Snakes and Ladders

Yasmin Bettles
Student ID: 2158064

December 2022

1 Introduction

Snakes and Ladders originated in ancient India, before being brought to the UK by John Jacques in 1892. [1] It is a game of chance, usually for two or more players, wherein the players roll a six-sided die, and move forwards on the board according to their roll. The players begin on square "zero", which in practice means they begin off the board, and enter the board when their first turn is taken. The board has "snakes" and "ladders" which are spaces that cause the player to move backwards or forwards respectively i.e. landing on the bottom of a ladder moves the player to the top, and vice versa for a snake.

There are many win conditions that can be applied to the game; the two most common conditions will be considered in this investigation. The aim of snakes and ladders is to reach the final square on the board, which is square 100 on a standard 10 x 10 board. If the player is expected to reach *at least* square 100, i.e. if they overshoot they are still considered to have won, then the winning condition is called "post". If the player is required to reach *exactly* square 100, and when they overshoot, they "bounce back" and move backwards by how many extra squares they rolled, then the win condition is called "bounce". The standard snakes and ladders board considered in this investigation is given in Figure 1.

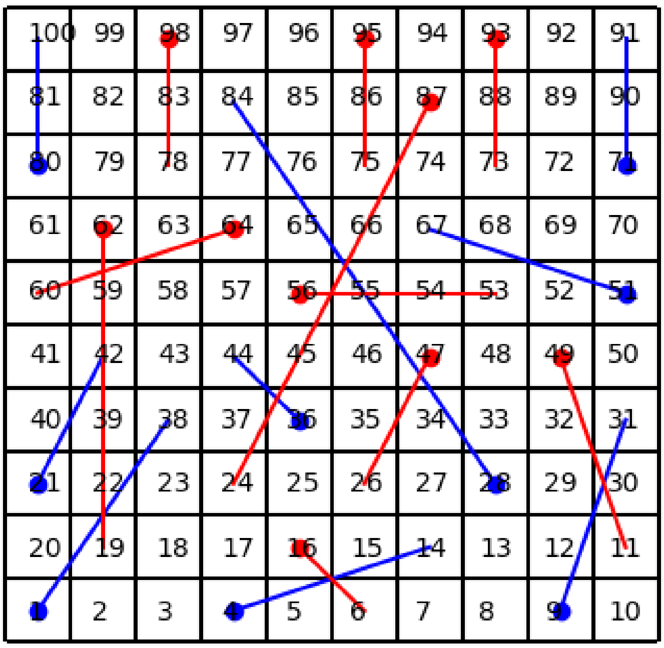

Figure 1: Default snakes and ladders board considered in this investigation. Snakes are represented in red, and ladders are in blue. The start of the snake or ladder is indicated by the circle at one end.

2 Methods

The problem was simulated using python to explore the behaviour. Analysis of behaviour is carried out using two methods: Monte Carlo experiments and absorbing Markov chains.

2.1 Monte Carlo

The Monte Carlo method consists of taking repeated random samples of the system, and making a statistical analysis of the results. In this case, a histogram of number of turns taken for the win condition to be satisfied is used to analyse the data.

In order to generate the dice rolls, a random number is generated, validated, and the player position is updated accordingly. Each roll, the game checks whether or not the win condition is satisfied, and if so, adds the number of turns into the data set to be plotted in the histogram. This is repeated O(10,000) times for the results to be statistically significant.

Similarly to how a binomial distribution approximates a Gaussian distribution for large N, where N is number of Bernoulli trials, the underlying probability distribution is approximated by the Monte Carlo method for sufficiently large numbers of iterations.

