Lower Bounds for the Total Stopping Time of $3x + 1$ Iterates Revisited

Kazuyuki Amano[*]

October 2021

Abstract

Let $T : \mathbb{N} \to \mathbb{N}$ be the $3x+1$ function defined by $T(n) = n/2$ if n is even and $T(n) = (3n+1)/2$ if n is odd. Let $\sigma_\infty(n)$ be the minimal k such that $T^{(k)}(n) = 1$ if one exists and $+\infty$ otherwise. By extending the computational efforts by Applegate and Lagarias (Math. Comp., 2003), we show that $\sigma_\infty(n) \geq (\frac{1}{2}\ln\frac{4}{3})^{-1}\ln n \approx 6.9521 \ln n$ for infinitely many n, improving the former bound of $\sigma_\infty(n) > 6.1413 \ln n$. The certificate of our proof consists of $19,065,883,794$ critical paths each for a different $3x+1$ tree of max-depth 74.

1 Introduction

Let $T : \mathbb{N} \to \mathbb{N}$ be the function defined by

$$T(n) \quad = \quad \begin{cases} \dfrac{n}{2}, & \text{if } n \equiv 0 \pmod 2, \\[2ex] \dfrac{3n+1}{2}, & \text{if } n \equiv 1 \pmod 2. \end{cases}$$

The Collatz conjecture (or $3x+1$ conjecture) asserts that, for every $n \geq 1$, there exists some iterate k such that $T^{(k)}(n) = 1$. At present, the conjecture remains unsolved. See e.g., Lagarias [4] and Chamberland [3] for an overview of the conjecture, and also Tao [7] for a recent exciting development.

This work focuses on the *stopping time* of n, which is defined by the minimal k such that $T^{(k)}(n) = 1$ and is denoted by $\sigma_\infty(n)$. If no such k exists, then we let $\sigma_\infty(n) = +\infty$. A *stopping time ratio* $\gamma(n)$ of n is defined by

$$\gamma(n) := \frac{\sigma_\infty(n)}{\ln n}.$$

If we assume that the parity of $T^{(1)}(n), T^{(2)}(n), \ldots, T^{(k)}(n)$ behaves like a random bit, then the process of iterating the function $T(\cdot)$ can be modeled as a random walk starting at the position $\ln n$ and making a step of length $\ln(1/2)$ or (approximately) $\ln(3/2)$ with probability $1/2$. This leads to the conjecture that an expected stopping time ratio should be $\gamma_0 := \left(\frac{1}{2}\ln\frac{4}{3}\right)^{-1} \approx 6.95212$. See e.g., [2, 5] for more detailed analysis of such stochastic models.

[*]Gunma University, amano@gunma-u.ac.jp

In [1], Applegate and Lagarias introduced a way for giving an infinite sequence of n such that $\gamma(n)$ is large. Then, by conducting a massive computation, they obtained a certificate showing that there is an infinite set of positive integers n such that $\gamma(n)$ is finite and

$$\gamma(n) \geq \frac{29}{29\ln 2 - 14\ln 3} \approx 6.141316. \tag{1}$$

The certificate of this lower bound is huge. It consists of 350,688,758 critical paths each for a different $3x+1$ tree of max-depth 60. (Some of the terminologies will be explained in Section 2.) However, this is not enough for proving a lower bound of $\gamma(n) \geq \gamma_0$ for infinitely many n, which is certainly satisfied in view of the random walk model. Lagarias [4, Section 8] said that the situation that we do not have a proof for the γ_0 lower bound is "scandalous".

This work is devoted to improving the lower bound of (1) to $\gamma_0 \approx 6.95212$ by extending their efforts two decades ago. We apply their framework in a parallel fashion together with some pruning techniques. After a computation of an about one CPU year, we finally obtain a certificate proving the desired bound of $\gamma(n) \geq \gamma_0$ (for infinitely many n). The certificate consists of $19,065,883,794$ critical paths each for a different $3x+1$ tree of max-depth 74.

The rest of this paper is organized as follows. In Section 2, we briefly sketch the framework for deriving a lower bound on $\gamma(n)$ developed by Applegate and Lagarias [1]. Then, in Section 3, we describe our computational efforts for deriving an improved bound. Finally, we make some concluding remarks in Section 4.

2 Certificate for Large Stopping Time

In order to prove $\sigma_\infty(n) \geq \gamma_0 \ln n$ for infinitely many n, we essentially follow the framework developed by Applegate and Lagarias [1]. Below we briefly sketch an idea in a somewhat intuitive way. See [1] for a clearer and more detailed exposition. In the following, the "threshold" value α appeared in Section 3 of [1] is fixed to 1/2 since our aim is to obtain a certificate of $\gamma(n) \geq \gamma_0$.

