The Fourth Tower of Hanoi

James Marks - U2209705

19/02/2025

Supervisor: Dr. Andrew Booker
Year 3 (20cp) Individual Project

Abstract

The Tower of Hanoi is a classic problem that highlights the wonder found uniquely in mathematics. It is a puzzle with a brilliantly simple set of rules, simple enough to be a common children's toy, and yet is able to generate questions that mathematicians still do not have all the answers to. While the three-peg version of the puzzle is solved and well documented, finding optimal solutions to problems with more pegs is still an area of active research. In this paper we investigate the generalisation of the puzzle to more pegs, focusing on fewest-move solutions. We examine the Frame-Stewart conjecture, and analyse Thierry Bousch's 2014 proof of its optimality.

Contents

Acknowledgement of Sources

For all ideas taken from other sources (books, articles, internet), the source of the ideas is mentioned in the main text and fully referenced at the end of the report. All material which is quoted essentially word-for-word from other sources is given in quotation marks and referenced. Pictures and diagrams copied from the internet or other sources are labelled with a reference to the web page or book, article etc.

Signed James Marks

Date 19/02/2025

1 Introduction

The creation of the Tower of Hanoi is accredited to Frenchman Edouard Lucas in 1883 [1]. It was designed as a children's toy, and consists of a wooden base plate with three vertical needles in an equilateral triangle. On one needle there are eight wooden discs of various diameters, increasing from bottom to top. The aim of the game is to move the stack of discs from one needle to another, while only moving one disc at a time and never placing a disc on top of a smaller one. It is well known that $2^n - 1$ moves are required to move n discs.

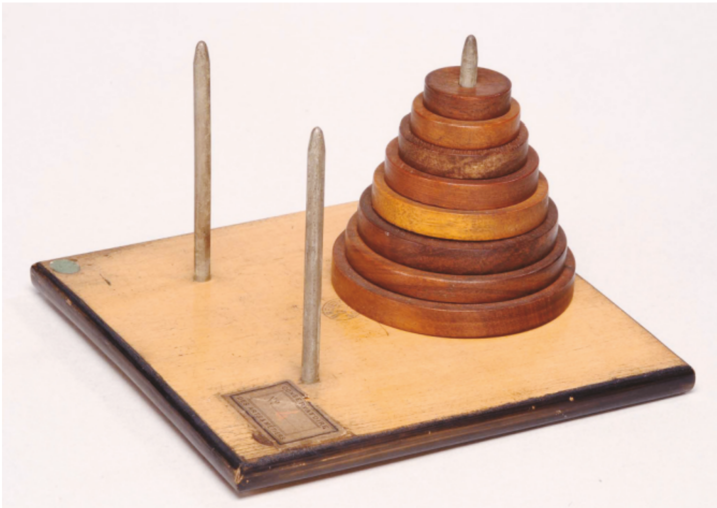

Figure 1: The original Tower of Hanoi [1]

Lucas had no direct connection to Hanoi or Vietnam, and it is thought that he named the puzzle after Hanoi simply because it was in the news at the time. The Battle of Gia Cuc was fought on 27 and 28 March 1883, during the Tonkin Campaign between the French and Vietnamese, and Lucas likely named it accordingly as a marketing ploy to sensationalise his product. The puzzle has since been represented in various stories under various names.

The Tower of Bramah was described in 1883 by Professor N. Claus (de Siam) of the Collège Li-Sou-Stian as "a brass-plate in which are fixed three diamond needles, each a cubit high and as thick as the body of a bee. On one of these needles, at the creation, God placed sixty-four discs of pure gold, the largest disc resting on the brass plate, and the others getting smaller and smaller up to the top one. Day and night unceasingly the priests transfer the discs from one diamond needle to another according to the fixed and immutable *laws of Bramah*, which require that the priest must not move more than one disc at a time and that he must place this disc on a needle so that there is no smaller disc below it. When the sixty-four discs shall have been thus transferred from the needle on which at the creation God placed them to one of the other needles, tower, temple, and Brahmins alike will crumble into dust, and with a thunderclap the world will vanish." With a small amount of calculation, we can conclude that if the priests were to make one movement every second, the time required to solve the puzzle would be about 585 billion years, after which the world and everything in it would cease to exist.

In Dudeney's "Canterbury Puzzles" [2], published in 1907, the "Reve's Puzzle" describes four bar stools, on one of which stood a pile of eight cheese wheels, each smaller than the one below. The rules again state that no cheese may be placed on top of a smaller one, and the aim is to move all cheeses from one stool to another, one wheel at a time. Interestingly, in the solutions section of the book Dudeney included a table of values for the minimum solution for varying numbers of stools and cheeses. He did not provide a proof for these claims, but closer inspection of the table reveals some interesting patterns.

Stools.	Number of Cheeses.						
3	1	2	3	4	5	6	7
4	1	3	6	10	15	21	28
5	1	4	10	20	35	56	84
	Number of Moves.						
3	1	3	7	15	31	63	127
4	1	5	17	49	129	321	769
5	1	7	31	111	351	1023	2815

Figure 2: Dudeney's solutions for the four-peg puzzle [2]

As we will see later in Section 3, the three-peg version of the puzzle has a simple solution for any $N \in \mathbb{N}$ discs, or cheeses in this case, and the table reflects this with the integers 1 to 7 making up the first row. The second row of the table is easily recognised as the triangle numbers, and the third the tetrahedral numbers [3]. The tetrahedral numbers can be generated by taking the sum of all the previous triangle numbers, much like how the triangle numbers are generated by the integers. Alongside the table, Dudeney gave an explanation for how he generated the numbers in the bottom half. "The fourth row contains successive powers of 2, less 1. The next series is found by doubling in turn each number of that series and adding the number that stands above the place where you write the result. The last row is obtained in the same way. This table will at once give solutions for any number of cheeses with three stools, for triangular numbers with four stools, and for pyramidal numbers with five stools." For four stools, the algorithm that Dudeney eloquently explained is as follows [4]:

Let $t_k = \frac{k(k+1)}{2}$ be the kth triangle number, and write $M(t_k)$ for the number of moves required to move t_k cheese wheels with four stools. Then

$$M(t_{k+1}) = 2M(t_k) + 2^{k+1} - 1, M(1) = 1$$

This algorithm will make a more mathematical appearance later in Section 4, but before that we must define what it means to be a Tower of Hanoi puzzle, by stating the Golden Rules and Fundamental Properties of all such puzzles.

2 Golden Rules and Fundamental Properties

2.1 The Golden Rules

The Tower of Hanoi appears in many forms and with many variants. As well as the original 3-peg game, we have the 4-peg Reve's puzzle, linear and cyclic versions where discs can only move to adjacent pegs, and variants with coloured discs that have to be sorted. Regardless, we arrive at the *golden rules* which define each Tower of Hanoi puzzle.

1. Only one disc is permitted to move at a time
2. Only the top disc of any pile can be moved
3. A disc may not be placed on a smaller disc

Lastly, the question we ask of each puzzle: how few moves are required to reach our goal?

2.2 Fundamental Properties

One of the most powerful mathematical tools in investigating Tower of Hanoi puzzles is recursion, which we will explore in detail in Section 3. A common theme for solving Hanoi puzzles is to reduce the problem of moving N discs to a problem of moving $N-1$, or some other number less than N. It would be a far easier task for the monks of Bramah to move one disc from needle to needle than the mammoth task of sixty-four, so for the sake of the world perhaps that is a good thing.

