Radovan Červený
cervera3@fit.cvut.cz
presented:

Unavoidable sets of constant length

Jean-Marc Champarnaud, Georges Hansel, Dominique Perrin
https://doi.org/10.1142/S0218196704001700

Definitions

- A is an *alphabet*; A^* is a *set of all words* on A; for $w \in A^*$: word *length* $|w|$; *prefix* p if $w = pu$; *proper prefix* p if $p \neq w$; *suffix* and *proper suffix* defined symmetrically; *factor* x if $w = pxq$.
- sequence $w = (a_n)_{n \in Z}$ is a *two-sided infinite word* on A; x is a factor of w if we have an index n such that $x = a_n a_{n+1} \dots a_{n+|x|-1}$; two-sided infinite word is *periodic* if it is a repetition of some finite word u.
- $<$ ordering on A^* is an *alphabetic* ordering.
- two words x, y are *conjugate* if there exist words u, v such that $x = uv$ and $y = vu$; *conjugacy* is an equivalence on A^*.
- a word is *primitive* if it is not of the form r^n for $r \in A^*$ and $n > 1$.
- a word is *minimal* if it is the least one in its conjugacy class.
- a word is a *Lyndon word* if it is both primitive and minimal
- M is a set of minimal words, P is a set of prefixes of minimal words, L is a set of Lyndon words.
- a word w is said to be a *sesquipower* of word x if it is of the form $w = x^n p$, $n > 0$ and p is a proper prefix of x.
- a *division* of word $w \in P$ is a pair (l^n, u) such that $w = l^n u$ where $l \in L$, $n > 0$ and $u \in A^*$, $|u| < |l|$.
- the *main division* of word $w \in P$ is a division (l^n, u) where l is the shortest Lyndon word that allows such division, the word l^n is the *principal part* of w, denoted by $p(w)$, and u is the *rest*, denoted by $r(w)$.
- a set $I \subset A^*$ is *unavoidable* if every two-sided infinite word admits at least one factor in I.

The Lemmas, the Theorem and the Consequence

Lemma 1: The following are equivalent for any non-empty word w:

1. w is a Lyndon word.
2. for any non-empty u, v such that $w = uv$, we have $w < vu$.
3. for any non-empty proper suffix s of w, we have $w < s$.

Lemma 2: A word w is minimal if and only if it is a power of a Lyndon word. This Lyndon word is uniquely determined.

Lemma 3: Let w be a prefix of a minimal word. Then any prefix of w is less than or equal to the suffix of the same length of w.

Lemma 4: The following are equivalent for any word w:

1. w is a non-empty prefix of a minimal word.
2. w is a sesquipower of a Lyndon word.

Lemma 5: The cardinality of I_k is at least the number of conjugacy classes of words of length k.

Lemma 6: The following are equivalent for any finite $I \subset A^*$ set of words:

1. I is unavoidable.
2. every two-sided infinite periodic word admits at least one factor in I.

Lemma 7: Let λ, l be Lyndon words, with λ prefix of l. Let s be a proper suffix of l, with $|s| < |\lambda|$. Then for all $n > 0$, the word $w = \lambda^n s$ is a Lyndon word.

Lemma 8: Let w be a prefix of a minimal word and let (λ^n, u) be its main division. Let u' be a word, with $|u'| = |u|$, such that $w' = \lambda^n u'$ is also a prefix of a minimal word. Then the main division of w' is (λ^n, u').

Lemma 9: Let l be a Lyndon word. Let $n > 0$ be the smallest integer such that $|l^n| > k$. Then the word l^{n+1} has a factor in I_k.

Theorem: The set $I_k = \{r(m)p(m) | m \in M_k\}$ is unavoidable.

Consequence: The cardinality of I_k is equal to the number of conjugacy classes of words of length k.