Ordered patterns in words generated by morphisms

Sergey Kitaev
Reykjavik University

Joint work with Toufik Mansour (University of Haifa) and
Patrice Seebold (Universite de Picardie Jules Verne)

Let $\mathcal{A} = \{a_1, a_2, \ldots, a_k\}$ be an alphabet of k letters.

A *word* in the alphabet $\mathcal{A}$ is a finite sequence of letters of the alphabet.

Any i consecutive letters of a word X generate a *factor* of length i.

The set $\mathcal{A}^*$ is the set of all the words on the alphabet $\mathcal{A}$.

Let Σ be an alphabet.

A map $\varphi : \Sigma^* \to \Sigma^*$ is called a *morphism*, if we have

$$\varphi(uv) = \varphi(u)\varphi(v)$$

for any $u, v \in \Sigma^*$.

A morphism φ can be defined by defining $\varphi(i)$ for each $i \in \Sigma$.

The *Thue-Morse sequence* is defined by the morphism μ:

$$\mu(0) = 01,$$
$$\mu(1) = 10.$$

0, 01, 0110, 01101001, ...

This sequence does not contain a factor of the form XXx, where X is itself a factor and x is the first letter in X.

Another way to define this sequence is

$$\mu^n(0) = \mu^{n-1}(0)C(\mu^{n-1}(0)),$$

where C is the *complement* (switching 0 and 1).

$u = u_1 u_2 \ldots$, where u_i's are letters over a finite alphabet

u's *complexity* $f_u(n)$ is the number of distinct words of the form $u_i u_{i+1} \ldots u_{i+n-1}$ and its *arithmetical complexity* $a_u(n)$ is the number of words of the form $u_k u_{k+d} \ldots u_{k+(n-1)d}$ for any i, k and d.

For example, if $u = 0110011$ then $f_u(3) = 4$ and $a_u(3) = 6$

1975 – subword (factor) complexity (Ehrenfeucht, Lee, Rozenberg)

1976 – Lempel-Ziv complexity (Lempel, Ziv)

1987 – d-complexity (Iványi)

1995 – palindrome (palindromic) complexity (Hof, Knill, Simon)

2000 – arithmetical complexity (Avgustinovich, Fon-Der-Flaass, Frid)

2002 – pattern complexity (Restivo, Salemi)

2002 – maximal pattern complexity (Kamae, Zamboni)

For arbitrary infinite word one has

$$P(n) \leq \frac{16}{n} f\left(n + \lfloor \frac{n}{4} \rfloor\right), \quad \text{for all } n \in \mathbb{N},$$

where $P(n)$ is the palindrome complexity and $f(n)$ is the factor complexity of the word. (Allouche, Baake, Cassaigne, and Damanik, 2003)

Sturmian words – binary words whose factor complexity is minimal among all non-periodic words and is equal to $n + 1$ for all n (study of Sturmian words dates back to J. Bernoulli III (1772), the first comprehensive study of them was by G. A. Hedlund and M. Morse in 1940). The *Fibonacci word* is an example of such a sequence.

An infinite word s is Sturmian if and only if for n even the number of palindrome factors of s is 1 and for n odd it is 2. (Droubay and Pirillo, 1999)

Arithmetical complexity is $O(n^3)$ (known upper and lower bounds differ by appr. 10.58 times). This complexity itself depends on choice of the Sturmian word. (Cassaigne and Frid, 2007)

Complexity of an infinite word generated by morphism has one of the following orders of growth 1, n, $n \log\log n$ or n^2. (Pansiot, 1984)

The variety of rates of growth of arithmetical complexity is not less than the variety of possible factor complexity rates of growth. (Frid, 2006)

Patterns in words:

XX – *squares*, $1\underbrace{02}\underbrace{02}010$

XXX – *cubes*, $11\underbrace{011}\underbrace{011}\underbrace{011}0$

XYXXY – $2\underbrace{11}\underbrace{323}\underbrace{11}\underbrace{11}\underbrace{323}3$

Permutation patterns:

The pattern 1-2 occurs in 34152 as 34, 35, 45, 15, and 12

The pattern 12, *ascent* or *rise*, occurs in 34152 as 34 and 15

The pattern 2-1, (*inversion*), occurs in 34152 as 31, 32, 41, 42, and 52

The pattern 21, *descent*, occurs in 34152 as 41 and 52

The pattern 1-23 occurs in 24135 as 235 and 135

Repetitions in patterns are allowed too while dealing with words!

