The Mathematician Educator
2020, Vol 1, No. 2, 62-74

An Analysis of a Solution

Derek Holton
Melbourne Graduate School of Education
University of Melbourne
Australia

With thanks to Duncan Symons, with whom I have had many useful conversations

How do mathematicians solve problems? In what follows we look at the Tower of Hanoi problem that is accessible to secondary students and show how a mathematician might tackle it. This 'showing' involves more than the final written solutions. It considers how the mathematician might be thinking and the body of mathematics that might have been used in the process. Now we have avoided taking a real mathematician's work on a research problem, simply because that would require the reader to know a large amount of mathematics relevant to that problem. Instead we will take an imaginary mathematician working on a problem that is easy to explain and does not rely on mathematics that is beyond the experience of secondary students. By doing this we hope that the reader is able to understand what happens while the fictitious mathematician is working on the problem. In the Discussion section, we will consider some of the things that a mathematician does on problems they are trying to solve.

Keywords: Tower of Hanoi problem, mathematician, analysis, mathematical problem solving

The Tower of Hanoi Problem

The Tower of Hanoi problem was invented by Édouard Lucas in 1883. (https://en.wikipedia.org/wiki/Tower_of_Hanoi). There is a background story to the game that involves either monks or Brahmin priests, depending on the version you happen to read. The religious men have 64 discs and they are required to move one of the discs in a certain way each day. When the last disc is put into place, the story says that the world will end. We'll come back to this later.

So what is the Tower of Hanoi problem? A Tower of Hanoi game consists of three vertical rods that have a number of discs of different sizes on them (Holton, 2013; Vakil & Heled, 2016). The starting position has all of the discs on one rod with the smallest on the top, the largest on the bottom with the sizes of the discs increasing from the top down. Figure 1 shows a Tower of Hanoi that has 7 discs. The aim of the game is to move all the discs from one rod to another. In the process no bigger disc is allowed to lie on top of a smaller disc. Examples of animated Towers of Hanoi can be found on the web (for example, in https://en.wikipedia.org/wiki/Tower_of_Hanoi).

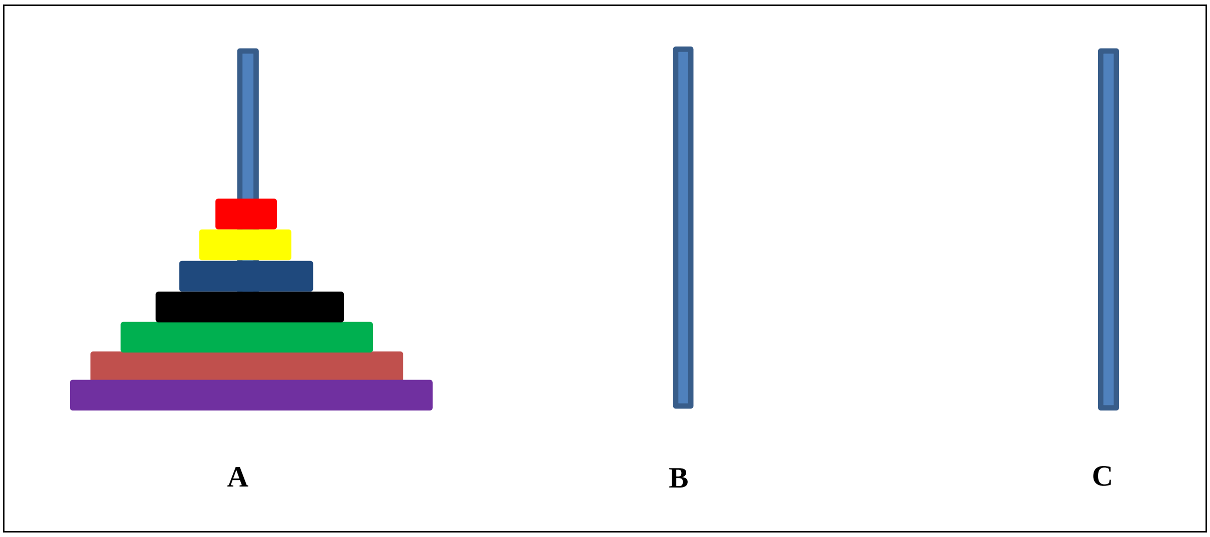

Figure 1. A Tower of Hanoi that has 7 discs

The mathematician, Freda, who we are assuming has never seen the problem before, is given a Tower of Hanoi game with 7 discs. Assuming the rods are in a line, she is asked to move the discs from the leftmost rod, A, to the rightmost rod C, in the fewest number of valid moves. Note that there are many problems that are similar in nature to the Tower of Hanoi and accessible to secondary students. Some sources are Gardiner (1987) and Holton (2013). In the Discussion we look at some generalisations of the Tower of Hanoi.

