

Morphisms and Morphic Words

Jeffrey Shallit

School of Computer Science

University of Waterloo

Waterloo, Ontario N2L 3G1

Canada

`shallit@graceland.uwaterloo.ca`

`http://www.cs.uwaterloo.ca/~shallit`

Transformations of Strings

The idea of transforming strings of symbols is ancient.
It is the basis for

- ▶ grammar
- ▶ cryptography (see H. Prodinger)
- ▶ mechanical composition of music
- ▶ automatic theorem proving
- ▶ computer algebra
- ▶ quasicrystals (see B. Solomyak)
- ▶ computer graphics

and many other subjects.

Morphisms

We focus on transformations where each symbol is replaced by a group of 0 or more symbols.

These transformations are called **morphisms**, but also go by other names:

- ▶ **substitutions** (mathematics)
- ▶ **inflation rules** (physics)
- ▶ **L-systems** (biology)

Morphisms Defined

Formally, a **morphism** is a map $h : \Sigma^* \rightarrow \Delta^*$, where Σ, Δ are alphabets, such that $h(xy) = h(x)h(y)$ for all strings $x, y \in \Sigma^*$.

If $\Sigma = \Delta$, we can iterate h .

By $h^2(x)$ I mean $h(h(x))$.

More generally $h^0(x) = x$ and $h^n(x) = h(h^{n-1}(x))$.

Types of Morphisms

- ▶ A morphism $h : \Sigma^* \rightarrow \Delta^*$ is **k -uniform** if $|h(a)| = k$ for all $a \in \Sigma$.
- ▶ A morphism is **uniform** if it is k -uniform for some k .
- ▶ A **coding** is a 1-uniform morphism.
- ▶ A morphism is **nonerasing** if $h(a) \neq \epsilon$ for all $a \in \Sigma$.
- ▶ A morphism is **primitive** if there exists an integer n such that for all letters $a, b \in \Sigma$, a appears in $h^n(b)$.

Types of Morphisms

Let $a \in \Sigma$ be a symbol. A morphism h is said to be **prolongable** if it satisfies the two conditions

- (i) $h(a) = ax$ for some $x \in \Sigma^*$;
- (ii) $h^n(x) \neq \epsilon$ for all $n \geq 0$.

If h is prolongable, we can define an infinite word in Σ^ω as follows:

$$h^\omega(a) = a x h(x) h^2(x) h^3(x) \dots$$

Note that

$$\begin{aligned} h(h^\omega(a)) &= h(a x h(x) h^2(x) h^3(x) \dots) \\ &= h(a) h(x) h^2(x) h^3(x) \dots \\ &= a x h(x) h^2(x) h^3(x) \\ &= h^\omega(a), \end{aligned}$$

so $h^\omega(a)$ is a **fixed point** of h .

The Thue-Morse Morphism

For example, consider the morphism μ mapping

$$0 \rightarrow 01;$$

$$1 \rightarrow 10.$$

If we start with 0 and iterate this morphism, we get successively

$$\mu^0(0) = 0$$

$$\mu^1(0) = 01$$

$$\mu^2(0) = 0110$$

$$\mu^3(0) = 01101001$$

$$\mu^4(0) = 0110100110010110$$

\vdots

The Thue-Morse Morphism

Since μ is prolongable, there is a well-defined infinite word

$$\mu^\omega(0) = 01101001100101101001 \dots$$

of which all the $\mu^i(0)$ are prefixes.

We call this word $\mathbf{t} = t_0 t_1 t_2 \dots$, the **Thue-Morse** word.

Thue and Morse



Figure: Axel Thue (1863–1922)



Figure: Marston Morse (1892–1977)

The Thue-Morse Sequence

Define a sequence of strings of 0's and 1's as follows:

$$X_0 = 0$$

$$X_{n+1} = X_n \overline{X_n}$$

where \overline{x} means change all the 0's in x to 1's and vice-versa.

For example, we find

$$X_0 = 0$$

$$X_1 = 01$$

$$X_2 = 0110$$

$$X_3 = 01101001$$

$$X_4 = 0110100110010110$$

$$\vdots$$

Then $\lim_{n \rightarrow \infty} X_n = \mathbf{t}$.

Let's prove $\mu^n(0) = X_n$ by induction on n .

