A Synergic Approach to the Minimal Uncompletable Words Problem

Sandrine Julia, Arnaud Malapert, **Julien Provillard**

Université Côte d'Azur, CNRS, I3S, France

Journée MDSC, Nice, France

Outline

1. Background
2. Theoretical results
3. Computation of a minimal uncompletable word
4. Experimental search using a genetic algorithm
5. Results

Outline

1. Background
2. Theoretical results
3. Computation of a minimal uncompletable word
4. Experimental search using a genetic algorithm
5. Results

Background

Field : Combinatorics of words

- An alphabet is a finite set whose elements are called letters.
- A word is a finite sequence of letters.
- A language is a set of words.

Background

- $|u|$ denotes the length of a word u and ε is the only word of length 0.
- uv denotes the concatenation of the words u and v, *i.e.* the word $u_0 \dots u_{|u|-1} v_0 \dots v_{|v|-1}$.
- XY denotes the concatenations of the languages X and Y, $XY = \{uv : u \in X, v \in Y\}$.
- X^* denotes the finite concatenations of words in X, $X^* = \cup_{n \geq 0} X^n$.

Background

- If $x = uvw$, then
  - u is a prefix of x,
  - w is a suffix of x,
  - v is a factor of x.
- $Pref(X), Suff(X), Fact(X)$ denote the sets of prefixes, suffixes and factors of words in a language X, respectively.

Background

- A language $\mathcal{L}$ is *factorial* if it is closed under factors.
- A *minimal forbidden word* is a word which does not appear in $\mathcal{L}$ but all of its proper factors do.
- The study of the set of minimal forbidden words of $\mathcal{L}$ gives a lot of informations about $\mathcal{L}$ itself.

Background

- A language X is *complete* if any word is a factor of X^*.
- A word which is not a factor of X^* is *uncompletable*.
- An uncompletable word whose proper factors are factors of X^* is a *minimal uncompletable word*.
- Minimal uncompletable words are special cases of minimal forbidden words.

Background

Restivo's conjecture [1981]

Let X be a non-complete finite set and k be the length of a longest word in X, then X admits an uncompletable word of length at most $2k^2$.

Background

Restivo's conjecture [1981]

Let X be a non-complete finite set and k be the length of a longest word in X, then X admits an uncompletable word of length at most $2k^2$.

- Counter-examples for $7 \leq k \leq 12$
  [Fici, Pribavkina and Sakarovitch 2010]

Background

Restivo's conjecture [1981]

Let X be a non-complete finite set and k be the length of a longest word in X, then X admits an uncompletable word of length at most $2k^2$.

- Counter-examples for $7 \leq k \leq 12$ [Fici, Pribavkina and Sakarovitch 2010]
- Counter-examples for $k \geq 4$ [Gusev and Pribavkina 2011]

Background

Restivo's conjecture [1981]

Let X be a non-complete finite set and k be the length of a longest word in X, then X admits an uncompletable word of length at most $2k^2$.

- Counter-examples for $7 \leq k \leq 12$ [Fici, Pribavkina and Sakarovitch 2010]
- Counter-examples for $k \geq 4$ [Gusev and Pribavkina 2011]
- In all cases, an uncompletable word of quadratic length is given.

Background

Reformulation of Restivo's conjecture

Let X be a non-complete finite set and k be the length of a longest word in X, then X admits an uncompletable word of length $O(k^2)$.

Our goal

To combine a theoretical approach and an experimental approach to obtain new clues and data on the minimal uncompletable words problem.

Outline

1. Background
2. Theoretical results
3. Computation of a minimal uncompletable word
4. Experimental search using a genetic algorithm
5. Results

Formal definitions

Let Σ be a finite alphabet and $X \subseteq \Sigma^*$ a language.

- X is *complete* if $Fact(X^*) = \Sigma^*$.
- $u \in \Sigma^*$ is *uncompletable* ($w.r.t.\ X$) if $u \notin Fact(X^*)$.
- An uncompletable word $u \in \Sigma^*$ is *minimal* if $Fact(u) \setminus \{u\} \subseteq Fact(X^*)$.

