projekt06

June 12, 2025

1 Problem wieloagentowy

1.0.1 Autorzy:

Diana Misiaczyńska, Emilia Myrta

Tematyką szóstego, i tym samym ostatniego, projektu są problemy wieloagentowe. Środowisko, które wybrałyśmy to `Simple Reference` z `PettingZoo`. Jest ono częścią nastawionego na komunikację agentów zestawu `Multi Particle Environments (MPE)`.

Wybranym przez nas algorytmem rozwiązującym problem jest `DQN` z biblioteki `stable_baselines3`.

[28]:
```
from pettingzoo.mpe import simple_reference_v3
```

[29]:
```
env = simple_reference_v3.env(render_mode="human")
env.reset(seed=42)

for agent in env.agent_iter():
    observation, reward, termination, truncation, info = env.last()

    if termination or truncation:
        action = None
    else:
        action = env.action_space(agent).sample()

    env.step(action)
env.close()
```

Dodajemy wszystkie potrzebne importy.

[30]:
```
import supersuit as ss
from stable_baselines3 import DQN
from pettingzoo.mpe import simple_reference_v3
from pettingzoo.utils.conversions import aec_to_parallel
import matplotlib.pyplot as plt
import numpy as np
```

Inicjalizujemy środowisko PettingZoo i konwertujemy do ParallelEnv. Następnie korzystamy z wrapperów supersuit do wyrównania przestrzeni obserwacji i akcji. Potrzebna jest także konwersja środowiska do formatu kompatybilnego z Gym.

[31]:
```
env = simple_reference_v3.env()
parallel_env = aec_to_parallel(env)

parallel_env = ss.pad_observations_v0(parallel_env)
parallel_env = ss.pad_action_space_v0(parallel_env)

env = ss.pettingzoo_env_to_vec_env_v1(parallel_env)
env = ss.concat_vec_envs_v1(env, 4, num_cpus=1, base_class="stable_baselines3")
```

Dostosowujemy hiperparametry modelu DQN oraz ustawiamy liczbę timesteps na 10 000.

