Problem 1a

Videep Ekbote, 2023MT10940

We will use a recursive approach for this problem. It is given that Alice wins the first match and Bob wins the second match. We shall use this information to decide our base cases.
Let $f(x,t)$ denote the probability that Alice wins x rounds out of t rounds played till now. This will be computed recursively as the sum of two terms:

$$f(x,t) = \frac{x}{t-1} f(x,t-1) + \frac{t-x}{t-1} f(x-1,t-1)$$

This recurrence consists of two terms since there are two possibilities for each round, which can make the situation in the round as "Alice has won x out of t rounds". One possibility is that Alice loses the t-th round, and the second possibility is that Alice wins the t-th round.

Possibility 1: Alice loses the t-th round.
This means that till the previous round, Alice still has x points. So, in the $(t-1)$-th round, the situation is as follows: Alice has won x rounds, and Bob has won $t-1-x$ rounds. Now, since in this case Alice loses the t-th round (or equivalently Bob wins the t-th round), we can say that the conditional probability that Bob wins the t-th round, given that Alice has won x out of $t-1$ rounds, is:

$$\frac{x}{t-1}$$

This probability comes from the definition of the Aggressive-Aggressive situation of the payoff matrix. To get the total probability of this case, we need to multiply the conditional probability with the probability of Alice winning x out of $t-1$ rounds, which is simply given by $f(x,t-1)$. This is exactly the first term of the recursive function $f(x,t)$. So the total probability for this case is given by

$$\frac{x}{t-1} f(x,t-1)$$

Possibility 2: Alice wins the t-th round.
This means that till the previous round, Alice has $x-1$ points. So, in the $(t-1)$-th round, the situation is as follows: Alice has won $x-1$ rounds, and Bob has won $(t-1-(x-1)) = t-x$ rounds. Since in this case Alice wins the t-th round, we can say that the conditional probability that Alice wins the t-th round, given that Alice has won $x-1$ out of $t-1$ rounds, is:

$$\frac{t-x}{t-1}$$

This probability comes from the payoff matrix. To get the total probability of this case, we need to multiply the conditional probability with the probability of Alice winning $x-1$ out of $t-1$ rounds, which is simply given by $f(x-1,t-1)$. This forms the second term of the recursive relation. So the total probability for this case is given by

$$\frac{t-x}{t-1} f(x-1,t-1)$$

Since Possibility 1 and Possibility 2 are independent of each other, the probabilities add, and we get the total probability as

$$f(x,t) = \frac{x}{t-1} f(x,t-1) + \frac{t-x}{t-1} f(x-1,t-1)$$

Explanation of the Base Cases:
The base case is fixed, as we are given the outcomes of the first two games. Specifically, we are told that Alice wins the first round and Bob wins the second round. Therefore, each has won 1 round by the second round. For $x = 0$ and $t = 2$, we return 0, as it is impossible for Alice to win 0 rounds given that she has won 1 round and Bob has won 1 round. For $x = 1$ and $t = 2$, we return 1, since the probability of Alice winning exactly 1 round (which is the only possible outcome) is 1. For $x = 2$

and $t = 2$, we return 0, as it is impossible for Alice to have won 2 rounds given that Bob has won 1 round.

Note: In the code, the cases where $x = 1$ and $x = t-1$ are handled separately to avoid index-out-of-range errors.

Problem 1b

Videep Ekbote, 2023MT10940

We start by defining a new random variable Y. Y represents the "Net" score of the game (from Alice's point of view) after T rounds.

Define:

$$Y = X_1 + X_2 + X_3 + \cdots + X_T = \sum_{i=1}^{T} X_i$$

Expectation:
We are required to calculate $E[Y]$.
Clearly, Y is a discrete random variable, so its expectation is given by:

$$E[Y] = \sum_i Y_i \cdot P(Y = Y_i)$$

The possible values of Y, denoted as Y_i, are:

$$\begin{cases} \text{All odd integers in the range } [-(T-2), (T-2)] & \text{if } T \text{ is odd} \\ \text{All even integers in the range } [-(T-2), (T-2)] & \text{if } T \text{ is even} \end{cases}$$

The argument is as follows. The minimum possible value of Y is $-T$, which happens when all X_i's are equal to -1. The maximum possible value of Y is $+T$, which happens when all X_i's are equal to 1. However, Y cannot take the value T, because this would mean Alice wins all matches, which is impossible since the initial condition requires Alice to lose the second match. Thus, the bounds for Y are $-T+1$ and $T-1$.