2.2 Absorbing Markov chains

The absorbing Markov chain method uses transition matrices, which store the probability of landing on each square from the square that is currently occupied. Specifically, the row represents the current square, and each column represents the possible destination squares. So for a 100x100 board with a 6-sided die, it would begin as below:

$$P = \begin{bmatrix} 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & \cdots \\ 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \cdots \\ \vdots & & \ddots & & & & & & & \end{bmatrix} \tag{1}$$

The leading diagonal is all zeroes as the player is assumed to always move forwards, and the first 6 elements after the leading diagonal in each row are 1/6. Once the snakes and ladders are factored in, we have a matrix which describes the behaviour of the board entirely. As a proof of concept, (2) is the transition matrix representing the board and dice rolls for a 6 squared board, with a ladder from square 1 to 4, and a snake from square 5 to 2, using post win condition:

$$\begin{bmatrix} 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{2}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{3}{6} \\ 0 & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{4}{6} \\ 0 & 0 & \frac{1}{6} & 0 & 0 & 0 & \frac{5}{6} \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

A Markov process is one whose probabilities are independent of previous outcomes, that is, the process can be thought of as "memoryless" as it does not "remember" the previous states that the system was in. [2] This translates well to the analysis of snakes and ladders as each turn is completely independent of the last. For a 10x10 board, the transition matrix has dimensions 101x101 to include all 100 spaces plus space zero. The transition matrix, $\mathbf{P}$, has a canonical form, in which the matrix is made up of a set of four sub-matrices.

$$P = \begin{pmatrix} Q & R \\ \mathbf{0} & I_r \end{pmatrix} \tag{3}$$

In the transition matrix, there are t transient states, and r absorbing states. Absorbing states are those which, once entered, the system cannot leave, and transient states are non-absorbing, and the system will reach an absorbing state in a finite number of steps from a transient state. The sub-matrix $\mathbf{I_r}$ is an identity matrix of order r x r and this represents the absorbing states which cannot be left once entered. $\mathbf{Q}$ is a t x t matrix of the transient states and their respective probabilities of transitioning to another transient state. $\mathbf{R}$ is a t x r matrix of probabilities of the system moving from a transient state to an absorbing state. There is an r x t zero-matrix in the bottom left corner of the matrix as this corresponds with the zero probability of leaving the absorbing states. [2] The fundamental matrix, denoted $\mathbf{N}$, can be evaluated as a sum of $\mathbf{Q^k}$ where $\mathbf{Q}$ is the aforementioned sub-matrix of **P**, which can be represented in a closed-form, analogous to that of the standard geometric series $\frac{1}{1-x}$. [3]

$$N = \sum_{k=1}^{\infty} Q^k = (I_t - Q)^{-1} \tag{4}$$

From the fundamental matrix, many properties of the system can be extracted. The mean number of times that the system is in a given transient state, which in this case is the mean number of times the player lands on a given square of the board, is given by the elements of the fundamental matrix **N**. [2]

2.3 Shannon entropy

The Shannon entropy, or statistical entropy, is a measure of the "randomness" or "unpredictability" of the system. It is evaluated using the following formula:

$$H = \sum_{i=0}^{\infty} p_i \log(p_i), \tag{5}$$

where i labels the non-zero-probability states of the system, p_i is the corresponding probability for state i, and H is the Shannon entropy. [4] Using the natural logarithm in this calculation gives the entropy in natural information units, or nats, and this corresponds to the information contained in an event when its probability is $\frac{1}{e}$. In the Markov chain analysis, p_i is the i^th element in the transition matrix. Applying this method to all non-zero elements of the transition matrix, P, raised to the power of the number of turns, n, allows the entropy to be calculated at all possible numbers of turns in the game. From this, we can find the number of turns taken that maximises the entropy, i.e. the number of turns at which the player's position is least certain.

For the "post" win-condition, we expect the entropy to be very low at the beginning of the game as there are very few positions the player could find themselves in, and as the game progresses, there are an increasing number of positions the player could occupy, so the entropy reaches a maximum, before decreasing again at larger n as we become more certain that the player will have reached the winning space. For the "bounce" win condition, however, we would not expect the entropy to fall off as turns increases. As the player approaches the end of the board, they move backwards due to the bounce condition, as well as moving forward from the dice roll, and hence the entropy remains high as turns increases.

