Theoretical Computer Science 412 (2011) 5409–5419

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

The convergence classes of Collatz function

Livio Colussi *

Department of Pure and Applied Mathematics, University of Padova, via Trieste, 63, 35121 Padova, Italy

ARTICLE INFO

Article history:
Received 4 August 2010
Received in revised form 13 May 2011
Accepted 25 May 2011
Communicated by D. Perrin

Keywords:
Collatz conjecture
$3n + 1$ problem

ABSTRACT

The Collatz conjecture, also known as the $3x + 1$ conjecture, can be stated in terms of the reduced Collatz function $R(x) = (3x + 1)/2^h$ (where 2^h is the larger power of 2 that divides $3x + 1$). The conjecture is: *Starting from any odd positive integer and repeating* $R(x)$ *we eventually get to* 1. G_k, the k-th convergence class, is the set of odd positive integers x such that $R^k(x) = 1$.

In this paper an infinite sequence of binary strings s_h of length $2 \cdot 3^{h-1}$ (the *seeds*) are defined and it is shown that the binary representation of all $x \in G_k$ is the concatenation of k periodic strings whose periods are $s_k, \dots, s_1$. More precisely $x = s_{k,d_{k,1}}^{[n_1]} \dots s_{1,d_{k,k}}^{[n_k]}$ where $s_{k,d_{k,i}}^{[n_i]}$ is the substring of length n_i that starts in position $d_{k,i}$ in a sufficiently long repetition of the seed s_i.

Finally, starting positions $d_{k,i}$ and lengths n_i for which $s_{k,d_{k,1}}^{[n_1]} \dots s_{1,d_{k,k}}^{[n_k]} \in G_k$ are defined, thus giving a complete characterization of classes G_k.

1. Introduction

The Collatz function is defined on all positive integers x by:

$$f(x) = \begin{cases} x/2 & x \text{ even} \\ 3x + 1 & x \text{ odd.} \end{cases}$$

Given any odd integer x, let $x' = (3x + 1)/2^h$ where 2^h is the highest power of 2 that divides $3x + 1$. The reduced form of the Collatz function is $R(x) = x'$ and is defined only for odd integers.

The Collatz conjecture says that for all integers $x > 0$ there exists i such that $f^i(x) = 1$ or, equivalently, that there exists k such that $R^k(x) = 1$.

Despite the efforts of many people for about seventy years, the conjecture is still undecided. The efforts are well documented in a very large literature. The problem has been attacked from many viewpoints. The Collatz function has been studied in large domains: Integer, rational, real and even complex numbers (where a beautiful fractal has been obtained) [5,3,4,9]. The Collatz conjecture has been also proved equivalent to many other conjectures in different contexts: Rewriting systems, tag systems, etc. [7,2,6].

Our bibliography contains only a very small and incomplete selection of papers; we refer interested readers to the large annotated bibliography in Lagarias [1]. The paper by Jean Paul Van Bendegem [10] is a philosophical essay on the $3x + 1$ problem.

The paper is organized as follows: Section 2 shows the direct computation of G_k, as sets of binary strings, for the first few values of k. Those computational experiments suggest that binary strings in G_k are the concatenation of k periodic strings whose periods, that we call *seeds*, are of length $2, 6, 18, \dots, 2 \cdot 3^{k-1}$. In Section 3 some useful (and beautiful) properties of seeds are proved. Section 4 contains the main result: A complete characterization of classes G_k as sets of binary strings.

* Tel.: +39 049 8271484.
E-mail address: colussi@math.unipd.it.

