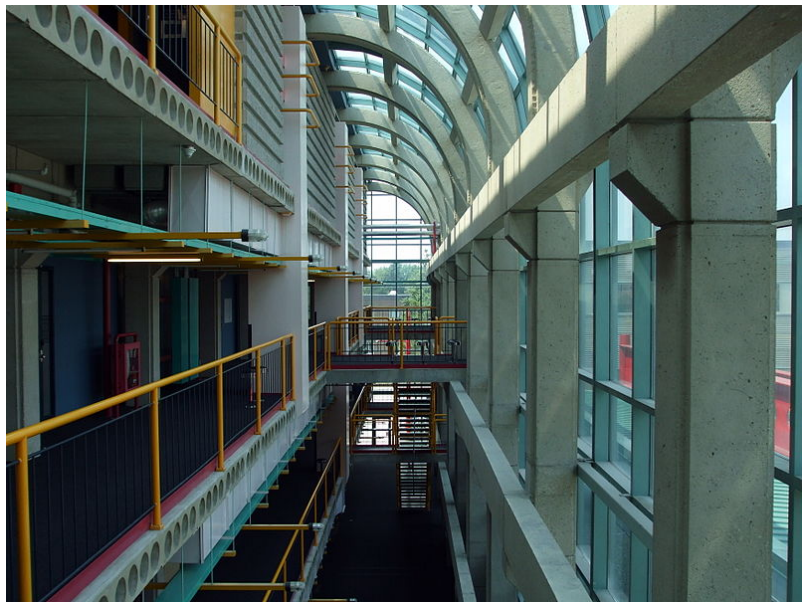
Doing Additive Number Theory with Logic and Automata

Jeffrey Shallit
(Joint work with Jean-Paul Allouche and Jason Bell)

School of Computer Science
University of Waterloo
Waterloo, ON N2L 3G1
Canada
shallit@uwaterloo.ca
https://cs.uwaterloo.ca/~shallit/



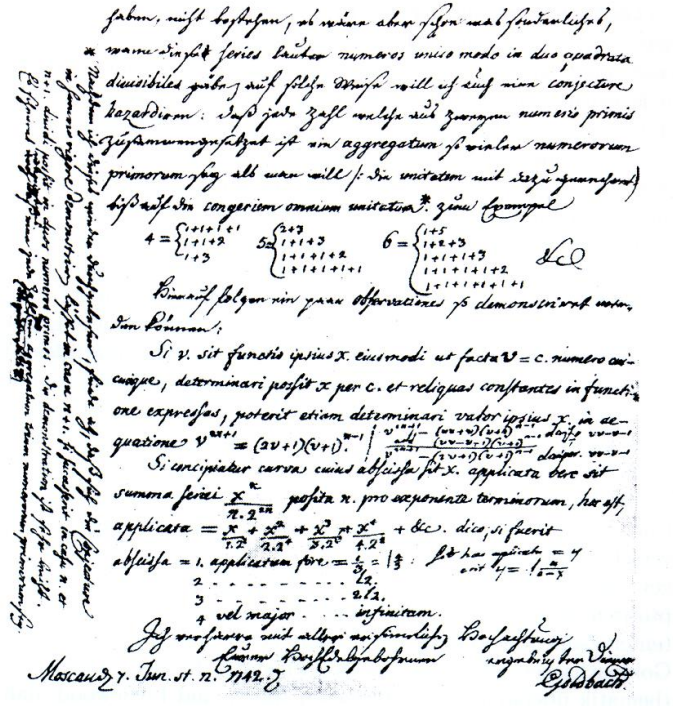
Davis Centre, U. Waterloo

Additive number theory

Additive number theory is the study of the additive properties of integers.

It poses simple-to-state questions that can be *very* hard to resolve.

Probably the most famous example is *Goldbach's conjecture* from 1742: every even number ≥ 4 is the sum of two primes.



Goldbach letter to Euler
June 7 1742

Additive number theory

Less well-known to the general public, but very well-known to additive number theorists, is the existence of an asymptotic formula that conjecturally predicts the *number* $G_2(n)$ of representations of n as the sum of two primes, due to Hardy and Littlewood in 1923:

$$G_2(n) \approx 2 \cdot \Pi_2 \cdot \left(\prod_{\substack{p \mid n \\ p \geq 3}} \frac{p-1}{p-2} \right) \frac{n}{(\log n)^2}$$

for n even, where

$$\Pi_2 = \prod_{p \geq 3} \left(1 - \frac{1}{(p-1)^2} \right) \doteq 0.66016$$

is the twin-prime constant.

G. H. Hardy

J. E. Littlewood

Additive number theory

So, given a set S, number theorists are interested in both

- *which* numbers are representable as sums of elements of S, and
- the *number* of such representations.

In this talk I focus on the second: the number of representations.

Let $A \subseteq \mathbb{N} = \{0, 1, 2, \ldots\}$ be a subset of the natural numbers. We define

$$r(k, A, n) := |\{(a_1, a_2, \ldots, a_k) \in A^k \ : \ n = a_1 + a_2 + \cdots + a_k\}|$$

$$r_<(k, A, n) := |\{(a_1, a_2, \ldots, a_k) \in A^k \ : \ n = a_1 + a_2 + \cdots + a_k, \ a_1 < a_2 < \cdots < a_k\}|$$

$$r_\leq(k, A, n) := |\{(a_1, a_2, \ldots, a_k) \in A^k \ : \ n = a_1 + a_2 + \cdots + a_k, \ a_1 \leq a_2 \leq \cdots \leq a_k\}|.$$

These functions were originally studied by Erdős, Turán, and co-authors starting in the 1940's.

