Digging for Diamonds in the Mines of Collatz

Hector J. Levesque
Dept. of Computer Science
University of Toronto

November 17, 2022

The *Collatz Conjecture*, sometimes called the $3x + 1$ problem [2], is one that appeals immediately to math hobbyists. Unlike other famous conjectures such as Goldbach or *P vs NP*, Collatz does not mention prime numbers, Turing machines, or any other rich topics. It thus lends itself to explorations using basic mathematics (essentially high-school algebra), as will be done here. While there is evidence that the conjecture will only be cracked by new techniques not yet in the professional mathematician's toolkit (see the 2011 blog by Fields Medal winner Terence Tao), we amateurs can still discover tantalizing things about it using more elementary means.

We begin with the presentation of the problem used by Terras [3]. Define the function T from $\mathbb{N}^+$ to $\mathbb{N}^+$ as follows:

$$T(n) = \begin{cases} n/2, & \text{if } n \text{ is even} \\ (3n+1)/2, & \text{if } n \text{ is odd} \end{cases}$$

We say that n *evolves* to m in k steps iff $T^k(n) = m$. A step is called *even* or *odd* according to whether the argument to T at that point is even or odd. The Collatz Conjecture is that every integer $n > 1$ evolves to 1. Terras uses the function $\chi(n)$, the *stopping time* of n, defined as the smallest $k \in \mathbb{N}^+$ such that $T^k(n) < n$. (If no such integer exists, set $\chi(n) = \infty$.) It is not hard to see that the Collatz Conjecture is true iff $\chi(n)$ is finite for all $n > 1$. The conjecture has been verified by computer to hold for all $1 < n < 2^{68}$ [1]. In this paper, we develop a variant way of confirming the conjecture up to some large bound N in terms of what we call *diamonds*.

The fundamental property used throughout is that once we know about the evolution of an integer, we immediately know the evolution of infinitely many larger ones:

Theorem 1: *Suppose r evolves to s after k steps of which m are odd. Then for every $b \in \mathbb{N}$, the integer $a = 2^k b + r$ evolves to $c = 3^m b + s$ with the same sequence of steps.*

Proof: By induction on k. If $k = 0$ then $r = s$ and the claim clearly holds. Otherwise, suppose $k > 0$, in which case $2^k b$ is even. There are two cases:

- If r is even, then it evolves in one step to $r' = r/2$ which evolves to s after $(k-1)$ steps of which m are odd. Then $a = 2^k b + r$ is also even and evolves in one step to $2^{k-1}b + r'$ which, by induction, evolves to c with the same steps as from r' to s. So a evolves to c with the same steps as from r to s.

- If r is odd, then it evolves in one step to $r' = (3r+1)/2$ which evolves to s after $(k-1)$ steps of which $(m-1)$ are odd. Then $a = 2^k b + r$ is also odd and evolves in one step to $[3(2^k b + r) + 1]/2 = 2^{k-1}3b + r'$ which, by induction, then evolves to $3^{m-1}3b + s = c$ with the same steps as from r' to s. So a evolves to c with the same steps as from r to s. ∎

Lemma 2: *Suppose r evolves to s after k steps of which m are odd. Then $s/r \geq 3^m/2^k$.*

Proof: By induction on k. If $k = 0$, then $r = s$ and the claim clearly holds. Otherwise, suppose there is a least one step, from r to some r' which then evolves to s. There are two cases:

- If r is even, then $r' = r/2$ and r' evolves to s after $(k-1)$ steps of which m are odd. By induction, $s/r' \geq 3^m/2^{k-1}$. But then we have that
$$\frac{s}{r} = \frac{s}{2r'} \geq \frac{3^m}{2 \cdot 2^{k-1}} = \frac{3^m}{2^k}.$$
- If r is odd, then $r' = (3r+1)/2$ and r' evolves to s after $(k-1)$ steps of which $(m-1)$ are odd. By induction, $s/r' \geq 3^{m-1}/2^{k-1}$. But then we have that
$$\frac{s}{r} = \frac{3s}{3r} > \frac{3s}{3r+1} = \frac{3s}{2r'} \geq \frac{3 \cdot 3^{m-1}}{2 \cdot 2^{k-1}} = \frac{3^m}{2^k}. \quad ∎$$

Corollary 3: *Suppose r, s, a, c are as in the theorem. If $r > s$ then $a > c$.*

Proof: If $r > s$ then $1 > s/r$ and so $1 > 3^m/2^k$ by Lemma 2. It then follows that $2^k > 3^m$ and therefore that $2^k b + r > 3^m b + s$. ∎

The converse to this corollary might appear to hold as well, but there are some (rare) counter-examples. Here is the smallest one: the integer $r = 7$ evolves to $s = 8$ in $k = 8$ steps with $m = 5$ odd steps. Consequently, $a = 256 + 7$ evolves to $c = 243 + 8$ with the same sequence of steps. As can be seen in this case, $a > c$ but $r < s$.

