Relative Watson-Crick Primitivity of Words

Manasi S. Kulkarni

Department of Mathematics
Indian Institute of Technology Madras, India

Joint work with Lila Kari and Kalpana Mahalingam

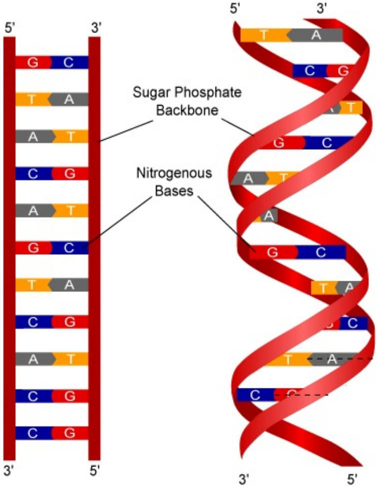
February 22, 2018

Introduction

Motivation

- Pattern matching and text compression algorithms: Detection of repetitions in a string
- Tandem repeats: Biological and medical significance
- Information equivalence between DNA strands
- Relatively prime numbers

Structure of DeoxyriboNucleic Acid



Notations

- Σ :- Alphabet
- λ :- Empty word
- A mapping $\theta : \Sigma^* \rightarrow \Sigma^*$ is said to be a morphism if $\theta(uv) = \theta(u)\theta(v)$, an antimorphism if $\theta(uv) = \theta(v)\theta(u)$, and an involution if $\theta(\theta(u)) = u$ for all $u, v \in \Sigma^*$.
- $|u|_a$:- Number of occurrences of $a \in \Sigma$ in $u \in \Sigma^+$
- $|u|_{a,\theta(a)} = |u|_a + |u|_{\theta(a)}$

Primitive Words

A word $u \in \Sigma^+$ is said to be primitive if $u = x^m$ for $x \in \Sigma^+$ implies $m = 1$ and $u = x$.
$\lambda(u)$: The primitive root of u

Primitive Words

A word $u \in \Sigma^+$ is said to be primitive if $u = x^m$ for $x \in \Sigma^+$ implies $m = 1$ and $u = x$.
$\lambda(u)$: The primitive root of u

θ-Primitive Words (Czeizler et al.)

A θ-power of u is a word of the form $u_1 u_2 \cdots u_n$ for some $n \geq 1$, where $u_1 = u$ and $u_i \in \{u, \theta(u)\}$ for $2 \leq i \leq n$.
A word is said to be θ-primitive if it's not a θ-power of another word.
$\rho_\theta(u)$: θ-primitive root of u
Q_θ: Set of all θ-primitive words

Primitive Words

A word $u \in \Sigma^+$ is said to be primitive if $u = x^m$ for $x \in \Sigma^+$ implies $m = 1$ and $u = x$.
$\lambda(u)$: The primitive root of u

θ-Primitive Words (Czeizler et al.)

A θ-power of u is a word of the form $u_1 u_2 \cdots u_n$ for some $n \geq 1$, where $u_1 = u$ and $u_i \in \{u, \theta(u)\}$ for $2 \leq i \leq n$.
A word is said to be θ-primitive if it's not a θ-power of another word.
$\rho_\theta(u)$: θ-primitive root of u
Q_θ: Set of all θ-primitive words

Example

Let $\Sigma = \{A, C, G, T\}$ and θ be an antimorphic involution such that $\theta(A) = T$, $\theta(G) = C$ and vice versa. Then $u_1 = ATT$ is not θ-primitive, however $u_2 = ACTG$ is a θ-primitive word.

