

Sieci Samouczące Się - Lab 2

Jakub Sachajko 179976

przeprowadziłem 4 testy, dla metody dynamicznej iteracji wartości. Wyniki prezentują się następująco:

parametr ogólny:

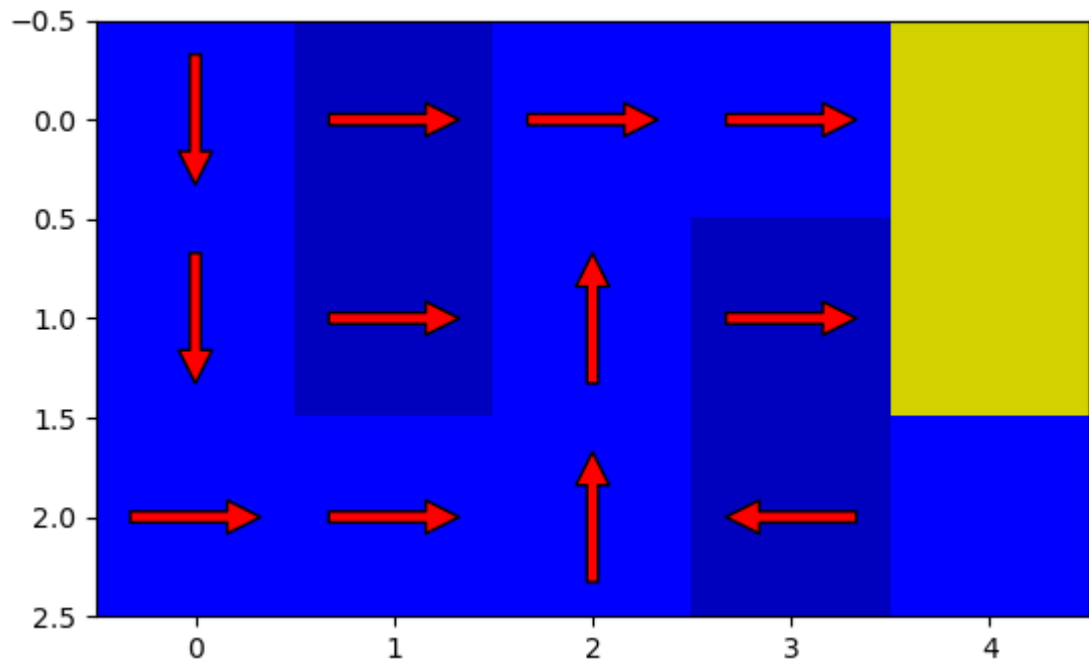
- $\gamma = 0.9$

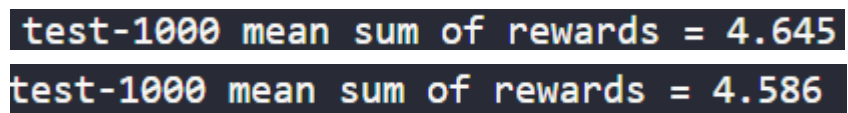
1. Mapa mała:

Dla parametrów:

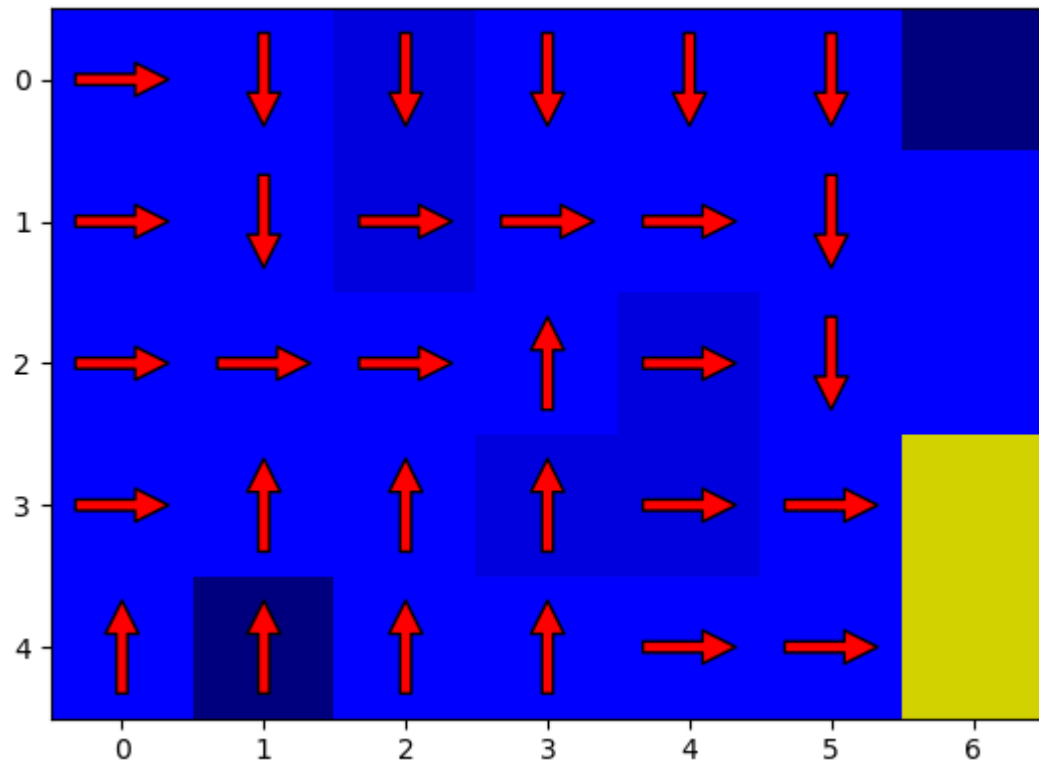
$\delta_{\max} = 4$

punkty ujemne za uderzenie w ścianę = -10





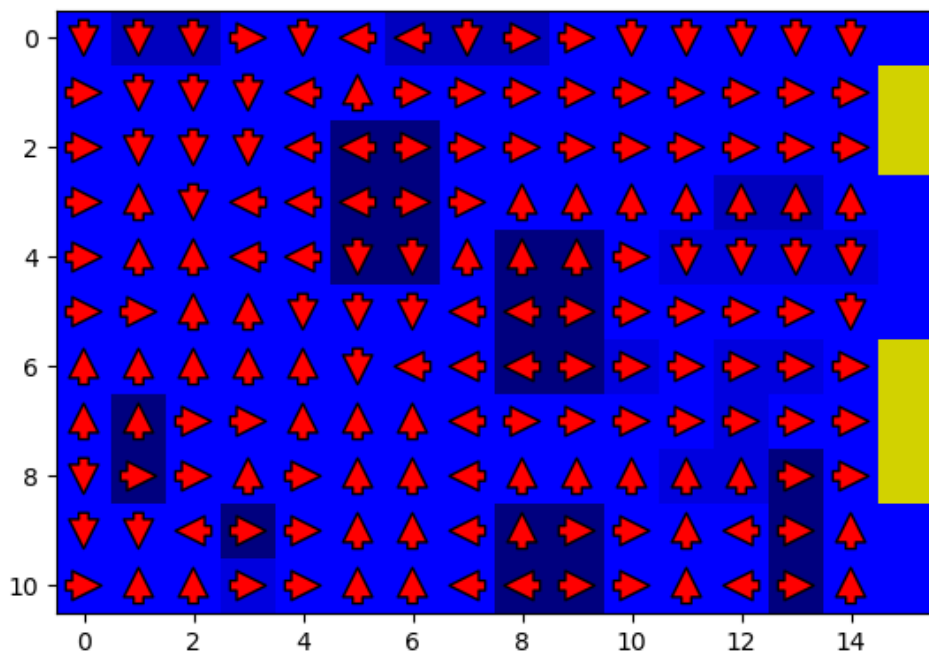
-
- The grid world environment is a 5x6 grid. The start state is at (0,4) and the goal state is at (6,3). Obstacles are located at (1,0), (1,1), (1,4), (2,0), (2,1), (3,3), (4,3), (5,0), (5,1), (5,2), (5,3), (5,4), (6,0), (6,1), (6,2), (6,3), and (6,4). Red arrows indicate the direction of movement from each state.