Actually, it turns out to be easier to prove this together with the claim $\mu^n(1) = \overline{X_n}$.

These claims are clearly true for $n = 0$. Now assume they are true for n ; let's prove them for $n + 1$. We have

$$\begin{aligned}\mu^{n+1}(0) &= \mu^n(\mu(0)) \\ &= \mu^n(01) \\ &= \mu^n(0)\mu^n(1) \\ &= X_n \overline{X_n} \\ &= X_{n+1}.\end{aligned}$$

Similarly

$$\begin{aligned}\mu^{n+1}(1) &= \mu^n(\mu(1)) \\ &= \mu^n(10) \\ &= \mu^n(1)\mu^n(0) \\ &= \overline{X_n} X_n \\ &= \overline{X_{n+1}}.\end{aligned}$$

Another Definition

Given a number n we can write it in base 2,

$$n = \sum_{0 \leq i \leq k} a_i 2^i.$$

We define the “sum of digits” function $s_2(n)$ to be the sum of the a_i . So

$$s_2(43) = 1 + 0 + 1 + 0 + 1 + 1 = 4.$$

Then $t_n = s_2(n) \bmod 2$.

Let's prove this by induction. It's evidently true for $n = 0$. Now assume it is true for all $n' < n$.

- ▶ Define k by $2^k \leq n < 2^{k+1}$.
- ▶ Then t_n is the n 'th symbol of X_{k+1} .
- ▶ So it is the $(n - 2^k)$ 'th symbol of $\overline{X_k}$.
- ▶ In other words, $t_n = (t_{n-2^k} + 1) \bmod 2$.
- ▶ By induction we have $t_{n-2^k} = s_2(n - 2^k) \bmod 2$.
- ▶ Since $2^k \leq n < 2^{k+1}$, we have $s_2(n) = s_2(n - 2^k) + 1$.
- ▶ It follows that $t_n = s_2(n) \bmod 2$.

The definition in terms of $s_2(n)$ is good because we can efficiently compute t_n without having to compute t_0, t_1, \dots, t_{n-1} .

Repetitions in Strings

A **square** is a string of the form xx . Examples include

- ▶ **murmur**
- ▶ **hotshots**

- ▶ **chercher**
- ▶ **couscous**
- ▶ **rentrent**

- ▶ **tenten**
- ▶ **kerker**

A word is **squarefree** if it contains no subword (block of consecutive symbols) that is a square. Note that **squarefree** is not squarefree, but **square** is.

Cubes

A **cube** is a string of the form xxx . Examples in English include

- ▶ hahaha
- ▶ shshsh

The name of a certain Dutch mathematician contains a cube. Do you know who?

Cornelis Lek**kerkerker**

A word is **cube-free** if it contains no subword that is a cube.

Fourth Powers

A **fourth power** is a string of the form $xxxx$. The only example I know of in English is

▶ tratra tratra

which is an extinct lemur from Madagascar.

Overlaps

A **overlap** is a string of the form $axaxa$ where a is a single letter and x is a string. Examples include

- ▶ alfalfa
- ▶ entente
- ▶ koekoek

A word is **overlap-free** if it contains no word that is an overlap.

No Squarefree Words Over $\{0, 1\}$

Theorem

There are no squarefree strings of 0's and 1's of length ≥ 4 .

Proof.

Assume x is squarefree and $|x| \geq 4$. Then without loss of generality we may assume the first symbol of x is 0. Then the second symbol must be 1, for otherwise we would have the square 00. Then the third symbol must be 0, for otherwise we would have the square 11. Thus the first three symbols are 010, and whatever symbol we choose next gives a square. Contradiction. \square

Avoiding Squares Over Larger Alphabets

But how about over larger alphabets? Are there large squarefree words over three symbols?

A backtracking algorithm gives

0102012021...

and seems to go on forever.

But how can we prove that there exists an infinite squarefree word?

This is what Thue did.

The Thue-Morse word plays a critical role.

Avoiding Overlaps

As a first step let's show that \mathbf{t} is overlap-free: it contains no subword of the form $axaxa$, with a a single letter and x a string.

Theorem

The Thue-Morse infinite word \mathbf{t} is overlap-free.

Proof. Observe that $t_{2n} = t_n$ and $t_{2n+1} = 1 - t_n$ for $n \geq 0$.

