

Planar digraphs for automatic complexity*

Achilles A. Beros, Bjørn Kjos-Hanssen^[0000–0002–6199–1755], and Daylan Kauí Yogi

University of Hawai‘i at Mānoa, Honolulu HI 96822, USA
 {beros,bjoernkh,dkyogi64}@hawaii.edu

Abstract. We show that the digraph of a nondeterministic finite automaton witnessing the automatic complexity of a word can always be taken to be planar. In the case of total transition functions studied by Shallit and Wang, planarity can fail.

Let $s_q(n)$ be the number of binary words x of length n having nondeterministic automatic complexity $A_N(x) = q$. We show that s_q is eventually constant for each q and that the eventual constant value of s_q is computable.

Keywords: Automatic complexity · planar graph · Möbius function · Nondeterministic finite automata.

1 Introduction

Automatic complexity, introduced by Shallit and Wang [7], is an automata-based and length-conditional analogue of Sipser’s CD complexity [8] which is in turn a computable analogue of the noncomputable Kolmogorov complexity. The nondeterministic case was taken up by Hyde and Kjos-Hanssen [3], who gave a table of the number of words of length n of a given complexity q for $n \leq 23$. The numbers in the table suggested (see Table 2) that the number may be eventually constant for each fixed q . Here we establish that that is the case (Theorem 9), and show that the limit is computable (in exponential time). Moreover, we narrow down the possible automata that are needed to witness nondeterministic automatic complexity: they must have planar digraphs, in fact their digraphs are trees of cycles in a certain sense.

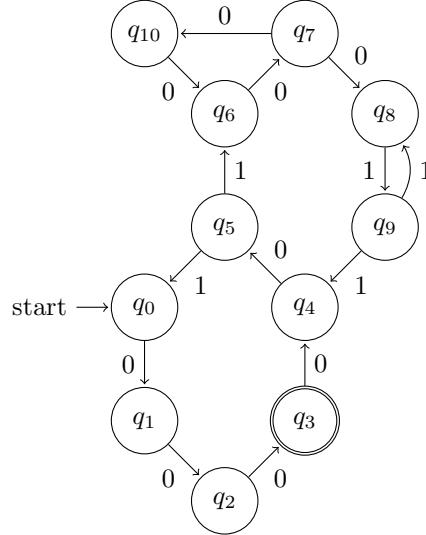
We recall our basic notion.

Definition 1 ([7]) *The nondeterministic automatic complexity $A_N(x)$ of a word x is the minimal number of states of a nondeterministic finite automaton M (without ϵ -transitions) accepting x such that there is only one accepting path in M of length $|x|$.*

* This work was partially supported by a grant from the Simons Foundation (#315188 to Bjørn Kjos-Hanssen). We are indebted to Jeff Shallit and Malik Younsi for helpful comments.

2 Automatic complexity as chains of trees of lumps

Consider the version of automatic complexity where the transition functions are not required to be total.¹ Then we claim that the digraphs representing the witnessing automata are planar, in fact they are “trees of cycles”. As an example, for the word $0^5 10^5 1^6 010^3$, we have the following witnessing automaton:



To explain this, first let us say that a cycle is a sequence of states that starts and ends with the same state. Let us say that a lump is the automaton whose transitions come from a given cycle. So if a cycle is repetitive, like 3456734567345673, then it generates the same lump as just 345673.

Consider the sequence of states visited during processing of a unique accepted word x of length n . Let us call the first visited state 0, the next distinct state 1, and so on. (So for example the permitted state sequences of length 3 are only 000, 001, 010, 011, 012.)

Then the state sequence starts $0, 1, \dots, q, q+1, \dots, q$ where q is the first state that is visited twice. Now the claim is that there will never, at a later point in the state sequence, be a transition (an edge) q_1, q_2 such that q_2 occurs within the lump generated by the cycle $q, q+1, \dots, q$ and such that the transition q_1, q_2 does not occur in that lump. Indeed, otherwise our state sequence would start

$$0, 1, \dots, \underbrace{q, \dots, q_2}_{\text{first}}, \dots, \underbrace{q, \dots, q_1, q_2}_{\text{second}}$$

and then there is a second accepting path of the same length where the first and second segments are switched.

¹ Whether determinism is required is not important in the following, but in the nondeterministic case we assume we require there to be only one accepting path, as usual.