Getting Started

Freda started moving the discs around. Her first attempts began by just moving the discs around in a more or less random way, but always subject to the rules and trying to keep in mind the final goal. As you would expect this problem was too difficult for her as it stood. Very soon she realised that there were many places she lost her way and had to start again or she recognised she was back at a position she had been in before. In fact, she got frustrated and moved the tower of discs all together in one move. Maybe this problem can't be solved in a reasonable time anyway.

At that point she decided all this experimenting wasn't getting her very far and she stopped dead (see Schoenfeld, 1992). There had to be a better way! One technique that she had found useful in other problems was to try a smaller case to help her understanding of the problem (Shepherd et al., 2012). For instance, what if she only had 4 discs to worry about? Better still, 1 disc. Sarcastically triumphant she moved 1 disc from A to C in one move. Then 2 discs in three moves. So, she checked out 3 discs (see Figure 2). But that took seven moves. Well she didn't actually get all the discs on to rod C, but she could fix that by moving the smallest disc to C to start with. Fixing rod C as the destination for all of the discs wasn't really changing the problem because in some sense rods B and C are equivalent. If she could get all discs from A to B, she could get it from A to C and in the same number of moves.

But how could she know that she couldn't do the 3 discs problem in fewer than 7 moves? After all, the problem asked her to be as efficient as possible.

