Quiz

CISC 7404 - Decision Making

Steven Morad

University of Macau

Exam

Exam

- After all students put away their computer/notes/phone I will hand out exams

Exam

- After all students put away their computer/notes/phone I will hand out exams
- If you have computer/notes/phone out after this point, it counts as cheating

Exam

- After all students put away their computer/notes/phone I will hand out exams
- If you have computer/notes/phone out after this point, it counts as cheating
- I will hand out exams face down, do not turn them over until I say so

Exam

- After all students put away their computer/notes/phone I will hand out exams
- If you have computer/notes/phone out after this point, it counts as cheating
- I will hand out exams face down, do not turn them over until I say so
- After turning them over, I will briefly explain each question

Exam

- After all students put away their computer/notes/phone I will hand out exams
- If you have computer/notes/phone out after this point, it counts as cheating
- I will hand out exams face down, do not turn them over until I say so
- After turning them over, I will briefly explain each question
- After my explanation, you will have 75 minutes to complete the exam

Exam

- After all students put away their computer/notes/phone I will hand out exams
- If you have computer/notes/phone out after this point, it counts as cheating
- I will hand out exams face down, do not turn them over until I say so
- After turning them over, I will briefly explain each question
- After my explanation, you will have 75 minutes to complete the exam
- After you are done, give me your exam and go relax outside, we resume class at 8:30

Exam

- After all students put away their computer/notes/phone I will hand out exams
- If you have computer/notes/phone out after this point, it counts as cheating
- I will hand out exams face down, do not turn them over until I say so
- After turning them over, I will briefly explain each question
- After my explanation, you will have 75 minutes to complete the exam
- After you are done, give me your exam and go relax outside, we resume class at 8:30

Exam

- There may or may not be different versions of the exam

Exam

- There may or may not be different versions of the exam
- If your exam has the answer for another version, it is cheating

Exam

- There may or may not be different versions of the exam
- If your exam has the answer for another version, it is cheating
- Instructions are in both english and chinese, english instructions take precedence

Exam

- There may or may not be different versions of the exam
- If your exam has the answer for another version, it is cheating
- Instructions are in both english and chinese, english instructions take precedence
- Good luck!

Exam

- 在所有学生收起电脑/笔记/手机后,我会分发试卷。
- 如果在此之后仍有电脑/笔记/手机未收,将视为作弊。
- 试卷会背面朝下发下,在我宣布开始前请勿翻面。
- 试卷翻面后,我会简要说明每道题的注意事项。
- 说明结束后,你们有 75 分钟完成考试。
- 交卷后请到教室外休息,8:30 恢复上课。
- 试卷可能存在不同版本,细节略有差异。
- 若你的试卷上出现其他版本的答案,将被判定为作弊。
- 试卷说明为中英双语,若内容冲突以英文为准。
- 祝各位考试顺利!

Exam

Our lab is always looking for smart students to work on RL problems

Exam

Our lab is always looking for smart students to work on RL problems

If you thought the exam was easy, come talk to me after class

Review

Coding

Coding

In this course, we will implement MDPs using **gymnasium**

Coding

In this course, we will implement MDPs using **gymnasium**

Developed by OpenAI for reinforcement learning

Coding

In this course, we will implement MDPs using **gymnasium**

Developed by OpenAI for reinforcement learning

Gymnasium provides an **environment** (MDP) API

Coding

In this course, we will implement MDPs using **gymnasium**

Developed by OpenAI for reinforcement learning

Gymnasium provides an **environment** (MDP) API

Must define:

Coding

In this course, we will implement MDPs using **gymnasium**

Developed by OpenAI for reinforcement learning

Gymnasium provides an **environment** (MDP) API

Must define:

- state space (S)

Coding

In this course, we will implement MDPs using **gymnasium**

Developed by OpenAI for reinforcement learning

Gymnasium provides an **environment** (MDP) API

Must define:

- state space (S)
- action space (A)

Coding

In this course, we will implement MDPs using **gymnasium**

Developed by OpenAI for reinforcement learning

Gymnasium provides an **environment** (MDP) API

Must define:
- state space (S)
- action space (A)
- step (Tr, R, terminated)

