



CHARUSAT
CHAROTAR UNIVERSITY OF SCIENCE AND TECHNOLOGY



Faculty of Technology and Engineering

Chandubhai S. Patel Institute of Technology

Date: / /

Practical Performa

Academic Year	:	2025-26	Semester	:	7 th
Course code	:	OCCSE4001	Course name	:	Reinforcement Learning

Practical- No. 8

Aim: Design and register a custom Gym environment from scratch, understand its structure, and train/test RL agents on custom tasks.

Code:

```

RL_Prac8.ipynb
File Edit View Insert Runtime Tools Help

Q Commands + Code + Text ▶ Run all ▼

[6] ✓ Os
import gymnasium as gym
from gymnasium import spaces
import numpy as np

/usr/local/lib/python3.12/dist-packages/jupyter_client/session.py:203: DeprecationWarning:
return datetime.utcnow().replace(tzinfo=utc)

[7] ✓ Os
class LineWorldEnv(gym.Env):
    metadata = {"render_modes": ["human"]} # Changed render_modes

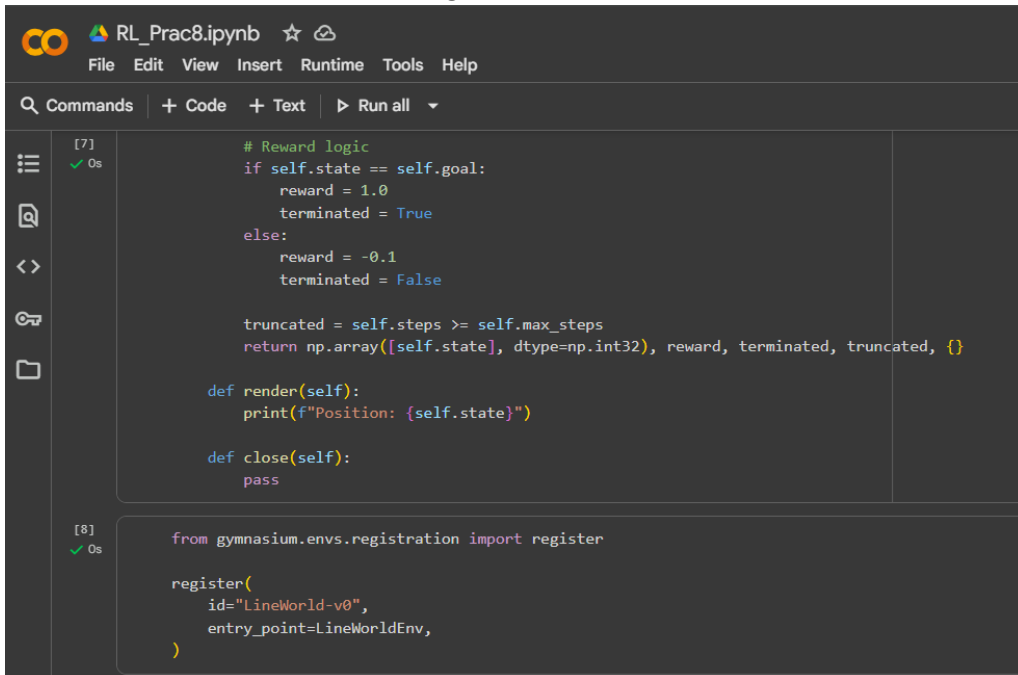
    def __init__(self):
        super(LineWorldEnv, self).__init__()
        # Action space: 0=left, 1=right
        self.action_space = spaces.Discrete(2)
        # Observation: agent's position (int, bounded)
        self.observation_space = spaces.Box(low=-10, high=10, shape=(1,), dtype=np.int32)

        self.state = 0
        self.goal = 10
        self.max_steps = 50
        self.steps = 0

    def reset(self, seed=None, options=None):
        super().reset(seed=seed)
        self.state = 0
        self.steps = 0
        return np.array([self.state], dtype=np.int32), {}

    def step(self, action):
        self.steps += 1
        if action == 0: # move left
            self.state -= 1
        elif action == 1: # move right
            self.state += 1

