Research Article
A New Approach on Proving Collatz Conjecture

Wei Ren [1,2,3]

[1]*School of Computer Science, China University of Geosciences, Wuhan 430074, China*
[2]*Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan, China*
[3]*Guizhou Provincial Key Laboratory of Public Big Data, GuiZhou University, Guizhou, China*

Correspondence should be addressed to Wei Ren; weirencs@cug.edu.cn

Collatz Conjecture (3x+1 problem) states any natural number x will return to 1 after $3*x+1$ computation (when x is odd) and x/2 computation (when x is even). In this paper, we propose a new approach for possibly proving Collatz Conjecture (CC). We propose Reduced Collatz Conjecture (RCC)—any natural number x will return to an integer that is less than x. We prove that RCC is equivalent to CC. For proving RCC, we propose exploring laws of Reduced Collatz Dynamics (RCD), i.e., from a starting integer to the first integer less than the starting integer. RCC can also be stated as follows: RCD of any natural number exists. We prove that RCD is the components of original Collatz dynamics (from a starting integer to 1); i.e., RCD is more primitive and presents better properties. We prove that RCD presents unified structure in terms of $(3*x+1)/2$ and x/2, because $3*x+1$ is always followed by x/2. The number of forthcoming $(3*x+1)/2$ computations can be determined directly by inputting x. We propose an induction method for proving RCC. We also discover that some starting integers present RCD with short lengths no more than 7. Hence, partial natural numbers are proved to guarantee RCC in this paper, e.g., 0 module 2; 1 module 4; 3 module 16; 11 or 23 module 32; 7, 15, or 59 module 128. The future work for proving CC can follow this direction, to prove that RCD of left portion of natural numbers exists.

1. Introduction

The Collatz Conjecture is a mathematical conjecture that is first proposed by Lothar Collatz in 1937. It is also known as the $3x + 1$ conjecture, the Ulam conjecture, the Kakutani's problem, the Thwaites conjecture, or the Syracuse problem [1–3]. "Mathematics may not be ready for such problems", Paul Erdos once speculated about the Collatz Conjecture [4].

The conjecture can be stated simply as follows: take any positive integer number x. If x is even, divide it by 2 to get $x/2$. If x is odd, multiply it by 3 and add 1 to get $3 * x + 1$. Repeat the process again and again. The Collatz Conjecture is that no matter what the number (i.e., x) is taken, the process will always eventually reach 1. The longest progressions for initial starting numbers of less than 10 billion and 100 quadrillion are calculated by Gary T. Leavens [5] and R. E. Crandall [6], respectively. Wei Ren et al. verified $2^{100000} - 1$ can return to 1 after 481603 times of $3 * x + 1$ computation, and 863323 times of $x/2$ computation, which is the largest integer being verified in the world [7]. So far no one has tried to figure out whether

all of the positive numbers eventually reach one, but we know that most of them do so. In particular, Krasikov and Lagarias proved that the number of integers finally reaching one in the interval $[1, x]$ is at least proportional to x 0.84 [8].

The paper is organized as follows: Section 2 presents our Reduced Collatz Conjecture. A mathematical induction for proving Collatz Conjecture is proposed in Section 3. Section 4 introduces $CODE(x)$ for representing reduced dynamics of x and explores its properties. Section 5 studies all starting numbers whose lengths of reduced dynamics are no more than 7. Finally, Section 6 concludes the paper.

Notations

(1) $\mathbb{Z}$: the set of integers.

(2) $\mathbb{N} = \{a \mid a \in \mathbb{Z}, a \geq 0\}$.

(3) $\mathbb{N}^* = \mathbb{N} \setminus \{0\} = \{a \mid a \in \mathbb{Z}, a \geq 1\}$.

(4) $[1]_2 = \{x \mid x \equiv 1 \bmod 2, x \in \mathbb{N}^*\}$.

(5) $[0]_2 = \{x \mid x \equiv 0 \bmod 2, x \in \mathbb{N}^*\}$.

(6) $[i]_m = \{x \mid x \equiv i \bmod m, x \in \mathbb{N}^*, m \geq 2, m \in \mathbb{N}^*, 0 \leq i \leq m - 1, i \in \mathbb{N}\}$.

(7) $CT, CT(\cdot)$: Collatz transformation. $CT(x) = TPO(x)$ or $H(x)$ according to the parity of current inputting x; $CT(x) = I(x)$ or $O(x)$ according to the parity of current inputting x.

(8) $TPO, TPO(\cdot)$: $TPO(x) = 3 * x + 1$ when $x \in [1]_2$.

(9) $H, H(\cdot)$: $H(x) = x/2$ when $x \in [0]_2$.

(10) CC: Collatz Conjecture.

(11) RCC: Reduced Collatz Conjecture.

(12) $x \in \mathscr{RTN}$: $x \in \mathbb{N}^*$ is Returnable.

(13) $I, O, I(\cdot), O(\cdot)$: $I(x) = H(TPO(x)) = (3 * x + 1)/2$, $O(x) = H(x) = x/2$, $I^0(x) = x$, $I^n(x) = I^{n-1}(I(x))$, $O^0(x) = x$, $O^n(x) = O^{n-1}(O(x))$, $n \in \mathbb{N}^*$.

(14) $\mathrm{CODE}(x)$: reduced dynamics or code for x.

(15) $CODE = \{c \mid c = \mathrm{CODE}(x), x \in \mathbb{N}^*\}$.

(16) $|code|$: the length of $code$; e.g., if a code consists of I or O, it will be the number of I, O in the ordered sequence $\{I, O\}^{\geq 1}$.

(17) $A \parallel B$: concatenation of $A \in \{I, O\}^{\geq a}$ and $B \in \{I, O\}^{\geq b}$. $a \in \mathbb{N}^*, b \in \mathbb{N}$, or $a \in \mathbb{N}, b \in \mathbb{N}^*$, $A \parallel \{I, O\}^0 = A$, $\{I, O\}^0 \parallel B = B$.

(18) $\parallel S = \{\cdots\} \parallel$ returns the number of elements in a set S.

(19) $\max(S = \{\})$, $\max(a, b)$ returns the maximal value in a set S, and the larger one in a, b, respectively.

(20) $[i_1, i_2, \ldots, i_n]_m = \bigcup_{j=1}^{n} [i_j]_m$.

(21) X_c: current transformed number.

2. Reduced Collatz Conjecture

Definition 1 (Collatz transformation).

$$CT(x) = \begin{cases} TPO(x) = 3 * x + 1 & (x \in [1]_2) \\ H(x) = \dfrac{x}{2} & (x \in [0]_2) \end{cases} \quad (1)$$

$TPO(x)$ can be simply denoted as TPO, and $H(x)$ can be simply denoted as H.

We assume $CT^0(x) = x$. Let $CT^n(x) = CT(CT^{n-1}(x)), n \in \mathbb{N}^*$.

Definition 2 (Collatz Conjecture (CC)). $\forall x \in \mathbb{N}^*, \exists L \in \mathbb{N}^*$, such that $CT^L(x) = 1$.

When $x = 1$, CC is held (i.e., $1 \longrightarrow 4 \longrightarrow 2 \longrightarrow 1$). Thus, $\exists L = 3, CT^3(1) = H(H(TPO(1))) = H(H(4)) = H(2) = 1$. $CT^3 \in \{TPO, H\}^3$. More specifically, here corresponding CT^3 for $x = 1$ is an *ordered sequence* $\langle TPO, H, H \rangle$. Here "ordered sequence" implies that the parity of $CT^{n-1}(x)$ determines whether the intermediately forthcoming CT is TPO or H.

In the following, we mainly are concerned with $x \geq 2, x \in \mathbb{N}^*$.

We give the Reduced Collatz Conjecture as follows.

Definition 3 (Reduced Collatz Conjecture (RCC)). $\forall x \in \mathbb{N}^*, x \geq 2, \exists L \in \mathbb{N}^*$, such that $CT^L(x) < x$ and $CT^i(x) \not< x, i = 0, 1, \ldots, L - 1$.

That is, the minimal L such that $CT^L(x) < x$ is of interest, since $CT^i(x) \not< x, i = 0, 1, \ldots, L - 1$.

Proposition 4. $CC \Longleftrightarrow RCC$.

Proof. $\forall x, L \in \mathbb{N}^*, x \geq 2$, and it is obvious that $CT^L(x) \in \mathbb{N}^*$, i.e., $CT^L(x) \geq 1$.

(1) Suppose CC is true. That is, $\forall x \in \mathbb{N}^*, x \geq 2, \exists L \in \mathbb{N}^*, CT^L(x) = 1 < x$. Thus, RCC is true.

(2) Inversely, suppose RCC is true. That is, $\forall x \in \mathbb{N}^*, x \geq 2, \exists q_0 \in \mathbb{N}^*, CT^{q_0}(x) < x$.

If $CT^{q_0}(x) = 1$, then CC is true.

If $CT^{q_0}(x) > 1$, then let $y_1 = CT^{q_0}(x)$. As RCC is true, $\exists q_1 \in \mathbb{N}^*, CT^{q_1}(y_1) < y_1$.

Let $y_0 = x$. Iteratively, if $y_i = CT^{q_{i-1}}(y_{i-1}) = 1, i \in \mathbb{N}^*$, then CC is true. If $y_i = CT^{q_{i-1}}(y_{i-1}) > 1$, then $\exists q_i \in \mathbb{N}^*, y_{i+1} = CT^{q_i}(y_i) < y_i$.

Thus, $y_{i+1} < y_i < \ldots < y_1 < y_0 = x$. $y_i, i \in \mathbb{N}^*$ is a *strictly decreasing serial*.

Besides, $y_{i+1} = CT^{q_0 + q_1 + q_2 + \cdots + q_i}(x) \geq 1$.

Therefore, after finite times of iterations, $\exists n \in \mathbb{N}^*, y_n = 1$. That is, $\exists q = q_0 + q_1 + \cdots + q_{n-1} = \sum_{i=0}^{n-1} q_i \in \mathbb{N}^*, CT^q(x) = 1$.

Thus, CC is true. □

Remark 5.