Coding

In this course, we will implement MDPs using **gymnasium**

Developed by OpenAI for reinforcement learning

Gymnasium provides an **environment** (MDP) API

Must define:
- state space (S)
- action space (A)
- step ($\mathrm{Tr}, R,$ terminated)
- reset (s_0)

Coding

In this course, we will implement MDPs using **gymnasium**

Developed by OpenAI for reinforcement learning

Gymnasium provides an **environment** (MDP) API

Must define:
- state space (S)
- action space (A)
- step (Tr, R, terminated)
- reset (s_0)

https://gymnasium.farama.org/api/env/

Coding

Gymnasium uses **observations** instead of **states**

Coding

Gymnasium uses **observations** instead of **states**

Question: What was the Markov condition for MDPs?

Coding

Gymnasium uses **observations** instead of **states**

Question: What was the Markov condition for MDPs?

The next Markov state only depends on the current Markov state

Coding

Gymnasium uses **observations** instead of **states**

Question: What was the Markov condition for MDPs?

The next Markov state only depends on the current Markov state

$$\Pr(s_{t+1} \mid s_t, s_{t-1}, ..., s_1) = \Pr(s_{t+1} \mid s_t)$$

Coding

Gymnasium uses **observations** instead of **states**

Question: What was the Markov condition for MDPs?

The next Markov state only depends on the current Markov state

$$\Pr\left(s_{t+1} \mid s_t, s_{t-1}, ..., s_1\right) = \Pr\left(s_{t+1} \mid s_t\right)$$

If the Markov property is broken, $s_t \in S$ is not a Markov state

Coding

Gymnasium uses **observations** instead of **states**

Question: What was the Markov condition for MDPs?

The next Markov state only depends on the current Markov state

$$\Pr\big(s_{t+1} \mid s_t, s_{t-1}, ..., s_1\big) = \Pr\big(s_{t+1} \mid s_t\big)$$

If the Markov property is broken, $s_t \in S$ is not a Markov state

Then, we change $s_t \in S$ to an **observation** $o_t \in O$ (more later)

Coding

Gymnasium uses **observations** instead of **states**

Question: What was the Markov condition for MDPs?

The next Markov state only depends on the current Markov state

$$\Pr(s_{t+1} \mid s_t, s_{t-1}, ..., s_1) = \Pr(s_{t+1} \mid s_t)$$

If the Markov property is broken, $s_t \in S$ is not a Markov state

Then, we change $s_t \in S$ to an **observation** $o_t \in O$ (more later)

Gymnasium uses observations, but for MDPs we treat them as states

Coding

```python
import gymnasium as gym

MyMDP(gym.Env):
  def __init__(self):
    self.action_space = gym.spaces.Discrete(3) # A
    self.observation_space = gym.spaces.Discrete(5) # S

  def reset(self, seed=None) -> Tuple[Observation, Dict]

  def step(self, action) -> Tuple[
    Observation, Reward, Terminated, Truncated, Dict
  ]
```

Coding

I stress MDPs in this course because it is the most important concept

Coding

I stress MDPs in this course because it is the most important concept

We will see an example of using RL to solve a task

Coding

I stress MDPs in this course because it is the most important concept

We will see an example of using RL to solve a task

You do not know any RL algorithms, but we will can still solve an MDP

Coding

I stress MDPs in this course because it is the most important concept

We will see an example of using RL to solve a task

You do not know any RL algorithms, but we will can still solve an MDP

If you can make an MDP, you can use existing code to solve it!

Coding

I stress MDPs in this course because it is the most important concept

We will see an example of using RL to solve a task

You do not know any RL algorithms, but we will can still solve an MDP

If you can make an MDP, you can use existing code to solve it!

For example, you can use stable-baselines3

https://stable-baselines3.readthedocs.io

Coding

https://colab.research.google.com/drive/1rDNik5oRl27si8wdtMLE7Y41U5J2bx-I#scrollTo=yVjbC_VQ-Wha