Consequently, the path can only return to states that are not yet in any lumps. This leaves only two choices whenever we decide to create a new edge leading to a previously visited state:

Case 1. Go back to a state that was first visited after the last completed lump so far seen, or Case 2. Go back to a state that was first visited at some earlier time, before some of the lumps so far seen started (and in general after some of them were complete).

This gives a tree of lumps where each new lump either (Case 1) creates a new sibling for the previous lump, or (Case 2) creates a new parent for a final segment of the so far seen top-level siblings. In this tree of lumps, only the leaves (the lumps that are not anybody's parents) can be traversed more than once by the uniquely accepted path of length n .

So if the first lump created is l_1 then next we can have two cases:

$$(l_1, \quad l_2) \quad (\text{Case 1})$$

$$l_1 \rightarrow l_2 \quad (\text{Case 2})$$

In Case 1, l_1 and l_2 are siblings ordered from first to second. In Case 2, \rightarrow denotes *is a child of*, which by definition is the same as *sub-digraph*. Now for the third lump l_3 , we have only the following possibilities:

$$(l_1, \quad l_2, \quad l_3) \quad (\text{Subcase 1.1})$$

$$(l_1, \quad l_2 \rightarrow l_3) \quad (\text{Subcase 1.2})$$

$$(l_1, l_2) \rightarrow l_3 \quad (\text{Subcase 1.3})$$

$$(l_1 \rightarrow l_2, \quad l_3) \quad (\text{Subcase 2.1})$$

$$l_1 \rightarrow l_2 \rightarrow l_3 \quad (\text{Subcase 2.2})$$

In Subcase 1.2, l_1 and l_3 are siblings and l_2 is a child of l_3 . In Subcase 1.3, l_3 is a common parent of l_1 and l_2 . In Subcase 2.1, l_3 is a new sibling for l_2 , and l_2 still has l_1 as its child. In Subcase 2.2, l_3 is a parent of l_2 .

For instance, the state sequence 01234567345673456720 has the structure of Subcase 2.2, with l_1 being the lump generated from 345673, l_2 being generated from 23456734567345672, and l_3 being generated from the whole sequence 01234567345673456720. The corresponding automaton is shown in an online tool.² Using this planarity result, we are able to increase the speed of our algorithm for calculating $A_N(x)$. Consequently, we have been able to extend the string length in our computations from $n = 23$ to $n = 25$. The number of maximally complex binary words of a given length are shown in Table 1. A similar table for $n \leq 23$ was given in [3].

² <http://math.hawaii.edu/wordpress/bjoern/complexity-of-0001110111011111/>

n	#	2^n	%complex	2^n -#
0	1	1	100.00%	0
1	2	2	100.00%	0
2	2	4	50.00%	2
3	6	8	75.00%	2
4	8	16	50.00%	8
5	24	32	75.00%	8
6	30	64	46.88%	34
7	98	128	76.56%	30
8	98	256	38.28%	158
9	406	512	79.30%	106
10	344	1,024	33.59%	680
11	1,398	2,048	68.26%	650
12	1,638	4,096	39.99%	2,458
13	5,774	8,192	70.48%	2,418
14	5,116	16,384	31.23%	11,268
15	23,018	32,768	70.25%	9,750
16	22,476	65,536	34.30%	43,060
17	86,128	131,072	65.71%	44,944
18	89,566	262,144	34.17%	172,578
19	351,250	524,288	67.00%	173,038
20	375,710	1,048,576	35.83%	672,866
21	1,461,670	2,097,152	69.70%	635,482
22	1,539,164	4,194,304	36.70%	2,655,140
23	5,687,234	8,388,608	67.80%	2,701,374
24	6,814,782	16,777,216	40.62%	9,962,434
25	24,031,676	33,554,432	71.62%	9,522,756
26	27,782,964	67,108,864	41.40%	39,325,900
27	97,974,668	134,217,728	73.00%	36,243,060

Table 1: Lengths n , number of words of length n of maximal $A_N(x)$, 2^n , percentage of maximally complex words, number of non-maximally complex words.

3 The asymptotic number of words of given complexity

In this section, we examine the asymptotic behavior of the number of words with automatic complexity q for a fixed $q \in \mathbb{N}$.

Definition 2 A binary word x is right inextendible if $A_N(x) < A_N(x0)$ and $A_N(x) < A_N(x1)$.

Inextendibility is closely related to volatility of the automatic complexity, as examined in the Complexity Option Game [5]. The number and proportion of right-inextendible words of length n and complexity q can be examined using an online database [4] and is shown in Table 2 for small q and n .