We also find that Hanoi puzzles tend to be very symmetrical. This is hardly surprising because the choice of initial and final pegs is arbitrary, and this property allows us to significantly reduce the computation required.

Other properties commonly observed in Hanoi puzzles include an exponential increase in complexity with the addition of more discs, the presence of a (not always unique) optimal strategy, and the notion that a player should not need to know any previous moves made in a puzzle to determine the optimal strategy.

2.3 Hanoi Graphs

The golden rules play an important role in creating a “good” mathematical puzzle. In this context, a good puzzle is one that has a well-defined set of possible configurations, a well-defined set of rules that allows movement between configurations, a clear objective, and a solvable but non-trivial solution. The rules achieve these properties, in particular creating a finite (for finite N and p) set of possible *game states*, and providing rules for determining whether these game states can be moved between by one move. The structure that we have described here is of course a graph, with the set of game states making the set of vertices and the set of valid moves making the set of edges. We call such a graph a *Hanoi Graph* and for N discs and p pegs, we denote the corresponding graph H_p^N. To avoid requiring a diagram for every game state on a graph, we encode states to an N digit long string as follows:

Definition 2.1. For any given game state on a Hanoi graph H_p^N, we write the game state as a concatenation of the peg that each disc is on, in ascending order of diameter. For example, in Figure 3, the state is `1102011` because the smallest two discs are on peg 1, the next smallest is on peg 0, the next smallest on peg 2, and so on. As we will see later in Section 3, this is sufficient to uniquely encode each state.

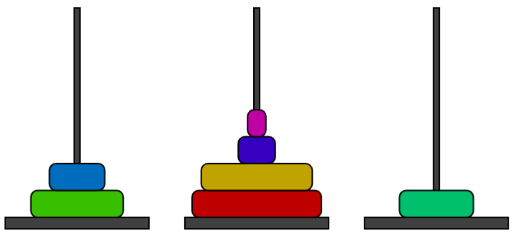

Figure 3: Game state `1102011` on Hanoi Graph H_3^7

This allows us to create graphs to represent the game space of a Hanoi puzzle, and these can be helpful in exploring possible paths between game states. Inspection of the 3-peg Hanoi graphs for 1, 2 and 3 discs reveals many of the properties already discussed. There is a clear recursive structure as

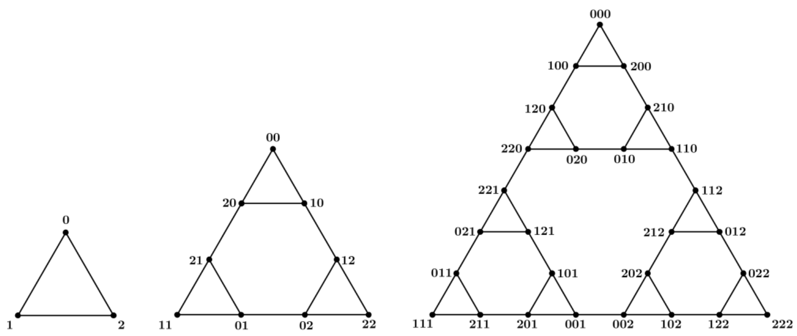

Figure 4: Hanoi Graphs H_3^1, H_3^2 and H_3^3[5]

the number of discs increases, with three copies of H_3^2 being visible in H_3^3, as well as noticeable reflective and rotational symmetries, which translate to various swapping of pegs.

The graphs also provide excellent insight into our next area of focus, the distance function and corresponding optimal path(s).

Definition 2.2. Let p, $N \in \mathbb{N}$ and let u and v be two game states on Hanoi graph H_p^N. Then we write $d(u, v)$ to denote the length of a minimal path between u and v, or equivalently the fewest legal moves required to get from u to v.

We call said minimal path an *optimal path between u and v.*

In Figure 4, it becomes obvious what the optimal path between initial state `000` and target state `222` is, following the graph down the right-hand side of H_3^3 gives us a path of length 7. We also see that the critical points on the graph are `220`, `110`, `112`, `002`, `001`, and `221`. These are the states which have the smallest two discs one one peg, the largest on a different peg, and one peg empty, and are the only possible game states that allow the largest disc to move.

3 The Three-Peg Solution

In this Section we will derive the fewest moves required to move N discs from one peg to another in the 3-peg variant without using Hanoi graphs.

We begin by outlining some notation. We will write $\mathcal{C} = \{0, 1, 2\}$ to represent the set of 3 pegs, and $[N] = \{0, 1, 2, ..., N-1\}$ to represent the set of n discs. Contrary to the definition of game state from before, we write 0 to represent the largest disc and $N-1$ to represent the smallest disc. Next, we define a game state u as a map from the set $[N]$ to the set $\mathcal{C}$. We are able to do this by virtue of the third golden rule: A disc may not be placed on a smaller disc. This means that it is sufficient to map each disc to the peg it is on, and the order of discs on each peg is entirely defined by their diameters.

To solve the puzzle, we first consider the most difficult disc to move, disc 0. In order to move disc 0, we have two requirements:

1. There are no smaller discs on the peg that disc 0 is on.
2. One of the other two pegs is entirely empty.

There are only six possible game states that allow disc 0 to move, and we can use the symmetrical properties of the puzzle to reduce this number to one. We begin our puzzle with all N discs stacked on peg 0. Without loss of generality we can assume that the final state will have all N discs stacked on peg 2. Thus in order to move disc 0 to peg 2, we must have all other discs stacked on peg 1. What we have described here is the first step in a recursive algorithm to solve the puzzle: Move the top $N-1$ discs to peg 1, move disc 0 to peg 2, move the top $N-1$ discs to peg 2. Because the above conditions are required for the completion of the puzzle, we can be sure that this solution is indeed optimal.

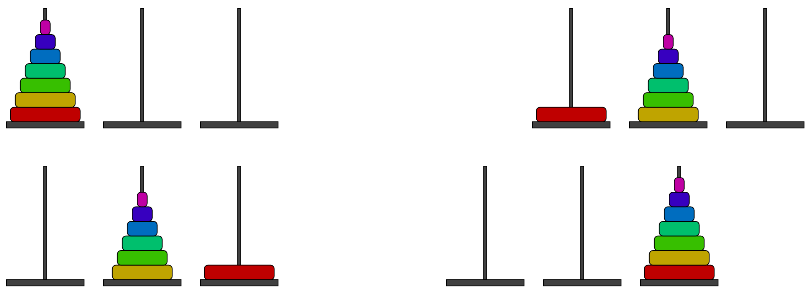

Figure 5: The recursive algorithm to solve the 3-peg Hanoi puzzle

We will use this algorithm to calculate the fewest moves required to solve the 3-peg Hanoi puzzle. Let $M(N)$ be the number of moves required to move N discs from one peg to another on 3 pegs. Then our algorithm gives us the recursive definition

$$M(N) = 2M(N-1) + 1$$

because we have moved discs $0, 1, ..., N-1$ from one peg to another twice, and moved peg 0 once. We also see that trivially $M(1) = 1$. Repeating this a few times gives us

n	1	2	3	4	5	6
$M(n)$	1	3	7	15	31	63

and solving the recurrence relation leaves $M(n) = 2^n - 1$.

4 The Frame-Stewart Conjecture

4.1 The Frame-Stewart Algorithm

We begin this section by exploring the limitations of generalising from the neatly solved 3-peg puzzle to the considerably more complicated 4-peg puzzle, and beyond. The reason that the 3-peg problem has a neat solution is the forcing nature of its setup. In order to move the largest disc, we must have all other discs stacked on one peg. This gives us the recursive structure that we need to define a clean formula for the number of moves required. On the other hand, moving the largest disc from peg 0 to peg 3 in the 4-peg problem can be achieved in 2^{N-1} different game states, corresponding to moving each disc 1 to $N - 1$ to either peg 1 or peg 2. Figure 6 gives an example of this for seven pegs, without claiming that it is an optimal solution.

