Remarks on Inferring Integer Sequences

Jeffrey Shallit
Department of Computer Science
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada
shallit@graceland.uwaterloo.ca

The slides for this talk can be found on my home page:
http://math.uwaterloo.ca/~shallit/
Introduction

What are the rules behind the following integer sequences?

- 1, 2, 3, 4, 5, 6, 7, 8, ...
- 2, 5, 10, 17, 26, 37, 50, 65, ...
- 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, ...
- 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, ...
- 0, 1, 1, 2, 1, 2, 2, 3, 1, 2, 2, 3, 2, 3, 4, 5, ...
- 0, 3, 5, 6, 9, 10, 12, 15, 17, 18, 20, 23, 24, 27, 29, 30, ...

Given an integer sequence, how can we determine what it is?

The question is ill-posed: can only look at a finite number of terms, and any such sequence has an infinite number of potential extensions.
Two Approaches to Sequence Recognition

• One can mathematically define a large class of sequences and then try to determine membership in that class
  – A. K. Dewdney (Scientific American, Mathematical Recreations, March 1986)
  – Bhansali and Skiena (Computational Support for Discrete Mathematics, 1994)
  – Sloane and Plouffe’s SuperSeeker program (superseeker@research.att.com)

• One can collect sequences from the literature and then try to express the target sequence in terms of known sequences
  – Peter Liu (Master’s Essay, University of Waterloo, 1994)
  – Sloane and Plouffe’s SuperSeeker program
Two Neglected Classes of Sequences

• The $k$-automatic sequences
  – form about 3% of the sequences in the Sloane-Plouffe table

• The $k$-regular sequences
  – form about 7% of the sequences in the Sloane-Plouffe table
Basics of Finite Automata

• If $\Sigma$ is a finite set of symbols, then by $\Sigma^*$ we mean the free monoid over $\Sigma$ (set of all finite strings of symbols chosen from $\Sigma$);

• A language is a subset of $\Sigma^*$.

• A finite automaton is a simple model of a computer

• formally it is a quintuple: $M = (Q, \Sigma, \delta, q_0, F)$ where:
  
  - $Q$ is a finite set of states;
  - $\Sigma$ is a finite set of symbols, called the input alphabet;
  - $q_0 \in Q$ is the start state;
  - $F \subseteq Q$ is the set of final states;
  - $\delta : Q \times \Sigma \rightarrow Q$ is the transition function

• The language accepted by $M$ is denoted by $L(M)$ and is given by $\{ w \in \Sigma^* \mid \delta(q_0, w) \in F \}$.
Example of a Finite Automaton
Automata as Computers of Sequences

• First, we can generalize our notion of automaton to provide an output, not simply accept/reject.

• Formally, we define a deterministic finite automaton with output (DFAO) as a sextuple: \((Q, \Sigma, \delta, q_0, \Delta, \tau)\), where \(\Delta\) is the finite output alphabet and \(\tau: Q \rightarrow \Delta\) is the output mapping.

• Next, we decide on an integer base \(k \geq 2\) and represent \(n\) as a string of symbols over the alphabet \(\Sigma = \{0, 1, 2, \ldots, k - 1\}\).

• To compute \(f_n\), given an automaton \(M\), express \(n\) in base-\(k\), say, \(a_r a_{r-1} \cdots a_1 a_0\), and compute \(f_n = \tau(\delta(q_0, a_0 a_1 \cdots a_{r-1} a_r))\).

• Any sequence that can be computed in this way is said to be \(k\)-automatic.
$k$-Automatic Sequences

A sequence $(a_n)_{n \geq 0}$ is said to be $k$-automatic if, $a_n$ is a finite-state ("automatic") function of the base-$k$ representation of $n$.

