

# Sieci Samouczące Się - Lab 1

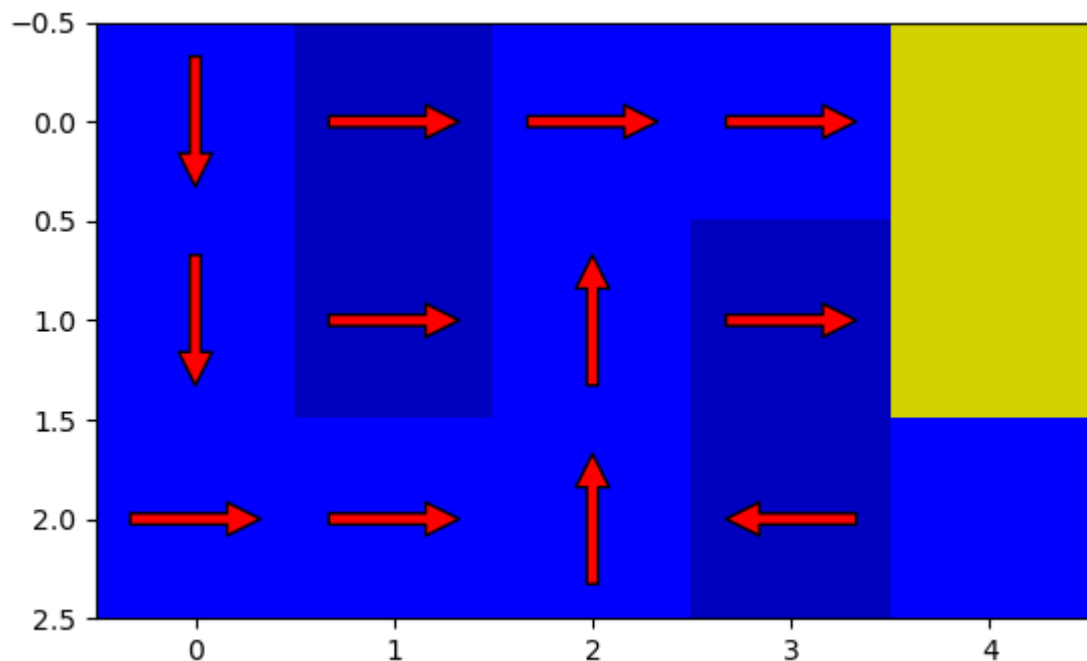
Jakub Sachajko 179976

Dla parametrów:

- number\_of\_episodes = 200
- number\_of\_iterations = 100
- gamma = 0.87

przeprowadziłem 8 testów. 4 dla każdej metody oraz 2 dla każdej mapy. Wyniki prezentują się następująco:

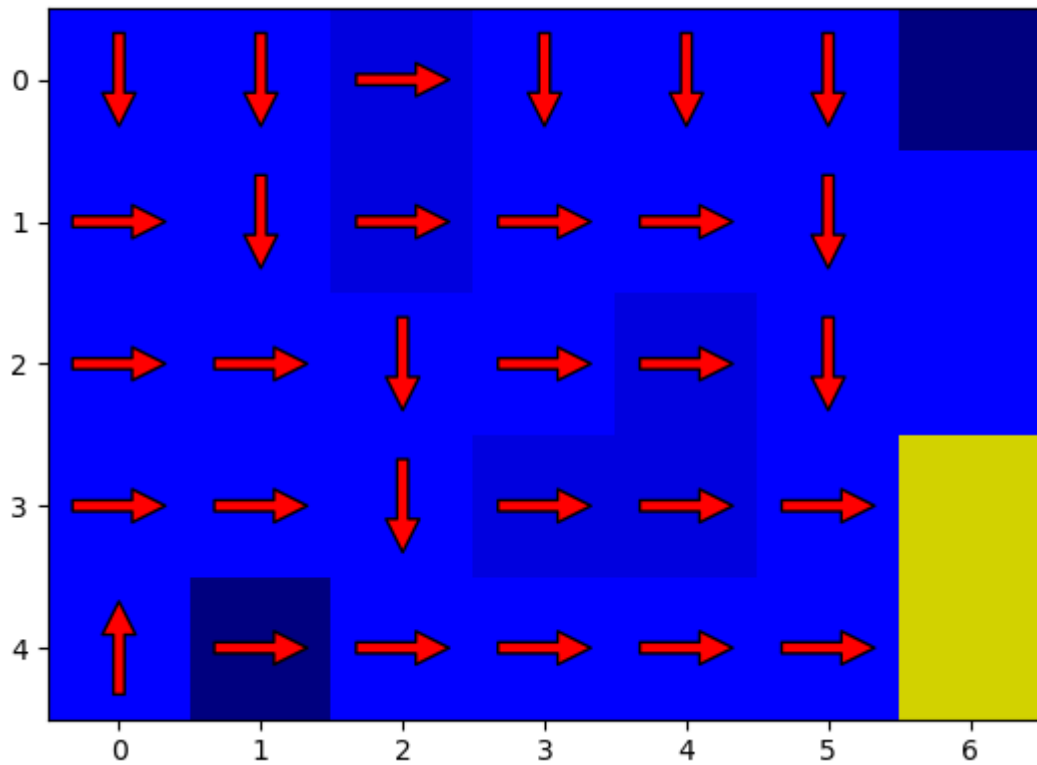
1. Mapa mała, metoda strategii:



test-1000 mean sum of rewards = 4.691

test-1000 mean sum of rewards = 4.775

2. mapa średnia, metoda strategii

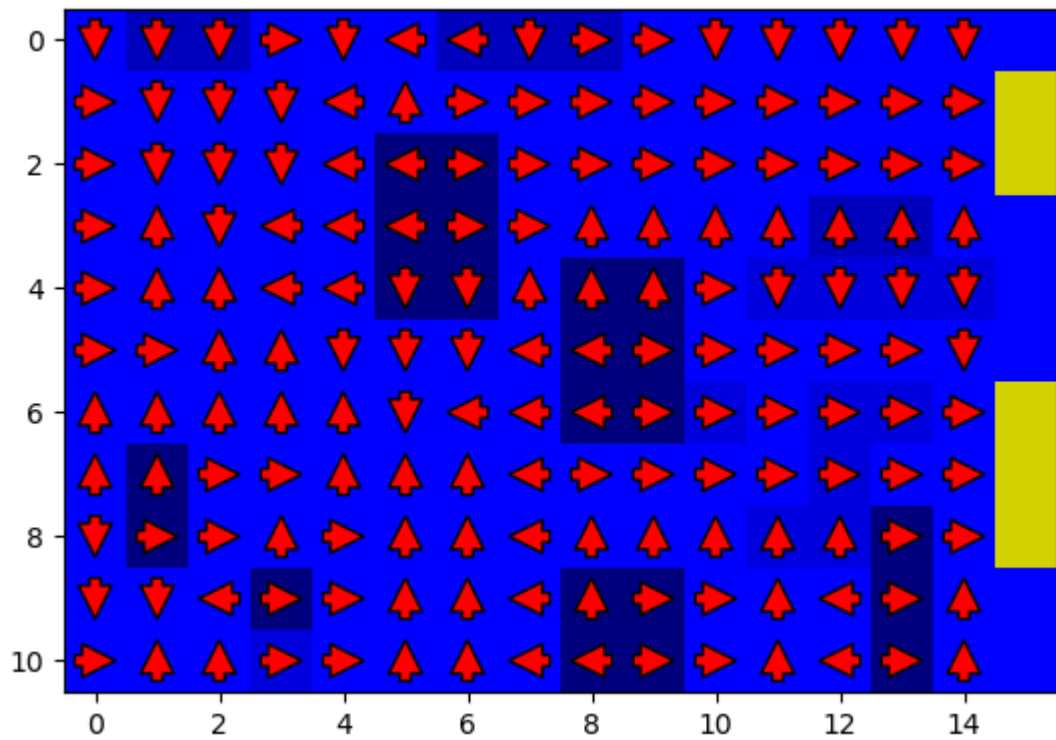


test-1000 mean sum of rewards = 6.992

test-1000 mean sum of rewards = 7.047

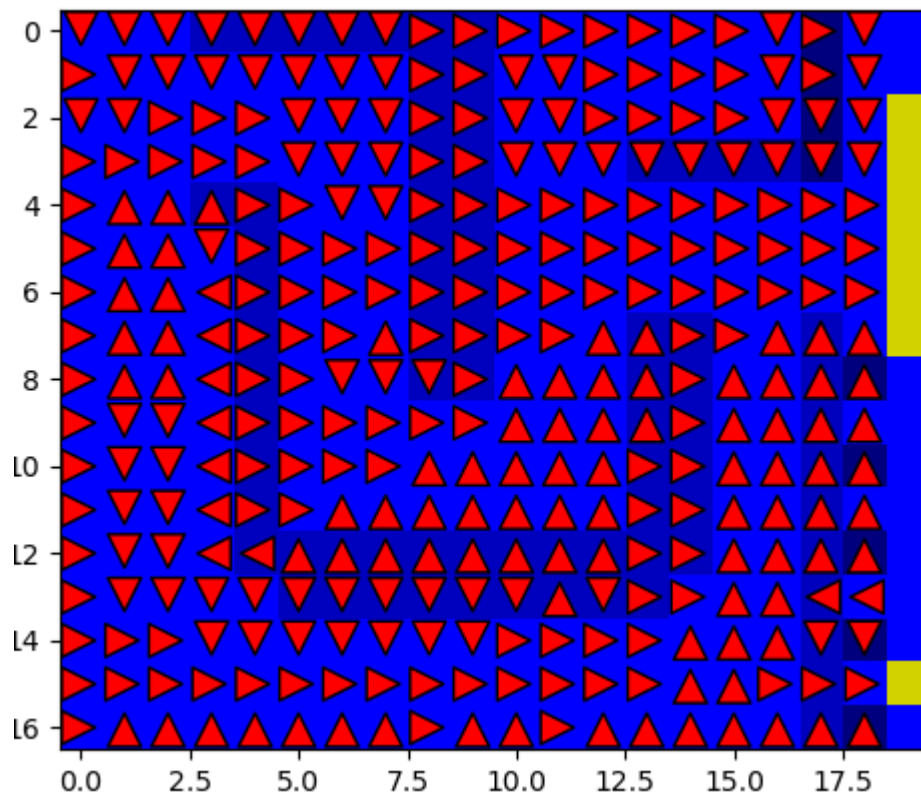
3. mapa duża, metoda strategii

przez zbyt długi czas obliczeniowy zmieniłem ilość iteracji na 50 i ilość epizodów na 80