Motivation for studying r: powers of power series

$$r(k, A, n) := |\{(a_1, a_2, \ldots, a_k) \in A^k \ : \ n = a_1 + a_2 + \cdots + a_k\}|$$

$r(k, A, n)$ has a nice interpretation in terms of the coefficients of a power series.

Given a set A, we can define its associated *characteristic sequence* $(a(n))_{n \geq 0}$ as follows:

$$a(n) = \begin{cases} 1, & \text{if } n \in A; \\ 0, & \text{otherwise.} \end{cases}$$

And we can define its associated *power series*:

$$A(X) = \sum_{n \geq 0} a(n) X^n.$$

Then $r(k, A, n)$ is just the coefficient of X^n in $A(X)^k$.

Example: Goldbach representations

Take $A = \{2, 3, 5, \ldots\}$ to be the prime numbers.

Then $A(X) = X^2 + X^3 + X^5 + \cdots$ and Goldbach's conjecture can be restated as *the coefficients of* X^{2n} *in*

$$A(X)^2 = X^4 + 2X^5 + X^6 + 2X^7 + 2X^8 + 2X^9 + 3X^{10} + 2X^{12} + \cdots$$

are all positive for $n \geq 2$.

Additive number theory is filled with seemingly simple statements that can be very hard to prove.

But there are also examples of long complicated proofs that were superseded by very simple arguments...

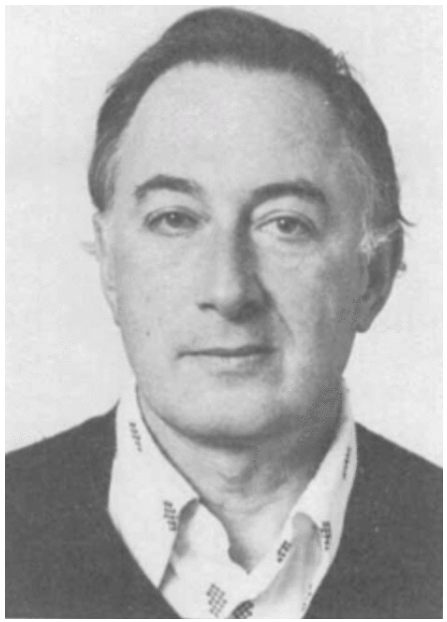
A result of Erdős and Turán

Theorem

Suppose $A = \{a_1, a_2, \ldots, \}$ is an infinite subset of $\mathbb{N}$. Then $(r(2, A, n))_{n \geq 0}$ cannot be eventually constant.

The original proof of Erdős and Turán used a big sledgehammer: the Fabry gap theorem.

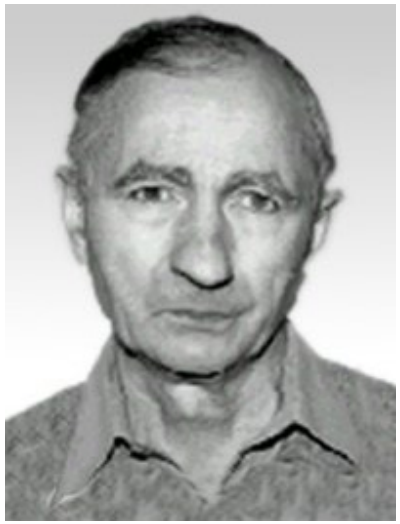
But G. A. Dirac (nephew of Eugene Wigner and stepson of the physicist Paul Dirac) observed this has a trivial proof: if $n = 2a_i$ then the number of representations of n must be odd, since $n = a_i + a_i$ and for all other representations, order matters, while if n is odd then the number of representations must be even.



G. A. Dirac

A result of Erdős, Sárközy, and Sós (1985)

Paul Erdős



András Sárközy

Vera Sós

Theorem

Let A be a subset of $\mathbb{N}$. If $(r(2, A, n))_{n \geq 0}$ is eventually increasing, then the complement set $\mathbb{N} \setminus A$ is finite.

Their proof was quite complicated (8 pages) and required case analysis.

But Balasubramanian found a much simpler 1-page proof in 1987.

Goal of the talk

My goal in this talk is to convince you that *tools from logic and automata theory can be used to prove interesting, non-trivial theorems in additive number theory, in relatively simple ways.*

This approach gives the additive number theorist new tools, and gives the specialist in automata theory applications for their theorems.

Automata

A deterministic finite automaton (DFA) is a very simple model of a computer.

It consists of a finite number of states.

