String attractors of Rote sequences

Veronika Hendrychová, Ľubomíra Dvořáková

Czech Technical University in Prague

February 26, 2024

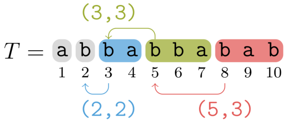
Outline

Outline

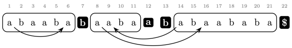
Motivation: Unifying repetitiveness measures

Motivation: Unifying repetitiveness measures

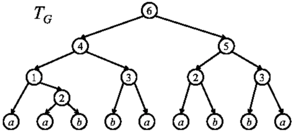
Dictionary compressors



Lempel-Ziv methods



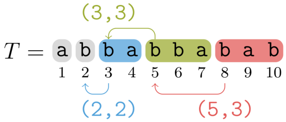
Pointer macro scheme



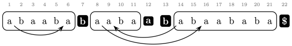
Grammars

Motivation: Unifying repetitiveness measures

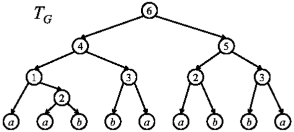
Dictionary compressors



Lempel-Ziv methods



Pointer macro scheme



Grammars

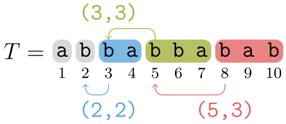
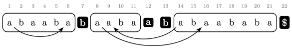
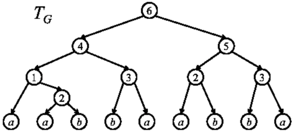
Induced repetitiveness measures

$\rightarrow z =$ size of the parsing

$\rightarrow b =$ size of the scheme

$\rightarrow g =$ size of the straight-line program

Motivation: Unifying repetitiveness measures

Dictionary compressors	Induced repetitiveness measures	Bounds via the smallest string attractors
 Lempel-Ziv methods	$\rightarrow z =$ size of the parsing	$z^* \in \mathcal{O}(\gamma^* \log^2(\frac{n}{\gamma^*}))$
 Pointer macro scheme	$\rightarrow b =$ size of the scheme	$b^* \in \mathcal{O}(\gamma^* \log(\frac{n}{\gamma^*}))$
 Grammars	$\rightarrow g =$ size of the straight-line program	$g^* \in \mathcal{O}(\gamma^* \log^2(\frac{n}{\gamma^*}))$ [Kempa & Prezza, 2018]

Motivation: Unifying repetitiveness measures

- Repetitiveness measures also upper bounds for the smallest string attractor

[Kempa & Prezza, STOC 2018]

Dictionary compressors can be interpreted as approximation algorithms for the smallest string attractor.

- Minimum size of an attractor gives us
  - lower and upper bounds for dictionary compression methods
  - direct stringological measure instead of the result of a specific compression method

Motivation: Unifying repetitiveness measures

- Repetitiveness measures also upper bounds for the smallest string attractor

[Kempa & Prezza, STOC 2018]

Dictionary compressors can be interpreted as approximation algorithms for the smallest string attractor.

- Minimum size of an attractor gives us
  - lower and upper bounds for dictionary compression methods
  - direct stringological measure instead of the result of a specific compression method
- Finding the smallest attractor size is NP-hard
  - $\rightarrow$ CoW approach:
    structural assumption (e.g., special classes of words) may make the computation tractable

Outline

String attractors: definition and example

Definition: Let $w = w_0 w_1 \dots w_n$ be a word, let $u = w_i w_{i+1} \dots w_j$ be its factor. Then $\{i, i+1, \dots, j\}$ is an *occurrence* of u in w.

String attractors: definition and example

Definition: Let $w = w_0 w_1 \ldots w_n$ be a word, let $u = w_i w_{i+1} \ldots w_j$ be its factor. Then $\{i, i+1, \ldots, j\}$ is an *occurrence* of u in w.