Parikh vectors

$A = \{a_1 < a_2 < a_3\}$

$w = a_2 a_1 a_3 a_1 a_3$

The Parikh vector of w is $(|w|_{a_1}, |w|_{a_2}, |w|_{a_3}) = (2, 1, 2)$

Introduced by R.J. Parikh in 1966

Parikh matrices

Introduced by A. Mateescu, A. Salomaa, K. Salomaa and S. Yu in 2001

$\Psi(aabbc) = \Psi(a)\Psi(a)\Psi(b)\Psi(b)\Psi(c) =$

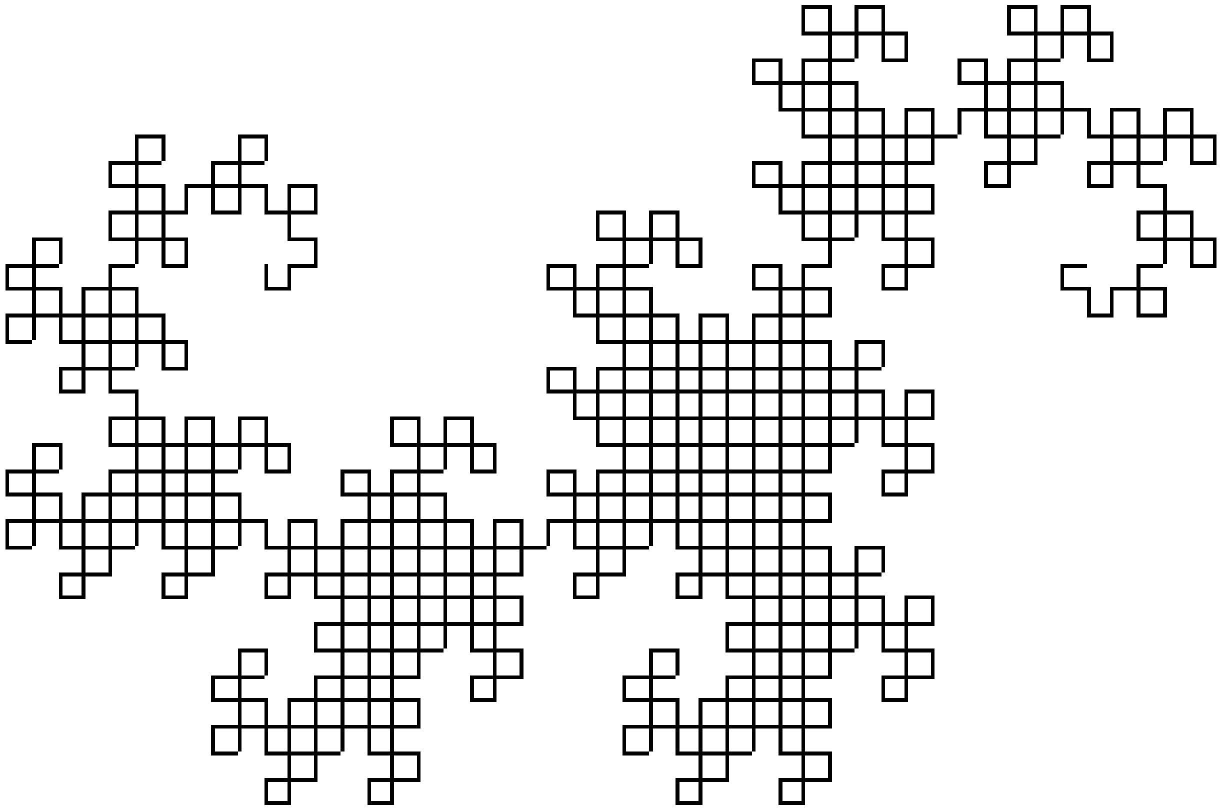
$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 & 4 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Parikh vector is the second diagonal, but we get much more!

The (Harter-Heighway) Dragon Curve, 1967

The *Dragon curve* (*paperfolding sequence*) was discovered by physicist John E. Heighway. (An example of a recursively generated *fractal shape*.)



Any natural number n can be presented unambiguously as $n = 2^t(4s + \sigma)$, where $\sigma < 4$, and t is the greatest natural number such that 2^t divides n.

If n runs through the natural numbers then σ runs through the sequence that we will call the *sequence of* σ.

We let w_σ denote that sequence. Obviously, w_σ consists of 1s and 3s.

The initial letters of w_σ are $11311331113313\ldots$.

An equivalent definition of the σ-sequence:

$$C_1 = 1, \qquad D_1 = 3$$

$$C_{k+1} = C_k 1 D_k, \qquad D_{k+1} = C_k 3 D_k$$

$$k = 1, 2, \ldots$$

and $w_\sigma = \lim_{k \to \infty} C_k$.