(1) x is called starting number, and x after transformation (e.g., $CT^i(x), i = 1, 2, \ldots, q$) is called transformed number.

(2) We call an ordered sequence $CT^q \in \{TPO, H\}^q$ in above proof as *original dynamics* (as $CT^q(x) = 1$). Simply speaking, original dynamics of a starting number x is represented by (or composed of) a serial of occurred Collatz transformations during the procedure from the starting number (i.e., x) to 1.

For example, the dynamics of 5 (i.e., occurred transformations during the procedure from 5 to 1) is $5 \longrightarrow 16 \longrightarrow 8 \longrightarrow 4 \longrightarrow 2 \longrightarrow 1$. That is, original dynamics of 5 is $\langle TPO, H, H, H, H \rangle$.

(3) In contrast, we call $CT^{q_0} \in \{TPO, H\}^{q_0}$ in the above proof as *reduced dynamics* (as $CT^{q_0}(x) < x$). Simply speaking, reduced dynamics of a starting number x is represented by (or composed of) a serial of occurred Collatz transformations during the procedure from the starting number (i.e., x) to the *first* transformed number that is less than the starting number (i.e., x).

For example, the reduced dynamics of 5 (i.e., occurred transformations during the procedure from 5 to the first transformed number less than the starting

number, namely, 4) is $5 \longrightarrow 16 \longrightarrow 8 \longrightarrow 4$, in other words, $\langle TPO, H, H \rangle$.

Note that reduced dynamics is more primitive than original dynamics, because original dynamics consists of reduced dynamics. It can be easily observed from the proof of Proposition 4. Nonetheless, we formally prove it in Proposition 6 as follows.

Proposition 6. $\forall x_1 \in \mathbb{N}^*, x_1 \geq 2$, if $\exists L \in \mathbb{N}^*$ such that $CT^L(x_1) = 1$; then, $\exists l_i \in \mathbb{N}^*, i = 1, 2, \ldots, n$, such that $CT^{(\sum_{i=1}^{n} l_i)=L}(x_1) = 1$ and $x_{i+1} = CT^{l_i}(x_i) < x_i, x_i \in \mathbb{N}^*$.

Proof. (1) $CT^L(x_1) = 1$ and $x_1 \geq 2$; thus, $\exists l_1 \in \mathbb{N}^*, l_1 \leq L$ such that $CT^{l_1}(x_1) < x_1$.

(2) Let $x_2 = CT^{l_1}(x_1)$. If $x_2 > 1$, then $\exists L' = L - l_1 \in \mathbb{N}^*$ such that $CT^{L'}(x_2) = 1$. Thus, similarly to (1), $\exists l_2 \in \mathbb{N}^*, l_2 \leq L'$ such that $CT^{l_2}(x_2) < x_2$.

(3) Iteratively, compute l_i in the above way. Thus, x_1, $x_2 = CT^{l_1}(x_1)$, $x_3 = CT^{l_1+l_2}(x_1)$,..., $CT^{l_1+l_2+\cdots}(x_1)$ is a strictly decreasing serial. Besides, $CT^{\geq 1}(x) \geq 1$, thus $\exists n \in \mathbb{N}^*$, $CT^{l_1+l_2+\ldots+l_n}(x_1) = 1$. Therefore, $CT^{(\sum_{i=1}^{n} l_i)=L}(x_1) = 1$ and $x_{i+1} = CT^{l_i}(x_i) < x_i, x_i \in \mathbb{N}^*, i = 1, 2, \ldots, n$. $\square$

Due to above proposition, we concentrate on reduced dynamics, which is a component of original dynamics.

3. Induction

This section is not preliminary for the rest of the paper, but it presents a formal mathematical induction related to Reduced Collatz Conjecture.

To simplify the statement for conjecture, we define "*Returnable*" as follows.

Definition 7 (Returnable). $x \in \mathbb{N}^*$ is Returnable (denoted as $x \in \mathcal{RTN}$), if and only if $\exists L \in \mathbb{N}^*$ such that $CT^L(x) = 1$.

The Collatz Conjecture will be true, if the following mathematical induction can be proved.

Induction (for Collatz Conjecture)

(1) $x = 1 \in \mathcal{RTN}$ (recall that $1 \longrightarrow 4 \longrightarrow 2 \longrightarrow 1$).

(2) If $\forall x \leq k$ (where $x, k \in \mathbb{N}^*$) is Returnable, then $x = k + 1$ will be Returnable. That is, if $x \in \mathcal{RTN}$ (where $x \leq k, x, k \in \mathbb{N}^*$), then $x = k + 1 \in \mathcal{RTN}$ can be proved.

In shorthand, the induction is as follows.

$$x \in \mathcal{RTN} \implies x = k + 1 \in \mathcal{RTN}, \quad \forall x \leq k, \quad (2)$$

where $x, k \in \mathbb{N}^*$.

Therefore, we *only* need to check whether current transformed number is less than designated starting number. Once current transformed number is less than the starting number, the starting number will be Returnable (i.e., $x = k + 1 \in$
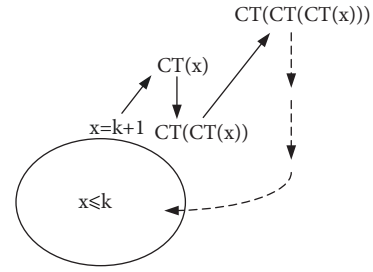

Figure 1: Induction rationale. Once current transformed number is less than the starting number, the starting number will be called "Returnable". That is, once $CT(CT(\ldots CT(x))) < x$, then $x \in \mathcal{RTN}$.

$\mathcal{RTN}$) due to the induction assumption ($\forall x \leq k, x \in \mathcal{RTN}$).

Figure 1 illustrates the rationale in our induction.

Proposition 8. *If the induction (especially, Step 2) can be proved, Collatz Conjecture is True.*

Proof. Straightforward. $\square$

Besides, it is trivial to check that $k = 1, x = k = 1 \in \mathcal{RTN}, x = k + 1 = 2 \in \mathcal{RTN}$.

If k is odd in the induction, the induction is trivial to be proved. We state it as a proposition as follows.

Proposition 9.

$$x \in \mathcal{RTN} \implies x = k + 1 \in \mathcal{RTN}, \quad \forall x \leq k, \quad (3)$$

where $x \in \mathbb{N}^*, k \in [1]_2$.

Proof. $k \in [1]_2$, thus $k + 1$ is even. That is, when $x = k + 1$, $CT(x) = (k + 1)/2 \leq k. \forall x \leq k, x \in \mathcal{RTN}$, so $CT(x) \in \mathcal{RTN}$. Thus, $x = k + 1 \in \mathcal{RTN}$. $\square$

Therefore, for the proof of induction we only need to prove the case that k is even.

If in the induction k is even with $k = 4t, t \in \mathbb{N}^*$, the induction is straightforward to be proved. We state it as a proposition as follows.

Proposition 10.

$$x \in \mathcal{RTN} \implies x = k + 1 \in \mathcal{RTN}, \quad \forall x \leq k, \quad (4)$$

where $x \in \mathbb{N}^*, k = 4t, t \in \mathbb{N}^*$.

Proof. $k = 4t, t \in \mathbb{N}^*$. Thus, $k + 1 = 4t + 1 \in [1]_2$. Next, let us check whether $x = k+1 \in \mathcal{RTN}$. $CT(k+1) = (3(4t+1)+1) = 12t + 4 \in [0]_2, CT(CT(x)) = (12t + 4)/2 = 6t + 2 \in [0]_2$, $CT(CT(CT(x))) = (6t + 2)/2 = 3t + 1 \leq 4t$. As $\forall x \leq k = 4t$ is Returnable, $CT(CT(CT(x))) \in \mathcal{RTN}$. Thus, $x = k + 1 \in \mathcal{RTN}$. $\square$

Therefore, we only need to prove the case that k is even with $k = 4t + 2, t \in \mathbb{N}$ in the induction due to the above

Propositions 9 and 10. We give reduced version of induction as follows.

Induction (Reduced Version of Induction for Collatz Conjecture)

(1) $x = 1, 2, 3, 4, 5, 6 \in \mathscr{RTN}$ (Straightforward).

(2) If $\forall x \leq k = 2t + 1$ $(x, k, t \in \mathbb{N}^*)$ is Returnable, then $x = k + 1$ will be Returnable. That is, if $x \in \mathscr{RTN}$ $(x \leq k = 2t + 1, x, k, t \in \mathbb{N}^*)$, then $x = k + 1 \in \mathscr{RTN}$ can be proved.

(3) If $\forall x \leq k = 4t$ $(x, k, t \in \mathbb{N}^*)$ is Returnable, then $x = k + 1$ will be Returnable. That is, if $x \in \mathscr{RTN}$ $(x \leq k = 4t, x, k, t \in \mathbb{N}^*)$, then $x = k + 1 \in \mathscr{RTN}$ can be proved.

(4) If $\forall x \leq k = 4t + 2$ $(x, k, t \in \mathbb{N}^*)$ is Returnable, then $x = k + 1$ will be Returnable. That is, if $x \in \mathscr{RTN}$ $(x \leq k = 4t + 2, x, k, t \in \mathbb{N}^*)$, then $x = k + 1 \in \mathscr{RTN}$ can be proved.

As Steps (2) and (3) can be proved by Propositions 9 and 10, respectively. In shorthand, the reduced version of induction for Collatz Conjecture that needs to be proved is only Step (4) as follows:

$$x \in \mathscr{RTN} \implies x = k + 1 \in \mathscr{RTN},$$
$$\forall x \leq k = 4t + 2, \tag{5}$$

where $x, k, t \in \mathbb{N}^*$.

4. CODE(x) and Its Properties

Theorem 11. *H always follows after TPO in $CT^{\geq 1}(x \in \mathbb{N}^*)$.*

Proof. In the definition of $CT(\cdot)$, when $x \in [1]_2$, $CT(x) = TPO(x) = 3 * x + 1$, which is even obviously. Thus, next $CT(\cdot)$ must be $H(\cdot)$ consequently. Therefore, H always follows after TPO. □

Therefore, we introduce new notations (i.e., I and O) for simplicity.