Definition of a string attractor [Prezza, ICTCS 2017]

Let $w = w_0 w_1 \ldots w_n$ be a finite word over alphabet $\mathcal{A}$. A *string attractor* of w is a set of positions $\Gamma \subseteq \{0, \ldots, n\}$ such that every substring of w has an occurrence containing an element of Γ.

String attractors: definition and example

Definition: Let $w = w_0 w_1 \dots w_n$ be a word, let $u = w_i w_{i+1} \dots w_j$ be its factor. Then $\{i, i+1, \dots, j\}$ is an *occurrence* of u in w.

Definition of a string attractor [Prezza, ICTCS 2017]

Let $w = w_0 w_1 \dots w_n$ be a finite word over alphabet $\mathcal{A}$. A *string attractor* of w is a set of positions $\Gamma \subseteq \{0, \dots, n\}$ such that every substring of w has an occurrence containing an element of Γ.

Example:

$$w = 012300123012$$

$$\Gamma = \{2, 3, 4, 8, 10\} \leftrightarrow w = 01\colorbox{lime}{230}012\colorbox{lime}{3}0\colorbox{lime}{1}2$$

String attractors: definition and example

Definition: Let $w = w_0 w_1 \dots w_n$ be a word, let $u = w_i w_{i+1} \dots w_j$ be its factor. Then $\{i, i+1, \dots, j\}$ is an *occurrence* of u in w.

Definition of a string attractor [Prezza, ICTCS 2017]

Let $w = w_0 w_1 \dots w_n$ be a finite word over alphabet $\mathcal{A}$. A *string attractor* of w is a set of positions $\Gamma \subseteq \{0, \dots, n\}$ such that every substring of w has an occurrence containing an element of Γ.

Example:

$$w = 012300123012$$

$$\Gamma = \{2, 3, 4, 8, 10\} \leftrightarrow w = 01\colorbox{lime}{230}012\colorbox{lime}{3}0\colorbox{lime}{1}2$$

$$\Gamma^* = \{2, 3, 4, 10\} \leftrightarrow w = 01\colorbox{lime}{230}01230\colorbox{lime}{1}2$$

String attractors: definition and example

Definition: Let $w = w_0 w_1 \dots w_n$ be a word, let $u = w_i w_{i+1} \dots w_j$ be its factor. Then $\{i, i+1, \dots, j\}$ is an *occurrence* of u in w.

Definition of a string attractor [Prezza, ICTCS 2017]

Let $w = w_0 w_1 \dots w_n$ be a finite word over alphabet $\mathcal{A}$. A *string attractor* of w is a set of positions $\Gamma \subseteq \{0, \dots, n\}$ such that every substring of w has an occurrence containing an element of Γ.

Example:

$$w = 012300123012$$

$$\Gamma = \{2, 3, 4, 8, 10\} \leftrightarrow w = 01\underline{230}012\underline{3}0\underline{1}2$$

$$\Gamma^* = \{2, 3, 4, 10\} \leftrightarrow w = 01\underline{230}01230\underline{1}2$$

$$\Gamma^* = \{3, 5, 7, 10\} \leftrightarrow w = 012\underline{3}0\underline{0}1\underline{2}30\underline{1}2$$

Γ^* = some attractor with the minimum length

Overview of attractors in CoW

In CoW, minimal attractors have been determined for

Overview of attractors in CoW

In CoW, minimal attractors have been determined for

- particular prefixes
  - of standard Sturmian sequences by Mantaci, Restivo, Romana, Rosone, Sciortino, 2021
  - of the Thue-Morse sequence by Kutsukake et al., 2020

Overview of attractors in CoW

In CoW, minimal attractors have been determined for

- particular prefixes
  - of standard Sturmian sequences by Mantaci, Restivo, Romana, Rosone, Sciortino, 2021
  - of the Thue-Morse sequence by Kutsukake et al., 2020
- for prefixes
  - of standard Sturmian sequences by Restivo, Romana, Sciortino, 2022
  - of the Tribonacci sequence by Schaeffer & Shallit, 2021
  - of the Thue-Morse sequence by Schaeffer & Shallit, 2021
  - of the period-doubling sequence by Schaeffer & Shallit, 2021
  - of the powers of two sequence by Schaeffer & Shallit, 2021