A basic procedure in our results will be the counting of periodic words, since a cycle containing a periodic word can be shortened and an automaton containing such a cycle will not be optimal.

Definition 3 A word x is periodic if there exists a subword $y \neq x$ and an integer n such that

$$\underbrace{yyy \cdots y}_n = x.$$

A non-periodic word [2] is also called a primitive word and one starting with 0, in our setting, is called a *Lyndon word* [6].

Definition 4 ([1]) Let n be a positive integer with $\omega(n)$ denoting the number of distinct prime factors of n and $\Omega(n)$ denoting the total number of prime factors (i.e., with repetition) of n . The Möbius function μ is defined as

$$\mu(n) := \begin{cases} (-1)^{\omega(n) \bmod 2} & \text{if } \Omega(n) = \omega(n), \\ 0 & \text{if } \Omega(n) > \omega(n). \end{cases}$$

Theorem 5 ([2]) The number of unique periodic binary words of length n is given by $Z(0) = 0$ and for $n \geq 1$,

$$Z(n) = 2^n - \sum_{d|n} \mu\left(\frac{n}{d}\right) \cdot 2^d.$$

Recall that a *necklace* is an equivalence class of non-periodic words under cyclic rotation. Thus, for instance, $\{0011, 0110, 1100, 1001\}$ is a necklace. Theorem 5 is a restatement of the following classical result.

Theorem 6 (Witt's Formula [9]) The number of necklaces of binary words of length n is

$$\frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) \cdot 2^d.$$

Definition 7 We define the set $S_q(n) = \{x \in \{0, 1\}^n : A(x) = q\}$ and $s_q(n) = |S_q(n)|$.

$n \setminus q$	3	4	5	6	7	8	9	10
22	8/20	28/58	86/164	322/502	1288/2846	6594/16024	44922/94732	220544/451368
21	8/20	28/58	98/176	292/496	1318/3168	8472/18720	52178/108042	266760/504794
20	8/20	28/58	86/164	238/430	1478/3814	11670/23328	54990/115896	278696/529148
19	8/20	28/58	86/164	402/582	2380/4996	12312/26542	78892/410668	134578/351250
18	8/20	28/58	110/188	356/598	2070/5692	14456/29990	68288/36024	0/0
17	8/20	28/58	104/200	262/514	2850/7102	20516/37042	30486/86128	
16	8/20	28/58	80/164	536/752	2908/7738	14230/34320	0/22476	
15	8/20	28/58	148/226	578/908	3338/8530	7524/23018		
14	8/20	28/58	112/244	774/1270	4442/9868	0/5116		
13	8/20	28/58	120/250	1396/2076	1736/5774			
12	8/20	28/58	158/282	1048/2090	0/1638			
11	8/20	28/58	384/564	576/1398				
10	8/20	34/64	244/588	0/344				
9	8/20	48/78	112/406					
8	8/20	82/130	0/98					
7	10/22	38/98						
6	14/26	0/30						
5	8/24							
4	0/8							

Table 2: Proportions $r_q(n)/s_q(n)$ of right-inextendible binary words of automatic complexity q and length n .

Definition 8 Given an automaton, G , whose set of states is Q , we define a detour to be a pair of finite non-trivial sequences of states, $\alpha, \beta \in Q^*$, such that $\alpha(0) = \beta(0)$, $\alpha(|\alpha| - 1) = \beta(|\beta| - 1)$ and $\alpha \neq \beta$. We call a detour minimal if $\{\alpha(i) : 0 < i < |\alpha| - 1\} \cap \{\beta(i) : 0 < i < |\beta| - 1\} = \emptyset$.

Consider an automaton with a single cycle (Figure 5). Suppose the automaton has i states before the cycle and ℓ states after the cycle (which implies that there are $q - (i + \ell)$ states within the cycle). We now obtain a formula for the limit of the number of binary words of given complexity q .

Theorem 9 s_q is eventually constant, with limiting value

$$\sum_{\substack{i, \ell \geq 0 \\ i + \ell < q}} 2^{(i-1)^+} \cdot [2^{q-(i+\ell)} - Z(q - (i + \ell))] \cdot 2^{(\ell-1)^+},$$

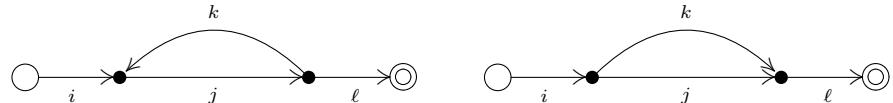
where Z was defined in Theorem 5 and

$$x^+ = \max\{x, 0\}.$$

Proof. Consider an arbitrary automaton G with q states. There are a finite number of such automata. We will prove that unless G has at most one minimal detour, there is an N such that, for all $n \geq N$, G cannot accept a unique word of length n .