```



RL_Prac8.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text Run all

```
[7]
✓ 0s
# Reward logic
if self.state == self.goal:
    reward = 1.0
    terminated = True
else:
    reward = -0.1
    terminated = False

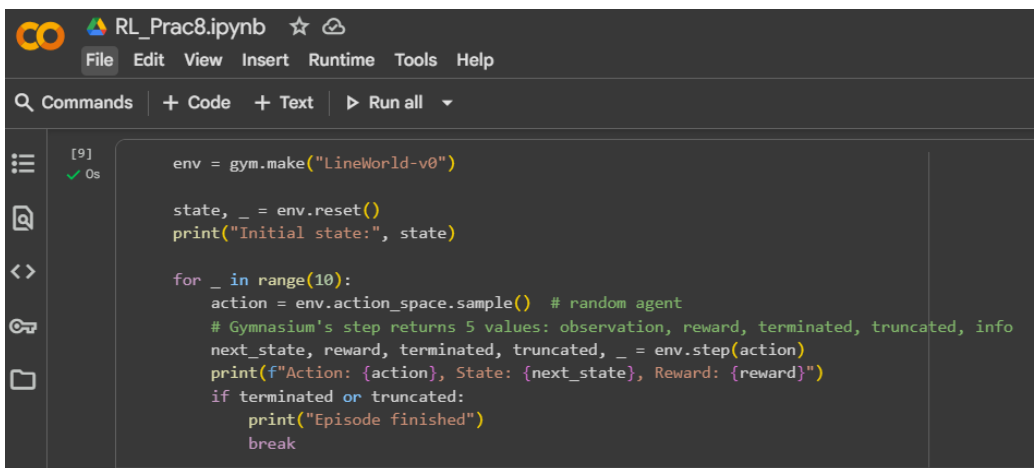
truncated = self.steps >= self.max_steps
return np.array([self.state], dtype=np.int32), reward, terminated, truncated, {}

def render(self):
    print(f"Position: {self.state}")

def close(self):
    pass
```

```
[8]
✓ 0s
from gymnasium.envs.registration import register

register(
    id="LineWorld-v0",
    entry_point=LineWorldEnv,
)
```



RL_Prac8.ipynb

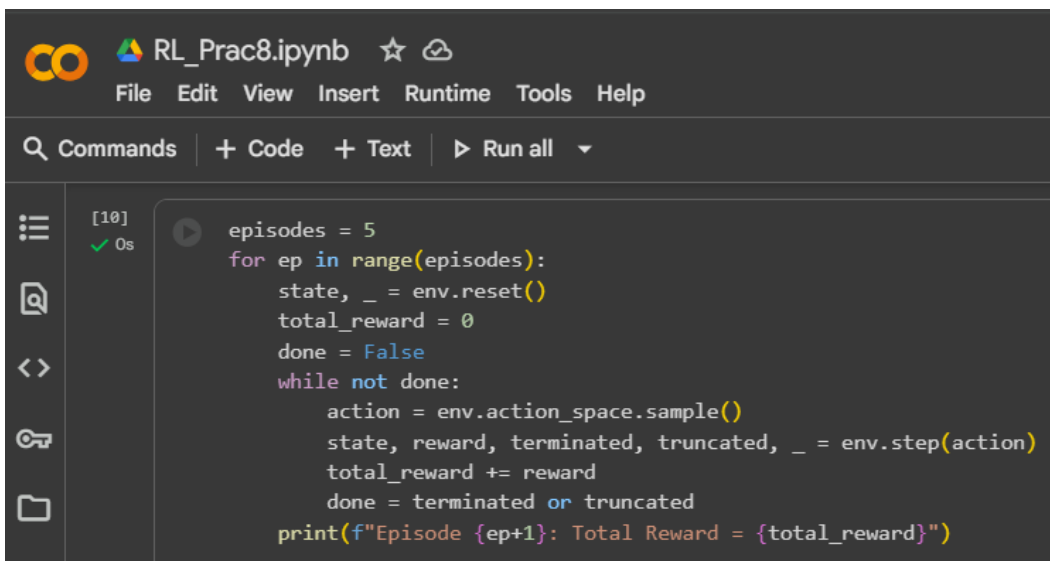
File Edit View Insert Runtime Tools Help

Commands + Code + Text Run all

```
[9]
✓ 0s
env = gym.make("LineWorld-v0")

state, _ = env.reset()
print("Initial state:", state)

for _ in range(10):
    action = env.action_space.sample() # random agent
    # Gymnasium's step returns 5 values: observation, reward, terminated, truncated, info
    next_state, reward, terminated, truncated, _ = env.step(action)
    print(f"Action: {action}, State: {next_state}, Reward: {reward}")
    if terminated or truncated:
        print("Episode finished")
        break
```