$\Omega = \{A, B, a, b\}$, and γ and h are the following morphisms.

$$
\begin{aligned}
\gamma : \Omega^* &\rightarrow \Omega^* \\
A &\mapsto AaB \\
B &\mapsto AbB \\
a &\mapsto a \\
b &\mapsto b
\end{aligned}
\qquad\qquad
\begin{aligned}
h : \Omega^* &\rightarrow \{1, 3\}^* \\
A &\mapsto 1 \\
B &\mapsto 3 \\
a &\mapsto 1 \\
b &\mapsto 3
\end{aligned}
$$

Theorem. The σ-sequence w_σ (the Dragon curve) is generated by the *tag-system* $(\Omega, A, \gamma, h, \{1, 3\})$, i.e., $w_\sigma = h(\gamma^\omega(A))$.

The sigma-sequence w_σ can be obtained from traveling along the Dragon curve:

Following the curve from beginning to end, each turn is either to the left or to the right. Thus, each generation of the dragon corresponds to a sequences of 1's (lefts) and 3's (rights). It turns out that we get exactly w_σ.

Application to the "Snake in the box" problem.

Theorem. [SK, 2003] There does not exist a morphism whose iteration defines the sequence of σ.

Theorem. [Allouche, 1997; Baake 1999] For the sequence of σ, the palindromic complexity, $P(n)$, is 0 for $n \geq 14$.

Samples of counting patterns in w_σ results, [SK, 2003-2004]

Among the first $2^n - 1$ symbols of w_σ:

$\underbrace{1-1-\cdots-1}_{k}$ occurs $\frac{2^n-k}{2^{n-1}-k}\binom{2^{n-1}-1}{k}$ times

1-2 occurs $2 \cdot 4^{n-2} + (n-2) \cdot 2^{n-2}$ times

221 occurs $3 \cdot 2^{n-4} - 1$ times

12-21 occurs $\frac{1}{2}4^{n-2} - 3 \cdot 2^{n-4}$ times

1-221 occurs $\frac{1}{2}4^{n-2} + 27 \cdot 2^{n-5} - n - 7$ times

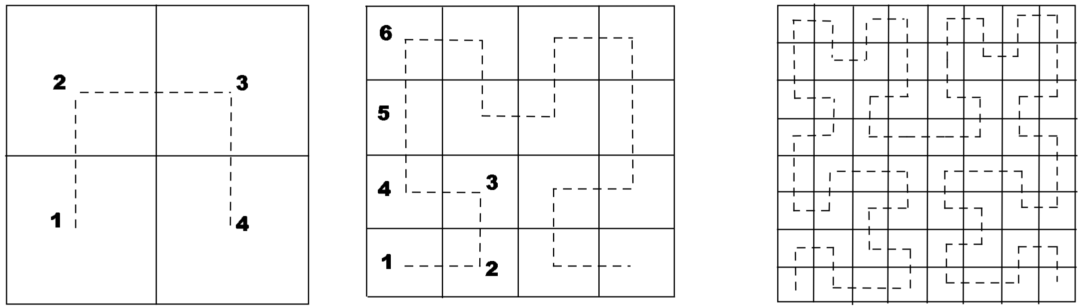
The number c_n^τ of occurrences of 2-1-221 among the first $2^n - 1$ symbols of w_σ can be calculated using

$$\begin{pmatrix} c_n^\tau \\ \\ d_n^\tau \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_{n-1}^\tau \\ \\ d_{n-1}^\tau \end{pmatrix} + \begin{pmatrix} \frac{5}{1024}8^n + \frac{25-3n}{256}4^n - \frac{171}{64}2^n + 9 \\ \\ \frac{5}{1024}8^n + \frac{21-3n}{256}4^n - 2^{n+1} \end{pmatrix}$$

with initial conditions $c_5^\tau = 70$ and $d_5^\tau = 74$.

The Peano (Hilbert) Curve, 1890 (1891)

The Peano curve is an example of *fractal space filling curves*.



x_1 x_2 x_3

A *Peano word* P_n is obtained by traveling along the Peano curve after the n-th iteration.

P_n is over $\Sigma = \{u, \bar{u}, r, \bar{r}\}$ where u stands for *up*, $\bar{u}$ stands for *down*, r stands for *right*, and $\bar{r}$ stands for *left*.

The *Peano infinite word* $P = \lim_{n\to\infty} P_{2n+1}$.

$\Omega = \{A, B, C, D, a, b, c, d\}$, and γ and h are the following morphisms.

$$\begin{aligned} \gamma : \Omega^* &\longrightarrow \Omega^* \\ A &\mapsto BaAbAcD \\ B &\mapsto AbBaBdC \\ C &\mapsto DcCdCaB \\ D &\mapsto CdDcDbA \\ a &\mapsto a \\ b &\mapsto b \\ c &\mapsto c \\ d &\mapsto d \end{aligned} \qquad \begin{aligned} h : \Omega^* &\longrightarrow \Sigma^* \\ A &\mapsto ur\bar{u} \\ B &\mapsto ru\bar{r} \\ C &\mapsto \bar{u}\bar{r}u \\ D &\mapsto \bar{r}\bar{u}r \\ a &\mapsto u \\ b &\mapsto r \\ c &\mapsto \bar{u} \\ d &\mapsto \bar{r} \end{aligned}$$

Theorem. [SK, Mansour, Séébold, 2003] P is the infinite word generated by the tag-system $(\Omega, A, \gamma^2, h, \Sigma)$, i.e., $P = h((\gamma^2)^\omega(A))$.