Moreover, Y cannot take the value $T-1$ either. Alice can win a maximum of $T-1$ rounds. In that case, Y would be $T - 1 + (-1) = T - 2$. Thus, the final bounds for Y are $-T+2$ and $T-2$.
Finally, we argue that the parity of all possible values of Y is the same as the parity of T. Let g be the number of rounds Alice wins out of T rounds, and let her lose $T - g$ rounds. The value of Y is then:

$$Y = g + (-1) \cdot (T - g) = 2g - T$$

Thus, $Y = 2g - T$, and hence $g = \frac{Y+T}{2}$. Since g is an integer (as it represents the number of rounds Alice wins), Y must have the same parity as T.
This completes our proof of demonstrating the allowed values of Y.

PMF of Y:
We now require the probability mass function (PMF) of Y, which gives the probability that Y takes on the value Y_i for each i.
We argue that $P(Y = Y_i)$ is the probability that Alice wins g rounds out of T total rounds, where $Y_i = 2g - T = g + (-1) \cdot (T - g)$.

Formally,

$$P(Y = Y_i) = f\left(\frac{Y_i + T}{2}, T\right)$$

where $f(x, t)$ is the function defined in problem 1a) that gives the probability Alice wins x rounds out of t rounds played.

We now show that $E[Y] = 0$.
Since Alice wins the first game and Bob wins the second game, the situation is symmetrical from the third game onwards, as both are playing aggressively. The probability of Alice winning x rounds is the same as Bob winning x rounds. Hence, the expectation is zero:

$$E[Y] = 0$$

Variance of Y:
For any random variable Y, the variance $\text{Var}(Y)$ is given by:

$$\text{Var}(Y) = E[Y^2] - (E[Y])^2$$

In our case, $E[Y] = 0$, so:

$$\text{Var}(Y) = E[Y^2]$$

This can be computed as:

$$\text{Var}(Y) = \sum_i Y_i^2 \cdot P(Y = Y_i)$$

We have already provided the arguments for the possible values of Y_i and $P(Y = Y_i)$.

Hence, we have successfully shown the arguments used to compute $E[Y]$ and $\text{Var}(Y)$.

Problem 2a

Videep Ekbote, 2023MT10940

The optimal strategy for Alice is a Greedy strategy. This can be argued since we only have to maximize the points we gain in this current round, and the points we gain in future rounds do not matter. Since we are given that Bob is predictable and plays according to the past game, Alice's strategy is entirely governed by the outcome of the last game, and Alice plays in such a way that she maximizes her points in this current game, thereby using a Greedy approach.

CASE 1:
If Bob wins the previous round, then he will play defensively in the current round. Given that Bob plays defensively, we want the move that will maximize Alice's points in the current round.

If Alice plays Balanced, expected number of points = $\frac{3}{10} + \frac{1}{2} \times \frac{1}{2} = 0.55$
If Alice plays Defense, expected number of points = $\frac{1}{10} + \frac{1}{2} \times \frac{4}{5} = 0.5$
If Alice plays Aggressive, expected number of points = $\frac{5}{11} = 0.4545$

Strategy: Alice plays Balanced.

CASE 2:
If Bob draws the previous round, then he plays balanced in the current round. Given that Bob plays balanced in the current round, we want the move that will maximize Alice's points in the current round.

If Alice plays Balanced, expected number of points = $\frac{1}{3} + \frac{1}{6} = 0.5$
If Alice plays Defense, expected number of points = $\frac{1}{5} + \frac{1}{4} = 0.45$
If Alice plays Aggressive, expected number of points = $\frac{7}{10} = 0.7$

Strategy: Alice plays Aggressive.

CASE 3:
If Bob loses the previous round, then he plays Aggressive in the current round. Given that Bob plays Attack in the current round, we want the move that will maximize Alice's points in the current round.