(1) H always occurs after TPO; thus, $H(TPO(x))$ can be written together and denoted as $I(x)$. That is, $I(x) = H(TPO(x)) = (3 * x + 1)/2$.

(2) $O(\cdot)$ is used to denote $H(\cdot)$ (for better vision contrastively).

(3) I and O may also be called Collatz transformations.

For example, reduced dynamics of $x = 5$ are $\langle TPO, H, H \rangle$. The transformation procedures are $TPO(x), H(TPO(x))$, and $H(H(TPO(x)))$. It can also be simplified as $I(x)$ and $O(I(x))$. Thus, reduced dynamics of $x = 5$ can be written as $\langle I, O \rangle$ or "IO" in short.

Besides, $O(I(x))$ can be simply written as $IO(x)$. That is, $IO(x) = O(I(x))$, where $IO(\cdot)$ is a composite function, e.g., $IO(5) = O(I(5)) = O((3 * 5 + 1)/2) = O(8) = 8/2 = 4 < 5$. Formally, $CT_1 \parallel CT_2 \parallel \cdots \parallel CT_n(x) = CT_n(CT_{n-1}(\ldots CT_2(CT_1(x))))$, where $CT_i \in \{I(\cdot), O(\cdot)\}, i = 1, 2, \ldots, n$.

Definition 12. $\forall x \in \mathbb{N}^*$, $x \geq 2$, if $\exists L \in \mathbb{N}^*$ such that $CT^L(x) < x$ and $CT^i(x) \not< x, i = 0, 1, \ldots, L - 1$, where $CT(\cdot) \in \{I(\cdot), O(\cdot)\}$, then let $c = CT^L \in \{I, O\}^L$ and c is called code (or reduced dynamics) for x, denoted as $\mathsf{CODE}(x) = CT^L = c$.

Note that $CT^L \in \{I, O\}^L$ is an ordered sequence consisting of I and O. Besides, $CT^L = CT^{L-1} \parallel CT, L \geq 2, CT^L(x) = CT(CT^{L-1}(x))$. Recall that $CT^0(x) = x$. Furthermore, this sequence implicitly matches the *parity* of all intermediate transformed numbers that are taken as input of $CT(\cdot)$.

For example, $\mathsf{CODE}(x = 5) = IO$ implies the following results:

(1) $IO(5) = (3 * 5 + 1)/2/2 = 4 < 5$;

(2) $I(5) = (3 * 5 + 1)/2 = 8 \not< 5$;

(3) "I" is due to $5 \in [1]_2$;

(4) "O" is due to $I(5) = (3 * 5 + 1)/2 = 8 \in [0]_2$.

Theorem 13. *$\forall x \in \mathbb{N}^*$, $x \geq 2$, if $\exists L \in \mathbb{N}^*$, such that $CT^L(x) < x$ and $CT^i(x) \not< x, i = 0, 1, \ldots, L - 1$, where $CT(\cdot) \in \{I(\cdot), O(\cdot)\}$, and letting $c = CT^L \in \{I, O\}^L$, then c is unique.*
That is,

$$\left\| \left\{ c \mid x, L \in \mathbb{N}^*, x \geq 2, CT^L(x) < x, CT^i(x) \not< x, i \right. \right.$$
$$= 0, 1, \ldots, L - 1, CT(\cdot) \in \{I(\cdot), O(\cdot)\}, c = CT^L \right\} \right\| \tag{6}$$
$$= 1.$$

Proof (straightforward). $\forall x \in \mathbb{N}^*$, $\exists L \in \mathbb{N}^*$, such that $CT^L(x) < x$, where $CT \in \{I, O\}$; let $c = CT^L \in \{I, O\}^L$. Given x, either $I(x)$ or $O(x)$ is deterministic and unique. Similarly, given x, $CT^i(x), i = 1, 2, \ldots, L$ is deterministic and unique. (Recall that, the parity of $CT^i(x)$ determines the intermediately forthcoming transformation). Thus, c is unique for any given x. □

Remark 14.

(1) We assume $\mathsf{CODE}(x = 1) = IO$, although $IO(1) = O((3 * 1 + 1)/2) = O(2) = 2/2 = 1 \not< x$. In other words, the code for $x = 1$ is IO. $\|\{c \mid c = \mathsf{CODE}(1)\}\| = \|\{c \mid c = IO\}\| = 1$. In the following, we are always concerned with $\{x \mid x \geq 2, x \in \mathbb{N}^*\}$.

(2) If $L = |c|$ is finite for x ($|\cdot|$ returns the length of c, or the number of I and O in the ordered sequence $c \in \{I, O\}^{\geq 1}$), then $\mathsf{CODE}(x)$ exists; if $\mathsf{CODE}(x)$ exists, then L is finite.

(3) If RCC is true, then $\forall x \in \mathbb{N}^*$, $\mathsf{CODE}(x)$ exists; if $\forall x \in \mathbb{N}^*$, $\mathsf{CODE}(x)$ exists, then RCC is true.

(4) In $\mathsf{CODE}(x)$, x is called starting number. $CT^i(x), i = 1, 2, \ldots, L, L = |c|$ are called transformed numbers. $CT^L(x)$ is the first transformed number that is less than the starting number x. That is, $CT^i(x) \not< x, i = 0, 1, \ldots, L - 1$, and $CT^L(x) < x$. Besides, the parity of

$CT^i(x)$ determines the selection of the intermediately next $CT(\cdot) \in \{I(\cdot), O(\cdot)\}$ after CT^i.

(5) Each $I(\cdot)$ includes one $TPO(\cdot)$ computation since $I(\cdot) = H(TPO(\cdot))$. We denote the count of $I(\cdot)$ in $\text{CODE}(x)$ as U. It equals the count of TPO or $3 * x + 1$ in the reduced dynamics of x. As $TPO(x) = 3 * x + 1 > x$, roughly speaking, U indeed equals the times of "going Up (becoming larger)" in the reduced dynamics of x.

(6) Each $I(\cdot)$ includes one $H(\cdot)$ computation, and $O(\cdot)$ itself is one $H(\cdot)$ computation. We denote the count of $H(\cdot)$ in $\text{CODE}(x)$ as D. It equals the count of H or $x/2$ in the reduced dynamics of x. As $H(x) = x/2 < x$, roughly speaking, D indeed equals the times of "going $Down$ (becoming smaller)" in the reduced dynamics of x. Note that the count of $I(\cdot)$ and $O(\cdot)$ in $\text{CODE}(x)$ also equals D. In other words, D equals the length of $\text{CODE}(x)$. That is, $D = |\text{CODE}(x)|$.

(7) We do not assume the existence of $\text{CODE}(x)$ for $\forall x \in \mathbb{N}^*$, which is exactly what needs to be proved in Reduced Collatz Conjecture.

We introduce notation $\text{CODE}(x)$ for the following reasons:

 (i) The presentation will be more convenient.
 (ii) $\text{CODE}(x \in [3, 99999999])$ are outputted by our computer program. We may discover some properties in them by observation, and they will be proved formally.
 (iii) We can explore inner laws for $\text{CODE}(x)$ without the detail of x (independent to x).

The following propositions again confirm Propositions 9 and 10.

Proposition 15. $\text{CODE}(\{x \mid x \in [0]_2\}) = O$, $\text{CODE}(\{x \mid x \in [1]_4\}) = IO$.

Proof. (1) $x \in [0]_2$, thus H occurs. $x/2 < x$, thus $\text{CODE}(x) = O$.

(2) If $x = 1$, $\text{CODE}(1) = IO$ (by assumption).

If $x \geq 2$, $x = 4t + 1 \in [1]_2$, where $t \in \mathbb{N}^*$. Thus, I occurs. $I(x) = (3 * x + 1)/2 = (3 * (4t + 1) + 1)/2 = (12t + 4)/2 = 2 * (3t + 1) \in [0]_2$. $2 * (3t + 1) > x = 4t + 1$, thus further transformation occurs. $O(I(x)) = 2 * (3t + 1)/2 = 3t + 1 < 4t + 1 = x (\because t \in \mathbb{N}^*)$, thus $\text{CODE}(x) = IO$.

In summary,

$$\text{CODE}(x) = \begin{cases} O & x \in [0]_2, \\ IO & x \in [1]_4. \end{cases} \qquad (7)$$

$\square$

(In the following, $\text{CODE}(\{x \mid x \in [i]_m\})$ is shortened as $(x \in [i]_m)$, $m \geq 2$, $m \in \mathbb{N}^*$, $0 \leq i \leq m - 1$, $i \in \mathbb{N}$.)

If $\text{CODE}(x \in \mathbb{N}^*)$ exists, they can be looked as a whole - $CODE = \{c \mid c = \text{CODE}(x), x \in \mathbb{N}^*\}$, and $CODE$ presents certain properties.

Proposition 16. *If* $\text{CODE}(x \in [3]_4)$ *exists, then* $\text{CODE}(x) = I^p O \parallel \{I, O\}^{\geq 1}$, $p \geq 2$.

Proof. Letting $x = 4t + 3$, $t \in \mathbb{N}$. Obviously, $x \in [1]_2$.

$I(x) = (3x + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$.

$I^2(x) = I(I(x)) = (3(6t + 5) + 1)/2 = 9t + 8$.

(1) If $t \in [0]_2 \cup \{0\}$, then $9t + 8 \in [0]_2$. Thus, the next transformation is "$x/2$". Thus, the first five Collatz transformations are "IIO" (i.e., "TPO, H, TPO, H, H"). Besides, $IIO(x) = (9t + 8)/2 = 4.5t + 4 > 4t + 3 = x$. Further transformation thus occurs. Hence, if $\text{CODE}(x)$ exists, then $\text{CODE}(x) = I^2 O \parallel \{I, O\}^{\geq 1}$.

(2) If $t \in [1]_2$, then $9t + 8 \in [1]_2$. Thus, the first six Collatz transformations are "III" (i.e., "TPO, H, TPO, H, TPO, H"). Besides, $III(x) = (3(9t + 8) + 1)/2 = (27t + 25)/2 = 13.5t + 12.5 > 4t + 3 = x$. Further transformation occurs. Hence, if $\text{CODE}(x)$ exists, then $\text{CODE}(x) = I^3 \parallel \{I, O\}^{\geq 1}$.

