Journal of Integer Sequences, Vol. 10 (2007),
Article 07.1.2

A Slow-Growing Sequence Defined by an Unusual Recurrence

Fokko J. van de Bult and Dion C. Gijswijt
Korteweg-de Vries Institute for Mathematics
University of Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
Netherlands

John P. Linderman, N. J. A. Sloane,[1] and Allan R. Wilks
AT&T Shannon Labs
180 Park Avenue
Florham Park, NJ 07932–0971
USA
fjvdbult@science.uva.nl
dion.gijswijt@gmail.com
jpl@research.att.com
njas@research.att.com
allan@research.att.com

Abstract

The sequence starts with $a(1) = 1$; to extend it one writes the sequence so far as XY^k, where X and Y are strings of integers, Y is nonempty and k is as large as possible: then the next term is k. The sequence begins 1, 1, 2, 1, 1, 2, 2, 2, 3, 1, 1, 2, 1, 1, 2, 2, 2, 3, 2, ... A 4 appears for the first time at position 220, but a 5 does not appear until about position $10^{10^{23}}$. The main result of the paper is a proof that the sequence is unbounded. We also present results from extensive numerical investigations of the sequence and of certain derived sequences, culminating with a heuristic argument that t (for $t = 5, 6, \ldots$) appears for the first time at about position

[1]All correspondence should be directed to this author.

$2 \uparrow (2 \uparrow (3 \uparrow (4 \uparrow (5 \uparrow \ldots \uparrow ((t-2) \uparrow (t-1))))))$, where $\uparrow$ denotes exponentiation. The final section discusses generalizations.

1 Introduction

This paper introduces an integer sequence $A = a(1),\ a(2),\ a(3), \ldots$ with some remarkable properties. Define the *curling number* $\mathcal{C}(U)$ of a string $U = u(1),\ u(2), \ldots, u(n)$ over some alphabet Ω to be the largest integer $k \geq 1$ such that

$$U = X \underbrace{YY \cdots Y}_{k \text{ copies}} = XY^k\,, \tag{1}$$

where X and Y are strings over Ω and Y is nonempty. Our sequence is defined by

$$a(1) = 1,\ a(n+1) = \mathcal{C}(a(1), \ldots, a(n)) \quad \text{for} \quad n \geq 1\,. \tag{2}$$

Then $a(2) = \mathcal{C}(1) = 1$, since we can only take X to be the empty string ϵ, $Y = 1$ and $k = 1$; $a(3) = \mathcal{C}(1,1) = 2$, by taking $X = \epsilon$, $Y = 1$, $k = 2$; $a(4) = \mathcal{C}(1,1,2) = 1$, by taking $X = 1,1$, $Y = 2$, $k = 1$ (as this example shows, there may be more than one choice for Y); and so on. The first 220 terms of A are shown in Tables 1 and 2.

To avoid any possible confusion, for example with the "Say What You See" sequence studied in [3], we emphasize that the curling number does *not* depend on the decimal representation of its arguments. For example, if $U = (8, 9, 10, 11, 11, 11)$, $\mathcal{C}(U) = 3$.

1	1	$\underline{\mathbf{2}}$																	
1	1	**2**	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$														
1	1	2																	
1	1	**2**	**2**	**2**	**3**	$\underline{\mathbf{2}}$													
1	1	2																	
1	1	2	2	2	3														
1	1	2																	
1	1	**2**	**2**	**2**	**3**	**2**	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{2}}$				
1	1	2																	
1	1	2	2	2	3														
1	1	2																	
1	1	2	2	2	3	2													
1	1	2																	
1	1	2	2	2	3														
1	1	2																	
1	1	**2**	**2**	**2**	**3**	**2**	**2**	**2**	**3**	**2**	**2**	**2**	**3**	**3**	**2**	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{2}}$

Table 1: The first 98 terms of the sequence. In the notation to be introduced in Section 2, the five underlined strings are the glue strings $S_1^{(1)}, S_2^{(1)}, \ldots, S_5^{(1)}$ and the five bold-face strings are $T_2^{(1)}, T_3^{(1)}, \ldots, T_6^{(1)}$.

1	1	2																									
1	1	2	2	2	3																						
1	1	2																									
1	1	2	2	2	3	2																					
1	1	2																									
1	1	2	2	2	3																						
1	1	2																									
1	1	2	2	2	3	2	2	2	3	2	2	2	3	3	2												
1	1	2																									
1	1	2	2	2	3																						
1	1	2																									
1	1	2	2	2	3	2																					
1	1	2																									
1	1	2	2	2	3																						
1	1	2																									
1	1	**2**	**2**	**2**	**3**	**2**	**2**	**2**	**3**	**2**	**2**	**2**	**3**	**3**	**2**	**2**	**2**	**3**	**2**	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{3}}$
		$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{2}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{4}}$										

Table 2: Terms 99 through 220 of the sequence, up to the point where the first 4 appears ($S_6^{(1)}$ is underlined, $T_7^{(1)}$ is shown in bold-face).

In Section 2 we describe the recursive structure of the sequence, in particular explaining the block structure visible in Tables 1 and 2. The proof that this structure is valid is postponed to Section 3, where we give the main results of the paper, Theorems 3.1 and 3.2. Corollary 3.4 shows that the sequence is unbounded.

In Section 4 we give empirical estimates for the lengths of the blocks in the recursive structure, culminating in the estimate that $t \geq 5$ appears in the sequence for the first time at about position

$$2^{2^{3^{4^{\cdot^{\cdot^{\cdot^{t-1}}}}}}}, \tag{3}$$

a tower of height $t-1$. These estimates are based on examination of the first two million terms of the sequence A and of the higher-order sequences $A^{(2)}$, $A^{(3)}$ and $A^{(4)}$ introduced in Section 2.

The final section is devoted to comments and generalizations. §5.1 discusses a certain plausible "Finiteness Conjecture" that arises from studying curling numbers. §5.2 discusses sequences that are obtained when the "curling number transform" (defined below) is applied to certain well-known sequences. Finally, §5.3 briefly mentions some generalizations of our sequence, including a broad class of extensions suggested by J. Taylor [9].

Although the sequence A grows very slowly, there are certainly familiar sequences with an even slower growth rate, such as the inverse Ackermann function [1], the Davenport-Schinzel sequences [7], or the inverse to Harvey Friedman's sequence [4]. Nevertheless, we think the combination of slow growth, an unusual definition, and a remarkable recursive structure makes the sequence noteworthy.

The sequence was invented by one of us (D.C.G.) while composing problems for the Dutch magazine *Pythagoras*. It now appears as sequence A90822 in [8].

Notation

If Ω is a set, Ω^n denotes the strings of length n from Ω, Ω^+ is the set of all nonempty finite strings from Ω, and Ω^* is the set of all finite or infinite strings from Ω, including the empty string ϵ. Strings will usually be denoted by uppercase letters. The elements of a string may or may not be separated by commas, and a string may or may not be enclosed in parentheses. A sequence is an infinite string. The length of $U \in \Omega^*$ (which may be ∞) will be denoted by $l(U)$.

Products in Ω^* represent concatenation: if $U \in \Omega^+$, $V \in \Omega^*$ then UV means U followed by V. We will usually not concatenate two infinite strings. A string $U = u(1), \ldots, u(i)$ is said to be a *substring* of $V = v(1), \ldots, v(j)$ if there is an r, $0 \le r \le j - i$, such that $u(k) = v(k + r)$ for $k = 1, \ldots, i$; that is, if the elements of U occur consecutively in V. We say V *contains* U to indicate that U is a substring of V. Terms such as *prefix*, *suffix*, etc., have their usual meanings — see [2] for formal definitions. A sequence U is said to be a *subsequence* of a sequence V if U can be obtained by deleting terms from V.

Usually Ω will be either the nonnegative integers $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$, the positive integers $\mathbb{P} = \{1, 2, 3, \ldots\}$, or the set $\mathbb{P}_m = \{m, m + 1, m + 2, \ldots\}$ for some integer $m \ge 1$.

Given a sequence $U = u(1), u(2), \ldots \in \Omega^*$, its *curling number transform* is the sequence $U_* = u_*(1), u_*(2), \ldots \in \mathbb{P}^*$ given by $u_*(1) = 1$ and

$$u_*(i) = \mathcal{C}(u(1), \ldots, u(i - 1)) \quad \text{for} \quad i \ge 2\,. \tag{4}$$

It is immediate from the definition (2) that our sequence A is equal to its curling number transform, and in fact is the unique sequence with this property.

2 The recursive structure

We introduce the notation in three stages: informally, more formally and—in Section 3—with a somewhat different emphasis that will be needed to prove the main theorems.

Informally, the sequence A is built up recursively from "blocks" $B_n^{(1)}$ that are doubled at each step and are joined together by "glue" strings $S_n^{(1)}$. When the glue strings alone are concatenated together they form a sequence $A^{(2)}$ which has a similar structure to A: it is built up recursively from blocks $B_n^{(2)}$ that are repeated three times at each step and are joined together by "second-order glue" strings $S_n^{(2)}$. When the second-order glue strings are concatenated together they form a sequence $A^{(3)}$ which in turn has a similar structure, but now the blocks $B_n^{(3)}$ are repeated four-fold at each step; and so on. The proof that this description is correct will be given in the next section.

We now make this description more precise. The following description is correct, and is the best way to think about the sequence. However, we will not know for certain that it is correct until the end of Section 3.

The sequence A is constructed from strings $B_n^{(1)}$ and $S_n^{(1)}$, $n \geq 1$, which we call "blocks" and "glue," respectively. The initial block is $B_1^{(1)} = 1$; the second block is $B_2^{(1)} = B_1^{(1)} B_1^{(1)} S_1^{(1)} = 1\ 1\ 2$, where $S_1^{(1)} = 2$; the third block is

$$\begin{aligned} B_3^{(1)} &= B_2^{(1)} B_2^{(1)} S_2^{(1)} \\ &= 1\ 1\ 2\ 1\ 1\ 2\ 2\ 2\ 3\,, \end{aligned}$$

where $S_2^{(1)} = 2\ 2\ 3$, and so on, the n-th block for $n \geq 2$ being

$$B_n^{(1)} = B_{n-1}^{(1)} B_{n-1}^{(1)} S_{n-1}^{(1)}\,, \tag{5}$$

where $S_{n-1}^{(1)}$ contains no 1's. Then for all $n \geq 1$, A begins with $B_n^{(1)}$ (and hence $A = \lim_{n\to\infty} B_n^{(1)}$).

