Self-Shuffling Words

Émilie Charlier

(Joint work with Teturo Kamae, Svetlana Puzynina and Luca Zamboni)

Département de mathématiques, Université de Liège

Challenges in Combinatorics on Words, Toronto, April 2013
A finite word $w$ is obtained as a shuffle of two words $x$ and $y$ if it can be factorized as

$$w = \prod_{i=1}^{n} a_i b_i$$

with

$$x = \prod_{i=1}^{n} a_i \quad \text{and} \quad y = \prod_{i=1}^{n} b_i.$$

Can $x$ be obtained as a shuffle of a word $y$ with itself?
Definition

An infinite word $x \in \mathbb{A}^\mathbb{N}$ is self-shuffling if it admits the factorizations:

$$x = \prod_{i=1}^{+\infty} U_i V_i = \prod_{i=1}^{+\infty} U_i = \prod_{i=1}^{+\infty} V_i$$

for some $U_i, V_i \in \mathbb{A}^+$. 
Self-shuffling and periodicity

Every purely periodic word \( x = u^\omega \) is self-shuffling:

\[
x = \prod_{i=1}^{+\infty} (uu) = \prod_{i=1}^{+\infty} u.
\]

Every letter of a self-shuffling word must occur an infinite number of times.

Hence the word \( 01^\omega \) is not self-shuffling.

The word \( 0110(0011)^\omega \) is self-shuffling but \( 011(10)^\omega \) is not.
The Fibonacci word $F = F_1 F_2 F_3 \cdots$, where $F_i \in \{0, 1\}$, is self-shuffling:

$$F = \prod_{i=1}^{+\infty} \sigma(F_i) F_i = \prod_{i=1}^{+\infty} \sigma(F_i) = \prod_{i=1}^{+\infty} F_i$$

where $\sigma: 0 \mapsto 01; 1 \mapsto 0$.

It suffices to observe that $\sigma^2(0) = \sigma(0)0$ and $\sigma^2(1) = \sigma(1)1$.

$$F = 010\ 01\ 010\ 010\ 01\ 010\ 01\ 010\ 010\ 010\ 010\ 010\ \cdots$$

$\sigma^2(0)\ \sigma^2(1)\ \sigma^2(0)\ \sigma^2(0)\ \sigma^2(1)\ \sigma^2(0)\ \sigma^2(0)\ \sigma^2(1)\ \sigma^2(0)\ \sigma^2(0)\ \sigma^2(0)$
The morphic image of a self-shuffling word is again self-shuffling:

If

\[ x = \prod_{i=1}^{+\infty} U_i V_i = \prod_{i=1}^{+\infty} U_i = \prod_{i=1}^{+\infty} V_i \]

then

\[ \mu(x) = \prod_{i=1}^{+\infty} \mu(U_i) \mu(V_i) = \prod_{i=1}^{+\infty} \mu(U_i) = \prod_{i=1}^{+\infty} \mu(V_i). \]

Can be used as a tool for showing that a word is not the morphic image of another word.
A necessary condition

A finite word is Abelian border-free if no proper suffix is Abelian equivalent to a proper prefix.

- A self-shuffling word must begin in only finitely many Abelian border-free prefixes.

\[ x = \text{---} \quad \text{---} \quad \text{---} \]
A necessary condition

A finite word is Abelian border-free if no proper suffix is Abelian equivalent to a proper prefix.

- A self-shuffling word must begin in only finitely many Abelian border-free prefixes.
A necessary condition

A finite word is Abelian border-free if no proper suffix is Abelian equivalent to a proper prefix.

- A self-shuffling word must begin in only finitely many Abelian border-free prefixes.
• The word $0F$ is not self-shuffling:

We know $F$ begins in arbitrarily large palindromes $B$ followed by 1.

Hence $0F$ begins in arbitrarily large words of the form $0B1$.

Those words $0B1$ are all Abelian border-free.
The paper-folding word is not self-shuffling:

\[
\begin{align*}
0_1\ 0_1\ 0_1\ 0_1\ 0_1\ 0_1\ 0_1\ 0_1\ 0_1\ 0_1\ 0_1\ 0_1\ \\
001_011_001_011_001_011_001_011\ \\
0010011_0011011_0010011_0011011\ \\
001001100011011_0010011_0011011\ \\
00100110001101100010011_0011011\ \\
0010011000110110001001110010111011\ \\
\end{align*}
\]

One can check that all prefixes of length \(2^j - 1\) are Abelian border-free:

\[
0,\ 001,\ 0010011,\ 001001100011011,\ldots
\]
Main results - Lyndon words

An infinite word $x$ is **Lyndon** if $x <_{\text{lex}} T^n(x)$ for all $n \geq 1$.

**Theorem 1**

Let $x$ and $y$ be in the shift orbit closure of an infinite word $z$, and suppose that $x$ is Lyndon. Then, if $w$ can be obtained as a shuffle of $x$ and $y$, then $w <_{\text{lex}} y$.

In particular, Lyndon infinite words are never self-shuffling.
Consequences:

- This gives another proof that $0F$ is not self-shuffling since it is Lyndon.
- It is well known that the first shift

$$x = 110100110010110 \cdots$$

of the Thue-Morse word is Lyndon, and hence is not self-shuffling. Yet it can be verified that $x$ begins in only a finite number of Abelian border-free words.
Thue-Morse word

Theorem 2
The Thue-Morse word $T = 011010011001 \cdots$ fixed by the morphism $\tau : 0 \mapsto 01; 1 \mapsto 10$ is self-shuffling.

$$T = 0110100110010110100101101001101 \cdots$$
Sturmian words

Here \( \rho(x) \) designates the intercept of the Sturmian word \( x \).

**Theorem 3**
Let \( S, M \) and \( L \) be Sturmian words of the same slope satisfying \( S \leq_{\text{lex}} M \leq_{\text{lex}} L \). Then \( M \) can be obtained as a shuffle of \( S \) and \( L \) iff the following conditions are verified:

- If \( \rho(M) = \rho(S) \) then \( \rho(L) \neq 0 \)
- If \( \rho(M) = \rho(L) \) then \( \rho(S) \neq 0 \).

In particular (taking \( S = M = L \)), we obtain

**Corollary**
A Sturmian word \( x \in \{0, 1\}^\mathbb{N} \) is self-shuffling iff \( x \) is not of the form \( axC \) where \( a \in \{0, 1\} \) and \( C \) is a characteristic Sturmian word.
As an application, we recover the following result:

**Theorem (Yasutomi 1999, Berthé-Ei-Ito-Rao 2007)**

Let $x \in \{0, 1\}^\mathbb{N}$ be a characteristic Sturmian word. If $y$ is a pure morphic word in the orbit of $x$, then $y \in \{x, 0x, 1x, 01x, 10x\}$. 
References

- Romeo Rizzi and Stéphane Vialette, *On recognizing words that are squares for the shuffle product*, to appear in LNCS (CSR 2013)