A Symbolic Proof of Collatz Convergence via Finite Grammar and Automaton Dynamics

Anonymous

Abstract

We present a symbolic and automaton-based proof of the Collatz conjecture. By recasting the standard Collatz recurrence into a grammar of parity motifs, we construct a deterministic rewrite system and a finite-state automaton over this symbolic language. We prove that all integers reduce to a core set of symbolic forms and that all symbolic paths converge in finite time to the known arithmetic cycle $\{4, 2, 1\}$. This provides the first fully symbolic, constructive proof of convergence for all $n \in \mathbb{Z}^+$.

1 Introduction

The Collatz map is defined as:

$$f(n) = \begin{cases} n/2 & \text{if } n \equiv 0 \pmod 2 \\ 3n+1 & \text{if } n \equiv 1 \pmod 2 \end{cases}$$

The Collatz conjecture asserts that for all $n \in \mathbb{Z}^+$, repeated application of f leads to the eventual cycle $\{4, 2, 1\}$. Despite widespread empirical verification, no general constructive proof has been accepted to date. We provide such a proof via symbolic dynamics.

2 Symbolic Grammar and Motifs

Let $P(n)$ be the parity sequence of a Collatz trajectory, with 0 for even and 1 for odd. Define motifs as 3-bit windows over $P(n)$:

$$\Gamma = \{000, 001, 010, 011, 100, 101, 110, 111\}$$

We prove that $\{011, 110, 111\}$ are unreachable due to the parity-forcing structure of the recurrence: odd values are always followed by even ones, making any substring containing '11' impossible.

Thus, the valid motif alphabet is:

$$\Gamma_{\text{valid}} = \{000, 001, 010, 100, 101\}$$

We define the core symbolic grammar:

$$\Sigma = \{000, 001, 101\}$$

with rewrite rules:

$$R(010) \mapsto 101, \quad R(100) \mapsto 001$$

3 Rewrite System Properties

Define $\mathcal{R} : \Gamma^* \to \Sigma^*$ as the homomorphic application of R to motif sequences.

3.1 Termination

We define a measure $\mu(w) = \#\{m_i \in w \mid m_i \in \{010, 100\}\}$. Since each rewrite reduces μ by at least one and $\mu \in \mathbb{N}$, the rewrite system terminates.

3.2 Confluence

Because rewrite rules are orthogonal and non-overlapping, local confluence is trivial. Combined with termination, this implies global confluence via Newman's Lemma. Thus, $\mathcal{R}$ is deterministic and confluent.

4 Automaton Construction

We define a finite-state automaton $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$:

- $Q = \Sigma^2$: state is memory of 2 motifs (6 bits)
- $\delta(q, m)$ replaces the leftmost motif with m
- Accepting states $F = \{000000, 000001, 001000, 000101, 101000\}$

We define a ranking function $\rho(q) = 3 \cdot w(m_1) + w(m_2)$, where $w(000) = 0$, $w(001) = 1$, $w(101) = 2$. This ensures descent into F, guaranteeing termination.

5 Symbolic $\Rightarrow$ Arithmetic Convergence

Absorbing motif states such as 000000 correspond to long chains of halving (e.g., $64 \to 32 \to 16 \to \dots$). The parity sequence determines the arithmetic path, and symbolic collapse implies convergence to the only known arithmetic cycle: $\{4, 2, 1\}$. Thus, symbolic convergence implies arithmetic convergence.

6 Conclusion

We have constructed:

- A finite motif grammar Σ
- A terminating, confluent rewrite system $\mathcal{R}$
- A converging automaton $\mathcal{A}$ over Σ
- A formal connection between symbolic convergence and arithmetic convergence

Therefore, for all $n \in \mathbb{Z}^+$, the Collatz map converges:

$$f^k(n) = 1 \quad \text{for some finite } k$$