

Artificial Intelligence & Its Applications

Laboratory 02

Submission Due: 31-July-2020

Submission Procedure:

1. Use Pickle to save all variables into a file named “Lab02_XX_YY.pkl”
2. Compress the programs of each question into a file named “Lab02_XX_YY.zip”
3. Send “Lab02_XX_YY.py” and “Lab02_XX_YY.zip” to your monitor.

XX is your student ID.

YY is your first name + last name in small letter without spacing.

For example:

Chen Tao and ID 20190123456789

The file names should be

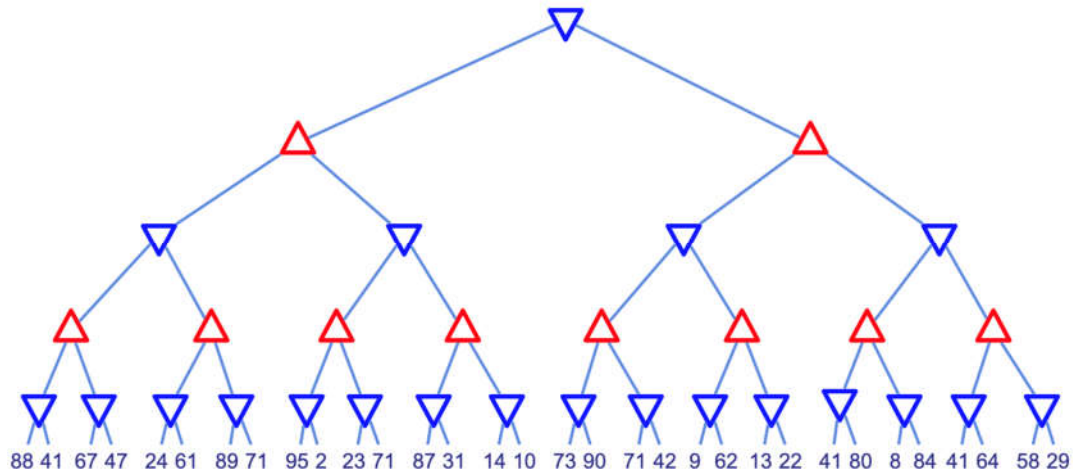
“Lab02_20190123456789_chentao.pkl” and

“Lab02_20190123456789_chentao.zip”

minimax and **gym** in Python are used in this laboratory. For detail, please refer to

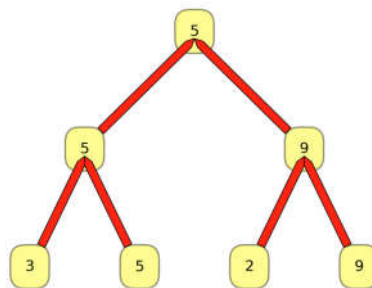
Lab2_sup.pdf

1. Consider a minimax game which has 2^n final states and paths to reach final state are shown in the following complete binary tree. Assume you are the maximizing player and you get the first chance to move. Find the values of all non-leaf nodes in the tree.



Answer Requirement:

- Save answer in **Q1** in the tree format of *minimax* library. Be noted that the tree format is a **BiTree** object, and can be visualized by **tree.view_in_graph()**



2. A and B are playing a game. There are N coins at the beginning, and a player can pick 7 or 4 or 1 coins in each move. A always starts the game. The player who picks the last coin wins the game. Assume the strategy of both are minimax.
- i) Who will win the game when $N = 100$?
 - ii) When N is from 0 to 200, in which cases A will win?

Answer Requirement:

- 1. Save the answer in **Q2_i** in this format: A
- 2. Save the sequences in **Q2_ii** in this format: [0, 1, 0, 1, ...]
Note: Set $\text{list}[0] = 0$ because $N=0$ is nonsense. $\text{list}[i]=i$ represents A will win if $N=i$, $1 \leq i \leq 200$. The list contains 201 elements in total.

3. A pole is attached by an un-actuated joint to a cart, which moves along a frictionless track. The system is controlled by applying a force of +1 or -1 to the cart. The pendulum starts upright, and the goal is to prevent it from falling over. A reward of +1 is provided for every timestep that the pole remains upright. The episode ends when the pole is more than 15 degrees from vertical, or the cart moves more than 2.4 units from the center.
- Repeat the experiment 30 times with increased gamma to observe that as gamma increases, the score will also increase until gamma reaches the tipping point. This tipping point is the best value for gamma.
 - Find the best gamma to maximize the score.



Answer Requirement:

- Save the answer in **Q4_i** in this format:
`[(q_table, 0.1, 196), (q_table, 0.2, 243), (q_table, 0.3, 201) ...]`
where (Q-table, gamma, score), and totally 30 tuples.
- Save the sequences in **Q4_ii** in this format: `(q_table, 0.2, 243)`
where (Q-table, gamma, score)