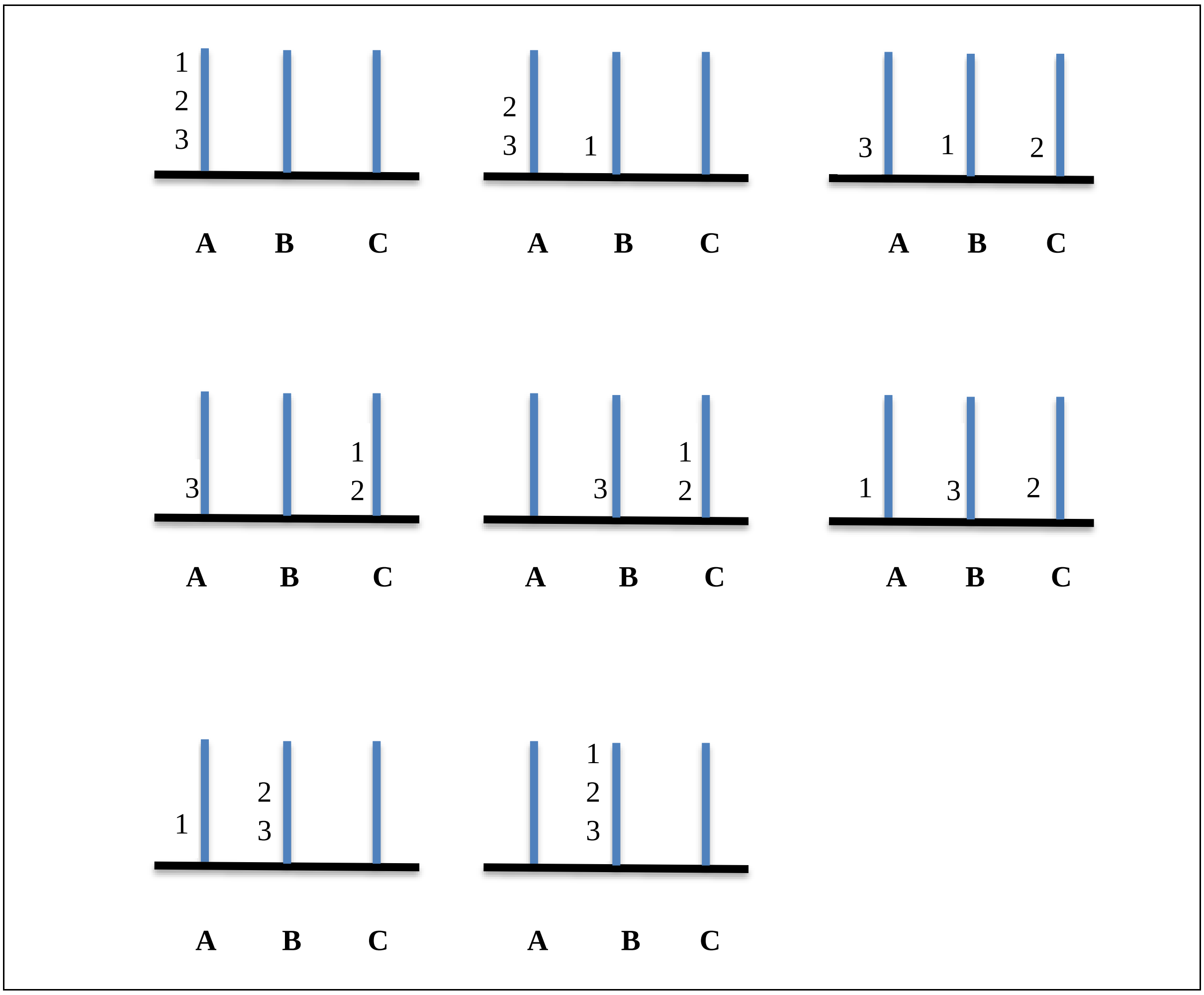

Figure 2. The moves for 3 discs

The other thing that she did on some paper was to replace the discs by numbers with 1 being the smallest disc and 3 the biggest. This was quicker than drawing discs. One of the other advantages of writing down the moves was that she then had a record of what she had done. It was a bit painful, but it would be worth being careful right from the early stages.

But 7 discs was a big jump from 3. Presumably 4 discs would take considerably more than 7 moves. Was the number of moves going to go up exponentially? It looked as if it was more than doubling each time. So, she worked a little harder and did the 4 discs problem, Figure 3. Eventually Freda could do it in 15 moves.

```
1
2            2
3            3            3
4  -  -      4  1  -      4  1  2


3     1            1      1
4  -  2      4  3  2      4  3  2


                1            1
1  2            2            2
4  3  -      4  3  -      -  3  4


   2  1            1      1
-  3  4      2  3  4      2  3  4

                                2
1     3            3            3
2  -  4      2  1  4      -  1  4

      1
      2
      3
-  -  4
```

Figure 3. The moves for 4 discs

But note that she had changed her notation removing the rods and lines underneath the rods as well as the A, B, and C. It is still clear what Freda is doing. It is certainly also a more efficient way to represent the situation and in some ways makes it easier to see what is happening.

A Side Track

Now Freda realised that the number of moves might be following a pattern. At this point she had started to go off on a side track. Though it might be valuable and interesting to find the number of moves, it did not directly help with finding how to make the moves. Of course it might be useful later; it was certainly an interesting idea to follow up. So, she made up a table of the results so far along with her attempt with 5 discs (Table 1). The entry for 5 discs was largely a guess given the pattern she was seeing with the other four entries. It seemed that she got the next number of moves by multiplying the previous case by 2 and adding 1. But she knew this was only a conjecture and she needed to prove it.

Table 1.
Number of moves

Number of discs	1	2	3	4	5
Number of moves	1	3	7	15	31??

So, she set up some more notation. Freda let T(d) be the *minimum* number of moves needed for d discs. So T(3) = 7 for example. This allowed her to formalise her conjecture, Conjecture 1.

***Conjecture 1*: ** For d > 1, $T(d) = 2 \cdot T(d - 1) + 1$.

If this were true, then

$$T(5) = 2 \cdot T(4) + 1 = 31.$$

Now proving this conjecture wasn't going to be easy without a new idea. Was there anything in Figures 2 and 3 that might help? Suddenly something came to Freda. Now the largest disc had to get out sometime. It could only get out if there was a spare rod. If there was a spare rod, all of the other d – 1 discs had to be on another rod. So just before the largest disc could move, she needed to move the tower of the smaller d – 1 discs away from the largest disc. When the largest disc was moved she then had to move the smaller disc back onto this largest disc again. Looking at Figure 3, she marked in yellow the 1, 2, 3 tower whenever it formed. She also underlined the move of the biggest disc, 4, with green. So Freda saw that the first motion of the smaller tower with the discs 1, 2 and 3, sent this tower to rod B. Then 4 moved to rod C. This was followed by the move of the 1, 2, 3 tower. This is all shown in Figure 4.