That is, for all $n \geq 2$, A begins with two copies of $B_{n-1}^{(1)}$ followed by a "glue" string $S_{n-1}^{(1)}$ that contains no 1's. $S_{n-1}^{(1)}$ is terminated by the first 1 that follows the initial $B_{n-1}^{(1)} B_{n-1}^{(1)}$. Table 1 shows $B_1^{(1)}$ through $B_6^{(1)}$ (the first row is $B_2^{(1)}$, the first two rows together form $B_3^{(1)}, \ldots$, and the whole table forms $B_6^{(1)}$), and Tables 1 and 2 together form $B_7^{(1)}$. The glue strings $S_1^{(1)}, S_2^{(1)}, \ldots, S_6^{(1)}$ are underlined. By iterating (5) we see that $B_n^{(1)}$ can also be written as

$$B_n^{(1)} = B_{n-1}^{(1)} B_{n-2}^{(1)} \cdots B_1^{(1)} B_1^{(1)} S_1^{(1)} S_2^{(1)} \cdots S_{n-1}^{(1)}\,. \tag{6}$$

The terminating string $S_1^{(1)} S_2^{(1)} \cdots S_{n-1}^{(1)}$ (denoted by $T_n^{(1)}$ in Section 3) is shown in bold-face in Tables 1 and 2 for $n = 2, \ldots, 7$.

In Section 4 we state some conjectures about the lengths of the blocks $B_n^{(1)}$ and of the glue strings $S_{n-1}^{(1)}$. Assuming these conjectures are correct, $l(S_{n-1}^{(1)})$ is much less than $l(B_{n-1}^{(1)})$, and consequently $l(B_n^{(1)})$ is roughly twice $l(B_{n-1}^{(1)})$.

2	2	2	$\underline{\mathbf{3}}$					
2	2	2	3					
2	2	2	**3**	$\underline{\mathbf{3}}$				
2	2	2	3					
2	2	2	3					
2	2	2	3	3				
2	2	2	3					
2	2	2	3					
2	2	2	**3**	**3**	$\underline{\mathbf{3}}$	$\underline{\mathbf{3}}$	$\underline{\mathbf{4}}$	
2	2	2	3					
2	2	2	3					
2	2	2	3	3				
2	2	2	3					
2	2	2	3					
2	2	2	3	3				
2	2	2	3					
2	2	2	3					
2	2	2	3	3	3	3	4	
2	2	2	3					
2	2	2	3					
2	2	2	3	3				
2	2	2	3					
2	2	2	3					
2	2	2	3	3				
2	2	2	3					
2	2	2	3					
2	2	2	**3**	**3**	**3**	**3**	**4**	$\underline{\mathbf{3}}$

Table 3: The first 127 terms of the second-order sequence $A^{(2)}$ (the successive second-order glue strings $S_1^{(2)}, S_2^{(2)}, S_3^{(2)}, S_4^{(2)}$ are underlined; the strings $T_2^{(2)}, T_3^{(2)}, T_4^{(2)}, T_5^{(2)}$ are shown in bold-face).

The above decomposition reduces the study of A to the study of the glue strings $S_n^{(1)}$. We define the "second-order sequence" $A^{(2)} = a^{(2)}(1), a^{(2)}(2), a^{(2)}(3), \ldots$ to be the concatenation $S_1^{(1)} S_2^{(1)} S_3^{(1)} \cdots \in \mathbb{P}_2^*$ of the glue strings. It will be shown later that $A^{(2)}$ can also be defined by

$$\begin{aligned} &a^{(2)}(1) = 2\,, \\ &a^{(2)}(n+1) = \mathcal{C}^{(2)}(a^{(2)}(1), a^{(2)}(2), \ldots, a^{(2)}(n)) \quad \text{for} \quad n \geq 1\,, \end{aligned} \tag{7}$$

where we define

$$\mathcal{C}^{(m)}(U) = \max\{m, \mathcal{C}(U)\} \tag{8}$$

for $m \geq 1$. That is, if $\mathcal{C}(U) = k$ is less than m it is "promoted" to m (we will say more about "promotion" at the end of Section 3). Of course $\mathcal{C}^{(1)} = \mathcal{C}$.

The first 127 terms of $A^{(2)}$ are shown in Table 3, and the reader can verify that they may indeed be obtained by starting with 2 and repeatedly applying the map $\mathcal{C}^{(2)}$.

It is remarkable that $A^{(2)}$ has a similar structure to A, only now the blocks are repeated three times. That is, if we define $B_1^{(2)} = 2$, then for all $n \geq 2$, $A^{(2)}$ begins with a block

$$B_n^{(2)} = B_{n-1}^{(2)} B_{n-1}^{(2)} B_{n-1}^{(2)} S_{n-1}^{(2)}, \tag{9}$$

consisting of three copies of $B_{n-1}^{(2)}$ followed by a "second-order glue" string $S_{n-1}^{(2)} \in \mathbb{P}_3^*$ that contains no 1's or 2's and is terminated by the first number less than 3 that follows the initial $B_{n-1}^{(2)} B_{n-1}^{(2)} B_{n-1}^{(2)}$. Table 3 shows $B_5^{(2)}$ (as well as $B_1^{(2)}$ through $B_4^{(2)}$). The glue strings $S_1^{(2)}, S_2^{(2)}, S_3^{(2)}, S_4^{(2)}$ are underlined. $B_n^{(2)}$ ends with the string $S_1^{(2)} S_2^{(2)} \cdots S_{n-1}^{(2)}$ (denoted by $T_n^{(2)}$ in Section 3); these strings are shown in bold-face in Table 3 for $n = 2, \ldots, 5$.

Again we have a conjectured estimate (see Section 4) for the lengths of the glue, which implies that $l(B_n^{(2)})$ is roughly three times $l(B_{n-1}^{(2)})$.

This analysis reduces the study of $A^{(2)}$ to the study of the second-order glue strings $S_n^{(2)}$, and these, when concatenated, form the third-order sequence $A^{(3)}$, which in turn has a similar structure. And so on!

3 The main theorems

In this section we establish our main theorems, which will show that the description of the sequence given in Section 2 is correct. To do this we must introduce our notation very carefully. The following definitions (temporarily) supersede those in Section 2. For $m \geq 1$, the mth-order sequence $A^{(m)} = a^{(m)}(1), a^{(m)}(2), a^{(m)}(3), \ldots \in \mathbb{P}_m^*$ is defined by

$$\begin{aligned} a^{(m)}(1) &= m\,, \\ a^{(m)}(i+1) &= \mathcal{C}^{(m)}(a^{(m)}(1), \ldots, a^{(m)}(i)) \quad \text{for} \quad i \geq 1\,, \end{aligned} \tag{10}$$

where $\mathcal{C}^{(m)}$ is defined in (8). Note that $A^{(1)}$ is our sequence A. Theorem 3.1 will show that $A^{(m+1)}$ is the concatenation of the glue strings for $A^{(m)}$.

For $m \geq 1$, $n \geq 1$, the blocks $B_n^{(m)} \in \mathbb{P}_m^*$ and the glue strings $S_n^{(m)} \in \mathbb{P}_{m+1}^*$ are defined recursively, and independently of the $A^{(m)}$. Corollary 3.4 will show that all the strings $B_n^{(m)}$, $S_n^{(m)}$ and $T_n^{(m)}$ (defined below) are in fact finite, but at this point we do not know that, and the definitions must allow for the possibility that some of these strings may be infinite.

The recursion for the blocks is

$$B_1^{(m)} = m\,, \tag{11}$$

and, for $n \geq 1$,

$$B_{n+1}^{(m)} = \begin{cases} (B_n^{(m)})^{m+1} S_n^{(m)}, & \text{if } l(B_n^{(m)}) < \infty\,; \\ B_n^{(m)}, & \text{if } l(B_n^{(m)}) = \infty\,; \end{cases} \tag{12}$$

where $S_n^{(m)}$ will be constructed from $B_n^{(m)}$. If $l(B_n^{(m)}) = \infty$, $S_i^{(m)} = \epsilon$ for $i \geq n$. If $l(B_n^{(m)}) < \infty$, consider the sequence $s_n^{(m)}(1), s_n^{(m)}(2), s_n^{(m)}(3), \ldots \in \mathbb{P}_m^*$ defined by

$$\begin{aligned} s_n^{(m)}(1) &= \mathcal{C}^{(m)}((B_n^{(m)})^{m+1})\,, \\ s_n^{(m)}(i+1) &= \mathcal{C}^{(m)}((B_n^{(m)})^{m+1} s_n^{(m)}(1) \cdots s_n^{(m)}(i)) \quad \text{for} \quad i \geq 1\,. \end{aligned} \tag{13}$$

Clearly $s_n^{(m)}(1) \geq m+1$. If there is an integer $i \geq 1$ such that $s_n^{(m)}(i+1) < m+1$, choose the smallest such i, and set

$$S_n^{(m)} = s_n^{(m)}(1), s_n^{(m)}(2), \ldots, s_n^{(m)}(i) \in \mathbb{P}_{m+1}^{+}, \tag{14}$$

but if no such i exists set

$$S_n^{(m)} = s_n^{(m)}(1), s_n^{(m)}(2), \ldots \in \mathbb{P}_{m+1}^{*}. \tag{15}$$

In the latter case $S_n^{(m)}$ and $B_{n+1}^{(m)}$ are infinite.

The $T_n^{(m)}$ are defined as follows. For $n \geq 1$, if $S_1^{(m)}, \ldots, S_n^{(m)}$ are finite we set

$$T_{n+1}^{(m)} = S_1^{(m)} \cdots S_n^{(m)} \in \mathbb{P}_{m+1}^{+}, \tag{16}$$

while if $S_1^{(m)}, \ldots, S_{n-1}^{(m)}$ are finite but $S_n^{(m)}$ is infinite, we still use (16) and define

$$T_i^{(m)} = T_{n+1}^{(m)} \tag{17}$$

for $i \geq n+2$. In the latter case all the $T_i^{(m)}$ for $i \geq n+1$ are infinite. Note that $T_1^{(m)}$ is always undefined.

The lengths of these strings (which may be infinite) are denoted by

$$\begin{aligned} \beta^{(m)}(n) &= l(B_n^{(m)}), & (18)\\ \sigma^{(m)}(n) &= l(S_n^{(m)}), & (19)\\ \tau^{(m)}(n) &= l(T_n^{(m)}). & (20) \end{aligned}$$

We also let $B^{(m)} = b^{(m)}(1), b^{(m)}(2), b^{(m)}(3), \ldots = \lim_{n\to\infty} B_n^{(m)}$. This is well defined since each $B_n^{(m)}$ starts with $B_{n-1}^{(m)}$.

We will require three lemmas.