Relative θ-Primitivity

Let $u, v \in \Sigma^*$ and let θ be an (anti)morphic involution on Σ^*. Then $(u, v)_\theta$ is defined as:

$$(u, v)_\theta = \begin{cases} x & \text{if } \rho_\theta(u) = \rho_\theta(v) = x,\ x \in \Sigma^+ \\ \lambda & \text{otherwise} \end{cases}$$

Relative θ-Primitivity

Let $u, v \in \Sigma^*$ and let θ be an (anti)morphic involution on Σ^*. Then $(u, v)_\theta$ is defined as:

$$(u, v)_\theta = \begin{cases} x & \text{if } \rho_\theta(u) = \rho_\theta(v) = x,\ x \in \Sigma^+ \\ \lambda & \text{otherwise} \end{cases}$$

If $(u, v)_\theta = \lambda$, then u and v are said to be relatively θ-primitive words, denoted by $u \perp_\theta v$.

Relative θ-Primitivity

Let $u, v \in \Sigma^*$ and let θ be an (anti)morphic involution on Σ^*. Then $(u, v)_\theta$ is defined as:

$$(u, v)_\theta = \begin{cases} x & \text{if } \rho_\theta(u) = \rho_\theta(v) = x, \; x \in \Sigma^+ \\ \lambda & \text{otherwise} \end{cases}$$

If $(u, v)_\theta = \lambda$, then u and v are said to be relatively θ-primitive words, denoted by $u \perp_\theta v$.
For $x \in \Sigma^+$, if $\rho_\theta(u) = \rho_\theta(v) = x$, then u and v are not relatively θ-primitive, denoted by $u \not\perp_\theta v$.

Example

Let $\Sigma = \{a, b, c\}$ and θ be an antimorphic involution such that $\theta(a) = b$ and vice versa, and $\theta(c) = c$. Let $u = abccababc$, $v_1 = abcc$, $v_2 = abcabc$. Then
$(u, v_1)_\theta = (abccababc, abcc)_\theta = \lambda$, i.e., $u \perp_\theta v_1$.
$(u, v_2)_\theta = (abccababc, abcabc)_\theta = abc$, since,
$\rho_\theta(u) = \rho_\theta(v_2) = abc$, i.e., $u \not\perp_\theta v_2$.

Equivalence Relation Properties

Let θ be an (anti)morphic involution over Σ^*.

The relation $\perp_\theta$ is symmetric on Σ^* and transitive on Q_θ. However the relation $\perp_\theta$ is not reflexive on Σ^*.

Equivalence Relation Properties

Let θ be an (anti)morphic involution over Σ^*.

The relation $\perp_\theta$ is symmetric on Σ^* and transitive on Q_θ. However the relation $\perp_\theta$ is not reflexive on Σ^*.

The relation $\not\perp_\theta$ is an equivalence relation on Σ^+.

Associativity and Commutativity

Associativity

For (anti)morphic involution θ over Σ^*, and $u, v, w \in \Sigma^+$,
$((u, v)_\theta, w)_\theta = (u, (v, w)_\theta)_\theta$

Associativity and Commutativity

Associativity

For (anti)morphic involution θ over Σ^*, and $u, v, w \in \Sigma^+$, $((u, v)_\theta, w)_\theta = (u, (v, w)_\theta)_\theta$

For $u, v \in \Sigma^+$, u is said to θ-commute with v if $uv = \theta(v)u$.

Commutativity

Let θ be a morphic involution over Σ^* and $u, v, w \in \Sigma^+$ be such that $vw = \theta(w)v$. Then,

1. $u \perp_\theta \{v, \theta(v)\}$ implies $u \perp_\theta w$
2. $u \perp_\theta \{w, \theta(w)\}$ implies $u \perp_\theta v$

Associativity and Commutativity

Associativity

For (anti)morphic involution θ over Σ^*, and $u, v, w \in \Sigma^+$, $((u, v)_\theta, w)_\theta = (u, (v, w)_\theta)_\theta$

For $u, v \in \Sigma^+$, u is said to θ-commute with v if $uv = \theta(v)u$.

Commutativity

Let θ be a morphic involution over Σ^* and $u, v, w \in \Sigma^+$ be such that $vw = \theta(w)v$. Then,

1. $u \perp_\theta \{v, \theta(v)\}$ implies $u \perp_\theta w$
2. $u \perp_\theta \{w, \theta(w)\}$ implies $u \perp_\theta v$

Does not hold for an antimorphic involution!

