

running reward: 10.93 at episode 530

running reward: 10.72 at episode 540

running reward: 10.54 at episode 550

running reward: 10.28 at episode 560

running reward: 10.50 at episode 570

running reward: 10.45 at episode 580

running reward: 10.67 at episode 590

running reward: 10.81 at episode 600

running reward: 10.63 at episode 610

running reward: 10.66 at episode 620

running reward: 10.92 at episode 630

running reward: 10.76 at episode 640

running reward: 13.04 at episode 650

running reward: 12.45 at episode 660

running reward: 12.32 at episode 670

running reward: 11.31 at episode 680

running reward: 11.72 at episode 690

running reward: 11.78 at episode 700

running reward: 11.13 at episode 710

running reward: 10.73 at episode 720

running reward: 10.50 at episode 730

running reward: 10.16 at episode 740

running reward: 10.04 at episode 750

running reward: 10.28 at episode 760

running reward: 10.28 at episode 770

running reward: 9.95 at episode 780

running reward: 9.57 at episode 790

running reward: 9.56 at episode 800

running reward: 9.52 at episode 810

running reward: 10.22 at episode 820

running reward: 10.15 at episode 830

running reward: 10.34 at episode 840

running reward: 10.56 at episode 850

running reward: 10.64 at episode 860

running reward: 10.59 at episode 870

running reward: 10.48 at episode 880

running reward: 10.04 at episode 890

running reward: 9.86 at episode 900

running reward: 9.54 at episode 910

running reward: 9.54 at episode 920

running reward: 9.49 at episode 930

running reward: 9.85 at episode 940

running reward: 9.73 at episode 950

running reward: 9.78 at episode 960

running reward: 9.79 at episode 970

running reward: 9.89 at episode 980

running reward: 9.94 at episode 990

running reward: 9.92 at episode 1000

running reward: 9.96 at episode 1010

running reward: 9.84 at episode 1020

running reward: 9.90 at episode 1030

running reward: 9.74 at episode 1040

running reward: 9.55 at episode 1050

running reward: 9.77 at episode 1060

running reward: 9.73 at episode 1070

running reward: 9.80 at episode 1080

running reward: 9.81 at episode 1090

running reward: 9.52 at episode 1100

running reward: 9.42 at episode 1110

running reward: 9.41 at episode 1120

running reward: 9.26 at episode 1130

running reward: 9.87 at episode 1140

running reward: 9.79 at episode 1150

running reward: 9.65 at episode 1160

running reward: 9.59 at episode 1170

running reward: 9.78 at episode 1180

running reward: 9.81 at episode 1190

running reward: 10.32 at episode 1200

running reward: 10.23 at episode 1210

running reward: 10.40 at episode 1220

running reward: 10.13 at episode 1230

running reward: 9.79 at episode 1240

running reward: 9.67 at episode 1250

running reward: 9.74 at episode 1260

running reward: 9.83 at episode 1270

running reward: 9.83 at episode 1280

running reward: 9.60 at episode 1290

running reward: 9.66 at episode 1300

running reward: 9.50 at episode 1310

running reward: 9.31 at episode 1320

running reward: 9.35 at episode 1330

running reward: 9.43 at episode 1340

running reward: 9.40 at episode 1350

running reward: 9.44 at episode 1360

running reward: 9.60 at episode 1370

running reward: 9.59 at episode 1380

running reward: 9.76 at episode 1390

running reward: 9.52 at episode 1400

running reward: 9.44 at episode 1410

running reward: 9.56 at episode 1420

running reward: 9.78 at episode 1430

running reward: 9.71 at episode 1440

running reward: 9.75 at episode 1450

running reward: 9.61 at episode 1460

running reward: 9.52 at episode 1470

running reward: 9.48 at episode 1480

running reward: 9.83 at episode 1490

running reward: 9.98 at episode 1500

running reward: 9.90 at episode 1510

running reward: 9.76 at episode 1520

running reward: 9.75 at episode 1530

running reward: 9.70 at episode 1540

running reward: 9.54 at episode 1550

running reward: 9.64 at episode 1560

running reward: 9.57 at episode 1570

running reward: 9.51 at episode 1580

running reward: 9.48 at episode 1590

running reward: 9.33 at episode 1600

running reward: 9.52 at episode 1610

running reward: 9.50 at episode 1620

running reward: 9.51 at episode 1630

running reward: 9.77 at episode 1640

running reward: 9.59 at episode 1650

running reward: 9.67 at episode 1660