Overview of attractors in CoW

In CoW, minimal attractors have been determined for

- particular prefixes
  - of standard Sturmian sequences by Mantaci, Restivo, Romana, Rosone, Sciortino, 2021
  - of the Thue-Morse sequence by Kutsukake et al., 2020
- for prefixes
  - of standard Sturmian sequences by Restivo, Romana, Sciortino, 2022
  - of the Tribonacci sequence by Schaeffer & Shallit, 2021
  - of the Thue-Morse sequence by Schaeffer & Shallit, 2021
  - of the period-doubling sequence by Schaeffer & Shallit, 2021
  - of the powers of two sequence by Schaeffer & Shallit, 2021
- for factors
  - of episturmian sequences by Dvořáková, 2022
  - of the Thue-Morse sequence by Dolce, 2023

Overview of attractors in CoW

- Schaeffer & Shallit, 2021: study of attractors in linearly recurrent and in automatic sequences
- Restivo, Romana, Sciortino, 2022: combinatorial properties of attractors (relation to factor complexity, recurrence function, etc.), study of attractors in fixed points of morphisms
- Romana: String Attractor: a Combinatorial Object from Data Compression, 2022
- Gheeraert, Romana, Stipulanti, 2023: study of attractors in fixed points of k-bonacci-like morphisms

Outline

Way to generate Sturmian sequences: Palindromic closures

Palindromes:
Word u is a *palindrome* if it reads the same forward and backward.

e.g. $1001, 11011, 10101$

Way to generate Sturmian sequences: Palindromic closures

Palindromes:
Word u is a *palindrome* if it reads the same forward and backward.

e.g. $1001, 11011, 10101$

Palindromic closure
Palindromic closure of u is the shortest palindrome having u as a prefix.

e.g. $100 \to 1001$, $1011 \to 101101$

Way to generate Sturmian sequences: Palindromic closures

Palindromes:
Word u is a *palindrome* if it reads the same forward and backward.

e.g. $1001, 11011, 10101$

Palindromic closure
Palindromic closure of u is the shortest palindrome having u as a prefix.

e.g. $100 \rightarrow 1001$, $1011 \rightarrow 101101$

[X. Droubay, J. Justin, G. Pirillo, 2001]

Algorithm for generating Sturmian sequences

- Take any binary sequence (= directive sequence)
- Add letters from directive sequence one by one to generated word
- After each letter addition, make a palindromic closure

Example: Fibonacci word

Directive sequence $\Delta = (01)^\omega = 0\,1\,0\,1\,0\,1\ldots$

Example: Fibonacci word

$$\text{Directive sequence } \Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1 \ldots$$

$$u_1 = 0$$

Example: Fibonacci word

$$\text{Directive sequence } \Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1\ \ldots$$

$$u_1 = 0$$

$$u_2 = 01$$

Example: Fibonacci word

$$\text{Directive sequence } \Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1 \ldots$$

$$u_1 = 0$$

$$u_2 = 010$$

Example: Fibonacci word

Directive sequence $\Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1 \ldots$

$$u_1 = 0$$
$$u_2 = 010$$
$$u_3 = 0100$$

Example: Fibonacci word

Directive sequence $\Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1 \ldots$

$$u_1 = 0$$
$$u_2 = 010$$
$$u_3 = 010010$$

Example: Fibonacci word

Directive sequence $\Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1 \ldots$

$$u_1 = 0$$
$$u_2 = 010$$
$$u_3 = 010010$$
$$u_4 = 0100101001 0$$

Example: Fibonacci word

Directive sequence $\Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1\ \ldots$