Lemma 3.1. *For $m \geq 1$, if $A^{(m)}$ contains a string $U^{t+1} \in \mathbb{P}_m^{+}$ for some $t \geq m$, then $U \in \mathbb{P}_t^{+}$.*

Proof. If $t = m$ the claim is trivially true, so we may assume $t \geq m+1$. Suppose, on the contrary, that $U \notin \mathbb{P}_t^{+}$. Then we may write $U = GiH$ for $G, H \in \mathbb{P}_m^{*}$ and some i with $m \leq i \leq t-1$. Thus A contains $GiH\,GiH\,\ldots\,GiH$ ($t+1$ copies). But the final i is preceded by t copies of iHG, so the final i must be at least t, by definition of $A^{(m)}$, a contradiction. ∎

Lemma 3.2. *For $m \geq 1$, $n \geq 2$, (a) $T_n^{(m)}$ is a suffix of $B_n^{(m)}$, and (b) this is the only occurrence of $T_n^{(m)}$ as a substring of $B_n^{(m)}$.*

Proof. Fix $m \geq 1$. It follows by iterating (12) that

$$\begin{aligned} B_{n+1}^{(m)} &= (B_n^{(m)})^m (B_{n-1}^{(m)})^m \cdots (B_1^{(m)})^m B_1^{(m)} S_1^{(m)} S_2^{(m)} \cdots S_{n-1}^{(m)} S_n^{(m)} \\ &= (B_n^{(m)})^m (B_{n-1}^{(m)})^m \cdots (B_1^{(m)})^m B_1^{(m)} T_{n+1}^{(m)}, \end{aligned} \tag{21}$$

provided all of $S_1^{(m)}, S_2^{(m)}, \ldots, S_n^{(m)}$ are finite. If $S_1^{(m)}, S_2^{(m)}, \ldots, S_{n-1}^{(m)}$ are finite but $S_n^{(m)}$ is infinite, (21) is still true, but

$$B_i^{(m)} = B_{n+1}^{(m)},\ T_i^{(m)} = T_{n+1}^{(m)} \text{ for } i \geq n+2\,. \tag{22}$$

Assertion (a) follows at once. To show (b) we use induction on n. The base case, $n = 2$, is true because $T_2^{(m)} = m+1$ and $B_2^{(m)} = (m^{m+1}, m+1)$. If $T_{n+1}^{(m)}$ is infinite and has two occurrences in $B_{n+1}^{(m)}$, they are both suffixes of $B_{n+1}^{(m)}$, implying that $T_{n+1}^{(m)}$ is a suffix of itself, and hence is a periodic sequence. But this is impossible: let M_0 be the maximal element of $T_{n+1}^{(m)}$. After sufficiently many terms the curling number given by (13) would produce a term exceeding M_0, a contradiction. On the other hand, suppose that all the $S_n^{(m)}$ are finite. If $T_{n+1}^{(m)}$ also occurs in $B_{n+1}^{(m)}$ other than as a suffix, it must be a substring of a block $B_j^{(m)}$ in (21), for some j with $2 \leq j \leq n$, for otherwise it would contain the m at the beginning of a block. Write $B_j^{(m)} = UT_{n+1}^{(m)}V = UT_j^{(m)}S_j^{(m)} \cdots S_n^{(m)}V$ for some $U \in \mathbb{P}_m^+$ and $V \in \mathbb{P}_m^*$. But $l(S_j^{(m)} \cdots S_n^{(m)}) > 0$, so $T_j^{(m)}$ occurs as a non-suffix in $B_j^{(m)}$, a contradiction to the induction hypothesis. ∎

Remark. It follows from the above proof that, for any r with $1 \leq r \leq m+1$, any finite substring $(B_n^{(m)})^r$ in (21) contains exactly r copies of $T_n^{(m)}$, each one occurring at the end of a $B_n^{(m)}$. The copies are disjoint.

Lemma 3.3. *For $m \geq 1$, $n \geq 2$, suppose that $k = b^{(m)}(i) \geq m+1$ with $1 \leq i \leq \beta^{(m)}(n)$. Then there exists a Y such that $b^{(m)}(1), \ldots, b^{(m)}(i-1) = XY^k$. Moreover, let Y satisfy this condition with $l(Y)$ minimal and suppose $m \in Y$. Then $Y = B_j^{(m)}$ for some j with $1 \leq j \leq n-1$.*

Proof. We fix $m \geq 1$, and will prove the result for all n by induction. The base case $n = 2$ is immediate, since $B_2^{(m)} = (m^{m+1}, m+1)$. Supposing the result holds for some $n \geq 2$, we will show it holds for $n+1$. If $B_n^{(m)}$ is infinite then the result also holds for $n+1$, by (12), so we may assume that $B_n^{(m)}$ is finite. Then $B_{n+1}^{(m)} = (B_n^{(m)})^{(m+1)} S_n^{(m)}$, by (12). We must show that the result holds for all positions $\beta^{(m)}(n) < i \leq \beta^{(m)}(n+1)$.

If i is a position in $(B_n^{(m)})^{m+1}$, we may write $i = r\beta^{(m)}(n) + s \leq (m+1)\beta^{(m)}(n)$, for $1 \leq r \leq m$, $1 \leq s \leq \beta^{(m)}(n)$. Then $b^{(m)}(i) = b^{(m)}(s)$ and by induction we know that in the first $B_n^{(m)}$ we can write $b^{(m)}(1), \ldots, b^{(m)}(s-1) = XY^k$, and if the minimal Y contains an m then it equals $B_j^{(m)}$ for some j with $1 \leq j \leq n-1$. Therefore this Y (and no shorter string) can also be used at position i, and thus the statement holds.

If $i = (m+1)\beta^{(m)}(n) + 1$ then the part preceding i is $(B_n^{(m)})^{m+1}$, and from (13) we have

$$k = b^{(m)}(i) = s_n^{(m)}(1) = \mathcal{C}^{(m)}((B_n^{(m)})^{m+1}) \geq m+1\,.$$

So certainly one Y exists with $b^{(m)}(1), \ldots, b^{(m)}(i-1) = XY^k$. We must show that if the minimal Y satisfying this property contains an m, then this $Y = B_j^{(m)}$ for some $j \leq n$. If Y contains an m, then it contains $T_n^{(m)}$ as a substring, since the last m in $B_{n+1}^{(m)}$ occurs before the $T_n^{(m)}$ in the last copy of $B_n^{(m)}$. Therefore the string Y^k contains at least $k \geq m+1$ copies of $T_n^{(m)}$. It follows from the Remark below Lemma 3.2 that $k = m+1$ and $Y = B_n^{(m)}$.

If $i > (m+1)\beta^{(m)}(n)+1$ we see by the definition of $S_n^{(m)}$ that again a Y exists. If Y contains an m, then it must properly contain the $T_n^{(m)}$ in the final copy of $B_n^{(m)}$. But in the last of the k copies of Y the copy of $T_n^{(m)}$ is followed by an integer larger than m, whereas in the earlier $k-1$ copies it was followed by the first element of $B_n^{(m)}$, which is m. This is a contradiction, and shows that in this case Y cannot contain an m. ∎

Note that, by definition of $S_{n-1}^{(m)}$, the Y^k for the first element of $S_n^{(m)}$ extends further back (towards the beginning of the sequence) than the start of $B_n^{(m)}$, and thus contains an m. Therefore we see that the situation described in the penultimate paragraph of the above proof is indeed the case and we may conclude that

$$s_n^{(m)}(1) = m+1 \text{ for all } m \geq 1, n \geq 1\,. \tag{23}$$

At this point we can already see that the concatenation of the glue strings is equal to the next A sequence:

Theorem 3.1. *Suppose $m \geq 1$. For all $n \geq 2$, $T_n^{(m)}$ is a prefix of $A^{(m+1)}$, or equals $A^{(m+1)}$ if $T_n^{(m)}$ is infinite.*

Proof. Again we fix m and use induction on n. For $n=2$ the result is trivial. Supposing the result holds for some $n \geq 2$, we will show it holds for $n+1$. If $T_n^{(m)}$ is infinite then clearly the result holds for $T_{n+1}^{(m)}$, so assume that $T_n^{(m)}$ is finite.

Write $(B_n^{(m)})^{m+1} = UT_n^{(m)}$ for some $U \in P_m^+$. We know that $S_n^{(m)}$ begins with $m+1 = s_n^{(m)}(1) = \mathcal{C}^{(m)}(UT_n^{(m)}) = \mathcal{C}^{(m+1)}(UT_n^{(m)}) = \mathcal{C}^{(m+1)}(T_n^{(m)})$. The last equality holds because dropping the U can only decrease the value, but it is already equal to its minimal value of $m+1$. By the induction hypothesis, $T_n^{(m)}$ is a prefix of $A^{(m+1)}$, and therefore $T_n^{(m)}s_n^{(m)}(1)$ is a prefix of $A^{(m+1)}$. For $i \geq 1$, as long as $s_n^{(m)}(i) \geq m+1$, we have

$$\begin{aligned} s_n^{(m)}(i+1) &= \mathcal{C}^{(m)}(UT_n^{(m)}s_n^{(m)}(1)\cdots s_n^{(m)}(i)) \\ &= \mathcal{C}^{(m+1)}(UT_n^{(m)}s_n^{(m)}(1)\cdots s_n^{(m)}(i)) \\ &= \mathcal{C}^{(m+1)}(T_n^{(m)}s_n^{(m)}(1)\cdots s_n^{(m)}(i))\,. \end{aligned}$$

The second equality holds because $s_n^{(m)}(i) \geq m+1$. The third equality holds because Y^k for $s_n^{(m)}(i+1)$ goes back no further than the beginning of $T_n^{(m)}$, as we saw in the proof of the previous lemma. Hence $T_{n+1}^{(m)} = T_n^{(m)}S_n^{(m)}$ is a prefix of $A^{(m+1)}$, as required. ∎

Theorem 3.2. *For all $m \geq 1$, the sequences $A^{(m)}$ and $B^{(m)}$ coincide.*

Proof. Fix $m \geq 1$. We will show by induction on n that, for all $n \geq 1$, $B_n^{(m)}$ is a prefix of $A^{(m)}$ or is all of $A^{(m)}$ if $B_n^{(m)}$ is infinite. This will establish the theorem.

The cases $n=1$ and $n=2$ are immediate, since $B_1^{(m)} = m$, $B_2^{(m)} = (m^{m+1}, m+1)$. So assume the truth of the induction hypothesis up to and including some $n \geq 2$.

If $B_n^{(m)}$ is infinite the result follows from (12), so we may assume that $B_n^{(m)}$ and hence $T_n^{(m)}$ are finite. We wish to show that $B_{n+1}^{(m)} = (B_n^{(m)})^{m+1}S_n^{(m)}$ is a prefix of $A^{(m)}$. If this is not true, the first discrepancy between $B_{n+1}^{(m)}$ and $A^{(m)}$ occurs in the substring $(B_n^{(m)})^{m+1}$, by the definition of $S_n^{(m)}$. Let $i \geq \beta^{(m)}(n)+1$ be the first position in $(B_n^{(m)})^{m+1}$ at which $a^{(m)}(i) \neq b^{(m)}(i)$. Our goal is to show that the existence of i leads to a contradiction.

