Branching and Nodes in Collatz Conjecture Tree

Aditya Vishwakarma (adityav31121999@gmail.com)

Abstract

The Collatz conjecture is a mathematical puzzle that determines how to operate on a positive integer by considering whether the number is even or odd. In this paper, approximation of generalized term is presented along with its derivation and stopping time for nodes. The values of trees are discussed as nodes and branch values, their behaviour, branching and non-branching branches with no-branching criteria, and relationship between branch values of two consecutive branches have been worked out.

Keywords: Branches, Nodes

Introduction

The Collatz conjecture is a famous unsolved problem in mathematics, and was proposed by Lothar Collatz in 1937. It involves a sequence defined as follows:

$$a(n) = \begin{cases} 3n+1, & n = odd \\ n/2, & n = even \end{cases} \qquad \text{-eqn. 1}$$

The conjecture states that you can always work your way up to the number 1 from any positive integer. The sequence continues in a loop at 1, 4, 2, 1, and so on. For example, 17, 52, 26, 13, 40, 20, 10, 5, 16, 8, 4, 2, 15, 4, 2, 1, 4, 2, 1, 4, 1, 4, 1, ... and so on. The stopping time is the number of steps required for an integer to equal one. It could be lower for larger numbers and higher for smaller values, or vice versa. In the case of 17, reaching 1 takes 12 steps, whereas 27 takes 111 steps. It takes exactly n steps for 2^n values.

Collatz Tree

Base Stem

It simply takes one condition to obtain an odd-valued result: it is divided by two when the input integer is even. It can, however, halve an even number ($44/2 = 22$). The first criterion, 3x+1, can even out an odd number (3*21+1 = 64). Therefore, to obtain one, the sequence involves multiple odds and even phrases, the last of which is 4-2-1. For values of 2^n, it just takes n steps to reach 1, thus it is the shortest path. If the order is reversed, most of the terms preceding one are even and often powers of two.

$$a(n) = \begin{cases} \frac{n-1}{3}, & mod(n-1,3) \equiv 0, n > 1 \\ 2n \end{cases} \qquad \text{-eqn. 2}$$

Nodes are values specified as the values resulting from the first condition. Through equation (2), the first condition may be considered only when n-1 is divisible by 3, and the number can be doubled every time, resulting in two solutions for each alternating node. Starting from 1, 1 becomes 2, 2 becomes 4, but because 4 satisfies the first condition, it leads to $\frac{4-1}{3} = 1$, 1 leads to 2, then 4, but then 4 doubles to 8, 8 doubles to 16, 16 satisfies the first condition, it leads to $\frac{16-1}{3} = 5$, 16 doubles to 32, 32 to 64, 64 leads to

21 and 128, 128 to 256, 256 to 85 and 512, 512 to 1024, 1024 to 341 and 2048, and so on. Hence, it can be assumed that the value of 2^n is the base stem.

Nodes and Branches

Nodes are values that are defined as the values formed from the first condition of eqn. 2; and are odd. Nodes formed through 2^n values are given as:

$$\boldsymbol{Node} = \frac{R-1}{3},\ R = 2^{2n+2}, n \geq 1 \qquad \text{-eqn. 3}$$

Nodes are always unique. Nodes give rise to branches. It follows that every single node value is thought to be at position 0 (value*2^n = value). Nodes are what gives a branch its value. Node values are used to construct branch values. Each of these branch values is a multiple of two numbers i.e., node and 2^n.

$$\boldsymbol{Branch\,Value} = Node * 2^n \qquad \text{-eqn. 4}$$

Again, if eqn. 2 is followed for all node values, new nodes are formed from these branches and new branches are created from newly formed nodes.