Assume, contrary to what we want to prove, that \mathbf{t} contains an overlap.

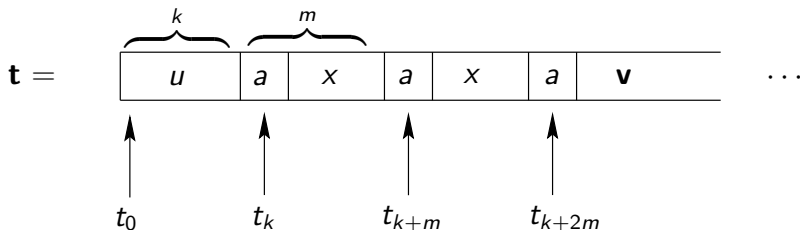


Figure: Hypothesized overlap in \mathbf{t}

Then we would be able to write $\mathbf{t} = uaxaxav$ for some finite strings u, x , an infinite string \mathbf{v} , and a letter a .

In other words, we would have $t_{k+j} = t_{k+j+m}$ for $0 \leq j \leq m$, where $m = |ax|$ and $k = |u|$. Assume $m \geq 1$ is as small as possible. Then there are two cases: (i) m is even; and (ii) m is odd.

(i) If m is even, then let $m = 2m'$. Again there are two cases: (a) k is even; and (b) k is odd.

(a) If k is even, then let $k = 2k'$. Then we know $t_{k+j} = t_{k+j+m}$ for $0 \leq j \leq m$, so it is certainly true that $t_{k+2j'} = t_{k+2j'+m}$ for $0 \leq j' \leq m/2$. Hence $t_{2k'+2j'} = t_{2k'+2j'+2m'}$ for $0 \leq j' \leq m'$, and so $t_{k'+j'} = t_{k'+j'+m'}$ for $0 \leq j' \leq m'$. But this contradicts the minimality of m .

(b) If k is odd, then let $k = 2k' + 1$. Then as before we have $t_{k+2j'} = t_{k+2j'+m}$ for $0 \leq j' \leq m/2$. Hence $t_{2k'+2j'+1} = t_{2k'+2j'+2m'+1}$ for $0 \leq j' \leq m'$, and so $t_{k'+j'} = t_{k'+j'+m'}$ for $0 \leq j' \leq m'$, again contradicting the minimality of m .

(ii) If m is odd, then there are three cases:

- (a) $m \geq 5$;
- (b) $m = 3$; and
- (c) $m = 1$.

For $n \geq 1$, we define $b_n = (t_n + t_{n-1}) \bmod 2$.

Note that $b_{4n+2} = (t_{4n+2} + t_{4n+1}) \bmod 2$.

Since the base-2 representations of $4n + 2$ and $4n + 1$ are identical, except that the last two bits are switched, we have $t_{4n+2} = t_{4n+1}$, and so $b_{4n+2} = 0$.

On the other hand, $b_{2n+1} = (t_{2n+1} + t_{2n}) \bmod 2$, and the base-2 representations of $2n + 1$ and $2n$ are identical except for the last bit; hence $b_{2n+1} = 1$.

(a) m odd, ≥ 5 . We have $b_{k+j} = b_{k+j+m}$ for $1 \leq j \leq m$. Since $m \geq 5$, we can choose j such that $k+j \equiv 2 \pmod{4}$. Then for this value of $k+j$, we have from above that $b_{k+j} = 0$, but $k+j+m$ is odd, so $b_{k+j+m} = 1$, a contradiction.

(b) $m = 3$. Again, $b_{k+j} = b_{k+j+3}$ for $1 \leq j \leq 3$. Choose j such that $k+j \equiv 2$ or $3 \pmod{4}$. If $k+j \equiv 2 \pmod{4}$, then the reasoning of the previous case applies. Otherwise $k+j \equiv 3 \pmod{4}$, and then $b_{k+j} = 1$, while $b_{k+j+3} = 0$.

(c) $m = 1$. Then $t_k = t_{k+1} = t_{k+2}$. Hence $t_{2n} = t_{2n+1}$ for $n = \lceil k/2 \rceil$, a contradiction.

This completes the proof.

Using the fact that \mathbf{t} is overlap-free, we may now construct a squarefree infinite word over the alphabet $\Sigma_3 = \{0, 1, 2\}$.