We may write $i = j\beta^{(m)}(n) + r$ with $1 \leq j \leq m$ and $1 \leq r \leq \beta^{(m)}(n)$. Then i is also minimal with respect to the condition that $a^{(m)}(i) \neq a^{(m)}(r)$. Let $a^{(m)}(1), \ldots, a^{(m)}(i-1) = XY^k$ with k maximal and $l(Y)$ minimal. Then $a^{(m)}(i) = \max\{m, k\}$.

We consider two cases, depending on whether or not $a^{(m)}(i)$ is at the beginning of one of the $B_n^{(m)}$ blocks, i.e. whether $r = 1$ or $r \geq 2$.

First, suppose $r = 1$; then we need to prove that $a^{(m)}(i) = a^{(m)}(1) = m$. This follows by definition of $S_{n-1}^{(m)}$ if $j = 1$; so assume $j \geq 2$, and that $k = a^{(m)}(i) \geq m+1$. Using (21) we may write $a^{(m)}(1), \ldots, a^{(m)}(i-1) = (B_n^{(m)})^j = (B_n^{(m)})^{j-1} U m T_n^{(m)}$ for some $U \in \mathbb{P}_m^*$. If $T_n^{(m)}$ is a proper suffix of Y^k then $m \in Y$, which implies that $T_n^{(m)}$ is a proper suffix of Y and therefore $(B_n^{(m)})^j$ contains at least $m+1$ copies of $T_n^{(m)}$, contradicting the Remark following Lemma 3.2. On the other hand, if Y^k were a suffix of $T_n^{(m)}$, this would contradict the fact that $S_{n-1}^{(m)}$ is followed by an element $\leq m$.

Second, suppose that $r \geq 2$. Let $L = a^{(m)}(1), \ldots, a^{(m)}(r-1)$ and write $L = X_* Y_*^{k_*}$ with k_* maximal and $l(Y_*)$ minimal. Then $a^{(m)}(r)$ $=\max\{m, k_*\}$. By the definition of i, $a^{(m)}(i) > a^{(m)}(r) \geq m$. Hence $a^{(m)}(i) = k \geq m+1$. To have $a^{(m)}(i) > a^{(m)}(r)$, L must be a suffix of Y^k, so $m \in Y$ and therefore, by Lemma 3.1, k is at most $m+1$ and therefore is equal to $m+1$. Hence $k_* \leq m$.

The situation, then, is that $(B_n^{(m)})^j L$ is a prefix of $A^{(m)}$. We are supposing that we can achieve $a^{(m)}(i) = m+1$ by allowing L to be a suffix of Y^{m+1}. Noting that $T_n^{(m)}$ is a suffix of $(B_n^{(m)})^j$, by (21), we distinguish two cases, depending on the relationship between $T_n^{(m)} L$ and Y^{m+1}.

(i) Suppose that $T_n^{(m)} L$ is a suffix of Y^{m+1}. We know $m \in Y$ and $m \notin T_n^{(m)}$, so Y^{m+1} contains at least m disjoint copies of $T_n^{(m)}$. Hence $j = m$, and there are exactly m disjoint copies, by the Remark following Lemma 3.2. This means that each copy of $T_n^{(m)}$ straddles the end of one copy of Y and the beginning of the next (if not, $T_n^{(m)}$ is wholly contained in Y, and so there are $m+1$ copies of $T_n^{(m)}$ in the sequence before position i, which is a contradiction since there are only m copies, one in each of the m copies of $B_n^{(m)}$ and none so far in the next copy of $B_n^{(m)}$ that we are building), and hence that Y is a proper suffix of $T_n^{(m)} L$. Write $T_n^{(m)} = VW$ where W is the intersection of $T_n^{(m)}$ and the last (or $(m+1)$-st) copy of Y, and write $B_n^{(m)} = UT_n^{(m)}$, using (21). If $m \geq 2$ it is easy to complete the proof. We have $Y = WL = WUV$, so $L = UV$ and therefore $i > l(Y^{m+1}) = (m+1)l(WUV) = (m+1)l(UVW) = (m+1)\beta^{(m)}(n)$, contradicting the definition of i.

Suppose then that $m = 1$. Again L is a proper suffix of Y and Y is a proper suffix of $T_n^{(1)} L$. Write $Y = WL$, and let $s \geq 2$ be the first element of Y. Let this element s in the second copy of Y be preceded by s copies of some string Y' with $l(Y')$ minimal.

Suppose that Y' does not contain a 1. Since Y does contain a 1 (L starts with a 1), Y'^s is a suffix of the first copy of Y, and hence also of the second copy of Y. This contradicts the minimality of $l(Y)$, since then $l(Y') < l(Y)$.

So we may assume that $1 \in Y'$, hence by Lemma 3.3 we know that

$$Y' = B_\kappa^{(1)}$$

for some $\kappa < n$. Now $T_\kappa^{(1)}$ is a suffix of Y^2 and since L starts with a 1, $T_\kappa^{(1)}$ is also a suffix of L. By Lemma 3.2, $B_\kappa^{(1)}$ is also a suffix of L. Suppose $B_\kappa^{(1)} = L$. Then WL is a suffix of $Y'Y' = LL$ (look at the first copy of $Y = WL$ and remember $Y'Y'$ begins with a 1) and hence W is a suffix of L. But then W^2 is a suffix of $T_n^{(1)}$, contradicting the fact that L starts with a 1.

So we may assume that $B_\kappa^{(1)}$ is a strict suffix of L. But now $l(Y) > l(L) \geq 2l(B_\kappa^{(1)})$. (Indeed, if $l(L) < 2l(B_\kappa^{(1)})$, then we know that L is a prefix of $B_n^{(1)}$, by definition, $B_\kappa^{(1)}B_\kappa^{(1)}$ is also a prefix of $B_n^{(1)}$, and so L is a strict prefix of $B_\kappa^{(1)}B_\kappa^{(1)}$; but L has $T_\kappa^{(1)}$ as a suffix, so by Lemma 3.2, $L = B_\kappa^{(1)}$, a contradiction.) But now $(B_\kappa^{(1)})^2$ is a suffix of Y, contradicting the minimality of Y.

(ii) Suppose on the other hand that Y^{m+1} is a suffix of $T_n^{(m)}L$. Since no Y is contained in $T_n^{(m)}$ (remember that $m \in Y$), Y^m is a suffix of L and the first element, t, of Y is in $T_n^{(m)}$ with $t \geq m+1$. Therefore the first element of the second Y is also t and since $(B_n^{(m)})^j L$ is a prefix of $A^{(m)}$, Y ends with U^t for some U. Hence U^t is a suffix of L, which contradicts the fact that $k_* = m$. This completes the proof. ∎

Corollary 3.3. *The sequence $A^{(m)}$ contains every integer $\geq m$.*

Proof. From Theorem 3.1 we know that, for $m \geq 2$, $n \geq 2$, $T_n^{(m-1)}$ is a prefix of $A^{(m)}$, so, for a given m, either

$$A^{(m)} = S_1^{(m-1)} S_2^{(m-1)} \cdots S_n^{(m-1)}$$

if some $S_n^{(m-1)}$ is infinite, or

$$A^{(m)} = S_1^{(m-1)} S_2^{(m-1)} S_3^{(m-1)} \cdots$$

if all $S_n^{(m-1)}$ are finite. Also, by Theorem 3.2, $B_{n+1}^{(m-1)}$ is a prefix of $A^{(m-1)}$, so from (12), if some $S_n^{(m-1)}$ is infinite, $A^{(m-1)}$ contains

$$S_1^{(m-1)}, S_2^{(m-1)}, \ldots, S_n^{(m-1)},$$

or if all $S_n^{(m-1)}$ are finite, $A^{(m-1)}$ contains $S_1^{(m-1)}, S_2^{(m-1)}, \ldots, S_n^{(m-1)}$ for all n. In either case (and this is the key point), every prefix of $A^{(m)}$ is a subsequence of $A^{(m-1)}$. Repeating this argument shows that every prefix of every $A^{(j)}$ is a subsequence of $A^{(m)}$ if $j \geq m$.

Since $A^{(j)}$ begins with j, $A^{(m)}$ contains every integer $j \geq m$. ∎

Corollary 3.4. *The strings $B_n^{(m)}$, $S_n^{(m)}$ and $T_n^{(m)}$ have finite length.*

Proof. The first occurrence of an integer in $A^{(m)}$ is necessarily followed by an m. Since we saw in the previous corollary that $A^{(m)}$ contains infinitely many different integers, it follows that all $S_n^{(m)}$ are finite. This implies that $B_n^{(m)}$ and $T_n^{(m)}$ are also finite. ∎

Promotion

In the definition of $A^{(m)}$, (10), let us say that $a^{(m)}(i)$ is *promoted* if either $i = 1$ or $\mathcal{C}(a^{(m)}(1), \ldots, a^{(m)}(i-1)) < m$. If we know which elements in $A^{(m+1)}$ are promoted, we can recover $A^{(m)}$ from $A^{(m+1)}$. To make this precise, we define the strings $D_i^{(m)} \in \mathbb{P}_m^+$ by $D_0^{(m)} = m$ and, for $i > 0$,

$$D_i^{(m)} = \begin{cases} D_{i-1}^{(m)} a^{(m+1)}(i)\,, & \text{if } a^{(m+1)}(i) \text{ is not promoted;} \\ \left(D_{i-1}^{(m)}\right)^{m+1} a^{(m+1)}(i)\,, & \text{if } a^{(m+1)}(i) \text{ is promoted.} \end{cases} \tag{24}$$

Since $D_i^{(m)}$ starts with $D_{i-1}^{(m)}$, we can define the limiting sequence $D^{(m)} = \lim_{i\to\infty} D_i^{(m)}$. Then it can be shown that:

Theorem 3.5. *For all $m \geq 1$, the sequences $A^{(m)}$ and $D^{(m)}$ coincide.*

Sketch of proof. By Theorem 3.1, $A^{(m+1)} = S_1^{(m)} S_2^{(m)} \cdots$. The promoted elements $a^{(m+1)}(i_1)$, $a^{(m+1)}(i_2), \ldots$ of $A^{(m+1)}$ are precisely the initial elements of the $S_k^{(m)}$. Indeed, by definition of $S_k^{(m)}$, the first element of $S_k^{(m)}$ is promoted.

