Neutral sets of arbitrary characteristic

Francesco Dolce

Felice e Lieto Anniversario al
migliore maestro e amico

In honour of Antonio Restivo

Napoli, 14 Gennaio 2016

Joint work with

Dominique Perrin

Overview

- Families of sets of words verifying a “classical” property (neutrality) and a less classical one (tree condition).
- Common generalization of Arnoux-Rauzy languages and interval exchange sets.
- Enumeration formulæ in these sets for bifix codes and return words.
- Unexpected result about recurrence and uniformly recurrence.

Outline

Introduction

0. Neutral sets

1. Bifix codes

2. Return words

3. Bifix decoding

Conclusions

Outline

Let A a finite alphabet and S be a *factorial* set on A.

For a word $w \in S$, we denote

$$\begin{array}{lcllll} \ell(w) & = & \text{the number of letters} & a & \text{such that} & aw \in S, \\ r(w) & = & \text{the number of letters} & a & \text{such that} & wa \in S, \\ e(w) & = & \text{the number of pairs} & (a, b) & \text{such that} & awb \in S. \end{array}$$

A word w is *left-special* if $\ell(w) \geq 2$, *right-special* if $r(w) \geq 2$ and *bispecial* if it is both left and right-special.

Let A a finite alphabet and S be a *factorial* set on A.

For a word $w \in S$, we denote

$$
\begin{aligned}
\ell(w) &= \text{the number of letters } a \text{ such that } aw \in S, \\
r(w) &= \text{the number of letters } a \text{ such that } wa \in S, \\
e(w) &= \text{the number of pairs } (a, b) \text{ such that } awb \in S.
\end{aligned}
$$

A word w is *left-special* if $\ell(w) \geq 2$, *right-special* if $r(w) \geq 2$ and *bispecial* if it is both left and right-special.

The *multiplicity* of a word w is the quantity

$$m(w) = e(w) - \ell(w) - r(w) + 1.$$

A word is called *neutral* if $m(w) = 0$.

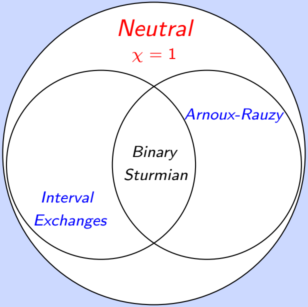
A set S is *neutral* if it is factorial and every word $w \in S \setminus \{\varepsilon\}$ is neutral.

The integer $\chi(S) = 1 - m(\varepsilon)$ ($= \ell(\varepsilon) + r(\varepsilon) - e(\varepsilon)$) is called the *characteristic* of S.

A set S is *neutral* if it is factorial and every word $w \in S \setminus \{\varepsilon\}$ is neutral.

The integer $\chi(S) = 1 - m(\varepsilon)$ ($= \ell(\varepsilon) + r(\varepsilon) - e(\varepsilon)$) is called the *characteristic* of S.

Proposition



Example (Fibonacci)

Factors of $\varphi^\omega(a) = abaababaaba\cdots$ of the morphism $\varphi : a \mapsto ab,\ b \mapsto a$.

The *factor complexity* of a factorial set $S \subset A^*$ is the sequence $p(n) = \text{Card}\,(S \cap A^n)$.

Proposition [*using* J. Cassaigne (1997)]

The factor complexity of a neutral set is given by $p(0) = 1$ and

$$p(n) = n(\text{Card}\,(A) - \chi(S)) + \chi(S).$$

Example

The Fibonacci set has factor complexity $p(n) = n + 1$.

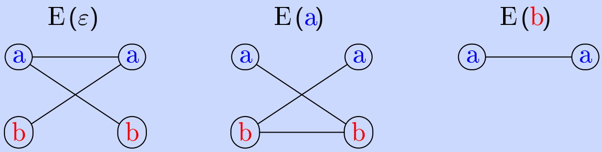
The *extension graph* of a word $w \in S$ is the undirected bipartite graph $G(w)$ with vertices the disjoint union of

$$L(w) = \{a \in A \mid aw \in S\} \quad \text{and} \quad R(w) = \{a \in A \mid wa \in S\},$$

and edges the pairs $E(w) = \{(a,b) \in A \times A \mid awb \in S\}$

Example (Fibonacci)

$S = \{\varepsilon, a, b, aa, ab, ba, aab, aba, baa, bab, \ldots\}$.



A factorial and biextendable set S is called a *tree set* of *characteristic* c if for any nonempty $w \in S$, the graph $E(w)$ is a tree and if $E(\varepsilon)$ is a union of c trees.