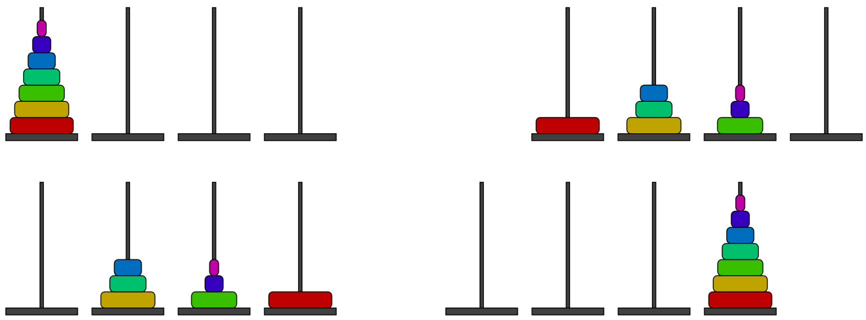

Figure 6: An example of the limitations of generalising to 4 pegs

Using this, we introduce the *Frame-Stewart Algorithm*. Consider a Hanoi problem with N discs and 4 pegs, where all discs are stacked on peg 0 in the initial state u. Then we can solve the puzzle by the following algorithm:

1. Choose $M \in \mathbb{N}$ with $0 \leq M \leq N$.
2. Move the top M discs from peg 0 to peg 1 using all four pegs.
3. Move the bottom $N - M$ discs from peg 0 to peg 2 without using peg 1.
4. Move the M discs on peg 1 to peg 2 using all four pegs.

We will write $\Phi_p(N)$ to denote the number of moves required to move N discs from one peg to another on p pegs. Through this algorithm we find that we have used $2\Phi_4(M) + \Phi_3(N - M)$ moves.

4.2 The Frame-Stewart Conjecture

We are now well-equipped to introduce the most famous result of this paper, the Frame-Stewart Conjecture. The conjecture states that an optimal choice of M will produce an optimal solution to the 4-peg problem, and we now state this concretely.

Theorem 4.1 (Frame-Stewart Conjecture)**.** *Consider a Tower of Hanoi puzzle with N discs and p pegs. Then the fewest moves required to move all N discs from one peg to another, $\Phi_p(N)$, is defined recursively by*

$$\Phi_p(N) = \begin{cases} 2^N - 1, & \text{if } p = 3, \\ min_{1 \leq M < N}\{2\Phi_p(M) + \Phi_{p-1}(N-M)\} & \text{if } p \geq 4. \end{cases}$$

The conjecture was made in 1941, but its optimality is still yet to be proven for general p pegs. American mathematician Donald Knuth, called "father of the analysis of algorithms", is reported to have written “I doubt if anyone will ever resolve the conjecture; it is truly difficult.” [6] In 2014, French mathematician Thierry Bousch proved the conjecture for $p = 4$ [7], and this is what we will be investigating.

Firstly, we will state our theorem specifically for $p = 4$, using our result from Section 3, $\Phi_3(N) = 2^N - 1$.

Theorem 4.2. *Consider a Tower of Hanoi puzzle with N discs and 4 pegs. Then the fewest moves required to move all n discs from one peg to another, $\Phi_4(N)$, is defined recursively by*

$$\Phi_4(N) = min_{1 \leq M < N}\{2\Phi_4(M) + 2^{N-M} - 1\}$$

For the remainder of the paper, we will write $\Phi(N) = \Phi_4(N)$ where there is no risk of confusion.

We will also formalise Dudeney's claim from Section 1.

Theorem 4.3. *Let $N \in \mathbb{N}$. Then*

$$\Phi(N) = 2^{\nabla 0} + 2^{\nabla 1} + ... + 2^{\nabla (N-1)}$$

satisfies Theorem 4.2, where ∇n denotes the “triangular root” of n, i.e. the largest $p \in \mathbb{N}$ that satisfies $\frac{p(p+1)}{2} \leq n$.

Proof. Let n, $p \in \mathbb{N}$ such that $p \leq n + 1$ and $N = \Delta n + p$. Suppose $M = N - n - 1$. Then we have

$$\begin{aligned} 2\Phi(M) + 2^{N-M} - 1 &= 2\Phi(M) + 2^{n+1} - 1 \\ &= 2\Phi(M) + 2^0 + 2^1 + ... + 2^n \\ &= 2^{\nabla 0+1} + 2^{\nabla 1+1} + 2^{\nabla 2+1} + ... + 2^{\nabla (M-1)+1} + 2^0 + 2^1 + ... + 2^n \\ &= 2^1 + 2^2 + 2^2 + 2^3 + 2^3 + 2^3 + ... + (p-2)2^n + 2^0 + 2^1 + ... + 2^n \\ &= 2^0 + 2^1 + 2^1 + 2^2 + 2^2 + 2^2 + ... + (p-1)2^n \\ &= 2^{\nabla 0} + ... + 2^{\nabla (\Delta n + p - 1)} \\ &= \Phi(N) \end{aligned}$$

□

Now we define a function $\Psi : \mathcal{P}_{\text{finite}}(\mathbb{N}) \to \mathbb{N}$, where $\mathcal{P}_{\text{finite}}$ denotes the set of all finite subsets of $\mathbb{N}$. Let E be a finite subset of $\mathbb{N}$. For any $L \in \mathbb{N}$, let

$$\Psi_L(E) = (1-L)2^L - 1 + \sum_{n \in E} 2^{\min(\nabla n, L)}$$

and define

$$\Psi(E) = \sup_{L \in \mathbb{N}} \Psi_L(E).$$

Because $|E|$ is finite, we must have $\Psi_L(E) \to -\infty$ for $L \to \infty$. This means that the supremum is not ∞, and thus $\Psi(E)$ is an integer. We note that $\Psi(\emptyset) = 0$ and both functions Ψ_L and Ψ are clearly increasing under inclusion.

5 Bousch's Proof for Four Pegs

Similar to Section 2, the solution to the 4-peg puzzle reduces game states to problems involving fewer discs. To formalise this notion, we introduce some new notation.

Definition 5.1. Consider a Hanoi puzzle with N discs and 4 pegs. Let $[N] = \{0, 1, 2, ..., N-1\}$ be the set of discs and $\mathcal{C} = \{0, 1, 2, 3\}$ be the set of pegs. Let $u : [N] \to \mathcal{C}$ be a game state. We write u' to represent the restriction of u to $[N-1]$. That is, we ignore the largest disc. This is useful because larger discs can never inhibit a smaller disc from moving, so we can effectively treat the game state the same.

We now have the notation required to begin to understand the main result of Thierry Bousch's paper, "La Quatrième Tour de Hanoï" [7]. It is worthwhile noting that Bousch refers to the pegs as columns in his proof, and as such we have adopted the same notation.

