Penney's game

Penney's game works as follow: two players, A and B, agree on an integer $k \geq 3$. then, A decides a sequence of length k consisting of heads (H) and tails (T). This sequence is shown to B, which can then determine a sequence of length k. (For example, $k = 4$, A's sequence = HTTH, B's sequence = TTHT. Therefore, a fair coin is tossed until one of the selected sequences first appears. The player whose sequence turns up first wins. Is this a fair game, i.e. do the players have the same chance to win? This will be investigated in this assignment.

1. Play the game few times with each others. Does the game seem to be fair, i.e. do all the sequences seem to have the same probability of winning?

2. We start by studying a single sequence. We flip a fair coin 100 times. Let X_S denote the number of times a certain sequence S shows up.

   (a) Suppose S has length 4. Consider the first four coin coin flips. Let

   $$I_1 = \begin{cases} 1, & \text{if the first four flips produce } S \\ 0, & \text{otherwise} \end{cases}$$

   I_1 is a random variable which is often called indicator variable. COmpute $\mathbb{E}[I_1]$.

   (b) We can compare the sequences of four flips with our sequence S. That is, we look at the coin flips $i, i+1, \ldots, i+3$ and we are interested in

   $$I_i = \begin{cases} 1, & \text{if the four flips } i, i+1, \ldots, i+3 \text{ produce } S \\ 0, & \text{otherwise} \end{cases}$$

   Use $I_1, I_2, \ldots$ to find an expression for X_S.

   (c) Compute $\mathbb{E}[X_S]$ for few different S. Use task 2(b).

   (d) Given your answer to point 2(c), do you think that the game is fair?

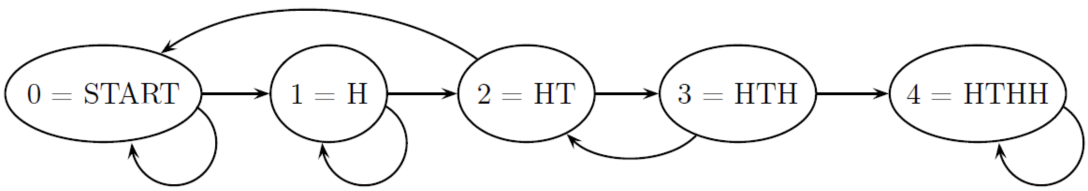
3. For a given sequence of length k consisting of H and T one can construct a Markov chain which illustrates the coin flips until the sequence shows up for the first time. Let the states be:

   - 0 = START,
   - 1 = letter 1 of the sequence,
   - 2 = letter 1 and 2 of the sequence,
   - $\vdots$
   - k = whole sequence.

We start at state 0, and for each toss of the coin we move to the state:

- state k if the entire sequence has appeared, otherwise
- state j, for $j = 0, \ldots, k-1$, if the latest j flips corresponds to the beginning of the sequence, but the $j+1$ latest flip does not.

It turns out (convince yourself!) that the next state depens only on the current state and the outcome of the next coin toss, which make so that we can model the walk between the states with a Markov chain. The figure shows the Markov chain in the case of the sequence HTHH. Each arrow corresponds to the outcome of a coin toss, i.e. H or T, and has probability $1/2$ (except the last arrow that has probability 1).



The chain will reach the absorbing state k the same time as the selected sequence appears for the first time. Draw the corresponding figures for the sequences HTHT and THTT. (if you don't want to use a computer, you can draw them by hand).

4. We will now use absorbing Markov chain (chapter 11.2 in GS) to study the time it takes before a specific sequence appears.
   (a) Let N_A be the number of flips until the sequence A appears. Use the figures from task 3 and Theorem 11.5 in GS to calculate $\mathbb{E}[N_A], \mathbb{E}[N_B]$, where A=HTHT and B=THTT. Choose also some other sequences.
   (b) Which of the sequences HTHT and THTT do you think will win, i.e. will appear first, given the above results?
5. Similarly as above, one can construct a Markov chain for two sequences. The state in which the chain is depends on how far both sequences have show up. There are also two absorbing states, one for each sequence. The state in which the chain is absorbed is the one of the sequence that first appeared.
   (a) What are the possible states for the Markov chain for the sequences HTHT and THTT? What states are absorbing?
   (b) Draw the corresponding figure.
6. If the Markov chain has multiple abrorbing states, we can calculate the probability that it will be absorbed in a specific absorbing state (it will always be absorbed in one of the absorbing states).

(a) Compute the probability that HTHT will win against THTT with the help of Theorem 11.6 in GS.

(b) Reflect on the answer and compare with those in tasks 2(d) and 4(b).

Optional questions

7. Study the game for other and preferably longer sequences. Compute exactly as above or simulate. Try to find a strategy for player B to choose his sequence. Hint: think of A=HHHH and B=THHH. In which case does A win?

8. Study the three sequences A=HTHT, B=THTT and C=HTTH. Estimate the probability that A wins over B, B wins over C, and C wins over A, by means of simulation. Then let all three sequences compete against each other, and estimate the probability that A, B or C win. Comment on the results.