On a group theoretic generalization of the Morse-Hedlund theorem

É. Charlier[1] S. Puzynina[2] L. Q. Zamboni[3]

[1]Université de Liège, Belgium

[2]Saint Petersburg State University, Russia

[3]Université Lyon 1, France

Outline

complexity functions of infinite words:

relations

$\Leftrightarrow$

periodicity

minimal complexity of aperiodic words:

Outline

complexity functions of infinite words:
factor,
abelian,
cyclic,
group
. . .

relations
$\Leftrightarrow$

periodicity

minimal complexity of aperiodic words:

Morse-Hedlund theorem and Sturmian words

Outline

complexity functions of infinite words:
factor,
abelian,
cyclic,
group
. . .

relations
$\Leftrightarrow$

periodicity

minimal complexity of aperiodic words:

Morse-Hedlund theorem and Sturmian words

- New notion of complexity by group actions.
- Broad generalization of Morse-Hedlund theorem via group complexity.

Complexity and periodicity

(Ultimately) periodic word x:

$$x = uvvvvv \cdots$$

Aperiodic word = not ultimately periodic.

Connection between periodicity and complexity:

Theorem of Morse and Hedlund, 1940

Let x be an infinite word

- x aperiodic $\Rightarrow \forall n$: $p_x(n) \geq n+1$
- $\forall n$: $p_x(n) = n+1 \Leftrightarrow x$ is Sturmian word

Generalizations and modifications

Generalizations of the notion of words complexity:

- abelian complexity
- maximal pattern complexity
- arithmetical complexity
- conjugate complexity
- etc.

Abelian complexity

- Two finite words are abelian equivalent if they contain the same numbers of occurrences of each letter: $00111 \sim_{ab} 01101$

Abelian complexity

- Two finite words are abelian equivalent if they contain the same numbers of occurrences of each letter: $00111 \sim_{ab} 01101$
- The abelian complexity $a_w(n)$ of an infinite word w is the function that counts the number of classes of abelian equivalence of its factors length n

Abelian complexity

- Two finite words are abelian equivalent if they contain the same numbers of occurrences of each letter: $00111 \sim_{ab} 01101$
- The abelian complexity $a_w(n)$ of an infinite word w is the function that counts the number of classes of abelian equivalence of its factors length n

Example (Thue-Morse word)

$$t = 0110100110010110\cdots$$

The abelian complexity of the Thue-Morse word t is

$$a_t(n) = \begin{cases} 2 & \text{if } n \text{ is odd} \\ 3 & \text{if } n \text{ is even} \end{cases}$$

E.g., we have two abelian classes of factors of length 3: $\{001, 010, 100\}$, $\{011, 101, 110\}$.

Abelian complexity

Properties of abelian complexity

- $a_x(n) \leq \binom{n+|\Sigma|-1}{|\Sigma|-1} = O(n^{|\Sigma|-1})$.
- Ultimate periodicity $\Rightarrow$ bounded abelian complexity.
- The converse is not true: e.g., Thue-Morse word is aperiodic and has abelian complexity bounded by 3.

Abelian complexity

Properties of abelian complexity

- $a_x(n) \le \binom{n+|\Sigma|-1}{|\Sigma|-1} = O(n^{|\Sigma|-1})$.
- Ultimate periodicity $\Rightarrow$ bounded abelian complexity.
- The converse is not true: e.g., Thue-Morse word is aperiodic and has abelian complexity bounded by 3.

Relations between periodicity and abelian complexity:

Theorem (abelian Morse-Hedlund)

Let x be an infinite word.

- *x aperiodic $\Rightarrow \forall n$: $a_x(n) \ge 2$.*
- *x aperiodic, $\forall n$ $a_x(n) = 2 \Leftrightarrow x$ is Sturmian.*

Cyclic complexity

- Two finite words u and v are conjugate if there exist words w_1, w_2 such that $u = w_1 w_2$ and $v = w_2 w_1$ (e.g., *ababba* and *babbaa*).
- The cyclic complexity $c_x(n)$ of a word x is the function counting the number of conjugacy classes of length n of x for each $n \geq 0$.

