Deadline: Mo. 2023/01/16, 11:59 pm
All submissions from now on will be code and implementations. Therefore please submit your commented code as **jupyter notebook** on Learnweb.

1 OpenAI Gym (20 points)

Gymnasium (former OpenAI Gym originally created by Open AI) is a toolkit to simplify the development and comparison of reinforcement learning algorithms.

1. Install the package via: *pip install gymnasium*

2. Familiarize yourself with the functionality of general gym environments, e.g. CartPole and MountainCar. You can find instructions at https://gymnasium.farama.org/ under "**Basic Usage**".

3. Use the `gym_env_template.ipynb` on learnweb to implement the simple game environment we used in the last tutorials. Suppose the environment dynamics are described by the following elements:

   Transition matrices:

$$P(s' \mid s, jump) = \begin{array}{c} \\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array} \begin{array}{c} \begin{array}{ccccc} s_0 & s_1 & s_2 & s_3 & s_4 \end{array} \\ \begin{bmatrix} 1/5 & 4/5 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 3/5 & 2/5 & 0 \\ 0 & 0 & 1/5 & 1/5 & 3/5 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

$$P(s' \mid s, walk) = \begin{array}{c} \\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array} \begin{array}{c} \begin{array}{ccccc} s_0 & s_1 & s_2 & s_3 & s_4 \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 4/5 & 1/5 & 0 \\ 0 & 0 & 9/10 & 1/10 & 0 \\ 0 & 0 & 2/5 & 2/5 & 1/5 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

   Direct rewards:

$$r(s, a) = \begin{array}{c} \\ s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array} \begin{array}{c} \begin{array}{cc} jump & walk \end{array} \\ \begin{bmatrix} -2 & 0 \\ -1 & -1 \\ -2 & 0 \\ -1 & -1 \\ 0 & 0 \end{bmatrix} \end{array}$$

   Note: You don't have to register the environment with the gym registry.

4. Finally implement a sampler to sample trajectories from the gym environment given our two known deterministic policies:
   - π_1: {always jump}
   - π_2: {first walk, then always jump}