$u_1 = 0$

$u_2 = 010$

$u_3 = 010010$

$u_4 = 0100101001 0$

$u_5 = 0100101001001010010$

Example: Fibonacci word

Directive sequence $\Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1 \dots$

$$u_1 = 0$$
$$u_2 = 010$$
$$u_3 = 010010$$
$$u_4 = 0100101001 0$$
$$u_5 = 0100101001001010010$$
$$\vdots$$

Highlights mark the *longest palindromic prefixes* followed by 0 and 1

Example: Fibonacci word

Directive sequence $\Delta = (01)^\omega = 0\ 1\ 0\ 1\ 0\ 1 \ldots$

$$u_1 = 0$$
$$u_2 = 010$$
$$u_3 = 010010$$
$$u_4 = 0100101001 0$$
$$u_5 = 0100101001001010010$$
$$\vdots$$

Highlights mark the *longest palindromic prefixes* followed by 0 and 1

Longest palindromic prefixes followed by distinct letters mark attractors for all palindromic prefixes of standard Sturmian (and episturmian) words [L. Dvořáková, 2022]

Attractors of episturmian sequences via palindromic closures

$\rightarrow$ generalizing for any alphabet size

Attractors of episturmian sequences via palindromic closures

$\rightarrow$ generalizing for any alphabet size

Example: Tribonacci word

Directive sequence $\Delta = (012)^{\omega} = 0\ 1\ 2\ 0\ 1\ 2\ 0\ 1\ 2 \ldots$

Attractors of episturmian sequences via palindromic closures

$\rightarrow$ generalizing for any alphabet size

Example: Tribonacci word

Directive sequence $\Delta = (012)^{\omega} = 0\,1\,2\,0\,1\,2\,0\,1\,2\ldots$

$u_1 = 0$

Attractors of episturmian sequences via palindromic closures

$\rightarrow$ generalizing for any alphabet size

Example: Tribonacci word

$$\text{Directive sequence } \Delta = (012)^{\omega} = 0\,1\,2\,0\,1\,2\,0\,1\,2\ldots$$

$$u_1 = 0$$

$$u_2 = 010$$

Attractors of episturmian sequences via palindromic closures

$\rightarrow$ generalizing for any alphabet size

Example: Tribonacci word

Directive sequence $\Delta = (012)^{\omega} = 0\,1\,2\,0\,1\,2\,0\,1\,2 \dots$

$u_1 = 0$

$u_2 = 010$

$u_3 = 0102010$

Attractors of episturmian sequences via palindromic closures

$\rightarrow$ generalizing for any alphabet size

Example: Tribonacci word

$$\text{Directive sequence } \Delta = (012)^{\omega} = 0\,1\,2\,0\,1\,2\,0\,1\,2\dots$$

$$u_1 = 0$$
$$u_2 = 010$$
$$u_3 = 0102010$$
$$u_4 = 0102010010201 0$$

Attractors of episturmian sequences via palindromic closures

$\rightarrow$ generalizing for any alphabet size

Example: Tribonacci word

$$\text{Directive sequence } \Delta = (012)^{\omega} = \boxed{0}\boxed{1}\boxed{2}\boxed{0}\boxed{1}\boxed{2}\boxed{0}\boxed{1}\boxed{2}\ldots$$

$$u_1 = \boxed{0}$$

$$u_2 = \boxed{0}\boxed{1}0$$

$$u_3 = \boxed{0}\boxed{1}0\boxed{2}010$$

$$u_4 = 0\boxed{1}0\boxed{2}010\boxed{0}102010$$

$$u_5 = 010\boxed{2}010\boxed{0}102010\boxed{1}020100102010$$

$$\vdots$$

The *longest palindromic prefixes* followed by $\boxed{0}$, $\boxed{1}$, and $\boxed{2}$ create attractors for these words. [L. Dvořáková, 2022]

Outline

1 Motivation

2 String attractors overview

3 Palindromic closures and Sturmian sequences

4 Pseudopalindromic closures and Rote sequences

5 Open questions

Our interest: Rote sequences

Definition

Complementary-symmetric (CS) Rote sequences are binary sequences having complexity $2n$ and such that their language is closed under letter exchange.