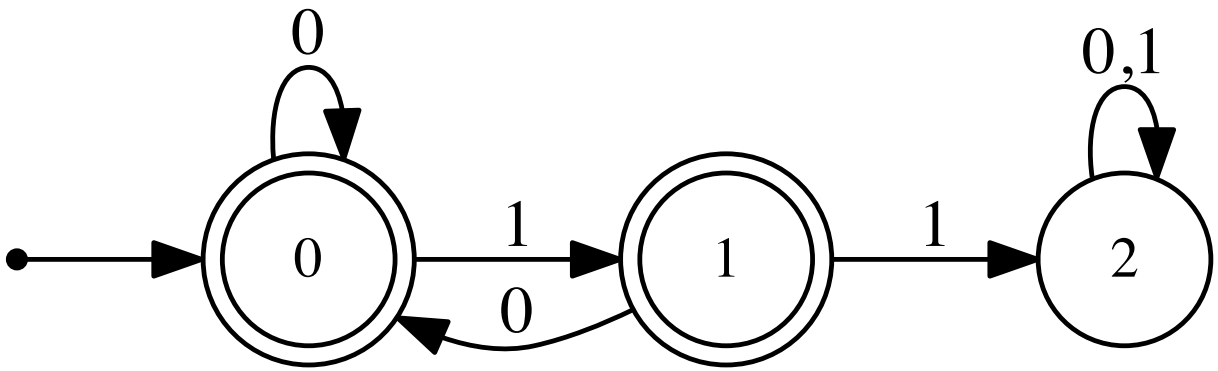
A finite automaton takes, as input, finite words (or strings) of symbols chosen from a finite alphabet Σ.

Each new symbol read causes a transition: a movement from one state to another, based on the current state and the symbol.

Some states are distinguished and called *accepting* or *final*. If, after reading the entire input, the automaton is in a final state, then the input is said to be *accepted*.

Example of an automaton

For example, here is an automaton that accepts binary strings having no two consecutive 1's:



The initial state is state 0.

The final states are state 0 and state 1.

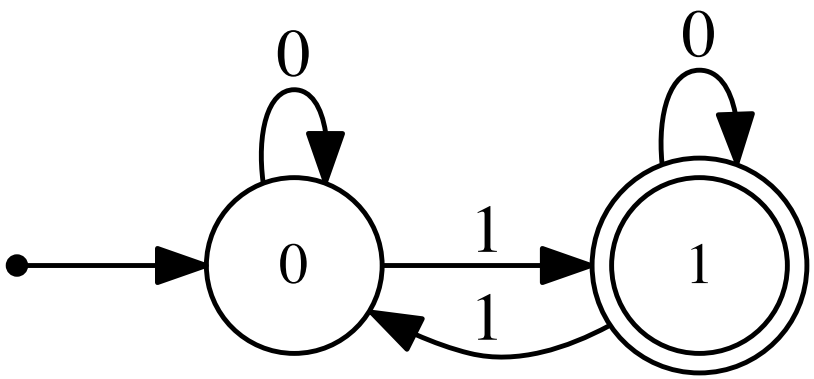
This automaton accepts 01001, but rejects 01101.

Automatic sets

A set A is said to be *b-automatic* if there is a finite automaton that recognizes exactly the set of base-b representations of members of A.

For example, consider the set $\mathcal{O}$ of *odious numbers*. These are the numbers having a base-2 representation with an odd number of 1's.

Then $\mathcal{O}$ is 2-automatic, and recognized by the following automaton.



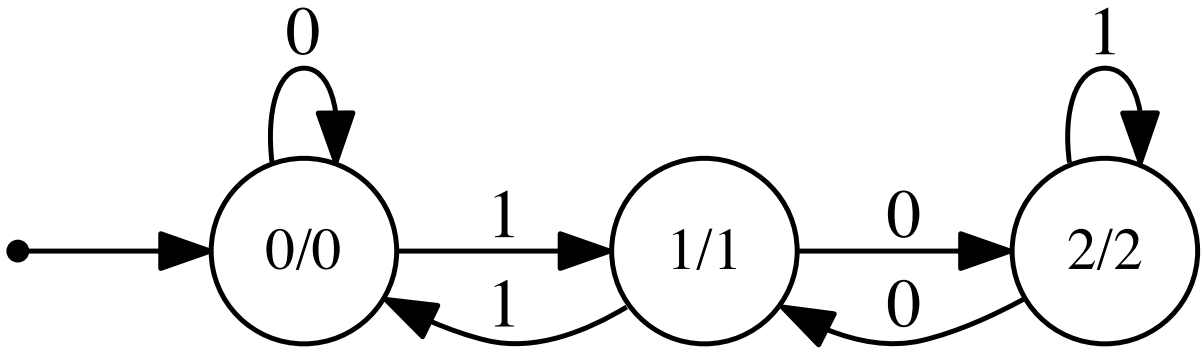
To use it, start in state 0, read the representation of n in base 2 and follow the arrows, accept iff you end up at state 1.

Automata with output

We can generalize the notion of automaton by adding an output associated with each state.

Now the output corresponding to an input is the output associated with the last state reached.

For example, here is an automaton that computes n mod 3, if the input represents n in base 2.



A sequence is said to be *b-automatic* if it is computed by an automaton with inputs represented in base b.

Walnut

Walnut is a free software tool, originally created by Hamoon Mousavi, that can prove or disprove assertions about automatic sequences.

One only has to state the claim in first-order logic, and then Walnut will prove or disprove it.

In some cases, its time and space usage can be extraordinary, so it's possible that some assertions can't be handled in practice.

Nevertheless, it has been used to prove results in 70 papers in the literature already.

Result of Lambek and Moser

Let $\mathcal{E} = \{0, 3, 5, 6, 9, 10, \ldots\}$ be the evil numbers (number of 1-bits in the binary representation of n is even) and $\mathcal{O} = \{1, 2, 4, 7, 8, 11, \ldots\}$ be the odious numbers (number of 1-bits is odd).