3 Results

3.1 Monte Carlo

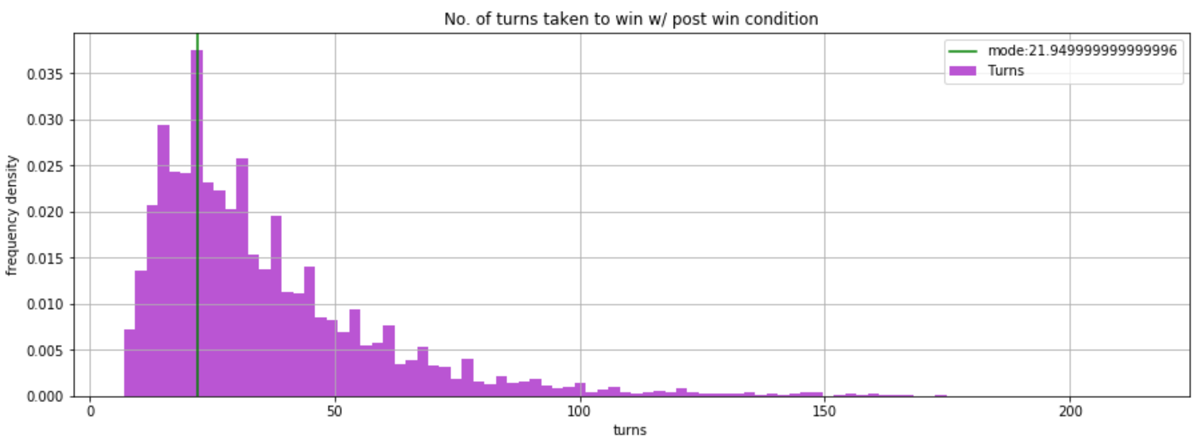

Figure 2: Histogram showing the frequency density of number of turns taken to finish the game with "post" win condition for 10,000 iterations.

By presenting the data from the Monte Carlo sampling in histograms, we see that the average number of turns taken to win the game is lower in "post" games than in "bounce" games. Both

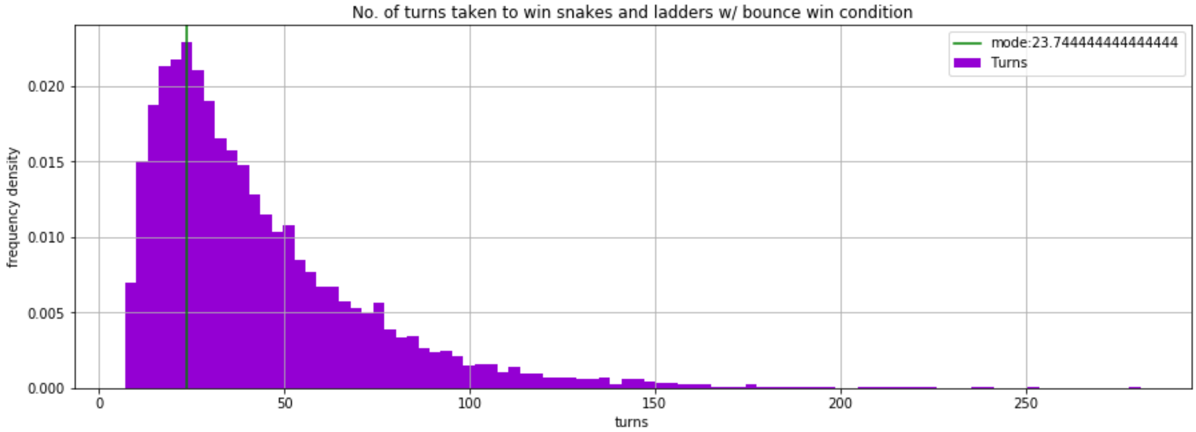

Figure 3: Histogram showing the frequency density of number of turns taken to finish the game with "bounce" win condition for 10,000 iterations.