[32]:
```
model = DQN('MlpPolicy', env, verbose=1,
            learning_rate=1e-3,
            gamma=0.99,
            exploration_fraction=0.1,
            exploration_final_eps=0.01,
            target_update_interval=1000,
            train_freq=1,
            gradient_steps=1,
            batch_size=32,
            buffer_size=100000)

model.learn(total_timesteps=10000)
```

```
Using cpu device
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.802    |
| time/               |          |
|    episodes         | 4        |
|    fps              | 1668     |
|    time_elapsed     | 0        |
|    total_timesteps  | 200      |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.453    |
|    n_updates        | 12       |
----------------------------------
-----------------------------
| time/              |      |
|    episodes        | 8    |
|    fps             | 1655 |
|    time_elapsed    | 0    |
|    total_timesteps | 200  |
-----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.604    |
```

```
| time/               |          |
|    episodes         | 12       |
|    fps              | 1521     |
|    time_elapsed     | 0        |
|    total_timesteps  | 400      |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.373    |
|    n_updates        | 37       |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 16   |
|    fps             | 1516 |
|    time_elapsed    | 0    |
|    total_timesteps | 400  |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.406    |
| time/               |          |
|    episodes         | 20       |
|    fps              | 1491     |
|    time_elapsed     | 0        |
|    total_timesteps  | 600      |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.155    |
|    n_updates        | 62       |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 24   |
|    fps             | 1487 |
|    time_elapsed    | 0    |
|    total_timesteps | 600  |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.208    |
| time/               |          |
|    episodes         | 28       |
|    fps              | 1462     |
|    time_elapsed     | 0        |
|    total_timesteps  | 800      |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.15     |
```

```
|    n_updates        | 87       |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 32   |
|    fps             | 1460 |
|    time_elapsed    | 0    |
|    total_timesteps | 800  |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 36       |
|    fps              | 1452     |
|    time_elapsed     | 0        |
|    total_timesteps  | 1000     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.0844   |
|    n_updates        | 112      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 40   |
|    fps             | 1450 |
|    time_elapsed    | 0    |
|    total_timesteps | 1000 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 44       |
|    fps              | 1454     |
|    time_elapsed     | 0        |
|    total_timesteps  | 1200     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.093    |
|    n_updates        | 137      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 48   |
|    fps             | 1452 |
|    time_elapsed    | 0    |
|    total_timesteps | 1200 |
```

```
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 52      |
|    fps              | 1449    |
|    time_elapsed     | 0       |
|    total_timesteps  | 1400    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.15    |
|    n_updates        | 162     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 56   |
|    fps             | 1448 |
|    time_elapsed    | 0    |
|    total_timesteps | 1400 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 60      |
|    fps              | 1411    |
|    time_elapsed     | 1       |
|    total_timesteps  | 1600    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.107   |
|    n_updates        | 187     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 64   |
|    fps             | 1410 |
|    time_elapsed    | 1    |
|    total_timesteps | 1600 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 68      |
|    fps              | 1417    |
|    time_elapsed     | 1       |
```

```
|    total_timesteps  | 1800     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.0757   |
|    n_updates        | 212      |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 72   |
|    fps             | 1416 |
|    time_elapsed    | 1    |
|    total_timesteps | 1800 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 76       |
|    fps              | 1418     |
|    time_elapsed     | 1        |
|    total_timesteps  | 2000     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.104    |
|    n_updates        | 237      |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 80   |
|    fps             | 1417 |
|    time_elapsed    | 1    |
|    total_timesteps | 2000 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 84       |
|    fps              | 1412     |
|    time_elapsed     | 1        |
|    total_timesteps  | 2200     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.174    |
|    n_updates        | 262      |
----------------------------------
----------------------------
| time/              |      |
```

```
|    episodes        | 88   |
|    fps             | 1411 |
|    time_elapsed    | 1    |
|    total_timesteps | 2200 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 92      |
|    fps              | 1410    |
|    time_elapsed     | 1       |
|    total_timesteps  | 2400    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.194   |
|    n_updates        | 287     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 96   |
|    fps             | 1409 |
|    time_elapsed    | 1    |
|    total_timesteps | 2400 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 100     |
|    fps              | 1407    |
|    time_elapsed     | 1       |
|    total_timesteps  | 2600    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.137   |
|    n_updates        | 312     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 104  |
|    fps             | 1406 |