We begin with the observation that we may assume that G has a unique initial state and a unique accepting state.

If G has at most one detour, then G has one of the the following forms.



If G is of the type on the right and G accepts a unique word σ of length n , then any accepting path for σ either uses the k states that comprise the top path of the detour, or uses the j states that comprise the bottom path, but no both. Thus, if both k and j are non-zero, there is an automaton with fewer states that accepts only σ among all words of length n . We conclude that in the case of automata with at most one minimal detour, we need only consider ones of the form on the left.

Now, we consider the possibilities for automata with at least two distinct minimal detours. Each of the twelve cases in Figure 1 falls into one of three cases.

1. On any accepting path, each detour can be used at most once ((1), (2) and (3)).
2. On any accepting path, one of the detours can be used at most once ((7), (8), (10), (11) and (12)).

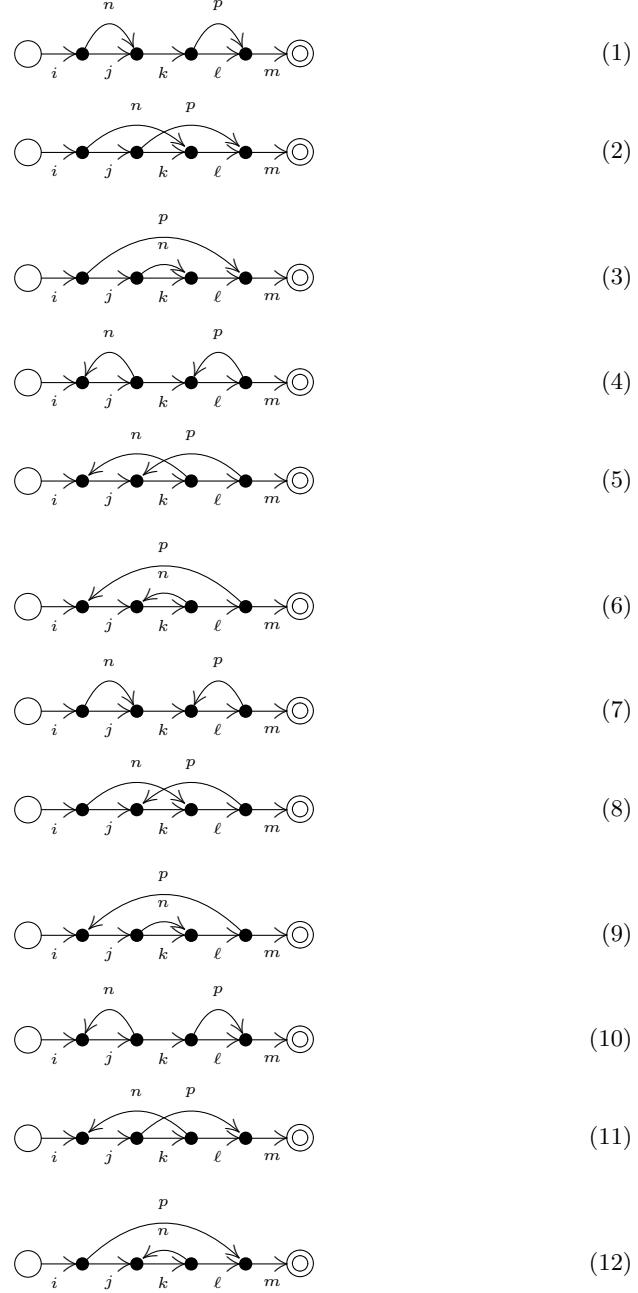


Fig. 1: The possibilities for automata with at least two distinct minimal detours.

3. There are accepting paths that use each of the detours an arbitrary number of times ((4), (5), (6) and (9)).

These further break down as follows:

- (1), (4), (7), (10) represent two separated cycles;
- (2), (5), (8), (11) represent overlapping cycles.
- (3), (6), (9), (12) represent nested cycles; and

If G falls into the first case, then σ is also uniquely accepted among words of length n by an automaton with at most q states and no detours. If G falls into the second case, then σ is uniquely accepted by an automaton with at most q states and at most one detour. If G falls into the third category, then there are two cycles (although they may have common transitions) which can each be traversed and independent and arbitrary number of times on an accepting path. Thus, for large enough n , the cycles can be traversed in different orders or different numbers of times and still reach an accepting state, thereby violating the requirement that G accept exactly one word of length n .