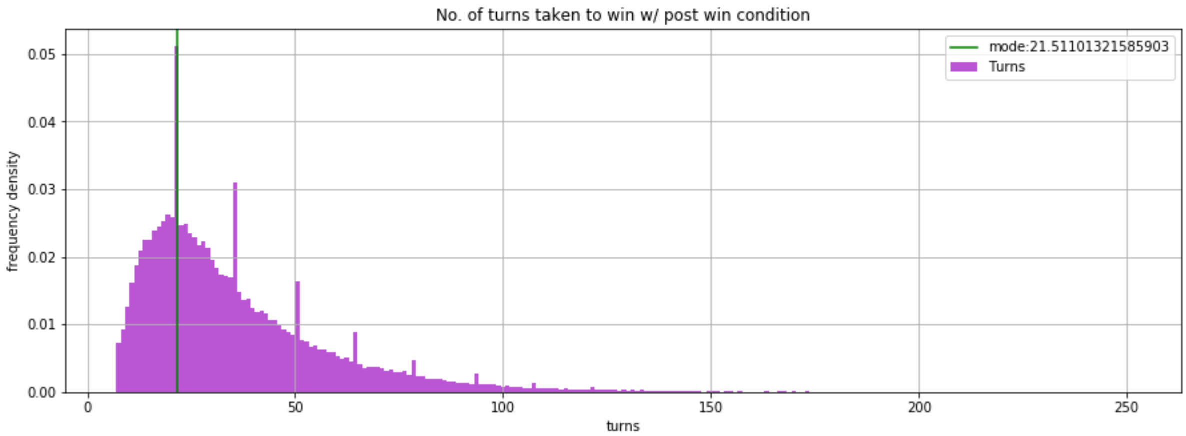

Figure 4: Histogram showing the frequency density of number of turns taken to finish the game with "post" win condition for 100,000 iterations.

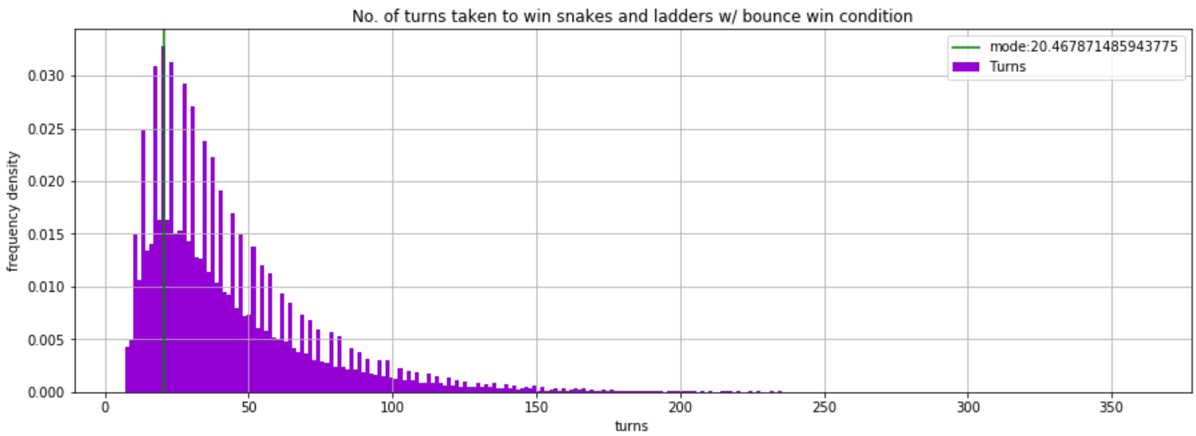

Figure 5: Histogram showing the frequency density of number of turns taken to finish the game with "bounce" win condition for 100,000 iterations.

histograms form a right-skewed Gaussian distribution, however, in the "post" condition histogram, the distribution has notable spikes. These are present regardless of whether the snakes and ladders are included, so they cannot be due to the snakes and ladders. Instead, there are more ways to combine the dice rolls to add to 100 in these numbers of turns corresponding to the spikes. In the "bounce" condition histogram, these spikes are still present, however the bouncing back adds entropy to the system, and this smooths them out into the continuum, and they are only notable at $\approx 100,000$ iterations of the game. The "bounce" game duration is also more right-skewed than the "post" game, and this is due to the fact that the bounce condition adds a chance of replaying a section of the board, thus adding more turns to the duration.