Example. The following automaton generates the Rudin-Shapiro sequence:

To compute $r_n$, expand $n$ in base-2, and then input the bits of $n$ into the automaton, starting with the least significant bit, transiting from state to state. When last state is encountered, output is specified in the state.
The Thue-Morse Sequence

- Introduced by Axel Thue (1863–1922).
- \( t_n = \text{sum of bits of } n \text{ (base 2), taken modulo 2} \).
- First few terms: 0 1 1 0 1 0 0 1 1 0 0 1 0 \cdots

Example 1. An unusual infinite product. Define \( a_n = (-1)^{t_n} \) for \( n \geq 0 \). Then
\[
\prod_{n \geq 0} \frac{(2n + 1)^{a_n}}{2n + 2} = \frac{1}{2} \cdot \frac{4}{3} \cdot \frac{6}{5} \cdot \frac{7}{8} \cdots = \frac{\sqrt{2}}{2}.
\]

Example 2. A converse of sorts to Example 1. Define \( b_0 = 1 \), and
\[
b_n = \begin{cases} 
1, & \text{if } \prod_{0 \leq i < n} \left( \frac{2i+1}{2i+2} \right)^{b_i} > \sqrt{2}/2; \\
-1, & \text{if } \prod_{0 \leq i < n} \left( \frac{2i+1}{2i+2} \right)^{b_i} < \sqrt{2}/2.
\end{cases}
\]
Then \( a_n = b_n \).
The Thue-Morse sequence \((u_n)\) continued

**Example 3.** Prouhet’s result of 1851 on “multigrades”. Separate the integers in the set

\[
S_n = \{0, 1, 2, \ldots, 2^n - 1\}
\]

into two subsets:

\[
T_n = \{i \in S_n : t_i = 0\}
\]

and

\[
U_n = \{i \in S_n : t_i = 1\}.
\]

Then

\[
\sum_{k \in T_n} k^j = \sum_{\ell \in U_n} \ell^j
\]

for \(j = 0, 1, \ldots, n - 1\).

**Example:**

\[
0^i + 3^i + 5^i + 6^i + 9^i + 10^i + 12^i + 15^i =
\]

\[
1^i + 2^i + 4^i + 7^i + 8^i + 11^i + 13^i + 14^i
\]

for \(i = 0, 1, 2, 3\).
The Rudin-Shapiro Sequence \((u_n)\)

- Define \(u_n = (-1)^{r_n}\), where \(r_n\) counts the number of (possibly overlapping) occurrences of the block ‘11’ in the binary representation of \(n\).
- This sequence was introduced by Rudin and Shapiro, independently.

**Example 1.** It is easy to prove that, for any sequence \((a_n)_{n\geq 0}\) of +1’s and -1’s, we have

\[
\sup_{\theta} \left| \sum_{0 \leq k < n} a_k e^{ik\theta} \right| \geq \sqrt{n}.
\]

On the other hand, it can be shown that for “almost all” sequences \((a_n)_{n\geq 0}\) we have

\[
\sup_{\theta} \left| \sum_{0 \leq k < n} a_k e^{ik\theta} \right| = O(\sqrt{n \log n}).
\]

Rudin and Shapiro (independently) proved in the 1950’s that

\[
\sup_{\theta} \left| \sum_{0 \leq k < n} u_k e^{ik\theta} \right| = O(\sqrt{n}).
\]
The Rudin-Shapiro Sequence, Continued

Example 2. Consider a path visiting lattice points in the plane. Start at the origin and make a first move to $(0, 1)$. At step $n$, turn “left” or “right” $90^\circ$ according to the following rule:

- “left”, if $r(n) - r(n - 1) + n \equiv 0 \pmod{2}$;
- “right”, if $r(n) - r(n - 1) + n \equiv 1 \pmod{2}$.

We get a spacefilling curve that visits every lattice point in $1/8$ of the plane exactly once.
Robustness of the Notion of Automatic Sequence

- the order in which the base-$k$ digits are fed into the automaton in does not matter (provided it is fixed for all $n$);

- other representations also work (such as expansion in base-$(-k)$);

- automatic sequences are closed under many operations, such as shift, periodic deletion, $q$-block compression, and $q$-block substitution.