Theorem. [SK, Mansour, Séébold, 2003] The infinite word P does not contain any factor $xyWxyWxy$ with x, y letters and W a word. In particular, the only cubes in P are x^3 with x a letter. Moreover, P is 4-power-free.

Corollary. [SK, Mansour, Séébold, 2003] The infinite word P cannot be generated by a *D0L-system*, and thus cannot be generated by a morphism.

Samples of counting patterns in P_n results, [SK, Mansour, Séébold, 2003]

$$12(P_{2k+1}) = \tfrac{2}{5}(4 \cdot 16^k + 1),$$

$$12(P_{2k+2}) = \tfrac{2}{5}(16^{k+1} - 1),$$

$$21(P_{2k+1}) = \tfrac{8}{5}(16^k - 1),$$

$$21(P_{2k+2}) = \tfrac{2}{5}(16^{k+1} - 1).$$

The number of occurrences of the pattern $(1-)^\ell$ in P_n is

$$\binom{4^{n-1} - 2^{n-1}}{\ell} + 2\binom{4^{n-1}}{\ell} + \binom{4^{n-1} + 2^{n-1} - 1}{\ell}.$$

Counting ordered patterns in words generated by morphisms [SK, Mansour, Séébold, 2008]

$\mathcal{A} = \{a_1 < a_2 < \cdots < a_k\}$. Let f be a morphism and $n \geq 0$. The *incidence matrix* of f^n is the $k \times k$ matrix

$$M(f^n) = (m_{n,i,j})_{1 \leq i,j \leq k}$$

where $m_{n,i,j}$ is the number of occurrences of the letter a_i in the word $f^n(a_j)$.

It is known that $M(f)^n = M(f^n)$.

The *vector of non-inversions* 1-2$(f^n) = (|f^n(a_i)|_{1-2})_{1\leq i\leq k}$.

The *vector of inversions* 2-1$(f^n) = (|f^n(a_i)|_{2-1})_{1\leq i\leq k}$.

The *vector of p-repetitions with gaps* of a letter $R_pG(f^n) = (|f^n(a_i)|_{(1-)^p})_{1\leq i\leq k}$.

Using $f^{n+1} = f^n \circ f = f \circ f^n$, we get the following result.

For each letter $a_\ell \in A$, let p_ℓ and q_ℓ be such that $f(a_\ell) = a_{\ell_1} \ldots a_{\ell_{p_\ell}}$ and $f^n(a_\ell) = a_{\ell'_1} \ldots a_{\ell'_{q_\ell}}$. Then, for all $n \in \mathbb{N}$,

$$|f^{n+1}(a_\ell)|_{1-2} = \sum_{1\le i<j\le p_\ell} \left(\sum_{r=1}^{k-1} (m_{n,r,\ell_i} \cdot \sum_{s=r+1}^{k} m_{n,s,\ell_j})\right) + \sum_{t=1}^{k} |f^n(a_t)|_{1-2} \cdot m_{1,t,\ell},$$

$$= \sum_{1\le i<j\le q_\ell} \left(\sum_{r=1}^{k-1} (m_{1,r,\ell'_i} \cdot \sum_{s=r+1}^{k} m_{1,s,\ell'_j})\right) + \sum_{t=1}^{k} |f(a_t)|_{1-2} \cdot m_{n,t,\ell},$$

$$|f^{n+1}(a_\ell)|_{2-1} = \sum_{1\le i<j\le p_\ell} \left(\sum_{r=2}^{k} (m_{n,r,\ell_i} \cdot \sum_{s=1}^{r-1} m_{n,s,\ell_j})\right) + \sum_{t=1}^{k} |f^n(a_t)|_{2-1} \cdot m_{1,t,\ell},$$

$$= \sum_{1\le i<j\le q_\ell} \left(\sum_{r=2}^{k} (m_{1,r,\ell'_i} \cdot \sum_{s=1}^{r-1} m_{1,s,\ell'_j})\right) + \sum_{t=1}^{k} |f(a_t)|_{2-1} \cdot m_{n,t,\ell}.$$

The following is obvious.