Theorem 5.2 ([7] Theorem 2.9)**.** *Write $\mathcal{C} = \{0, 1, 2, 3\}$ for the set of pegs. Let $N \in \mathbb{N}$ be the number of discs and $u, v : [N] \to \mathcal{C}$, both game states. Suppose that in v the first two columns are empty, that is, $v[N] \subseteq \{2, 3\}$. Then*

$$d(u, v) \geq \Psi\{k \in [N] : u(k) = 0\}. \tag{1}$$

Before we begin the proof, it is important to have an intuitive idea of what the theorem is stating. In Section 2.3 we looked at the critical points on H_3^3, noting that they had the smallest two discs on one peg, the largest disc on a different peg, and one peg empty. We can think of Ψ in a similar manner, being the distance function from a game state to the nearest critical point. This is not rigorous, but helpful to think about when going through the proof.

Proof. Write $E = \{k \in [N] : u(k) = 0\}$. This represents the first column in position u. As defined previously, u' and v' represent restrictions of u and

v to $[N-1]$. We will use this to perform induction on N. Assume that our Theorem holds for $N-1$ discs and board positions $u', v' : [N-1] \to \mathcal{C}$. Clearly $v'[N-1] \subseteq \{2, 3\}$. Then we have

$$\begin{aligned} d(u,v) \geq d(u',v') &\geq \psi\{k \in [N-1] : u'(k) = 0\} \\ &= \psi\{k \in [N-1] : u(k) = 0\} \text{ (by inductive hypothesis)} \\ &= \psi\{E - \{N-1\}\} \end{aligned}$$

We have that $\Psi(\emptyset) = 0 \leq d(u,v)$ so the inequality holds.

If E doesn't contain $N-1$, we have $d(u,v) \geq \psi E$ as required. Thus we can assume for the rest of the proof that E contains $N-1$, i.e. the largest disc in u is in column 0. Clearly in v the largest disc is in either column 2 or column 3, and without loss of generality, we say that it is in column 2. Thus we have $u(N-1) = 0$ and $v(N-1) = 2$.

Write $D = d(u,v)$ for the distance between u and v, and let $\gamma : [D+1] \to \mathcal{C}^{[N]}$ be an optimal path going from u to v in $\mathcal{C}^{[N]}$. That is, $\gamma(0) = u$, $\gamma(D) = v$, and $d(\gamma(i), \gamma(j)) = |i-j|$ for all i, $j \in [D+1]$. Here it is important to note that each $\gamma(i)$ is a game state, which in itself is a map, so we write $\gamma(i) = \gamma_i$. Now we write E' for the set of discs which start in column 0 and at some point on γ pass through column 3, and we see that $E' = \{k \in E : \exists t \in [D+1], \gamma_t(k) = 3\}$. With reference to the Frame-Stewart algorithm in Section 4, this translates to the set of M pegs that are moved to an auxiliary peg in the first step.

We now introduce a few lemmas which will be useful later on.

Lemma 5.3. *Let n, $p \in \mathbb{N}$ such that $p \leq n+1$. Then we have*

$$\Phi(\Delta n + p) = 1 + (n + p - 1)2^n.$$

Proof. Let $N = \Delta n + p$. We have that

$$\begin{aligned} \Phi(N) &= 2^{\nabla 0} + 2^{\nabla 1} + ... + 2^{\nabla(N-1)} \\ &= (2^{\nabla 0}) + (2^{\nabla 1} + 2^{\nabla 2}) + (2^{\nabla 3} + 2^{\nabla 4} + 2^{\nabla 5}) + ... \\ &\quad + (2^{\nabla(\Delta(n-1))} + 2^{\nabla(\Delta(n-1)+1)} + ... + 2^{\nabla(\Delta(n-1)+n-1)}) \\ &\quad + (2^{\nabla(\Delta(n))} + 2^{\nabla(\Delta(n)+1)} + ... + 2^{\nabla(\Delta(n)+p-1)}) \end{aligned}$$

$$\begin{aligned} = 2^0 + 2^1 + 2^2 + 2^3 + ... + 2^{n-1} &+ 2^0 \\ +2^1 + 2^2 + 2^3 + ... + 2^{n-1} &+ 2^1 \\ +2^2 + 2^3 + ... + 2^{n-1} &+ 2^2 \\ +2^3 + ... + 2^{n-1} &+ 2^3 \\ +... + 2^{n-1} &+ 2^4 \\ ... & \\ +2^{n-1} &+ 2^{n-1} \\ +p2^n - (2^0 + 2^1 + 2^2 + 2^3 + ... &+ 2^{n-1}) \end{aligned}$$

$$
\begin{aligned}
&= n2^n + p2^n - (2^n - 1)\\
&= 1 + (n + p - 1)2^n
\end{aligned}
$$

□

This lemma demonstrates nicely the importance that triangle numbers play in solving the puzzle.

Lemma 5.4. *For all $n \in \mathbb{N}$,*

$$\Psi[n] = \frac{\Phi(n+1) - 1}{2} = \frac{1}{2}(2^{\nabla 1} + 2^{\nabla 2} + ... + 2^{\nabla n})$$

Proof. Trivially, we have that $\Psi[0] = 0 = \frac{\Phi(1)-1}{2}$. Suppose therefore that $n \geq 1$ and write $n = \Delta m + p$, with $m = \nabla n$, $m \geq 1$ and $0 \leq p \leq m$. By Lemma 5.3, $\Phi(\Delta m) = 1 + (m-1)2^m$, and $\Phi(n+1) = 1 + (m+p)2^m$.
To find $\Psi[n] = \sup_{L \in \mathbb{N}} \Psi_L[n]$, we compare $\Psi_{L+1}[n]$ and $\Psi_L[n]$, with the maximum value being achieved when the difference turns positive.

For $L \in \mathbb{N}_0$, we have

$$
\begin{aligned}
\Psi_{L+1}[n] - \Psi_L[n] &= -(L+1)2^L + \sum_{k \in [n]} 2^{\min(\nabla k, L+1)} - 2^{\min(\nabla k, L)}\\
&= 2^L(\#\{k \in [n] : k \geq \Delta(L+1)\} - (L+1))\\
&= 2^L((n - \Delta(L+1)) - (L+1))
\end{aligned}
$$

This expression is strictly positive if and only if $n \geq \Delta(L+1) + L + 2 = \Delta(L+2)$, or equivalently $\nabla n \geq L+2$, or $L < m-1$. Thus,

$$
\begin{aligned}
\Psi[n] = \Psi_{m-1}[n] &= (2-m)2^{m-1} - 1 + \sum_{0 \leq k < \Delta m} 2^{\nabla k} + \sum_{\Delta m \leq k < n} 2^{m-1}\\
&= (2-m)2^{m-1} - 1 + \Phi(\Delta m) + (n - \Delta m)2^{m-1}\\
&= (2-m)2^{m-1} + (m-1)2^m + p2^{m-1}\\
&= (m+p)2^{m-1}\\
&= \frac{\Phi(n+1) - 1}{2}
\end{aligned}
$$

□

The connection between Ψ and Φ is of note. We can rearrange the first equality to get

$$\Phi(N) = 2\Psi[N-1] + 1$$

which looks remarkably similar to the Frame-Stewart algorithm for four pegs. Intuitively, we can picture $\Psi[n]$ as being the number of moves required to make it to the "halfway point" of the puzzle, i.e. moving the largest disc from its starting peg to its final destination. It is clear that in an optimal solution the largest disc should never move twice, and this is what gives us this relationship.