3.2 Shannon Entropy

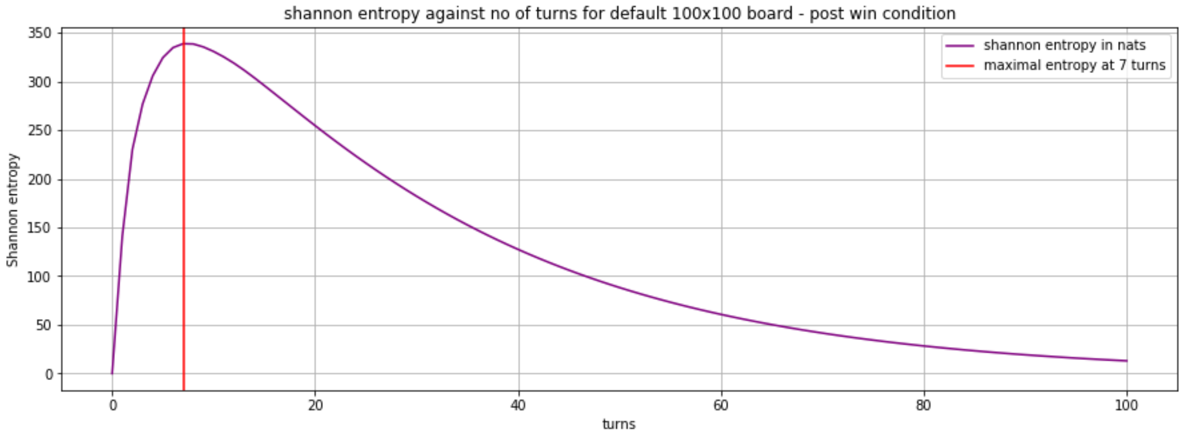

Figure 6: Shannon entropy in natural information units against number of turns for the post win condition.

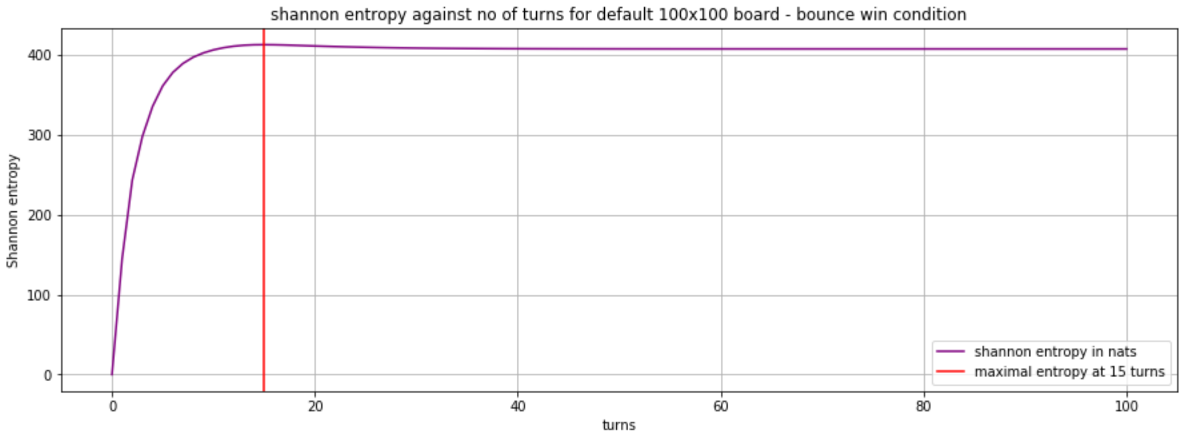

Figure 7: Shannon entropy in natural information units against number of turns for the bounce win condition.

In the "post" win condition game, the entropy is shown in Figure 6 to be very low initially as the player has very limited spaces that they can occupy. As the player takes more turns, the number of spaces they could occupy increases and this leads to a steep increase in entropy. The player's position is most unpredictable for this board after 7 turns, which is the point at which the entropy is maximised. After this point, as the player continues to take more turns, the entropy decreases as we become more certain that they will have crossed the finishing space. This behaviour is in line with the predictions made previously.