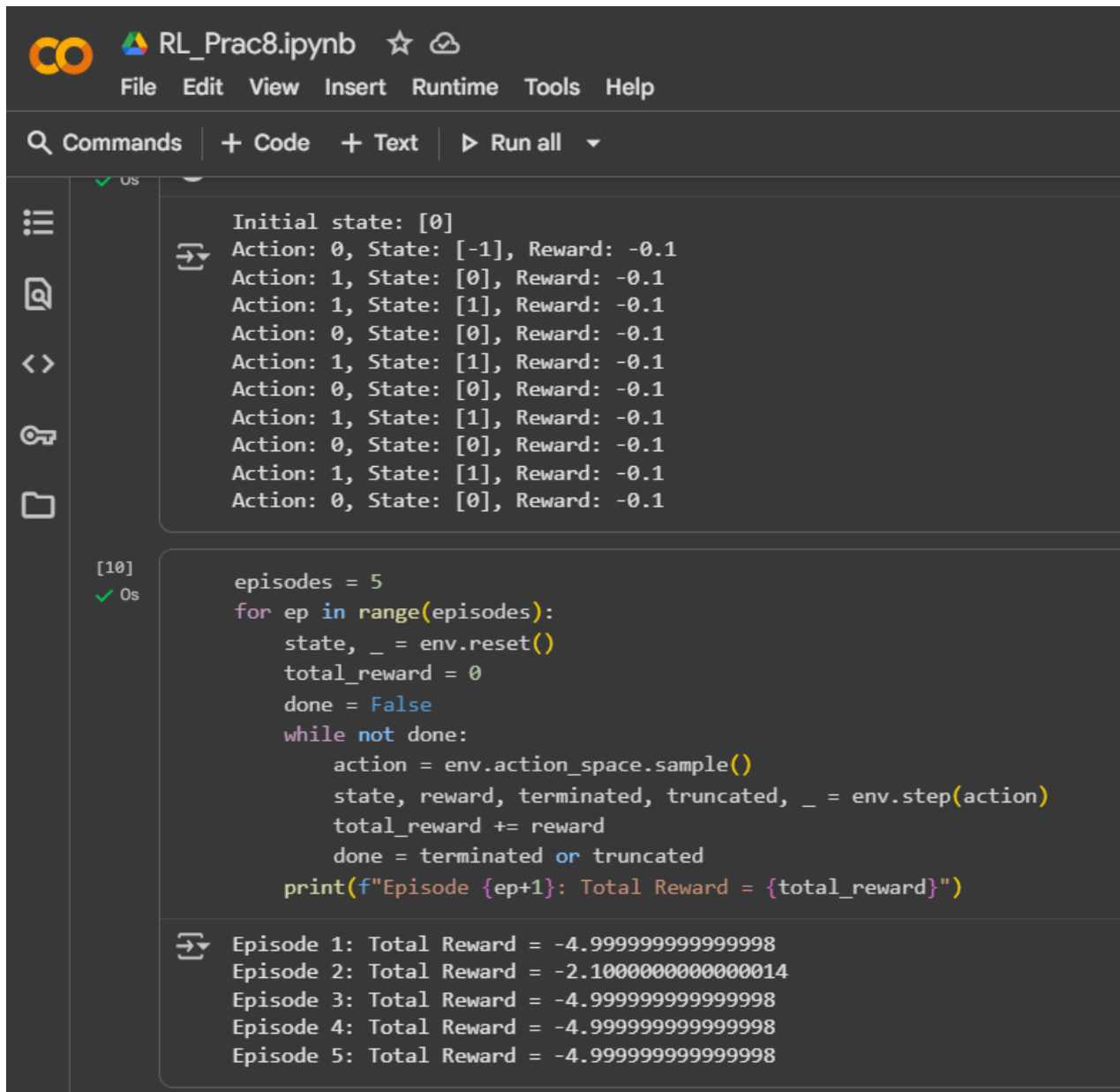


RL_Prac8.ipynb

File Edit View Insert Runtime Tools Help

Commands + Code + Text Run all

```
[10]
✓ 0s
episodes = 5
for ep in range(episodes):
    state, _ = env.reset()
    total_reward = 0
    done = False
    while not done:
        action = env.action_space.sample()
        state, reward, terminated, truncated, _ = env.step(action)
        total_reward += reward
        done = terminated or truncated
    print(f"Episode {ep+1}: Total Reward = {total_reward}")
```

Output:

```
Initial state: [0]
Action: 0, State: [-1], Reward: -0.1
Action: 1, State: [0], Reward: -0.1
Action: 1, State: [1], Reward: -0.1
Action: 0, State: [0], Reward: -0.1
Action: 1, State: [1], Reward: -0.1
Action: 0, State: [0], Reward: -0.1
Action: 1, State: [1], Reward: -0.1
Action: 0, State: [0], Reward: -0.1
Action: 1, State: [1], Reward: -0.1
Action: 0, State: [0], Reward: -0.1

[10]
✓ 0s
episodes = 5
for ep in range(episodes):
    state, _ = env.reset()
    total_reward = 0
    done = False
    while not done:
        action = env.action_space.sample()
        state, reward, terminated, truncated, _ = env.step(action)
        total_reward += reward
        done = terminated or truncated
    print(f"Episode {ep+1}: Total Reward = {total_reward}")

Episode 1: Total Reward = -4.999999999999998
Episode 2: Total Reward = -2.1000000000000014
Episode 3: Total Reward = -4.999999999999998
Episode 4: Total Reward = -4.999999999999998
Episode 5: Total Reward = -4.999999999999998
```

Grade/Marks

(____ / 10)

Sign of Lab Teacher with Date