Innerer palindromischer Abschluss

Jürgen Dassow[1] Florin Manea[2] Robert Mercaş[1,2] Mike Müller[2]

[1]Otto-von-Guericke-Universität Magdeburg

[2]Christian-Albrechts-Universität zu Kiel

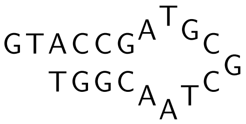
27. September 2013

Similar problems

GTACCGATGCGCTAACGGT

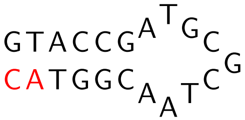
Similar problems

GTACCGATGCGCTAACGGT $\rightarrow$



Similar problems

GTACCGATGCGCTAACGGTAC→



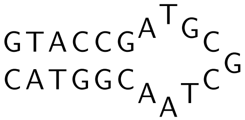
Similar problems

GTACCGATGCGCTAACGGTAC→

GTACCG$^{A\,T}$G$_C$
CATGGC$_{A\,A}$T^C G

$xYz\overline{Y}^R$

Similar problems

GTACCGATGCGCTAACGGTAC→ 

$xYz\overline{Y}^R \rightarrow xYz\overline{Y}^R\overline{x}^R$

Similar problems

GTACCGATGCGCTAACGGTAC→

GTACCG A T G C
CATGGC A A T C G

$xYz\overline{Y}^R \rightarrow xYz\overline{Y}^R\overline{x}^R$

hairpin completion

Similar problems

GTACCGATGCGCTAACGGTAC→

GTACCG^{A} ^{T} G$_{C}$
CATGGC$_{A}$ $_{A}$ T C ^{G}

$$xY z \overline{Y}^R \to xY z \overline{Y}^R \overline{x}^R$$
$$Y z \overline{Y}^R x \to \overline{x}^R Y z \overline{Y}^R x$$

hairpin completion

Similar problems

GTACCGATGCGCTAACGGTAC$\rightarrow$

GTACCG$^{A\,T}$G$_{C}$ ^{G}
CATGGC$_{A\,A}$T^{C}

$$xYz\overline{Y}^R \rightarrow xYz\overline{Y}^R\overline{x}^R$$
$$Yz\overline{Y}^R x \rightarrow \overline{x}^R Yz\overline{Y}^R x$$

hairpin completion

X

Similar problems

GTACCGATGCGCTAACGGTAC→

GTACCG^{A} ^{T} G$_{C}$ ^{G}
CATGGC$_{A}$ $_{A}$ T^{C}

$xYz\overline{Y}^R \rightarrow xYz\overline{Y}^R\overline{x}^R$

$Yz\overline{Y}^R x \rightarrow \overline{x}^R Yz\overline{Y}^R x$

hairpin completion

$X \rightarrow XX$

Similar problems

GTACCGATGCGCTAACGGTAC$\rightarrow$

GTACCG A T G C
CATGGC A A T C G

$xYz\overline{Y}^R \rightarrow xYz\overline{Y}^R\overline{x}^R$ hairpin completion

$Yz\overline{Y}^R x \rightarrow \overline{x}^R Yz\overline{Y}^R x$

$yXz \rightarrow yXXz$ duplication languages

Basics

palindrome: $w = w_1 \ldots w_n = w_n \ldots w_1 = w^R$ (*rentner*)

Basics

palindrome: $w = w_1 \ldots w_n = w_n \ldots w_1 = w^R$ (*rentner*)

DEFINITION (DE LUCA)

For a word u, the left (right) palindromic closure of u is a palindrome vu (uv) with v non-empty, such that any other palindrome with u as proper suffix (prefix) has length greater than $|uv|$.

Basics

palindrome: $w = w_1 \ldots w_n = w_n \ldots w_1 = w^R$ (*rentner*)

Definition (de Luca)

For a word u, the left (right) palindromic closure of u is a palindrome vu (uv) with v non-empty, such that any other palindrome with u as proper suffix (prefix) has length greater than $|uv|$.

aab

Basics

palindrome: $w = w_1 \ldots w_n = w_n \ldots w_1 = w^R$ (*rentner*)

Definition (de Luca)

For a word u, the left (right) palindromic closure of u is a palindrome vu (uv) with v non-empty, such that any other palindrome with u as proper suffix (prefix) has length greater than $|uv|$.

$aab \rightarrow \{aabaa, baab\}$

Basics

palindrome: $w = w_1 \ldots w_n = w_n \ldots w_1 = w^R$ (rentner)

Definition (de Luca)

For a word u, the left (right) palindromic closure of u is a palindrome vu (uv) with v non-empty, such that any other palindrome with u as proper suffix (prefix) has length greater than $|uv|$.