|    time_elapsed    | 1    |
|    total_timesteps | 2600 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
```

```
| time/               |          |
|    episodes         | 108      |
|    fps              | 1380     |
|    time_elapsed     | 2        |
|    total_timesteps  | 2800     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.268    |
|    n_updates        | 337      |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 112  |
|    fps             | 1378 |
|    time_elapsed    | 2    |
|    total_timesteps | 2800 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 116      |
|    fps              | 1314     |
|    time_elapsed     | 2        |
|    total_timesteps  | 3000     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.19     |
|    n_updates        | 362      |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 120  |
|    fps             | 1312 |
|    time_elapsed    | 2    |
|    total_timesteps | 3000 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 124      |
|    fps              | 1280     |
|    time_elapsed     | 2        |
|    total_timesteps  | 3200     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.17     |
```

```
|    n_updates        | 387      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 128  |
|    fps             | 1280 |
|    time_elapsed    | 2    |
|    total_timesteps | 3200 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 132      |
|    fps              | 1262     |
|    time_elapsed     | 2        |
|    total_timesteps  | 3400     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.117    |
|    n_updates        | 412      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 136  |
|    fps             | 1261 |
|    time_elapsed    | 2    |
|    total_timesteps | 3400 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 140      |
|    fps              | 1241     |
|    time_elapsed     | 2        |
|    total_timesteps  | 3600     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.141    |
|    n_updates        | 437      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 144  |
|    fps             | 1241 |
|    time_elapsed    | 2    |
|    total_timesteps | 3600 |
```

```
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 148     |
|    fps              | 1223    |
|    time_elapsed     | 3       |
|    total_timesteps  | 3800    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.306   |
|    n_updates        | 462     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 152  |
|    fps             | 1221 |
|    time_elapsed    | 3    |
|    total_timesteps | 3800 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 156     |
|    fps              | 1202    |
|    time_elapsed     | 3       |
|    total_timesteps  | 4000    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.188   |
|    n_updates        | 487     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 160  |
|    fps             | 1201 |
|    time_elapsed    | 3    |
|    total_timesteps | 4000 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 164     |
|    fps              | 1179    |
|    time_elapsed     | 3       |
```

```
|    total_timesteps  | 4200     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.278    |
|    n_updates        | 512      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 168  |
|    fps             | 1178 |
|    time_elapsed    | 3    |
|    total_timesteps | 4200 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 172      |
|    fps              | 1165     |
|    time_elapsed     | 3        |
|    total_timesteps  | 4400     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.309    |
|    n_updates        | 537      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 176  |
|    fps             | 1164 |
|    time_elapsed    | 3    |
|    total_timesteps | 4400 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 180      |
|    fps              | 1150     |
|    time_elapsed     | 3        |
|    total_timesteps  | 4600     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.259    |
|    n_updates        | 562      |
---------------------------------
----------------------------
| time/              |      |
```

```
|    episodes        | 184  |
|    fps             | 1149 |
|    time_elapsed    | 4    |
|    total_timesteps | 4600 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 188     |
|    fps              | 1130    |
|    time_elapsed     | 4       |
|    total_timesteps  | 4800    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.29    |
|    n_updates        | 587     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 192  |
|    fps             | 1129 |
|    time_elapsed    | 4    |
|    total_timesteps | 4800 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 196     |
|    fps              | 1110    |
|    time_elapsed     | 4       |
|    total_timesteps  | 5000    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.376   |
|    n_updates        | 612     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 200  |
|    fps             | 1109 |
|    time_elapsed    | 4    |
|    total_timesteps | 5000 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
```

```
| time/               |          |
|    episodes         | 204      |
|    fps              | 1118     |
|    time_elapsed     | 4        |
|    total_timesteps  | 5200     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.369    |
|    n_updates        | 637      |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 208  |
|    fps             | 1118 |
|    time_elapsed    | 4    |
|    total_timesteps | 5200 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 212      |
|    fps              | 1127     |
|    time_elapsed     | 4        |
|    total_timesteps  | 5400     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.131    |
|    n_updates        | 662      |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 216  |
|    fps             | 1127 |
|    time_elapsed    | 4    |
|    total_timesteps | 5400 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 220      |
|    fps              | 1136     |