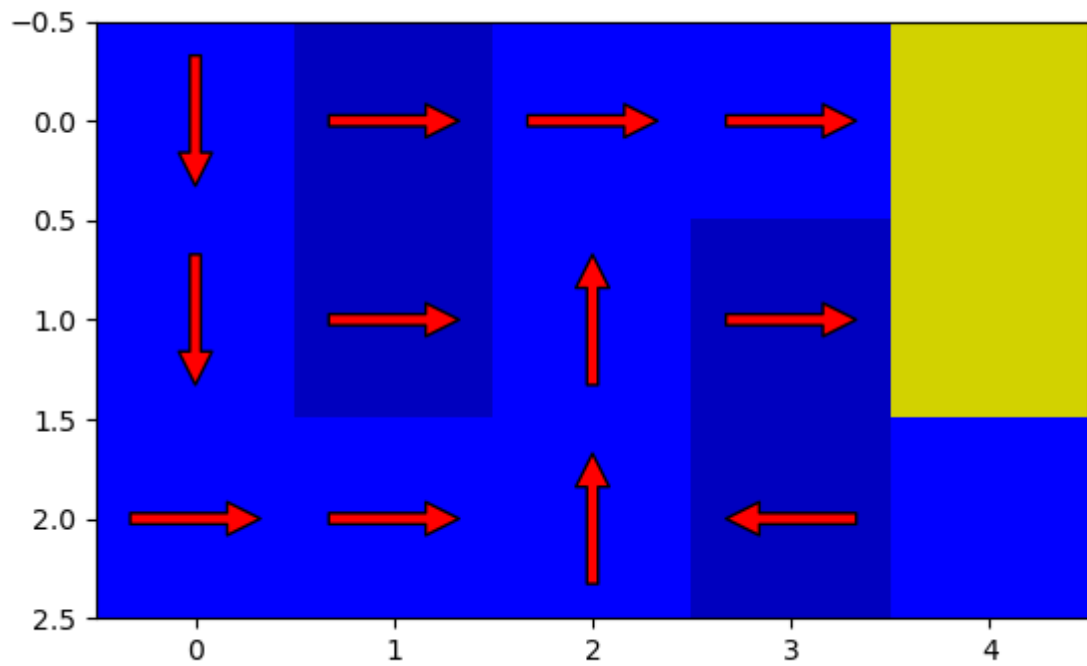
test-1000 mean sum of rewards = -3.581

4. mapa spiralna, metoda strategii  
przez zbyt długi czas obliczeniowy zmieniłem ilość iteracji na 50 i ilość epizodów na 80