Lemma 5.5. *We have $\Psi[n+2] \geq 2^{(\nabla n)+1}$ for all $n \in \mathbb{N}$.*

Proof. Let $s = \nabla n$. Because $\Psi[\cdot]$ is increasing and $\Delta s \leq n$, it is enough to show that $\Psi[\Delta s + 2] \geq 2^{s+1}$. This inequality is satisfied (with equality) for $s = 0, 1$; $\Psi[2] = 2, \Psi[3] = 4$. For $s \geq 2$, Lemmas 5.4 and 5.3 give us $\Psi[\Delta s + 2] = (s+2)2^{s-1}$, which is greater than or equal to 2^{s+1} since $s + 2 \geq 4$. $\square$

Lemma 5.6. *For all finite subsets E of $\mathbb{N}$, we have*

$$n \leq \Psi[n] \leq \Psi(E) \leq 2^n - 1$$

where $n = |E|$.

Proof. First inequality: We have that $\Psi[n] \geq \Psi_0[n] = n$.

Second inequality: Let $L \in \mathbb{N}$. We have

$$\begin{aligned}
\Psi_L(E) &= (1-L)2^L - 1 + \sum_{k \in E} 2^{\min(\nabla k, L)} \\
&\geq (1-L)2^L - 1 + \sum_{k=0}^{n} 2^{\min(\nabla k, L)} \\
&= \Psi_L([n])
\end{aligned}$$

Since this holds for all L, it must hold for the supremum.
Third inequality: Let $L \in \mathbb{N}$. Again we have

$$\begin{aligned}
\Psi_L(E) &= (1-L)2^L - 1 + \sum_{k \in E} 2^{\min(\nabla k, L)} \\
&\leq (1-L)2^L - 1 + \sum_{k \in E} 2^L \\
&= (1+n-L)2^L - 1
\end{aligned}$$

Lastly we note that

$$\begin{aligned}
2^{n-L} &\geq (n-L) + 1 \\
2^n &\geq (1+n-L)2^L \\
&\geq (1+n-L)2^L - 1
\end{aligned}$$

which completes the proof. $\square$

Proof of Theorem 5.2 continued. Consider the case where E' is empty. Then all discs finish their path in column 2. Since column 3 is not used, they must take a least $2^{|E|} - 1$ moves, so $D \geq 2^{|E|} - 1$. We know that $\Psi E \leq 2^{|E|} - 1$

so the inequality $D \geq \Psi E$ is clearly satisfied in this case. Therefore we can assume that E' is non-empty, and write $T = \max(E)$ as the largest element. We write $E'' = \{k \in E : k > T\}$. Let $K = |E''|$ and $b_1 < b_2 < ... < b_K$ the elements of E''. We see that

$$E \subseteq [T] \sqcup T \sqcup \{b_1, b_2, ..., b_K\} \subseteq [N]$$

and in particular

$$T + K + 1 \leq N.$$

Let t_0 be the first state in our optimal path that disc T is not in column 0, that is the smallest element of $[D+1]$ such that $\gamma_{t_0}(T) \neq 0$. Because t_0 is minimal, we must have $\gamma_{t_0-1}(T) = 0$, so clearly disc T is the disc that moves between $t_0 - 1$ and t_0. Thus column 0 must be clear of any discs smaller than T at γ_{t_0-1}.

Write $x_0 = \gamma_{t_0-1}$, the game state just before disc T moves from column 0 for the first time. Here column 0 contains disc T and no smaller discs, and one other column also contains no discs smaller than T.

Let t_1 be the first time that T is in column 3, that is the smallest element of $[D+1]$ such that $\gamma_{t_1}(T) = 3$, and write $x_3 = \gamma_{t_1}$ the corresponding game state. By a similar argument, in x_3, column 3 contains T and no smaller discs, and there is one other column with no discs smaller than T. Clearly we have $1 \leq t_0 \leq t_1 \leq D$.

Let t_2 be the first time $N-1$ is not in column 0 and $z_0 = \gamma_{t_2-1}$. Here column 0 contains only disc $N-1$ and one other column is completely empty. Lastly, let t_3 be the last time $N-1$ is not in column 2 and write $z_2 = \gamma_{t_3+1}$. Here column 2 only contains disc $N-1$ and one other column is completely empty. Like before we get the inequality $1 \leq t_2 \leq t_3 + 1 \leq D$.

Continuing previous notation we write x'_a and z'_b to represent restrictions of x_a and z_b to $[N-1]$ and u'', v'', x''_a and z''_b to represent the corresponding restrictions to $[T]$. We observe that in z'_0 column 0 is empty as well as one other, which allows us to use similar inductive reasoning to that used previously.

$$\begin{aligned} d(u, z_0) \geq d(u', z'_0) &\geq \Psi\{k \in [N-1] : u(k) = 0\} \\ &= \Psi(E - \{N-1\}) \end{aligned}$$

Similarly

$$\begin{aligned} d(u, x_0) \geq d(u'', x''_0) &\geq \Psi\{k \in [T] : u(k) = 0\} \\ &= \Psi(E \cap [T]) \end{aligned} \tag{2}$$

Interestingly, when $T = N-1$, x_0 and z_0 are the same game state and the two inequalities are the same.

Before we move into the depths of the proof, it might be helpful to recap some of the notation seen so far.

5.1 Recap of notation

u: The initial game state
v: The final game state, with pegs 0 and 1 empty
$D = d(u, v)$: The fewest moves required to go from state u to state v
γ_t: The game state at time $t \in [D + 1]$
E: The set of discs initially in column 0
E': The set of discs that start in column 0 and pass through column 3
T: The largest disc in set E'
E'': Discs larger than T that start in column 0 but do not pass through column 3
$K = |E''|$: The number of discs larger than T that start in column 0 but do not pass through column 3
$b_1, ..., b_K$: The elements in E''
x': State x, ignoring disc $N - 1$
x'': State x, ignoring all discs larger than or equal to T
t_0: The first time that disc T is not in column 0
t_1: The first time that disc T is in column 3
t_2: The first time that disc $N - 1$ is not in column 0
t_3: The last time that disc $N - 1$ is not in column 2
$x_0 = \gamma_{t_0-1}$: The game state just before disc T first leaves column 0
$x_3 = \gamma_{t_1}$: The game state when disc T first arrives in column 3
$z_0 = \gamma_{t_2-1}$: The game state just before disc $N - 1$ first leaves column 0
$z_2 = \gamma_{t_3+1}$: The game state just after disc $N - 1$ arrives in column 2 for the last time

Proof of Theorem 5.2 continued. In our endeavour to prove that $D \geq \Psi E$, we find it useful to split into two cases: $\Delta K > T$ and $\Delta K \leq T$. Before that, we introduce two more lemmas.

Lemma 5.7. *Let A, B be two finite subsets of $\mathbb{N}$. Then we have*

$$\Psi(A) - \Psi(B) \leq \sum_{k \in A-B} 2^{\nabla k}$$

Proof. Let $L \in \mathbb{N}$ be the value which achieves $\Psi(A) = \Psi_L(A)$. Then we must have $\Psi_L(B) \leq \Psi(B)$, and clearly $\Psi_L(B) \geq \Psi_L(A \cap B)$. Thus

$$\begin{aligned}
\Psi(A) - \Psi(B) &\leq \Psi_L(A) - \Psi_L(B) \\
&\leq \Psi_L(A) - \Psi_L(A \cap B) \\
&= \sum_{k \in A-B} 2^{\min(\nabla k, L)} \\
&\leq \sum_{k \in A-B} 2^{\nabla k}
\end{aligned}$$

□

Lemma 5.8. *Let A be a finite subset of $\mathbb{N}$, and s an integer such that the set $A - [\Delta s]$ contains at most s elements. Then*

$$\Psi(A) - \Psi(A - \{a\}) \leq 2^{s-1}$$

for all $a \in A$.