To see the reverse, let $a^{(m)}(i) = t \geq m+1$, with $(m+1)\beta^{(m)}(k)+1 < i \leq \beta^{(m)}(k+1)$. That is, we are in the first copy of $S_k^{(m)}$. It suffices to show that when $a^{(m)}(1)\cdots a^{(m)}(i) = XY^t$, Y^t lies completely inside $T_{k+1}^{(m)}$; that is, Y does not contain an m. This follows since otherwise XY^t contains $t \geq m+1$ copies of T_k, contradicting Lemma 3.2.

We now show by induction that $D_{i_k-1}^{(m)} = B_k^{(m)}$ $(k = 1, 2, \ldots)$. For $k = 1$ this is obvious: $D_0^{(m)} = B_1 = m$. Let $k \geq 2$. Then

$$D_{i_k}^{(m)} = (D_{i_k-1}^{(m)})^{m+1} a^{(m+1)}(i_k) = (B_k^{(m)})^{m+1} a^{(m+1)}(i_k)\,. \tag{25}$$

Hence

$$D_{i_{k+1}-1}^{(m)} = (B_k^{(m)})^{m+1} a^{(m+1)}(i_k) \cdots a^{(m+1)}(i_{k+1}-1) \tag{26}$$

$$= (B_k^{(m)})^{m+1} S_k^{(m)} = B_{k+1}^{(m)}\,. \tag{27}$$

■

4 Estimates for the rate of growth

In this section we take an experimental approach, and record a series of observations about the sequence. These observations appear to be correct, but we have been unable to prove them, so we state them as conjectures. In §4.1 we study the lengths of the glue strings $S_n^{(m)}$. Although these lengths are somewhat irregular, it appears that they can be "smoothed" so as to become much more regular "ruler" sequences, whose peak values will be denoted by $\rho^{(m)}(n)$. In §4.2 we describe a "tabular" construction for the higher-order sequences $A^{(2)}, A^{(3)}, \ldots$ which leads to a recurrence relating the $\rho^{(m)}(n)$, $\beta^{(m)}(n)$ and $\sigma^{(m)}(n)$. Sections 4.3, 4.4 and 4.5 contain estimates for $\beta^{(m)}(n)$, $\rho^{(m)}(n)$ and $\tau^{(m)}(n)$. Finally, in §4.6, we use these estimates to determine where each number $t \geq 1$ appears for the first time in our sequence A.

4.1 Ruler sequences and smoothing

It appears that the sequence $\sigma^{(m)} = \sigma^{(m)}(1), \sigma^{(m)}(2), \sigma^{(m)}(3), \ldots$ giving the lengths of the glue strings $S_n^{(m)}$ is essentially a "ruler" sequence, in the sense that $\sigma^{(m)}(n)$ essentially depends only on the $(m+1)$-adic valuation of n.

For positive integers m, n, define the *m-adic valuation* of n, $|n|_m$, to be the highest power of m dividing n. The classical example of a ruler sequence is the sequence $r = r(1), r(2), r(3), \ldots$ given by

$$r(n) = |n|_2 + 1 \,. \tag{28}$$

The first 32 terms are

1	**2**	1	**3**	1	2	1	**4**	1	2	1	3	1	2	1	**5**
1	2	1	3	1	2	1	4	1	2	1	3	1	2	1	**6**

where the new record entries, shown in bold-face, occur at powers of 2. For much more about this sequence, including an extensive bibliography, see entry A1511 in [8].

The initial values of $\sigma^{(1)}, \ldots, \sigma^{(4)}$ are shown in Table 4, and the record entries in $\sigma^{(1)}, \ldots, \sigma^{(10)}$ in Table 5. Let $\pi^{(m)}(j)$ $(j \geq 0)$ denote the j-th record in $\sigma^{(m)}$.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\sigma^{(1)}(n)$	**1**	**3**	1	**9**	4	**24**	1	3	1	9	4	**67**	1	3	1	9
$\sigma^{(2)}(n)$	**1**	1	**3**	1	1	3	1	1	**9**	1	1	3	1	1	3	1
$\sigma^{(3)}(n)$	**1**	1	1	**3**	1	1	1	3	1	1	1	3	1	1	1	**10**
$\sigma^{(4)}(n)$	**1**	1	1	1	**3**	1	1	1	1	3	1	1	1	1	3	1

n	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
$\sigma^{(1)}(n)$	4	24	1	3	1	9	4	**196**	3	1	9	4	24	1	3	1
$\sigma^{(2)}(n)$	1	9	1	1	3	1	1	3	1	1	**32**	1	3	1	1	3
$\sigma^{(3)}(n)$	1	1	1	3	1	1	1	3	1	1	1	3	1	1	1	10
$\sigma^{(4)}(n)$	1	1	1	**11**	1	1	1	1	3	1	1	1	1	3	1	1

n	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
$\sigma^{(1)}(n)$	9	4	68	3	1	9	4	24	1	3	1	9	4	**581**	3	1

Table 4: Values of $\sigma^{(1)}(n)$ for $n \leq 48$ and $\sigma^{(2)}(n)$, $\sigma^{(3)}(n)$, $\sigma^{(4)}(n)$, for $n \leq 32$, with record entries shown in bold-face.

$m \setminus j$	0	1	2	3	4	5	6	7	8	9
1	1	3	*9*	24	67	*196*	*581*	*1731*	*5180*	*15534*
2	1	3	9	*32*	119	463	1837	*7332*	29307	117203
3	1	3	10	42	*200*	983	4892	24434	122141	
4	1	3	11	55	315	*1872*	11205	67195		
5	1	3	12	70	471	3273	*22883*			
6	1	3	13	87	673	5355	42805			
7	1	3	14	106	927	8309	74740			
8	1	3	15	127	1239	12351	123463			
9	1	3	16	150	1615	17721				
10	1	3	17	175	2061	24683				

Table 5: Values of $\pi^{(m)}(j)$, the j-th record in sequence $\sigma^{(m)}$. The smoothed record values $\rho^{(m)}(j)$ are obtained by reducing the italicized entries by 1. The next three terms in the first row are *46578*, *139713*, *419116*, and the next term in the $m = 2$ row is 468785. The missing entries in this table have not been calculated, although we predict that the entries on or below the diagonal $m = j$ are given by (43) and the entries just above this diagonal by (44).

As can be seen from Table 4, $\sigma^{(1)}$ is not quite as regular as the ruler sequence r. However:

Conjecture 4.1. *If the sequence $\sigma^{(1)}$ is "smoothed" by replacing every instance of 4 by the pair of numbers 3, 1, every 9 by 8, 1, every 25 by 24, 1, and so on, $\sigma^{(1)}$ becomes a ruler sequence $r^{(1)}$ given by*

$$r^{(1)}(n) = \rho^{(1)}(|n|_2)\,, \tag{29}$$

in which the first 64 terms are

1	**3**	1	**8**	1	3	1	**24**	1	3	1	8	1	3	1	**67**
1	3	1	8	1	3	1	24	1	3	1	8	1	3	1	**195**
1	3	1	8	1	3	1	24	1	3	1	8	1	3	1	67
1	3	1	8	1	3	1	24	1	3	1	8	1	3	1	**580**

and where the record values (shown in bold-face) $\rho^{(1)}(0), \rho^{(1)}(1), \ldots$ are

$$1, 3, 8, 24, 67, 195, 580, 1730, 5179, 15533, 46578, 139712, 419115, \ldots\,. \tag{30}$$

The numbers i in $\sigma^{(1)}$ that are to be replaced by $i-1$, 1 to get $r^{(1)}$ are

$$4, 9, 25, 68, 196, 581, 1731, 5180, 15534, 46579, 139713, 419116, \ldots\,, \tag{31}$$

The numbers that need to be smoothed, given in (31), are one greater than the numbers in (30), except that 2 is missing. The records in the smoothed sequence $r^{(1)}$, (30), either agree with or are one less than the terms in the first row of Table 5.

The sequences $\sigma^{(m)}$ for $m \ge 2$ appear to need less smoothing than $\sigma^{(1)}$ to make them into ruler sequences. In the range of our tables, $\sigma^{(2)}$ needs to be smoothed by replacing every 32 by 31, 1, and every 7332 by 7331, 1; $\sigma^{(3)}$ by replacing every 200 by 199, 1; $\sigma^{(4)}$ by replacing every 1872 by 1871, 1; and so on. If $r^{(m)}$ denotes the smoothed version of $\sigma^{(m)}$ and $\rho^{(m)}(j)$ the j-th record in the smoothed version (see Table 5) then we have, for all $m \ge 1$, $n \ge 1$,

$$r^{(m)}(n) = \rho^{(m)}(|n|_{m+1})\,. \tag{32}$$

The lengths $\beta^{(m)}(n)$ of the blocks are given by (from (12), (18))

$$\begin{aligned} \beta^{(m)}(1) &= 1\,, \\ \beta^{(m)}(n+1) &= (m+1)\beta^{(m)}(n) + \sigma^{(m)}(n) \text{ for } n \ge 1\,. \end{aligned} \tag{33}$$

The initial values of $\beta^{(1)}(n), \ldots, \beta^{(6)}(n)$ are shown in Table 6.

$m \setminus n$	1	2	3	4	5	6	7	8
1	1	3	9	19	47	98	220	441
2	1	4	13	42	127	382	1149	3448
3	1	5	21	85	343	1373	5493	21973
4	1	6	31	156	781	3908	19541	97706
5	1	7	43	259	1555	9331	55989	335935
6	1	8	57	400	2801	19608	137257	960802

Table 6: Lengths $\beta^{(m)}(n)$ of the blocks $B_n^{(m)}$.

4.2 The tabular construction

The appearance of ruler sequences can be partially explained if we present the construction of the higher-order sequences $A^{(2)}, A^{(3)}, \ldots$, in a tabular format. In this construction we keep track not only of the actual value $A^{(m)}(n) = \max\{m, k\}$ (cf. (10)) but also whether the promotion rule was invoked (if $k < m$ we indicate this by drawing a circle around the entry) and the length of the shortest Y that was used to compute k if $k \geq m$ (shown as a subscript; if the promotion rule was invoked the subscript is 0). This tabular construction will also suggest a recurrence that relates $\rho^{(m)}(n+1)$, $\beta^{(m+1)}(n+1)$ and $\sigma^{(m+1)}(n+1)$.

We will construct $A^{(2)}$ as an example. We start by making a small table of the glue strings $S_n^{(2)}$ for $n \leq 10$ — see Table 7. (We already saw $S_1^{(2)}, \ldots, S_4^{(2)}$ in Table 3.)

n	$S_n^{(2)}$
1	3
2	3
3	3 3 4
4	3
5	3
6	3 3 4
7	3
8	3
9	3 3 4 3 3 3 3 4 4
10	3

Table 7: The first few glue strings $S_n^{(2)}$.