For each letter $a_\ell \in A$ and for all $n \in \mathbb{N}$,

$$|f^n(a_\ell)|_{(1-)^p} = \sum_{t=1}^{k} \binom{m_{n,t,\ell}}{p}.$$

The Thue-Morse morphism

$\mu(0) = 01,$
$\mu(1) = 10.$

$$M(\mu^n) = \begin{bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{bmatrix}$$

For any integer $n \geq 2$,

$$\text{1-2}(\mu^n) = \text{2-1}(\mu^n) = \begin{bmatrix} 2^{2n-3} & 2^{2n-3} \end{bmatrix} \text{ and}$$

$$R_p G(\mu^n) = \begin{bmatrix} 2 \cdot \binom{2^{n-1}}{p} & 2 \cdot \binom{2^{n-1}}{p} \end{bmatrix}$$

The Fibonacci morphism

$\phi(a_1) = a_1 a_2,$
$\phi(a_2) = a_1.$

It generates *Fibonacci sequence* $\varphi^\omega(a_1)$.

$$M(\varphi^n) = \begin{bmatrix} F_n & F_{n-1} \\ F_{n-1} & F_{n-2} \end{bmatrix}$$

For every integer $n \geq 0$,

$$|\varphi^{n+2}(a_1)|_{\text{2-1}} = \sum_{p=0}^{n} F_p F_{n-p}^2 \,,$$

$$|\varphi^{n+2}(a_1)|_{\text{1-2}} = |\varphi^{n+2}(a_1)|_{\text{2-1}} + F_n + \begin{cases} 1 & \text{if } n \text{ is odd,} \\ -1 & \text{if } n \text{ is even.} \end{cases}$$

A particular family of morphisms

f involving at least 2 letters has the following properties:

1. $\exists$ a positive integer m such that $|f(a_1)|_{a_i} = m$, $1 \leq i \leq k$,

2. $\exists$ a positive integer d such that $|f(a_2 \ldots a_k)|_{a_i} = d$, $1 \leq i \leq k$,

3. $\forall\ i, j$, $1 \leq i, j \leq k$, $|f(a_i a_j)|^{ext}_{\text{1-2}} = |f(a_j a_i)|^{ext}_{\text{1-2}}$. [For example, for Thue-Morse morphism, $|\mu(a_1 a_2)|^{ext}_{\text{1-2}} = |a_1 a_2 a_2 a_1|^{ext}_{\text{1-2}} = 1$ $= |a_2 a_1 a_1 a_2|^{ext}_{\text{1-2}} = |\mu(a_2 a_1)|^{ext}_{\text{1-2}}$]

For every positive integer n,

$$
\begin{aligned}
|f^{n+1}(a_1)|_{\textsf{1-2}} = \quad & m(d+m)^{n-1} \textstyle\sum_{i=1}^{k} |f(a_i)|_{\textsf{1-2}} \\
+ \quad & \tfrac{[m(d+m)^{n-1}-1]m(d+m)^{n-1}}{2} \textstyle\sum_{j=1}^{k} |f(a_j a_j)|_{\textsf{1-2}}^{ext} \\
+ \quad & m^2(d+m)^{2n-2} \textstyle\sum_{1\le i<j\le k} |f(a_i a_j)|_{\textsf{1-2}}^{ext}
\end{aligned}
$$

$$
\begin{aligned}
|f^{n+1}(a_2 \dots a_k)|_{\textsf{1-2}} = \quad & d(d+m)^{n-1} \textstyle\sum_{i=1}^{k} |f(a_i)|_{\textsf{1-2}} \\
+ \quad & \tfrac{[d(d+m)^{n-1}-1]d(d+m)^{n-1}}{2} \textstyle\sum_{j=1}^{k} |f(a_j a_j)|_{\textsf{1-2}}^{ext} \\
+ \quad & d^2(d+m)^{2n-2} \textstyle\sum_{1\le i<j\le k} |f(a_i a_j)|_{\textsf{1-2}}^{ext}
\end{aligned}
$$

The Istrail morphism, 1977

The morphism h on $A = \{a_1 < a_2 < a_3\}$:

$$h(a_1) = a_1a_2a_3, \qquad h(a_2) = a_1a_3, \qquad h(a_3) = a_2$$

h generates a square-free infinite word, $h^\omega(a_1)$, but is not a square-free morphism: $h(a_1a_2a_1) = a_1a_2a_3a_1a_3a_1a_2a_3$ contains $a_3a_1a_3a_1$.

The word $h^\omega(a_1)$ is closely related to the Thue-Morse word T. If

$$\begin{aligned} \delta : \quad a_1 &\mapsto a_1 \\ a_2 &\mapsto a_1a_2 \\ a_3 &\mapsto a_1a_2a_2 \end{aligned}$$

then $T = \delta(h^\omega(a_1))$ (Lothaire, 1983).