Lambek and Moser (1959) proved the following theorem:

$$r_<(2, \mathcal{E}, n) = r_<(2, \mathcal{O}, n)$$

for $n \geq 0$.

An example of the theorem: the representations of 9 as sums of $\mathcal{E}$ are $(0, 9)$ and $(3, 6)$. The representations as sums of $\mathcal{O}$ are $(1, 8)$ and $(2, 7)$.

This theorem was later proved again by Dombi (2002), Lev (2004), and others.

Joachim "Jim" Lambek

Leo Moser

Detour: linear representations

A *linear representation* for a sequence $(f(n))_{n\geq 0}$ is a triple (v, γ, w), where

- v is a t-element row vector;
- γ is a $t \times t$-matrix-valued morphism;
- w is a t-element column vector

and

$$f(n) = v\,\gamma(x)\,w$$

whenever x is the base-b representation of n.

Here $\gamma(x) = \gamma(a_1)\cdots\gamma(a_i)$ if $x = a_1\cdots a_i$.

The integer t is called the *rank* of the representation.

Example of a linear representation

Here is a linear representation for the Stern sequence $a(n)$, defined by $a(2n) = a(n)$ and $a(2n+1) = a(n) + a(n+1)$, with initial values $a(0) = 0$ and $a(1) = 1$:

$$v^T = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad \gamma(0) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}; \quad \gamma(1) = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix}; \quad w = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

For example, let's compute $a(27)$. Express 27 in base 2 as 11011. Then

$$a(27) = v\gamma(11011)w = v\gamma(1)\gamma(1)\gamma(0)\gamma(1)\gamma(1)w$$

$$= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} -5 & 8 \\ -7 & 11 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 8.$$

Computing linear representations

Theorem

Let $A \subseteq \mathbb{N}$ be a b-automatic set (i.e., an automaton recognizes representations of A in base b).

Then $r(k, A, n)$ (resp., $r_<(k, A, n)$; $r_\leq(k, A, n)$) has a linear representation that can be computed directly from the automaton for A.

Proof.

By a theorem of Büchi-Bruyère, it suffices to write first-order logical formulas for $r(k, A, n)$ (resp., $r_<(k, A, n)$; $r_\leq(k, A, n)$). But these are given by the definitions of these functions. □

J. Richard Büchi

Véronique Bruyère

Comparing linear representations

If we have a linear representation (v_f, γ_f, w_f) for $f(n)$ and a linear representation (v_g, γ_g, w_g) for $g(n)$, we can form a linear representation (v, γ, w) for the linear combination $\alpha f(n) + \beta g(n)$ by using block matrices, as follows:

$$v = [\alpha v_f \quad \beta v_g]$$

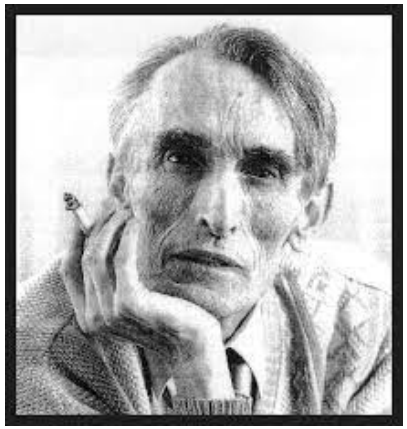
$$\gamma(a) = \begin{bmatrix} \gamma_f(a) & \mathbf{0} \\ \mathbf{0} & \gamma_g(a) \end{bmatrix}$$

$$w = \begin{bmatrix} w_f \\ w_g \end{bmatrix}.$$

Comparing linear representations

Furthermore, if we have a linear representation (v, γ, w) there is an algorithm, due to Schützenberger, for finding an equivalent linear representation of minimum rank.

Putting these two ideas together, we have the following theorem:



M.-P. Schützenberger

Theorem

Given a linear representation (v_f, γ_f, w_f) for $f(n)$ and a linear representation (v_g, γ_g, w_g) for $g(n)$, it is decidable if $f(n) = g(n)$ for all n.

Proof.

Form the linear representation for $f(n) - g(n)$, and then minimize it. Then $f(n) = g(n)$ for all n iff the linear representation is of rank 0 computing the 0 function. □

Lambek and Moser: proof via Walnut

To prove the Lambek-Moser result that

$$r_<(2, \mathcal{E}, n) = r_<(2, \mathcal{O}, n)$$

for $n \geq 0$, we just need to find a linear representation for both sides and then use the theorem on the previous slide.

This can be done using the Walnut software package as follows:

```
def evil_sum n "T[i]=@0 & T[j]=@0 & i<j & n=i+j":
def odious_sum n "T[i]=@1 & T[j]=@1 & i<j & n=i+j":
```

Here T[i] is Walnut's way of writing the Thue-Morse sequence, t_i, the parity of the number of 1-bits of i.

These create two linear representations of rank 8, and we can use the ideas above to demonstrate they compute the same function.

Another result for the evil and odious numbers

That was the result for $r_<$. How about r?

Theorem

We have

$$r(2, \mathcal{E}, n) - r(2, \mathcal{O}, n) = [n \text{ even}] \cdot (-1)^{t_n},$$

where $[P]$ *is Iverson notation, evaluating to* 1 *if* P *is true and* 0 *otherwise.*

Proof.

