



Politechnika
Wrocławska

Solving the optimal path planning of a mobile robot using improved Q-learning - omówienie artykułu

Przemysław Jaskuła 269995@student.pwr.edu.pl

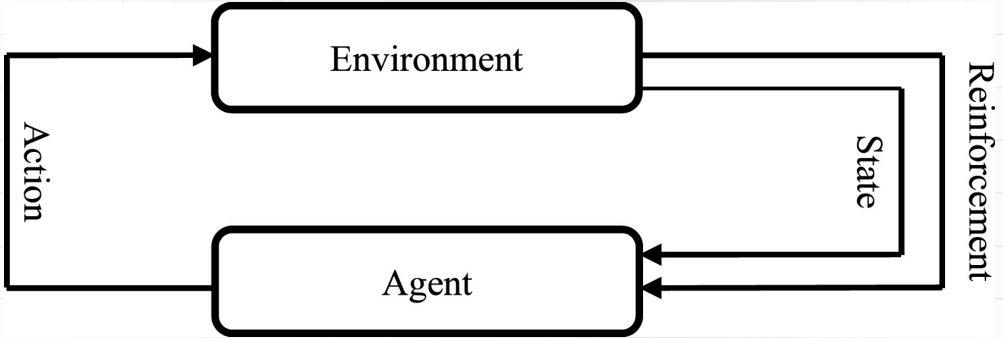
WYDZIAŁ ELEKTRONIKI, FOTONIKI I MIKROSYSTEMÓW

Projekt Specjalnościowy

- 1 Reinforcement learning
- 2 Q learning
- 3 Problemy związane z Q learningiem
- 4 Co to jest FPA
- 5 FPA jako lepsze zainicjowanie macierzy Q
- 6 Q & A

Section 1

Reinforcement learning



Section 2

Q learning

► Stan

- ▶ Stan
- ▶ Akcja

- ▶ Stan
- ▶ Akcja
- ▶ Agent

- ▶ Stan
- ▶ Akcja
- ▶ Agent
- ▶ Nagroda i Kara

$$Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s, a) + \alpha \left[r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) \right]$$

s_t : Obecny stan

a_t : Akcja podjęta w stanie s_t

r_{t+1} : Nagroda po stanie s_t

α : Współczynnik nauki

γ : Współczynnik dyskontowy

Algorithm 1: Classical Q-learning algorithm

Begin

Initiate all Q-values, $Q(s, a)$ in Q-table to zero

Select a starting state, $Q(s_1, a_1)$

while (iteration < Max iteration)

while goal is not achieved

 Select an action, a within the available actions in the current state according to the highest Q-value in the next state

 Perform the selected action, a and reward or penalty, r will be given

 Update the Q-value using Equation (1)

 Move the state to new state, s'

end while

end while

end

Section 3

Problemy związane z Q learningiem

1. Złożoność obliczeniowa wraz z wzrostem przestrzeni poszukiwań

1. Złożoność obliczeniowa wraz z wzrostem przestrzeni poszukiwań
2. Początkowy etap eksploracji

Section 4

Co to jest FPA

- ▶ Zapylenie globalne zawiera zapylenie biotyczne i krzyżowe, a zapylacze używają dystrybucji Levy-iego przy przenoszeniu pyłków.

- ▶ Zapylenie globalne zawiera zapylenie biotyczne i krzyżowe, a zapylacze używają dystrybucji Levy-iego przy przenoszeniu pyłków.
- ▶ Zapylenie lokalne będące abiotyczne i samozapylające.

- ▶ Zapylenie globalne zawiera zapylenie biotyczne i krzyżowe, a zapylacze używają dystrybucji Levy-iego przy przenoszeniu pyłków.
- ▶ Zapylenie lokalne będące abiotyczne i samozapylające.
- ▶ Szansa na reprodukcję zależy od tego jak dwa kwiaty są podobne.

- ▶ Zapylenie globalne zawiera zapylenie biotyczne i krzyżowe, a zapylacze używają dystrybucji Levy-iego przy przenoszeniu pyłków.
- ▶ Zapylenie lokalne będące abiotyczne i samozapylające.
- ▶ Szansa na reprodukcję zależy od tego jak dwa kwiaty są podobne.
- ▶ Prawdopodobieństwo zmiany decyduje między zapyleniem lokalnym a globalnym.

$$x_i^{t+1} = x_i^t + \gamma L(\lambda)(x_i^t - g^*)$$

x_i^t : Pyłek i w iteracji t

g^* : Najlepsze rozwiązanie w obecnej iteracji

γ : Współczynnik skalowania

$L(\lambda)$: Siła zapylania

$$x_i^{t+1} = x_i^t + \epsilon(x_j^t - x_k^t)$$

x_i^t : Pyłek i w iteracji t

$x_j^t - x_k^t$: Pyłki z tego samego gatunku, lecz innych kwiatów

ϵ : Liczba losowa z dystrybucji jednostajnej

Algorithm 2: Flower pollination algorithm

Begin

Define objective function $f(x), x = (x_1, x_2, x_3, \dots, x_d)^T$

Do initialization of a population of n flowers in random positions.