We know from Section 3 that $A^{(2)} = \lim_{n\to\infty} B_n^{(2)} = \lim_{n\to\infty} T_n^{(1)} = S_1^{(1)} S_2^{(1)} \cdots$ and that $B_{n+1}^{(2)} = (B_n^{(2)})^3 S_n^{(2)}$. Table 8 shows the beginning of the construction of $A^{(2)}$.

The aim is to understand how $A^{(2)}$ breaks into the consecutive $S_n^{(1)}$ glue strings for $A^{(1)}$. To do this, a version of $A^{(2)}$ is produced in which terms that are obtained by promotion are circled, and where the subscript on each term is either 0 for a circled term or else gives the length of the shortest Y that can be used to compute that term. The circled terms will be the first terms of each of the glue strings $S_n^{(1)}$ of $A^{(1)}$. Most of the circling and subscripting work is done by a few simple rules. However, the rules occasionally give the wrong answer and a few corrections may need to be made by hand at the end of each round. It is the presence of these adjustments that makes our sequence hard to analyze.

We start with $B_1^{(2)} = 2_0$. The rules for going from $B_n^{(2)}$ to $B_{n+1}^{(2)}$ are as follows:

(i) Write $B_n^{(2)}$ as a single string, and construct a three-rowed array in which each row is a copy of $B_n^{(2)}$, omitting all circles from the third row. This three-rowed array (after $S_n^{(2)}$ is appended in step (iii)) will form $B_{n+1}^{(2)}$ when read as a single string. (When constructing $A^{(m)}$ we make m copies of $B_n^{(m)}$ and omit the circles from the m-th copy.)

(ii) The subscripts in rows 2 and 3 are the same as in row 1, except that terms in row 3 that are under circled terms in row 2 have their subscripts changed to $l(B_n^{(2)})$.

(iii) Append $S_n^{(2)}$ to the end of row 3. The first term of $S_n^{(2)}$ receives the subscript $l(B_n^{(2)})$. The subscripts on the remaining terms of $S_n^{(2)}$ must be computed separately—they can be obtained from the tabular construction of $A^{(m+1)}$.
(iv) Finally, a few circles in row 2 may need to be omitted and their subscripts recomputed, as well as the subscripts on the same terms in row 3.

In Table 8, rules (i)–(iii) give the correct answers for $B_2^{(2)}$ and $B_3^{(2)}$. But in $B_4^{(2)}$ four terms (marked with asterisks in Table 8) must be corrected. The first entry in row 2 of $B_4^{(2)}$ is $\textcircled{2}_0$. However, row 1 ends with $3\ 3 = S_1^{(2)} S_2^{(2)} = Y^2$, with a Y of length 1, so that 2 did not need to be promoted and we must change $\textcircled{2}_0$ to 2_1. The fifth entry in row 2 of $B_4^{(2)}$ is $\textcircled{2}_0$. But it is preceded by

$$32223 \;\; 32223 \;=\; S_1^{(2)} B_2^{(2)} S_2^{(2)} B_2^{(2)} \;=\; Y^2,$$

with a Y of length 5, so we must change $\textcircled{2}_0$ to 2_5. The corresponding entries in row 3, presently both equal to 2_{13}, also get changed to 2_1 and 2_5 respectively.

$$
\begin{array}{ll}
B_1^{(2)} = & \textcircled{2}_0 \\[1ex]
B_2^{(2)} = & \begin{array}{ll} \textcircled{2}_0 & \\ \textcircled{2}_0 & \\ 2_1 & 3_1 \end{array} \\[1ex]
B_3^{(2)} = & \begin{array}{lllll} \textcircled{2}_0 & \textcircled{2}_0 & 2_1 & 3_1 & \\ \textcircled{2}_0 & \textcircled{2}_0 & 2_1 & 3_1 & \\ 2_4 & 2_4 & 2_1 & 3_1 & 3_4 \end{array} \\[1ex]
B_4^{(2)} = & \begin{array}{llllllllllllllll} \textcircled{2}_0 & \textcircled{2}_0 & 2_1 & 3_1 & \textcircled{2}_0 & \textcircled{2}_0 & 2_1 & 3_1 & 2_4 & 2_4 & 2_1 & 3_1 & 3_4 & & & \\ \textcircled{2}_0^* & \textcircled{2}_0 & 2_1 & 3_1 & \textcircled{2}_0^* & \textcircled{2}_0 & 2_1 & 3_1 & 2_4 & 2_4 & 2_1 & 3_1 & 3_4 & & & \\ 2_{13}^* & 2_{13} & 2_1 & 3_1 & 2_{13}^* & 2_{13} & 2_1 & 3_1 & 2_4 & 2_4 & 2_1 & 3_1 & 3_4 & 3_{13} & 3_1 & 4_1 \end{array}
\end{array}
$$

Table 8: Tabular construction of $A^{(2)}$.

When we extend Table 8 to $B_{10}^{(2)}$, we find that in all only ten circles need to be removed. After $B_4^{(2)}$, the next changes are at $B_7^{(2)}$, where two circles get removed because of the splittings $S_4^{(2)} S_5^{(2)} = 3\ 3 = Y^2$, with a Y of length 1, and $S_4^{(2)} B_5^{(2)} S_5^{(2)} B_5^{(2)} = Y^2$ with a Y of length 128. But not all instances of such splittings cause circles in the table to be removed, and not all circle-removals arise in this way. It seems difficult to explain exactly where corrections to the table are required.

However, the corrections are rare, and still fewer corrections are needed for larger values of m.

Since $A^{(2)}$ is also $\lim_{n\to\infty} T_n^{(1)}$, we can read off the lengths of the glue strings $S_n^{(1)}$ from the table. Look at the lengths of the strings (in $B_4^{(2)}$) between one circle and the next: these are 1, 3, 1, 9, 4, 24, ... exactly the values of $\sigma^{(1)}(1), \sigma^{(1)}(2), \ldots$ (cf. Table 4). If we do not make the corrections needed in step (iv), we instead get the smoothed lengths 1, 3, 1, 8, 1, 3, 1, 24, ... These observations lead to our conjectured recurrence. For example, note that the string in $B_4^{(2)}$ from the last circled entry to the end has length 24 (which is $\rho^{(1)}(3)$) and is made up of the last string in $B_3^{(2)}$ (length 8, which is $\rho^{(1)}(2)$) plus the whole of $B_3^{(2)}$ (length 13, which is $\beta^{(2)}(3)$), plus $S_3^{(2)}$ (length 3, which is $\sigma^{(2)}(3)$). More generally, we have:

Conjecture 4.2. *For $m \geq 1$,*

$$\begin{aligned} \rho^{(m)}(0) &= 1\,, \\ \rho^{(m)}(n+1) &= \rho^{(m)}(n) + \beta^{(m+1)}(n+1) + \sigma^{(m+1)}(n+1) \text{ for } n \geq 0\,. \end{aligned} \tag{34}$$

This recurrence is supported by all the data, although we do not have a proof.

4.3 Estimates for the lengths $\beta^{(m)}(n)$ of the blocks

In this section we first prove formulas (36), (37), (38), which give the exact value of $\beta^{(m)}(n)$ in the parabolic region $1 \leq n \leq (m+1)^2 - 1$ for $m \geq 1$. We then give *conjectural* estimates for $\beta^{(m)}(n)$ (indicated by $\approx$), (39) and (40), which apply for all m and n.

We take (33) as our starting point. For $m \geq 1$, $\sigma^{(m)}(n)$ and $r^{(m)}(n)$ coincide for $1 \leq n \leq (m+1)^2 - 1$, and in that range are given by

$$\sigma^{(m)}(n) = r^{(m)}(n) = \begin{cases} 1\,, & \text{if } m+1 \text{ does not divide } n\,; \\ 3\,, & \text{if } m+1 \text{ divides } n\,. \end{cases} \tag{35}$$

By iterating (33) and using (35) we find:

Lemma 4.1. *For $m \geq 1$ and $1 \leq n \leq (m+1)^2 - 1$,*

$$\beta^{(m)}(n) = \frac{(m+1)^n - 1}{m} + 2\,\frac{(m+1)^{n-1} - (m+1)^v}{(m+1)^{m+1} - 1}\,, \tag{36}$$

where $v \in \{0, 1, \ldots, m\}$ is given by $n - 1 \equiv v \bmod m+1$.

In particular, for $1 \leq n \leq m+1$ we have $v = n-1$ and so

$$\beta^{(m)}(n) = \frac{(m+1)^n - 1}{m}\,, \tag{37}$$

and for $m+2 \leq n \leq 2m+2$ we have $v = n-m-2$ and

$$\beta^{(m)}(n) = \frac{(m+1)^n + 2(m+1)^{n-m-1} - 2(m+1)^{n-m-2} - 1}{m}\,. \tag{38}$$

Equation (37) explains the entries on or below the diagonal $n = m+1$ in Table 6, (38) explains the entries in the "wedge" $m+2 \leq n \leq 2m+2$, and (36) the entries in the whole parabolic region bounded by $n \leq (m+1)^2 - 1$.

The right-hand side of (36) is also a good approximation to $\beta^{(m)}(n)$ for fixed $m \geq 2$ and $n \geq (m+1)^2$. The case $m = 1$ is special, because of the greater differences between $\sigma^{(m)}(n)$ and $r^{(m)}(n)$ when $m = 1$. However, calculations show that $\beta^{(1)}(n)$ is well-approximated by

$$\beta^{(1)}(n) \approx \varepsilon_1\, 2^{n-1}\,, \tag{39}$$

where $\varepsilon_1 = 3.48669886\ldots$.

For our applications, the approximation

$$\beta^{(m)}(n) \approx \varepsilon_m\, (m+1)^{n-1} \text{ for } m \geq 1, n \geq 1 \tag{40}$$

(consistent with (36)–(39)) will be adequate, where ε_m is a constant on the order of 1.