Let now us a define an odd integer $p > 1$ to be *χ-prime* iff p does not evolve to an odd integer n with $1 < n < p$. So for example, 3 and 5 are χ-prime, but 7 is not, since 7 evolves to 5. Intuitively, an integer is χ-prime if it evolves through a sequence of larger integers until it hits a power of 2 (or continues forever). The reason we care about χ-primes is that they seem to be rare. It is easy to check with a computer that there are only 34 of them less than 2^{40}. Moreover they are "sufficient" in the following sense:

Lemma 4: *Let N be any positive integer. Suppose that every χ-prime $p < N$ has finite stopping time. Then every integer a where $1 < a < N$ has finite stopping time.*

Proof: We prove this by induction on a. For $a = 2$, $\chi(a) = 1$. Now suppose that $2 < a < N$. If a is even, then $1 < a/2 < a$. By induction, $a/2$ has finite stopping time and so a does too. If a is χ-prime, then a is given to have finite stopping time. If a is odd but not χ-prime, then a evolves to some odd b such that $1 < b < a$. By induction, b has finite stopping time, so a does too. ∎

So to confirm that the Collatz conjecture holds up to some bound N, it is sufficient to confirm that the χ-primes up to N have finite stopping time. (Note that although χ-primes are rare, with a bit of algebra, it can be shown that there are infinitely many of them. See the Appendix.)

We can go further to a special sort of χ-prime. Let us call a positive integer q *basic* iff there is no χ-prime $p > 3$ where for some positive integers u and v, $q = 2^u v + p$ where $p < 2^u$. So all even numbers are basic (since $2^u v$ is even and p is odd). Among the odd numbers, 5 and 7 are basic, but $13 = 2^3 + 5$ and $21 = 2^4 + 5$ are not basic. (Note that 5 and 21 are χ-prime, but 7 and 13 are not.) Finally, let us define a *diamond* to be any integer that is both χ-prime and basic.

These numbers are appropriately named in that they seem to be *very* rare. There are only 8 diamonds between 5 and 2^{40}:

5, 75, 151, 227, 184111, 276167, 13256071, and 26512143.

It is not even clear that there are infinitely many of them. (Observe that there are none at all between 2^{25} and 2^{40}.) Nonetheless, they have the following property:

Lemma 5: *Let N be any positive integer. Suppose that every diamond p where $3 < p < N$ satisfies $\chi(p) < \log_2(p)$. Then every χ-prime p where $3 < p < N$ satisfies $\chi(p) < \log_2(p)$.*

Proof: The proof is by by induction on the χ-primes. When $p = 5$, we have $\chi(5) = 2$. If $5 < p < N$, then there are two cases. If p is a diamond, then p is given to satisfy $\chi(p) < \log_2(p)$. Otherwise, p is not basic, and so there is a χ-prime $p' > 3$, and positive integers u and v such that $p = 2^u v + p'$, where $p' < 2^u$. Since $3 < p' < p$, p' satisfies $\chi(p') < \log_2(p')$ by induction. Let $k = \chi(p')$. Then $k < \log_2(p') < u < \log_2(p)$. So p' evolves to some $q' < p'$ in k steps of which some m are odd. By Theorem 1, for any positive integer b, $2^k b + p'$ also evolves to $3^m b + q'$ in the same steps. Now since $k < u$, let $b = 2^{u-k} v$, so that $p = 2^k b + p'$ and let $q = 3^m b + q'$. So p evolves to q in k steps and $q < p$ by Corollary 3. So $\chi(p) \leq k < \log_2(p)$. ∎

So from Lemmas 4 and 5, to confirm that the Collatz conjecture holds up to some bound N, it is sufficient to confirm that every diamond p where $3 < p < N$ has logarithmic stopping time. In other words, if someone were able to prove that *all* diamonds had logarithmic stopping time, this would completely solve the Collatz problem. Here is the conjecture for the record:

Diamond Conjecture: *Every diamond $p > 3$ satisfies $\chi(p) < \log_2(p)$.*

Note that this is different from the Collatz conjecture: it could be false even if Collatz is true. (And recall that there are integers that are known not to have small stopping times. For example, $\chi(27) = 59$.) At any rate, it is easy to check that the 8 numbers above do in fact have small stopping times, and therefore that the Collatz conjecture holds up to 2^{40} at least.

So this is where we conclude our study. In testing the Collatz conjecture up to some large bound like $N = 2^{80}$ say, we are suggesting a slightly different way to proceed. Instead of looking for integers with no stopping time, we can instead go digging for diamonds beyond the 8 listed above. For future work, there are some interesting patterns among the 8 diamonds above to follow up on.

For one thing, there are close pairs, x and $2x+1$, that occur twice: for $x = 75$ and $x = 13256071$. There are even closer pairs, x and $(3x+1)/2$, that also occur twice: for $x = 151$ and $x = 184111$. Do either of these patterns persist? Do they occur infinitely often? Another curious fact is that 5 of the 8 diamonds are actually primes in the traditional sense (all but 75, 276167, and 26512143). A coincidence? Of course, if someone could show that there were only finitely many diamonds in total, the Collatz Conjecture might then be confirmed to hold for all $n > 1$ in a finite way.