Example (Thue-Morse word)

$$t = 0110100110010110 \cdots$$

We have four conjugacy classes of length 4:
$\{0010, 0100\}$,
$\{0110, 1001, 1100, 0011\}$,
$\{0101, 1010\}$,
$\{1011, 1101\}$.

Extension of Morse-Hedlund Theorem:

Theorem (Cassaigne, Fici, Sciortino, Zamboni, 2014)

Ultimate periodicity $\Leftrightarrow$ *bounded cyclic complexity.*

Minimal cyclic complexity and Sturmian words

Extension of Morse-Hedlund Theorem:

Theorem (Cassaigne, Fici, Sciortino, Zamboni, 2014)

Ultimate periodicity $\Leftrightarrow$ *bounded cyclic complexity.*

- $c_x(n) = 1$ for some $n \geq 1 \Rightarrow$ periodicity
- consider $\liminf c_x(n)$.

$\liminf c_x(n)$ and Sturmian words

- For Sturmian words $\limsup c_x(n) = \infty$, but $\liminf c_x(n) = 2$.
- This is not a characterization of Sturmian words: for example, for the period-doubling word $\liminf c_x(n) = 2$.

Generalization via group actions

$G \leq S_n$ subgroup of a symmetric group
G-action on $\{1, 2, \ldots, n\}$ is given by $g : i \mapsto g(i)$

G acts on words of length n by permutation of symbols:

For $g \in G$, $u \in \Sigma^n$ we define the action by

$$g * u = u_{g^{-1}(1)} u_{g^{-1}(2)} \cdots u_{g^{-1}(n)}.$$

Example

$g = (123)(45)$, $g : 1 \mapsto 2, 2 \mapsto 3, 3 \mapsto 1, 4 \mapsto 5, 5 \mapsto 4$

$$abcab \overset{(123)(45)}{\curvearrowright} cabba$$

In particular we have $g * u \sim_{ab} u$ for all $g \in G$.

G-equivalence of words of length n:

u, v words of length n, $u \sim_G v$ if $\exists g \in G$ such that $g(u) = v$.

x infinite word

$\omega = (G_n)_{n\geq 1}$, $G_n \leq S_n$ a sequence of subgroups

The group complexity $p_{\omega,x}(n)$ of x is the function which counts the number of classes of G_n-equivalence of factors of length n.

Example (Thue-Morse word)

$$t = 0110100110010110\cdots$$

For $G_4 = \langle (13),(24) \rangle$ we have $p_{\omega,t}(4) = 7$.

We have six singleton classes of length 4:

$$[0010], [0100], [0101], [1010], [1011], [1101]$$

and one class of order 4:

$$[0110 \overset{(13)(24)}{\curvearrowright} 1001 \overset{(24)}{\curvearrowright} 1100 \overset{(13)}{\curvearrowright} 0011].$$

Group actions: generalization of factor, abelian and cyclic complexities

Particular cases:

- factor complexity: $\omega = (Id_n)_{n\geq 1}$, $p_{\omega,x}(n) = p_x(n)$
- abelian complexity: $\omega = (S_n)_{n\geq 1}$, $p_{\omega,x}(n) = a_x(n)$
- cyclic complexity: $\omega =< (12\cdots n) >_{n\geq 1}$, $p_{\omega,x}(n) = c_x(n)$

Remark

Group and cyclic complexities are between abelian and classic complexity:

$$a_x(n) \leq p_{\omega,x}(n), c_x(n) \leq p_x(n).$$

Complexity by group actions: $\varepsilon(G)$

$G \leq S_n$

We consider the G-action on $\{1, 2, \ldots, n\}$ given by $g : i \mapsto g(i)$
$G(i) = \{g(i) \mid g \in G\}$ denotes the G-orbit of i.

Let $\varepsilon(G)$ denote the number of distinct G-orbits:

$$\varepsilon(G) = \sharp\{G(i) \mid i \in \{1, 2, \ldots, n\}\}$$

Example

For $n = 6$, $G = \langle (13), (256) \rangle$, we have $\varepsilon(G) = 3$:

123456

- If $G = Id$, then $\varepsilon(G) = n$.
- If G contains an n-cycle, then $\varepsilon(G) = 1$.