(Sketch.) Form linear representations for both sides. For the right side, use the fact that $(-1)^i = 1 - 2i$ for $i \in \{0, 1\}$. □

Chen and Wang result

Define the analogues of the evil and odious numbers, where we consider the parity of the number of 0-bits in the binary representation of n, instead of the number of 1-bits:

$$\mathcal{E}' := \{1, 3, 4, 7, 9, 10, 12, \ldots\}$$
$$\mathcal{O}' := \{0, 2, 5, 6, 8, 11, 13, \ldots\}$$

Chen and Wang (2003) proved

$$r_{\leq}(2, \mathcal{E}', n) = r_{\leq}(2, \mathcal{O}', n)$$

for $n \geq 1$.

Also proved later by Lev (2004).

Chen and Wang result

Shifting indices in the equation on the previous page gives

$$r_{\leq}(2, \mathcal{E}', n+1) = r_{\leq}(2, \mathcal{O}', n+1)$$

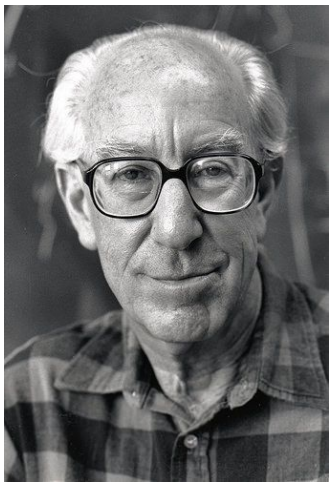
for $n \geq 0$. We can prove this as follows:

```
eval evil2_sum n "TT[i]=@0 & TT[j]=@0 & i<=j & n+1=i+j":
eval odious2_sum n "TT[i]=@1 & TT[j]=@1 & i<=j & n+1=i+j":
```

Here `TT[i]` is `Walnut`'s way of representing the twisted Thue-Morse sequence, counting the parity of the number of 0's in the base-2 representation of i.

These commands compute linear representations (of rank 20). We can then use the theorem above to show that they represent the same function.

The Rudin-Shapiro set



Walter Rudin

Harold S. Shapiro

Let $\mathcal{R} = \{3, 6, 11, 12, 13, 15, \ldots\}$ be the Rudin-Shapiro set: the numbers n where the number of 11's (possibly overlapping) in the binary expansion of n is odd.

Dombi (2002) proved that for $k \geq 5$, the function $r(k, \mathcal{R}, n)$ is an eventually increasing function of n.

He conjectured this is also true for $k = 4$, but still no proof is known.

The Rudin-Shapiro set

We can prove that $r(3, \mathcal{R}, n)$ is *not* eventually increasing as follows.

The first step is to create a linear representation for the difference sequence

$$d(n) := r(3, \mathcal{R}, n) - r(3, \mathcal{R}, n-1).$$

We can do that with the following `Walnut` code:

```
def rudin3 n "RS[i]=@1 & RS[j]=@1 & RS[k]=@1 & n=i+j+k":
def rudin3m1 n "RS[i]=@1 & RS[j]=@1 & RS[k]=@1 & n=i+j+k+1":
```

and then combine them with the block matrix trick to get a linear representation (v, γ, w) for $d(n)$.

The goal is to find infinitely many n such that $d(n) < 0$.

Closed forms for linear representations along subsequences

In general a function $f(n)$ given by a linear representation (v, γ, w) will not have a simply-describable behavior.

However, we can always obtain a formula for f evaluated at a *subsequence* $(n_i)_i$ for which the base-b representation is of the form

$$x \overbrace{yy \cdots y}^{i \text{ copies}} z$$

where x, y, z are strings of digits.

This is because

$$v\,\gamma(n_i)\,w = v\,\gamma(x)\,\gamma(y)^i\,\gamma(z)\,w,$$

and each entry of $\gamma(y)^i$ can be expressed as a linear combination of the i'th powers of the zeros of the minimal polynomial of $\gamma(y)$.

We can then solve for the coefficients of this linear combination from the first few values of f, giving an exact closed-form formula for f.

The Rudin-Shapiro set

The n that we choose have a base-2 representation of the form

$$z_t := \overbrace{10\,10\,\cdots\,10}^{t+1 \text{ copies}} = (2^{2t+3} - 2)/3.$$

Now $\gamma(10)$ has minimal polynomial

$$X^2(X-1)(X-2)(X-4)(X^3-5X^2+12X-16)(X^4-13X^3+72X^2-196X+256)$$

and hence there exist constants

$$a, b, c, \alpha, \gamma,\ \alpha_i, \gamma_i\ (1 \le i \le 2), \beta_i, \delta_i\ (1 \le i \le 4)$$

such that

$$d(z_t) = a + b \cdot 2^t + c \cdot 4^t + \alpha_1\gamma_1^t + \alpha_2\gamma_2^t + \alpha\gamma^t + \beta_1\delta_1^t + \beta_2\delta_2^t + \zeta_1\eta_1^t + \zeta_2\eta_2^t$$

where $\gamma, \gamma_1, \gamma_2$ are the zeros of $X^3 - 5X^2 + 12X - 16$ and the δ_i, η_i are the zeros of $X^4 - 13X^3 + 72X^2 - 196X + 256$.

The Rudin-Shapiro set

Here the α_i are complex conjugates, as are the γ_i, the β_i, the δ_i, the ζ_i, and the η_i.