$aab \rightarrow \{aabaa, baab\}$ aaa

Basics

palindrome: $w = w_1 \ldots w_n = w_n \ldots w_1 = w^R$ (rentner)

Definition (de Luca)

For a word u, the left (right) palindromic closure of u is a palindrome vu (uv) with v non-empty, such that any other palindrome with u as proper suffix (prefix) has length greater than $|uv|$.

$aab \rightarrow \{aabaa, baab\}$

$aaa \rightarrow \{aaaa\}$

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

aab

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

$\spadesuit_\ell$ $\spadesuit_r$

$aab \rightarrow \{aaab,$

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

$\spadesuit_\ell$ $\spadesuit_r$

$aab \rightarrow \{aaab,$

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

$\spadesuit_\ell$ $\spadesuit_r$

$aab \rightarrow \{aaab, aabb,$

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

$\spadesuit_\ell$

$aab \rightarrow \{aaab, aabb, baab,$

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

$\spadesuit_r$

$aab \rightarrow \{aaab, aabb, baab, aabaa,$

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

$\spadesuit_r$

$aab \rightarrow \{aaab, aabb, baab, aabaa, aaba,$

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

$\spadesuit_\ell$

$$aab \rightarrow \{aaab, aabb, baab, aabaa, aaba, abab\}$$

Definitions

Definition

For a word w, the left (right) inner palindromic closure of w, denoted by $\spadesuit_\ell(w)$ ($\spadesuit_r(w)$), is the set of all words $xvuy$ ($xuvy$) for any factorisation $w = xuy$ with possibly empty x, y and non-empty u, v, such that vu (uv) is the left (right) palindromic closure of u. The inner palindromic closure $\spadesuit(w)$ is the union of $\spadesuit_\ell(w)$ and $\spadesuit_r(w)$.

$$aab \rightarrow \{aaab, aabb, baab, aabaa, aaba, abab\}$$

Definition

For a language L, let $\spadesuit(L) = \bigcup_{w \in L} \spadesuit(w)$. We set $\spadesuit^0(L) = L$, $\spadesuit^n(L) = \spadesuit(\spadesuit^{n-1}(L))$ for $n \geq 1$, $\spadesuit^*(L) = \bigcup_{n \geq 0} \spadesuit^n(L)$.

Observations

Lemma

For every word w, if $u \in \spadesuit^*(w)$, then $w \preccurlyeq u$.

Observations

Lemma

For every word w, if $u \in \spadesuit^*(w)$, then $w \preccurlyeq u$.

Lemma

[Propagation rule] For a word $w = a^n b^m$ with positive integers n and m, the set $\spadesuit(w)$ contains all words of length $n + m + 1$ with a letter $x \in \{a, b\}$ inserted at any position i of w, where $0 \leq i < n + m$.

Corollary

For any binary words w and u, $w \preccurlyeq u$ if and only if $u \in \spadesuit^*(w)$.

Regularity of closure

Theorem

The iterated inner palindromic closure of a binary language is regular.

Regularity of closure

Theorem

The iterated inner palindromic closure of a binary language is regular.

Theorem (Higman)

If L is a language such that any two words in L are incomparable with respect to the scattered factors partial order, then L is finite.

Regularity of closure

Theorem

The iterated inner palindromic closure of a binary language is regular.

Theorem (Higman)

If L is a language such that any two words in L are incomparable with respect to the scattered factors partial order, then L is finite.

Theorem

The finite inner palindromic closure of a regular language is not necessarily regular.

Duplication idea

Duplication idea

Duplication idea

Duplication idea

Duplication idea

Duplication idea

Take prefixes of an infinite square-free word, and show that they are from different equivalent classes.

Why it does not work?!

Why it does not work?!

Why it does not work?!

Why it does not work?!

Why it does not work?!

Why it does not work?!

Why it does not work?!

Why it does not work?!

Why it does not work?!

$\color{red}{abc}\underline{d}abfa$

Why it does not work?!

$abc\underline{d}abfa$

$abcba\underline{d}abfa$

Why it does not work?!

$abc\underline{d}abfa$

$abc\color{red}{ba\underline{d}abfa}$

Why it does not work?!

abcdabfa

abcbadabfa

abcafbadabfa

Bigger alphabets

Bigger alphabets

Consider now the word abc and the language $(abc)^*$ that contains no palindromes of length greater than one. However, $babc \in \spadesuit(abc)$ can generate at the beginning as many abc's as we want, $(abc)^*babc$.

Bigger alphabets

Consider now the word *abc* and the language $(abc)^*$ that contains no palindromes of length greater than one. However, $babc \in ♠(abc)$ can generate at the beginning as many *abc*'s as we want, $(abc)^* babc$.