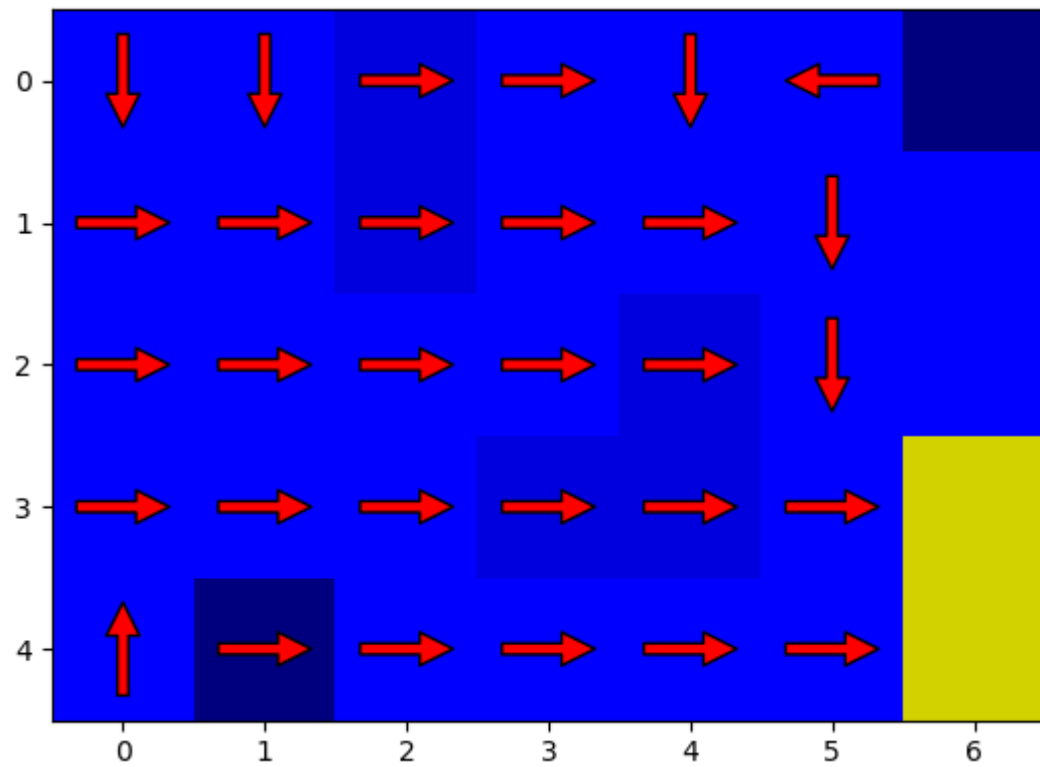
test-1000 mean sum of rewards = 181.99

5. mapa mała, metoda wartości



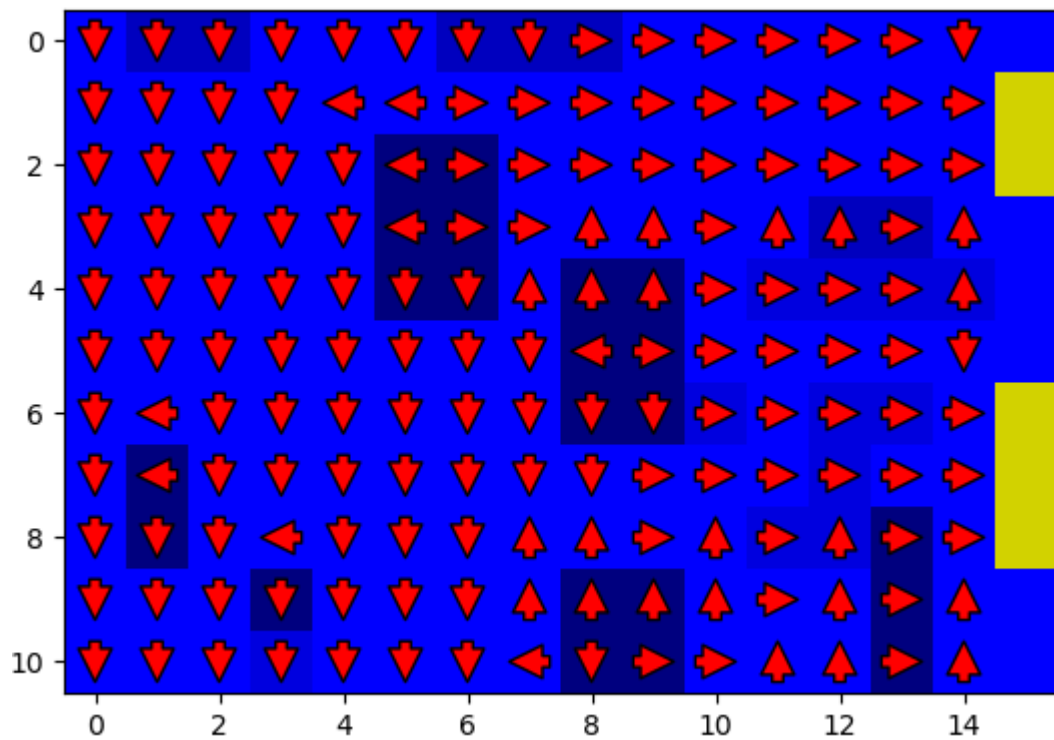
test-1000 mean sum of rewards = 4.702

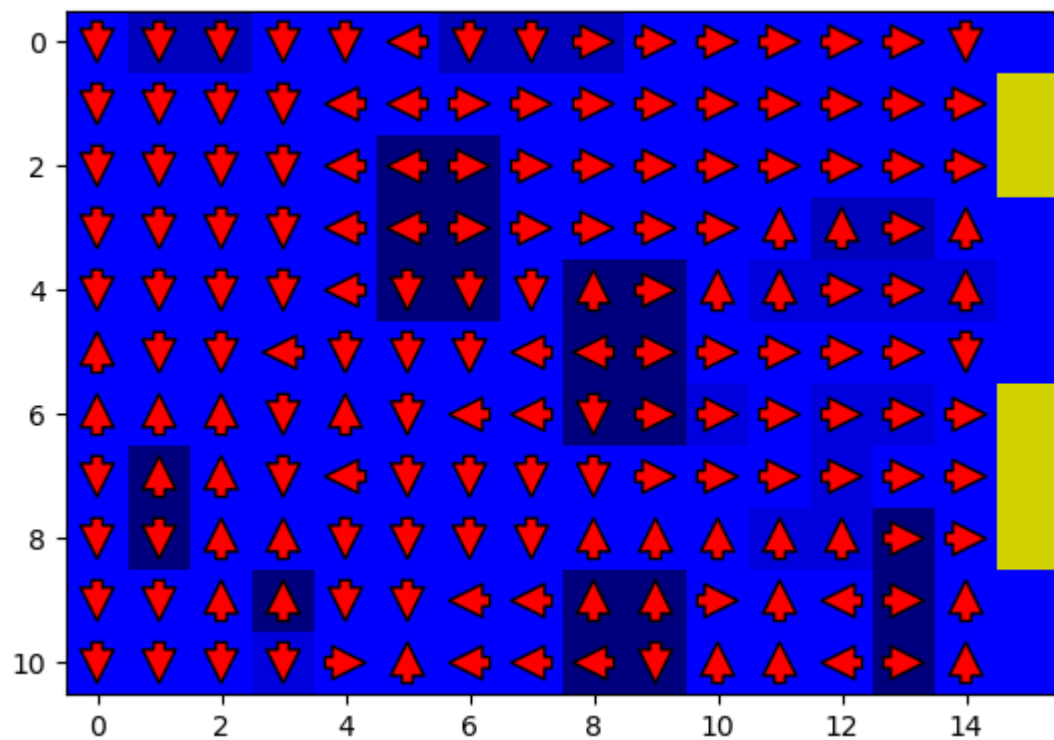
6. mapa średnia, metoda wartości



test-1000 mean sum of rewards = 6.104