If $III(x) \in [1]_2$, then more "I" occurs. Obviously, $IIII(x) > x$. Further transformation occurs. Hence, if $\text{CODE}(x)$ exists, then $\text{CODE}(x) = I^4 \parallel \{I, O\}^{\geq 1}$. If $III(x) \in [0]_2$, then $IIIO(x) = (13.5t + 12.5)/2 = 6.75t + 6.25 > 4t + 3 = x$. Further transformation occurs consequently. Hence, if $\text{CODE}(x)$ exists, then $\text{CODE}(x) = I^3 \parallel O \parallel \{I, O\}^{\geq 1}$.

Suppose $\text{CODE}(x) = CT^L \in \{I, O\}^L$. There exists at least one "O" in L transformations; otherwise, $CT^L(x) = I^L(x) > I^{L-1}(x) > \ldots > I(x) > x$, which contradicts with $CT^L(x) < x$. Besides, $I^p O(x) = I^p(x)/2 > I^{p-1}(x)/2 = I^{p-1}O(x) > \ldots > IIO(x) > x$, $p \geq 3$; thus, after $I^p O (p \geq 3)$ further transformation occurs.

In summary, if $\text{CODE}(x \in [3]_4)$ exists, then $\text{CODE}(x) = I^p O \parallel \{I, O\}^{\geq 1}$, $p \geq 2$. $\square$

Next corollary states that $\text{CODE}(x)$ (or $CODE$) presents unified format.

Proposition 17. $\text{CODE}(x) \in FORMAT = I^p \parallel O \parallel \{I, O\}^q$, *where* $p, q \in \mathbb{N}$. *More specifically,*

$$\begin{aligned} & p = 0, \\ & q = 0 \\ & \qquad x \in [0]_2, \\ & p = 1, \\ & q = 0 \\ & \qquad x \in [1]_4, \\ & p = 2, \\ & q \geq 1 \\ & \qquad x \in [3]_8, \\ & p \geq 3, \\ & q \geq 1 \\ & \qquad x \in [7]_8. \end{aligned} \qquad (8)$$

Proof. We here assume $I^0 \parallel O = O$ and $O \parallel \{I, O\}^0 = O$. According to Propositions 15 and 16, we have

$$p = 0,$$

$$q = 0$$

$$x \in [0]_2,$$

$$p = 1,$$

$$q = 0$$

$$x \in [1]_4,$$

$$p = 2,$$

$$q \geq 1$$

$$x = 4t + 3, \ t \in [0]_2 \cup \{0\},$$

$$p \geq 3,$$

$$q \geq 1$$

$$x = 4t + 3, \ t \in [1]_2.$$

(9)

It can be written as follows.

(1) $p = 0, q = 0$ as CODE$(x \in [0]_2) = O$; $p = 1, q = 0$ as CODE$(x \in [1]_4) = IO$.

(2) $p = 2, q \geq 1$ when $x \in [3]_8$ as $\{x \mid x = 4t + 3, t \in [0]_2 \cup \{0\}\} = \{x \mid x = 4 * 2k + 3, k \in \mathbb{N}\} = [3]_8$.

(3) $p \geq 3, q \geq 1$ when $x \in [7]_8$, as $\{x \mid x = 4t + 3, t \in [1]_2\} = \{x \mid x = 4(2k + 1) + 3, k \in \mathbb{N}\} = [7]_8$.

Thus, (9) can be rewritten as (8). $\qquad\square$

More specifically, we have the following theorem that give more details on p.

Theorem 18 (format theorem). CODE$(x) \in FORMAT = I^p \|O\| \{I, O\}^q$, where $p, q \in \mathbb{N}$. Besides,

$$p = \begin{cases} 0 & x \in [0]_2, \\ 1 & x \in [1]_4, \\ 2 & x \in [3]_8 \ (x = 4t + 3, t \in [0]_2 \cup \{0\}), \\ \alpha + 2 & t + 1 = 2^\alpha * A, \ A \in [1]_2, \alpha \in \mathbb{N}^*, \ x \in [7]_8 \ (x = 4t + 3, t \in [1]_2). \end{cases}$$

(10)

$q = 0$ when $p = 0, 1$; $q \geq 1$ when $p \geq 2$.

Proof. q is trivial due to Proposition 17, so we are mainly concerned with p.

By Proposition 15 or (7), if $x \in [0]_2$, then CODE$(x) = O \in FORMAT$ and $p = 0$. If $x \in [1]_4$, then CODE$(x) = IO \in FORMAT$ and $p = 1$.

Next, we concentrate on $x \in [3]_4$.

CODE$(3) = IIOO = I^2 O^2$, which can be manually and easily verified.

Let $x = 4t + 3, t \in \mathbb{N}^*$.

(1) Case I: $t \in [0]_2$. As $x = 4t + 3 \in [1]_2$, $I(\cdot)$ is conducted consequently. As $I(x) = (3x + 1)/2 = 1.5x + 1.5 > x$ and $II(x) > I(x) > x$, the checking on whether current transformed number is less than starting number may be omitted in some straightforward cases.

$I(x) = (3x + 1)/2 = (3(4t + 3) + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$; thus, transformation $I(\cdot)$ is conducted consequently.

$I(I(x)) = II(x) = (3(6t + 5) + 1)/2 = (18t + 16)/2 = 9t + 8 \in [0]_2$. Thus, $O(\cdot)$ is conducted consequently.

$O(I(I(x))) = IIO(x) = (9t + 8)/2 = 4.5t + 4 > 4t + 3$. Thus, further transformation is conducted consequently.

Therefore, CODE$(x) = I^2 O \| \{I, O\}^{\geq 1} \in FORMAT$.

(2) Case II: $t \in [1]_2$. As $x = 4t + 3 \in [1]_2$, $I(\cdot)$ is conducted consequently.

$I(x) = (3x + 1)/2 = (3(4t + 3) + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$; thus, $I(\cdot)$ is conducted consequently.

$II(x) = (3(6t + 5) + 1)/2 = (18t + 16)/2 = 9t + 8 \in [1]_2$. Thus, $I(\cdot)$ is conducted consequently.

$III(x) = (3(9t + 8) + 1)/2 = (27t + 25)/2$. It depends on the partition of t (more specifically, $t \in [1]_4$ or $[3]_4$) whether $(27t + 25)/2$ is even or odd.

(It comes from following observations: $(27t + 25)/2 \in [0]_2 \Longleftarrow (27t + 25)/2 = 2k, k \in \mathbb{N}^* \Longleftarrow 27t + 25 = 4k \Longleftarrow 27t \in [3]_4 \Longleftarrow 3t \in [3]_4 \Longleftarrow t \in [1]_4$.)

(2.1) If $t \in [1]_4$, $III(x) = (27t + 25)/2 = (27 * (4 * k + 1) + 25)/2 = (108k + 52)/2 = 54k + 26 \in [0]_2 \ (k \in \mathbb{N}^*)$; thus, $O(\cdot)$ will occur consequently.

(2.2) If $\in [3]_4$, $III(x) = (27t + 25)/2 = (27 * (4 * k + 3) + 25)/2 = (108k + 106)/2 = 54k + 53 \in [1]_2$; thus, $I(\cdot)$ will occur consequently.

Besides, suppose current transformed number is denoted as X_c.

$X_c = IIIO(x) = (27t + 25)/2/2 = (54k + 26)/2 = 27k + 13$, whose parity depends on the parity of k.

$X_c = IIII(x) = (3(27t + 25)/2 + 1)/2 = (3(54k + 53) + 1)/2 = (162k + 160)/2 = 81k + 80$, whose parity depends on the parity of k.

In other words, the judgement on the parity of X_c is undecidable, unless the domain $(t \in [1]_4$ or $t \in [3]_4)$ is *partitioned* further.

For exploring more general results, we put it in another way as follows.

Suppose there exist at most p times of "I" at code head (i.e., $I^p \|O\| \ldots$) for $x \in [3]_4$. Observing following equation for $X_c = I^p(x)$ after consecutive $p \geq 2$ times of "I":

$$X_c = I^P(x) = \frac{(3(\ldots(3(3x+1)/2)+1)/2\ldots)+1)}{2} = \frac{3}{2}\left(\frac{3}{2}\left(\ldots\frac{3}{2}\left(\frac{3}{2}x + \frac{1}{2}\right) + \frac{1}{2}\right) + \cdots + \frac{1}{2}\right) + \frac{1}{2}$$

$$= \left(\frac{3}{2}\right)^P x + \frac{1}{2}\left(\left(\frac{3}{2}\right)^{P-1} + \left(\frac{3}{2}\right)^{P-2} + \cdots + 1\right) = \left(\frac{3}{2}\right)^P x + \frac{1}{2}\left(\frac{(3/2)^P - 1}{3/2 - 1}\right) = \left(\frac{3}{2}\right)^P x + \left(\frac{3}{2}\right)^P - 1 \quad (11)$$

$$= \left(\frac{3}{2}\right)^P (x+1) - 1 \quad (\because x = 4t + 3,\ t \in \mathbb{N}) = \left(\frac{3}{2}\right)^P (4t + 3 + 1) - 1 = \left(\frac{3^P}{2^{P-2}}\right)(t+1) - 1 \in \mathbb{N}^*.$$

Note that above computation implicitly includes two requirements due to p times of consecutive $I(\cdot)$ as follows.

(i) All intermediate transformed numbers during processes (i.e., computing p times of consecutive "I") satisfy

$$X_c = \left(\frac{3^i}{2^{i-2}}\right)(t+1) - 1 \in [1]_2, \quad (12)$$

where $2 \le i \le p - 1$, $i \in \mathbb{N}^*$.

(ii) Besides,

$$X_c = \left(\frac{3^i}{2^{i-2}}\right)(t+1) - 1 \in [0]_2, \quad (13)$$

where $i = p$, as only (or at most) p consecutive $I(\cdot)$ occur.