We consider a "prunned" inverse map $T^{*(-1)}$ on the set of positive integers $n \not\equiv 0 \pmod 3$ given by

$$T^{*(-1)}(n) = \begin{cases} \{2n\}, & \text{if } n \equiv 1, 4, 5 \text{ or } 7 \pmod 9, \\ \left\{2n, \dfrac{2n-1}{3}\right\}, & \text{if } n \equiv 2 \text{ or } 8 \pmod 9. \end{cases}$$

The word "prunned" represents that we discard integers with $n \equiv 0 \pmod 3$ since they will only produce a linear chain of integers with $n' \equiv 0 \pmod 3$ (see [5]).

Given an integer $n \in \mathbb{N}$, we can generate a tree such that the root node is n and a node labelled m at depth d of the tree is connected to a node labelled $T^{*(-1)}(m)$ of depth $d+1$. Then, we label each edge by 1 if an integer labelled at the sink of the edge is odd, and by 0 otherwise. The *weight* $w(n)$ is defined by the number of edges that are labeled 1 on the path between the root and the leaf labeled n, and the *ones-ratio* $\rho(n)$ of this path is defined by $w(n)$ divided by the length of the path.

By recalling the definition of the function T, it is easy to see that

$$\gamma(n) \geq \frac{1}{\ln 2 - \rho(n)\ln 3}.$$

Hence, in order to obtain a good lower bound on $\gamma(n)$, our task is to obtain a good lower bound on $\rho(n)$. A bound of $\rho(n) \geq 1/2$ corresponds to the bound of $\gamma(n) \geq \gamma_0$. Roughly speaking, what we will see is that, there is an infinitely long path satisfying that the path can be divided into chunks of small length such that the ones-ratio of each chunk is at least 1/2.

A path from a leaf to the root in a tree truncated at some depth is called a *critical path* if the ones-ratio of the path is at least 1/2. The key observation is that, if $a \equiv b \pmod{3^{l+1}}$, the tree with root a and the one with root b are isomorphic (including the labels of edges) if we truncate the trees at depth k at which it has a critical path of weight l for the first time. For example, the following is a critical path of weight $l = 3$ starting from 5.

$$5 \xrightarrow{0} 10 \xrightarrow{0} 20 \xrightarrow{1} 13 \xrightarrow{0} 26 \xrightarrow{1} 17 \xrightarrow{1} 11.$$

One can see that every integer a such that $a \equiv 5 \pmod{3^4}$ has an isomorphic path starting from a.

By these arguments, our task can be reduced to find a certificate consists of (i) a finite list $\mathcal{C}$ of residue classes $a \pmod{3^{l+1}}$ for various length $l = l(a)$ which forms a ternaly prefix code, and (ii) a list of critical paths in the pruned tree with root a, over all a in $\mathcal{C}$. Once we have such a certificate, we can make a path consists of an infinite number of chunks such that each chunk has ones-ratio at least 1/2 by chaining critical paths in an appropriate way. See Section 3 in [1] for a formal definition of a certificate and a proof on why it yields a lower bound.

3 Finding a Certificate

For searching a certificate, we apply a greedy algorithm developed by Applegate and Lagarias [1, Section 4]. For a given $a \pmod{3^{l+1}}$, we generate a tree for the map $T^{*(-1)}(\cdot)$ starting from a in the breadth first manner to find a critical path of length $2l$. If no such path exists, we add $3a, 3a+1, 3a+2$ to the list of open vectors with l incremented by one. However, a naive approach seems infeasible since the maximum depth of a tree that needed to be enumerated has been estimated around 72 to 76 [1], and the expected number of leaves of a tree of depth k is the order of $(4/3)^k$ [5, Section 3].

In order to reduce the computation time, we conduct a search in parallel and also introduce some pruning techniques. First, we compute a list of critical paths for small values of l (say, $l \leq 14$), and keep the list of open vectors for $l = 15$. The list has 2,563,281 entries. Then, we distribute these vectors to many threads and continue the computation in parallel.

For pruning, we made look-up tables for a possible maximum "gain" of various starting integers, i.e., the maximum value of the number of ones minus the number of zeros of the edge labels over all paths in a tree of some fixed depth. This works since a tree of depth d can be fully characterized by its root modulo 3^{d+1}. In our computation, we used tables for various depths up to 15 to cut off unnecessary nodes efficiently. This speeds up the computation by roughly 10 times.

We run the code using three standard PCs. The whole computation took about two weeks using maximum of about 40 threads. This means that if we run our code on a single thread machine, then the computation takes about one year. Finally, we obtain the desired certificate consists of $19,065,883,794$ critical paths each for a different $3x+1$ tree. The max-depth of a tree that needed to be enumerated, which is equal to the max-length of a critical path, is 74.

This establishes the following theorem.

Theorem 1 *There is an infinite set of positive integers n such that $\gamma(n)$ is finite and $\gamma(n) \geq \left(\frac{1}{2}\ln\frac{4}{3}\right)^{-1} \approx 6.95212$.* □

A part of the certificate is shown in Table 1, and the breakdown of the size of the certificate with respect to l is shown in Table 2. The code that we used for finding them together with some data is available at `https://gitlab.com/KazAmano/collatz`.