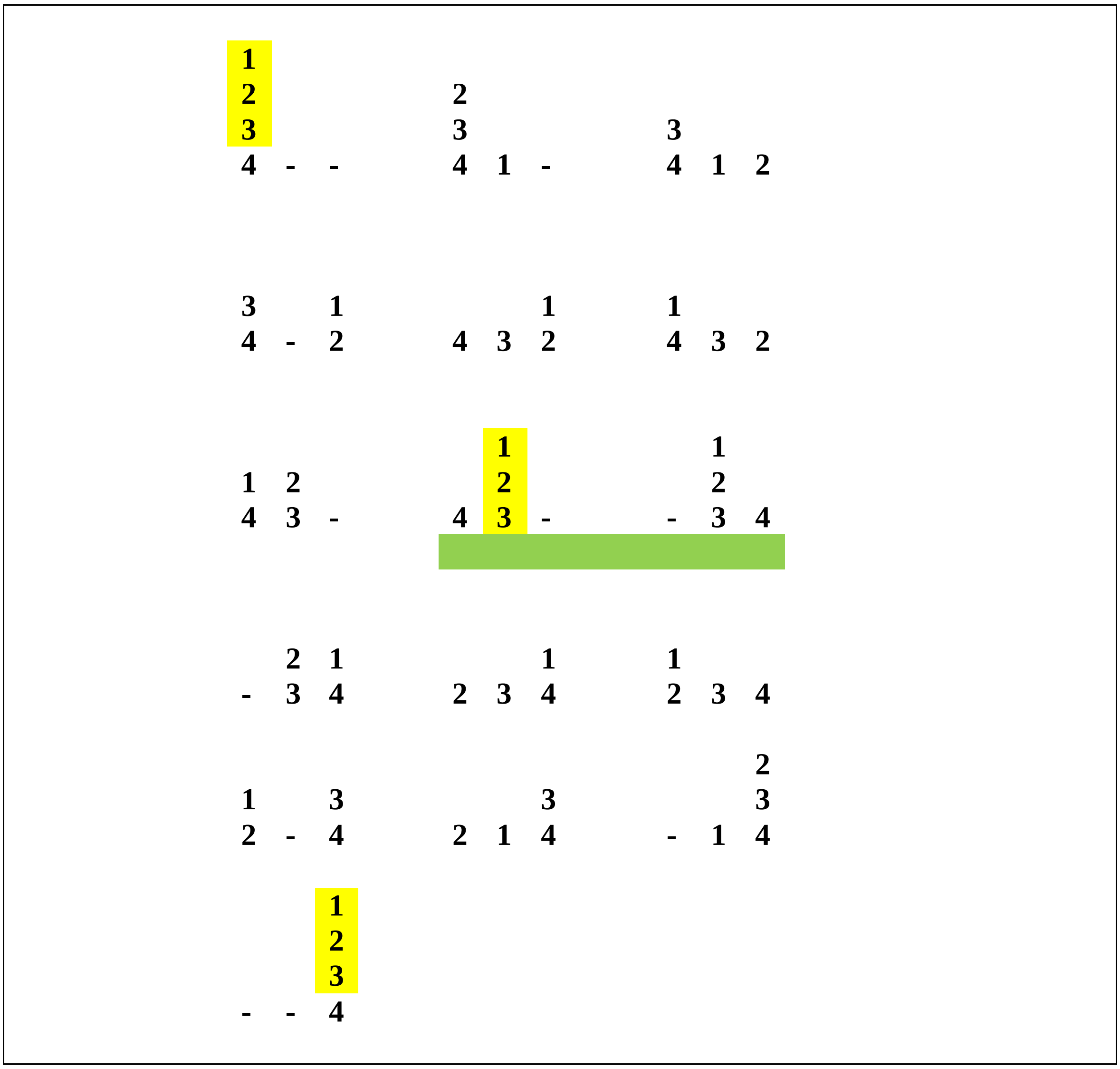

Figure 4: Highlighted moves of the 4 discs game

Then Freda drew the final drawing that cleared the whole thing up (Figure 5). The conjecture had become a Theorem!

***Theorem 1*:** For d > 1, T(d) = 2·T(d – 1) + 1.

<u>Proof</u>**:** After seeing the moves from Figure 5, all she needed to establish was that the final result was a minimum number of moves for d discs. In Figure 5, the d – 1 tower needs to get to rod B (so that the largest disc can get to rod C). So the d – 1 tower needs T(d – 1) moves to do this. (Freda knew that T(d – 1) was the fewest moves to move d – 1 discs.) Then she moved the largest disc to rod C in one move. From here the d – 1 tower moved to rod C in T(d – 1) moves. So T(d) = T(d – 1) + 1 + T(d – 1) = 2·T(d – 1) + 1.