$$\boldsymbol{Node} = \frac{Branch\,Value - 1}{3}$$

-eqn. 5

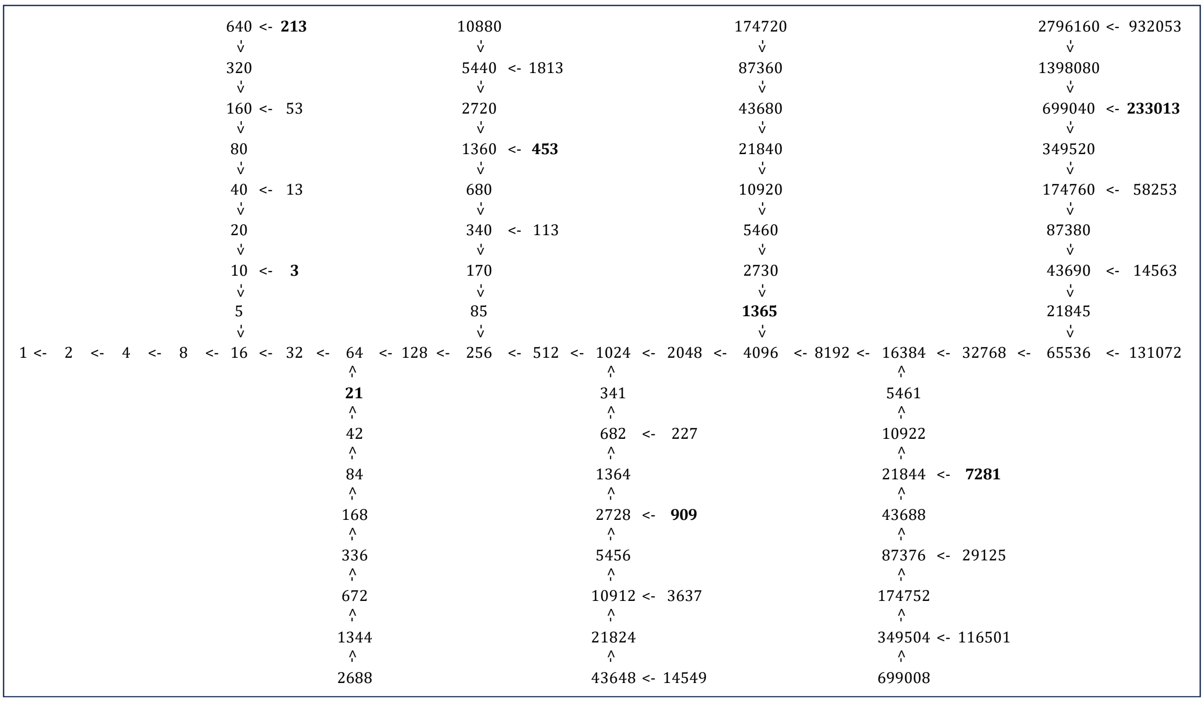

Fig. 1: *Collatz Tree Expansion through eqn. 2*

Criteria for No-branching

Since these values are multiples of 3, some nodes created from the base stem by the first condition of equation 2 result in branches that either fail to branch further or do not meet the first requirement. These are values of 2^{6n} (64^n), such as 64, 4096, 262144, ...

In a comparable manner, branches made from branch value nodes result in branchless branches if the first requirement is met and the new node is a multiple of three.

$$\boldsymbol{Criteria\ for\ no-branching\ node}: mod(node-1,9)\equiv 0$$

Consecutive Branches and Relationship

When consecutive branches are taken into consideration, it is observed that values of consecutive branches are related to each other. These are vertical and diagonal relations. All three relations are applicable on any two consecutive branches over the Collatz Tree.