Using Maple we can find the estimates

$$\text{(zeros of } X^3 - 5X^2 + 12X - 16) \quad \begin{cases} |\gamma_1|, |\gamma_2| & \doteq 2.41114 \\ \gamma & \doteq 2.75217 \end{cases}$$

$$\text{(zeros of } X^4 - 13X^3 + 72X^2 - 196X + 256) \quad \begin{cases} |\delta_1|, |\delta_2| & \doteq 4.88015 \\ |\eta_1|, |\eta_2| & \doteq 3.27859 \end{cases}$$

The dominant roots are clearly the δ_i and the corresponding coefficients are

$$\beta_1 \doteq -.03881 + .00706i$$
$$\beta_2 \doteq -.03881 - .00706i$$

The Rudin-Shapiro set

For t large enough, then, the value of $d(z_t)$ is dominated by

$$\beta_1\delta_1^t + \beta_2\delta_2^t = 2\Re(\beta_1\delta_1^t),$$

which is large and negative when (say)

$$3\pi/4 < \arg(\beta_1\delta_1^t) = (\arg(\beta_1) + t\arg(\delta_1)) \bmod 2\pi < 5\pi/4.$$

Since $\beta_1/|\beta_1|$ is not a root of unity, this will occur for infinitely many t.

Hence $d(z_t) < 0$ infinitely often.

Hence $r(3, \mathcal{R}, n)$ is not eventually increasing.

Powers of Thue-Morse power series

We can also study powers of the Thue-Morse power series

$$T(X) := \sum_{n \geq 0} t_n X^n = X + X^2 + X^4 + X^7 + X^8 + \cdots .$$

Allouche recently proved, using complex analysis and following ideas of Dombi, that the coefficients of $T^{10}(X)$ are eventually increasing.

Jean-Paul Allouche

Powers of Thue-Morse power series

More precisely, suppose $(q_n)_{n\geq 0}$ is a sequence of ± 1, and define $Q_n(z) = \sum_{0\leq j\leq n} q_j z^j$ and $A = \{n \geq 1 : q_{n-1} = 1\}$.

Theorem (Allouche)

Suppose there exist constants $C > 0$ and $0 < \alpha < 1$ such that for all complex z with $|z| = 1$ and all $n \geq 1$ one has $|Q_n(z)| \leq Cn^\alpha$. Then $(r(k, A, n))_{n\geq 0}$ is eventually strictly increasing for all $k > 2/(1-\alpha)$.

For Thue-Morse we can take $\alpha = (\log 3)/(\log 4) \doteq 0.79248$. Since $10 > 2/(1-\alpha) \doteq 9.63768$, Allouche's result follows.

Powers of Thue-Morse power series

On the other hand, we can prove (just as we did for Rudin-Shapiro) that the coefficients of $T^5(X)$ are *not* eventually increasing.

The status of T^6, T^7, T^8, T^9 is still unknown. It seems likely that T^6 has eventually increasing coefficients.

Dombi's conjecture

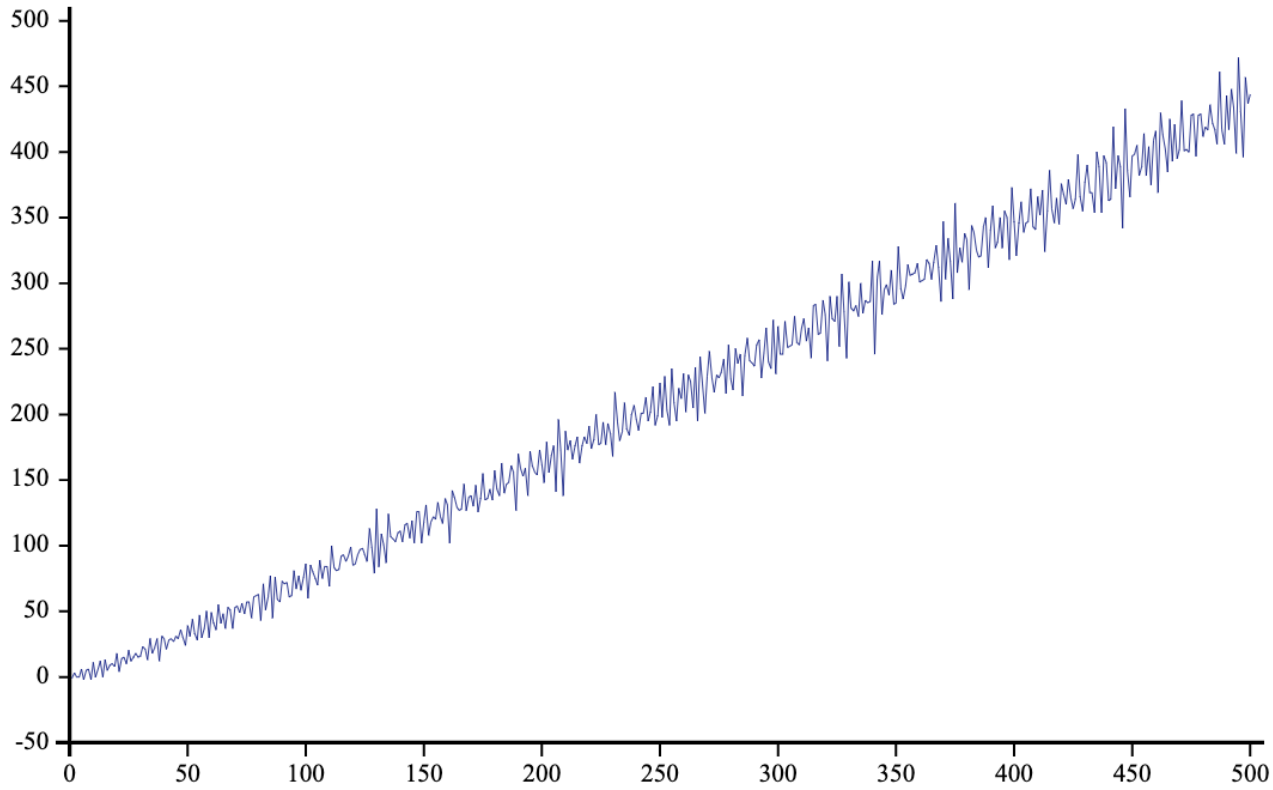
Let $A \subset \mathbb{N}$ be a set. Dombi studied the properties of $r(3, A, n)$—in particular the first difference of this sequence.