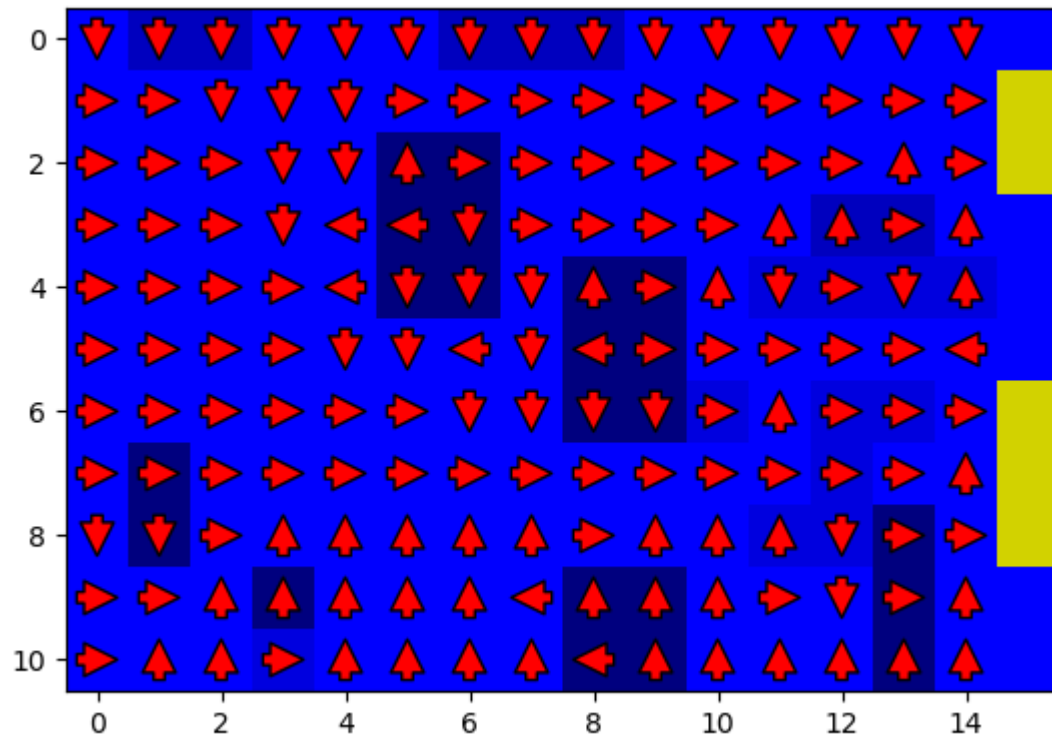


test-1000 mean sum of rewards = 6.229

test-1000 mean sum of rewards = 6.485

3. mapa duża
delta max = 4
punkty ujemne za uderzenie w ścianę = -5

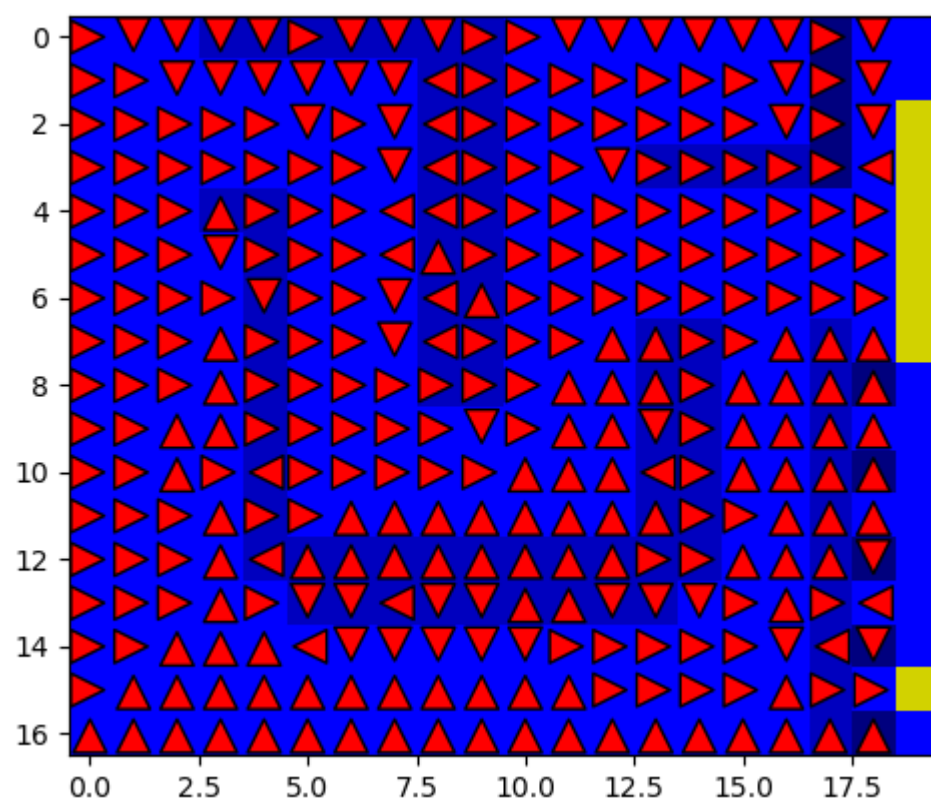
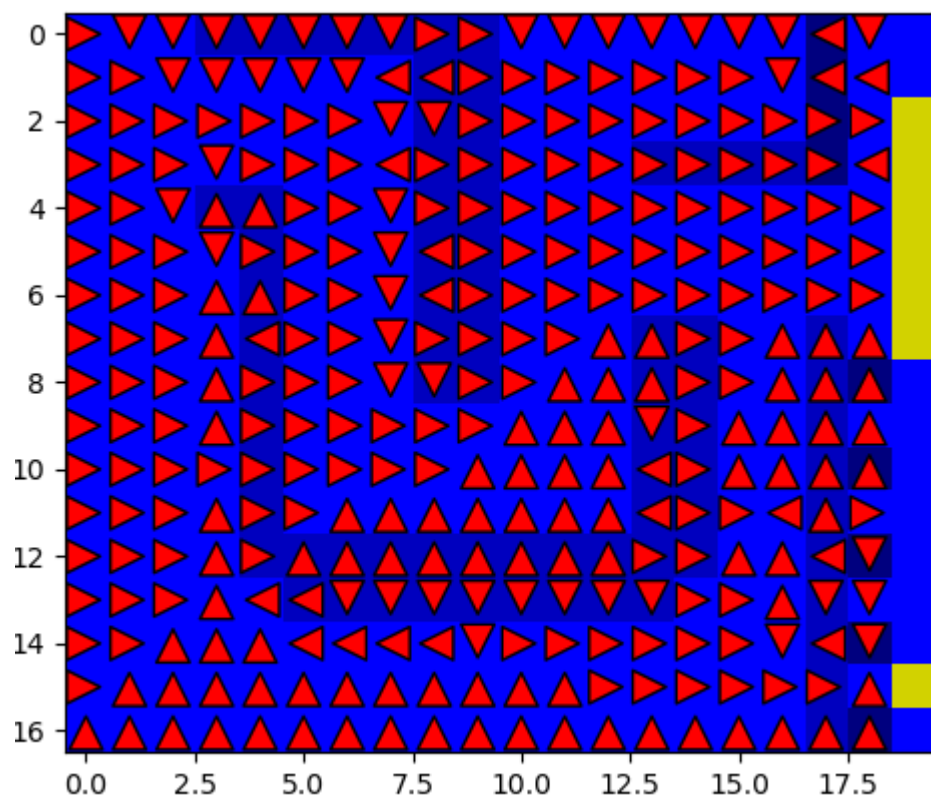