The Istrail morphism, 1977

For $n \geq 1$,

$$|h^{n+1}(a_1)|_{\textsf{1-2}} = |h^{n+1}(a_2 a_3)|_{\textsf{1-2}} = 3 \cdot 2^{2n-1} + 2^n.$$

For $n \geq 1$,

$$|h^{n+1}(a_1)|_{\textsf{2-1}} = |h^{n+1}(a_2 a_3)|_{\textsf{2-1}} = 3 \cdot 2^{2n-1} - 2^n.$$

The Prouhet morphisms, 1851

(A generalization of Thue-Morse morphism)

Let $k \geq 2$ and $A = \{a_1 < \cdots < a_k\}$. The *Prouhet morphism* π_k is

$$\pi_k(a_i) = a_i a_{i+1} \ldots a_k a_1 \ldots a_{i-1}, \qquad 1 \leq i \leq k.$$

Let $k = 6$. The morphism π_6 is given by

$$\begin{array}{lcl}
a_1 & \mapsto & a_1 a_2 a_3 a_4 a_5 a_6 \\
a_2 & \mapsto & a_2 a_3 a_4 a_5 a_6 a_1 \\
a_3 & \mapsto & a_3 a_4 a_5 a_6 a_1 a_2 \\
a_4 & \mapsto & a_4 a_5 a_6 a_1 a_2 a_3 \\
a_5 & \mapsto & a_5 a_6 a_1 a_2 a_3 a_4 \\
a_6 & \mapsto & a_6 a_1 a_2 a_3 a_4 a_5
\end{array}$$

The Prouhet morphisms, 1851

For every i, $1 \leq i \leq k$, and for every positive integer n,

$$|\pi_k^{n+1}(a_i)|_{\text{1-2}} = \frac{(k-1)k^n}{12}\left(3k^{n+1} + k - 2\right),$$

$$|\pi_k^{n+1}(a_i)|_{\text{2-1}} = \frac{(k-1)k^n}{12}\left(3k^{n+1} - k + 2\right).$$

For example,

$$|\pi_6^{n+1}(a_i)|_{\text{1-2}} = \tfrac{5\cdot 6^n}{12}\left(3\cdot 6^{n+1} + 6 - 2\right) = 6^{n-1}\cdot(45\cdot 6^n + 10),$$

$$|\pi_6^{n+1}(a_i)|_{\text{2-1}} = 6^{n-1}\cdot(45\cdot 6^n - 10).$$

The Arshon morphisms, β_k^n, 1937

$\mathbf{A} = \{1, 2, \ldots, k\}$.

Let $w_1 = 1$. For $n \geq 1$, w_{n+1} is obtained by replacing the letters of w_n:

in odd positions	in even positions
$1 \to 123\ldots(k-1)k$	$1 \to k(k-1)\ldots 321$
$2 \to 234\ldots(k-1)k1$	$2 \to 1k(k-1)\ldots 432$
...	...
$k \to k12\ldots(k-2)(k-1)$	$k \to (k-1)(k-2)\ldots 21k$

Then $w_2 = 123\ldots(k-1)k$ and each w_i is the initial subword of w_{i+1}, so $w = \lim\limits_{i\to\infty} w_i$ is well defined.

When $k = 3$, w is called the *Arshon sequence*.

$\mathbf{A} = \{1, 2, 3, 4\}$.

Let $w_1 = 1$. For $n \geq 1$, w_{n+1} is obtained by replacing the letters of w_n:

in odd positions	in even positions
$1 \rightarrow 1234$	$\underline{1 \rightarrow 4321}$
$\underline{2 \rightarrow 2341}$	$2 \rightarrow 1432$
$3 \rightarrow 3412$	$\underline{3 \rightarrow 2143}$
$\underline{4 \rightarrow 4123}$	$4 \rightarrow 3214$

Then $w_2 = 1234$, $w_3 = 1234143234123214$, ...

Theorem. [Berstel 1979, SK, 2003] There does not exist a morphism, whose iteration defines the Arshon sequence.

Theorem. [Currie, 2002] No Arshon sequence of odd order can be generated by an iterated morphism.