$G \leq S_n$

$\varepsilon(G)$: the number of G-orbits of $\{1, \ldots, n\}$.

Example

Klein group $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$:

$$G = \{\epsilon, (12), (34), (12)(34)\}$$

G-orbits : $\{\{12\}, \{34\}\} \Rightarrow \varepsilon(G) = 2$

$$G' = \{\epsilon, (12)(34), (13)(24), (14)(23)\}$$

G'-orbit : $\{1, 2, 3, 4\} \Rightarrow \varepsilon(G') = 1$

$\varepsilon(G)$ depends on the embedding of G into S_n!

Generalisation of theorem of Morse and Hedlund

Theorem 1, Charlier, P., Zamboni, 2017

Let x be an infinite word, $\omega = (G_n)_{n\geq 1}$, $G_n \leq S_n$.

- x aperiodic $\Rightarrow \forall n\ p_{\omega,x}(n) \geq \varepsilon(G_n) + 1$
- $\forall n\ p_{\omega,x}(n) = \varepsilon(G_n) + 1 \Rightarrow x$ Sturmian.

Theorem 2, Charlier, P., Zamboni, 2017

Let x be a Sturmian word, $\omega = (G_n)_{n\geq 1}$, $G_n \leq S_n$ abelian, then $\exists \omega' = (G'_n)_{n\geq 1}$, $G'_n \leq S_n$ isomorphic to G_n: $p_{\omega',x}(n) = \varepsilon(G'_n) + 1$.

Particular cases:

- Theorem of Morse and Hedlund: $\omega = (Id_n)_{n\geq 1}$, $p_{\omega,x}(n) = p_x(n)$, $\varepsilon(G_n) = n$
- abelian complexity: $\omega = (S_n)_{n\geq 1}$, $p_{\omega,x}(n) = a_x(n)$, $\varepsilon(G_n) = 1$

Theorem 2: We cannot always take $G' = G$

Theorem 2, Charlier, P., Zamboni, 2017

Let x be a Sturmian word, $\omega = (G_n)_{n \geq 1}$, $G_n \leq S_n$ abelian, then $\exists \omega' = (G'_n)_{n \geq 1}$, $G'_n \leq S_n$ isomorphic to G_n: $p_{\omega',x}(n) = \varepsilon(G'_n) + 1$.

Example

$G_4 = \langle \sigma \rangle$, $\sigma = (1234)$, $\varepsilon(G_4) = 1$, Fibonacci word:

$$F = 0100101001001010010100100101 0010 \cdots$$

$$\{[0100 \overset{\sigma}{\curvearrowright} 0010], [0101 \overset{\sigma}{\curvearrowright} 1010], [1001]\}$$

$$p_{\omega,F}(4) = 3 > \varepsilon(G_4) + 1 = 2$$

But we can take $G'_4 = \langle \sigma' \rangle$, $\sigma' = (1324)$, $\varepsilon(G'_4) = 1$.

$$\{[0010 \overset{\sigma'}{\curvearrowright} 0010], [0101 \overset{\sigma'}{\curvearrowright} 1001 \overset{\sigma'}{\curvearrowright} 1010]\}$$

$$p_{\omega,F}(4) = 2 = \varepsilon(G'_4) + 1 = 2$$

Theorem 2: we cannot replace "isomorphic" by "conjugate"

Theorem 2, Charlier, P., Zamboni, 2017

Let x be a Sturmian word, $\omega = (G_n)_{n\geq 1}$, $G_n \leq S_n$ abelian, then $\exists \omega' = (G'_n)_{n\geq 1}$, $G'_n \leq S_n$ isomorphic to G_n: $p_{\omega',x}(n) = \varepsilon(G'_n) + 1$.

We cannot replace "isomorphic" by "conjugate":

Example

$G = \langle (123)(456) \rangle \leq S_6$ cyclic of order 3.
Then $\varepsilon(G) = 2$.
We can show that if x is the Fibonacci word, then

$$\mathrm{Card}\left(\mathrm{Fact}_x(6)/\sim_{G'}\right) \geq 4$$

for each subgroup G' of S_6 conjugate to G.

