

# Zadanie 6 - Raport

Jan Stusio

Czerwiec 2024

## 1 Wstęp

Celem zadania jest implementacja algorytmu Q-Learning oraz analizy wpływu parametrów  $\alpha$  (współczynnik uczenia),  $\gamma$  (współczynnik dyskontowania) i  $\epsilon$  (eksploracja w polityce  $\epsilon$ -zachłannej) oraz parametr T na rozkładzie Boltzmana na zbieżność algorytmu w środowisku FrozenLake-v1 z biblioteki gym.

## 2 Implementacja

### 2.1 Algorytm Q-Learning

Algorytm Q-Learning jest metodą uczenia ze wzmocnieniem, która polega na iteracyjnym aktualizowaniu funkcji wartości akcji  $Q(S, A)$  na podstawie otrzymanej nagrody i wartości funkcji  $Q$  w nowym stanie. Wzór aktualizacji funkcji wartości akcji przedstawia się następująco:

$$Q^{new}(S_t, A_t) \leftarrow (1 - \alpha) \cdot Q(S_t, A_t) + \alpha(R_{t+1} + \gamma \cdot \max_a Q(S_{t+1}, a))$$

Ja w implementacji zmodyfikowałem ten wzór.

## 3 Wyniki

Pomiary na wynikach są dla siatki 4x4, ponieważ nie udało mi się dobrać parametrów, które by dobrze obrazowały zbieżność algorytmu dla siatki 8x8.