4.4 Estimates for the records $\rho^{(m)}(n)$

We now apply (34) to estimate $\rho^{(m)}(n)$. Except for (41), the formulas in this section are conjectural. Again using (35), we find that

$$\rho^{(m)}(n) = \frac{m(n+1+2u) + \beta^{(m+1)}(n+1)}{m+1} \tag{41}$$

for $0 \le n \le (m+2)^2 - 1$, where $u = \lfloor n/(m+2) \rfloor$. Eliminating $\beta^{(m+1)}(n+1)$ from (34) and (41) we obtain

$$\rho^{(m)}(n+1) = (m+2)\rho^{(m)}(n) + \sigma^{(m+1)}(n+1) - m(n+1+2u)\,. \tag{42}$$

In particular, for $0 \le n \le m+1$, $u = 0$ and so (41) gives

$$\rho^{(m)}(n) = \frac{(m+2)^{n+1} + (n+1)m(m+1) - 1}{(m+1)^2}\,, \tag{43}$$

while for $m+2 \le n \le 2m+3$, $u = 1$ and

$$\rho^{(m)}(n) = \frac{(m+2)^{n+1} + 2(m+1)(m+2)^{n-m-2} + (n+3)m(m+1) - 1}{(m+1)^2}\,. \tag{44}$$

Equation (43) matches the smoothed values on or below the diagonal $n = m+1$ in Table 5, (44) matches the values in the "wedge" $m+2 \le n \le 2m+3$, and (41) matches the values in the whole region bounded by the "parabola" $n \le (m+2)^2 - 1$. Equation (44) is in fact a good estimate of $\rho^{(m)}(n)$ for all m and n.

The greatest differences between the exact values $\pi^{(m)}(n)$ and the smoothed values $\rho^{(m)}(n)$ occur in the first row of Table 5. The ratio of terms $\pi^{(m)}(n+1)/\pi^{(m)}(n)$ in row m of that table rapidly approaches $m+2$, and for fixed m we find that

$$\pi^{(m)}(n) \approx \lambda_m (m+2)^n\,, \tag{45}$$

where approximate values of λ_m are

m	:	1	2	3	4	5	6	7	$\cdots$
λ_m	:	.778	.447	.312	.240	.194	.163	.140	$\cdots$

Curve-fitting suggests that

$$\lambda_m \approx \frac{.956m + 2.11}{(m+1)^2}$$

which we approximate by

$$\lambda_m \approx \frac{m+2}{(m+1)^2}\,, \tag{46}$$

leading to

$$\pi^{(m)}(n) \approx \frac{(m+2)^{n+1}}{(m+1)^2} \tag{47}$$

for m fixed and n large.

Since the leading terms in (41), (43) and (44) agree with (47), we will take (47) as our approximation to both $\pi^{(m)}(n)$ and $\rho^{(m)}(n)$ for all m and n.

4.5 An estimate for $\tau^{(m)}(n)$

From equations (16), (20) we have

$$\tau^{(m)}(n+1) = \sum_{i=1}^{n} \sigma^{(m)}(i), \quad n \ge 1\,. \tag{48}$$

To simplify the analysis (we are only seeking a crude estimate in this section) we suppose we have reached the end of block $B_n^{(m)}$ in $A^{(m)}$, where $n = (m+1)^\mu + 1$ for some $\mu \ge 1$. This block ends with the string $T_n^{(m)}$ of length $\tau^{(m)}(n)$.

Of the $(m+1)^\mu$ strings $S_i^{(m)}$, $1 \le i \le (m+1)^\mu$, that appear in $T_n^{(m)}$, a fraction $\frac{m}{m+1}$ have $|i|_{m+1} = 0$ and contribute $\pi^{(m)}(0)$ to the sum; a fraction $\frac{m}{(m+1)^2}$ have $|i|_{m+1} = 1$ and contribute $\pi^{(m)}(1)$; and so on. Therefore, from (48),

$$\tau^{(m)}((m+1)^\mu + 1) = (m+1)^\mu \sum_{i=0}^{\mu-1} \frac{m}{(m+1)^{i+1}} \pi^{(m)}(i) + \pi^{(m)}(\mu)\,,$$

where the last term accounts for the final glue string $S_{n-1}^{(m)}$. Using (47) this becomes

$$\tau^{(m)}((m+1)^\mu + 1) \approx \frac{m+2}{m+1} \left((m+2)^\mu - m(m+1)^{\mu-1} \right). \tag{49}$$

We summarize the discussion in the last two sections in the following:

Conjecture 4.3. *After smoothing (cf. Conjecture 4.1), the sequence* $\sigma^{(m)} = \sigma^{(m)}(1), \sigma^{(m)}(2), \ldots$ *of the lengths of the glue strings in* $A^{(m)}$ *is a ruler sequence given by (32), where the record values are given by (41) (exact, for* $n \le (m+2)^2 - 1$*) and (47) (approximate, for all* m *and* n*). Equation (49) gives an estimate for* $\tau^{(m)}(n)$*.*

4.6 The first occurrence of t

We use the above estimates to determine where a number $t \ge 1$ appears for the first time in A. We already know from Tables 1 and 2 that a 1 appears at position 1, a 2 at position 3, a 3 at position 9, and a 4 at position 220, so we may assume $t \ge 5$.

For fixed t, let $x(m)$ be the position where t appears for the first time in the sequence $A^{(m)}$, for $1 \le m \le t$. We will successively estimate $x(t), x(t-1), \ldots, x(1)$, working backwards from

$$A^{(t)} = \underbrace{t, t, \ldots, t}_{t+1 \text{ copies}}, t+1, \ldots,$$

where t appears as the leading term, $t+1$ appears for the first time at position $t+2$, and $t+2$ appears for the first time at position

$$\frac{(t+1)^{t+2} + 2t - 1}{t}\,. \tag{50}$$

For example, 3 appears in A at position 9, 4 in $A^{(2)}$ at position 42, and 5 in $A^{(3)}$ at position 343. Thus $x(t) = 1$, $x(t-1) = t+1$, and

$$x(t-2) = \frac{(t-1)^t + 2t - 5}{t-2}\,. \tag{51}$$

We first consider the case $t = 5$. Since $x(3) = 343$, a 5 appears in $A^{(2)}$ for the first time at the end of block $B_i^{(2)}$, where i is such that $\tau^{(2)}(i) = 343$. That is, i is determined (see (48) and Table 4) by the equation

$$\tau^{(2)}(i) = \underbrace{1+1+3+1+1+3+1+1+9+\cdots}_{i-1 \text{ terms}} = 343\,.$$

By direct calculation, $i = 80$, and again by direct calculation from (33), a 5 appears in $A^{(2)}$ at position

$$x(2) = \beta^{(2)}(80) = 77709404388415370160829246932345692180\,,$$

which is $10^{37.9\ldots}$. So 5 appears in $A = A^{(1)}$ at the end of block $B_i^{(1)}$, where i is such that $\tau^{(1)}(i) = x(2)$. Setting $m = 1$ in (49) we get

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

hence $\mu = 79.0\ldots$, $i = 2^{79.0\ldots}$. Setting $m = 1$, $n = 2^{79.0\ldots}$ in (39) we finally obtain

$$x(1) = \beta^{(1)}(2^{79.0\ldots}) = \varepsilon_1 2^{2^{79.0\ldots}} = 10^{10^{23.3\ldots}}$$

for the position of the first 5.

Consider now a general value of $t \ge 6$. To find $x(t-3)$, we must solve (from (51))

$$\tau^{(t-3)}(i) = \frac{(t-1)^t + 2t - 5}{t-2}\,.$$

Setting $i = (t-2)^\mu + 1$ and using (49) we get

$$\frac{t-1}{t-2}\left((t-1)^\mu - (t-3)(t-2)^{\mu-1}\right) = \frac{(t-1)^t + 2t - 5}{t-2}\,,$$

hence $\mu \approx t-1$, and so

$$x(t-3) = \beta^{(3)}((t-2)^\mu + 1) = \varepsilon_3 (t-2)^{(t-2)^{t-1}}\,.$$

The ε_3 may be ignored, since it can be absorbed into the tower of exponentials. The next iteration gives

$$x(t-4) = (t-3)^{(t-3)^{(t-2)^{t-1}}}$$

and eventually we obtain, for $t \ge 5$,

$$x(1) = 2^{2^{3^{4^{\cdot^{\cdot^{\cdot^{t-1}}}}}}}\,, \tag{52}$$

a tower of height $t-1$. We formalize this as our final

Conjecture 4.4. *The number $t \geq 5$ appears for the first time in the sequence A at about position (52).*

Remark. The position of the first 5 can be estimated more accurately using the tabular construction of Section 4.2, thus avoiding the simplifying assumptions made in Section 4.5. This analysis predicts that the first 5 will appear at about position

$$\varepsilon_1 \times 2^{4180901959526919227883 53} = 10^{10^{23.09987\ldots}}.$$

where ε_1 is given in §4.3. We omit the details.

5 Comments and generalizations

5.1 The Finiteness Conjecture

The proof of Theorem 3.2 would have been simpler if we had known in advance that the glue strings $S_n^{(m)}$ were finite. This would follow from the following:

Finiteness Conjecture. *For integers $m \geq 2$ and $r \geq 1$, let $x(1), x(2), \ldots, x(r)$ be a string from $\mathbb{P}_m^+$. Let $x(n+1) = \mathcal{C}(x(1), x(2), \ldots, x(n))$ for $n \geq r$. Then for some $n \geq r+1$, $x(n) < m$.*

In other words, there is no finite starting string from $\mathbb{P}_m^r$ which extends under repeated application of the map $\mathcal{C}$ to an infinite sequence from $\mathbb{P}_m^*$. Sooner or later a term less than m must appear.

Although this conjecture seems very plausible, we have not been able to find a proof. If one tries to construct a starting string which extends for a long time without dropping below m one quickly runs into difficulties. Let m_1 and m_2 be respectively the smallest and largest values in the starting string. Then no number outside the range $[m_1, m_2]$ — or in fact any number not in the starting string — may appear in the resulting sequence, for such a number is immediately followed by a 1, terminating the sequence. So if the sequence is infinite it must be bounded.

As an experiment we considered all 2^n starting strings of length $n \leq 30$ consisting just of 2's and 3's, and calculated the lengths of the resulting strings until just before the appearance of the first 1. The maximum and average lengths are shown in Table 9 below.

n	Max.	Ave.
1	1	1
2	4	2.75
3	5	3.75
4	8	5.125
5	9	6.2187
6	14	7.5
7	15	8.5703
8	66	10.2734
9	68	11.3828
10	70	12.5293
11	123	13.6099
12	124	14.6658
13	125	15.6683
14	132	16.6957
15	133	17.7047
16	134	18.7168
17	135	19.7206
18	136	20.7278
19	138	21.7304
20	139	22.7341
21	140	23.7353
22	142	24.7372
23	143	25.7379
24	144	26.7388
25	145	27.7391
26	146	28.7396
27	147	29.7398
28	148	30.74
29	149	31.7401
30	150	32.7402

Table 9: Maximum and average length of string produced by any starting sequence of n 2's and 3's, stopping when first 1 is reached.

The average length seems to approach $n+$ constant, but the maximum length is harder to understand, and it would be nice to have more data. Does the maximum length continue to grow linearly, or are there further jumps of ever-increasing size? We do not know.