Theorem

For $n \geq 1$, define c_n to be the number of 1's between the n th and $(n+1)$ st occurrence of 0 in the word \mathbf{t} . Set $\mathbf{c} = c_1 c_2 c_3 \dots$. Then $\mathbf{c} = 210201\dots$ is an infinite squarefree word over the alphabet Σ_3 .

Proof.

First, observe that \mathbf{c} is over the alphabet $\{0, 1, 2\}$. For if there were three or more 1's between two consecutive occurrences of 0 in \mathbf{t} , then \mathbf{t} would not be overlap-free, a contradiction.

Next, assume that \mathbf{c} is not squarefree. Then it contains a square of the form xx , with $x = x_1 x_2 \dots x_n$ and $n \geq 1$. Then, from the definition of \mathbf{c} , the word \mathbf{t} would contain a subword of the form

$$01^{x_1}01^{x_2}0\dots01^{x_n}01^{x_1}01^{x_2}0\dots01^{x_n}0$$

which constitutes an overlap, a contradiction. □

Avoidability in Words

Thue's result is just the first step in a large and active theory:
avoidability in words.

Obvious generalization: avoid fractional powers. We say x is a p/q power if x is of length p and period q . Thus **outshout** is an $8/5$ -power.

The **critical exponent** of a word \mathbf{w} is the supremum over all α , such that \mathbf{w} contains an α -power.

Recent Results in Avoidability

Some recent results:

- ▶ (Karhumäki & JOS) There are only polynomially many binary words of length n that avoid α -powers, for $\alpha \leq 7/3$;
- ▶ (Karhumäki & JOS) There are exponentially many binary words of length n that avoid $(7/3 + \epsilon)$ -powers.
- ▶ (Rampersad) The Thue-Morse word and its complement are the only words that are the fixed points of a non-trivial morphism and avoid α -powers for $2 < \alpha \leq 7/3$.
- ▶ (Brown, Rampersad, Vasiga, & JOS) If you change any finite number of bits in the Thue-Morse word, it has an overlap.
- ▶ (Krieger) The critical exponent for uniform binary morphisms is rational, and for non-erasing morphisms is algebraic.
- ▶ (Kreiger & JOS) For each real $\alpha > 1$ there exists an infinite word with α as critical exponent.

More About Thue-Morse: An Amazing Infinite Product

Consider the sequence

$$\frac{1}{2}, \quad \frac{1/2}{3/4}, \quad \frac{\frac{1/2}{3/4}}{\frac{5/6}{7/8}}, \quad \frac{\frac{1/2}{3/4}}{\frac{5/6}{7/8}} / \frac{9/10}{\frac{11/12}{13/14}}, \dots$$

What does this converge to?

An Amazing Infinite Product

We find

$$\frac{1}{2} \doteq .500$$

$$\frac{1/2}{3/4} = 2/3 \doteq .666$$

$$\frac{\frac{1/2}{3/4}}{5/6} = 7/10 \doteq .700$$

Numerically, the limit seems to be $0.7071\dots$.

Let's prove that this sequence converges to $\frac{\sqrt{2}}{2}$.

An Amazing Infinite Product

First, we observe that the limit is

$$\prod_{n \geq 0} \left(\frac{2n+1}{2n+2} \right)^{(-1)^{t_n}} \quad (1)$$

where t_n is the sum of the bits (mod 2) in the binary expansion of n .

An Amazing Infinite Product

We now use a trick of Allouche: let

$$P = \prod_{n \geq 0} \left(\frac{2n+1}{2n+2} \right)^{(-1)^{t_n}}$$

and define

$$Q = \prod_{n \geq 1} \left(\frac{2n}{2n+1} \right)^{(-1)^{t_n}}.$$

Clearly

$$PQ = \frac{1}{2} \prod_{n \geq 1} \left(\frac{n}{n+1} \right)^{(-1)^{t_n}}.$$

An Amazing Infinite Product

Now break this infinite product into separate products over odd and even indices; we find

$$\begin{aligned}PQ &= \frac{1}{2} \prod_{n \geq 1} \left(\frac{n}{n+1} \right)^{(-1)^{t_n}} \\&= \frac{1}{2} \prod_{n \geq 0} \left(\frac{2n+1}{2n+2} \right)^{(-1)^{t_{2n+1}}} \prod_{n \geq 1} \left(\frac{2n}{2n+1} \right)^{(-1)^{t_n}} \\&= \frac{1}{2} P^{-1} Q.\end{aligned}$$

It follows that $P^2 = \frac{1}{2}$.

