Department of Computer Science

# New Results in Curling Number Sequences

Riwaz Poudyal'18, Adly Templeton, and Duane Bailey

## The Curling Number Conjecture

### Curling Number

The *curling number* of a finite, non-empty sequence of integers S is the maximum number of times any suffix repeats, starting from the right. If we write S as $AB^K$ such that B is the suffix that repeats the greatest number of times, K is the curling number of S.

### The Curling Number Conjecture

If we repeatedly extend S by appending its curling number to the end, we generate a curling number sequence. We terminate this growth process when we encounter a curling number of 1, and call the initial sequence the *starting sequence* and the curling numbers we append to it its *tail*. The conjecture proposes that every starting sequence will have a finite tail.

For the sequence 323232, the Curling Number is 3 as the suffix '32' repeats 3 times. If we keep on appending the Curling Number to the end, we reach a 1 in the fourth step, and the tail-length is 3.

**323232-3-3-2-1**

Research has focused on the simplest case when there are only 2s and 3s in the starting sequence.

## Especially Good Starting Sequences

EGSs are sequences with the longest tail for a given starting length. Sequences with long tails help us understand how even good sequences reach a 1.

Direct exhaustive search to find these sequences is computationally intractable as the search space grows exponentially.

For example, there are 280 trillion possible sequences of length 48 and analyzing all these takes dozens of CPU years.

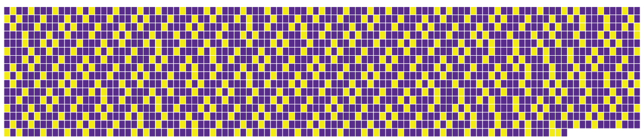
### Algorithm to Find Good Starting Sequences

Analyzing the structure of EGSs up to length 78, which were computed by authors of [1], we created a parallel rewriting grammar that can be used to generate sequences with relatively long tails.

This sequence of length 197, which we diagram using purple as 2 and gold as 3, has tail-length 1668 and tail-length/starting-length ratio of 8.47:

**Sequence:**

**Tail:**



## Forcing to Find Good Sequences

Using a grammar to describe patterns in sequences allows us to identify how curling numbers are 'forced' to specific values.

For Example:

**-2-2-2-2322--22--22XXX222322232232-2-2-2-2**

Whenever a '323' pattern, denoted by 'XXX' above, appears in any sequence, a pattern '222322232232' must appear after it. The '323' pattern must also be preceded by a '22'.

This property can be used to construct sequences that have a relatively long tail.

## High Level Forcing

Using forcing to generate the grammar below and narrow down the search space, we were able to generate better candidates for EGSs (Proof of their uniqueness remains intractable).

$$P \longrightarrow XYZ$$

$$Y \longrightarrow X$$

$$ZX \longrightarrow YX$$

where P is '223', X is '2223', Y is '232223', and Z is '2232223', and the character before the arrow forces the character after it.

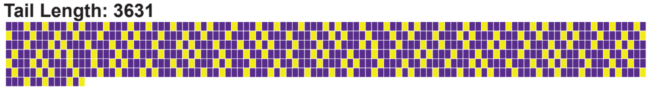
## Best Sequences So Far

We found more than half a dozen sequences with a tail longer than 3000 characters and a significant tail-length to starting-length ratio.

**Sequence Length: 290** **Tail/Starting-seq ratio: 11.67**
**Tail Length: 3383**

**Sequence Length: 385** **Tail/Starting-seq ratio: 9.24**
**Tail Length: 3559**

**Sequence Length: 643** **Tail/Starting-seq ratio: 5.64**
**Tail Length: 3631**



Before this, the sequence with the longest known tail had a starting length of 227 and a tail length of 369. The sequence with the best tail-length to starting-length ratio of 10.18 was:

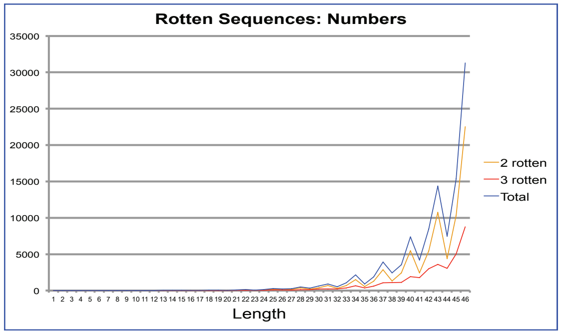
**Sequence Length: 11**
**Tail Length: 112**

## Rotten Sequences

A sequence S is *rotten* if the sequence formed by appending either a 2 or a 3 prefix to S (2S or 3S) has a *shorter* tail length than S. *2-rotten* sequences have shorter tail length when prefixed by 2 and *3-rotten* sequences have shorter tail length when prefixed by 3.

These sequences are rare, but help us understand how prefixes and subsequences affect a sequence's tail length.

Using dynamically optimized algorithm and spending about 196 CPU years, we have computed all rotten sequences with starting length less than or equal to 48

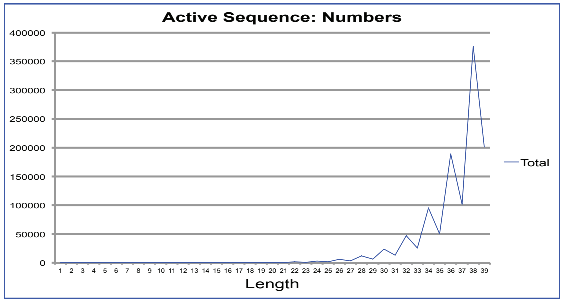



As conjectured by authors of [1], we found no doubly rotten sequences--sequences that are both 2-rotten and 3-rotten--up to length 48.

## Active Sequences

*Active sequences* are the opposites of rotten sequences: their tail length increase when prefixed by a 2 or 3.

We have also computed all active sequences up to the starting length of 41.




No doubly-active sequences were found, which means for every sequence up to length 41, either one or both of 2 or 3 prefix is neutral. The relation between active and rotten sequences remains a subject of research.

**References:**
[1] On Curling Number of Sequences, http://arxiv.org/abs/1212.6102

**Acknowledgment:**
We would like to thank Division III research funding committee and the department of Computer Science for computing resources