The Arshon morphisms, β_k^n, 1937

Let k be any even positive integer. For every i, $1 \le i \le k$, and for every positive integer n,

$$|\beta_k^{n+1}(a_i)|_{\text{1-2}} = \frac{k^{n-1}}{4}\left[k^{n+2}\cdot(k-1)+2k\right],$$

$$|\beta_k^{n+1}(a_i)|_{\text{2-1}} = \frac{k^{n-1}}{4}\left[k^{n+2}\cdot(k-1)-2k\right].$$

For example,

$$\begin{aligned}|\beta_6^{n+1}(a_i)|_{\text{1-2}} &= \tfrac{6^{n-1}}{4}\cdot\left(6^{n+2}\cdot 5+2\cdot 6\right)\\ &= 6^{n-1}\cdot(45\cdot 6^n+3),\end{aligned}$$

$$|\beta_6^{n+1}(a_i)|_{\text{2-1}} = 6^{n-1}\cdot(45\cdot 6^n-3).$$

More examples of morphisms satisfying the three conditions, but not linked with Thue-Morse sequence:

$$
\begin{aligned}
f : a_1 &\mapsto a_1 a_3 a_2 a_4 \\
a_2 &\mapsto \varepsilon \\
a_3 &\mapsto a_1 a_4 \\
a_4 &\mapsto a_2 a_3
\end{aligned}
$$

$$
\begin{aligned}
|f^{n+1}(a_1)|_{\text{1-2}} &= |f^{n+1}(a_3 a_4)|_{\text{1-2}} = 3 \cdot 2^{n-1} \cdot (2^{n+1} + 1), \\
|f^{n+1}(a_1)|_{\text{2-1}} &= |f^{n+1}(a_3 a_4)|_{\text{2-1}} = 3 \cdot 2^{n-1} \cdot (2^{n+1} - 1), \\
|f^{n+1}(a_2)|_{\text{1-2}} &= |f^{n+1}(a_2)|_{\text{2-1}} = 0.
\end{aligned}
$$

More examples of morphisms satisfying the three conditions, but not linked with Thue-Morse sequence:

$$\begin{aligned} h : a &\mapsto aba\,cab\,cac\,bab\,cba\,cbc \\ b &\mapsto aba\,cab\,cac\,bca\,bcb\,abc \\ c &\mapsto aba\,cab\,cba\,cbc\,acb\,abc \end{aligned}$$

This morphism is square-free (Brandenburg, 1983)

For every $x \in A = \{a < b < c\}$ and for every positive integer n,

$$|h^{n+1}(x)|_{\text{1-2}} = 6 \cdot 18^{n-1} \cdot (9 \cdot 18^{n+1} + 40),$$

$$|h^{n+1}(x)|_{\text{2-1}} = 6 \cdot 18^{n-1} \cdot (9 \cdot 18^{n+1} - 40).$$

Consecutive patterns and morphisms

The *vector of rises* of f^n is

$$R(f^n) = (|f^n(a_i)|_{12})_{1 \le i \le k}.$$

The *vector of descents* of f^n is

$$D(f^n) = (|f^n(a_i)|_{21})_{1 \le i \le k}.$$

The *vector of squares of one letter* of f^n is

$$R_2(f^n) = (|f^n(a_i)|_{11})_{1 \le i \le k}.$$

We define two sequences of k vectors, $(F(f^n))_{n\in\mathbb{N}}$ and $(L(f^n))_{n\in\mathbb{N}}$, where $F(f^n)[i]$ is the first letter of $f^n(a_i)$ and $L(f^n)[i]$ is the last letter of $f^n(a_i)$ if $f^n(a_i) \neq \varepsilon$, and $F(f^n)[i] = L(f^n)[i] = 0$ if $f^n(a_i) = \varepsilon$.

These two sequences take their values in a finite set: they are ultimately periodic. Thus they can be computed *a priori* from f.

We define

$$C_n^{12}(i,j) = \begin{cases} 1, & \text{if } L(f^n)[i] < F(f^n)[j] \\ 0, & \text{if } L(f^n)[i] \geq F(f^n)[j]. \end{cases}$$

$$C_n^{21}(i,j) = \begin{cases} 1, & \text{if } L(f^n)[i] > F(f^n)[j] \\ 0, & \text{if } L(f^n)[i] \leq F(f^n)[j], \end{cases}$$

$$C_n^{11}(i,j) = \begin{cases} 1, & \text{if } L(f^n)[i] = F(f^n)[j] \\ 0, & \text{if } L(f^n)[i] \neq F(f^n)[j]. \end{cases}$$

For any morphism f on A, there exists the least integer M_f such that, for every $a \in A$, if $f^n(a) = \varepsilon$ for some n then $f^{M_f}(a) = \varepsilon$. By convention, if f is a nonerasing morphism then $M_f = 0$. The integer M_f is the *mortality exponent* of f.