running reward: 9.60 at episode 1670

running reward: 9.74 at episode 1680

running reward: 9.52 at episode 1690

running reward: 9.47 at episode 1700

running reward: 9.85 at episode 1710

running reward: 9.75 at episode 1720

running reward: 9.62 at episode 1730

running reward: 9.69 at episode 1740

running reward: 9.48 at episode 1750

running reward: 9.49 at episode 1760

running reward: 9.65 at episode 1770

running reward: 9.50 at episode 1780

running reward: 9.50 at episode 1790

running reward: 9.78 at episode 1800

running reward: 9.64 at episode 1810

running reward: 9.44 at episode 1820

running reward: 9.41 at episode 1830

running reward: 9.57 at episode 1840

running reward: 9.36 at episode 1850

running reward: 9.33 at episode 1860

running reward: 9.48 at episode 1870

running reward: 9.56 at episode 1880

running reward: 9.40 at episode 1890

running reward: 9.31 at episode 1900

running reward: 9.30 at episode 1910

running reward: 9.46 at episode 1920

running reward: 9.61 at episode 1930

running reward: 9.38 at episode 1940

running reward: 9.46 at episode 1950

running reward: 9.38 at episode 1960

running reward: 9.38 at episode 1970

running reward: 9.56 at episode 1980

running reward: 9.37 at episode 1990

running reward: 9.50 at episode 2000

running reward: 9.47 at episode 2010

running reward: 9.40 at episode 2020

running reward: 9.22 at episode 2030

running reward: 9.27 at episode 2040

running reward: 9.30 at episode 2050

running reward: 9.39 at episode 2060

running reward: 9.59 at episode 2070

running reward: 9.48 at episode 2080

running reward: 9.73 at episode 2090

running reward: 9.84 at episode 2100

running reward: 9.50 at episode 2110

running reward: 9.44 at episode 2120

running reward: 9.38 at episode 2130

running reward: 9.29 at episode 2140

running reward: 9.31 at episode 2150

running reward: 9.45 at episode 2160

running reward: 9.50 at episode 2170

running reward: 9.54 at episode 2180

running reward: 9.40 at episode 2190

running reward: 9.29 at episode 2200

running reward: 9.43 at episode 2210

running reward: 9.49 at episode 2220

running reward: 9.52 at episode 2230

running reward: 9.49 at episode 2240

running reward: 9.53 at episode 2250

running reward: 9.48 at episode 2260

running reward: 9.43 at episode 2270

running reward: 9.54 at episode 2280

running reward: 9.52 at episode 2290

running reward: 9.35 at episode 2300

running reward: 9.43 at episode 2310

running reward: 9.37 at episode 2320

running reward: 9.49 at episode 2330

running reward: 9.68 at episode 2340

running reward: 9.58 at episode 2350

running reward: 9.80 at episode 2360

running reward: 10.01 at episode 2370

running reward: 9.65 at episode 2380

running reward: 9.46 at episode 2390

running reward: 9.58 at episode 2400

running reward: 9.61 at episode 2410

running reward: 9.48 at episode 2420

running reward: 9.48 at episode 2430

running reward: 9.18 at episode 2440

running reward: 9.27 at episode 2450

running reward: 9.34 at episode 2460

running reward: 9.44 at episode 2470

running reward: 9.34 at episode 2480

running reward: 9.34 at episode 2490

running reward: 9.35 at episode 2500

running reward: 9.38 at episode 2510

running reward: 9.52 at episode 2520

running reward: 9.55 at episode 2530

running reward: 9.64 at episode 2540

running reward: 9.35 at episode 2550

running reward: 9.34 at episode 2560

running reward: 9.23 at episode 2570

running reward: 9.32 at episode 2580

running reward: 9.56 at episode 2590

running reward: 9.51 at episode 2600

running reward: 9.63 at episode 2610

running reward: 9.76 at episode 2620

running reward: 9.66 at episode 2630

running reward: 9.68 at episode 2640

running reward: 9.59 at episode 2650

running reward: 9.46 at episode 2660

running reward: 9.49 at episode 2670

running reward: 9.40 at episode 2680

running reward: 9.27 at episode 2690

running reward: 9.26 at episode 2700

running reward: 9.27 at episode 2710

running reward: 9.32 at episode 2720

running reward: 9.41 at episode 2730

running reward: 9.25 at episode 2740

running reward: 9.46 at episode 2750

running reward: 9.50 at episode 2760

running reward: 9.33 at episode 2770

running reward: 9.34 at episode 2780

running reward: 9.13 at episode 2790

running reward: 9.40 at episode 2800