As an example of the third case, suppose that G is of the type shown in (9). G has two independent cycles, one of length $p + j + k + \ell$ and the other of length $p + j + n + \ell$. Let $N = i + a(p + j + k + \ell) + m = i + b(p + j + n + \ell) + m$, where $a, b \in \mathbb{N}$. There are at least two words of length N that G accepts, and for any $M \geq N$ such that G accepts a word of length M , G must accept at least two words of length M .

In conclusion, we may assume our automata have at most one detour. Thus they consist of a chain of states, followed by a single (in general multi-state) cycle, followed by another chain. Let i be the number of states before the cycle, ℓ the number of states after the cycle, so that $q - (i + \ell)$ if the number of states within the cycle. If the bits read within the cycle do not form a necklace, we can reduce the number of states. Thus there are $[2^{q-(i+\ell)} - Z(q - (i + \ell))]$ states within the cycle. The an upper bound for the total number of binary words with $A_N(x) = q$ is

$$2^i \cdot 2^\ell \cdot [2^{q-(i+\ell)} - Z(q - (i + \ell))].$$

Let ξ be the bit that advances the automaton from the i th state to the $(i + 1)$ th state (i.e. the transition that takes the automaton into the cycle) and η be the bit that advances that automaton from the $q - (i + \ell)$ th state to the $(i + 1)$ th state (i.e, the transition that completes the cycle). If $\xi = \eta$, then it is possible to create an automaton with fewer states that accepts the same word and no other of length n . A similar consideration applies upon leaving the cycle. Thus, we have

$$2^{(i-1)^+} \cdot [2^{q-(i+\ell)} - Z(q - (i + \ell))] \cdot 2^{(\ell-1)^+}$$

possible words.

Finally, to conclude that $s_q(n)$ is eventually constant, note that while the single cycle will have to be exited at different points depending on $n \bmod k$, where k is the length of the main cycle, there will always be exactly one value of $n \bmod k$ and hence exactly one automaton contributed from the cycle and the

given “head” and “tail” words. See Figures 2, 3, and 4 for illustrations of the cases $q = 2, 3, 4$, respectively.

Remark 10 *Here is perhaps a simpler view of the classification of detours in Figure 1. Suppose A is an NFA that uniquely accepts some word. Now consider some shortest directed path P from q_0 to the unique final state q_f . Let us say that an alternate route is any simple directed path, edge-disjoint from P , joining two vertices of P .*

Suppose there are two alternate routes, Q and R , joining q_i and q_j , and q_k and q_l , respectively. If we do not worry about the direction of the paths for the moment, we may assume $i \leq j$ and $k \leq l$. Then there are three possibilities:

1. $j \leq k$: Q precedes R ;
2. $k \leq i$ and $j \leq l$: Q encompasses R ;
3. $i \leq k \leq j \leq l$: Q and R overlap.

Furthermore, for Q and R one can choose the direction of the edges independently. This gives $3 \cdot 4 = 12$ possibilities to consider.

Count	Regex	Automaton		
1	01^*	<pre> graph LR start(()) --> q1((q1)) q1 -- 0 --> q2((q2)) q2 -- 1 --> q2 </pre>		
2	0^*1	<pre> graph LR start(()) --> q1((q1)) q1 -- 0 --> q1 q1 -- 1 --> q2((q2)) </pre>		
3	$(01)^*$	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center; width: 50%;"> $n \text{ odd}$ <pre> graph LR start(()) --> q1((q1)) q1 -- 0 --> q2((q2)) q2 -- 1 --> q1 </pre> </td> <td style="text-align: center; width: 50%;"> $n \text{ even}$ <pre> graph LR start(()) --> q1((q1)) q1 -- 1 --> q2((q2)) q2 -- 0 --> q1 </pre> </td> </tr> </table>	$n \text{ odd}$ <pre> graph LR start(()) --> q1((q1)) q1 -- 0 --> q2((q2)) q2 -- 1 --> q1 </pre>	$n \text{ even}$ <pre> graph LR start(()) --> q1((q1)) q1 -- 1 --> q2((q2)) q2 -- 0 --> q1 </pre>
$n \text{ odd}$ <pre> graph LR start(()) --> q1((q1)) q1 -- 0 --> q2((q2)) q2 -- 1 --> q1 </pre>	$n \text{ even}$ <pre> graph LR start(()) --> q1((q1)) q1 -- 1 --> q2((q2)) q2 -- 0 --> q1 </pre>			

Fig. 2: The witnessing automata for $\lim_n s_q(n)/2 = 3$, $q = 2$. The first two are used at any length n , whereas the bottom two are each used only for one value of $n \bmod 2$, illustrating Theorem 11.