Example

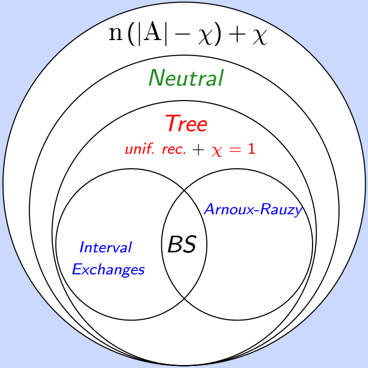
The Fibonacci set is a tree set of characteristic 1.

A factorial and biextendable set S is called a *tree set* of *characteristic* c if for any nonempty $w \in S$, the graph $E(w)$ is a tree and if $E(\varepsilon)$ is a union of c trees.

Example

The Fibonacci set is a tree set of characteristic 1.

Proposition [Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone (2014, 2015)]



Outline

A set $X \subset A^+$ of nonempty words over an alphabet A is a *bifix code* if it does not contain any proper prefix or suffix of its elements (as defined by *Sabrina* a couple of hours ago).

Example

- $\{aa, ab, ba\}$
- $\{aa, ab, bba, bbb\}$
- $\{ac, bcc, bcbca\}$

A set $X \subset A^+$ of nonempty words over an alphabet A is a *bifix code* if it does not contain any proper prefix or suffix of its elements (as defined by *Sabrina* a couple of hours ago).

Example

- $\{aa, ab, ba\}$
- $\{aa, ab, bba, bbb\}$
- $\{ac, bcc, bcbca\}$

$X \subset S$ is S-*maximal* if it is not properly contained in a bifix code $Y \subset S$ (more general than *Sabrina* and *Arturo*'s definition).

Example (Fibonacci)

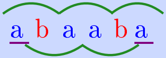
The set $X = \{aa, ab, ba\}$ is an S-maximal bifix code.
It is not an A^*-maximal bifix code, indeed $X \subset Y = X \cup \{bb\}$.

A *parse* of a word w w.r.t. a bifix code X is a triple (q, x, p) with $w = qxp$ and such that q has no suffix in X, $x \in X^*$ and p has no prefix in X.

Example

Let $X = \{aa, ab, ba\}$ and $w = abaaba$. The two possible parses of w are

- $(\varepsilon, ab \cdot aa \cdot ba, \varepsilon)$,
- $(a, ba \cdot ab, a)$.

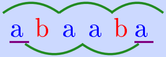


A *parse* of a word w w.r.t. a bifix code X is a triple (q,x,p) with $w = qxp$ and such that q has no suffix in X, $x \in X^*$ and p has no prefix in X.

Example

Let $X = \{aa, ab, ba\}$ and $w = abaaba$. The two possible parses of w are

- $(\varepsilon, ab \cdot aa \cdot ba, \varepsilon)$,
- $(a, ba \cdot ab, a)$.



The S-*degree* of X is the maximal number of parses w.r.t. X of a word of S.

Example

- For $S =$ Fibonacci, the set $X = \{aa, ab, ba\}$ has S-degree 2;
- The set $X = S \cap A^n$ has S-degree n.

Theorem [D., Perrin, (2016)]

Let S be a neutral set of characteristic χ.
For any finite S-maximal bifix code X of S-degree d, one has

$$\mathrm{Card}\,(X) = d\,(\mathrm{Card}\,(A) - \chi) + \chi.$$

Theorem [D., Perrin, (2016)]

Let S be a neutral set of characteristic χ.
For any finite S-maximal bifix code X of S-degree d, one has

$$\mathrm{Card}(X) = d(\mathrm{Card}(A) - \chi) + \chi.$$

Example (Fibonacci)

The set S-maximal bifix code $X = \{aa, ab, ba\}$ of S-degree 2 verifies

$$\mathrm{Card}(X) = 2(2-1) + 1.$$

Outline

A *complete return word* to a set $X \subset S$ is a word of S which has a proper prefix in X and a proper suffix in X and no internal factor in X.

The set $\mathcal{CR}(X)$ of complete return words to X is a bifix code.

Example (Fibonacci)

$\mathcal{R}(\{aa\}) = \{aabaa, aababaa\}$.

$$\varphi(a)^\omega = abaababaabaababaababaabaababaabaab\cdots$$

A *complete return word* to a set $X \subset S$ is a word of S which has a proper prefix in X and a proper suffix in X and no internal factor in X.

The set $\mathcal{CR}(X)$ of complete return words to X is a bifix code.