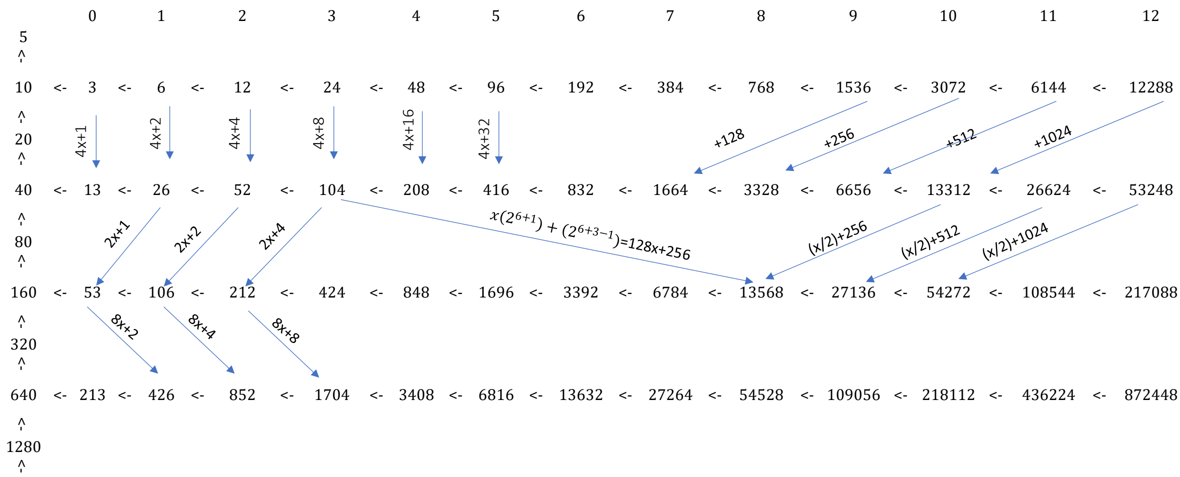

Fig. 2: Example for Fig. 2

Vertical relations in the consecutive branches are direct, its multiple of 4 and addition of power of 2, power is the location of value. If it is node, i = 0. Diagonal relationship is of two kinds left to right and right to left. Here number of terms between two values are taken as n, including the location of the terms also. For rightward diagonal, it is multiple and for leftward it is division by power of 2 along with addition of power of two either based on location or with location and number of terms.