Words with Common θ-Primitive Root

For words u and v, $uv = vu$ if and only if u and v share the common primitive root

Words with Common θ-Primitive Root

For words u and v, $uv = vu$ if and only if u and v share the common primitive root

Proposition

Let θ be a morphic involution over Σ^* and let $u, v \in \Sigma^+$. If for some $x \in \Sigma^+$ we have that $(uv, vu)_\theta = x$ then $(u, v)_\theta = x$, and conversely.

Decidability

Theorem

Let θ be (anti)morphic involution over Σ^* and $u, v \in \Sigma^+$ be two words with $u \neq v$. It is decidable, in $\Theta(n^2 lg\ n)$ time, whether $u \perp_\theta v$, where $n = max\{|u|, |v|\}$.

Language Properties

$$L_{\theta,\lambda}(x) = \{w \in \Sigma^* | w \perp_\theta x\}$$

$$L_\theta(x) = \{w \in \Sigma^* | w \not\perp_\theta x\}$$

Language Properties

$$L_{\theta,\lambda}(x) = \{w \in \Sigma^* | w \perp_\theta x\}$$

$$L_\theta(x) = \{w \in \Sigma^* | w \not\perp_\theta x\}$$

Theorem

For a given (anti)morphic involution θ and $x \in \Sigma^+$, the languages $L_\theta(x)$ and $L_{\theta,\lambda}(x)$ are regular.

Binary Word (BW) Operation - θ-catenation

Binary Word (BW) Operation - θ-catenation

For (anti)morphic involution θ on Σ^* and two words $u, v \in \Sigma^*$, the binary word operation of θ-catenation is defined as

$$u \odot v = \{uv, u\theta(v)\}$$

Binary Word (BW) Operation - θ-catenation

For (anti)morphic involution θ on Σ^* and two words $u, v \in \Sigma^*$, the binary word operation of θ-catenation is defined as

$$u \odot v = \{uv, u\theta(v)\}$$

Proposition

Let θ be (anti)morphic involution over Σ^* and $u, v \in \Sigma^+$ be such that $u \perp_\theta \{v, \theta(v)\}$. Then for all $x \in u \odot v$, we have that $x \perp_\theta u$.

Proposition

Let θ be (anti)morphic involution over Σ^* and $u, v \in \Sigma^+$ be such that $u \perp_\theta v$. Then for all $x \in u \odot v$, we have that $x \perp_\theta v$.

BW Operation - Shuffle

Given $u, v \in \Sigma^+$, their shuffle $u \sqcup\!\sqcup v$ is defined as the set of all words of the form $u_1 v_1 \cdots u_k v_k$ such that $u = u_1 \cdots u_k, v = v_1 \cdots v_k$ where $u_i, v_i \in \Sigma^*$ for $1 \leq i \leq k$.

BW Operation - Shuffle

Given $u, v \in \Sigma^+$, their shuffle $u \sqcup\!\sqcup\, v$ is defined as the set of all words of the form $u_1 v_1 \cdots u_k v_k$ such that $u = u_1 \cdots u_k, v = v_1 \cdots v_k$ where $u_i, v_i \in \Sigma^*$ for $1 \leq i \leq k$.

Proposition

Let θ be (anti)morphic involution over Σ^*, and let $u, v \in \Sigma^+$ such that $u \perp_\theta v$, $|u| = |v|$, and there exists $a \in \Sigma$ such that $|u|_{a,\theta(a)} = |v|_{a,\theta(a)}$. Then $(u \sqcup\!\sqcup\, v) \perp_\theta \{u, v\}$.

BW Operation - Perfect Shuffle

Given $u, v \in \Sigma^+$, their perfect shuffle $u \sqcup\!\sqcup_p v$ is defined as the set of all words of the form $a_1 b_1 \cdots a_k b_k$ such that $u = a_1 \cdots a_k, v = b_1 \cdots b_k$ where $a_i, b_i \in \Sigma^*$ for $1 \leq i \leq k$.