Example (Fibonacci)

$\mathcal{R}(\{aa\}) = \{aabaa, aababaa\}$.

$$\varphi(a)^\omega = abaababaabaababaababaabaababaabaab\cdots$$

Example

For any $n \geq 1$ one has $\mathcal{CR}(S \cap A^n) = S \cap A^{n+1}$.

Theorem [D., Perrin, (2016)]

Let S be a recurrent neutral set and $X \subset S$ a bifix code with empty kernel. We have

$$\mathrm{Card}\,(\mathcal{CR}(X)) = \mathrm{Card}\,(X) + \mathrm{Card}\,(A) - \chi.$$

Theorem [D., Perrin, (2016)]

Let S be a recurrent neutral set and $X \subset S$ a bifix code with empty kernel. We have

$$\text{Card}\,(\mathcal{CR}(X)) = \text{Card}\,(X) + \text{Card}\,(A) - \chi.$$

Corollary

For any $w \in S$, one has $\text{Card}\,(\mathcal{R}(w)) = \text{Card}\,(A) - \chi + 1$.

Theorem [D., Perrin, (2016)]

Let S be a recurrent neutral set and $X \subset S$ a bifix code with empty kernel. We have

$$\mathrm{Card}\,(\mathcal{CR}(X)) = \mathrm{Card}\,(X) + \mathrm{Card}\,(A) - \chi.$$

Corollary

For any $w \in S$, one has $\mathrm{Card}\,(\mathcal{R}(w)) = \mathrm{Card}\,(A) - \chi + 1$.

Remark.(recall *Aldo de Luca*'s talk)
A recurrent set S is uniformly recurrent $\iff$ $\mathcal{R}(w)$ is finite for every $w \in S$.

Corollary

A recurrent neutral set is uniformly recurrent.

Outline

A *coding morphism* for a bifix code $X \subset A^+$ is a morphism $f : B^* \to A^*$ which maps bijectively B onto X.

Example

Let's consider the bifix code $X = \{aa, ab, ba\}$ on $A = \{a, b\}$ and let $B = \{u, v, w\}$. The map

$$f : \begin{cases} u & \mapsto & aa \\ v & \mapsto & ab \\ w & \mapsto & ba \end{cases}$$

is a coding morphism for X.

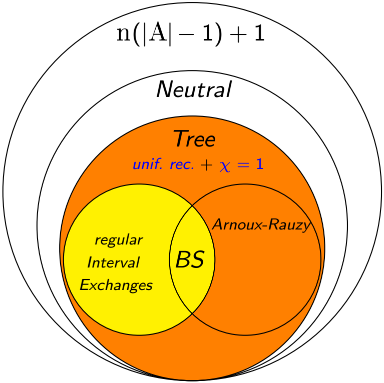
If S is factorial and X is an S-maximal bifix code, we call the set $f^{-1}(S)$ a *maximal bifix decoding* of S.

Example (Fibonacci, $\varphi^\omega(a) = ab\,aa\,ba\,ba\,ab\,ab\cdots$)

The set of factors of $f^{-1}(\varphi^\omega(a)) = vuwwvv\cdots$ is a maximal bifix decoding of S

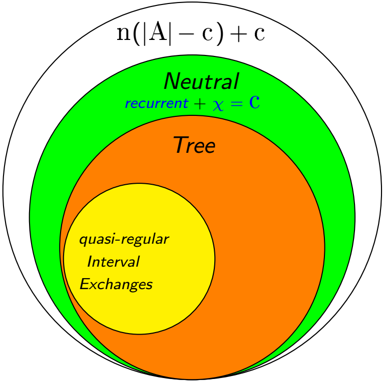
Theorem [Berthé, De Felice, D., Leroy, Perrin, Reutenauer, Rindone (2014, 2015)]

The family of uniformly recurrent tree sets of characteristic 1 is closed under maximal bifix decoding (and so is the family of regular interval exchange sets).



Theorem [D., Perrin, (2016)]

The family of *recurrent* *neutral* sets of characteristic c is closed under maximal bifix decoding (and so is the family of recurrent *tree* sets of characteristic c).



Further research directions

$\rightarrow$ Sets with a finite number of elements satisfying $\mathrm{m}(\mathrm{w}) \neq 0$.

$\rightarrow$ Bifix decoding for general bifix codes (and for a general neutral set).

$\rightarrow$ Return words and basis of the free group.

[S tree of $\chi = 1 \implies$ for every w, $\mathcal{R}(\mathrm{w})$ is a basis of F_{A}.]


Gràzie assàje
e...


...auguri Prof !