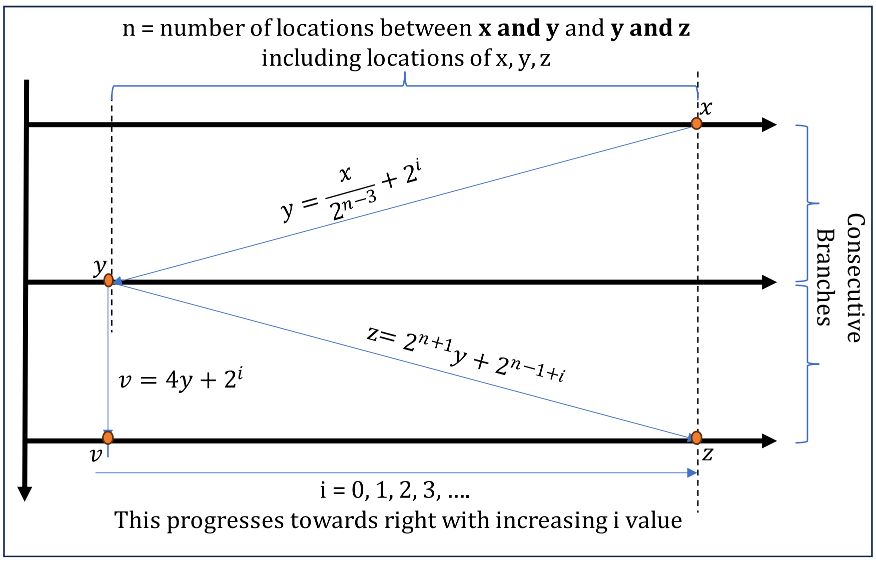

Fig. 2: Diagonal and Vertical Relations

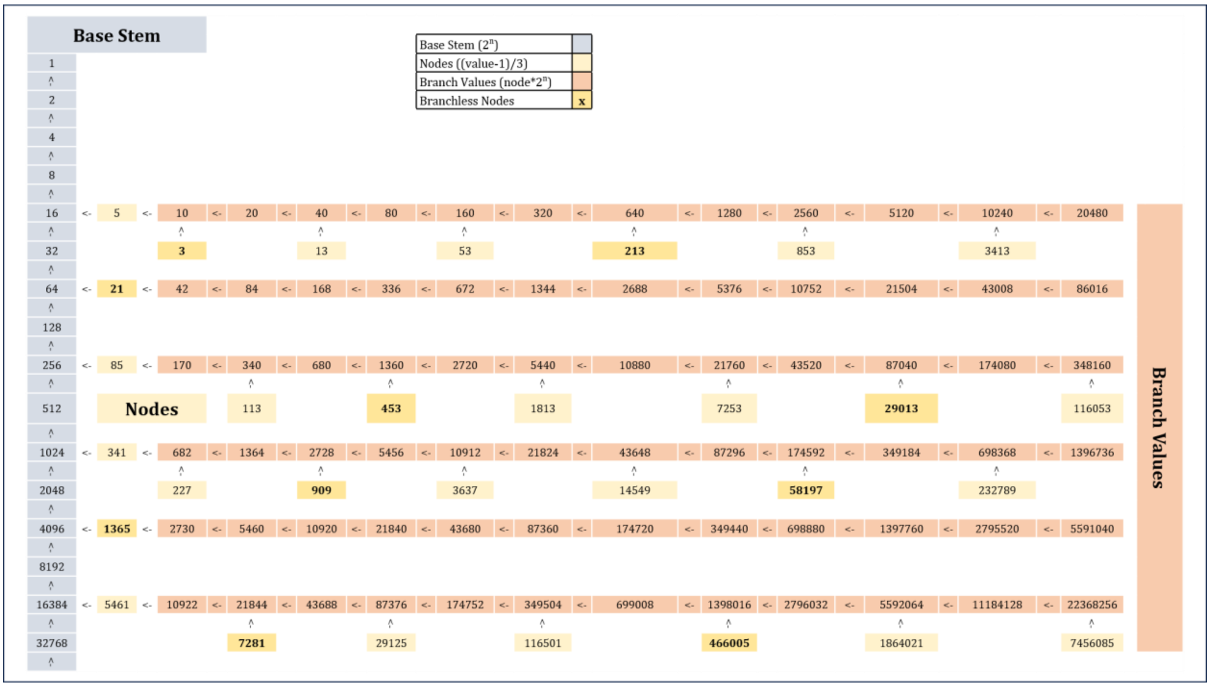

Fig. 4: Collatz Tree Structure

Generalised Term for Node Values

Derivation

For base stem, values that have branching can be given as follows:

$$R = \begin{cases} 2^{6n-2} \\ 2^{6n+2} \end{cases}, \; n \geq 1, \boldsymbol{Node} = \frac{R-1}{3} \; or \; \frac{Branch\,Value-1}{3}$$

Here 2^{6n} values are not taken since these values create branches with no branching. For branches, it can be said that they cause branching when an odd value is obtained.

$$\frac{R-1}{3} * 2^n \; or \; \frac{Branch\,Value-1}{3} * 2^n$$

Again, Nodes are at zeroth position. The first value of branch is node*2, and so on given by node*2^n, where n is position of each branch value. So, a new node is to be created from m_1 position of branch of node $\frac{R-1}{3}$. This node can be expressed as:

$$\frac{\left(\frac{R-1}{3} * 2^{m_1} - 1\right)}{3} = \frac{(R-1)2^{m_1} - 3}{3^2} = \frac{(R-1)2^{m_1}}{3^2} - \frac{1}{3}$$

Now for m_2 position of branch of this node, new node value can be expressed as:

$$\frac{\frac{\left(\frac{R-1}{3} * 2^{m_1} - 1\right)}{3} * 2^{m_2} - 1}{3} = \frac{\left((R-1)2^{m_1} - 3\right)2^{m_2} - 3^2}{3^3} = \frac{(R-1)}{3^3} 2^{m_1+m_2} - \frac{2^{m_2}}{3^2} - \frac{1}{3}$$