running reward: 9.25 at episode 2810

running reward: 9.54 at episode 2820

running reward: 10.15 at episode 2830

running reward: 10.26 at episode 2840

running reward: 10.46 at episode 2850

running reward: 10.60 at episode 2860

running reward: 10.94 at episode 2870

running reward: 11.55 at episode 2880

running reward: 13.03 at episode 2890

running reward: 17.74 at episode 2900

running reward: 19.04 at episode 2910

running reward: 16.29 at episode 2920

running reward: 14.22 at episode 2930

running reward: 12.64 at episode 2940

running reward: 11.68 at episode 2950

running reward: 10.84 at episode 2960

running reward: 10.12 at episode 2970

running reward: 9.87 at episode 2980

running reward: 9.94 at episode 2990

running reward: 10.00 at episode 3000

running reward: 9.89 at episode 3010

running reward: 9.97 at episode 3020

running reward: 9.96 at episode 3030

running reward: 9.81 at episode 3040

running reward: 9.77 at episode 3050

running reward: 9.59 at episode 3060

running reward: 9.70 at episode 3070

running reward: 9.40 at episode 3080

running reward: 9.42 at episode 3090

running reward: 9.52 at episode 3100

running reward: 9.37 at episode 3110

running reward: 9.37 at episode 3120

running reward: 9.42 at episode 3130

running reward: 9.98 at episode 3140

running reward: 9.76 at episode 3150

running reward: 10.00 at episode 3160

running reward: 9.78 at episode 3170

running reward: 9.53 at episode 3180

running reward: 9.38 at episode 3190

running reward: 9.53 at episode 3200

running reward: 9.47 at episode 3210

running reward: 9.47 at episode 3220

running reward: 9.45 at episode 3230

running reward: 9.68 at episode 3240

running reward: 9.64 at episode 3250

running reward: 9.86 at episode 3260

running reward: 9.88 at episode 3270

running reward: 9.92 at episode 3280

running reward: 9.65 at episode 3290

running reward: 9.76 at episode 3300

running reward: 9.71 at episode 3310

running reward: 9.71 at episode 3320

running reward: 9.58 at episode 3330

running reward: 9.78 at episode 3340

running reward: 9.77 at episode 3350

running reward: 9.58 at episode 3360

running reward: 9.57 at episode 3370

running reward: 9.37 at episode 3380

running reward: 9.60 at episode 3390

running reward: 9.58 at episode 3400

running reward: 10.00 at episode 3410

running reward: 10.32 at episode 3420

running reward: 10.41 at episode 3430

running reward: 13.02 at episode 3440

running reward: 12.72 at episode 3450

running reward: 11.75 at episode 3460

running reward: 10.94 at episode 3470

running reward: 10.35 at episode 3480

running reward: 10.13 at episode 3490

running reward: 10.30 at episode 3500

running reward: 9.88 at episode 3510

running reward: 9.76 at episode 3520

running reward: 9.70 at episode 3530

running reward: 9.77 at episode 3540

running reward: 9.77 at episode 3550

running reward: 9.81 at episode 3560

running reward: 9.73 at episode 3570

running reward: 9.58 at episode 3580

running reward: 9.46 at episode 3590

running reward: 9.50 at episode 3600

running reward: 9.41 at episode 3610

running reward: 9.42 at episode 3620

running reward: 9.41 at episode 3630

running reward: 9.48 at episode 3640

running reward: 9.61 at episode 3650

running reward: 9.54 at episode 3660

running reward: 9.63 at episode 3670

running reward: 9.73 at episode 3680

running reward: 9.42 at episode 3690

running reward: 9.60 at episode 3700

running reward: 9.49 at episode 3710

running reward: 9.42 at episode 3720

running reward: 9.29 at episode 3730

running reward: 9.36 at episode 3740

running reward: 9.42 at episode 3750

running reward: 9.56 at episode 3760

running reward: 9.63 at episode 3770

running reward: 9.68 at episode 3780

running reward: 9.52 at episode 3790

running reward: 9.58 at episode 3800

running reward: 9.64 at episode 3810

running reward: 9.83 at episode 3820

running reward: 9.88 at episode 3830

running reward: 10.03 at episode 3840

running reward: 9.68 at episode 3850

running reward: 9.48 at episode 3860

running reward: 9.96 at episode 3870

running reward: 10.09 at episode 3880

running reward: 10.08 at episode 3890

running reward: 10.19 at episode 3900

running reward: 10.86 at episode 3910

Niestety ale przy gammie 0.1 rozwiązanie jest nie możliwe.