### 3.1 Wpływ parametru $\alpha$

Figure 1

### 3.2 Wpływ parametru $\gamma$

Figure 2

### 3.3 Wpływ parametru $\epsilon$

Figure 3

### 3.4 Wpływ parametru $T$

Figure 4

### 3.5 Tabela wyników

Table 1

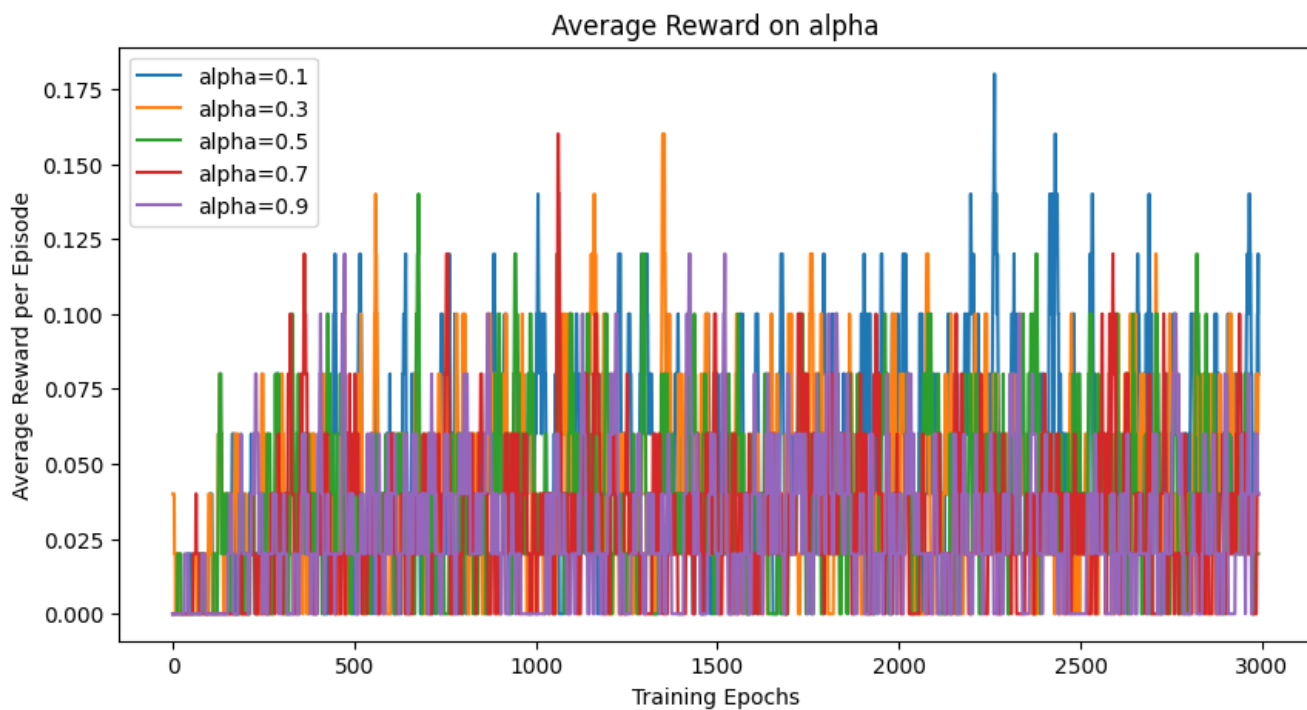


Figure 1: Wpływ parametru  $\alpha$  na zbieżność algorytmu Q-Learning

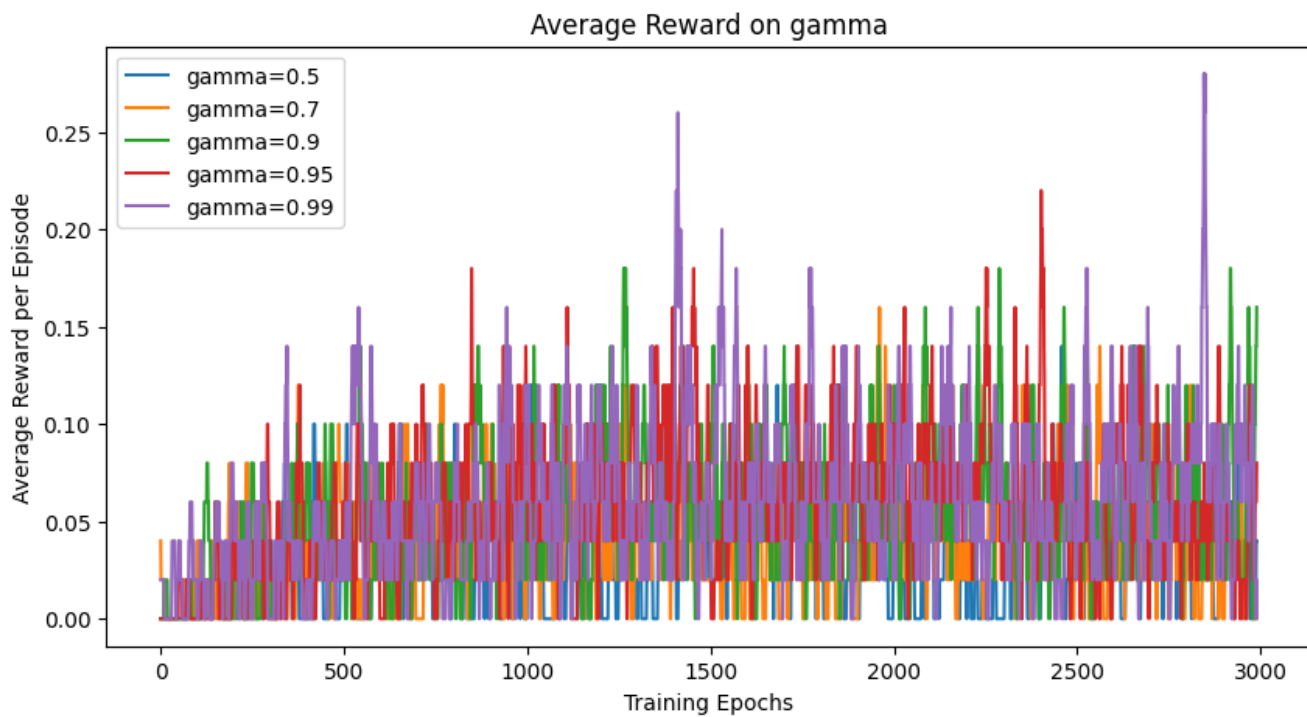


Figure 2: Wpływ parametru  $\gamma$  na zbieżność algorytmu Q-Learning

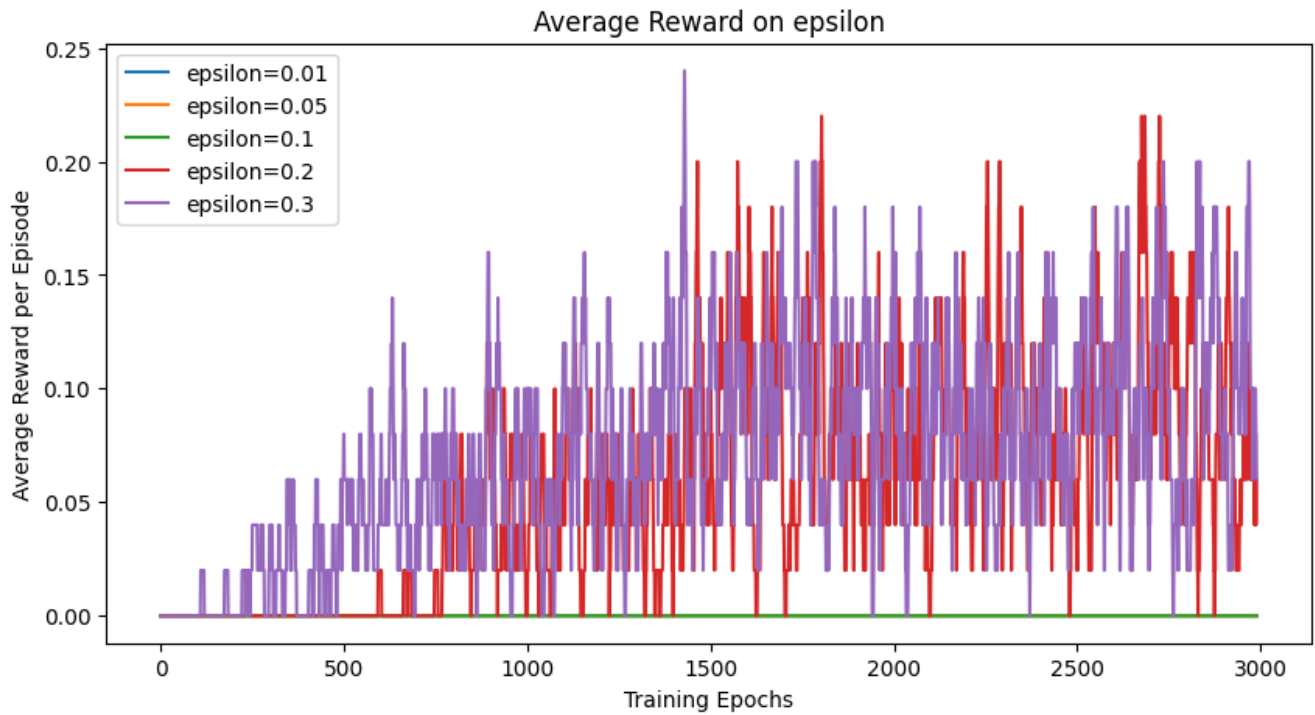


Figure 3: Wpływ parametru  $\epsilon$  na zbieżność algorytmu Q-Learning

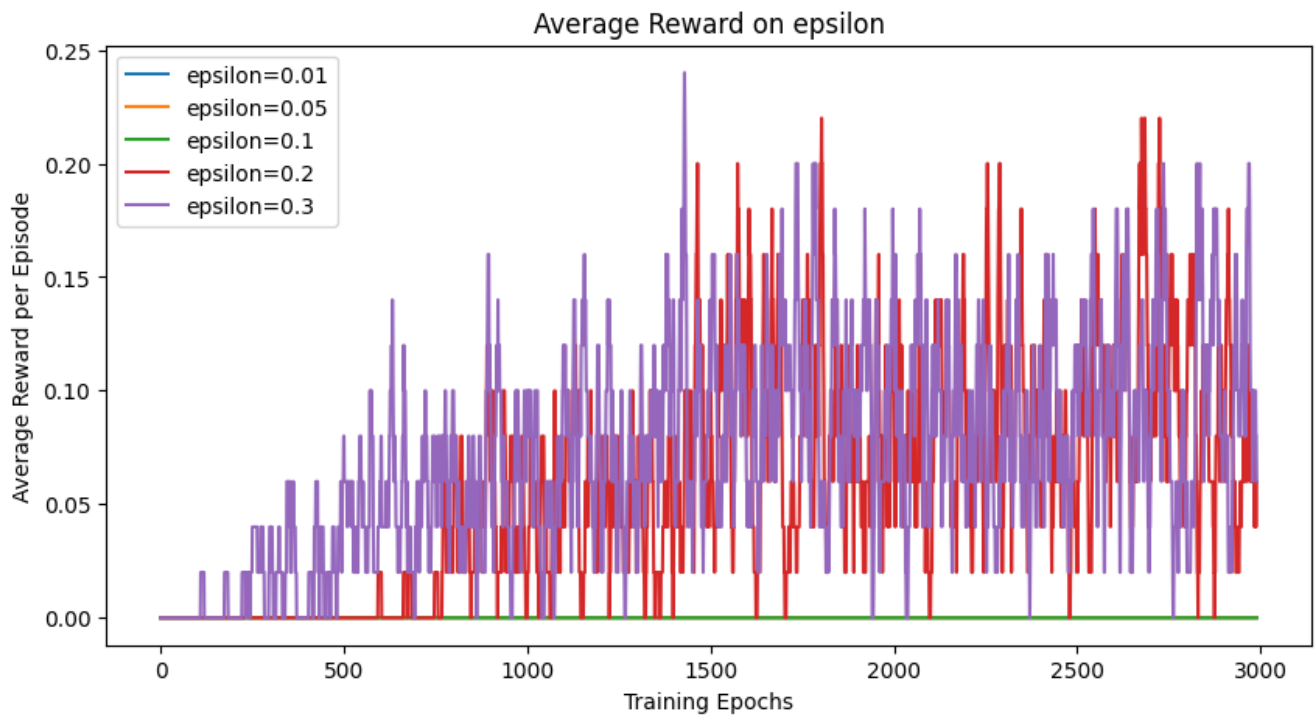