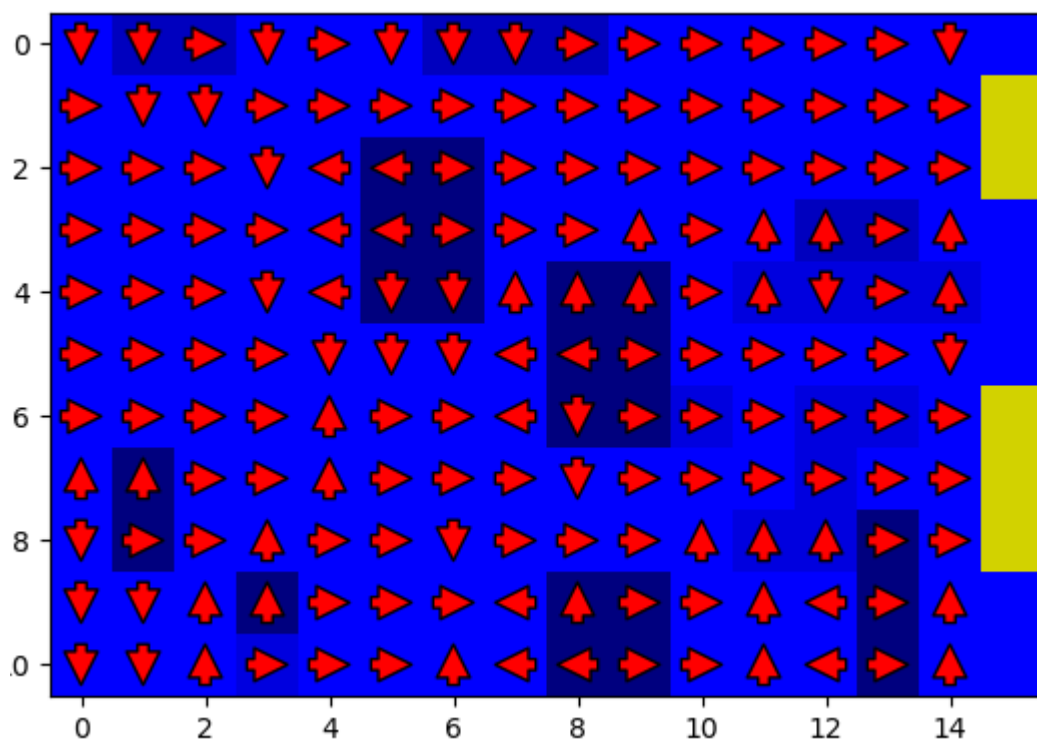
7. mapa duża, metoda wartości





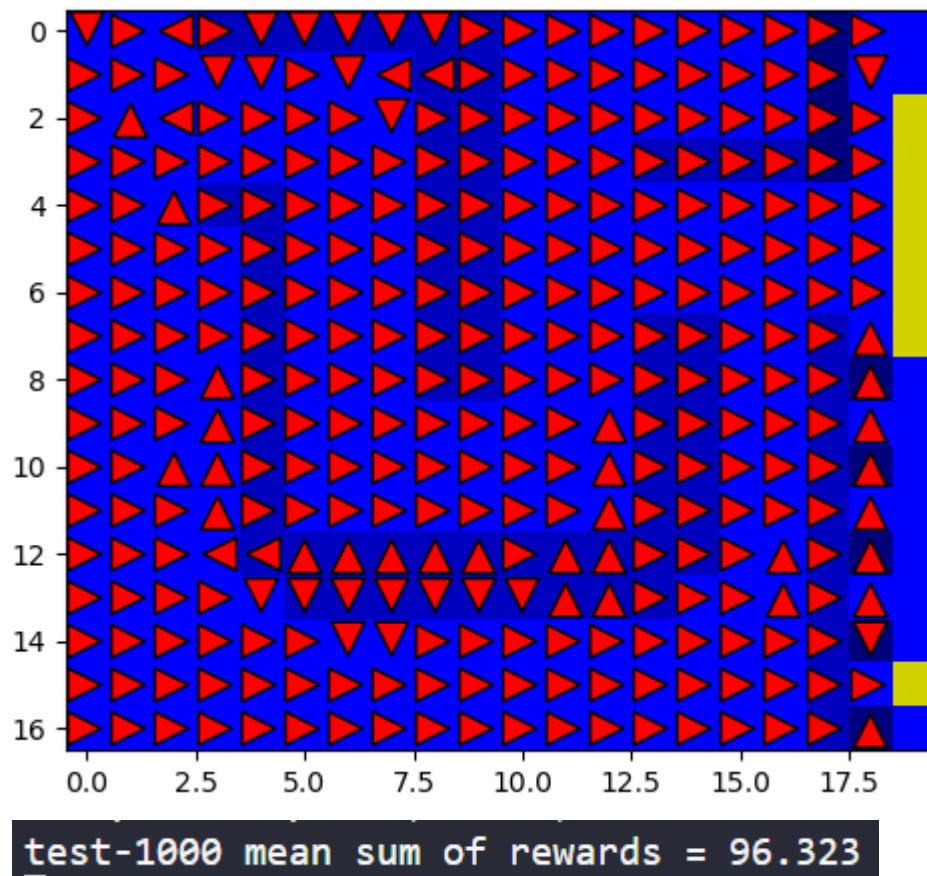
test-1000 mean sum of rewards = -207.711

test-1000 mean sum of rewards = -57.951



test-1000 mean sum of rewards = -17.051

## 8. mapa spiralna, metoda wartości



### Czego się nauczyłem:

- Metoda wartości czasem nie znajduje poprawnych wyników (jest niestabilna)
- Metoda wartości jest znacząco krótsza względem metody strategii
- Metoda wartości czasem potrzebuje wielu startów i daje bezsensowne wyniki jak dla przykładu nr 7. Dodatkowo pewnie przy użyciu większej ilości epok oraz epizodów dałbym radę osiągnąć wyniki drugiego algorytmu.