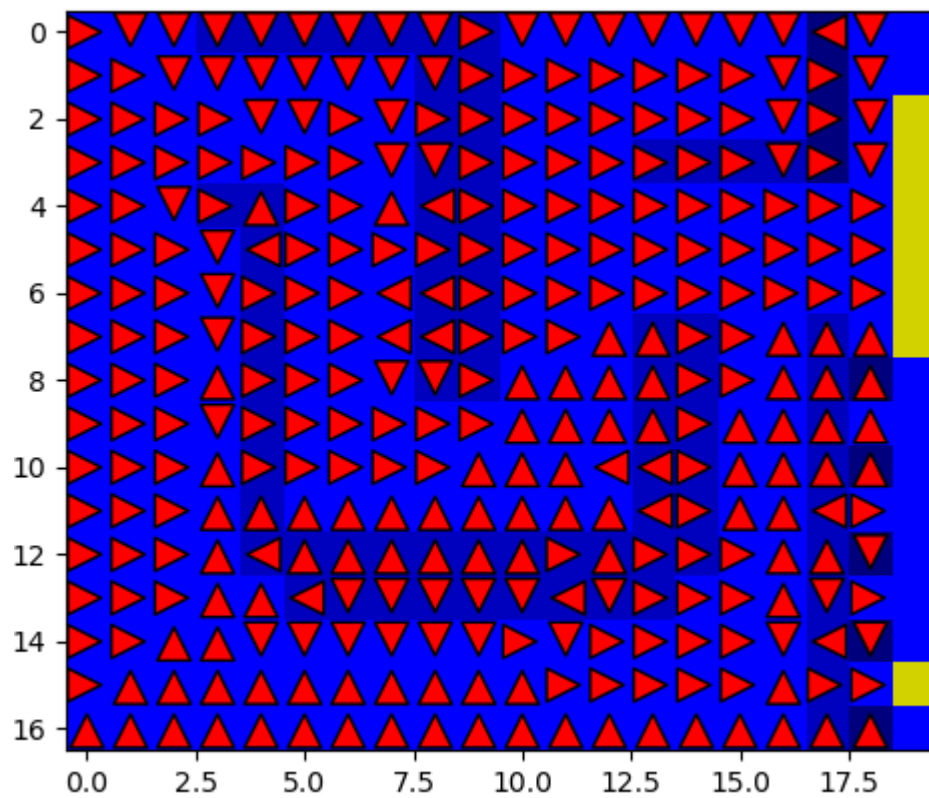


test-1000 mean sum of rewards = -3.581

test-1000 mean sum of rewards = -1.942

4. mapa spiralna
delta max = 20
punkty ujemne za uderzenie w ścianę = -10





test-1000 mean sum of rewards = 171.635

test-1000 mean sum of rewards = 166.66

test-1000 mean sum of rewards = 174.629