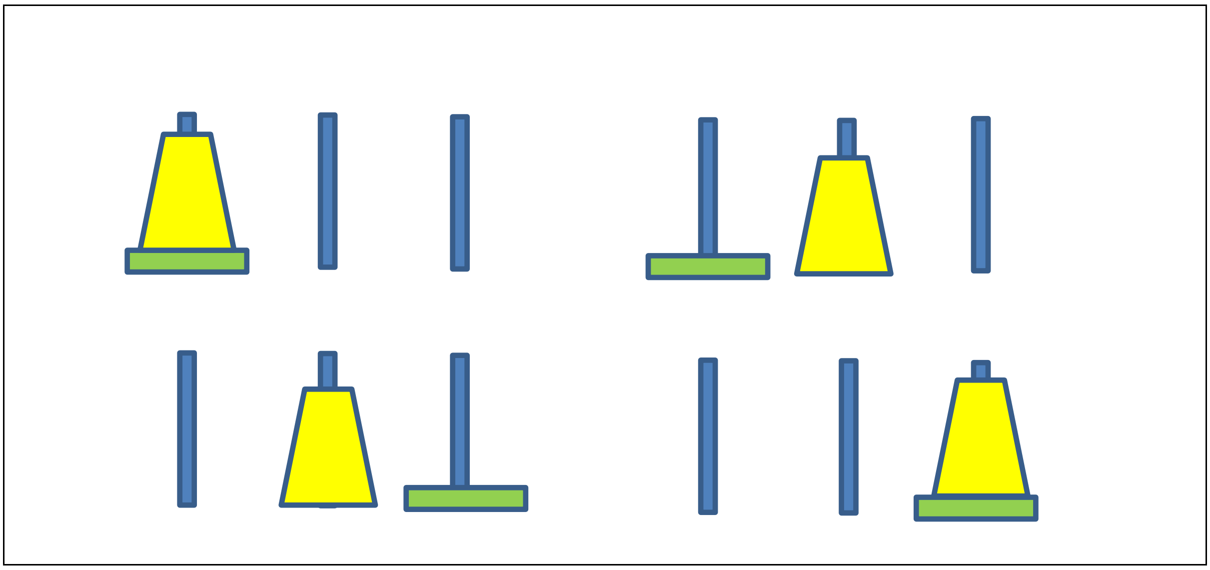

Figure 5. Highlighted moves of d discs

A Closed Form for the Number of Moves

For many discrete problems, it is often the case that a recursive form of the generalisation, as Freda has found in Theorem 1, is often the first result. Here a recursive form is one that depends on the earlier sized discs. We have seen that recurrence relation in the last section (Theorem 1). However, an algebraic form (closed form or formula) is usually more efficient as it reduces the number of calculations needed to get, say the minimum number of moves required for a given number of discs. For the monks to find out how many moves they require to complete their travail of the 64 discs, they need to find the number of moves for 1, 2, 3, …, 63 discs. It would be better for them if they had a closed algebraic form that would produce T(64) using some formula. To get this formula, Freda was going to have a conjecture to work with. This is going to have to be produced by some good guessing.

Slight panic! In a boring staff meeting Freda suddenly realised that she had made a small error in Theorem 1. Thinking about it again she definitely had to have $d > 1$. Otherwise $T(1) = 2 \cdot T(0) + 1$. But what is T(0)? It doesn't necessarily make any sense. Fine, but there are lots of sequences with $T(d) = 2 \cdot T(d - 1) + 1$. For instance: 4, 9, 19, 39, and so on. For this sequence $T(1) = 4$. To make the sequence for the Tower problem, she knew she had to specify $T(1) = 1$. So she adjusted her original theorem slightly to give Theorem 1'.

***Theorem 1'*:** For $d = 1$, $T(1) = 1$ and for $d > 1$, $T(d) = 2 \cdot T(d - 1) + 1$.

The proof is the same as that of Theorem 1 plus the observation that $T(1) = 1$. Freda had fixed the sequence she was looking at and gave the recursive form of T(d) precisely.
Then Freda went back to try to find a closed form for T(d). She started out by looking at the recursive form applied to early values of T.

$T(1) = 1$;
$T(2) = 2 \times 1 + 1 = 2 + 1$;
$T(3) = 2(2 + 1) + 1 = 2^2 + 2 + 1$; and
$T(4) = 2(2^2 + 2 + 1) + 1 = 2^3 + 2^2 + 2 + 1$.