For the "bounce" win condition game, the entropy in Figure 7 is seen to begin similarly to the "post" condition game as there are initially very few spaces the player can occupy. It steeply increases as the player takes more turns and the number of potential spaces they can visit increases. However, after this peak, the behaviour diverges from the "post" condition game. The "bounce" game's entropy plateaus, because as turns increase, we cannot know whether the player has landed on space 100 or whether they have bounced back and replayed some area of the board. As there are also snakes and ladders within the region accessible by bouncing back ie within 6 spaces of 100, the player's position after bouncing back becomes even less predictable as they could feasibly bounce to a snake, and be sent further down to replay a larger area of the board.

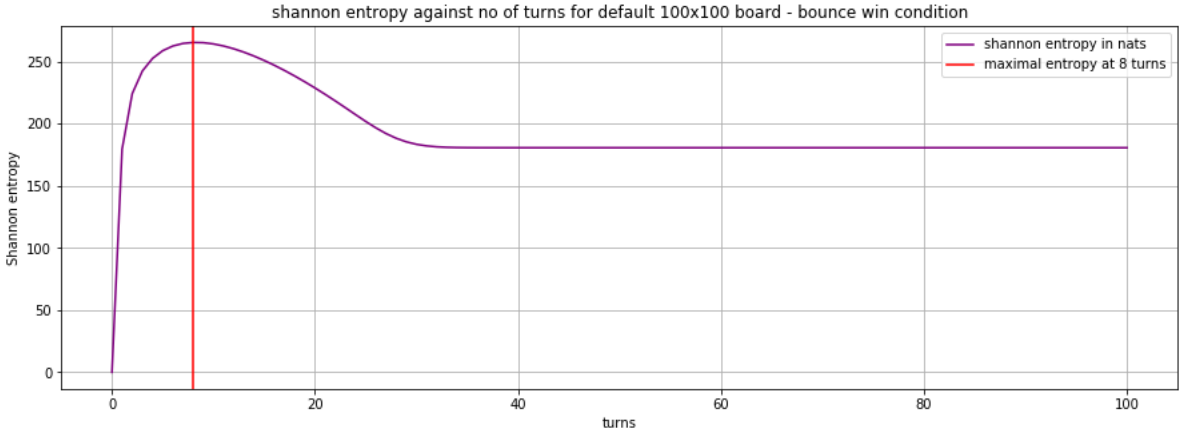

Figure 8: Shannon entropy in natural information units against number of turns for the bounce win condition with no snakes and ladders on the board.

If snakes did not exist within the final 6 squares of the board, then the value at which the entropy plateaus would be lower, as seen in Figure 8, as there is still no certainty that the player has won the game, but the player is limited to only bouncing around the last 6 spaces, and the position is therefore more predictable.