Table 1: Certificate for proving $\gamma(n) \geq \gamma_0$ for infinitely many n. We only show critical paths for $l \leq 5$ and the last three for $l = 37$ (of length $2l = 74$). The columns represent $a \pmod{3^{l+1}}$ (in reverse ternaly and then in decimal), l, critical path and an integer at the terminal of the critical path when starting from a, respectively.

a	l	crit.path	terminal
10 (1)	1	01	1
20 (2)	1	1	1
11 (4)	1	01	5
22 (8)	1	1	5
2100 (5)	3	001011	11
2102 (59)	3	001011	139
21100 (14)	4	00101011	43
21101 (95)	4	00101011	299
12111 (123)	4	00010111	391
12112 (205)	4	00010111	647
210110 (113)	5	0000111011	475
210111 (356)	5	0000111011	1499
211101 (284)	5	0010101011	1195
211102 (527)	5	0010101011	2219
212110 (131)	5	0000110111	551
212112 (617)	5	0000110111	2599
120021 (412)	5	0001010111	1735
120022 (655)	5	0001010111	2759
121210 (151)	5	0001011011	635
121212 (637)	5	0001011011	2683

$\cdots$ 19,065,883,771 lines are omitted due to the space restriction. $\cdots$

a	l	crit.path	terminal
12222201121111021122111012022122102010	37	000100010110001101011000101100100100 1→	7848063225965413510427
(187080808736798062)		→101100000011011100100111010111111011	
12222201121111021122111012022122102011	37	0001000101100011010100100110001100000→	2673752915744399436535 9
(637364714627795425)		→101101100111100110101110111100010111 1	
12222201121111021122111012022122102012	37	0001000101100011010110001011001001001→	4562699508892257521999 5
(1087648620518792788)		→1011000000110111001001111010111111011	

Table 2: The breakdown of the certificate with respect to the value of l.

l	# of crit.paths	l	# of crit.paths	l	# of crit.paths
1	4	14	159,336	27	3,196,943,303
2	0	15	453,340	28	3,416,750,430
3	2	16	1,260,293	29	2,942,079,419
4	4	17	3,401,901	30	1,948,191,245
5	10	18	8,925,553	31	943,786,096
6	28	19	22,496,469	32	315,669,869
7	82	20	54,348,400	33	67,543,670
8	242	21	124,688,261	34	8,465,899
9	714	22	269,621,335	35	563,978
10	2,119	23	542,983,077	36	16,852
11	6,298	24	1,006,813,061	37	186
12	18,660	25	1,688,508,366	-----	------------
13	54,992	26	2,502,130,300	Total	19,065,883,794

4 Concluding Remarks

In this work, we give a certificate of the proof that $\gamma(n) \geq \gamma_0 \approx 6.95212$ infinitely often, which corresponds to an expected hitting time of the random walk model. By using the Chernoff Bounds, we can see that such a model predicts that $\limsup_{n\to\infty} \gamma(n)$ is a certain constant about 41.678 (see [5] for the proof). It would be challenging to find an integer n such that $\gamma(n)$ is larger than this limit (if one exists). To the best of our knowledge, the current record is $n = 7,219,136,416,377,236,271,195$ with $\gamma(n) \approx 36.7619$ [6, $3x + 1$ Completeness and Gamma Records].

Acknowledgements

The author would like to thank Kento Kimura for setting up the remote computer servers that are quite useful under the pandemic. This work was supported in part by JSPS Kakenhi No. JP21K19758, JP18K11152 and JP18H04090.

References

[1] D. Applegate and J. C. Lagarias, Lower bounds for the total stopping time of $3x + 1$ iterates, Math. Comp. **72** (2003) 1035–1049.

[2] K. Borovkov and D. Pfeifer, Estimates for the Syracuse problem via a probabilistic model, Theory Probab. Appl. **45** (2000) 300–310.

[3] M. Chamberland, A $3x + 1$ survey: number theory and dynamical systems, in *The ultimate challenge: the* $3x+1$ *problem*, Amer. Math. Soc., Providence, RI (2010) 57–78.

[4] J. C. Lagarias, The $3x+1$ problem: An overview, in *The ultimate challenge: The* $3x+1$ *Problem*, Amer. Math. Society, Providence, RI (2010) 3–29.

[5] J. C. Lagarias and A. Weiss, The $3x + 1$ problem: Two stochastic models, Annals of Applied Probability, **2(1)** (1992) 229–261.

[6] E. Roosendaal, On the $3x + 1$ problem, `http://www.ericr.nl/wondrous` (accessed 2021.9.2).

[7] T. Tao, Almost all orbits of the Collatz map attain almost bounded values, `arXiv:1909.03562` (2019).