An Artificial Life View of the Collatz Problem

Hiroki Sayama*
Collective Dynamics of Complex Systems Research Group, Binghamton University

Abstract This letter presents a new, artificial-life-based view of the Collatz problem, a well-known mathematical problem about the behavior of a series of positive integers generated by a simple arithmetical rule. The Collatz conjecture asserts that this series always falls into a 4 → 2 → 1 cycle regardless of its initial values. No formal proof has been given yet. In this letter, the behavior of the series is considered an ecological process of artificial organisms (1s in bit strings). The Collatz conjecture is then reinterpreted as the competition between population growth and extinction. This new interpretation has made it possible to analytically calculate the growth and extinction speeds of bit strings. The results indicate that the extinction is always faster than the growth, providing an ecological explanation for the conjecture. Future research directions are also suggested.

Keywords
Collatz problem, ecological interpretation, spatiotemporal patterns, bit strings

In artificial life, it is usually the case that mathematical and computational models are developed for the understanding of biological problems. This letter attempts to demonstrate that the opposite is also possible—to develop biological models for the understanding of mathematical and computational problems. I consider the Collatz problem as a first such example.

The Collatz problem, also known as the $3x + 1$ problem [3], discusses the behavior of a series that starts with an arbitrary positive integer x_0 and develops according to the following rule:

$$x_{t+1} = \begin{cases} 3x_t + 1 & \text{if } x_t \text{ is odd,} \\ x_t/2 & \text{if } x_t \text{ is even.} \end{cases} \tag{1}$$

The Collatz conjecture asserts that this series always falls into a $4 \rightarrow 2 \rightarrow 1$ cycle regardless of x_0, which is believed to be true by many but has defied any formal proof for more than 70 years [4, 5].

Here I propose a new perspective on the Collatz problem by considering it an ecological process of artificial organisms (1s in bit strings) and studying the spatiotemporal dynamics of their patterns. To make this approach easier, I ignore the second condition in Equation 1, because it only shifts bit strings rightward with no influence on their patterns. Ignoring it converts the series into a simpler iterative map with no ifs:

$$x_{t+1} = 3x_t + \mathrm{LSNB}(x_t). \tag{2}$$

Here LSNB(x) is the *least significant nonzero bit* of x (e.g., LSNB(172) = LSNB(*10101100*) = *100* = 4; *italics* are binary representations).

* Departments of Bioengineering & Systems Science and Industrial Engineering, Binghamton University, State University of New York, P.O. Box 6000, Binghamton, New York 13902-6000, USA. E-mail: sayama@binghamton.edu

The above formula can be interpreted in ecological terms. A bit string of x_t represents the population distribution at time t, where 1s are living organisms and 0s are empty sites. $3x_t$ in Equation 2 represents the replication of those organisms because it literally replicates each single bit (Figure 1a). This causes leftward growth of the bit string as well as overcrowding of bits whose effects propagate leftward, depending on the carry rule. Also, LSNB(x_t) in Equation 2 represents an external perturbation continuously introduced into the population, which causes extinction of the living organisms residing at the rightmost end, making the nonzero region of the bit string shrink from the right (Figure 1b).

These interpretations suggest that the Collatz problem is about a competition between growth and extinction of the nonzero region in their speeds (Figure 2a).

As seen in Figure 2, the leftward growth of the nonzero region is almost linear, and its speed, L (bits/step), is approximated by

$$L \approx \langle \log_2 x_{t+1} - \log_2 x_t \rangle \tag{3}$$

$$= \left\langle \log_2 \frac{3x_t + \mathrm{LSNB}(x_t)}{x_t} \right\rangle \tag{4}$$

$$= \left\langle \log_2 \left(3 + \frac{\mathrm{LSNB}(x_t)}{x_t} \right) \right\rangle, \tag{5}$$