- automatic sequences are also closed under uniform transduction.
  - a uniform finite-state transducer is like an automaton, but outputs $s$ symbols at each transition
Properties of Automatic Sequences

Definition.  
The $k$-kernel of a sequence $(a_n)_{n \geq 0}$ is the set of subsequences

$$\{(a_{kn+c})_{n \geq 0} : r \geq 0, 0 \leq c < k^r\}.$$  

Cobham’s 1st Theorem.  A sequence is $k$-automatic if and only if its $k$-kernel is finite.

Definition.  
A homomorphism $\varphi : \Sigma^* \to \Sigma^*$ is a map satisfying $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in \Sigma^*$. If $|\varphi(a)| = k$ for all $a \in \Sigma$, then we say $\varphi$ is $k$-uniform. A coding is a 1-uniform homomorphism.

Cobham’s 2nd Theorem.  A sequence is $k$-automatic if and only if it is the image (under a coding) of a fixed point of a $k$-uniform homomorphism.

Example.  The Thue-Morse sequence is the fixed point of the map $0 \to 01$, $1 \to 10$ that starts with 0.
The Theorem of
Christol-Kamae-Mendès France-Rauzy

Theorem. (Christol, Kamae, Mendès France, Rauzy, 1980). Let \((u_n)_{n \geq 0}\) be a sequence over
\[
\Sigma = \{0, 1, \ldots, p - 1\},
\]
where \(p\) is a prime. Then the formal power series \(U(X) = \sum_{n \geq 0} u_n X^n\) is algebraic over \(GF(p)[X]\) if and only if \((u_n)_{n \geq 0}\) is \(p\)-automatic.

Example.

Let, as before, \((t_n)_{n \geq 0}\) denote the Thue-Morse sequence, i.e., \(t_n = \) sum of the bits in the binary expansion of \(n\), mod 2. Then \(t_{2n} \equiv t_n\) and \(t_{2n+1} \equiv t_n + 1\). If we set \(A(X) = \sum_{n \geq 0} t_n X^n\), then

\[
A(X) = \sum_{n \geq 0} t_{2n} X^{2n} + \sum_{n \geq 0} t_{2n+1} X^{2n+1}
= \sum_{n \geq 0} t_n X^{2n} + X \sum_{n \geq 0} t_n X^{2n} + X \sum_{n \geq 0} X^{2n}
= A(X^2) + XA(X^2) + X/(1 - X^2)
= A(X)^2(1 + X) + X/(1 + X)^2.
\]

Hence \((1 + X)^3 A^2 + (1 + X)^2 A + X = 0\).
Inferring Automatic Sequences

• Can one infer a $k$-automatic sequence, given the first few terms?

• If a sequence is $k$-automatic, and is generated by an automaton with $\leq r$ states, then given the first $k^{2r-2}$ terms, one can correctly and efficiently predict all future terms of the sequence.

• In practice $k$ and $r$ are usually small, and the correct automaton can often be guessed with far fewer terms.

• The automaton can be inferred purely mechanically, by examining the $k$-kernel, and declaring two members to be equal if they agree on the terms actually known.

• If a sequence is not $k$-automatic, then it is possible to have two genuinely different elements of the $k$-kernel agree on thousands or millions of terms before a distinguishing element is found.

• However, this rarely occurs in practice.
An Amazing Non-Automatic Sequence

Take the Thue-Morse sequence

$$(t_n)_{n \geq 0} = 011010011001 \ldots,$$

and create a new sequence

$$(u_n)_{n \geq 0} = 12112221121 \ldots$$

that counts the lengths of blocks of identical symbols in $$(t_n)_{n \geq 0}.$$ Then it can be shown that $$(u_n)$$ is not a 2-automatic sequence, (but the proof is not easy at all).
An Amazing Non-Automatic Sequence

However, the sequence \((u_n)\) comes very “close” to being 2-automatic, in that to distinguish two sequences in the kernel, one must look at very large values of \(n\). For example, \(u_{8n} = u_{32n}\) for \(0 \leq n \leq 14562\), but not for \(n = 14563\). Similarly, \(u_{16n+1} = u_{64n+1}\) for \(0 \leq n \leq 1864134\), but not for \(n = 1864135\).