Closely connected to Sturmian sequences by words' sum:

Our interest: Rote sequences

Definition

Complementary-symmetric (CS) Rote sequences are binary sequences having complexity $2n$ and such that their language is closed under letter exchange.

Closely connected to Sturmian sequences by words' sum:

Definition: Let $w = w_0 \dots w_n$ be a binary word. Its *sum* is defined as $S(w) = u = u_0 \dots u_{n-1}$, where $u_i = w_i + w_{i+1} \mod 2$.

$$w = 0011100$$
$$S(w) = 010010$$

Our interest: Rote sequences

Definition

Complementary-symmetric (CS) Rote sequences are binary sequences having complexity $2n$ and such that their language is closed under letter exchange.

Closely connected to Sturmian sequences by words' sum:

Definition: Let $w = w_0 \dots w_n$ be a binary word. Its *sum* is defined as $S(w) = u = u_0 \dots u_{n-1}$, where $u_i = w_i + w_{i+1} \mod 2$.

$$\begin{array}{rcll} w & = & 0011100 & \text{Rote} \\ & & \bigvee\bigvee\bigvee\bigvee\bigvee\bigvee & \\ S(w) & = & 010010 & \text{Sturmian} \end{array}$$

Structural theorem [G. Rote, 1994]

A binary sequence w is a CS Rote sequence if and only if the sequence $S(w)$ is a Sturmian sequence.

How can we obtain Rote attractors from Sturmian ones?

How can we obtain Rote attractors from Sturmian ones?

It seems that we cannot - not with known attractors of palindromic prefixes for Sturmian words.

How can we obtain Rote attractors from Sturmian ones?

It seems that we cannot - not with known attractors of palindromic prefixes for Sturmian words.

Example:

Rote: $w = 0011\underline{\mathbf{10}}0011$ - unique factor underlined

Sturmian: $u = 0100\underline{\mathbf{1}}0010$ - attractor should contain this position

How can we obtain Rote attractors from Sturmian ones?

It seems that we cannot - not with known attractors of palindromic prefixes for Sturmian words.

Example:

Rote: $w = 0011\underline{\mathbf{10}}0011$ - unique factor underlined

Sturmian: $u = 0100\underline{\mathbf{1}}0010$ - attractor should contain this position

Currently known Sturmian attractors:

$u = 0\boxed{1}0010\boxed{0}10$ $\qquad$ $u = 010010\boxed{01}0$

No straightforward way how to obtain the necessary position from these.

Back to closures: Generalized pseudostandard sequences

Antipalindromes (on binary alphabet):
Word w is an *antipalindrome* if it reads forward and backward the same, only with letter exchange ($\overline{1} = 0$, $\overline{0} = 1$).

e.g. $1010, 110100, 10110010$

Antipalindromes (on binary alphabet):
Word w is an *antipalindrome* if it reads forward and backward the same, only with letter exchange ($\overline{1} = 0$, $\overline{0} = 1$).

e.g. $1010, 110100, 10110010$

Antipalindromic closure
Antipalindromic closure of w is the shortest antipalindrome having w as a prefix.

e.g. $100 \rightarrow 100110$, $101 \rightarrow 1010$

Back to closures: Generalized pseudostandard sequences

Antipalindromes (on binary alphabet):
Word w is an *antipalindrome* if it reads forward and backward the same, only with letter exchange ($\overline{1} = 0$, $\overline{0} = 1$).

e.g. $1010, 110100, 10110010$

Antipalindromic closure
Antipalindromic closure of w is the shortest antipalindrome having w as a prefix.

e.g. $100 \to 100110$, $101 \to 1010$

Algorithm for generating generalized pseudostandard sequences

- Take any binary bisequence (= directive bisequence) specifying letters $\{0, 1\}$ and closures $\{R, E\}$
- Add letters from directive bisequence one by one to generated word
- After each letter addition, make an (anti)palindromic closure