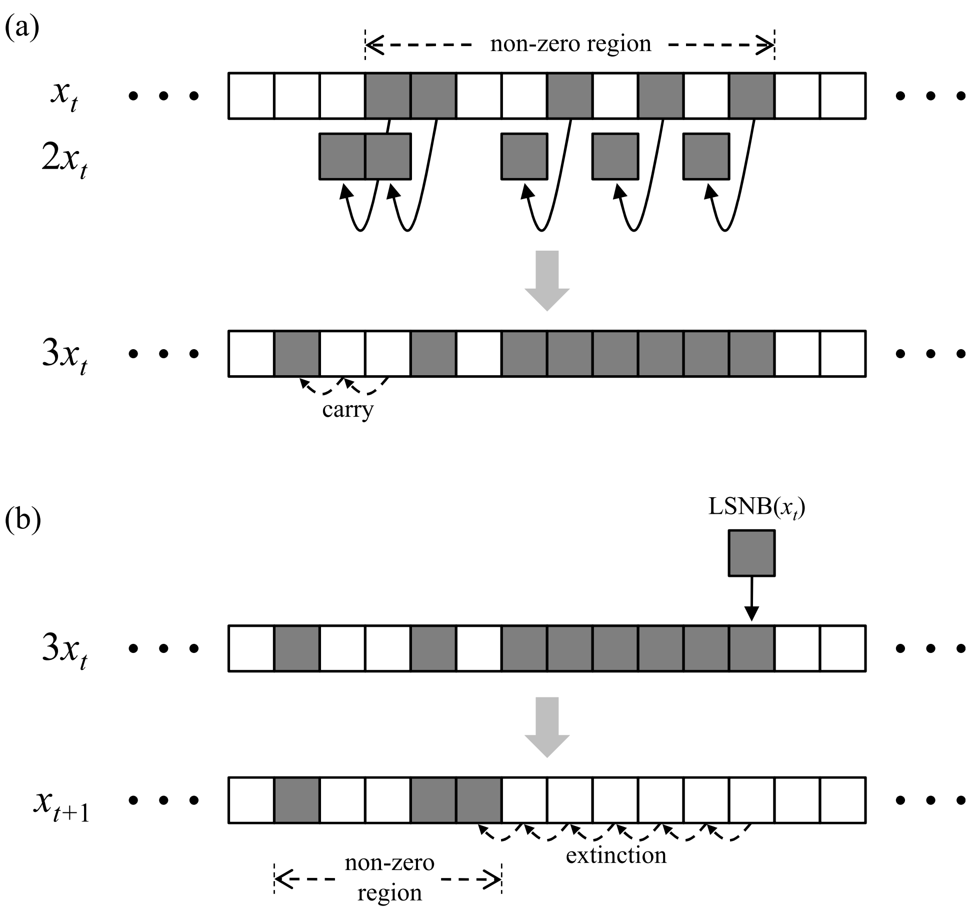

Figure 1. The Collatz problem as an ecological process of artificial organisms represented in bit strings. (a) Replication of 1s (gray cells) and growth of the nonzero region caused by $3x_t$. (b) Extinction from the right caused by LSNB(x_t).