|    time_elapsed     | 4        |
|    total_timesteps  | 5600     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.136    |
```

```
|    n_updates        | 687      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 224  |
|    fps             | 1135 |
|    time_elapsed    | 4    |
|    total_timesteps | 5600 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 228      |
|    fps              | 1142     |
|    time_elapsed     | 5        |
|    total_timesteps  | 5800     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.38     |
|    n_updates        | 712      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 232  |
|    fps             | 1142 |
|    time_elapsed    | 5    |
|    total_timesteps | 5800 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 236      |
|    fps              | 1151     |
|    time_elapsed     | 5        |
|    total_timesteps  | 6000     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.233    |
|    n_updates        | 737      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 240  |
|    fps             | 1151 |
|    time_elapsed    | 5    |
|    total_timesteps | 6000 |
```

```
-----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 244     |
|    fps              | 1156    |
|    time_elapsed     | 5       |
|    total_timesteps  | 6200    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.394   |
|    n_updates        | 762     |
---------------------------------
-----------------------------
| time/              |      |
|    episodes        | 248  |
|    fps             | 1156 |
|    time_elapsed    | 5    |
|    total_timesteps | 6200 |
-----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 252     |
|    fps              | 1161    |
|    time_elapsed     | 5       |
|    total_timesteps  | 6400    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.105   |
|    n_updates        | 787     |
---------------------------------
-----------------------------
| time/              |      |
|    episodes        | 256  |
|    fps             | 1161 |
|    time_elapsed    | 5    |
|    total_timesteps | 6400 |
-----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 260     |
|    fps              | 1168    |
|    time_elapsed     | 5       |
```

```
|    total_timesteps  | 6600     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.268    |
|    n_updates        | 812      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 264  |
|    fps             | 1168 |
|    time_elapsed    | 5    |
|    total_timesteps | 6600 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 268      |
|    fps              | 1177     |
|    time_elapsed     | 5        |
|    total_timesteps  | 6800     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.218    |
|    n_updates        | 837      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 272  |
|    fps             | 1176 |
|    time_elapsed    | 5    |
|    total_timesteps | 6800 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 276      |
|    fps              | 1183     |
|    time_elapsed     | 5        |
|    total_timesteps  | 7000     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.295    |
|    n_updates        | 862      |
---------------------------------
----------------------------
| time/              |      |
```

```
|    episodes        | 280  |
|    fps             | 1183 |
|    time_elapsed    | 5    |
|    total_timesteps | 7000 |
----------------------------
-----------------------------------
| rollout/            |           |
|    exploration_rate | 0.01      |
| time/               |           |
|    episodes         | 284       |
|    fps              | 1187      |
|    time_elapsed     | 6         |
|    total_timesteps  | 7200      |
| train/              |           |
|    learning_rate    | 0.001     |
|    loss             | 0.318     |
|    n_updates        | 887       |
-----------------------------------
----------------------------
| time/              |      |
|    episodes        | 288  |
|    fps             | 1187 |
|    time_elapsed    | 6    |
|    total_timesteps | 7200 |
----------------------------
-----------------------------------
| rollout/            |           |
|    exploration_rate | 0.01      |
| time/               |           |
|    episodes         | 292       |
|    fps              | 1190      |
|    time_elapsed     | 6         |
|    total_timesteps  | 7400      |
| train/              |           |
|    learning_rate    | 0.001     |
|    loss             | 0.229     |
|    n_updates        | 912       |
-----------------------------------
----------------------------
| time/              |      |
|    episodes        | 296  |
|    fps             | 1190 |
|    time_elapsed    | 6    |
|    total_timesteps | 7400 |
----------------------------
-----------------------------------
| rollout/            |           |
|    exploration_rate | 0.01      |
```

```
| time/               |          |
|    episodes         | 300      |
|    fps              | 1197     |
|    time_elapsed     | 6        |
|    total_timesteps  | 7600     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.457    |
|    n_updates        | 937      |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 304  |
|    fps             | 1197 |
|    time_elapsed    | 6    |
|    total_timesteps | 7600 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 308      |
|    fps              | 1198     |
|    time_elapsed     | 6        |
|    total_timesteps  | 7800     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.716    |
|    n_updates        | 962      |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 312  |
|    fps             | 1198 |
|    time_elapsed    | 6    |
|    total_timesteps | 7800 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 316      |
|    fps              | 1203     |
|    time_elapsed     | 6        |
|    total_timesteps  | 8000     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.329    |
```

```
|    n_updates        | 987      |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 320  |
|    fps             | 1203 |
|    time_elapsed    | 6    |