Proof. The proof is trivial for $A = \emptyset$, so we assume $A \neq \emptyset$, and so $s \geq 1$. Thus for all $L \geq s-1$,

$$\begin{aligned}
\Psi_{L+1}(A) - \Psi_L(A) &= \left((1-(L+1))2^{L+1} - 1 + \sum_{n\in A} 2^{\min(\nabla n, L+1)}\right) \\
&\quad - \left((1-L)2^L - 1 + \sum_{n\in A} 2^{\min(\nabla n, L)}\right) \\
&= -L2^{L+1} - (1-L)2^L + \sum_{n\in A} 2^{\min(\nabla n, L+1)} - \sum_{n\in A} 2^{\min(\nabla n, L)} \\
&= 2^L(-L-1) + \#\{\min(\nabla n, L+1) > \min(\nabla n, L)\}(2^{L+1} - 2^L) \\
&= 2^L(-L-1) + \#\{n \in A : n \geq \Delta(L+1)\}2^L \\
&\leq 2^L(\#\{n \in A : n \geq \Delta s\} - s) \text{ (by assumption)} \\
&\leq 0
\end{aligned}$$

What we have shown is that Ψ_L is decreasing for $L \geq s-1$. Thus there exists some $L \leq s-1$ such that $\Psi(A) = \Psi_L(A)$, and

$$\begin{aligned}
\Psi(A) - \Psi(A - \{a\}) &\leq \Psi_L(A) - \Psi_L(A - \{a\}) \\
&= 2^{\min(\nabla a, L)} \\
&\leq 2^L \\
&\leq 2^{s-1}
\end{aligned}$$

□

Proof of Theorem 5.2 continued. We now consider the first case.

5.2 Case 1: $\Delta K > T$

This is the simplest case. Firstly, it follows that $K \geq 1$, and $T < b_K = N-1$; the largest disc never passes through column 3. Secondly, the set $\{x \in E : x \geq \Delta K\}$ is contained within $\{x \in E : x > T\} = E''$ and is therefore of size at most K. This gives us

$$\Psi E - \Psi(E - \{N-1\}) \leq 2^{K-1}$$

and thus

$$d(u, z_0) \geq \Psi E - 2^{K-1}.$$

Because $t_2 \leq t_3 + 1$, the path γ passes through the game states u, z_0, z_2, v in the following order. Note that some game states can occur simultaneously.

$$u \to z_0 \to z_2 \to v$$

In z_2, column 2 and one other, which we will call column $c = \gamma_{t_3}(N-1) \in \{0, 1\}$, do not contain any discs smaller than $N-1$. The discs $b_1, ..., b_{K-1}$ are therefore in column $1-c$.

In the final game state v, all discs must be in column 2. Moving said discs, which do not pass through column 3, reduces to a 3-peg problem and therefore takes at least $2^{K-1} - 1$ movements.

$$d(z_2, v) \geq 2^{K-1} - 1.$$

Finally, $d(z_0, z_2) \geq 1$, because disc $N-1$ must pass from column 0 to column 2, and therefore

$$\begin{aligned} D &= d(u, v) \\ &= d(u, z_0) + d(z_0, z_2) + d(z_2, v) \\ &\geq (\Psi E - 2^{K-1}) + 1 + (2^{K-1} - 1) \\ &= \Psi E \end{aligned}$$

as required.

We note here that these are not particularly strict inequalities, so this is likely not the case in the optimal path.

5.3 Case 2: $\Delta K \leq T$

We no longer have that $\Psi E - \Psi(E - \{N-1\}) \leq 2^{K-1}$, but instead get

$$\Psi E - \Psi(E - \{N-1\}) \leq 2^{\nabla(T+K+1)-1}$$

which can be achieved in a similar fashion to before, using 5.8 and showing that $E - [\Delta s]$ has at most s elements, where $s = \nabla(T + K + 1)$. Because E is a subset of $[T+1] \cup E''$, $E - [\Delta s]$ has at most $T + K + 1 - \Delta s$ elements, and it is enough to show that this number is less than s. This boils down to the following inequalities.

$$K \leq s$$

$$T + 1 - \Delta s + K \leq s$$

The second inequality is equivalent to $T + K + 1 < \Delta(s + 1)$, a condition that is clear from the definition of s. The first inequality comes from the fact that $s \geq \nabla T \geq K$. This demonstrates the above inequality.

With (2) we get

$$d(u, z_0) \geq d(u', z_0') \geq \Psi E - 2^{\nabla(T+K+1)-1}.$$

It is possible to have $K = 0$, which is equivalent to saying $T = N - 1$. We will first look at this case in particular. Since $T \in E$, we have that disc $N-1$ starts in column 0, passes at least once through column 3, and finally arrives in column 2. Here T and $N-1$ are the same disc, so $t_1 \leq t_3$. Thus we have the following order of game states.

$$u \to z_0 = x_0 \to x_3 \to z_2 \to v$$

Write $c = \gamma_{t_3}(N-1)$ to represent the peg that disc $N-1$ last arrives at peg 2 from, clearly not equal to 2. In z_2', columns 2 and c are empty, and all the discs are in the two other columns. We write

$$\{0, 1, 2, 3\} - \{2, c\} = \{a, b\}$$

where $a \in \{0, 3\}$ and $b \in \{0, 1\}$, according to the following table:

c	0	1	3
a	3	3	0
b	1	0	1

In order to progress we introduce the final two lemmas required for the proof.

Lemma 5.9. *Let n, $s \in \mathbb{N}$ such that $s \geq 1$ and $n \geq \Delta(s-1)$, and let A be a subset of $[n]$. Then*

$$\Psi(A \cup \{b_1, ..., b_s\}) - \Psi(A) \leq \Psi[n+s] - \Psi[n]$$

for any $b_1, ..., b_s$ (not necessarily distinct) in $\mathbb{N}$.

Proof. Let $A_t = A \cup \{b_1, ..., b_t\}$ for $0 \leq t \leq s$. We see that

$$\begin{aligned}\Psi(A \cup \{b_1, ..., b_s\}) - \Psi(A) = \Psi(A_s) - \Psi(A_{s-1}) + \Psi(A_{s-1}) - \Psi(A_{s-2}) \\ + \Psi(A_{s-2}) - \Psi(A_{s-3}) + ... + \Psi(A_1) - \Psi(A_0)\end{aligned}$$

and

$$\begin{aligned}\Psi[n+s] - \Psi[n] = \Psi[n+s] - \Psi[n+s-1] + \Psi[n+s-1] - \Psi[n+s-2] \\ + ... + \Psi[n+1] - \Psi[n]\end{aligned}$$

so it is sufficient to show that $\Psi(A_t) - \Psi(A_{t-1}) \leq \Psi[n+t] - \Psi[n+t-1]$ for all $1 \leq t \leq s$. Using the fact that $A_{t-1} = A_t - \{b_t\}$, we see that we will be able to use Lemma 5.8.