For each letter $a_\ell \in A$, $f(a_\ell) = a_{\ell_1} \ldots a_{\ell_{p_\ell}}$, and for all $n \geq M_f$, let $\ell'_1 \ldots \ell'_{p'_\ell}$ be the subsequence of $\ell_1 \ldots \ell_{p_\ell}$ such that $f^{n+1}(a_\ell) = f^n(a_{\ell'_1} \ldots a_{\ell'_{p'_\ell}})$ and $f^n(a_{\ell'_i}) \neq \varepsilon$, $1 \leq i \leq p'_\ell$. Then

$$|f^{n+1}(a_\ell)|_{12} = \sum_{t=1}^{k} |f^n(a_t)|_{12} \cdot m_{1,t,\ell} + \sum_{i=1}^{p'_\ell - 1} C_n^{12}(\ell'_i, \ell'_{i+1}),$$

$$|f^{n+1}(a_\ell)|_{21} = \sum_{t=1}^{k} |f^n(a_t)|_{21} \cdot m_{1,t,\ell} + \sum_{i=1}^{p'_\ell - 1} C_n^{21}(\ell'_i, \ell'_{i+1}),$$

$$|f^{n+1}(a_\ell)|_{11} = \sum_{t=1}^{k} |f^n(a_t)|_{11} \cdot m_{1,t,\ell} + \sum_{i=1}^{p'_\ell - 1} C_n^{11}(\ell'_i, \ell'_{i+1}).$$

The Thue-Morse morphism

For any integer $n \geq 0$,

$$
\begin{aligned}
R(\mu^{2n}) &= \left[\begin{array}{cc} \frac{4^n-1}{3} & \frac{4^n-1}{3} \end{array}\right] = D(\mu^{2n}) = R_2(\mu^{2n}) \\
R(\mu^{2n+1}) &= \left[\begin{array}{cc} \frac{2(4^n-1)}{3}+1 & \frac{2(4^n-1)}{3} \end{array}\right] \\
D(\mu^{2n+1}) &= \left[\begin{array}{cc} \frac{2(4^n-1)}{3} & \frac{2(4^n-1)}{3}+1 \end{array}\right] \\
R_2(\mu^{2n+1}) &= \left[\begin{array}{cc} \frac{2(4^n-1)}{3} & \frac{2(4^n-1)}{3} \end{array}\right].
\end{aligned}
$$

The Fibonacci morphism

For any integer $n \geq 1$,

$$
\begin{aligned}
R(\varphi^n) &= \left[\begin{array}{cc} F_{n-1} & F_{n-2} \end{array}\right] \\
D(\varphi^{2n}) &= \left[\begin{array}{cc} F_{2n-1} & F_{2n-2}-1 \end{array}\right] = R_2(\varphi^{2n+1}) \\
R_2(\varphi^{2n}) &= \left[\begin{array}{cc} F_{2n-2}-1 & F_{2n-3} \end{array}\right] = D(\varphi^{2n-1}).
\end{aligned}
$$

Erasing morphisms

$$\begin{aligned} f(a_1) &= a_1a_3a_2a_4 \\ f(a_2) &= \varepsilon \\ f(a_3) &= a_1a_4 \\ f(a_4) &= a_2a_3 \end{aligned}$$

One has $M_f = 1$.

For any integer $n \geq 1$, $R_2(f^n) = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$ and

$$\text{if } n \text{ is even } \begin{cases} R(f^n) & = \begin{bmatrix} 2^n & 0 & \frac{2^{n+1}+1}{3} & \frac{2^n-1}{3} \end{bmatrix} \\ D(f^n) & = \begin{bmatrix} 2^n - 1 & 0 & \frac{2^{n+1}-2}{3} & \frac{2^n-4}{3} \end{bmatrix}, \end{cases}$$

$$\text{if } n \text{ is odd } \begin{cases} R(f^n) & = \begin{bmatrix} 2^n & 0 & \frac{2^{n+1}-1}{3} & \frac{2^n+1}{3} \end{bmatrix} \\ D(f^n) & = \begin{bmatrix} 2^n - 1 & 0 & \frac{2^{n+1}-4}{3} & \frac{2^n-2}{3} \end{bmatrix}. \end{cases}$$

Erasing morphisms

$$\begin{aligned} g(a_1) &= a_1a_2a_4a_3 \\ g(a_2) &= a_3 \\ g(a_3) &= \varepsilon \\ g(a_4) &= a_1a_2a_4 \end{aligned}$$

Here we have $M_g = 2$

$R(g) = \begin{bmatrix} 2 & 0 & 0 & 2 \end{bmatrix}$, $D(g) = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$,

$R_2(g) = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$, and, for any integer $n \geq 2$,

$$
\begin{aligned}
R(g^n) &= \left[\begin{array}{cccc} 2^n & 0 & 0 & 2^n \end{array} \right] \\
D(g^n) &= \left[\begin{array}{cccc} 2^{n-1}+2^{n-2}-1 & 0 & 0 & 2^{n-1}+2^{n-2}-1 \end{array} \right] \\
R_2(g^n) &= \left[\begin{array}{cccc} 2^{n-2} & 0 & 0 & 2^{n-2} \end{array} \right].
\end{aligned}
$$

Thank you for your attention!