|    total_timesteps | 8000 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 324      |
|    fps              | 1208     |
|    time_elapsed     | 6        |
|    total_timesteps  | 8200     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.778    |
|    n_updates        | 1012     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 328  |
|    fps             | 1208 |
|    time_elapsed    | 6    |
|    total_timesteps | 8200 |
----------------------------
---------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 332      |
|    fps              | 1212     |
|    time_elapsed     | 6        |
|    total_timesteps  | 8400     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.525    |
|    n_updates        | 1037     |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 336  |
|    fps             | 1212 |
|    time_elapsed    | 6    |
|    total_timesteps | 8400 |
```

```
-----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 340     |
|    fps              | 1214    |
|    time_elapsed     | 7       |
|    total_timesteps  | 8600    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.285   |
|    n_updates        | 1062    |
---------------------------------
-----------------------------
| time/              |      |
|    episodes        | 344  |
|    fps             | 1214 |
|    time_elapsed    | 7    |
|    total_timesteps | 8600 |
-----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 348     |
|    fps              | 1218    |
|    time_elapsed     | 7       |
|    total_timesteps  | 8800    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.396   |
|    n_updates        | 1087    |
---------------------------------
-----------------------------
| time/              |      |
|    episodes        | 352  |
|    fps             | 1218 |
|    time_elapsed    | 7    |
|    total_timesteps | 8800 |
-----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 356     |
|    fps              | 1222    |
|    time_elapsed     | 7       |
```

```
|    total_timesteps  | 9000     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.248    |
|    n_updates        | 1112     |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 360  |
|    fps             | 1222 |
|    time_elapsed    | 7    |
|    total_timesteps | 9000 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 364      |
|    fps              | 1223     |
|    time_elapsed     | 7        |
|    total_timesteps  | 9200     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.958    |
|    n_updates        | 1137     |
----------------------------------
----------------------------
| time/              |      |
|    episodes        | 368  |
|    fps             | 1223 |
|    time_elapsed    | 7    |
|    total_timesteps | 9200 |
----------------------------
----------------------------------
| rollout/            |          |
|    exploration_rate | 0.01     |
| time/               |          |
|    episodes         | 372      |
|    fps              | 1225     |
|    time_elapsed     | 7        |
|    total_timesteps  | 9400     |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.314    |
|    n_updates        | 1162     |
----------------------------------
----------------------------
| time/              |      |
```

```
|    episodes        | 376  |
|    fps             | 1224 |
|    time_elapsed    | 7    |
|    total_timesteps | 9400 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 380     |
|    fps              | 1228    |
|    time_elapsed     | 7       |
|    total_timesteps  | 9600    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.301   |
|    n_updates        | 1187    |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 384  |
|    fps             | 1228 |
|    time_elapsed    | 7    |
|    total_timesteps | 9600 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
| time/               |         |
|    episodes         | 388     |
|    fps              | 1231    |
|    time_elapsed     | 7       |
|    total_timesteps  | 9800    |
| train/              |         |
|    learning_rate    | 0.001   |
|    loss             | 0.445   |
|    n_updates        | 1212    |
---------------------------------
----------------------------
| time/              |      |
|    episodes        | 392  |
|    fps             | 1230 |
|    time_elapsed    | 7    |
|    total_timesteps | 9800 |
----------------------------
---------------------------------
| rollout/            |         |
|    exploration_rate | 0.01    |
```

```
| time/               |          |
|    episodes         | 396      |
|    fps              | 1233     |
|    time_elapsed     | 8        |
|    total_timesteps  | 10000    |
| train/              |          |
|    learning_rate    | 0.001    |
|    loss             | 0.147    |
|    n_updates        | 1237     |
----------------------------------
------------------------------
| time/              |       |
|    episodes        | 400   |
|    fps             | 1232  |
|    time_elapsed    | 8     |
|    total_timesteps | 10000 |
------------------------------
```

[32]:
```
<stable_baselines3.dqn.dqn.DQN at 0x78fb1b617250>
```

Następnie testujemy wytrenowany model. Oprócz tego wygładzamy wyniki za pomocą średniej kroczącej.

[33]:
```
obs = env.reset()
total_rewards = []
num_episodes = 1000
for i in range(num_episodes):
    total_reward = 0
    done = [False for _ in range(env.num_envs)]
    while not all(done):
        action, _states = model.predict(obs)
        obs, rewards, done, info = env.step(action)
        total_reward += sum(rewards)
    total_rewards.append(total_reward)

env.close()

def moving_average(data, window_size):
    return np.convolve(data, np.ones(window_size)/window_size, mode='valid')

window_size = 50
smoothed_rewards = moving_average(total_rewards, window_size)
```

Na koniec wizualizacjemy krzywą uczenia.

[34]:
```
plt.plot(total_rewards, alpha=0.3, label='Total Reward per Episode')
plt.plot(smoothed_rewards, label=f'Moving Average (window size={window_size})')
plt.xlabel('Episode')
```

```
plt.ylabel('Total Reward')
plt.title('Learning Curve')
plt.legend()
plt.show()
```