3.3 Heat Maps

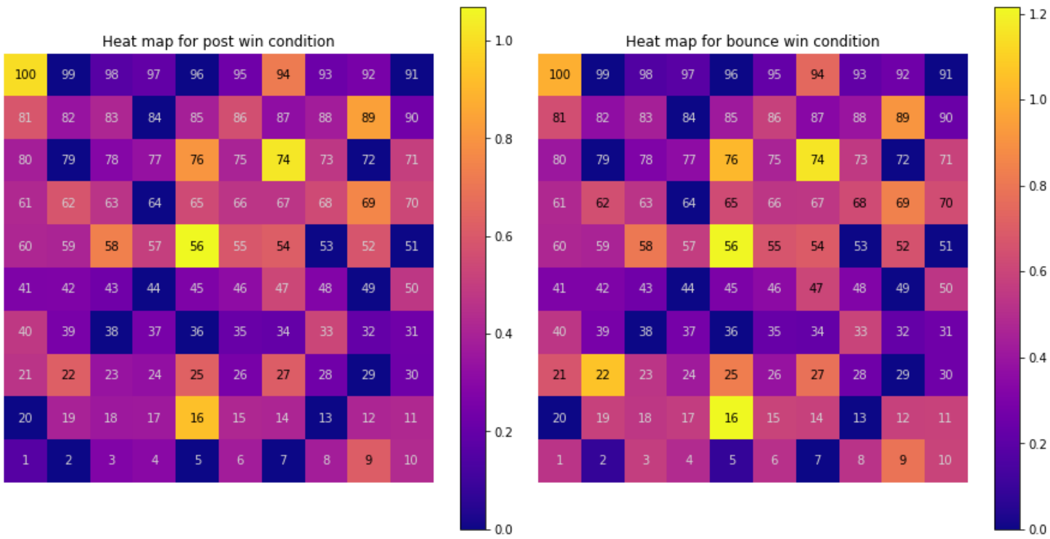

Figure 9: Heat maps representing mean number of times the player lands on each square in one game. Yellow representing more visits, and purple representing fewer.

Using the fundamental matrix extracted from the transition matrix, we can map the expected number of times to visit each square in one game. Figure 9 presents this in a heat map for the "post" win condition, and upon comparison with the board in Figure 1, we clearly see that the likelihood of landing on a square immediately after the end of a snake or ladder is notably higher than other surrounding spaces. Additionally, the likelihood of landing on the initial 5 squares is distinctly lower than most squares, as there is no way to return to them once the player has passed square 6. These observations align with the results that we expect.

In Figure 9, the heat map for the "bounce" win condition in the same board, and the expected number of times to land on most spaces is notably higher than in the "post" win condition, as the "bounce back" gives the player another opportunity to land on more spaces.

Interestingly, there is a region present in both heat maps where the squares are landed on much less frequently. Squares 28 to 49 are almost all notably lower than the rest of the board, and on comparing this with the board itself, it appears to be due to the higher density of snakes and ladders that cross this area, so these spaces are more likely to be skipped, and therefore are less frequently visited. There is a similar effect at the top of the board between spaces 91 to 99, and even in the bounce win condition, the snakes and ladders across this area make it more likely to be skipped over.

4 Further exploration

This investigation could be expanded to consider other variations on the winning condition. For example, we could consider a board where the spaces act as if joined circularly, i.e. if you overshoot past 100, you go back to square 1 and continue from there. This would presumably increase entropy significantly, like in the "bounce" win condition, and it would be interesting to analyse how it affects the likelihood of visiting each space, and the duration of the game. The winning space could also be moved, for example, to halfway up the board, provided that there are snakes above the win space, and ladders below the win space to allow the player to traverse from one side to the other. Winning would require the player to land exactly on the win space, and this would likely increase the duration of games dramatically.

Other variables could also be considered, for example, by varying the number of sides on the die and comparing the average duration of the game, the optimal number of dice sides to minimise duration could be determined. Additionally, we could explore how the number of dice used affects the game duration. The rules regarding the multiple dice could also be varied, e.g. take the sum of both dice rolls, or choose highest dice roll.

Another way to expand the investigation is to put constraints on the snakes and ladders to determine the magnitude of their effect on the duration. These constraints could be, for example, all snakes and ladders have fixed length, all snakes and ladders must cross the middle of the board (i.e. go from top half to bottom half or vice versa), or varying the ratio of snakes to ladders. Analysing randomly generated boards and averaging the behaviours would provide insight into how the snakes and ladders' placements affect the outcomes of the game, and what behaviours are common to all snakes and ladders board, regardless of snake and ladder placement. Adjusting the parameters of the board e.g. number of spaces, number of snakes/ladders, would also make for an interesting extension to the investigation as the relation between duration and board size could be determined, and whether it scales linearly with board size as we might expect.

References

[1] Anonymous, *The history of snakes & ladders*, `https://www.theforgottentoyshop.co.uk/blogs/news/15235037-the-history-of-snakes-ladders`, (accessed: 10/12/2022).

[2] J. G. Kemeny and J. L. Snell, *Finite markov chains* (Van Nostrand, Princeton, NJ, 1976), pp. 24–26, 44–47.

[3] J. K. Grewal, M. Krzywinski, and N. Altman, "Markov models - markov chains", Nature Methods **16**, 663–664 (2019), (accessed: 10/12/2022).

[4] R. Pathria and P. D. Beale, *Statistical mechanics*, 3rd ed. (Boston: Academic Press, 2011), pp. 51–52.