A complete understanding of the behaviour of this sequence is still not at hand, but it depends on the fact that the sequence is the fixed point of the map \(1 \to 121; 2 \to 12221\), and the associated matrix of the map is

\[
\begin{bmatrix}
2 & 1 \\
2 & 3
\end{bmatrix}
\]

whose characteristic polynomial is \((X - 1)(X - 4)\).
Automaticity

- One can study how “close” a non-automatic sequence comes to being automatic.
- To do this, compute \((a_i)_{0 \leq i \leq n}\) and then form the \(k\)-kernel.
- Then \((a_i)\) is known to \(n + 1\) terms, \((a_{2i})\) to \([n/2] + 1\) terms, etc. Call two elements of the (partially-computed) \(k\)-kernel the same if they coincide on the terms on which they are known. The size of the \(k\)-kernel, as a function of \(n\), is called the “automaticity” of the sequence \((a_n)\).

**Theorem.** A sequence has automaticity \(O(1)\) if and only if it is automatic.

**Theorem.** If a sequence is not automatic, then its automaticity is \(\Omega_\infty(\log n)\).
Automaticity (continued)

**Question.** Is there a homomorphism whose fixed point is quasi-automatic, but not automatic?

**Answer.** Yes, the homomorphism that sends $c \rightarrow cba$; $a \rightarrow aa$; and $b \rightarrow b$ has a fixed point

$$cbabaabaaaabaaaaaaab\cdots$$

in which the $b$’s are in positions $2^n + n$ for $n \geq 0$. This is not a 2-automatic sequence, but it is 2-quasiautomatic. An automaton with $\leq 6 \log_2 n$ states suffices to compute the sequence correctly to $n$ terms.

**Open Question.** Is the fixed point of the homomorphism $1 \rightarrow 121$; $2 \rightarrow 12221$ quasi-automatic?
Automaticity (continued)

• Let $0 < \alpha < 1$ be a real irrational number with bounded partial quotients in its continued fraction expansion.

• Then it can be shown (JOS, 1995) that the automaticity of the Sturmian sequence $(s_n)_{n \geq 1}$ defined by
  
  $$s_n = \lfloor (n + 1)\alpha \rfloor - \lfloor n\alpha \rfloor$$

  is $\Omega(n^{1/5})$.

• The proof uses basic techniques of Diophantine approximation.

• In particular, can show that for any integer $r \geq 2$, and all pairs $(c, d)$ with $c \neq d$ and $0 \leq c, d < r$, there exists an $n = O(r^3)$ such that $s_{rn+c} \neq s_{rn+d}$.

• Open Question: is the $O(r^3)$ bound best possible?
Generalization of Automatic Sequences

- Automatic sequences must take their values in a finite set.
- This is too restrictive; we would like to define “automatic sequences” over the integers.
- Need the correct definition to generalize.
- Recall the $k$-kernel of a sequence:
  \[
  K_k(a) = \{(a_{k^n i + j})_{n \geq 0} : i \geq 0, \ 0 \leq j < k^i\}.
  \]
- What is the proper generalization of the finiteness property?
\textbf{$k$-regular Sequences}

- An integer sequence $(a_n)_{n \geq 0}$ is said to be \textit{$k$-regular} if the $\mathbb{Z}/k\mathbb{Z}$-module generated by the sequences in the $k$-kernel is \textit{finitely generated}.

- Example: $a_n = s_2(n)$, the total number of 1’s in the binary expansion of $n$.