If Alice plays Balanced, expected number of points = $\frac{3}{10} = 0.3$
If Alice plays Defense, expected number of points = $\frac{6}{11} = 0.545$
If Alice plays Aggressive, expected number of points = $\frac{n_B}{n_B+n_A}$

Strategy: If $\frac{n_B}{n_B+n_A} > \frac{6}{11}$, Alice plays Aggressive. Otherwise, Alice plays Defense.

Since Alice bases her strategy solely on what Bob played in the previous round, we can conclude that her strategy is Greedy. Additionally, Monte Carlo simulations have shown the superiority of the greedy strategy.

Problem 2b

Videep Ekbote, 2023MT10940

We will prove that it is not optimal for Alice to play the Greedy strategy in every round to maximize the total number of points. We shall provide a proof of existence by empirical observations. We will show that there exists some strategy which is Non-Greedy but outperforms the Greedy strategy in one Monte Carlo simulation.

We shall run each simulation for 10 rounds and run 100 such simulations. In each simulation, we will use 81 different strategies and show that in at least one of the 100 simulations, there exists a Non-Greedy strategy that scores more points than the Greedy strategy. However, it is not necessary that in every simulation, the same Non-Greedy strategy will outperform the Greedy strategy. Instead, across 100 simulations, there will exist at least one Non-Greedy strategy that is more optimal than the Greedy strategy. The Non-Greedy strategy may vary from one simulation to another.

These 81 strategies are obtained by modifying the Greedy strategy. Since the Greedy strategy is unique, any modification to it will result in a Non-Greedy strategy. We modify the Greedy strategy as follows:

Whenever Bob plays defensively, according to the Greedy strategy we return 1, but in general, we can return any number from the set $\{0, 1, 2\}$, and we call this number d.
Whenever Bob plays aggressively, according to the Greedy strategy we return 0 if $\frac{n_B}{n_A+n_B} > \frac{6}{11}$ and 2 otherwise. However, we can return any numbers from the set $\{0, 1, 2\}$, which we call a and b.
Whenever Bob plays balanced, according to the Greedy strategy, Alice plays 0. But in general, we can return any number from the set $\{0, 1, 2\}$, and we call this number c.

The tuple (a, b, c, d) forms a 4-tuple, and there are 81 such 4-tuples. This follows from the fact that each of a, b, c, and d can take values from the set $\{0, 1, 2\}$, resulting in $3 \times 3 \times 3 \times 3 = 81$ possibilities. In the 100 simulations of 10 rounds each, if we find any one 4-tuple $(a, b, c, d) \neq (0, 2, 0, 1)$ (i.e., not equal to the Greedy 4-tuple) that returns more points than the Greedy 4-tuple, we can conclude that there exists a Non-Greedy strategy that outperforms the Greedy strategy. Thus, the Greedy strategy is not always the most optimal when it comes to maximizing the total number of points.

If the Greedy strategy were optimal, then in all 100 simulations of 10 rounds each, the Greedy strategy would achieve the highest number of points, and the number of wins for the Greedy strategy would be 100. However, after running the code of problem 2b, we find that the number of wins for the Greedy strategy in these 100 simulations is less than 100, and hence, the Greedy strategy is not optimal.

Problem 2c

Videep Ekbote, 2023MT10940

We are given a random variable τ which represents the number of rounds it takes for Alice to get T wins. We need to estimate the expected value of τ by running simulations.

We shall calculate the expected value of τ many times and take the mean of all those values. This mean will give us the estimated value of τ. Every time we are in a simulation, we need to initialize the conditions to Alice's 1 win and Bob's 1 win, and also reinitialize the payoff matrix. This is done because we are calculating τ for each simulation and taking the mean. In each simulation, while the number of wins of Alice in that simulation is less than T, we check whether Alice wins or not. If she wins, then the win counter is updated by 1, and the number of rounds that Alice takes to get T wins also gets updated by 1. If she does not win, only the number of rounds that Alice takes to get T wins get updated by 1.

By following this procedure, we get the mean value of τ.

Problem 3a

Videep Ekbote, 2023MT10940

The optimal strategy for Alice is a Greedy strategy. This can be argued since we only have to maximize the points we gain in this current round, and the points we gain in future rounds do not matter. Here, we are given that Bob is unpredictable, so Alice has to choose her moves keeping in mind this unpredictability.