Rote sequences are subset of generalized pseudostandard sequences

Theorem [Blondin-Massé A. et al., 2013]

Let (Δ, Θ) be a directive bisequence. Then w generated by this bisequence is a standard CS Rote sequence if and only if w is aperiodic and no factor of the directive bisequence is in the following sets:
$\{(ab, EE) : a, b \in \{0,1\}\}$,
$\{(aa, RR) : a \in \{0,1\}\}$,
$\{(aa, RE) : a \in \{0,1\}\}$.

Omitting these pairs in the bisequence, we can generate Rote sequences using pseudopalindromic closures!

Rote sequences via closures

CS Rote sequences generated by pseudopalindromic closures omitting the following patterns:
$\{(ab, EE) : a, b \in \{0,1\}\} \cup \{(a\bar{a}, RR) : a \in \{0,1\}\} \cup$
$\{(aa, RE) : a \in \{0,1\}\}$

Example:

$\Delta = \boxed{0}\ \boxed{0}\ \boxed{1}\ \boxed{1}\ \boxed{0}\ \boxed{0}\ \mathbf{....}$
$\Theta = \boxed{\mathrm{R}}\ \boxed{\mathrm{R}}\ \boxed{\mathrm{E}}\ \boxed{\mathrm{R}}\ \boxed{\mathrm{E}}\ \boxed{\mathrm{R}}\ \mathbf{....}$

Rote sequences via closures

CS Rote sequences generated by pseudopalindromic closures omitting the following patterns:
$\{(ab, EE) : a, b \in \{0,1\}\} \cup \{(a\bar{a}, RR) : a \in \{0,1\}\} \cup$
$\{(aa, RE) : a \in \{0,1\}\}$

Example:

$\Delta = 0\ 0\ 1\ 1\ 0\ 0 \ \ldots.$

$\Theta = R\ R\ E\ R\ E\ R \ \ldots.$

$w_1 = 0$

Rote sequences via closures

CS Rote sequences generated by pseudopalindromic closures omitting the following patterns:
$\{(ab, EE) : a, b \in \{0,1\}\} \cup \{(a\bar{a}, RR) : a \in \{0,1\}\} \cup$
$\{(aa, RE) : a \in \{0,1\}\}$

Example:

$\Delta = 0\ 0\ 1\ 1\ 0\ 0\ \ldots.$

$\Theta = R\ R\ E\ R\ E\ R\ \ldots.$

$w_1 = 0$

$w_2 = 00$

CS Rote sequences generated by pseudopalindromic closures omitting the following patterns:
$\{(ab, EE) : a, b \in \{0, 1\}\} \cup \{(a\bar{a}, RR) : a \in \{0, 1\}\} \cup \{(aa, RE) : a \in \{0, 1\}\}$

Example:

$\Delta = 0\ 0\ 1\ 1\ 0\ 0\ \ldots.$

$\Theta = R\ R\ E\ R\ E\ R\ \ldots.$

$w_1 = 0$

$w_2 = 00$

$w_3 = 0011$

CS Rote sequences generated by pseudopalindromic closures omitting the following patterns:
$\{(ab, EE) : a, b \in \{0,1\}\} \cup \{(a\bar{a}, RR) : a \in \{0,1\}\} \cup$
$\{(aa, RE) : a \in \{0,1\}\}$

Example:

$\Delta = \boxed{0}\ \boxed{0}\ \boxed{1}\ \boxed{1}\ \boxed{0}\ \boxed{0}$ ....

$\Theta = \boxed{R}\ \boxed{R}\ \boxed{E}\ \boxed{R}\ \boxed{E}\ \boxed{R}$ ....