For m_3 position of this branch, node value can be expressed as:

$$\cfrac{\cfrac{\cfrac{\left(\frac{R-1}{3} * 2^{m_1} - 1\right)}{3} * 2^{m_2} - 1}{3} 2^{m_3} - 1}{3} = \frac{\left(\left((R-1)2^{m_1} - 3\right)2^{m_2} - 3^2\right)2^{m_3} - 3^3}{3^4}$$

$$= \frac{(R-1)}{3^4} 2^{m_1+m_2+m_3} - \frac{2^{m_2+m_3}}{3^3} - \frac{2^{m_3}}{3^2} - \frac{1}{3}$$

For m_4 position of this branch, node value can be expressed as:

$$\cfrac{\cfrac{\cfrac{\cfrac{\left(\frac{R-1}{3} * 2^{m_1} - 1\right)}{3} * 2^{m_2} - 1}{3} 2^{m_3}}{3} 2^{m_4} - 1}{3}$$

$$= \frac{\left(\left(\left((R-1)2^{m_1} - 3\right)2^{m_2} - 3^2\right)2^{m_3} - 3^3\right)2^{m_4} - 3^4}{3^5}$$

$$= \frac{(R-1)}{3^5} 2^{m_1+m_2+m_3+m_4} - \frac{2^{m_2+m_3+m_4}}{3^4} - \frac{2^{m_3+m_4}}{3^3} - \frac{2^{m_4}}{3^2} - \frac{1}{3}$$

A generalisation can be made using this division ladder:

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\left(\frac{R-1}{3}\right)2^{m_1} - 1}{3} 2^{m_2} - 1}{3} 2^{m_3} - 1}{3} 2^{m_4} - 1}{3} - 1}{3} \cdots\cdots\cdots 2^{m_k} - 1}{\begin{matrix}\ddots \\ \ddots \\ 3\end{matrix}}$$

-eqn. 6

After simplification, following expression is obtained:

$$\frac{\left(\cdots\left(\left(\left(\left((R-1)2^{m_1} - 3\right)2^{m_2} - 3\right)2^{m_3} - 3\right)2^{m_4} - 3\right)2^{m_5} - \cdots\cdots\right)2^{m_k} - 3^k}{3^{k+1}}$$

$$= \frac{1}{3^k + 1}\left((R-1)2^{\sum_{i=1}^{k} m_i} + 3 * 2^{\sum_{i=2}^{k} m_i} - \sum_{i=1}^{k}\left(3^i * 2^{\sum_{j=i}^{k} m_{j+2}}\right) - 3^k\right)$$

Following is the complete equation with summation symbol:

$$T(k, \{m_i\}) = \left(\frac{(R-1)}{3} 2^{m_1} - 1\right)\frac{2^{\sum_{i=2}^{k} m_i}}{3^k} - \sum_{i=1}^{k-2} \frac{2^{\sum_{j=i}^{k} m_{j+2}}}{3^{k-i}} - \frac{1}{3}$$