She then saw what was going on. Was it true that $T(d) = 2^{d-1} + 2^{d-2} + ... + 2 + 1$? If so how could she prove it? What ammunition did she have to do this? Not much. Could Theorem 1' help? Why not the same way as we got T(2) from T(1) and T(3) from T(2) and T(4) from T(3)? If that works then why not

$$\begin{aligned} T(d) &= 2 \cdot T(d-1) + 1 \\ &= 2(2^{d-2} + 2^{d-3} + ... + 2 + 1) + 1 \\ &= 2^{d-1} + 2^{d-2} + ... + 2 + 1. \end{aligned}$$

She suddenly realised that she had solved her number of moves problem. What's more she had used the method of Proof by Mathematical Induction (see https://en.wikipedia.org/wiki/Mathematical_induction). This method works by first showing that one or more earlier numbers work. In Freda's case she had shown that T(1) = 1. That's not very hard, but there are times when more starting numbers might have to be tested, depending on the complexity of the recurrence relation. The only other Proof by Induction step she required for any number of discs, d – 1 say, was that she could show that the required closed form was true for the next case, d here. Freda had done all of that in Theorem 1'.

The proof method works like dominoes falling as she had implied above. From what she said if something is true for 1, then putting 2 for d in Theorem 1' she will get the value of T(2). From d = 3, she will get T(3) and so on. It's exactly like dominoes falling. Once we know that the first domino falls, it will knock over the second domino which will knock over the third domino and so on until all dominoes have fallen. So she had another theorem, Theorem 2.

***Theorem 2*:** For all $d \geq 1$, $T(d) = 2^{d-1} + 2^{d-2} + ... + 2 + 1$.

Hmm, the monks had at least 2^{63} moves to make. Freda did a bit of scrappy approximations and found that even if the monks made a move per second, then it would take them about 290 billion years to cause the Earth to end. With the relief she got from that, "I'm unlikely to last that long", she went back to her original problem.

How to Move the Discs

It was clear to Freda that she had also completed the moves problem. All she had to do was to make the first 2 discs move by the three steps she knew well. Then follow that by using the recurrence relation to make the T(3) moves. This could go on for T(4) moves and so on. Then she realised how useless that was! It was the same old recurrence problem. It was OK for small towers but not very good if you have a large tower. It would be good if she could find what to do at any stage for any height tower. One of the advantages of that would be that if she decided it was time for a coffee, she could pick up where she left off. It might also help her to make a program to move the towers.

There was nothing for it now, but to experiment. She looked at towers with up to 7 discs and made careful notes on them all the moves in the hope of finding a pattern. The first surprising thing she noticed was that the biggest disc always seemed to move only once. That wasn't really a surprise as she had used that in Theorem 1'. But then the next smallest disc always took 2 moves, the next 2^2, the next 2^3 and so on. That was where all of the powers of 2 in Theorem 2 came from! And she could prove that by using Proof By Induction again.

Corollary to Theorem 1': Let the discs be numbered from 1 to d in increasing size and let M(i) be the number of moves for disc i. Then $M(i) = 2^{d-i}$.

Then picking up her children from school, Freda suddenly realised that 2 – 1 = 1! But more than that she should have known that $(2 - 1)(2 + 1) = 2^2 - 1$, and $(2 - 1)(2^2 - 2 - 1) = 2^3 - 1$, and so on. So $(2 - 1)(2^{d-1} + 2^{d-2} + ... + 2 + 1) = 2^d - 1$. Ah! $T(d) = 2^d - 1$! She thought that this was a much nicer closed form than the one she came up with in Theorem 2. It was simpler and much easier to compute with. Why hadn't she thought about that before? It was so obvious when she looked at all of the positions of the discs including the original position of the discs (see for example, Figures 2 and 4). They were always a power of 2. And worse if she added 1 to the entries in Table 1, she always got a power of 2. She was so annoyed, but glad that she had seen this before she had written up her work for publication.

Good as it was, this had nothing to do with moving individual discs. But she didn't get very far with that, basically because she didn't see how to prove her observations. She summarised her findings in Table 2. (Note that we number the individual discs as in the Corollary to Theorem 1'.) The letters tell the rod movements of a disc.

Table 2.

Conjectured types of moves

Number of discs	Moves: Even numbered discs	Moves: Odd numbered discs
Even	A → C→ B	A → B → C
Odd	A → B → C	A → C→ B

Consider the tower in Figure 4, where the number of discs is 4 (even). Disc 2 moves from rod A to C to B to A to C and disc 4 goes from A to C. The odd numbered discs go from A to B to C and so on until they reach their final positions.

Freda felt that this provided, along with the rules of the game, sufficient information to solve any number of discs. However, she was not quite sure how to prove it. It did occur to her that the bottom disc d, had to get to C straight away. And then discs d – 1 had to be at B for the previous move. If she kept following this argument upwards, then disc 1 would need to go straight to B if d was even and C if d was odd. But she couldn't ever quite see how to make a proof of Table 2 out of this. But she did know that it always happened for d small. If it did work it would be a quite nice solution to the problem. Perhaps a Proof By Induction?