If $\mathbb{N} \setminus A$ is sparse, then we expect $r(3, A, n)$ to grow roughly like n^2 (because there are two choices for the first two summands, both with about n possibilities, and then the third is fixed).

So we expect the first difference $r(3, A, n) - r(3, A, n-1)$ to grow roughly like order n. But there might be fluctuations...

Dombi's conjecture

For example, here is how $r(3, A, n) - r(3, A, n-1)$ behaves when $A = \mathbb{N} \setminus \{1, 4, 9, 16, 25, \ldots\}$.



Dombi's conjecture refuted

Dombi (2002) conjectured that there is no set A such that $\mathbb{N} \setminus A$ is infinite and $r(3, A, n)$ is eventually increasing. But we have:

Theorem

Let $F = \{3 \cdot 2^n : n \geq 0\} = \{3, 6, 12, 24, \ldots\}$. Set $A := \mathbb{N} \setminus F$. Then $r(3, A, n)$ is strictly increasing right from the start.

Proof.

(Sketch.) Using `Walnut`, we generate a linear representation for $d(n) := r(3, A, n) - r(3, A, n-1)$, guess a closed form for it, and then verify the closed form with `Walnut`. The closed form is strong enough to show that $d(n)$ is always positive. □

Dombi's conjecture refuted

The closed form for

$$d(n) := r(3, A, n) - r(3, A, n-1)$$

looks like

$$d(3n+i) = 3n - 3\lceil \log_2 n \rceil - f_i(n),$$

for $i \in \{0, 1, 2\}$, where $(f_i(n))_{n \geq 0}$ is an automatic sequence.

How I found this counterexample

I found this counterexample by "intelligent guessing", namely:

- Do a breadth-first search on the tree of all possible finite characteristic sequences
- Reject sequences such that $r(3, A, n)$ is not strictly increasing right from the start
- Using the Myhill-Nerode theorem, find the size of the smallest automaton compatible with potential examples, and reject if it is too large.
- After some computation, I was left with potential counterexamples of automatic sequences generated by DFAO's with a small number of states.
- One of these worked right away.

A Dombi counterexample of positive density

The example of the previous slide corresponds to a sparse set: $F = \{3 \cdot 2^n : n \geq 0\}$.

This suggestions the question of whether there is an example where F has positive density.

Indeed there is such an example:

Theorem

Let $F = \{3, 12, 13, 14, 15, 48, 49, 50, \ldots\}$ be the set of natural numbers whose base-2 expansion is of even length and begins with 11.
Then F is of positive lower density and $r(3, \mathbb{N} \setminus F, n)$ is strictly increasing.

Proof.

Like before, using automata and the fact that F is a 2-automatic set. □

Dombi's conjecture refuted

Jason Bell found a different approach to refuting Dombi's conjecture:

Theorem (Bell & JOS, 2022)

Let $k \geq 3$ be an integer. Let $F \subseteq \mathbb{N}$ and assume $0 \notin F$. Let $(f(n))_{n\geq 0}$ be its associated characteristic sequence and $F(X)$ its associated power series $\sum_{i\geq 0} f(i)X^i$. Define $\sigma_f(n) = \sum_{0\leq i\leq n} f(i)$. Suppose $\sigma_f(n) = o(n^\alpha)$ for some $\alpha \leq (k-2)/k$ and $A = \mathbb{N} \setminus F$.

Then $(r(k, A, n))_{n\geq 0}$ is eventually strictly increasing.

Proof.

(Sketch.) $d(n) := r(k, A, n) - r(k, A, n-1)$ is the coefficient of X^n of

$$(1-X)\left(\frac{1}{1-X} - F(X)\right)^k.$$

Now expand using the binomial theorem and estimate the size of the coefficients. □

New advances on Dombi's theorem

News flash: the condition $\alpha \leq (k-2)/k$ was recently improved by Sándor Z. Kiss, Csaba Sándor, and Quan-Hui Yang to $\alpha \leq (k-2)/(k-1)$.

See https://arxiv.org/abs/2303.01314.

A Fibonacci example

Let $\mathbf{f} = 0100101001001\cdots$ be the infinite Fibonacci word. It has many different definitions, but one is as the fixed point of the map $0 \to 01$, $1 \to 0$.