0304-3975/$ – see front matter
doi:10.1016/j.tcs.2011.05.056

2. Computational experiments

Define the inverse $R^{-1}(x)$ of the reduced Collatz function as the set of odd integers such that $y \in R^{-1}(x)$ iff $R(y) = x$. We can easily see that

$$R^{-1}(x) = \begin{cases} \emptyset & \text{if } x \equiv 0 \pmod 3 \\ \left\{ \dfrac{x2^{2m+2} - 1}{3} \ : \ m \geq 0 \right\} & \text{if } x \equiv 1 \pmod 3 \\ \left\{ \dfrac{x2^{2m+1} - 1}{3} \ : \ m \geq 0 \right\} & \text{if } x \equiv 2 \pmod 3. \end{cases}$$

Let G_k the class of odd integers x that converge to 1 in k steps, i.e. such that $R^k(x) = 1$.
The class G_k can be defined inductively by

$$G_0 = \{1\}$$
$$G_k = \bigcup_{x \in G_{k-1}} R^{-1}(x).$$

For a binary string s let $[\![s]\!]$ be the non-negative integer whose binary representation is s. In what follows we see classes G_k as sets of binary strings.

Clearly $G_0 = \{1\}$: The singleton set that contains only the binary string 1.

Let us compute first G_1

$$G_1 = \bigcup_{x_0 \in G_0} R^{-1}(x_0) = R^{-1}(1) = \left\{ \frac{4^{m_1+1} - 1}{3} \ : \ m_1 \geq 0 \right\} = \left\{ \sum_{i=0}^{m_1} 4^i \ : \ m_1 \geq 0 \right\}.$$

If we represent $x_1 = \sum_{i=0}^{m_1} 4^i$ as a binary string of length $2m_1 + 2$ we obtain 01^{m_1+1}, i.e. the concatenation of one or more copies of the binary string $s_1 = 01$ of length 2. Thus

$$G_1 = \left\{ [\![s_1^{m_1+1}]\!] \ : \ m_1 \geq 0 \right\}.$$

Now we can compute G_2 from G_1.

$$G_2 = \bigcup_{x_1 \in G_1} R^{-1}(x_1) = \bigcup_{m_1=0}^{\infty} R^{-1}\left([\![s_1^{m_1+1}]\!] \right).$$

Since $x_1 = \sum_{i=0}^{m_1} 4^i \equiv m_1 + 1 \pmod 3$ we obtain

$$G_2 = \left\{ \frac{[\![s_1^{3k_1+1}]\!] \, 4^{m_2+1} - 1}{3} \ : \ k_1, m_2 \geq 0 \right\} \cup \left\{ \frac{2\,[\![s_1^{3k_1+2}]\!] \, 4^{m_2} - 1}{3} \ : \ k_1, m_2 \geq 0 \right\}.$$

Compute first

$$\frac{[\![s_1^3]\!]}{3} = \frac{\sum_{i=0}^{2} 4^i}{3} = \frac{4^3 - 1}{3^2} = 7$$

and let $s_2 = 000111$ be the binary representation of 7 as a string of length 6.

A simple computation shows that $[\![s_1^{3k_1+1}]\!]/3 = [\![s_2^{k_1} s_2^{[2]}]\!]$, where $s_2^{[2]} = 00$ is the prefix of length 2 of s_2 and that $[\![s_1^{3k_1+2}]\!]/3 = [\![s_2^{k_1} s_2^{[4]}]\!]$, where $s_2^{[4]} = 0001$ is the prefix of length 4 of s_2. Moreover, $[\![s_1^{3k_1+1}]\!] \bmod 3 = 1$ and $[\![s_1^{3k_1+2}]\!] \bmod 3 = 2$.

Then

$$G_2 = \left\{ [\![s_2^{k_1} s_2^{[2]}]\!] \, 4^{m_2+1} + \frac{4^{m_2+1} - 1}{3} \ : \ k_1, m_2 \geq 0 \right\}$$
$$\cup \left\{ 2\,[\![s_2^{k_1} s_2^{[4]}]\!] \, 4^{m_2} + \frac{4^{m_2+1} - 1}{3} \ : \ k_1, m_2 \geq 0 \right\}$$

We can write $(4^{m_2+1} - 1)/3 = \sum_{i=0}^{m_2} 4^i$ in binary both as $[\![s_1^{m_2} s_1^{[2)}]\!]$ and $[\![s_{1,1}^{m_2} s_{1,1}^{[1)}]\!]$, where $s_{1,1} = 10$ is the left rotation of s_1 by 1 position.