Extensions and Other Ideas

While all of this was going on, Freda's mind was always thinking of new ideas, but she didn't always find time to go much further than that – she never made any more progress with them. For example, she thought about what might happen if the moves were restricted to those between neighbouring rods? This meant that disc 1 had to first go to B and then C before disc 2 could move. Clearly that would mean that the minimum number of moves would increase, but by how much? What would the moves be overall?

What if the rods were arranged in a circle where the only moves allowed were in a clockwise direction or in an anti-clockwise direction? How many moves would be required then? What would the moves be?

But why not start at any valid position of the discs on the rods and aim to get the discs to some other valid nominated position? Could this be done? How many moves would be needed? What were the actual moves?

What if there were more than three rods? For four rods it was easy to see in her head that if there were more than 3 discs then the required number of moves would be less than T(d). Would the four rod number of moves get closer to ½T(d) as d got bigger? But Freda never really spent any time on this extension to the original problem.

What if different discs were coloured in red or white or blue? Would it be possible to move them so that the original mixture of colour discs ended up on monochromatic towers? Maybe it would be better to start this off with just red and white discs?

One day when she was at a red light, it suddenly occurred to her that 2^n was the number of binary numbers from 1 to 2^n. Was every move in the three rod problem equivalent to a binary number? Would this give a nice way to represent all the moves? Would this help with the method of moving them? But the light changed and she didn't ever get back to the idea.

All of the ideas in this section are considered in https://en.wikipedia.org/wiki/Tower_of_Hanoi.

Discussion

First, we should remind the reader that Freda is not as dumb as you might think. Most mathematicians could solve the moving problem. But you have to remember that she has been written as if this is a problem that no one has ever solved before. It is important to note that research mathematics is not easy. Mathematicians are not able to solve every problem that they try to solve. Some problems take decades or even centuries before someone is able to complete them. See for example Goldbach's conjecture (https://en.wikipedia.org/wiki/Goldbach%27s_conjecture). So it is not surprising that there is something that Freda can't prove here.

Then note that mathematics isn't just created in a musty office somewhere. We have seen that Freda's mind was always ticking away, maybe in a meeting or maybe when driving a car. She didn't seem to have full control over her thoughts or even when ideas might come to her. But she did know that they always needed her to work quite hard before they deigned to appear.

In 'serious' problems that a mathematician might tackle, there are many side-tracks that researchers are tempted along in just the same way that Freda moves from finding the sequence of moves to finding how many moves are required. And mathematicians often see what is correct, but in a less sophisticated way, such as $T(d) = 2^{d-1} + 2^{d-2} + ... + 2 + 1$, rather than the simpler form of $T(d) = 2^d - 1$. This kind of error, if it can be called that, often brings embarrassment even if no one else knows about it. But this kind of simple error is often avoided when mathematicians work together.

And mathematicians feel for their work. It's not only embarrassment that they feel, but joy and failure and so on. They usually get a buzz if a paper of theirs is accepted and failure when a paper is not accepted. Similarly, if they find a solution or can't find a solution. They also have emotional reactions to the mathematics. Students also can have similar feelings, see Tony feeling 'silly' in Tay, Toh, Dindyal & Deng, (2014). Certain work is felt to be elegant, nice, and even beautiful, in much the same way that people feel about works of art. In the same way too, it is hard to define what is elegant and what other such terms really mean. On the other hand, some mathematics can be ugly, such as in a Brute Force and Ignorance approach when every single case out of many has to be considered. It is definitely not true to say that mathematicians work dispassionately.

And self-interrupting the flow of work is not uncommon. Freda interrupted the disc moving search, by pursuing the pattern for T(d) in the 'A Side Track' section. She does it again, when she suddenly sees that she there is a simpler form for T(d).

The number of moves that Freda found took two forms. There was the recursive form $T(d) = 2 \cdot T(d-1) + 1$ and the closed forms $T(d) = 2^{d-1} + 2^{d-2} + ... + 2 + 1$ and $T(d) = 2^d - 1$. It is important to note the large advantage that the closed forms have over the iterative form. If Freda needs to find the number of moves for 100 discs, iteration means a large number of calculations. This is not the case with the closed forms. And it is clear that the simpler form $T(d) = 2^d - 1$, is easier for calculations than the other closed form. But the iterative form does have value too in the current situation. It allows us to prove that the other two forms hold.