In other words, p can also be looked as the minimal value to let current transformed number X_c be in $[0]_2$. Thus, we need to explore the requirement on p for given t such that

$$\left(\frac{3^i}{2^{i-2}}\right)(t+1) - 1 \in [1]_2 \quad 2 \le i \le p - 1,$$
$$\left(\frac{3^i}{2^{i-2}}\right)(t+1) - 1 \in [0]_2 \quad i = p. \quad (14)$$

We call this requirement (i.e., (14)) as REQ.

Represent $t + 1$ as $2^\alpha * A$, $A \in [1]_2$, $\alpha \in \mathbb{N}^*$. That is, $t + 1 = 2^\alpha * A$. Obviously, this representation is unique. We

thus need to prove that REQ is satisfied if and only if $p = \alpha + 2$. Note that we will see that here p is indeed determined by α.

For $2 \le i < p = \alpha + 2$, $i \in \mathbb{N}^*$, we have $\alpha + 2 - i > 0$.

$$X_c = \left(\frac{3^i}{2^{i-2}}\right)(t+1) - 1 = \left(\frac{3^i}{2^{i-2}}\right) * 2^\alpha * A - 1$$
$$= 3^i * 2^{\alpha - i + 2} * A - 1. \quad (15)$$

$\alpha + 2 - i > 0 \implies 2^{\alpha - i + 2} \in [0]_2 \implies 3^i * 2^{\alpha - i + 2} * A \in [0]_2 \implies 3^i * 2^{\alpha - i + 2} * A - 1 \in [1]_2.$

When $i = p = \alpha + 2$, we have exactly

$$X_c = \left(\frac{3^P}{2^{P-2}}\right)(t+1) - 1 = \left(\frac{3^P}{2^{P-2}}\right) * 2^\alpha * A - 1$$
$$= 3^P * 2^{\alpha - P + 2} * A - 1 = 3^P * A - 1 \in [0]_2. \quad (16)$$

$$\because A,\ 3^P \in [1]_2$$

It is easy to see that $p = \alpha + 2$ is the one and only one for REQ, as desired. □

Corollary 19 (t determine p corollary). *Given starting number* $x \in [3]_4$ *(i.e.,* $x = 4t + 3$, $t \in \mathbb{N}$*), the number of consecutive "I" (denoted as p) is determined by t according to (17) as follows:*

$$p = \begin{cases} 2 & t \in [0]_2, \\ \alpha + 2 & \alpha = \log_2 \dfrac{t+1}{A} \in \mathbb{N}^*,\ A = \max\left(\{a \mid a \in [1]_2, a \mid (t+1)\}\right), t \in [1]_2. \end{cases} \quad (17)$$

Corollary 20. *Given starting number* $x \in [3]_4$, $t = (x-3)/4 \in \mathbb{N}$*, the first p times of Collatz transformations must be "I" and p can be determined by t by (17), and the transformed number after I^P transformations is*

$$I^P(x) = \left(\frac{3}{2}\right)^P (x+1) - 1 = \left(\frac{3^P}{2^{P-2}}\right)(t+1) - 1 \quad (18)$$
$$\in \mathbb{N}^*, \quad p \ge 2.$$

Proof. It is straightforward due to (11), Theorem 18, and Corollary 19.

Alternatively, only to compute the transformed number $I^P(x)$, we can prove it by induction as follows.

(1) $p = 2$.

$$I^P(x) = I^2(x) = \frac{(3(3x+1)/2 + 1)}{2}$$
$$= \frac{3}{2}\left(\frac{3}{2}x + \frac{1}{2}\right) + \frac{1}{2} = \left(\frac{3}{2}\right)^2 x + \frac{3}{2} * \frac{1}{2} + \frac{1}{2}$$

$$= \left(\frac{3}{2}\right)^2 x + \left(\frac{3}{2}\right)\left(\frac{3}{2} - 1\right) + \frac{1}{2}$$

$$= \left(\frac{3}{2}\right)^2 x + \left(\frac{3}{2}\right)^2 - \frac{3}{2} + \frac{1}{2} = \left(\frac{3}{2}\right)^2 (x + 1) - 1$$

$$= \left(\frac{3}{2}\right)^p (x + 1) - 1.$$

$$(19)$$

(2) Suppose $p = i$; we have $I^i(x) = (3/2)^i(x + 1) - 1$.

$$I^{i+1}(x) = I\left(I^i(x)\right) = \frac{\left(3\left((3/2)^i(x+1) - 1\right) + 1\right)}{2}$$

$$= \frac{3}{2}\left(\left(\frac{3}{2}\right)^i (x+1) - 1\right) + \frac{1}{2}$$

$$= \left(\frac{3}{2}\right)^{i+1}(x+1) - \frac{3}{2} + \frac{1}{2} \qquad (20)$$

$$= \left(\frac{3}{2}\right)^{i+1}(x+1) - 1.$$

$\square$

Remark 21. (1) Note that, due to Corollary 19, p for I^p in $\mathsf{CODE}(x)$ can be computed by $t = (x - 3)/4$ and $\log_2((t+1)/A)$ directly without conducting concrete Collatz transformations, which can accelerate the computation of dynamics.

(2) Besides, by (11) or Corollary 20, if $t = 0$, we then have $p = 2$, because $I^p(x) = (3^p/2^{p-2})(t+1) - 1 = (3^p/2^{p-2}) - 1 \in \mathbb{N}^*$, which matches with the result $\mathsf{CODE}(3) = IIOO$ by manually computing.

(3) Indeed, (11) can be extended to include all cases (i.e., for $p = 0, 1$). If $p = 0$, by assuming $I^0(x) = x$, $I^0(x) = x = (3/2)^0(x+1) - 1 = (3/2)^p(x+1) - 1$; if $p = 1$, $I^1(x) = I(x) = (3*x+1)/2 = (3/2)^1(x+1) - 1 = (3/2)^p(x+1) - 1$. Therefore, $I^p(x) = (3/2)^p(x+1) - 1$ for $p \in \mathbb{N}^*$.

Corollary 22. *Given starting number $x \in [1]_2$, the first p times of Collatz transformations must be "I" and p can be determined by x, and the transformed number after I^p transformations is*

$$I^p(x) = \left(\frac{3}{2}\right)^p (x + 1) - 1, \quad p \geq 1. \qquad (21)$$

Proof. It is straightforward due to (11), Theorem 18, and Corollaries 19 and 20.

If $x \in [1]_4$, then $p = 1$, $I^p(x) = I(x) = (3x+1)/2 = (3/2)^1(x+1) - 1 = (3/2)^p(x+1) - 1$.

If $x \in [3]_4$, then $p \geq 2$, $I^p(x) = (3/2)^p(x+1) - 1$ due to Corollary 20. $\square$

Next corollary states that the head of code for $x = 2^n - 1, n \in \mathbb{N}^*$ is $I^n O$.

Corollary 23.

$$\mathsf{CODE}\left(x = 2^n - 1, n \in \mathbb{N}^*\right) \in I^n \| O \| \{I, O\}^{\geq 0}. \qquad (22)$$

Proof. Suppose $\mathsf{CODE}(x) \in I^p \| O \| \{I, O\}^{\geq 0}$, due to Theorem 18.

(1) $n = 1, x = 2 - 1 = 1 \in [1]_4$; thus, $p = 1$ due to Theorem 18. Thus, $\mathsf{CODE}(2^1 - 1) \in I^1 \| O \| \{I, O\}^{\geq 0}$.

Indeed, $\mathsf{CODE}(x = 1) = IO$ can be manually computed (recall that $1 \longrightarrow 4 \longrightarrow 2 \longrightarrow 1$.)

(2) $n \in \mathbb{N}^*, n \geq 2, x = 2^n - 1 \in [3]_4$.

(2.1) $n = 2, x = 2^2 - 1 = 3 \in [3]_8$, or $(x - 3)/4 = 0 \in [0]_2 \cup \{0\}$, thus $p = 2$ due to Theorem 18. Thus, $\mathsf{CODE}(2^2 - 1) \in I^2 \| O \| \{I, O\}^{\geq 1}$.

Indeed, $\mathsf{CODE}(x = 3) = IIOO = I^2OO$ can be manually computed (recall that $3 \longrightarrow 10 \longrightarrow 5 \longrightarrow 16 \longrightarrow 8 \longrightarrow 4 \longrightarrow 2 < 3$.)

(2.2) $n > 2, t = (x - 3)/4 = (2^n - 1 - 3)/4 = 2^{n-2} - 1 \in [1]_2$. $t + 1 = 2^{n-2}$. Thus, $A = 1$. $\alpha = \log_2(t+1)/A = \log_2 2^{n-2} = n - 2$. Thus, $p = \alpha + 2 = n - 2 + 2 = n$. Hence, $\mathsf{CODE}(x) \in I^n \| O \| \{I, O\}^{\geq 1}$ owing to Theorem 18.

Summarizing (1) and (2), $\mathsf{CODE}(x = 2^n - 1, n \in \mathbb{N}^*) \in I^n \| O \| \{I, O\}^{\geq 0}$. $\square$

Example 24. (1) $x = 7$. $7 \longrightarrow 22 \longrightarrow 11 \longrightarrow 34 \longrightarrow 17 \longrightarrow 52 \longrightarrow 26 \longrightarrow 13 \longrightarrow 40 \longrightarrow 20 \longrightarrow 10 \longrightarrow 5$. Thus, $\mathsf{CODE}(7) = IIIOIOO = I^3 \| O \| IOO$.

(2) $x = 15$. $15 \longrightarrow 46 \longrightarrow 23 \longrightarrow 70 \longrightarrow 35 \longrightarrow 106 \longrightarrow 53 \longrightarrow 160 \longrightarrow 80 \longrightarrow 40 \longrightarrow 20 \longrightarrow 10$. Thus, $\mathsf{CODE}(15) = IIIIOOO = I^4 \| O \| OO$.

Proposition 25. *$\mathsf{CODE} \in \{I, O\}^{\geq 0} \| O$. (That is, CODE must end with "O", or the last transformation in CODE is "O".)*

Proof (straightforward). It can be easily understood intuitively. That is, reduced dynamics should be ended by "going Down", not "going Up" (recall items (5) and (6) in Remark 14).