Corollary: Conjugate

But in case of relatively prime cycle lengths we can:

Corollary

Let $\sigma \in S_n$ and $G = \langle \sigma \rangle$. Writing $\sigma = \sigma_1 \cdots \sigma_k$ as a product of disjoint cycles, suppose $|\sigma_1|, \ldots, |\sigma_k|$ are pairwise relatively prime. Then for every Sturmian word x there exists $G' \leq S_n$ conjugate to G such that $Card(Fact_x(n)/\sim_{G'}) = \varepsilon(G) + 1$.

Theorem 1.2

x aperiodic $\forall n\ p_{\omega,x}(n) = \varepsilon(G_n) + 1 \Rightarrow x$ Sturmian.

We show that x is binary and balanced (hence Sturmian).

x is balanced: For $u, v \in F(x)$ with $|u| = |v|$ the numbers of occurrences of 0 in u and v differs by at most 1.

Since $\varepsilon(G_1) = 1$, then $p_{\omega,x}(1) = 2$, and hence x is binary.

We use:

Lemma

Let $x \in \{0,1\}^{\mathbb{N}}$ be aperiodic. Then either x is Sturmian or there exist an integer $n \geq 2$, a Sturmian word y and a bispecial factor $u \in \{0,1\}^{n-2}$ of y such that $Fact_x(n) = Fact_y(n) \cup \{0u0, 1u1\}$.

v is a bispecial factor of u if $v0$, $v1$, $0v$, $1v$ are factors of u.

Theorem 2: first construct a cycle

Theorem 2

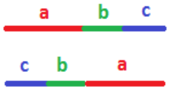
x Sturmian, $\omega = (G_n)_{n\geq 1}$, $G_n \leq S_n$ abelian $\Rightarrow \exists \omega' = (G'_n)_{n\geq 1}$, $G'_n \leq S_n$ isomorphic to G_n: $p_{\omega',x}(n) = \varepsilon(G'_n) + 1$.

First we prove Theorem 2 for a cycle.

abc-permutation

The numbers $1, 2, \ldots, n$ are divided into three subintervals of length a, b and c which are rearranged in the order c, b, a:

$$1, 2, \ldots, n \mapsto c+b+1, c+b+2, \ldots, n, c+1, c+2, \ldots, c+b, 1, 2, \ldots, c$$



In other words: A discrete 3-interval exchange transformation $(a, b, c) \to (c, b, a)$ on $\{1, 2, \ldots, n\}$ (where $n = a + b + c$).

Theorem 2: construction of cycle

To construct the cycle as (abc)-permutation, we make use of:

- bispecial factors of x of the next possible length.
- smallest periods of the bispecial factors to calculate a, b and c.
- prove that the (abc)-permutation is an n-cycle using [Pak, Redlich, 2008].
- Lexicographic arrays for the proofs.

Theorem 2: example of cycle

Example

$m = 6$ in Fibonacci word

- $w = 010010$ bispecial
- $p = 5$ and $q = 3$ periods of w
- $a = 1$, $b = 2$, $c = 3$ define the *abc*-permutation

0	0	1	0	0	1
0	0	1	0	1	0
0	1	0	0	1	0
0	1	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	1
1	0	1	0	0	1

Theorem 2: example of cycle

Example

$m = 6$ in Fibonacci word

- $w = 010010$ bispecial
- $p = 5$ and $q = 3$ periods of w
- $a = 1$, $b = 2$, $c = 3$ define the *abc*-permutation

0	0	1	0	0	1
0	0	1	0	1	0
0	1	0	0	1	0
0	1	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	1
1	0	1	0	0	1

Theorem 2: example of cycle

Example

$m = 6$ in Fibonacci word

- $w = 010010$ bispecial
- $p = 5$ and $q = 3$ periods of w
- $a = 1$, $b = 2$, $c = 3$ define the *abc*-permutation

0	0	1	0	0	1
0	0	1	0	1	0
0	1	0	0	1	0
0	1	0	1	0	0
1	0	0	1	0	0
1	0	0	1	0	1
1	0	1	0	0	1

Theorem 2: Corollary on cycles

Corollary

Let $x \in \{0,1\}^{\mathbb{N}}$ be a Sturmian word. Then for each positive integer n there exists a cyclic group G_n generated by an n-cycle such that $Card(Fact_x(n)/\sim_{G_n}) = 2$.