You should also note that what she is doing in most of her work is scribbling things down on paper. And we mean scribbling. Mathematicians try to be as careful as possible to get things right. However, they may spread bits and pieces all over the place, even the proverbial back of envelopes. Some draw on the door of their shower. Things are jotted down all over the place on paper or white board or computer. It is only at the stage of writing a paper for publication that great effort is given to formal writing along with checking that all that they have found is true.

Checking, of course, is continually being done. Is the power of 2 expression correct? Should the first term be 2^d or 2^{d-1}? Is $T(d) = 2^d - 1$, or 2^{d-1}? Do all the even discs in the even case really go A → C→ B or do some of them move A → B → C?

Notice too that not only do ideas and directions change as the work progresses but so too does notation. Freda found, between seeing the original problem and Figure 2, that it was better, more efficient, to use numbers rather than drawing different sized discs. By Figure 3 she has noticed that the vertical rods and the horizontal bases are superfluous. This is part of mathematicians searching for simplicity. The simpler the notation, the easier it is to work with as well as remembering what it is being used for. For example, if there is a problem on cakes that needs algebra, it is better to use c (for cake) than some other member of the alphabet. This is something that might be encouraged in school.

What we have not shown here is that mathematicians generally work in groups, at least in pairs. The notion of someone working alone for hours in an ivory tower is not a true picture of how much of mathematics is done.

On another level, mathematicians work according to certain, often unwritten laws. Implicitly they will experiment with a problem to make sure that they understand it and to get ideas/conjectures. These conjectures may be wrong but they give the worker clues on where to head next. Proving is a fundamental idea in mathematics. Because the subject is based on axioms, mathematicians are able to prove things far beyond 'reasonable doubt'. Subject only to the axioms, proofs give results that are undeniable. When a conjecture is found to be false, then the error in the conjecture can be adjusted until a correct conjecture is found or until the person(s) involved cannot find a proof and give up. Freda got to this stage when she couldn't see how to prove how discs would need to be moved. When a proof is found, the next thing to look for may well be an extension of the problem in some way. Then mathematicians go around a loop of experiment, conjecture and proof again. Although Freda doesn't mention this as she goes along, the steps follow Polya's four stages of problem solving (see Polya, 1945). For example, Polya's 'look back' encourages alternative solutions as well as extensions of the original problem. Looking for extensions and generalisations of a problem are one way that mathematics moves forward.

There are, of course, resources that mathematicians have recourse to. There are always books and published papers where they can find previously proved results. In the case where sequences of answers are found, as in Table 1, they can check the Online Encyclopedia of Integer Sequences (http://oeis.org/) to see if anyone has found that sequence before. If the sequence they have found is of different origin to those that are known, it might help people to prove the sequence in the new application.

It is no accident that much of what applies to Freda applies to anyone doing mathematical problem solving at every level. But the importance of what is said here is that the much of it is true for *anyone* solving *any* problem.

References

Gardiner, A. (1987). *Discovering mathematics: The art of investigation*. Oxford: Clarendon Press.

Goldbach's Conjecture in https://en.wikipedia.org/wiki/Goldbach%27s_conjecture.

Holton, D. (2013). *More problem solving: The creative side of mathematics*. Leicester, UK: The Mathematical Association.

Integer Sequences in http://oeis.org/.

Mathematical Induction in https://en.wikipedia.org/wiki/Mathematical_induction.

Polya, G. (1945). *How to solve it*. Princeton, New Jersey: *Princeton University Press*.

Schoenfeld, A. (1992). *Learning to Think Mathematically: Problem solving, metacognition, and sense making in mathematics.* In D. Grouws, (Ed.) *Handbook for research on mathematics teaching and learning* (pp. 334-370). New York: Macmillan.

Shepherd, M. D., Selden, A. & Selden, J. (2012). University students' reading of their first-year mathematics textbooks. *Mathematical Thinking and Learning, 14:3,* 226-256.

Tay, E. G., Toh, P. C., Dindyal, J., & Deng, F. (2014). Tony's story: Reading mathematics through problem solving. In C. Nicol, S. Oesterle, P. Liljedahl & D. Allan (Eds.), *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education.* (pp. 225-232). Vancouver, Canada: Psychology of Mathematics Education (PME).

Tower of Hanoi in https://en.wikipedia.org/wiki/Tower_of_Hanoi.

Vakil, E., & Heled, E. (2016). The effect of constant versus varied training on transfer in a cognitive skill learning task: The case of the Tower of Hanoi Puzzle. *Learning and Individual Differences, 47*, 207-214.

Author

Derek Holton (e-mail: derek.holton@bigpond.com), Melbourne Graduate School of Education, University of Melbourne, Australia.