If $CODE$ ends by "I", suppose $\exists x \in \mathbb{N}^*, x \geq 2, c(x) \not< x$, $\mathsf{CODE}(x) = c \| I$.

$\{c \| I\}(x) = I(c(x)) = (3 * c(x) + 1)/2 = 1.5c(x) + 0.5 = c(x) + 0.5c(x) + 0.5 > c(x)$; thus, $\mathsf{CODE}(x) = \{c \| I\}(x) \not< x$. Contradiction occurs.

Put it in another way, if $\{c \| I\}(x) < x$, then $I(c(x)) < x$. Together with $I(c(x)) > c(x)$, we have $c(x) < x$. Thus, reduced dynamics of x ends after c and $\mathsf{CODE}(x) = c \neq c \| I$. $\square$

Next corollary gives more details on $\mathsf{CODE}(x \in [3]_4)$ that has a unified form as

$$I^{p_0 \geq 2} \| O^{q_0 \geq 1} \| I^{p_1 \geq 1} \| O^{q_1 \geq 1} \| \cdots \| I^{p_n \geq 1} \| O^{q_n \geq 1}. \qquad (23)$$

That is, each code consists of one or more segments, and each segment has a unified form as $I^p O^q, p \geq 1, q \geq 1$.

Corollary 26.

$$\mathsf{CODE}\left(x \in [3]_4\right) \in I^{p_0 \geq 2} O^{q_0} \| \{I^{p_i} O^{q_i}\}^{\geq 0}, \qquad (24)$$

where $p_0, q_0, i, p_i, q_i \in \mathbb{N}^*$.

Proof (straightforward). $x \in [1]_2$, thus I occurs. After p times of I transformations, $I^p(x) \in [0]_2$ and thus O follows. After q times of O transformations, $I^p O^q(x) \in [1]_2$, thus I occurs.

Indeed, q can be determined by $I^p(x)$ by $q = \log_2(I^p(x)/B)$, $B = \max(\{b \mid b \in [1]_2, b \mid I^p(x)\})$.

Iteratively, each segment has a unified form I^pO^q, where $p, q \in \mathbb{N}^*$.

The first segment is listed solely, because the distinction between the first segment and the other segments is that $p_0 \geq 2$ but $p_i \geq 1, i \in \mathbb{N}^*$. (In other words, when and only when an intermediate transformed number $X_c \in [1]_4$ occurs, $p_i = 1$. Otherwise, $p_i \geq 2$.) □

Put it in another way, if intermediate transformed numbers in $[1]_4$ during reduced dynamics are tackled explicitly, we have the following corollary.

Corollary 27.

$$\text{CODE}\left(x \in [3]_4\right) \in I^{p_0 \geq 2}O^{q_0} \parallel \left\{IO^{q_i}, I^{p_i \geq 2}O^{q_i}\right\}^{\geq 0}. \quad (25)$$

where $p_0, q_0, i, p_i, q_i \in \mathbb{N}^*$.

5. Short Codes - $\{(x,c) \mid c=\text{CODE}(x), |c| \leq 7\}$

In this section, we explore x and their codes whose length is less than 7, called short codes. The exploration of this section helps build empirical understanding for Sections 3 and 4 (e.g., induction and $\text{CODE}(x)$). Before the exploration, a lemma is given as preliminaries as follows.

Lemma 28. $m \in \mathbb{N}^*$, $m \geq 2$, $0 \leq i \leq m - 1, a, b, x \in \mathbb{N}^*$.

(1) $x \in [i]_m \Longrightarrow a * x + b \in [(a * i + b) \bmod m]_m$;
(2) $m, i \in [0]_2, x \in [i]_m \Longrightarrow x/2 \in [i/2]_{m/2}$;
(3) $m, (a * i + b)/2 \in [0]_2, x \in [i]_m \Longrightarrow (a * x + b)/2 \in [((a * i + b) \bmod m)/2]_{m/2}$.

Proof. (1) $x \in [i]_m \Longrightarrow x = k * m + i, k \in \mathbb{N}$
$\Longrightarrow a * x + b = a * (k * m + i) + b = a * k * m + a * i + b$
$\Longrightarrow a * x + b \equiv a * i + b \bmod m \because a * k \in \mathbb{N}$
$\Longrightarrow a * x + b \in [(a * i + b) \bmod m]_m$.
(2) $x \in [i]_m \Longrightarrow x = k * m + i, k \in \mathbb{N}$
$\Longrightarrow x/2 = k * m/2 + i/2 \because m, i \in [0]_2$
$\Longrightarrow x/2 \in [i/2]_{m/2}$.
(Besides, $0 \leq i \leq m - 1, m, i \in [0]_2 \Longrightarrow 0 \leq i \leq m - 2 \Longrightarrow 0 \leq i/2 \leq (m - 2)/2 = m/2 - 1$.)
(3) The proof is the combination of above (1) and (2). Specifically, $x \in [i]_m \Longrightarrow a * x + b \in [(a * i + b) \bmod m]_m$

$\Longrightarrow (a * x + b)/2 \in [((a * i + b) \bmod m)/2]_{m/2}$
$\because m, (a * i + b)/2 \in [0]_2 \Longrightarrow ((a * i + b) \bmod m) \in [0]_2$. □

Proposition 29. $\text{CODE}(x) = IIOO$, if

$$x \in \left\{x \mid x \in [3]_4, t = \frac{(x - 3)}{4} \in \{0\} \cup [0]_4\right\}$$
$$= \{x \mid x = 4 * 4k + 3 = 16k + 3, k \in \mathbb{N}\} = [3]_{16}. \quad (26)$$

Proof. $x = 4t + 3 \in [1]_2$.
$I(x) = (3x + 1)/2 = (3(4t + 3) + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$, and $I(x) = 6t + 5 > 4t + 3 = x$. Thus, "$I$" occurs consequently.

$II(x) = I(6t + 5) = (3(6t + 5) + 1)/2 = 9t + 8 \in [(9 * 0 + 8) \bmod 4]_4 = [0]_4$. Thus, "$OO$" occurs consequently.

As $IIOO(x) = OO(9t + 8) = (9t + 8)/2/2 = 2.25t + 2 < 3t + 2 < 4t + 3 = x$, the code *ends* hereby with "$IIOO$". That is, $\text{CODE}(x \in [3]_{16}) = IIOO$. □

Example 30. $115 \longrightarrow 346 \longrightarrow 173 \longrightarrow 520 \longrightarrow 260 \longrightarrow 130 \longrightarrow 65 < 115$. Thus, $\text{CODE}(115) = IIOO$. It can be verified that $x = 115 \in [3]_{16}, t = (115 - 3)/4 = 28 \in [0]_4$.

Proposition 31. $\text{CODE}(x) = IIOIO$, if

$$x \in \left\{x \mid x \in [3]_4, t = \frac{(x - 3)}{4} \in [2]_8\right\}$$
$$= \{x \mid x = 4(8k + 2) + 3 = 32k + 11, k \in \mathbb{N}\} \quad (27)$$
$$= [11]_{32}.$$

Proof. $x = 4t + 3 \in [1]_2$.
$I(x) = (3x + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$.
$II(x) = I(6t + 5) = (3(6t + 5) + 1)/2 = 9t + 8 \in [(2 * 9 + 8) \bmod 8]_8 = [2]_8 \subset [0]_2$ due to Lemma 28 (1). "O" occurs consequently.

$IIO(x) = (9t + 8)/2 \in [2/2]_{8/2} = [1]_4 \subset [1]_2$. Thus, next transformation is "I". Besides, $(9t + 8)/2 = 4.5t + 4 > 4t + 3 = x$.

$IIOI(x) = (3 * (9t + 8)/2 + 1)/2. \ 3 * (9t + 8)/2 + 1 \in [(3 * 1 + 1) \bmod 4]_4 = [0]_4. \ (3 * (9t + 8)/2 + 1)/2 \in [0]_2$. Thus, "$O$" occurs consequently.

$IIOIO(x) = (3 * (9t + 8)/2 + 1)/2/2 = (13.5t + 13)/2/2 = (6.75t + 6.5)/2 = 3.375t + 3.25 = 4t + 3 + (0.25 - 0.625t) < 4t + 3 = x$; the code ends with "$IIOIO$". That is, $\text{CODE}(x \in [11]_{32}) = IIOIO$. □

Example 32. $11 \longrightarrow 34 \longrightarrow 17 \longrightarrow 52 \longrightarrow 26 \longrightarrow 13 \longrightarrow 40 \longrightarrow 20 \longrightarrow 10 < 11$; $\text{CODE}(11) = IIOIO$, $x = 11 \in [11]_{32}, (11 - 3)/4 = 2 \in [2]_8$.

Proposition 33. $\text{CODE}(x) = IIIOO$, if

$$x \in \left\{x \mid x \in [3]_4, t = \frac{(x - 3)}{4} \in [5]_8\right\}$$
$$= \{x \mid x = 4(8k + 5) + 3 = 32k + 23, k \in \mathbb{N}\} \quad (28)$$
$$= [23]_{32}.$$

Proof. $x = 4t + 3 \in [1]_2, t \in [5]_8$.
$I(x) = (3x + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$, thus next transformation is "I".

$II(x) = I(6t + 5) = (3(6t + 5) + 1)/2 = 9t + 8 \in [(9 * 5 + 8) \bmod 8]_8 = [5]_8 \subset [1]_2$, thus next transformation is "I".

$III(x) = I(9t + 8) = (3(9t + 8) + 1)/2. \ 3(9t + 8) + 1 \in ([(3 * 5 + 1) \bmod 8]_8 = [0]_8. \ (3(9t + 8) + 1)/2 \in [0]_4$, thus next transformations are double "O".

Check whether current transformed number is less than the starting number as follows:

$IIIO(x) = (27t + 25)/2/2 = (13.5t + 12.5)/2 = 6.75t + 6.25t > 4t + 3 = x.$

$IIIOO(x) = (6.75t + 6.25)/2 = 3.375t + 3.125 = 4t + 3 + (0.125 - 0.625t) < 4t + 3$, as $t \geq 1$. $\qquad \square$

Example 34. $55 \longrightarrow 166 \longrightarrow 83 \longrightarrow 250 \longrightarrow 125 \longrightarrow 376 \longrightarrow 188 \longrightarrow 94 \longrightarrow 47 < 55$; CODE(55) = *IIIOO*, $x = 55 \in [23]_{32}, (55-3)/4 = 13 \in [5]_8$.