References

[1] David, Barina. Convergence verification of the Collatz problem. *The Journal of Supercomputing*, **77** (2021): 2681–2688.

[2] Jeffrey Lagarias (editor). *The Ultimate Challenge: The 3x+1 Problem*, American Mathematical Society, 2010, 189–207.

[3] Rino Terras. A stopping time problem on the positive integers. *Acta Arithmetica*, **XXX** (1976): 241–252.

Appendix

In this appendix, we prove that there are infinitely many χ-prime numbers.

Lemma 6: *For every $n \geq 2$, the number $2^n - 1$ evolves to $3^n - 1$ in n steps all of which are odd.*

Proof: The proof is by induction on n. For $n = 2$, we have that 3 evolves to 5 to 8. For $n > 2$, assume that $2^n - 1$ evolves to $3^n - 1$ in n steps all of which are odd. By Theorem 1, when $b = 1$, $2^n + 2^n - 1$ evolves to $3^n + 3^n - 1$ in n steps all of which are odd. So $2^{n+1} - 1$ evolves to $2 \cdot 3^n - 1$ in n steps all of which are odd. The number $2 \cdot 3^n - 1$ evolves in one odd step to $3 \cdot 3^n - 1 = 3^{n+1} - 1$. So $2^{n+1} - 1$ evolves to $3^{n+1} - 1$ in $(n+1)$ steps all of which are odd. ∎

Corollary 7: *For every $n \geq 2$ and for every $b \in \mathbb{N}^+$, the number $2^n b - 1$ evolves to $3^n b - 1$ in n steps all of which are odd.*

Proof: Immediate from Lemma 6, Theorem 1, and the facts that $2^n b + 2^n - 1 = 2^n(b+1) - 1$, and $3^n b + 3^n - 1 = 3^n(b+1) - 1$. ∎

Lemma 8: *For every $n \in \mathbb{N}^+$, $2^{3^{n-1}} \equiv -1 \bmod 3^n$.*

Proof: The proof is by induction on n. For $n = 1$, $2 \equiv -1 \bmod 3$. Now let $v = 3^{n-1}$ and assume that $2^v \equiv -1 \bmod 3^n$. Therefore, for some $d \in \mathbb{N}$, $2^v = 3^n d - 1$. So for some $e \in \mathbb{N}$,

$$2^{3^n} = (2^v)^3 = (3^n d - 1)^3 = (3^n d)^3 - 3(3^n d)^2 + 3(3^n d) - 1 = 3^{n+1} e - 1.$$

Consequently, $2^{3^n} \equiv -1 \bmod 3^{n+1}$. ∎

Theorem 9: *For every $n \geq 2$, the set $\{ x \in \mathbb{N} \mid x \equiv -1 \bmod 2^n \}$ contains infinitely many χ-prime numbers.*

Proof: By Corollary 7, for every $b \in \mathbb{N}^+$, the integer $a = 2^n b - 1$ evolves to $c = 3^n b - 1$ in n steps all of which are odd, and thus, where the numbers up to and including c are all greater than a. Let $v = 3^{n-1}$. By Lemma 8, $2^v \equiv -1 \bmod 3^n$, and so for every $i \in \mathbb{N}$, $2^{v(2i+1)} \equiv -1 \bmod 3^n$. So letting

$$b = \frac{1}{3^n} \cdot [2^{v(2i+1)} + 1],$$

then $b \in \mathbb{N}^+$. Furthermore, for these values of b,

$$c = 3^n b - 1 = 2^{v(2i+1)}.$$

So for these values of b, the integer c is a power of 2, and evolves to 1 passing only through even numbers. Thus, for each choice of i, we get a value of b for which $a = 2^n b - 1$ is χ-prime. ∎

A few observations about the theorem. First, note that the χ-prime numbers generated in the proof, for various choices of n and i, tend to be large. For $n = 3$, the smallest χ-primes (for $i = 0, 1, 2$) are 151 (a diamond), 39768215 (a non-diamond), and 10424999137431 (larger than 2^{40}). For $n = 5$, the smallest χ-prime generated for $i = 0$ is larger than 2^{78}. The next thing to notice is that there are χ-primes beyond those claimed to exist by the theorem, such as $5 \not\equiv -1 \bmod 2^n$ for any $n \geq 2$. Finally, the formula in the proof generates χ-primes that evolve to 1 according to a specific pattern: all the odd steps come before any of the even steps. Of course, not all χ-primes work this way (if they did, this would settle the Collatz Conjecture): the χ-prime 75 (a diamond) does not, and neither does 19884107 (a non-diamond). Nonetheless, what the theorem shows is that there are infinitely many χ-primes that do follow this strict pattern.