$w_1 = 0$

$w_2 = 00$

$w_3 = 0011$

$w_4 = 0011100$

CS Rote sequences generated by pseudopalindromic closures omitting the following patterns:
$\{(ab, EE) : a, b \in \{0,1\}\} \cup \{(a\bar{a}, RR) : a \in \{0,1\}\} \cup$
$\{(aa, RE) : a \in \{0,1\}\}$

Example:

$\Delta = 0\ 0\ 1\ 1\ 0\ 0\ \ldots.$
$\Theta = R\ R\ E\ R\ E\ R\ \ldots.$

$w_1 = 0$

$w_2 = 00$

$w_3 = 0011$

$w_4 = 0011100$

$w_5 = 00111000011$

CS Rote sequences generated by pseudopalindromic closures omitting the following patterns:
$\{(ab, EE) : a, b \in \{0,1\}\} \cup \{(a\bar{a}, RR) : a \in \{0,1\}\} \cup \{(aa, RE) : a \in \{0,1\}\}$

Example:

$\Delta =$ 0 0 1 1 0 0 ....
$\Theta =$ R R E R E R ....

$w_1 = 0$

$w_2 = 00$

$w_3 = 0011$

$w_4 = 0011100$

$w_5 = 00111000011$

$w_6 = 00111000011000011100$

Rote sequences via closures

CS Rote sequences generated by pseudopalindromic closures omitting the following patterns:
$\{(ab, EE) : a, b \in \{0,1\}\} \cup \{(a\bar{a}, RR) : a \in \{0,1\}\} \cup$
$\{(aa, RE) : a \in \{0,1\}\}$

Example:

$\Delta =$ 0 0 1 1 0 0 ....
$\Theta =$ R R E R E R ....

$w_1 = 0$

$w_2 = 00$

$w_3 = 0011$

$w_4 = 0011100$

$w_5 = 00111000011$

$w_6 = 00111000011000011100$

Can we use the longest pseudopalindromic prefixes followed by distinct letters to obtain attractors of pseudopalindromic prefixes of Rote sequences?

Result: Attractors of Rote sequences

Theorem [Dvořáková L., Hendrychová V., 2023]

Assume (Δ, Θ) is the directive bisequence of a standard CS Rote sequence w, and w_n contains both letters. Then

1. If w_n is antipalindromic, w_n has an attractor $\Gamma = \{|w_i|, |w_{n-1}|\}$, where w_i is the longest antipalindromic prefix followed by $\overline{\Delta[n]}$ in w.

Result: Attractors of Rote sequences

Theorem [Dvořáková L., Hendrychová V., 2023]

Assume (Δ, Θ) is the directive bisequence of a standard CS Rote sequence w, and w_n contains both letters. Then

1. If w_n is antipalindromic, w_n has an attractor $\Gamma = \{|w_i|, |w_{n-1}|\}$, where w_i is the longest antipalindromic prefix followed by $\overline{\Delta[n]}$ in w.
2. If w_n is palindromic and w_{n-1} is antipalindromic, w_n has an attractor $\Gamma = \{|w_j|, |w_{n-1}|\}$, where w_j is the longest palindromic prefix followed by $\overline{\Delta[n]}$ in w.

Result: Attractors of Rote sequences

Theorem [Dvořáková L., Hendrychová V., 2023]

Assume (Δ, Θ) is the directive bisequence of a standard CS Rote sequence w, and w_n contains both letters. Then

1. If w_n is antipalindromic, w_n has an attractor $\Gamma = \{|w_i|, |w_{n-1}|\}$, where w_i is the longest antipalindromic prefix followed by $\overline{\Delta[n]}$ in w.
2. If w_n is palindromic and w_{n-1} is antipalindromic, w_n has an attractor $\Gamma = \{|w_j|, |w_{n-1}|\}$, where w_j is the longest palindromic prefix followed by $\overline{\Delta[n]}$ in w.
3. If w_n is palindromic and w_{n-1} is palindromic, w_n has the same attractor as w_{n-1}.

$\rightarrow$ The form of attractor depends not only on the current closure, but also on the preceding one.