Proposition 35. CODE$(x) = IIIOIOO$, *if*
$x \in \{x \mid x \in [3]_4, t = (x-3)/4 \in [1]_{32}\} = \{x \mid x = 4(32k+1) + 3 = 128k + 7, k \in \mathbb{N}\} = [7]_{128}$.

Proof. $x = 4t + 3 \in [1]_2$, $t \in [1]_{32}$.
$I(x) = (3x + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$.
$II(x) = (3(6t + 5) + 1)/2 = 9t + 8 \in [9 * 1 + 8]_{32} = [17]_{32} \subset [1]_2$.
$III(x) = (3(9t+8)+1)/2 \in ([((3*17+1) \bmod 32)/2]_{32/2} = [10]_{16} \subset [0]_2$.
$IIIO(x) \in [10/2]_{16/2} = [5]_8 \subset [1]_2$.
$IIIOI(x) = I(IIIO(x)) \in [((3*5+1) \bmod 8)/2]_{8/2} = [0]_4$. Thus, double "*O*" follow immediately. In summary, above occurred dynamics to current transformed number is thus *IIIOIOO*.

Next, we check whether it is the final code (or reduced dynamics is terminated at this transformed number) by checking whether current transformed number is less than the staring number (i.e., x).
$III(x) = (3(9t + 8) + 1)/2 = (27t + 25)/2 = 13.5t + 12.5$.
$IIIO(x) = O(I^3(x)) = (13.5t + 12.5)/2 = 6.75t + 6.25 > 4t + 3 = x$.
$IIIOI(x) = I(IIIO(x)) = (3(6.75t + 6.25) + 1)/2 = (20.25t + 18.75 + 1)/2 = (20.25t + 19.75)/2 = 10.125t + 9.875$.

$IIIOIO(x) = O(IIIOI(x)) = (10.125t + 9.875)/2 = 5.0625t + 4.9375 > 4t + 3 = x$.
$IIIOIOO(x) = O(IIIOIO(x)))) = (5.0625t + 4.9375)/2 = 2.53125t + 2.46875 < 3t + 3 < 4t + 3 = x$.
Therefore, CODE$(x \in [7]_{128}) = IIIOIOO$, which can be also written as I^3OIO^2. $\qquad \square$

Remark 36. Indeed, (11) can be used for computing current transformed number (denoted as X_c) after $I^3(x)$ to simplify above process in the proof. After $I^3(x)$,
$X_c \Longleftarrow (3/2)^3(x+1) - 1 = 1.5^3 x + 1.5^3 - 1 = 3.375x + 2.375$,
$X_c \Longleftarrow O(X_c) = 0.5(3.375x + 2.375) = 1.6875x + 1.1875$,
$X_c \Longleftarrow I(X_c) = 1.5(1.6875x + 1.1875) + 0.5 = 2.53125x + 2.28125$,
$X_c \Longleftarrow O(X_c) = 0.5(2.53125x + 2.28125) = 1.265625x + 1.140625$,
$X_c \Longleftarrow O(X_c) = 0.5(1.265625x + 1.140625) = 0.6328125x + 0.5703125 = x + (0.5703125 - 0.3671875x) < x$, as $x \geq 2$ due to $x \in [3]_4$.

Example 37. $135 \longrightarrow 406 \longrightarrow 203 \longrightarrow 610 \longrightarrow 305 \longrightarrow 916 \longrightarrow 458 \longrightarrow 229 \longrightarrow 688 \longrightarrow 344 \longrightarrow 172 \longrightarrow 86 < 135$; CODE(135) = *IIIOIOO*, $x = 135 \in [7]_{128}, (x-3)/4 = (135-3)/4 = 33 \in [1]_{32}$.

The following two propositions originally stem from our observations on codes outputted by our computer programs.

Proposition 38. CODE$(x) = IIOIIOO$, *if*

$$x \in \left\{ x \mid x \in [3]_4, t = \frac{(x-3)}{4} \in [14]_{32} \right\}$$
$$= \{x \mid x = 4(32k + 14) + 3 = 128k + 59, k \in \mathbb{N}\} \qquad (29)$$
$$= [59]_{128}.$$

Proof. $x = 4t + 3 \in [1]_2, t \in [14]_{32}$.
$I(x) = (3x + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$, thus "*I*" will follow consequently.
$I^2(x) = (3(6t+5)+1)/2 = 9t+8 \in [(9*14+8) \bmod 32]_{32} = [6]_{32} \subset [0]_2$, thus "*O*" will follow.
$O(I^2(x)) \in [6/2]_{32/2} = [3]_{16} \subset [1]_2$, thus "*I*" will follow.
$I(O(I^2(x))) \in ([(3 * 3 + 1)/2]_{16/2} = [5]_8 \subset [1]_2$, thus "*I*" will follow.
$I^2(O(I^2(x))) \in [((3 * 5 + 1) \bmod 8)/2]_{8/2} = [0]_4$, thus double "*O*" will follow.

Next, current transformed number will be compared with starting number x.
$I^2(x) = 9t + 8$.
$O(I^2(x)) = (9t + 8)/2 = 4.5t + 4 > 4t + 3 = x$.
$I(O(I^2(x))) = (3(4.5t + 4) + 1)/2 = (13.5t + 13)/2 = 6.75t + 6.5$.
$I^2(O(I^2(x))) = (3(6.75t+6.5)+1)/2 = (20.25t+20.5)/2 = 10.125t + 10.25$.
$O(I^2(O(I^2(x)))) = (10.125t+10.25)/2 = 5.0625t+5.125 > 4t + 3 = x$.
$O^2(I^2(O(I^2(x)))) = (5.0625t + 5.125)/2 = 2.53125t + 2.5625 < 3t + 3 < 4t + 3 = x$.

Therefore, the reduced dynamics ends with "*IIOIIOO*". That is, CODE$(x \in [59]_{128}) = IIOIIOO = I^2OI^2O^2$. $\qquad \square$

Proposition 39. CODE$(x) = IIIIOOO$, *if*

$$x \in \left\{ x \mid x \in [3]_4, t = \frac{(x-3)}{4} \in [3]_{32} \right\}$$
$$= \{x \mid x = 4(32k + 3) + 3 = 128k + 15, k \in \mathbb{N}\} \qquad (30)$$
$$= [15]_{128}$$

Proof. $x = 4t + 3 \in [1]_2, t \in [3]_{32}$.
$I(x) = (3x + 1)/2 = (12t + 10)/2 = 6t + 5 \in [1]_2$.
$I^2(x) = (3(6t+5)+1)/2 = 9t+8 \in [(9*3+8) \bmod 32]_{32} = [3]_{32} \subset [1]_2$.
$I^3(x) \in [(3 * 3 + 1)/2]_{32/2} = [5]_{16} \subset [1]_2$.
$I^4(x) \in [((3 * 5 + 1) \bmod 16)/2]_{16/2} = [0]_8$, thus triple "*O*" will follow.
$I^2(x) = 9t + 8$.
$I^3(x) = (3(9t + 8) + 1)/2 = (27t + 25)/2 = 13.5t + 12.5$.
$I^4(x) = (3(13.5t+12.5)+1)/2 = (40.5t+38.5)/2 = 20.25t + 19.25$.
$O(I^4(x))) = (20.25t + 19.25)/2 = 10.125t + 9.625$.
$O^2(I^4(x))) = (10.125t + 9.625)/2 = 5.0625t + 4.8125 > 4t + 3 = x$.
$O^3(I^4(x))) = 2.53125t + 2.40625 < 3t + 3 < 4t + 3 = x$.
Therefore, CODE$(x \in [15]_{128}) = IIIIOOO = I^4O^3$. $\qquad \square$

In summary, aforementioned codes that have short lengths are listed in (31) as follows:

$$\text{CODE}(x)$$
$$= \begin{cases} IIOO & x \in [3]_{16}, \left(\dfrac{(x-3)}{4} = t \in [0]_4 \cup \{0\} \right), \\ IIOIO & x \in [11]_{32}, \left(\dfrac{(x-3)}{4} = t \in [2]_8 \right), \\ IIIOO & x \in [23]_{32}, \left(\dfrac{(x-3)}{4} = t \in [5]_8 \right), \\ IIIOIOO & x \in [7]_{128}, \left(\dfrac{(x-3)}{4} = t \in [1]_{32} \right), \\ IIIIOOO & x \in [15]_{128}, \left(\dfrac{(x-3)}{4} = t \in [3]_{32} \right), \\ IIOIIOO & x \in [59]_{128}, \left(\dfrac{(x-3)}{4} = t \in [14]_{32} \right). \end{cases} \quad (31)$$

Note that here short length means $|\text{CODE}(x)|$ is short, where $|\cdot|$ is the length of $\text{CODE}(x)$ measured by the count of "I" and "O" in $\text{CODE}(x)$. For example, $\text{CODE}(x \in [3]_{16}) = |IIOO| = 4$, $\text{CODE}(x \in [11]_{32}) = |IIOIO| = 5$.

In summary, (7) and (31) are presented together in (32) ($L = |\text{CODE}(x)|$). It justifies that $CODE$ is unique (recall Theorem 13), as all intersection sets for x in (32) are empty. The Format Theorem (Theorem 18) is confirmed as well.

$$\text{CODE}(x) = \begin{cases} O & L = 1, \ x \in [0]_2, \\ IO & L = 2, \ x \in [1]_4, \\ IIOO & L = 4, \ x \in [3]_{16}, \\ IIOIO & L = 5, \ x \in [11]_{32}, \\ IIIOO & L = 5, \ x \in [23]_{32}, \\ IIIOIOO & L = 7, \ x \in [7]_{128}, \\ IIIIOO & L = 7, \ x \in [15]_{128}, \\ IIOIIOO & L = 7, \ x \in [59]_{128}. \end{cases} \quad (32)$$

We discover that (32) enumerates all possible codes for $\{x \mid L = |\text{CODE}(x)| \le 7\}$.