But how about Q ? Is it irrational? Transcendental?

I offer 20 Euros for the answer to this question.



The Multigrades Problem

The **multigrades problem** is the following: let I and J be disjoint sets. Can one find “short” solutions to the system of equations

$$\sum_{i \in I} i^k = \sum_{j \in J} j^k$$

for $k = 0, 1, 2, \dots, t$?

For example, one solution for $t = 2$ is the identity

$$0^k + 3^k + 5^k + 6^k = 1^k + 2^k + 4^k + 7^k$$

for $k = 0, 1, 2$.

In 1851, the French mathematician Étienne Prouhet gave the following general solution.

The Multigrades Problem

Theorem

The Thue-Morse sequence $\mathbf{t} = (t_n)_{n \geq 0}$ has the following property.

Define

$$I = \{0 \leq i < 2^N : t_i = 0\}$$

$$J = \{0 \leq j < 2^N : t_j = 1\}$$

Then for $0 \leq k < N$ we have

$$\sum_{i \in I} i^k = \sum_{j \in J} j^k.$$

For example, for $N = 3$ we have the partition obtained before:

$$0^k + 3^k + 5^k + 6^k = 1^k + 2^k + 4^k + 7^k$$

for $k = 0, 1, 2$.

The Multigrades Problem

Proof.

We actually prove a more general theorem by induction on N . We prove that if p is any polynomial of degree $< N$, then

$$\sum_{\substack{0 \leq i < 2^N \\ t_i = 0}} p(i) = \sum_{\substack{0 \leq j < 2^N \\ t_j = 1}} p(j)$$

The desired result then follows by successively considering the case $p(i) = 1$, $p(i) = i$, $p(i) = i^2$, etc.

The base case is $N = 1$. Then p is a constant, the result clearly follows.

The Multigrades Problem

Proof.

Now assume the result true for all polynomials of degree $< N$. We try to prove it for a polynomial $p(x)$ of degree N . Consider the polynomial $p(x + 2^N) - p(x)$. If

$$p(x) = a_N x^N + a_{N-1} x^{N-1} + \cdots + a_1 x + a_0,$$

then $p(x + 2^N) = a_N(x + 2^N)^N + \text{smaller degree terms}$, which by the binomial theorem, is $a_N(x^N + \text{smaller degree terms})$. So

$p(x + 2^N) - p(x)$ is actually a polynomial of degree $< N$. So we can apply induction to it. We get

The Multigrades Problem

$$\sum_{\substack{0 \leq i < 2^N \\ t_i=0}} (p(i + 2^N) - p(i)) = \sum_{\substack{0 \leq j < 2^N \\ t_j=1}} (p(j + 2^N) - p(j))$$

So, rearranging, we get

$$\begin{aligned} & \sum_{\substack{0 \leq i < 2^N \\ t_i=0}} p(i + 2^N) + \sum_{\substack{0 \leq j < 2^N \\ t_j=1}} p(j) \\ &= \sum_{\substack{0 \leq j < 2^N \\ t_j=1}} p(j + 2^N) + \sum_{\substack{0 \leq i < 2^N \\ t_i=0}} p(i). \end{aligned}$$

Hence

$$\sum_{\substack{0 \leq i < 2^{N+1} \\ t_i=0}} p(i) = \sum_{\substack{0 \leq j < 2^{N+1} \\ t_j=1}} p(j).$$

We're done.

Exercise: find the appropriate generalization for bases larger than 2.

An Open Problem

The Thue-Morse partition of $\{0, 1, \dots, 2^N - 1\}$ is **not** the **unique** partition such that

$$\sum_{i \in I} i^k = \sum_{j \in J} j^k$$

for $k = 0, 1, \dots, N - 1$.

But is it the partition that minimizes $\left| \sum_{i \in I} i^N - \sum_{j \in J} j^N \right|$?

And if so, is it unique?