By Lemma 5.4,

$$\begin{aligned}\Psi[n+t] - \Psi[n+t-1] &= \frac{1}{2}(2^{\nabla 1} + 2^{\nabla 2} + ... + 2^{\nabla(n+t)}) - \frac{1}{2}(2^{\nabla 1} + 2^{\nabla 2} + ... + 2^{\nabla(n+t-1)}) \\ &= \frac{1}{2} 2^{\nabla(n+t)} \\ &= 2^{\nabla(n+t)-1}\end{aligned}$$

We write $\sigma = \nabla(n+t)$ and claim that $A_t - [\Delta\sigma]$ contains at most σ elements. First, note that $\Delta(\sigma+1) > n+t \implies \Delta\sigma + \sigma \geq n+t$. Thus

$$\begin{aligned}
\#(A_t - [\Delta\sigma]) &\leq t + \#([n] - [\Delta\sigma]) \\
&= t + n - \Delta\sigma \\
&= \max(t, t+n-\Delta\sigma) \\
&\leq \max(t, \sigma)
\end{aligned}$$

With the final inequality coming by virtue of $\Delta\sigma \geq t+n-\sigma$. It remains to show that $t \leq \sigma$. Observe that

$$\begin{aligned}
\Delta t - t &= \Delta(t-1) \\
&\leq \Delta(s-1) \\
&\leq n
\end{aligned}$$

because $1 \leq t \leq s$, we get

$$\begin{aligned}
\Delta t &\leq n+t \\
\nabla(\Delta t) &\leq \nabla(n+t) \\
t &\leq \sigma.
\end{aligned}$$

Thus we are able to use Lemma 5.8 and we arrive at our result

$$\begin{aligned}
\Psi(A_t) - \Psi(A_{t-1}) &\leq 2^{\sigma-1} \\
&= \Psi[n+t] - \Psi[n+t-1]
\end{aligned}$$

□

Lemma 5.10. *Let A, B be finite subsets of $\mathbb{N}$. Then we have*

$$\Psi(A)+\Psi(B) \geq \frac{\Phi(n+3)-5}{4} = \frac{1}{2}\Psi[n+2]-1 = \frac{1}{4}(2^{\nabla 3}+2^{\nabla 4}+...+2^{\nabla(n+2)})$$

where $n = |A \cup B|$.

Proof. Write $E = A \cap B$. Let $L \in \mathbb{N}$. By Lemma 5.6, $\Psi_L(E) \geq \Psi_L[n]$, so

$$\begin{aligned}
\Psi(A) + \Psi(B) &\geq \Psi_L(A) + \Psi_L(B) \\
&= \Psi_L(A \cap B) + \Psi_L(A \cup B) \\
&\geq \Psi_L(\emptyset) + \Psi_L(E) \\
&\geq \Psi_L[0] + \Psi_L[n].
\end{aligned}$$

The above equality comes from the definition of Ψ and presents an interesting notion in that decomposing a set $A \subset \mathbb{N}$ into n disjoint subsets $\{A_1, ..., A_n\}$ and $\{A'_1, ..., A'_n\}$ will yield $\Psi(A_1) + ... + \Psi(A_n) = \Psi(A'_1) + ... + \Psi(A'_n)$.

As in previous proofs, we write $n+3 = \Delta m + p$, with $m = \nabla(n+3) \geq 2$ and $0 \leq p \leq m$. We see that $n \geq \Delta(m-2)$, and

$$\Phi(n+3) = 1 + (m+p-1)2^m$$
$$\Phi(\Delta(m-2)) = 1 + (m-3)2^{m-2}.$$

Take $L = m-2$. We have

$$\begin{aligned}
\Psi_L[0] + \Psi_l[n] &= (1-L)2^{L+1} - 2 + \sum_{0\leq k<n} 2^{\min(\nabla k, L)} \\
&= (3-m)2^{m-1} - 2 + \sum_{0\leq k<\Delta(m-2)} 2^{\nabla k} + \sum_{\Delta(m-2)\leq k<n} 2^{m-2} \\
&= (3-m)2^{m-1} - 2 + \Phi(\Delta(m-2)) + (n-\Delta(m-2))2^{m-2} \\
&2(3-m)2^{m-2} - 1 + (m-3)2^{m-2} + (p+2m-4)2^{m-2} \\
&= (m+p-1)2^{m-2} - 1 \\
&= \frac{\Phi(n+3)-5}{4}
\end{aligned}$$

The other two equalities follow quickly from previous results. □

Proof of Theorem 5.2 continued. We can now define

$$A = \{k \in [N-1] : z_2(k) = a\}$$
$$B = \{k \in [N-1] : z_2(k) = b\}$$

as the sets of discs in columns a and b immediately after disc $N-1$ arrives in column 2 for the last time. Clearly we have $A \sqcup B = [N-1]$, and thus by Lemma 5.10,

$$\begin{aligned}
\Psi A + \Psi B &\geq \frac{1}{2}\Psi[N+1] - 1 \\
&= \frac{1}{4}(2^{\nabla(N+1)} + 2^{\nabla N}) + \frac{1}{2}\Psi[N-1] - 1 \\
&\geq 2^{\nabla(T+K+1)-1} + \frac{1}{2}\Psi[N-1] - 1.
\end{aligned}$$

In x'_a, column a is empty as well as one other, which allows us to apply the induction hypothesis

$$d(z'_2, x'_a) \geq \Psi A$$

and similarly in v, column b is empty as well as column $1-b$, so

$$d(z_2, v) \geq d(z'_2, v') \geq \Psi B$$

Between states z_0 and z_2, the discs smaller than $N-1$ make at least $d(x'_a, z'_2)$ moves, and disc $N-1$ makes at least two moves, passing from column 0 to column 3 and then to column 2. Thus

$$d(z_0, z_2) \geq \Psi A + 2$$

and finally

$$\begin{aligned} D = d(u,v) &= d(u,z_0) + d(z_0,z_2) + d(z_2,v) \\ &\geq \Psi E - 2^{\nabla(T+K+1)-1} + \Psi A + 2 + \Psi B \\ &\geq \Psi E + 1 + \frac{1}{2}\Psi[N-1] \geq \Psi E. \end{aligned}$$

This proves the Theorem for $K = 0$, so all that remains is to prove it for $K \geq 1$. Now we are unable to compare t_1 and $t_3 + 1$, the times of states x_3 and z_2 respectively. We are able to say that they are not equal, because the most recent disc moved in each state is different. Thus we are led to one final bifurcation. We begin with assuming $t_1 > t_3 + 1$.

It follows that the optimal path γ passes through game states in the following order:

$$u \to z_0 \to z_2 \to x_3 \to v$$

Here disc $N-1$ leaves column 0, then disc $N-1$ last arrives in column 2, only after which disc T first arrives in column 3. In x_3, column 3 does not contain any discs smaller than T, and neither does one other column, which we write as $d = \gamma_{t_1-1}(T)$. All discs smaller than T are thus in the other two columns.

Write $c = \gamma_{t_3}(N-1) \in \{0,1\}$ to represent the column that disc $N-1$ moves from to get to column 2 for the last time. We see that $c \in \{0,1\}$ is justified because disc $N-1$ never enters column 3. Then we have

$$\{0,1,2,3\} - \{3,d\} = \{a,b\}$$

with $a \in \{2,c\}$ and $b \in \{0,1\}$, according to the following table:

d	0 or 1	2
a	2	c
b	$1-d$	$1-c$

Now we can define the following sets:

$$A = \{k \in [T] : x_3(k) = a\}$$

$$B = \{k \in [T] : x_3(k) = b\}$$

Since A and B don't contain any discs smaller than T, we have $A+B = [T]$. In z_2'', column a is empty as well as one other, so we have

$$d(x_3,z_2) \geq d(x_3'',z_2'') \geq \Psi A$$

and similarly in v'', column b is empty as well as column $1-b$, so

$$d(x_3,v) \geq d(x_3'',v'') \geq \Psi B.$$

Again $d(z_0, z_2) \geq 1$, so

$$\begin{aligned} D = d(u,v) &= d(u,z_0) + d(z_0,z_2) + d(z_2,x_3) + d(x_3,v) \\ &\geq \Psi E - 2^{\nabla(T+K+1)-1} + 1 + \Psi A + \Psi B \\ &\geq \Psi E - 2^{\nabla(T+K+1)-1} + \frac{1}{2}\Psi[T+2]. \end{aligned}$$

Write $s = \nabla(T+K+1)$. Since we have assumed that $T \geq \Delta K$, we see that $T+K+1 \geq \Delta(K+1)$, so $s \geq K+1$. Since $T = (T+K+1) - (K+1) \geq \Delta s - s = \Delta(s-1)$, we have $\nabla T \geq s-1$.