-eqn. 7

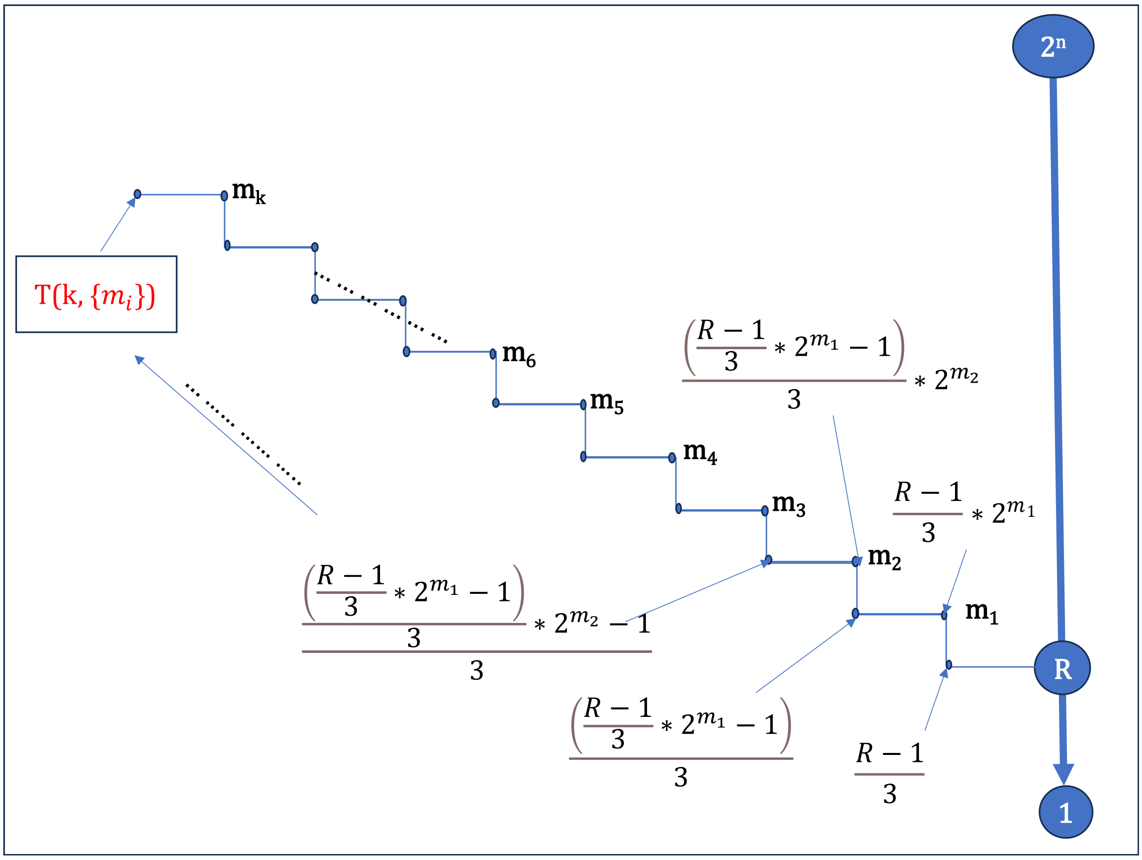

Fig. 5: Node Positions w.r.t. eqn. 6

Here:

- k = number of branches to be traversed for the node value
- $\{m_i\}$ = set of values of positions of nodes in each branch for traversing (total k values)

Stopping Time

Figure 3 and Equation 6 demonstrate the relationship between stopping time and node position. Each node has a placement or position of 0. However, each of these distinct nodes must be counted as one. Therefore, the number of nodes traversed is part of the stopping time. The nodes traversed, or the collection of sites {mi}, are also included in the stopping time. Since R is in the form of 2n and only requires n steps, the number of steps is precisely log2R. The last remaining step is the node itself.

Following is the expression for stopping time:

$$stopping\ time = \sum_{i=1}^{k} m_i + k + \log_2 R + 1 = \sum_{i=1}^{k}(m_i + 1) + \log_2 R + 1$$

$$= \sum_{i=1}^{k}(m_i + 1) + \log_2 2R$$

The given information is suitable for only odd values since equation 6 (and equation 7) produce odd values.