Let $\mathcal{F}$ be the associated set $\{1, 4, 6, 9, 12, \ldots\}$, corresponding to the positions of the 1's in $\mathbf{f}$. This is a simple variation on the *upper Wythoff sequence*.

Yet another way to express $\mathcal{F}$ is via *Fibonacci representations*: we write n as a sum of non-adjacent Fibonacci numbers $\sum_{2 \le i \le t} a_i F_i$ with $a_i \in \{0, 1\}$. Then $n \in \mathcal{F}$ iff $a_2 = 1$.

A Fibonacci example

Theorem

The equalities

$$r(2, \mathcal{F}, n) = r(2, \mathcal{F}, n-1)$$
$$r(2, \mathbb{N} \setminus \mathcal{F}, n) = r(2, \mathbb{N} \setminus \mathcal{F}, n-1)$$

hold for infinitely many n.

Proof. We can compute a linear representation (v, γ, w) for $r(2, \mathcal{F}, n)$ using `Walnut`. The idea is to show that

$$r(2, \mathcal{F}, n) = r(2, \mathcal{F}, n-1) = (F_{6i+3} - 2)/4 \tag{1}$$

for $n = (F_{6i+5} - 5)/2$ and $i \geq 1$.

A Fibonacci example

The Fibonacci representation of $(F_{6i+5} - 5)/2$ is $(100)^{2i+1}10000$, and the Fibonacci representation of $(F_{6i+5} - 7)/2$ is $(100)^{2i+1}01010$.

So both $r(2, \mathcal{F}, (F_{6i+5} - 5)/2)$ and $r(2, \mathcal{F}, (F_{6i+5} - 7)/2)$ can be expressed as a linear combination of the i'th powers of the zeros of the minimal polynomial of $\gamma(100)$.

This minimal polynomial is $X^2(X-1)(X+1)(X^2-4X-1)$. Solving for the coefficients and simplifying gives Eq. (1).

Using exactly the same ideas, we can prove that

$$r(2, \mathbb{N} \setminus \mathcal{F}, n) = r(2, \mathbb{N} \setminus \mathcal{F}, n-1) = (L_{6i+6} - 2)/4$$

for $n = (F_{6i+8} - 7)/2$ and $i \geq 1$. □

Another Fibonacci example

Theorem

For $i \geq 1$ we have

$$r(3, \mathcal{F}, n) = r(3, \mathcal{F}, n-1) = (F_{2i+3}^2 - 3F_{2i+3} + 2)/4$$

for $n = F_{2i+5} - 4$ and $i \geq 1$.

Proof.

(Sketch.) The Fibonacci representation of $F_{2i+5} - 5$ is $(10)^i000$ and the Fibonacci representation of $F_{2i+5} - 4$ is $(10)^i001$.

We find the linear representation (v, γ, w) for $r(3, \mathcal{F}, n)$ and compute the minimal polynomial for $\gamma(10)$. It is $X^2(X-1)(X^2-3X+1)(X^2-7X+1)$. Solving for constants and simplifying gives the formulas above. □

Proofs for families of sequences

We can also use `Walnut` to prove results for certain families of *uncountably* many sequences, for example, the *paperfolding* numbers.

These are sequences describable from the iterated folding of a piece of paper, introducing at each step either a hill (1) or a valley (0). This gives a characteristic sequence of a set $S_{\mathbf{f}}$ depending on the sequence $\mathbf{f}$ of folding choices.

Theorem

For all paperfolding sequences $\mathbf{f}$*, every* $n \geq 15$ *is the sum of three elements of* $S_{\mathbf{f}}$*. The bound* 15 *is optimal.*

Three conjectures

Conjecture

- For the Rudin-Shapiro set $\mathcal{R}$ we have $r(4, \mathcal{R}, n) > r(4, \mathcal{R}, n-1)$ for $n \geq 196$.
- For the odious numbers $\mathcal{O}$ we have $r(6, \mathcal{O}, n) > r(6, \mathcal{O}, n-1)$ for $n \geq 6$.
- For the evil numbers $\mathcal{E}$ we have $r(6, \mathcal{E}, n) > r(6, \mathcal{E}, n-1)$ for $n \geq 38$.

Wrapping things up

- Automata and combinatorics on words can be used to prove new theorems about additive number theory.
- In some cases the theorems can be proven "purely mechanically", just by doing a computation.
- Sometimes the computations require a *lot* of space and time.
- However, the techniques I presented *cannot* be used for more traditional sequences, like primes and squares—at least directly.

For further reading

- G. Dombi. Additive properties of certain sets. *Acta Arith.* **103** (2002), 137–146.
- J. Lambek and L. Moser. On some two way classifications of integers. *Canad. Math. Bull.* **2** (1959), 85–89.
- J. P. Bell and J. Shallit, Counterexamples to a conjecture of Dombi in additive number theory, arXiv:2212.12473 [math.NT].
- J.-P. Allouche and J. Shallit, Additive properties of the evil and odious numbers and similar sequences, arXiv:2112.13627 [math.NT].
- J. Shallit, A Dombi counterexample with positive lower density, arXiv:2302.02138 [math.NT].
- J. Shallit, *The Logical Approach to Automatic Sequences: Exploring Combinatorics on Words with* `Walnut`, Cambridge Univ. Press, 2022.