Another Definition

Yet another definition of the Thue-Morse sequence \mathbf{t} can be given in terms of power series. Let X be an indeterminate. We have

$$\begin{aligned}\prod_{i \geq 0} (1 - X^{2^i}) &= (1 - X)(1 - X^2)(1 - X^4) \dots \\ &= 1 - X - X^2 + X^3 - X^4 \dots \\ &= \sum_{j \geq 0} (-1)^{t_j} X^j.\end{aligned}$$

Another Definition

Something even more interesting arises when we consider Laurent series over $GF(2)$, the finite field with two elements. Basically, we do all arithmetic operations as usual, but reduce modulo 2.

For example, consider the Laurent series

$$G(X) = X^{-1} + X^{-2} + X^{-4} + X^{-8} + \dots .$$

It turns out that this series is *algebraic* over $GF(2)(X)$. By this we mean that G is the analogue of an algebraic number, a number satisfying an algebraic equation.

Formal Power Series

Let's try to find the equation that

$$G(X) = X^{-1} + X^{-2} + X^{-4} + X^{-8} + \dots$$

satisfies.

What is $G(X)^2$? If we compute it over the integers, we get

$$X^{-2} + 2X^{-3} + X^{-4} + 2X^{-5} + 2X^{-6} + X^{-8} + 2X^{-9} + \dots .$$

Reduced mod 2, this is just

$$X^{-2} + X^{-4} + X^{-8} + X^{-16} + \dots .$$

More generally, if we have a power series $H(X)$, then $H(X)^p = H(X^p)$ over $GF(p)$, where p is a prime number.

To see this, it suffices to remember that

$$(a + b)^p \equiv a^p + b^p \pmod{p}.$$

Formal Power Series

For $G(X) = X^{-1} + X^{-2} + X^{-4} + X^{-8} + \dots$ we get
 $G(X)^2 = G(X) - X^{-1}$, and so

$$G^2 + G + X^{-1} = 0.$$

Thus G is quadratic.

The Thue-Morse Power Series

Theorem

Let $F(X) = \sum_{n \geq 0} t_n X^{-n}$. Then, over $GF(2)$, the Laurent series F satisfies a quadratic equation with coefficients that are polynomials in X . More precisely, we have

$$(1 + X)^3 F^2 + X(1 + X)^2 F + X^2 = 0. \quad (2)$$

The Thue-Morse Power Series

Proof.

$$\begin{aligned} F &= \sum_{n \geq 0} t_n X^{-n} \\ &= \sum_{n \geq 0} t_{2n} X^{-2n} + \sum_{n \geq 0} t_{2n+1} X^{-2n-1} \\ &= \sum_{n \geq 0} t_n X^{-2n} + X^{-1} \sum_{n \geq 0} (1 + t_n) X^{-2n} \\ &= F^2 + X^{-1} \left(\frac{X^2}{1 + X^2} + F^2 \right) \\ &= \left(\frac{1 + X}{X} \right) F^2 + \frac{X}{1 + X^2} \\ &= \left(\frac{1 + X}{X} \right) F^2 + \frac{X}{(1 + X)^2}. \end{aligned}$$

The Thue-Morse Power Series

We have

$$F = \left(\frac{1+X}{X} \right) F^2 + \frac{X}{(1+X)^2}.$$

Hence, multiplying through by $X(1+X)^2$, we obtain

$$(1+X)^3 F^2 + X(1+X)^2 F + X^2 = 0.$$

The fact that F is not a rational function is an easy consequence of the overlap-free property of the sequence \mathbf{t} .

Thue-Morse and Continued Fractions

Recall that an expression of the form

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \cdots + \frac{1}{a_n}}}$$

is called a **continued fraction**.

It is usually abbreviated as $[a_0, a_1, a_2, \dots, a_n]$.

We can also consider infinite expressions of the form

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \cdots}}$$

which is abbreviated $[a_0, a_1, a_2, \dots]$.

The terms a_i are positive integers, except for a_0 , which may be any integer.

Thue-Morse and Continued Fractions

A continued fraction where the terms include 0 can be converted into the ordinary form using the identity

$$[\dots, a, 0, b, \dots] = [\dots, a + b, \dots].$$

Now consider the continued fraction where the terms are the Thue-Morse sequence:

$$[0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, \dots].$$

Using the collapsing rule gives

$$[0, 1, 2, 1, 1, 2, \dots].$$

This number is transcendental.