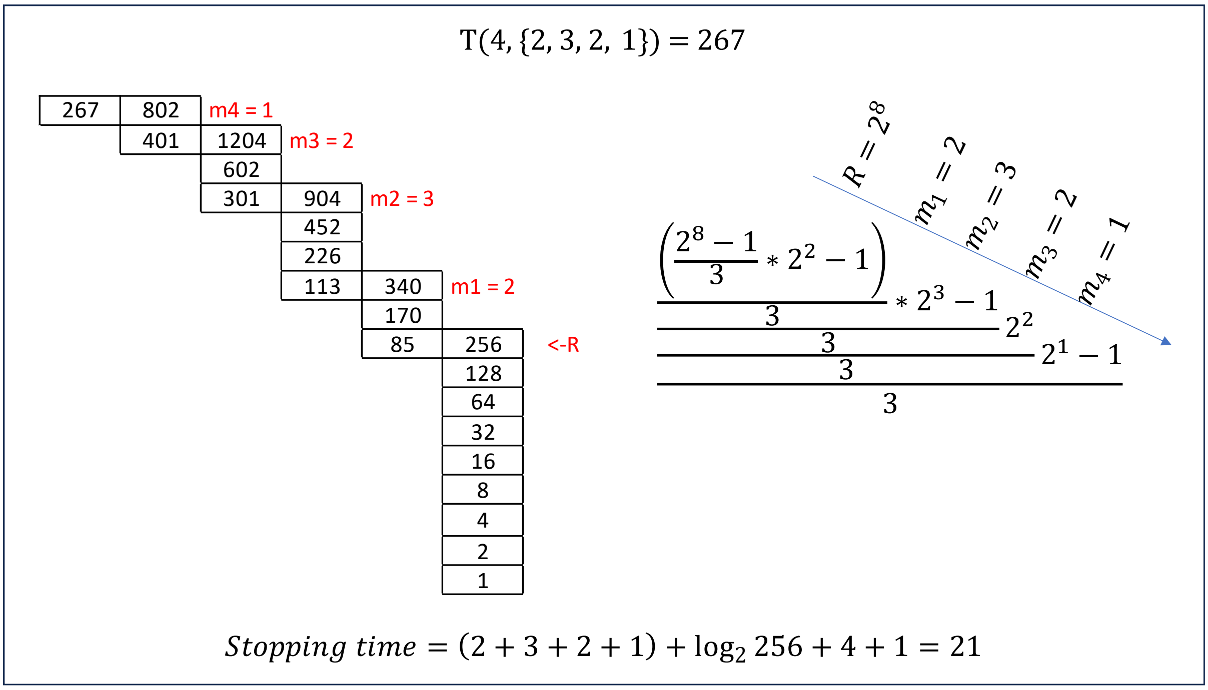

Fig. 6: Explanatory example of eqn. 6

Conclusion

Following are the conclusions of the above:

1. The values starting from 4 give nodes on every alternating value. $R = 2^{2n+2}, n \geq 1$.
2. A values 2^{4n} or powers of 2 which are multiples of 4 (i.e., 16^n) leads to branch which is a multiple of 5 ($\frac{2^{4n}-1}{3}$ is a multiple of 5).
3. The values that are 2^{6n-2} and 2^{6n+2} form branches that further branch on odd and even nodes.
4. The tree's structure can be visualized starting from the base stem, which follows the powers of 2, and the branches follow a similar pattern.
5. The position pattern is either odd or even, but the nodes for even, odd, and no-branching branches appear randomly.
6. When the values $\frac{2^{n}-1}{3}$ are divisible by 3, they lead to values or form branches which do not branch further. These are the 2^{6n} (64^n) values of the base stem.
7. The criteria for no branching: $mod(node - 1,9) \equiv 0$
8. The base stem and branch values are even (except for 1), whereas the node values are odd.
9. The generalization of node values can be designed with a root term that follows a known pattern, while the branches have random values. The formula in the division ladder is in its simplest form.
10. The stopping time of any odd value can be calculated with the positions, number of nodes and power of 2 in the base stem.

Reference

1. Collatz conjecture, https://en.wikipedia.org/wiki/Collatz_conjecture
2. The Ultimate Challenge: The 3x+1 Problem; Jeffery C. Lagarias
3. On the existence of cycles of given length in integer sequences like $x_{n+1} = x_n/2$ if x_n even, and $x_{n+1} = 3x_n + 1$; Corrado Böhm, Giovanna Sontacchi