- Then $a_{2n} = a_n$ and $a_{2n+1} = a_n + 1$. It follows that $\langle K_2(a) \rangle$ is generated by $(a_n)_{n \geq 0}$ and the constant sequence 1.

- $k$-regular sequences appear in many different fields of mathematics: numerical analysis, topology, number theory, combinatorics, analysis of algorithms, and fractal geometry.
Examples of $k$-regular Sequences

Example 1. The Stern-Brocot Tree

In the limit, the sequence $(s(n))_{n \geq 0}$ of numerators one gets at level $n$ is

$$1, 2, 3, 3, 4, 5, 5, 4, 5, 7, 8, 7, 8, 7, 5, \ldots$$

which satisfies the relations $s(2n + 1) = 3a(n) - a(2n)$; $s(4n) = 2a(2n) - a(n)$; $s(4n + 2) = 4a(n) - a(2n)$. 
Examples of $k$-regular Sequences

Example 2. Minimum cost of addition chains. An addition chain to $n$ is a sequence of pairs of positive integers

$$(a_1, b_1), (a_2, b_2), (a_3, b_3), \ldots, (a_r, b_r)$$

where

(i) $a_r + b_r = n$;

(ii) for all $s$, either $a_s = 1$, or $a_s = a_i + b_i$ for some $i < s$, and the same requirement holds for $b_s$.

Example: here is an addition chain to 21:

$$(1, 1), (2, 2), (4, 1), (5, 5), (10, 10), (20, 1)$$

- The cost of the addition chain is $\Sigma_{1 \leq i \leq r} a_i b_i$.
- Denote the cost of the minimum addition chain to $n$ as $c(n)$.
- Graham, Yao, and Yao showed that $c(2n) = c(n) + n^2$ and $c(2n + 1) = c(n) + n(n + 2)$ for $n \geq 1$.
- It follows that $(c(n))_{n \geq 0}$ is 2-regular.
Examples of $k$-regular Sequences

Example 3. Subword Complexity. Let $w = w_0w_1w_2 \ldots$ be an infinite word over a finite alphabet, and let $\rho_w(n)$ be the number of distinct subwords of length $n$ in $w$. Then $\rho_w(n)$ is frequently $k$-regular, especially when $w$ is the fixed point of a $k$-uniform homomorphism. For example, when $w$ is the Thue-Morse word $01101001 \ldots$, then $\rho_w(n)$ is 2-regular.

Example 4. Mergesort. To sort a list of $n$ integers recursively, first sort the left half (recursively), then sort the right half, and then merge the two halves together. Then $T(n)$, the total number of comparisons used in the worst case, is given by the recurrence

$$T(n) = T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + n - 1.$$  

It follows that $T(n)$ is 2-regular, and one can obtain the closed form

$$T(n) = n \lfloor \log_2 n \rfloor - 2^{\lceil \log_2 n \rceil} + 1.$$  

26
Properties of $k$-regular Sequences

- Every $k$-automatic sequence is also $k$-regular.
- If a $k$-regular sequence is bounded, then it is $k$-automatic.
- The $k$-regular sequences are closed under shift, and periodic deletion.
- A sequence is $k$-regular iff it is $k^r$-regular for any $r \geq 2$.
- The $k$-regular sequences are closed under (termwise) sum and product.
- If $f(X) = \sum_{n \geq 0} f_n X^n$ and $g(X) = \sum_{n \geq 0} g_n X^n$ are formal power series with $k$-regular coefficients, then so is $f(X)g(X)$.

Conjecture: if $(f_i)_{i \geq 0}$ and $(g_i)_{i \geq 0}$ are both $k$-regular sequences, and $f_i/g_i \in \mathbb{Z}$ for all $i \geq 0$, then $(f_i/g_i)_{i \geq 0}$ is also $k$-regular.

