

# エージェント工学 最終レポート課題

澤 祐里 (21-1-037-0801)

2021 年 7 月 14 日

目次

1	目的	1
1.1	課題 1 . . . . .	1
1.2	課題 2 . . . . .	1
1.3	課題 3 . . . . .	1
2	手順	2
2.1	課題 1 . . . . .	2
2.2	課題 2 . . . . .	2
2.2.1	計測対象 . . . . .	2

## 1 目的

リサイクルロボットについての強化学習サンプルプログラムを用いて、パラメータを変更することなどで強化学習について考察する。

### 1.1 課題 1

リサイクルロボットの MDP を 3 種類作成する。

### 1.2 課題 2

ε Greedy 手法と Softmax 手法の概要を説明する。

### 1.3 課題 3

ε Greedy 手法と softmax 手法の比較を行う。

## 2 手順

### 2.1 課題 1

recycleRobotMDP.csv の case1 を参考にして遷移確率や報酬を変更した case2, case3 を作成する。作成した報酬の与え方がエージェントの目標になっていることを示す。

表 1 元の MDP(case1)

状態	行動	次状態	遷移確率	報酬
START	N	high	1	0
high	search	high	0.8	12
high	search	low	0.2	4
low	search	GOAL	0.1	-3
low	search	low	0.9	2
high	wait	high	1	1
high	wait	low	0	1
low	wait	GOAL	0	1
low	wait	low	1	1
low	recharge	GOAL	1	0
low	recharge	low	0	0
GOAL	N	high	1	0

表 2 作成した MDP(case2)

状態	行動	次状態	遷移確率	報酬
START	N	high	1	0
high	search	high	0.8	10
high	search	low	0.2	4
low	search	GOAL	0.1	-3
low	search	low	0.9	-5
high	wait	high	1	1
high	wait	low	0	-2
low	wait	GOAL	0	0
low	wait	low	1	1
low	recharge	GOAL	1	4
low	recharge	low	0	0
GOAL	N	high	1	0

表 3 作成した MDP(case3)

状態	行動	次状態	遷移確率	報酬
START	N	high	1	0
high	search	high	0.8	10
high	search	low	0.2	4
low	search	GOAL	0.1	-7
low	search	low	0.9	-2
high	wait	high	0.9	2
high	wait	low	0.1	-3
low	wait	GOAL	0.1	-7
low	wait	low	0.9	2
low	recharge	GOAL	0.8	7
low	recharge	low	0.2	-2
GOAL	N	high	1	0

### 2.2 課題 2

#### 2.2.1 計測対象

## 参考文献

- [1] [https://www.codereading.com/algo\\_and\\_ds/algo/quick\\_sort.html](https://www.codereading.com/algo_and_ds/algo/quick_sort.html)
- [2] <http://www.ics.kagoshima-u.ac.jp/~fuchida/edu/algorithm/sort-algorithm/quick-sort.html>