The main proviso to Theorem 9 may be that while *the number of words with given complexity reaches a limit, the set of witnessing automata does not quite*. To wit:

Theorem 11 *There is a q such that there is no set of automata M_1, \dots, M_s such that for all sufficiently large n ,*

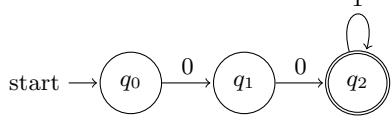
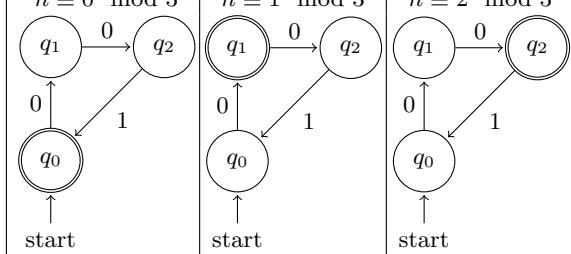
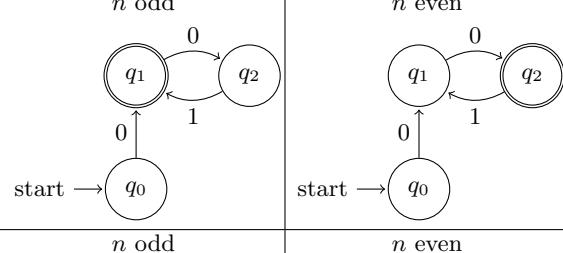
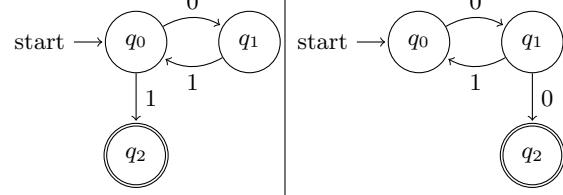
Count	Regex	Automata		
1-5	001* (shown) 010*, 01*0 0*10, 0*11			
6-8	(001)* (shown) (010)*, (011)*			
9	0(01)*			
10	(01)*x			

Fig. 3: Automata and regular expressions witnessing $\lim_n s_q(n)/2 = 10$ for $q = 3$. The exponents indicated by * are not necessarily integers (so that for instance $abcd^{1.5} = abcdab$). The letter x indicates 0 or 1, chosen so as to break a pattern.

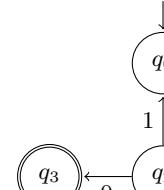
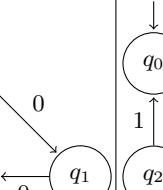
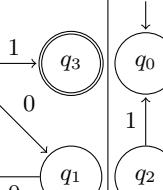
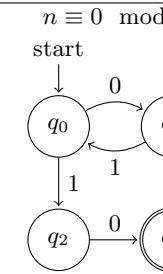
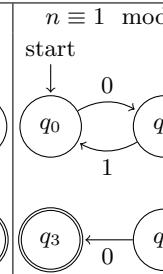
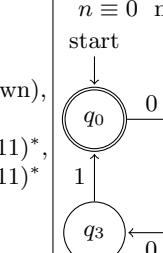
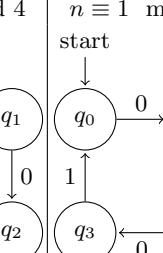
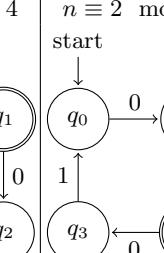
Count	Regex	Automata		
1–3	$(001)^*x$ (shown), $(010)^*x$, $(011)^*x$	$n \equiv 0 \pmod{3}$ start 	$n \equiv 1 \pmod{3}$ start 	$n \equiv 2 \pmod{3}$ start 
4–6	$0(001)^*$, $0(011)^*$, $0(101)^*$	edge followed by cycle of length 3		
7–8	$(01)^*x0$ (shown), $(01)^*x1$	$n \equiv 0 \pmod{2}$ start 	$n \equiv 1 \pmod{2}$ start 	
9	$0(01)^*x$	edge followed by cycle of length 2 followed by edge		
10–15	$(0001)^*$ (shown), $(0010)^*$, $(0100)^*$, $(0011)^*$, $(0110)^*$, $(0111)^*$	$n \equiv 0 \pmod{4}$ start 	$n \equiv 1 \pmod{4}$ start 	$n \equiv 2 \pmod{4}$ start 
16–17	$00(01)^*$, $01(10)^*$	two edges followed by cycle of length 2		
18–21	0^*101 , 0^*110 , 0^*111 , 0^*100	loop followed by chain		
22–25	0010^* , 0001^* , 0110^* , 0101^*	chain followed by loop		
26–29	010^*1 , 001^*0 , 01^*00 , 01^*01	chain of edges with a single loop near middle		