Open Question. Show that $\left\lfloor \frac{1}{2} + \log_2 n \right\rfloor$ is not a 2-regular sequence.
The Pattern Transform

• Let \( e_P(n) \) denote the number of (possibly overlapping) occurrences of the pattern \( P \) in the base-2 expansion of \( n \). Then \( e_P(n) \) is 2-regular. Furthermore, every sequence \((f_n)_{n \geq 0}\) can be expanded as a sum of such pattern sequences, and the coefficients in this sum are 2-regular if and only if \((f_n)_{n \geq 0}\) is 2-regular.

• Example:

\[
e_1(3n) = 2e_1 - 2e_{11}(n) + e_{111}(n) - 2e_{1011}(n) + \cdots
\]

\[
= 2e_1(n) - 2 \sum_{i \geq 0} e_{(10)i11}(n) + \sum_{i \geq 0} e_{11(01)i1}(n).
\]

• It had previously been observed by Newman that the first few values of \((e_1(3n))_{n \geq 0}\) are almost all even.
Inferring $k$-regular sequences

- given a sequence $(s_n)_{n \geq 0}$, how can we determine if it is $k$-regular?

- construct a matrix in which the rows are elements of the $k$-kernel, and attempt to do row reduction

- as elements further out in the $k$-kernel are examined, the number of columns of the matrix that are known in all entries decreases

- if rows that are previously linearly independent suddenly become dependent with the elimination of terms further out in the sequence, then no relation can be accurately deduced; stop and retry after computing more terms

- if the subsequence $(s_{kj(n+c)})_{n \geq 0}$ is not linearly dependent on the previous sequences, try adding the subsequences $(s_{kj((kn+a)+c)})_{n \geq 0}$ for $0 \leq a < k$

- when no more linearly independent sequences can be found, you have found relations for the sequence

29
Inferring $k$-regular Sequences

- (N. Strauss, 1988) Define
  \[ r(n) = \sum_{0 \leq i < n} \binom{2i}{i}, \]
  \[ let \nu_3(n) be the exponent of the highest power of 3 that divides n. \]

- The first few terms of $\nu_3(r(n))$ are:
  \[ 0, 1, 2, 0, 2, 3, 1, 2, 4, 0, 1, 2, 0, 3, 4, 2, 3, 5, 1, 2, \ldots \]

- A 3-regular sequence recognizer easily produces the following conjectured relations (where $f(n) = \nu_3(r(n + 1))$):
  \[ f(3n + 2) = f(n) + 2; \]
  \[ f(9n) = f(9n + 3) = f(3n); \]
  \[ f(9n + 1) = f(9n + 4) = f(9n + 7) = f(3n) + 1. \]

- With a little more work, one arrives at the conjecture
  \[ \nu_3(r(n)) = \nu_3\left(\binom{2n}{n}\right) + 2\nu_3(n). \]

- proved by Allouche and JOS.
• A beautiful proof of this identity using 3-adic analysis was also given by Don Zagier.

• Zagier showed that if we set

\[ F(n) = \frac{\sum_{0 \leq k \leq n-1} \binom{2k}{k}}{n^2 \binom{2n}{n}}, \]

then \( F(n) \) extends to a 3-adic analytic function from \( \mathbb{Z}_3 \) to \(-1 + 3\mathbb{Z}_3\), and has the expansion:

\[ F(-n) = -\frac{(2n - 1)!}{(n!)^2} \sum_{0 \leq k \leq n-1} \frac{(k!)^2}{(k - 1)!}. \]
The Automatic Real Numbers

• We say that a real number $r$ is $(k, b)$-automatic if the base-$b$ representation of its fractional part is a $k$-automatic sequence.

• For example, the number

$$0.110100010000001000000000000000001\cdots_{(b)}$$

with 1’s in the 1st, 2nd, 4th, 8th, etc., positions (sometimes called the Fredholm number, although Fredholm never studied it!) is $(2, b)$-automatic.