Obtain the best solution g^* in the initial population

Define the range of switch probability $p \in [0, 1]$

while ($t < \text{Max Generation}$)

for $i = 1:n$ (all n flowers in the population)

if ($\text{rand} > p$)

 Draw a (d -dimensional) state vector L from Levy distribution

 Obtain global population using the relation

$$x_i^{t+1} = x_i^t + \gamma L(\lambda)(x_i^t - g^*)$$

else

Draw ϵ from a uniform distribution in $[0,1]$

Obtain local pollination using $x_i^{t+1} = x_i^t + \epsilon (x_j^t - x_k^t)$

end if

Evaluate new solutions

If new solutions are better than update population

end for

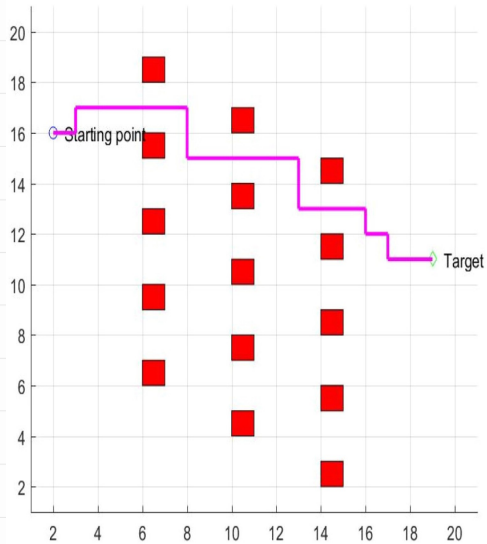
Find the current best solution g^*

end while

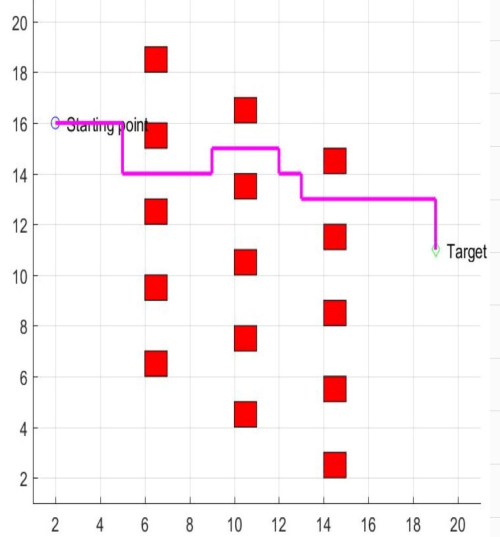
end

Section 5

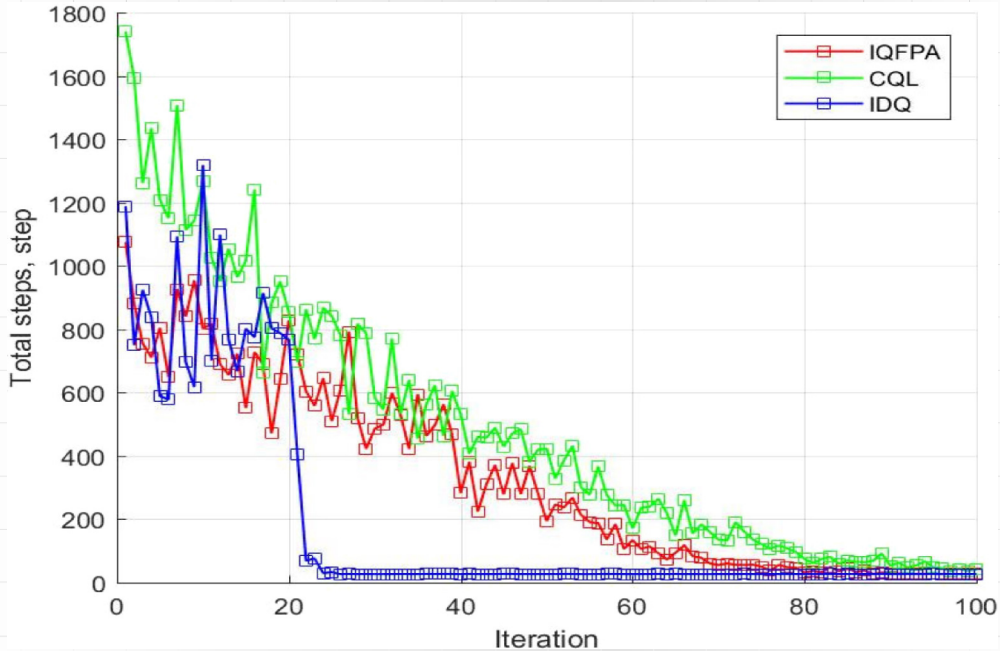
FPA jako lepsze zainicjowanie macierzy Q