Lemma

Let $\Sigma = \{a_1, a_2, \ldots, a_k\}$ *and define the recursive sequences*

$$w'_0 = \varepsilon \text{ and } w_0 = \varepsilon,$$
$$w'_i = w_{i-1} w'_{i-1} \text{ and } w_i = w'_i a_i \text{ for } 1 \leq i \leq k.$$

Then for $1 \leq i \leq k$, $alph(w_i)^* w_i \subseteq ♠^*(w_i)$.

Bigger alphabets

Consider now the word *abc* and the language $(abc)^*$ that contains no palindromes of length greater than one. However, $babc \in \spadesuit(abc)$ can generate at the beginning as many *abc*'s as we want, $(abc)^* babc$.

Lemma

Let $\Sigma = \{a_1, a_2, \ldots, a_k\}$ and define the recursive sequences

$$w'_0 = \varepsilon \text{ and } w_0 = \varepsilon,$$
$$w'_i = w_{i-1} w'_{i-1} \text{ and } w_i = w'_i a_i \text{ for } 1 \leq i \leq k.$$

Then for $1 \leq i \leq k$, $alph(w_i)^* w_i \subseteq \spadesuit^*(w_i)$.

$$a, ab, abac, abacabad, abacabadabacabae, \ldots$$

Parametrized inner closure

Parametrized inner closure

Definition

For a word u and $m, n \in \mathbb{N}$, we define the sets

$$L_{m,n}(w) = \{u \mid u = u^R, u = xw \text{ for } x \neq \varepsilon, |x| \geq n,\ m \geq |w| - |x| \geq 0\},$$
$$R_{m,n}(w) = \{u \mid u = u^R, u = wx \text{ for } x \neq \varepsilon, |x| \geq n,\ m \geq |w| - |x| \geq 0\}.$$

The left (right) (m, n)-palindromic closure of w is the shortest word of $L_{m,n}(w)$ (resp., $R_{m,n}(w)$), or undefined if $L_{m,n}(w)$ (resp., $R_{m,n}(w)$) is empty.

Definition

For non-negative integers n, m with $n > 0$, we define the $\spadesuit_{(m,n)}$ one step inner palindromic closure of some word w as

$$\spadesuit_{(m,n)}(w) = \{u \mid u = xy'z,\ w = xyz,\ \text{and } y' \text{ is obtained by left or right } (m,n)\text{-palindromic closure from } y\}.$$

Example

aab

Example

$L_{(1,2)}$ $R_{(1,2)}$

$\spadesuit_{(1,2)}(aab) = \{aaaab,$

Example

$R_{(1,2)}$

$\spadesuit_{(1,2)}(aab) = \{aaaab, aabba,$

Example

$L_{(1,2)}$

$\spadesuit_{(1,2)}(aab) = \{aaaab, aabba, abaab,$

Example

$L_{(1,2)}$

$\spadesuit_{(1,2)}(aab) = \{aaaab, aabba, abaab, baaab,$

Example

$R_{(1,2)}$

$\spadesuit_{(1,2)}(aab) = \{aaaab, aabba, abaab, baaab, aabaa\}$

Example

$$\spadesuit_{(1,2)}(aab) = \{aaaab, aabba, abaab, baaab, aabaa\}$$

$$\spadesuit(aab) = \{aaab,\ aabb,\ baab,\ aabaa,\ aaba,\ abab\}$$

Example

$$\spadesuit_{(1,2)}(aab) = \{aaaab, aabba, abaab, baaab, aabaa\}$$

$$\spadesuit(aab) = \{aaab, aabb, baab, aabaa, aaba, abab\}$$

Note that $L_{m,n}(w)$ and $R_{m,n}(w)$ are empty if and only if $|w| < n$.

Initial results

Proposition

For any word w with $|w| \geq n$ and positive integer m, the language $\spadesuit^*_{(m,n)}(w)$ is dense with respect to the alphabet $alph(w)$.

Lemma

Let Σ be an alphabet with $||\Sigma|| \geq 2$, $a \notin \Sigma$, and m and n positive integers. Let $w = a^m y_1 a \cdots y_{p-1} a y_p$ be a word such that $alph(w) = \Sigma \cup \{a\}$, $m, p > 0$, $y_i \in \Sigma^*$ for $1 \leq i \leq p$, $|y_1| > 0$, and such that there exists $1 \leq j \leq p$ with $|y_j| \geq n$. Then, for each $v \in \Sigma^*$ with $|v| \geq n$, there exists $w' \in \spadesuit^*_{(m,n)}(w)$ such that v is a prefix of w' and $|w'|_a = |w|_a$.