Fig. 4: Witnessing automata for $\lim_n s_q(n)/2 = 29$, $q = 4$. The exponents indicated by * are not necessarily integers, and the letter x indicates 0 or 1, chosen so as to break a pattern.

- for each i there is some x of length n such that $A_N(x) = q$ and M_i witnesses the inequality $A_N(x) \leq q$, and
- for all x of length n , $A_N(x) = q$ iff the inequality $A_N(x) \leq q$ is witnessed by one of the M_i .

Proof. Let $q = 2$. The limiting value of s_q is 6 as witnessed by the patterns: 0^*1 , 01^* , $(01)^*$. However, for $(01)^*$, different states will be the final state depending on the length $n \bmod 2$; see Figure 2.

Theorem 12 (Number of right-inextendible words) *For $q \geq 1$, define a function r_q by*

$$r_q(n) = \#\{x \in \{0, 1\}^n \mid A_N(x) + 1 = A_N(x0) = A_N(x1)\}.$$

Then r_q is eventually constant, with limiting value

$$\sum_{\substack{i \geq 0, \ell > 0 \\ i + \ell < q}} 2^{(a-1)^+} \cdot [2^{q-(i+\ell)} - Z(q - (i + \ell))] \cdot 2^{\ell-1},$$

where $Z(n)$ refers to the function defined in Theorem 2, and $(x-y)^+ := \max\{(x-y), 0\}$.

Proof. Let x be a binary word such that its accepting automaton has a single cycle, as in Figure 5. As shown in Theorem 9, we need only consider this particular case. Let ℓ be the number of states between the cycle and the accepting state of the automaton.

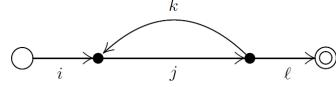


Fig. 5: Schematic of an automaton with a single cycle.

Suppose $\ell = 0$. Then the accepting state must be one of the states within the cycle. Without loss of generality, suppose the path out of the accepting state is triggered by a 0 input. Then $x0$ must have the same automatic complexity as x , as appending 0 to x does not require the addition of any additional states, and x is thus not inextendible. Thus, for a word to be inextendible, it is necessary that $\ell > 0$.

Theorem 13 $s_q(n)$ is eventually bounded by $2^{q-2} \left(\frac{q(q+5)}{2} + 1 \right)$.

Proof. By Theorem 9, we can upper bound the sum by

$$\sum_{i, \ell \geq 0, i + \ell < q} 2^q = \binom{q+1}{2} 2^q.$$

In fact, by considering the four possible truth values for the cases $i = 0$, $\ell = 0$, we get the upper bound

$$\begin{aligned} \sum_{i=\ell=0} 2^q + \sum_{i\ell=0, i+\ell>0} 2^{q-1} + \sum_{i>0, \ell>0} 2^{q-2} &= 2^q + 2(q-1)2^{q-1} + \binom{q-1}{2} 2^{q-2} \\ &= 2^{q-2} \left(4q + \binom{q-1}{2} \right) = 2^{q-2} \left(\frac{q(q+5)}{2} + 1 \right). \end{aligned}$$

Remark 14 A comparison of s_q with the bound in Theorem 13 can be done using the computer code in Figure 6. The number in the title of this section was calculated using that Python script and using a table of values of Z from the OEIS database. Table 3 shows an initial segment of the resulting sequence. There we count only words starting with 0, so that the full number would be twice that, matching the impression that $\lim_n s_3(n) = 20$ given by Table 2.