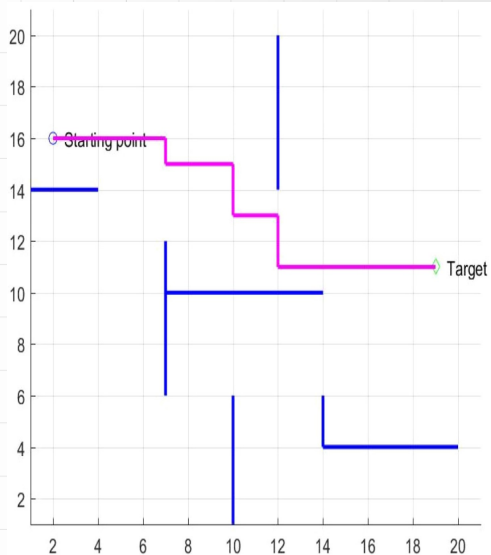


(a)

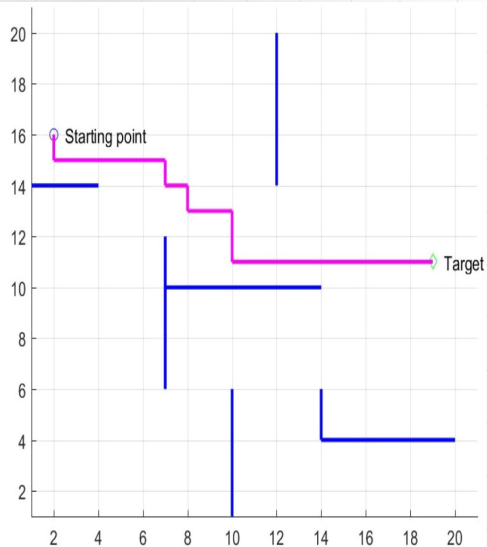


(b)

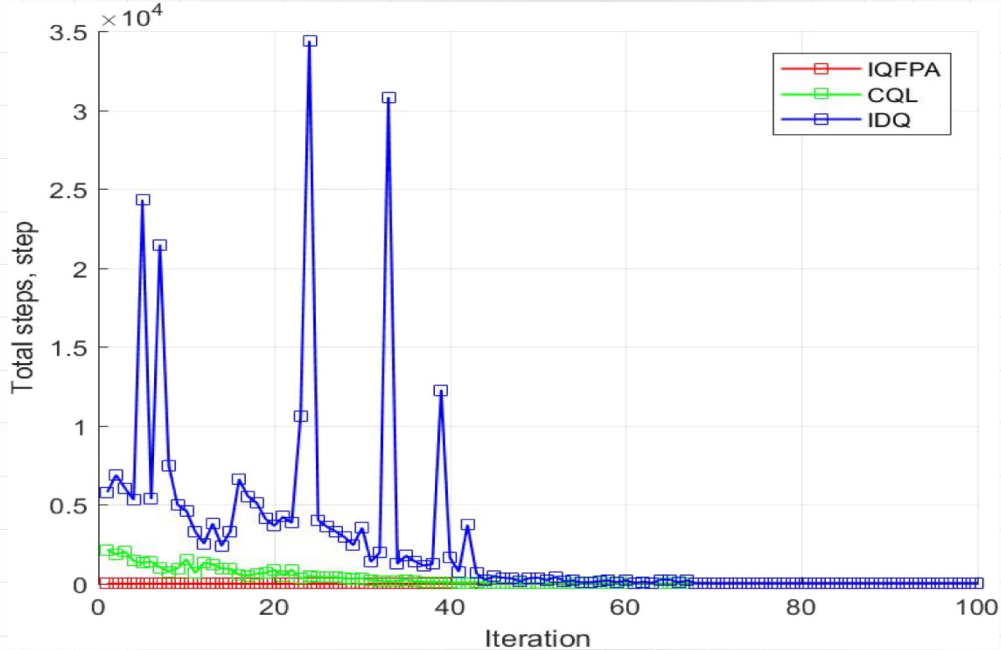




(a)



(b)



Section 6

Q & A

Dziękuję za uwagę