By Lemma 2.3, $\Psi[T+2] \geq 2^s$, so we have $D \geq \Psi E$ as required.

The final step of the proof is to prove the case $t_1 < t_3 + 1$. Here, the path γ passes through game states in the following order:

$$u \to x_0 \to x_3, z_0 \to z_2 \to v$$

In z_2' columns 2 and $c = \gamma_{t_3}(N-1) \in \{0,1\}$ are empty, so all discs are in columns 3 and $b = 1-c$. In particular, discs $b_1, ..., b_{K-1}$ are in column b because discs larger than T never pass through column 3.

In x_3'' column 3 is empty as well as one other, so $d(z_2'', x_3'') \geq \Psi A$, where

$$A = \{k \in [T] : z_2(k) = 3\}$$

and in v', columns b and $1-b$ are empty, so $d(z_2, v) \geq d(z_2', v') \geq \Psi B$, where

$$\begin{aligned} B &= \{k \in [N-1] : z_2(k) = b\} \\ &\supseteq \{k \in [T] : z_2(k) = b\} \sqcup \{b_1, ..., b_{K-1}\}. \end{aligned}$$

The set $A \cup B$ contains $[T] \sqcup \{b_1, ..., b_{K-1}\}$, so has at least $T+K-1$ elements, and

$$\Psi A + \Psi B \geq \frac{1}{2}\Psi[T+K+1] - 1.$$

Between x_3 and z_2, discs smaller than T make at least ΨA moves, and between u and x_0, the make at least $\Psi(E \cap [T])$ moves. Additionally, in u the discs $b_1, ..., b_{K-1}$ are all in column 0, whereas in z_0 they have all left column 0 and are in columns 1 or 2. Because they do not enter column 3, they must make at least $2^{K-1} - 1$ moves between u and z_0, as per the 3-peg problem. Lastly, there is at least one move between x_0 and x_3, and at least one move between z_0 and z_2. Thus, we achieve the inequality

$$d(u, z_2) \geq \Psi(E \cap [T]) + \Psi A + 2^{K-1} + 1$$

which gives us

$$\begin{aligned} D = d(u,v) &= d(u,z_2) + d(z_2,v) \\ &\geq \Psi(E \cap [T]) + \Psi A + 2^{K-1} + 1 + \Psi B \\ &\geq \Psi(E \cap [T]) + \frac{1}{2}\Psi[T+K+1] + 2^{K-1}. \end{aligned}$$

On the other hand, we have $E = (E \cap [T]) \sqcup \{T, b_1, ..., b_K\}$ and $T \geq \Delta K$, so

$$\Psi E - \Psi(E \cap [T]) \leq \Psi[T + K + 1] - \Psi[T]$$

by Lemma 5.9. Consequently,

$$D - \Psi E \geq \Psi[T] + 2^{K-1} - \frac{1}{2}\Psi[T + K + 1].$$

This final expression is positive by definition of Ψ, so we have arrived at our proof. $\square$

The final result of the paper goes as follows.

Corollary 5.11. *In the 4-peg variant of the Tower of Hanoi, at least $\Phi(N)$ moves are required to transfer N discs from one peg to another.*

Proof. We can suppose that $N \geq 1$. Let $u : [N] \to \mathcal{C}$ be the game state with all discs in column 0, and v the state with all discs in column 2. We define the path γ and intermediate states z_0, z_1 as previously; they must therefore occur in the following order:

$$u \to z_0 \to z_2 \to v$$

In z_0', column 0 is empty as well as one other. Thus by Theorem 5.2,

$$\begin{aligned} d(u, z_0) \geq d(u', z_0') &\geq \Psi\{k \in [N-1] : u(k) = 0\} \\ &= \Psi[N-1]. \end{aligned}$$

Similarly, in z_2' column 2 is empty as well as one other, so by the same Theorem

$$\begin{aligned} d(v, z_2) \geq d(v', z_2') &\geq \Psi\{k \in [N-1] : v(k) = 2\} \\ &= \Psi[N-1]. \end{aligned}$$

Between z_0 and z_2 disc $N-1$ must move, so $d(z_0, z_2) \geq 1$. Combining these inequalities, we get

$$\begin{aligned} d(u, v) = d(u, z_0) + d(z_0, z_2) + d(z_2, v) &\geq 1 + 2\Psi[N-1] \\ &= \Phi(N) \text{ (by Lemma 5.4)} \end{aligned}$$

$\square$

6 Conclusions

Bousch's proof successfully demonstrates that the Frame-Stewart algorithm is optimal for Hanoi puzzles with four pegs. This of course begs the question: does the Frame-Stewart Conjecture hold for $p > 4$? At time of writing this is still an open problem, and there are doubts as to whether it will ever be solved. It is interesting to look at how quickly the 4-peg solution grows for larger and larger n when compared to the 3-peg puzzle. [4] states that $\Phi(n)$ is "always within 6.2% of $\sqrt{n}2^{\sqrt{2n}}$ for large n." This is a significantly greater rate of change than that of the 3-peg puzzle, always being close to 2^n, and is suggestive of the difficulty that mathematicians will face when attempting to tackle the generalised problem: graph theoretic approaches will become exponentially more complicated for larger p. Computation can only get one so far with questions like this, greater mathematical insight is likely required to truly crack the puzzle.

To conclude, the Tower of Hanoi epitomises the world of mathematical puzzles. A baseplate, three metal pegs, and eight wooden discs, equipped with an appropriately succinct set of rules, is enough to take an inquisitive mind on an adventure through number theory, graph theory, and beyond, and still leaves an unanswered question at the end. If a problem with such a simple surface is able to stump mathematicians, it serves as a perfect reminder of the endless depth and intrigue of mathematics.

References

[1] A. Hinz, *The Tower of Hanoi– Myths and maths.* Basel: Birkhäuser, 2013.

[2] H. E. Dudeney, *The Canterbury Puzzles, and Other Curious Problems.* London, UK: Thomas Nelson and Sons, 1907.

[3] O. F. Inc., "Tetrahedral (or triangular pyramidal) numbers," 2010. Published electronically at `https://oeis.org/A000292`.

[4] P. K. Stockmeyer.

[5] K. Rock, "On the planarity and hamiltonicity of hanoi graphs," *MTH 501 Research Project at Portland State University*, 2016.

[6] W. F. Lunnon, "Correspondence," *The Computer Journal*, vol. 29, pp. 478–478, 01 1986.

[7] T. Bousch, "La quatrième tour de hanoï," *Bulletin of the Belgian Mathematical Society - Simon Stevin*, vol. 21, no. 5, pp. 895 – 912, 2014.

[8] D. Kao, "Reves-puzzle-solver." `https://github.com/Darren-Kao/Reves-Puzzle-Solver`, 2023.

7 Appendix

Using Python code from [8] and self-written JavaScript graphics renderer, we can witness the Frame-Stewart algorithm in action for a 4-peg 7-disc puzzle.