CASE 1:
Alice chooses to play aggressive when Bob's move is random. Since Bob's moves are uniformly random, this means that over a large number of rounds, $\frac{1}{3}$ of the time Bob will choose to play aggressively, $\frac{1}{3}$ of the time Bob will choose to play balanced, and $\frac{1}{3}$ of the time Bob will choose to play defense. So, if Alice plays aggressive, her expected payoff points will be given by:

$$E_{\text{aggression}} = \frac{1}{3}\left(\frac{n_B}{n_A+n_B}\right) + \frac{1}{3}\left(\frac{7}{10}\right) + \frac{1}{3}\left(\frac{5}{11}\right) = \frac{1}{3}\left(\frac{n_B}{n_A+n_B}\right) + \frac{254}{660}$$

CASE 2:
Alice chooses to play balanced when Bob's move is random. Since Bob's moves are uniformly random, this means that over a large number of rounds, $\frac{1}{3}$ of the time Bob will choose to play aggressively, $\frac{1}{3}$ of the time Bob will choose to play balanced, and $\frac{1}{3}$ of the time Bob will choose to play defense. So, if Alice plays balanced, her expected payoff points will be given by:

$$E_{\text{balanced}} = \frac{1}{3}\left(\frac{3}{10}\right) + \left(\frac{1}{3}\left(\frac{1}{3}\right) + \frac{1}{3}\left(\frac{1}{3}\right)\cdot 0.5\right) + \left(\frac{1}{3}\left(\frac{3}{10}\right) + \frac{1}{3}\cdot\left(\frac{1}{2}\right)\cdot 0.5\right) = \frac{297}{660}$$

CASE 3:
Alice chooses to play Defense when Bob's move is random. Since Bob's moves are uniformly random, this means that over a large number of rounds, $\frac{1}{3}$ of the time Bob will choose to play aggressively, $\frac{1}{3}$ of the time Bob will choose to play balanced, and $\frac{1}{3}$ of the time Bob will choose to play defense. So, if Alice plays defense, her expected payoff points will be given by:

$$E_{\text{defense}} = \frac{1}{3}\left(\frac{6}{11}\right) + \left(\frac{1}{3}\left(\frac{1}{5}\right) + \frac{1}{3}\left(\frac{1}{2}\right)\cdot 0.5\right) + \left(\frac{1}{3}\left(\frac{1}{10}\right) + \frac{1}{3}\left(\frac{4}{5}\right)\cdot 0.5\right) = \frac{329}{660}$$

Considering cases 1, 2, and 3, we can conclude that over a large number of rounds, given that Bob plays uniformly random, Alice should never play balanced, since her expected payoff from playing balanced in each round is less than her expected payoff from playing defense. Also, Alice will play defense until her expected payoff per round by playing aggressive ($E_{\text{aggression}}$) becomes more than her expected payoff per round by playing defense (E_{defense}).
More formally, whenever

$$\frac{1}{3}\left(\frac{n_B}{n_A+n_B}\right) + \frac{254}{660} > \frac{329}{660}$$

Alice will play aggressive. So whenever

$$\frac{n_B}{n_A+n_B} > \frac{15}{44}$$

Alice plays aggressive otherwise Alice plays defensive.

Since Alice has to just maximize the number of points in the current round and not care about future rounds, this is a greedy strategy, and the superiority of this strategy is also verified by the Monte Carlo simulations.

Problem 3b

Videep Ekbote, 2023MT10940

Since Bob is playing at random, we have to try out all the possibilities of Alice's play style, i.e whether she plays aggressive, balanced or defensive in each round. We require a strategy that maximizes the expected number of points in the next T rounds, so we will have to develop a strategy based on dynamic updation, due to the randomness of Bob's moves. The expected value for next T rounds can be broken into two parts as follows. Firstly, we calculate the expected number of points for this round which we shall get from the payoff matrix. Secondly, we calculate the maximum expected value we will get in the next t-1 rounds. This will be defined recursively and we will approach by solving the expected values subproblems from 1 to T.

This recursive strategy will give us the maximum number of expected rounds that Alice will have at the end of T future. The strategy, which yields this expectation is the strategy that Alice follows to maximize the expected number of points for T future rounds.