Main result

Theorem

Let $m > 0$ and $k \geq 2$ be two integers and define $n = \max\{\frac{q_k}{2}, p_k\}$. Let Σ be a k-letter alphabet with $a \notin \Sigma$ and $w = a^m y_1 a y_2 \cdots a y_{r-1} a y_r$ be a word such that $alph(w) = \Sigma \cup \{a\}$, $r > 0$, $y_i \in \Sigma^*$ for all $1 \leq i \leq r$, and there exists j with $1 \leq j \leq r$ and $|y_j| \geq n$. Then $\spadesuit^*_{(m,n)}(w)$ is not regular.

Consequence

The following theorem follows immediately from the previous results.

Theorem

Let $w = a^p y_1 a \cdots y_{r-1} a y_r$, where $a \notin alph(y_i)$ for $1 \leq i \leq r$.
(1) If $\|alph(w)\| \geq 3$ and $|y_j| \geq 3$ for some $1 \leq j \leq r$, then for every positive integer $m \leq p$ we have that $\spadesuit^*_{(m,3)}(w)$ is not regular.
(2) If $\|alph(w)\| \geq 4$ and $|y_j| \geq 2$ for some $1 \leq j \leq r$, then for every positive integer $m \leq p$ we have that $\spadesuit^*_{(m,2)}(w)$ is not regular.
(3) If $\|alph(w)\| \geq 5$, then for every positive integer $m \leq p$ we have that $\spadesuit^*_{(m,1)}(w)$ is not regular.
(4) For every positive integers m and n there exists u with $\spadesuit^*_{(m,n)}(u)$ not regular.

Proof

Theorem (Proof of (1))

There exist infinitely long binary words avoiding both palindromes of length 6 and longer, and squares of words with length 3 and longer.

Proof

Theorem (Proof of (1))

There exist infinitely long binary words avoiding both palindromes of length 6 and longer, and squares of words with length 3 and longer.

Rampersad, Shallit, and Wang gave an infinite word w, that is square-free and has no factors from the set
$\{ac, ad, ae, bd, be, ca, ce, da, db, eb, ec, aba, ede\}$.
The morphism γ, defined by

$$\gamma(a) = abaabbab, \quad \gamma(b) = aaabbbab, \quad \gamma(c) = aabbabab,$$
$$\gamma(d) = aabbbaba, \quad \gamma(e) = baaabbab,$$

maps this word w to a word with the desired properties.

Proof

Theorem (Proof of (1))

There exist infinitely long binary words avoiding both palindromes of length 6 and longer, and squares of words with length 3 and longer.

Theorem (Proof of (2))

There exist infinitely long ternary words avoiding both palindromes of length 3 and longer, and squares of words with length 2 and longer.

Proof

Theorem (Proof of (1))

There exist infinitely long binary words avoiding both palindromes of length 6 and longer, and squares of words with length 3 and longer.

Theorem (Proof of (2))

There exist infinitely long ternary words avoiding both palindromes of length 3 and longer, and squares of words with length 2 and longer.

The morphism ψ, that is defined by

$$\psi(a) = abbccaabccab, \quad \psi(b) = bccaabbcaabc, \quad \psi(c) = caabbccabbca,$$

maps all infinite square-free ternary words h to words with the desired properties.

Proof

Theorem (Proof of (1))

There exist infinitely long binary words avoiding both palindromes of length 6 and longer, and squares of words with length 3 and longer.

Theorem (Proof of (2))

There exist infinitely long ternary words avoiding both palindromes of length 3 and longer, and squares of words with length 2 and longer.

Theorem (Proof of (3) - Pansiot)

There exist infinitely long words on a four letter alphabet that avoid $\frac{7}{5}$-powers (thus palindormes longer than 1 and squares).

Binary bounded case

Theorem

For any word $w \in \{a, b\}^+$ *and integer* $m \geq 0$, $\spadesuit^*_{(m,1)}(w)$ *is regular.*

Binary bounded case

Theorem

For any word $w \in \{a, b\}^+$ and integer $m \geq 0$, $\spadesuit^*_{(m,1)}(w)$ is regular.

We express $\spadesuit^*_{(m,1)}(w)$ as a finite union and concatenation of several languages $\spadesuit^*_{(m,1)}(w')$ with w' being a word strictly shorter than w, and some other simple regular languages.

Binary bounded case

Theorem

For any word $w \in \{a,b\}^+$ and integer $m \geq 0$, $\spadesuit^*_{(m,1)}(w)$ is regular.

We express $\spadesuit^*_{(m,1)}(w)$ as a finite union and concatenation of several languages $\spadesuit^*_{(m,1)}(w')$ with w' being a word strictly shorter than w, and some other simple regular languages.

A series of basic cases:

- words that have no maximal unary power greater than m,
- words of the form xy^qx
- words of the form xy^q or y^qx (x, y are letters)

Acknowledgement

Thank YOU for your attention

and the
Alexander von Humboldt Foundation.