The terms of this continued fraction, disregarding the initial 0, form the **run-lengths** of the symbols in the Thue-Morse sequence.

An Open Problem

Characterize all infinite sequences of 0's and 1's such that when 0's are collapsed, the result is the run lengths of the original sequence, shifted by one.

The Thue-Morse Sequence and Chess

According to official rule (10.12) of the game of chess, a player can claim a draw if “at least 50 consecutive moves have been made by each side without the capture of any piece and without the movement of any pawn”. Actually, this is not enough for certain positions, such as King + Rook + Bishop versus King + 2 Knights, so the rule also stipulates that “This number of 50 moves can be increased for certain positions, provided that this increase in number and these positions have been clearly announced by the organisers before the event starts.”

The Thue-Morse Sequence and Chess

Another rule (10.10) allows a draw to be claimed if the same position occurs for the third time. By “same position” we mean that the pieces are in the same position, including the rights to castle or capture a pawn *en passant*. Without these two rules, infinite games are clearly possible. However, can rule (10.10) be weakened and still disallow infinite chess games?

The Thue-Morse Sequence and Chess

Consider the following alternative rule: a draw occurs if the same sequence of moves occurs twice in succession and is immediately followed by the first move of a third repetition.

Can an infinite game of chess occur under this rule?

The question was answered by Max Euwe, the Dutch chess master (and world champion from 1935–1937) in 1929.



Figure: Max Euwe (1901–1981)

The Thue-Morse Sequence and Chess

Euwe's construction used the Thue-Morse sequence! (He discovered it independently.)

One way to do this is to take the Thue-Morse sequence and map 0 to a sequence of moves, and 1 to another sequence of moves. For example, one way is as follows:

0	→	Ng1 – f3	Ng8 – f6
		Nf3 – g1	Nf6 – g8
1	→	Nb1 – c3	Nb8 – c6
		Nc3 – b1	Nc6 – b8

The Thue-Morse Sequence and Music

The Thue-Morse sequence has even been used in composing music!

Tom Johnson, a Paris-based composer, has used the Thue-Morse sequence and other sequences formed by iterated morphisms, in his work.

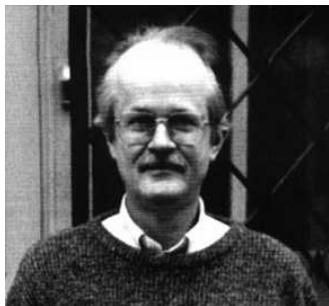


Figure: Tom Johnson

The Thue-Morse Sequence and Music



Figure: Composition by Tom Johnson

Per Nørgård



Some compositions based on the “infinity series”

$$0, 1, -1, 2, 1, 0, -2, 3, -1, 2, 0, 1, 2, -1, -3, 4, 1, \dots$$

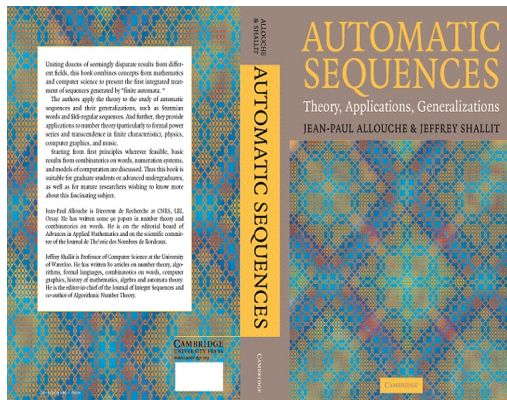
defined by

$$\begin{aligned}c_0 &= 0 \\c_{2n} &= -c_n \\c_{2n+1} &= c_n + 1.\end{aligned}$$

Note that $c_n \equiv t_n \pmod{2}$.

For Further Reading

Jean-Paul Allouche and Jeffrey Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, 2003.



For Further Reading

Jean-Paul Allouche and Jeffrey Shallit, The ubiquitous Prouhet-Thue-Morse sequence, in C. Ding, T. Helleseth, and H. Niederreiter, eds., *Sequences and Their Applications: Proceedings of SETA '98*, Springer-Verlag, 1999, pp. 1-16. Also available at <http://www.cs.uwaterloo.ca/~shallit/papers.html>

http://www.swets.nl/jnmr/vol24_2.html

<http://tom.johnson.org>

<http://www.pernoergaard.dk>