Example: Attractor of Rote sequence

"LPPn" = longest palindromic prefix followed by n
"LAPn" = longest antipalindromic prefix followed by n

Example:

$\Delta =$ 0 0 1 1 0 0 0 ....

$\Theta =$ R R E R E R R ....

w_i	attractor
$w_1 = 0$	-
$w_2 = 0\,0$	-
$w_3 = \underline{\mathbf{0}}\,0\,\underline{\mathbf{1}}\,1$	$\lvert \text{LAP } 0 \rvert$, $\lvert w_2 \rvert$

Example: Attractor of Rote sequence

"LPPn" = longest palindromic prefix followed by n
"LAPn" = longest antipalindromic prefix followed by n

Example:

$\Delta = 0\ 0\ 1\ 1\ 0\ 0\ 0 \ldots.$
$\Theta = R\ R\ E\ R\ E\ R\ R \ldots.$

w_i		attractor
w_1	$=0$	-
w_2	$=00$	-
w_3	$=\underline{\mathbf{0}}0\underline{\mathbf{1}}1$	$\lvert$LAP 0$\rvert$, $\lvert w_2\rvert$
w_4	$=0\underline{\mathbf{0}}11\underline{\mathbf{1}}00$	$\lvert$LPP 0$\rvert$, $\lvert w_3\rvert$

Example: Attractor of Rote sequence

"LPPn" = longest palindromic prefix followed by n
"LAPn" = longest antipalindromic prefix followed by n

Example:

Δ = 0 0 1 1 0 0 0 ....
Θ = R R E R E R R ....

w_i		attractor
w_1	$= 0$	-
w_2	$= 00$	-
w_3	$= \underline{\mathbf{0}}0\underline{\mathbf{1}}1$	$\lvert \text{LAP } 0\rvert$, $\lvert w_2\rvert$
w_4	$= 0\underline{\mathbf{0}}11\underline{\mathbf{1}}00$	$\lvert \text{LPP } 0\rvert$, $\lvert w_3\rvert$
w_5	$= 0011\underline{\mathbf{1}}00\underline{\mathbf{0}}11$	$\lvert \text{LAP } 1\rvert$, $\lvert w_4\rvert$
w_6	$= 00\underline{\mathbf{1}}110001 1\underline{\mathbf{0}}0011100$	$\lvert \text{LPP } 1\rvert$, $\lvert w_5\rvert$
w_7	$= 00\underline{\mathbf{1}}1100011\underline{\mathbf{0}}00111000110001 1100$	same as previous

Outline

Open problems

- What are the attractors of prefixes of **generalized pseudostandard sequences**?
  - i.e. what if we don't omit any (anti)palindromic combinations in the generating bisequence?

Open problems

- What are the attractors of prefixes of **generalized pseudostandard sequences**?
  - i.e. what if we don't omit any (anti)palindromic combinations in the generating bisequence?
  - For Thue-Morse word - min. size 4
  - For pseudostandard sequences (only E closures) - min. size 3
  - But generally it is unknown

Open problems

- What are the attractors of prefixes of **generalized pseudostandard sequences**?
  - i.e. what if we don't omit any (anti)palindromic combinations in the generating bisequence?
  - For Thue-Morse word - min. size 4
  - For pseudostandard sequences (only E closures) - min. size 3
  - But generally it is unknown
- What about attractors of (generalized) pseudostandard sequences over **larger alphabets**?

Open problems

- What are the attractors of prefixes of **generalized pseudostandard sequences**?
  - i.e. what if we don't omit any (anti)palindromic combinations in the generating bisequence?
  - For Thue-Morse word - min. size 4
  - For pseudostandard sequences (only E closures) - min. size 3
  - But generally it is unknown
- What about attractors of (generalized) pseudostandard sequences over **larger alphabets**?
- How does the minimum attractor size affect the **form of examined words compressed** by dictionary compressors? Do they also remain constant?

Thank you for your attention!