BW Operation - Perfect Shuffle

Given $u, v \in \Sigma^+$, their perfect shuffle $u \sqcup\!\sqcup_p v$ is defined as the set of all words of the form $a_1 b_1 \cdots a_k b_k$ such that $u = a_1 \cdots a_k, v = b_1 \cdots b_k$ where $a_i, b_i \in \Sigma^*$ for $1 \leq i \leq k$.

A word $w \in \Sigma^+$ is said to be a θ-palindrome if $w = \theta(w)$ for antimorphic involution θ.

BW Operation - Perfect Shuffle

Given $u, v \in \Sigma^+$, their perfect shuffle $u \sqcup\!\sqcup_p v$ is defined as the set of all words of the form $a_1 b_1 \cdots a_k b_k$ such that $u = a_1 \cdots a_k, v = b_1 \cdots b_k$ where $a_i, b_i \in \Sigma^*$ for $1 \leq i \leq k$.

A word $w \in \Sigma^+$ is said to be a θ-palindrome if $w = \theta(w)$ for antimorphic involution θ.

Proposition

Let θ be an antimorphic involution over Σ^*, and let $u, v \in \Sigma^+$ be two equi-length θ-palindromes. If $u \perp_\theta v$, then $u \sqcup\!\sqcup_p v$ cannot be a θ palindrome.

Observation

$$u \perp_\theta v \overset{?}{\rightarrow} \theta(u) \perp_\theta \theta(v)$$

Observation

$u \perp_\theta v \xrightarrow{?} \theta(u) \perp_\theta \theta(v)$

Holds for morphic involution
Doesn't hold for antimorphic involution
Counter example: $u = x\theta(x)x$, $v = \theta(x)x$. Then $u \perp_\theta v$; but $\theta(u) \not\perp_\theta \theta(v)$

Observation

$u \perp_\theta v \xrightarrow{?} \theta(u) \perp_\theta \theta(v)$

Holds for morphic involution
Doesn't hold for antimorphic involution
Counter example: $u = x\theta(x)x$, $v = \theta(x)x$. Then $u \perp_\theta v$; but $\theta(u) \not\perp_\theta \theta(v)$

$(u, v)_\theta = x \xrightarrow{?} (\theta(u), \theta(v))_\theta = \theta(x)$

Observation

$u \perp_\theta v \xrightarrow{?} \theta(u) \perp_\theta \theta(v)$

Holds for morphic involution
Doesn't hold for antimorphic involution
Counter example: $u = x\theta(x)x$, $v = \theta(x)x$. Then $u \perp_\theta v$; but $\theta(u) \not\perp_\theta \theta(v)$

$(u, v)_\theta = x \xrightarrow{?} (\theta(u), \theta(v))_\theta = \theta(x)$

Holds for morphic involution
Doesn't hold for antimorphic involution
Counter example: $u = x\theta(x)$, $v = x$ Then $(u, v)_\theta = x$; but $\theta(u) \perp_\theta \theta(v)$.

Future Work

- Relative θ-primitivity in set - pairwise relatively θ-primitive
- Notion of (strong) relative θ-primitivity
- Studying other BW operations


THANK YOU

Proposition (Gawrychowski et al.)

Let $\theta : \Sigma^* \to \Sigma^*$ be an (anti)morphism and $w \in \Sigma^*$ be a given word with $|w| = n$.

1. One can identify in time $\mathcal{O}(n^{3.5})$ the triples (i, j, k) with $w[i..j] \in \{t, \theta(t)\}^k$ for a proper factor of t of $w[i..j]$.
2. One can identify in time $\mathcal{O}(n^2 k)$ the pairs (i, j) such that $w[i..j] \in \{t, \theta(t)\}^k$ for a proper factor t of $w[i..j]$, when k is also given as input.

For a non-erasing θ we solve (1) in $\Theta(n^3)$ time and (2) in $\Theta(n^2)$ time. For a literal θ we solve (1) in $\Theta(n^2 lg\ n)$ time and (2) in $\Theta(n^2)$ time.