Remark [Cassaigne, Fici, Sciortino, Zamboni, 2017]

In contrast, if we set $G_n = \langle(1,2,\ldots,n)\rangle$, then $\limsup_{n\to\infty} p_{\omega,x} = +\infty$, while $\liminf_{n\to\infty} p_{\omega,x} = 2$.

Theorem 2: construction for abelian groups

Theorem (Fundamental Theorem of Finite Abelian Groups)

Every finite abelian group G can be written as a direct product of cyclic groups $\mathbb{Z}/m_1\mathbb{Z} \times \mathbb{Z}/m_2\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}$ where the m_i are prime powers.

- The unordered sequence $(m_1, m_2, \ldots, m_k)$ determines G up to isomorphism.
- The trace of G is given by $T(G) = m_1 + m_2 + \cdots + m_k$.

Proposition (Hoffman, 1987)

If an Abelian group G is embedded in S_n, then $T(G) \leq n$.

Question: non-abelian groups

Does Theorem 2 hold for non-abelian groups?

Question

Let x be a Sturmian word, $\omega = (G_n)_{n\geq 1}$, $G_n \leq S_n$. Does there exist $\omega' = (G'_n)_{n\geq 1}$, $G'_n \leq S_n$ isomorphic to G_n: $p_{\omega',x}(n) = \varepsilon(G'_n) + 1$?

Sturmian words and minimal complexity

- For most complexity functions Sturmian words possess minimal complexity;
- In some cases it give a characterization of Sturmian words, but not always.

For which complexities Sturmian words have minimal complexity? When are they the only ones?

Pattern complexity

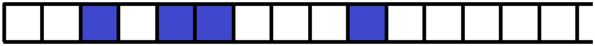
x an infinite word

Maximal pattern complexity $p_x^*(n)$ is defined by

$$p_x^*(n) = \sup_{\tau} \sharp\{x_{k+\tau(0)}x_{k+\tau(1)}\cdots x_{k+\tau(n-1)} | k = 0, 1, 2, \ldots\},$$

where the supremum is taken over all sequences of integers $\tau(0), \tau(1), \cdots \tau(n-1)$ of length n.

Pattern (0,2,3,7):



Minimal maximal pattern complexity

Theorem (Kamae, Zamboni, 2002)

An infinite word x is aperiodic if and only if $p^*_x(n) \geq 2n$ for every $n = 1, 2, \ldots$.

Minimal maximal pattern complexity

Theorem (Kamae, Zamboni, 2002)

An infinite word x is aperiodic if and only if $p_x^*(n) \geq 2n$ for every $n = 1, 2, \ldots$.

Words of complexity $2n + 1$:

- Sturmian
- certain rotational words
- certain Toeplitz words
- ...

No characterization.

Arithmetical complexity

x an infinite word

Arithmetical complexity $ar_x(n)$ is defined by

$$ar_x(n) = \sharp\{x_k x_{k+d} \cdots x_{k+d(n-1)} | k = 0, 1, 2, \dots, d = 1, 2, \dots\}.$$

I.e., arithmetical complexity counts the number of subwords in arithmetic progressions.

Minimal arithmetical complexity

- Minimal arithmetical complexity of aperiodic uniformly recurrent words is linear.
- Words of asymptotically minimal arithmetical complexity are Toeplitz words.
- Arithmetical complexity of Sturmian words is $\Theta(n^3)$.

[Avgustinovich, Cassaigne, Frid, 2006]

Minimal complexity

complexity type	minimal complexity	words family
factor	n+1	Sturmian
abelian	2	Sturmian
cyclic	$\liminf = 2$	Sturmian+
group	$\varepsilon(G_n) + 1$	Sturmian
maximal pattern	2n+1	Sturmian+
arithmetical	linear	(asymptotically) Toeplitz