## Time Allowed: 3 Hours

# Reinforcement Learning Assignment

## Module 3 & 4

MGM Casino at Las Vegas has introduced a modified version of Poker game that has just 5 cards (10, Jack, Queen, King and Ace) and just 2 types (Diamonds and Spades). In this new game there are 2 cards that are given to each player and 2 cards are opened / drawn on the table and there is a guaranteed return on whatever amount the player bets. There are 25 combinations of 2 cards that are dealt in the beginning and 625 combinations of rest of the cards being drawn with each of them having different probabilities. This has been calculated from the past data and is contained in 2 excel sheets – “Probabilistic\_Rewards.csv” contains the rewards for the 25 combinations possible and “Transition\_Probabilities.csv” contains the joint probabilities of a. hand that you have and b. probability of other 2 cards being drawn. The total duration of the game is 1 hour. The calculation of balance, WTB, amount after risk reduction etc. is done as per the formulae given by the casino. The amount that a person can start to play with is not fixed. However, the minimum bet is 50% of the amount that any player starts with. So if a player starts with $5000, then the minimum bet is $2500.

Description of Terms used:

1. Hand – the 2 cards that any player has
2. Cards drawn / opened – the 2 cards that are opened on the table
3. OptionsAvailable – the options available to any player given the cards that the player has
4. Current\_balance – the amount that the player brought initially to play with
5. WTB\_amount – the amount that the player is willing to bet
6. guaranteed\_win – the amount that serves as a guaranteed return in the game
7. time\_remaining – the time remaining for each bet for each hand
8. pay\_out – the amount that is expected to be received if the player wins
9. amount\_after\_risk\_reduction – the amount that the player will receive including the unpredictability (possibility of loss)
10. BASP – Betting Amount System Prediction – the amount predicted by the system for each hand that the player should bet

**Task 1:** You are required to identify what is the optimal policy (set of rules defined by states (pair of cards that may be drawn in the beginning) and actions (play or fold) that will give the highest probability of winning.

*Hint:* Places where you have to write the code are denoted by blanks “\_\_\_\_\_\_\_\_\_” in the file “Task\_1\_Question.py”.

**Expected Result:** This task should answer the question – out of all the possible combinations of the 10 different cards, what hand should I get so that I maximize my chances of winning?

**Task 2:** Use the optimal policy derived in the first task to identify the amount that a user should bet on each hand out of his total amount to win the game. Note that even if the optimal policy suggests to fold, a player can choose to ignore and bluff. So you are required to suggest the amount in this case also where the amount to be bet will be the bluff amount.

*Hint:* Places where you have to write the code are denoted by blanks “\_\_\_\_\_\_\_\_\_” in the “Task\_2\_Question.py”.

**Expected Result:** This task should answer the question – how much should I bet if I get the hands in the best policy given I have “x” amount?

**Note:** These tasks make use of Value Iteration, Bellman Equation Markov Decision Process and Monte Carlo.