q	$\lim_n s_q(n)/2$	q	$\lim_n s_q(n)/2$
1	1	21	64 594 576
2	3	22	141 046 655
3	10	23	306 858 874
4	29	24	665 342 837
5	82	25	1 438 134 475
6	215	26	3 099 548 927
7	556	27	6 662 442 946
8	1 385	28	14 285 118 725
9	3 391	29	30 557 828 119
10	8 135	30	65 225 030 201
11	19 261	31	138 937 277 596
12	44 963	32	295 385 810 819
13	103 906	33	626 867 939 224
14	237 719	34	1 328 075 901 017
15	539 458	35	2 809 126 944 436
16	1 214 993	36	5 932 793 909 801
17	2 718 760	37	12 511 847 996 740
18	6 047 426	38	26 350 575 690 893
19	13 380 766	39	55 423 630 773 538
20	29 463 632	40	116 429 658 505 697

Table 3: The number of binary words $0x$ of length n with $A_N(0x) = q$, for sufficiently large n . The value for $q = 7$ is surprisingly small when comparing with Table 2.

References

1. Bender, E.A., Goldman, J.R.: On the applications of Möbius inversion in combinatorial analysis. *Amer. Math. Monthly* **82**(8), 789–803 (1975). <https://doi.org/10.2307/2319793>
2. Choi, J.S.: Counts of unique periodic binary strings of length n. <http://oeis.org/A152061> (Sep 2011)
3. Hyde, K.K., Kjos-Hanssen, B.: Nondeterministic automatic complexity of overlap-free and almost square-free words. *Electron. J. Combin.* **22**(3) (2015), paper 3.22, 18
4. Kjos-Hanssen, B.: Complexity lookup. <http://math.hawaii.edu/wordpress/bjoern/complexity-of-0110100110010110/>
5. Kjos-Hanssen, B.: Complexity option game. <http://math.hawaii.edu/wordpress/bjoern/complexity-option-game/>
6. Lyndon, R.C.: On Burnside's problem. *Trans. Amer. Math. Soc.* **77**, 202–215 (1954). <https://doi.org/10.2307/1990868>, <https://doi.org/10.2307/1990868>
7. Shallit, J., Wang, M.W.: Automatic complexity of strings. *J. Autom. Lang. Comb.* **6**(4), 537–554 (2001), 2nd Workshop on Descriptive Complexity of Automata, Grammars and Related Structures (London, ON, 2000)
8. Sipser, M.: A complexity theoretic approach to randomness. In: Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing. pp. 330–335. STOC '83, ACM, New York, NY, USA (1983). <https://doi.org/10.1145/800061.808762>, <http://doi.acm.org/10.1145/800061.808762>
9. Witt, E.: Treue Darstellung Liescher Ringe. *J. Reine Angew. Math.* **177**, 152–160 (1937). <https://doi.org/10.1515/crll.1937.177.152>, <https://doi.org/10.1515/crll.1937.177.152>

```

oeisValues = [
    0, 0, 2, 2, 4, 2, 10, 2, 16, 8, 34, 2, 76, 2, 130, 38, 256, 2,
    568, 2, 1036, 134, 2050, 2, 4336, 32, 8194, 512, 16396, 2, 33814,
    2, 65536, 2054, 131074, 158, 266176, 2, 524290, 8198, 1048816, 2,
    2113462, 2, 4194316, 33272, 8388610, 2, 16842496, 128, 33555424
]# from \protect\vrule width0pt\protect\ href{http://oeis.org/A152061}{http://oeis.org/}
def Z(n): # number of periodic binary strings of length n
    return oeisValues[n]
def plus(k):
    if k<0:
        return 0
    return k
def limS(q): #limitingNumberOfStringsWithNFAComplexity(q):
    num = 0
    print "."
    for i in range(0, q):
        for l in range(0, q):
            if i+l<q:
                left = 2**plus(i-1))
                right = 2**plus(l-1))
                middle = (2***(q-(i+l))-Z(q-(i+l)))
                num += left * middle * right
    return num
def answer(q):
    bound = 2***(q-2)*(1+q*(q+5)/2)
    print "q=" + str(q) + ", " + str(limS(q)),
    print ", bound=" + str(bound) + ", " +
    print str(limS(q)/float(bound)))
for q in range(3, len(oeisValues)):
    answer(q)

```

Fig. 6: Python code which when run hints at the sharpness of Theorem 13.