Proposition 40. $\{x \mid L = |\text{CODE}(x)| \le 7\} = \{x \mid x \in [0]_2 \cup [1]_4 \cup [3]_{16} \cup [11]_{32} \cup [23]_{32} \cup [7]_{128} \cup [15]_{128} \cup [59]_{128}\}$.

Proof. If $|\text{CODE}(x)| = 3$, then $\text{CODE}(x) = IIO$ because Format Theorem (Theorem 18) and $CODE$ end by "O" (Proposition 25). However, $IIO(x) = (3(3x+1)/2+1)/2/2 = (3(1.5x+1.5)+1)/4 = 4.5/4x + 5.5/4 = 1.125x + 1.375 > x$. Thus, $|\text{CODE}(x)| \ne 3$.

If $|\text{CODE}(x)| = 4$, then $\text{CODE}(x) = II\|\{I,O\}\|O$ because Format Theorem and $CODE$ end by "O" (Proposition 25). Thus, $\text{CODE}(x) = IIIO$ or $\text{CODE}(x) = IIOO$. As $IIIO(x) = 6.75t + 6.25 > 4t + 3 = x$ (recall Proposition 35), $\text{CODE}(x) = IIIO$ is impossible. Hence, there exists only one type of code for $|\text{CODE}(x)| = 4$ (i.e., $\text{CODE}(x) = IIOO$).

Similarly, we can prove that there exists exactly two types of codes for $|\text{CODE}(x)| = 5$, since $\text{CODE}(x) = II\|\{I,O\}^2\|O$. $\{I,O\}^2 = \{IO, II, OO, OI\}$, but $\forall x \in [3]_4$, $IIIIO(x) > x$

and $IIOO(x) < x$. Thus, only "OI" and "IO" are possible. Similarly, we can prove that there exists three types of codes for $|\text{CODE}(x)| = 7$.

Besides, $|\text{CODE}(x)| = 6$ is impossible, as $\forall x \in [3]_4$, $IIIOIO(x) > x$, $IIIIOO(x) > x$, $IIOIIO(x) > x$, $IIOIO(x) < x$, and $IIIIOO(x) < x$ (and by Format Theorem and Proposition 25). $\qquad \square$

In the following, we use $[i_1, i_2, \ldots, i_n]_m$ to denote $\bigcup_{j=1}^n [i_j]_m$.

Corollary 41. *Let* $\overline{S} = [1]_2 \cup [0]_4 \cup [2]_{16} \cup [10, 22]_{32} \cup [6, 14, 58]_{128}$. $\forall x \le k, k \in \overline{S}, x \in \mathcal{RTN} \implies x = k + 1 \in \mathcal{RTN}$.

Proof. Let $S = [0]_2 \cup [1]_4 \cup [3]_{16} \cup [11, 23]_{32} \cup [7, 15, 59]_{128}$.

$k \in \overline{S} \implies k + 1 \in S \implies x \in S \because x = k + 1$
$\implies \exists \text{CODE}(x) \implies \exists L \in \mathbb{N}^*, CT^L(x) < x \because \text{CODE}(x) = CT^L$
$\implies CT^L(x) < k + 1 \because x = k + 1$
$\implies CT^L(x) \le k$
$\implies CT^L(x) \in \mathcal{RTN} \because \forall x \le k, x \in \mathcal{RTN}$
$\implies \exists L' \in \mathbb{N}^*, CT^{L'}(CT^L(x)) = 1$
$\implies \exists L'' = L' + L \in \mathbb{N}^*, CT^{L''}(x) = 1$
$\implies x \in \mathcal{RTN}. \qquad \square$

Indeed, $k \in [1]_2 \cup [0]_4$ have already been discussed in Propositions 9 and 10. Recall that $k \in [2]_4$ is the major concern in reduced version of induction for Collatz Conjecture; that is, $\forall x \le k = 4t + 2, x \in \mathcal{RTN} \implies x = k + 1 \in \mathcal{RTN}$, where $x, k, t \in \mathbb{N}^*$. We can prove that portion of $k = 4t + 2$ in induction is Returnable as follows.

Corollary 42. *Let* $\overline{S} = [2]_{16} \cup [10, 22]_{32} \cup [6, 14, 58]_{128}$.
$\forall x \le k, k \in \overline{S} \subset [2]_4, x \in \mathcal{RTN} \implies x = k + 1 \in \mathcal{RTN}$.

In other words, we can prove that portion of t is Returnable as follows:

Corollary 43. *Let* $T = [2, 5]_8 \cup [1, 3, 14]_{32}$.
$\forall x \le k, k = 4t + 2, t \in T, x \in \mathcal{RTN} \implies x = k + 1 \in \mathcal{RTN}$.

Thus, we have already proved that portion of induction cases is Returnable. For the proof of Collatz Conjecture, we need to prove that all cases are Returnable. That is, all cases for $k = 4t + 2$ are Returnable to make $\overline{S} = [2]_4$, or all cases for t are Returnable to make $T = \mathbb{N}^*$.

Note that, indeed, without relying on the induction in Section 3, we only need to prove $\forall x \in \mathbb{N}^*, x \in \mathcal{RTN}$. As CC is equivalent to RCC (recall Proposition 4), we thus solely need to prove RCC is true. In other words, $\forall x \in \mathbb{N}^*, \exists \text{CODE}(x)$ (recall Remark 14 (3)).

6. Conclusion

In this paper, we propose a new direction for proving Collatz Conjecture, by proving Reduced Collatz Conjecture. Reduced Collatz Conjecture is equivalent to Collatz Conjecture but

easier to explore for inherent properties. It is because all dynamics going to transformed number 1 will consist of multiple reduced dynamics, in which transformed number is less than the corresponding starting number instead of 1.

We also present an induction method for Collatz Conjecture for better understanding of Reduced Collatz Conjecture and reduced version of the induction that is easier to tackle.

We denoted reduced dynamics as $\mathsf{CODE}(x)$ (called code) and explore some fundamental properties of it, especially the structure of reduced dynamics (i.e., unified format concatenated by regular segments).

The starting numbers and their codes whose lengths are no more than 7 (i.e., $\{(x, c) \mid c = \mathsf{CODE}(x), |c| \leq 7\}$) are also given and proved; thus, portion of $x \in [3]_4$ has already guaranteed Reduced Collatz Conjecture. That is, $x \in [3]_{16} \cup [11, 23]_{32} \cup [7, 15, 59]_{128} \implies \exists \mathsf{CODE}(x))$. (Recall that codes for $x \in [1]_4$ and $x \in [0]_2$ are trivial - $\mathsf{CODE}(x \in [1]_4) = IO$ and $\mathsf{CODE}(x \in [0]_2) = O$.)

The future work for the proof of Collatz Conjecture can follow this direction, to just prove that $\mathsf{CODE}(x)$ exists for the left portion of $x \in [3]_4$. Indeed, we also discovered the bound between the counts of I and the counts of O for any valid code is related to $\log_2 3$, the period T in terms of $\mathsf{CODE}(x + T) = \mathsf{CODE}(x)$ is related to the length of $\mathsf{CODE}(x)$, and how to compute residue class of x directly when $\mathsf{CODE}(x)$ is given, which will be presented in our other papers.

Data Availability

The data (Source Code in C, Computer Program Outputs) used to support the findings of this study have been deposited in the [1-6] repository. (1) Wei Ren, Exploring properties in Reduced Collatz Dynamics, IEEE Dataport, 2018. https://doi.org/10.21227/ge9h-8a77 (2018). (2) Wei Ren, Verifying whether extremely large integer guarantees Collatz Conjecture (can return to 1 finally), IEEE Dataport, https://doi.org/10.21227/fs3z-vc10 (2018). (3) Wei Ren, Exploring the ratio between the count of x/2 and the count of (3∗x+1)/2 in original dynamics for extremely large starting integers asymptotically, IEEE Dataport. https://doi.org/10.21227/rxx6-8322 (2018). (4) Wei Ren, Exploring the inverse mapping from a dynamics to a residue class - inputting a reduced dynamics or partial dynamics and outputting a residue class, IEEE Dataport. https://doi.org/10.21227/qmzw-gn71 (2018). (5) Wei Ren, Reduced Collatz Dynamics for Integers from 3 to 999999, IEEE Dataport. https://doi.org/10.21227/hq8c-x340 (2018) (6) Wei Ren, Collatz Automata and Compute Residue Class from Reduced Dynamics by Formula, IEEE Dataport. https://doi.org/10.21227/7z84-ms87 (2018).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] T. Oliveira e Silva, "Maximum excursion and stopping time record-holders for the 3x + 1 problem: computational results," *Mathematics of Computation*, vol. 68, no. 225, pp. 371–384, 1999.

[2] L. Colussi, "The convergence classes of Collatz function," *Theoretical Computer Science*, vol. 412, no. 39, pp. 5409–5419, 2011.

[3] P. C. Hew, "Working in binary protects the repetends of 1/3h: Comment on Colussi's 'The convergence classes of Collatz function'," *Theoretical Computer Science*, vol. 618, pp. 135–141, 2016.

[4] R. K. Guy, "Don't try to solve these problems," *Computers and Mathematics with Applications*, vol. 90, no. 1, pp. 35–41, 1983.

[5] G. T. Leavens and M. Vermeulen, "3x + 1 search programs," *Computers & Mathematics with Applications. An International Journal*, vol. 24, no. 11, pp. 79–99, 1992.

[6] R. E. Crandall, "On the "3x+1" problem," *Mathematics of Computation*, vol. 32, no. 144, pp. 1281–1292, 1978.

[7] W. Ren, S. Li, R. Xiao, and W. Bi, "Collatz Conjecture for 2100000-1 is true - algorithms for verifying extremely large numbers," in *Proceedings of the IEEE UIC 2018*, pp. 411–416, Guangzhou, China, October 2018.

[8] I. Krasikov and J. C. Lagarias, "Bounds for the 3x + 1 problem using difference inequalities," *Acta Arithmetica*, vol. 109, no. 3, pp. 237–258, 2003.