• The set of all $(k, b)$-automatic numbers is denoted by $L(k, b)$. 
The Automatic Reals form a Vector Space over $\mathbb{Q}$

- The proof is rather technical.
- It suffices to show that if $x, y$ are in $L(k, b)$, then so are $x/n$ (for any integer $n$) and $x + y$.
- For the first, we can express division by $n$ as a finite-state transducer that just mimics long-division as done by hand. For example, here is a transducer for division by 3 for numbers written in base-2:

  \[
  \begin{align*}
  &\text{For the first, we can express division by } n \text{ as a finite-state transducer that just mimics long-division as done by hand. For example, here is a transducer for division by 3 for numbers written in base-2:} \\
  &\text{Now a theorem of Cobham implies that } x/n \in L(k, b). \\
  &\text{For } x + y, \text{ a more complicated proof is necessary, since carries can come from arbitrarily far to the right.} \\
  &\text{More generally, we have a Normalization Lemma: if } (a_i)_{i \geq 1} \text{ is a } k\text{-automatic sequence taking values in } \mathbb{Z}, \\
  &\text{then } \Sigma_{i \geq 1} a_i b^{-i} \in L(k, b). \\
  \end{align*}
\]
What is the Dimension of $L(k, b)$ over $\mathbb{Q}$?

- Now that we know $L(k, b)$ is a vector space over $\mathbb{Q}$, a natural question is, what is the dimension of that vector space?
- A simple argument shows that it is infinite:
- For example, define
  \[ f(X) = X + X^2 + X^4 + X^8 + X^{16} + \cdots. \]
  Then clearly $f(1/b^r) \in L(2, b)$ for all odd integers $r \geq 1$.
- But the numbers
  \[ \{f(1/b^r) : r \text{ odd, } \geq 1\} \]
  are linearly independent over $\mathbb{Q}$.
- For if not, then we would have
  \[ \sum_{0 \leq i \leq s} d_i f(1/b^{2i+1}) = \sum_{0 \leq i \leq s} e_i f(1/b^{2i+1}) \]
  with $0 \leq d_i, e_i \leq M$ and $d_i e_i = 0$ for $0 \leq i \leq s$.
- Now for $n$ sufficiently large, the base-$b$ digits to the left of position $(2i + 1)2^n$ on the left-hand side are $(d_i)_b$, while those in the same position on the right-hand side are $(e_i)_b$. It follows that $d_i = e_i = 0$. 

34
Automatic Reals are Not Closed Under Product

Theorem (Lehr, Shallit, and Tromp, 1994). The automatic reals are not closed under product.

Proof. We showed that

\[ f = \sum_{r \geq 0} 2^{-2^r} \]

and

\[ g = \sum_{m \geq 1, n \geq 0} 2^{-(2^m-1)2^n} \]

are both in \( L(2, 2) \), but their product is not.
Is There an Automatic Ring Strictly Containing $\mathbb{Q}$?

- Let $f(X) = X + X^2 + X^4 + X^8 + \cdots$.
- Then it can be shown that $y = f(1/b)$ is transcendental over $\mathbb{Q}$.
- Hence $\mathbb{Q}[y]$ is a ring with $[\mathbb{Q}[y] : \mathbb{Q}] = \infty$.
- If we could show all positive powers of $y$ are in $L(k, b)$, we’d be done.
- By previous results, it suffices to show that, for any fixed $r$, the coefficients of $f(X)^r$ are bounded.
- Let $W_r = \max_{n \geq 0} [X^n](f(X)^r)$. Consider the following Catalan tree:

```
1
   / \   \
1   2
  / \  / \  \
1 1 3
```

- Then Tromp has shown that $W_r$ is bounded by the sum, over all vertices $v$ at level $r - 1$, of the product of all vertex labels in the path from the root to $v$. It follows that $W_r \leq (2r)!/(2^r \cdot r!)$.
For Further Reading