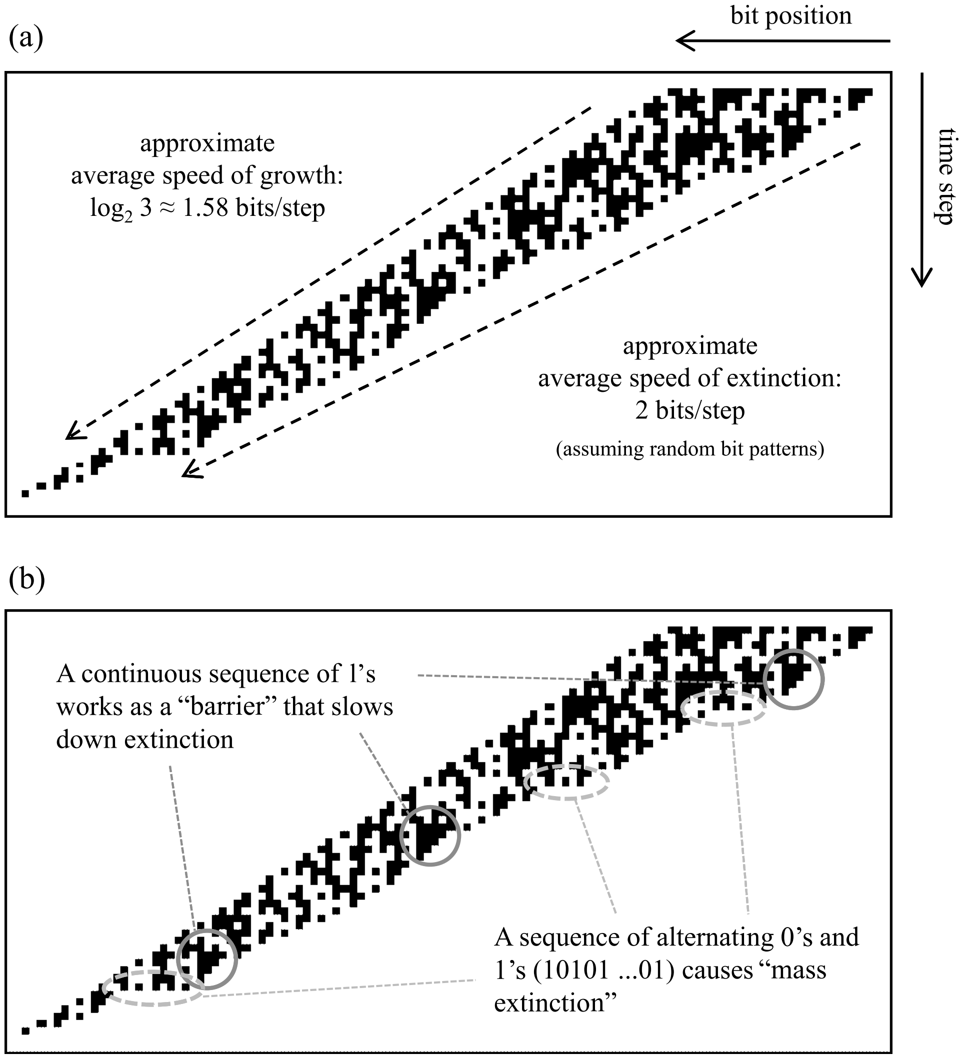

Figure 2. (a) Spatiotemporal dynamics of a sample series (starting with x_0 = 111111111) visualized as bit patterns. (b) Some features commonly observed in spatiotemporal patterns that influence the speed of extinction.

where $\langle \cdots \rangle$ represents time average. This shows that the maximal value of L is $\log_2 4 = 2$ bits/step, which is sustainable if and only if LSNB(x_t) = x_t always, that is, if the population consists of a single nonzero bit. Usually LSNB(x_t) is much smaller than x_t when the nonzero region is reasonably long (LSNB(x_t) $\ll$ x_t)). Therefore, an approximate value of L for sufficiently long nonzero bit patterns is given by $L_{\text{app}} = \log_2 3 \approx 1.58$ bits/step, which was confirmed by computer simulations.

On the contrary, the extinction of the nonzero region from the right is more complex, because some features in the spatiotemporal patterns influence the speed of extinction significantly (Figure 2b). For example, a sequence of alternating 0s and 1s (e.g., 10101…01) disappears at once no matter how long it is, causing a "mass extinction" of the population. This means that there is no theoretical maximum for the speed of extinction. Also, a continuous sequence of 1s works as a kind of "barrier" that slows down the extinction. It is generally difficult to predict when and where these features arise in the spatiotemporal patterns, because of the chaotic nature of the Collatz series.

However, by assuming that 0s and 1s appear in bit strings with equal probability, one can approximate the probabilities for the extinction to stop or continue to be 50/50 at each location in the bit string. With this assumption, the approximate speed of extinction, R_{app} (bits/step), is analytically calculated as

$$R_{app} = \sum_{l=1}^{\infty} l \left(\frac{1}{2}\right)^{l} = 2 \qquad \text{(bits/step)}, \tag{6}$$

which was also confirmed by computer simulations. These results indicate that the extinction from the right is "faster" than the population growth to the left, providing an ecological explanation of why the series always fall into a single-bit cycle.

In conclusion, I provided a new, biologically inspired perspective to the Collatz problem. Ecological interpretations and visual representations are proposed to facilitate intuitive understanding of the dynamics and to encourage interdisciplinary participation by nonmathematicians in this interesting mathematical problem. Another approach to the same problem inspired by artificial life has recently been reported by Eubanks [1], supporting my claim that artificial life has potential to contribute to this and other mathematical problems in multiple ways.

Finally, I would like to make it clear that the argument about the competition of the growth and extinction speeds developed above is still not a rigorous proof of the Collatz conjecture, because it assumed stochasticity in bit patterns. The artificial life community could also contribute to the Collatz problem by searching for counterexamples to the conjecture. It may be possible to design, either manually or evolutionally, specific bit strings that slow down the extinction by continuously producing "barriers," which might be possible with very long bit lengths. Similar phenomena (minimal complexity for nontrivial behavior) have been discovered and studied in various contexts, such as the self-reproduction ability realized by complex configurations in von Neumann's cellular automata [6] and the self-simulation ability realized by complex rules in Gács's cellular automata [2].

References

1. Eubanks, D. (2010). Collatz ecologies. *Higher Ed* (blog), November 21, 2010. http://highered.blogspot.com/2010/11/collatz-ecologies.html
2. Gács, P. (2001). Reliable cellular automata with self-organization. *Journal of Statistical Physics*, *103*(1–2), 45–267.
3. Lagarias, J. C. (1985). The 3*x* + 1 problem and its generalizations. *American Mathematical Monthly*, *92*(1), 3–23.
4. Lagarias, J. C. (2003). The 3*x* + 1 problem: An annotated bibliography (1963–1999). http://arxiv.org/abs/math/0309224
5. Lagarias, J. C. (2006). The 3*x* + 1 problem: An annotated bibliography, II (2000–). http://arxiv.org/abs/math/0608208
6. von Neumann, J. (1966). *Theory of self-reproducing automata*. Champaign, IL: University of Illinois Press.