Reinforcement Learning Exercise

Instructions: This exercise involves implementing a very simple reinforcement learning algorithm to learn a net-positive strategy in a simple card game. While the player will start off playing randomly, and poorly, they will quickly converge on a strategy that has positive overall expectation.

1. **Setting Up**: Two decks of cards are filled with an equal number of face cards (Jack, Queen, King, Ace), where cards can repeat multiple times. The two decks are shuffled, and placed side by side. (Code is supplied that will generate a simulation of this process and display cards.) Each player is given a copy of the worksheet. The player should fill in each of the boxes in the top two columns with an initial Q-value of 0.

2. On each round of the game, the players turn over a card from the left-hand deck. They then choose an action (either `Higher` or `Lower`). This choice is handled as follows:

   - Each player compares the two values in the row for the face-value of the card seen. If one value is greater than the other, the corresponding action is chosen. For instance, if the card is an `Ace` and the values under `Higher` and `Lower` are 1 and 0, respectively, the player chooses `Higher`. (Note: these numbers can be positive or negative.)
   - If the two values are the same, then the player tosses a coin: if the coin comes up Heads, they choose `Higher`, and else they choose `Lower`.

   (The player may want to write down their choice.)

3. After the players choose their actions, the next card in the right-hand (opponent) deck is turned over. Scoring is handled as follows:

   - If the player has chosen `Higher`, then:
     - If the opponent card is higher than the player card, the player wins. They record a score of $+1$ in the next unfilled box at the bottom of the sheet, and add 1 to the Q-value for `Higher` in the row with the face-value of the **player's card**.
     - If the opponent card is lower than the player card, the player loses. They record -1 in the next unfilled box at the bottom of the sheet, and subtract 1 from the Q-value for `Higher` in the row associated with the face-value of the **player's card**.
     - If the two cards are the same, this is a tie. The player records a 0 in the next unfilled box at the bottom of the sheet. No Q-value updates are needed.
   - If the player has chosen `Lower`, then:
     - If the opponent card is lower than the player card, the player wins. They record a score of $+1$ in the next unfilled box at the bottom of the sheet, and add 1 to the Q-value for `Lower` in the row with the face-value of the **player's card**.
     - If the opponent card is higher than the player card, the player loses. They record -1 in the next unfilled box at the bottom of the sheet, and subtract 1 from the Q-value for `Lower` in the row associated with the face-value of the **player's card**.
     - If the two cards are the same, this is a tie. The player records a 0 in the next unfilled box at the bottom of the sheet. No Q-value updates are needed.