Figure 4: Wpływ parametru  $T$  na zbieżność algorytmu Q-Learning

Table 1: Średnie nagrody w ostatnich 10 epizodach dla różnych wartości badanych parametrów

Parameter	Value	Mean Reward	Std Reward	Success Count
alpha	0.100000	0.080000	0.132665	3
alpha	0.300000	0.080000	0.132665	3
alpha	0.500000	0.020000	0.060000	1
alpha	0.700000	0.040000	0.080000	2
alpha	0.900000	0.040000	0.080000	2
gamma	0.500000	0.040000	0.080000	2
gamma	0.700000	0.000000	0.000000	0
gamma	0.900000	0.160000	0.174356	6
gamma	0.950000	0.080000	0.097980	4
gamma	0.990000	0.020000	0.060000	1
epsilon	0.010000	0.000000	0.000000	0
epsilon	0.050000	0.000000	0.000000	0
epsilon	0.100000	0.000000	0.000000	0
epsilon	0.200000	0.060000	0.091652	3
epsilon	0.300000	0.060000	0.091652	3
T	0.500000	0.000000	0.000000	0
T	1.000000	0.000000	0.000000	0
T	2.000000	0.000000	0.000000	0
T	5.000000	0.040000	0.120000	1
T	10.000000	0.040000	0.080000	2

## 4 Wnioski

Na podstawie przeprowadzonych eksperymentów można zauważyć, że wszystkie trzy parametry ( $\alpha$ ,  $\gamma$  i  $\epsilon$ ) mają istotny wpływ na zbieżność algorytmu Q-Learning.

Wyniki potwierdzają, że odpowiedni dobór parametrów jest kluczowy dla efektywnego działania algorytmu Q-Learning.