Table 10 shows the starting strings of lengths 2, 4, 6, 8 and 11 (when there are jumps in the maximum length) and the strings of record lengths 4, 8, 14, 66 and 123 that they produce. These five starting strings are unique.

$n = 2$	Starting string 2 2 produces length 4: 2 2 2 3
$n = 4$	Starting string 2 3 2 3 produces length 8: 2 3 2 3 2 2 2 3
$n = 6$	Starting string 2 2 2 3 2 2 produces length 14: 2 2 2 3 2 2 2 3 2 2 2 3 3 2
$n = 8$	Starting string 2 3 2 2 2 3 2 3 produces length 66: 2 3 2 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 2 2 3 2 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 2 2 3 2 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 2 3 3 2
$n = 11$	Starting string 2 2 3 2 3 2 2 2 3 2 2 produces length 123: 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 2 2 3 2 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 2 2 3 2 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 2 2 3 2 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 2 2 3 2 2 2 3 2 3 2 2 2 3 2 2 2 3 2 2 3 2 3 2 2 2 3

Table 10: Starting strings of n 2's and 3's which extend under the map $\mathcal{C}$ for a record number of steps before reaching a 1. Each entry also shows the final string (until just before the first 1 is reached).

5.2 Curling number transforms of other sequences

It is interesting to apply the curling number transform to other sequences, particularly those for which the definition involves properties of substrings. For example, the binary Thue-Morse sequence

$$\begin{array}{cccccccccccccccccc} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & \ldots \end{array}$$

(A10060 in [8]) has the property that it contains no cubes UUU as substrings (see [2], [5], [6], [8] for further information). Its curling number transform, which naturally contains only 1's and 2's, is

$$\begin{array}{cccccccccccccccccc} 1 & 1 & 1 & 2 & 1 & 1 & 2 & 2 & 1 & 2 & 1 & 2 & 2 & 1 & 2 & 2 & \\ 1 & 2 & 2 & 2 & 1 & 2 & 2 & 2 & 2 & 2 & 1 & 2 & 2 & 1 & 2 & 2 & \ldots \end{array}$$

(A93914). We leave it to the interested reader to investigate the properties of this sequences and of the other new sequences mentioned below.

There are many examples of ternary sequences which contain no squares, and of course their curling number transforms are simply the all-ones sequence $\mathbf{1}$. However, the lexicographically earliest sequence from $\mathbb{P}^*$ whose transform is $\mathbf{1}$ is the ruler sequence r (A1511) mentioned in Section 4.4.

We give one further example. The Kolakoski sequence is a sequences of 1's and 2's defined by $K(1) = 1$, $K(n) =$ length of n-th run:

$$\begin{array}{cccccccccccccccccc} 1 & 2 & 2 & 1 & 1 & 2 & 1 & 2 & 2 & 1 & 2 & 2 & 1 & 1 & 2 & 1 & \\ 1 & 2 & 2 & 1 & 2 & 1 & 1 & 2 & 1 & 2 & 2 & 1 & 1 & 2 & 1 & 1 & \ldots \end{array}$$

(A2 in [8]). This also contains no cubes. The transformed sequence is

$$\begin{array}{cccccccccccccccccc} 1 & 1 & 1 & 2 & 1 & 2 & 1 & 1 & 2 & 2 & 1 & 2 & 2 & 2 & 2 & 1 & \\ 1 & 2 & 2 & 2 & 1 & 1 & 2 & 2 & 1 & 2 & 2 & 2 & 1 & 2 & 1 & 1 & \ldots \end{array}$$

(A93921).

5.3 Generalizations

In this final section we briefly mention a few of the possible generalizations of the sequence A.

(i) The recurrence (2) may be replaced by $a(1) = 1$, $a(n+1) = f(\mathcal{C}(a(1), \ldots, a(n)))$ for $n \geq 1$, for any suitable function f. For example, $f(x) = \text{floor}(x/2)$ produces

$$\begin{array}{cccccccccccccccccc} 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 0 & 0 & 1 & 0 & 0 & 1 & \\ 1 & 1 & 1 & 2 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 0 & \ldots \end{array}$$

(A91970), which presumably has an even slower rate of growth than A.

(ii) $a(1) = a(2) = 1$, $a(n+2) = \mathcal{C}(a(1), \ldots, a(n))$ for $n \geq 1$ produces

$$\begin{array}{cccccccccccccccccc} 1 & 1 & 1 & 2 & 3 & 1 & 1 & 1 & 2 & 3 & 1 & 2 & 2 & 1 & 2 & 1 & \\ 1 & 2 & 2 & 1 & 2 & 1 & 1 & 2 & 2 & 2 & 2 & 3 & 4 & 1 & 1 & 1 & \ldots \end{array}$$

(A94006). This has the property that its curling number transform is the same sequence but shifted one place to the left.

(iii) A greedy version of $A^{(2)}$. Let $g(1) = 2$. For $n \geq 1$, let $k = \mathcal{C}(g(1), \ldots, g(n))$. If $k > 1$, $g(n+1) = k$ (as in $A^{(2)}$), but if $k = 1$, choose $g(n+1)$ so that $\mathcal{C}(g(1), \ldots, g(n+1))$ is maximized. If there is more than one choice for $g(n+1)$, pick the smallest. The resulting sequence (A94321) is:

$$\begin{array}{cccccccccccccccccc} 2 & 2 & 2 & 3 & 3 & 2 & 2 & 2 & 3 & 3 & 2 & 2 & 2 & 3 & 2 & 2 & \\ 2 & 3 & 2 & 2 & 2 & 3 & 3 & 2 & 2 & 2 & 3 & 2 & 2 & 2 & 3 & 2 & \ldots \end{array}$$

(iv) A two-dimensional version of A. Define $t(i, j)$, $i \geq 1$, $j \geq 1$, as follows: $t(i, 1) = t(1, i) = a(i)$. For $i, j \geq 1$, $t(i+1, j+1) = \max\{k_1, k_2\}$, where $k_1 = \mathcal{C}(t(i+1, 1), t(i+1, 2), \ldots, t(i+1, j))$, $k_2 = \mathcal{C}(t(1, j+1), t(2, j+1), \ldots, t(i, j+1))$ — see Table 11 (A94781). The first two rows (or columns) give A and the third row (or column) is $A^{(2)}$.

1	1	2	1	1	2	2	2	3	1	1	2	1	1	...
1	1	2	1	1	2	2	2	3	1	1	2	1	1	...
2	2	2	3	2	2	2	3	2	2	2	3	3	2	...
1	1	3	1	1	3	3	2	1	1	2	1	1	2	...
1	1	2	1	1	2	2	2	3	1	2	1	1	2	...
2	2	2	3	2	1	1	2	1	2	3	2	2	3	...
.	.	.	.	.	.	.	.	.	.	.	.	.	.	...

Table 11: A two-dimensional version of the sequence.

(v) J. Taylor [9] has suggested two broad generalizations of the original recurrence. Let $\sim$ be an equivalence relation on strings of integers of each fixed length. Write

$$a(1)a(2)\cdots a(n) = XY_1Y_2\cdots Y_k\,, \tag{53}$$

where the Y_i are nonempty strings with $Y_1 \sim Y_2 \sim \cdots \sim Y_k$ and k is maximal; then $a(n+1) = k$. Choosing $\sim$ to be the identity relation gives A. Taylor has contributed several interesting generalizations of A to [8] obtained from other equivalence relations. For example, if two strings are equivalent if one is a permutation of the other, the resulting sequence is

$$\begin{array}{cccccccccccccccccc} 1 & 1 & 2 & 1 & 1 & 2 & 2 & 2 & 3 & 1 & 1 & 2 & 1 & 1 & 2 & 2 & \\ 2 & 3 & 2 & 2 & 2 & 3 & 2 & 2 & 2 & 3 & 3 & 2 & 2 & 4 & 1 & 1 & \ldots \end{array}$$

(A91976), which agrees with A for the first 19 terms. But after

$$1\ \ 1\ \ 2\ \ 1\ \ 1\ \ 2\ \ 2\ \ 2\ \ 3\ \ 1\ \ 1\ \ 2\ \ 1\ \ 1\ \ 2\ \ 2\ \ 2\ \ 3\ \ 2$$

the next term is now 2, not 1, since we can take $U = 1\ 1\ 2\ 1\ 1$, $Y_1 = 2\ 2\ 2\ 3\ 1\ 1\ 2$, $Y_2 = 1\ 1\ 2\ 2\ 2\ 3\ 2$, where Y_2 is a permutation of Y_1.
(vii) More generally, Taylor suggests using a partial order $\prec$ on integer strings of all lengths (not just strings of the same length), and requiring the Y_i in (53) to be nonempty and satisfy $Y_1 \prec Y_2 \prec \cdots \prec Y_k$ where k is maximal. For further examples of Taylor's sequences the reader is referred to the entries A91975 and A92331–A92335 in [8].

Acknowledgements

We thank J. Taylor for telling us about his generalizations of the sequence.

References

[1] W. Ackermann, Zum Hilbertschen Aufbau der reellen Zahlen, *Math. Ann.* **99** (1928), 118–133.

[2] J.-P. Allouche and J. Shallit, *Automatic Sequences*, Cambridge Univ. Press, 2003.

[3] J. H. Conway, The weird and wonderful chemistry of audioactive decay, in T. M. Cover and Gopinath, eds., *Open Problems in Communication and Computation*, Springer-Verlag, NY, 1987, pp. 173–188.

[4] H. M. Friedman, Long finite sequences, *J. Combin. Theory*, Ser **A**, **95** (2001), 102–144.

[5] M. Lothaire, *Combinatorics on Words*, Addison-Wesley, Reading, MA, 1983.

[6] A. Salomaa, *Jewels of Formal Language Theory*, Computer Science Press, Rockville, MD, 1981.

[7] M. Sharir and P. K. Agarwal, *Davenport-Schinzel Sequences and Their Geometric Applications*, Cambridge Univ. Press, 1995.

[8] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, published electronically at http://www.research.att.com/~njas/sequences/, 1996–2006.

[9] J. Taylor, personal communication.

2000 *Mathematics Subject Classification*: Primary 11B37.
Keywords: recurrence, Friedman sequence, Ackermann function.

(Concerned with sequences A000002, A001511, A010060, A090822, A091408, A091409, A091410, A091411, A091412, A091413, A091579, A091586, A091587, A091588, A091787, A091799, A091839, A091970, A091970, A091975, A091976, A092331, A092332, A092333, A092334, A092335, A093914, A093921, A093955, A093956, A093957, A093958, A094006, A094321, and A094781.)

Received June 26 2004; revised versions received February 22 2006 and September 13 2006. Published in *Journal of Integer Sequences*, December 16 2006.

Return to Journal of Integer Sequences home page.