Notations

- U is the set of uncompletable words $w.r.t.$ X.
- M is the set of minimal uncompletable words $w.r.t.$ X.
- $uwl(X) = \min\{|u| : u \in M\}$.
- $UWL(k,d) = \max\left\{uwl(X) : X \subseteq \Sigma^{\leq k}, |\Sigma| = d\right\}$.
- $UWL(k) = UWL(k,2)$

Notations

- U is the set of uncompletable words $w.r.t.\ X$.
- M is the set of minimal uncompletable words $w.r.t.\ X$.
- $uwl(X) = \min \{|u| : u \in M\}$.
- $UWL(k, d) = \max \left\{ uwl(X) : X \subseteq \Sigma^{\leq k}, |\Sigma| = d \right\}$.
- $UWL(k) = UWL(k, 2)$

Restivo's conjecture

For all $d \geq 2$, $UWL(k, d) = O(k^2)$.

The case of binary alphabet

Proposition [Carpi, D'Alessandro 2014]

For all $k \geq 1$, $d > 2$,

$$UWL(k, d) \leq \left\lceil \frac{UWL\left(\lceil \log_2(d) \rceil\, k\right)}{\lfloor \log_2(d) \rfloor} \right\rceil$$

The case of binary alphabet

Proposition [Carpi, D'Alessandro 2014]

For all $k \geq 1$, $d > 2$,

$$UWL(k,d) \leq \left\lceil \frac{UWL\left(\lceil \log_2(d) \rceil \, k\right)}{\lfloor \log_2(d) \rfloor} \right\rceil$$

Restivo's conjecture

$UWL(k) = O(k^2)$.

The case of binary alphabet

Proposition [Carpi, D'Alessandro 2014]

For all $k \geq 1$, $d > 2$,

$$UWL(k,d) \leq \left\lceil \frac{UWL\left(\lceil \log_2(d) \rceil \, k\right)}{\lfloor \log_2(d) \rfloor} \right\rceil$$

Restivo's conjecture

$UWL(k) = O(k^2)$.

From now on, Σ is the binary alphabet $\{a, b\}$.

Structure of M

Given a set X, we define the following sets of words:

- $S = Suff(X) \setminus Suff(X)X^+$.
- $P = Pref(X) \setminus X^+ Pref(X)$.
- $Z = P\Sigma \setminus Pref(X)$.
- $Z_\triangleleft = \Sigma S \setminus Suff(X)$.

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b a a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of $baabaaabaab$:

b a a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of $baabaaabaab$:

b a a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b a a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a a b a a a b a a b

b a	a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of $baabaaabaab$:

[b] a a b a a a b a a b

[b a] a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of *baabaaabaab*:

[b] [a a b a] a a b a a b

[b a] a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a a b a	a a b a a b

b a	a b a a a b a a b

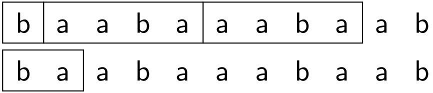
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a a b a	a a b a	a b

b a	a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a a b a	a a b a	a b

b a	a b a a a b a a b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a a b a	a a b a	a b	
b a	a b a a a b a a b			

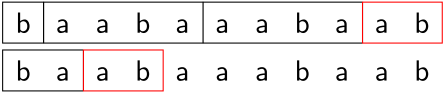
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of *baabaaabaab*:



Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b

Proposition

$M \subseteq P^{-1}Z \cup SX^*Z$.

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b

Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b

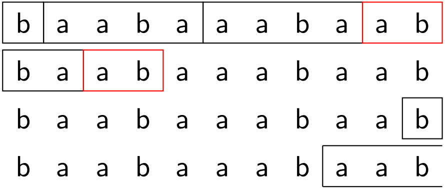
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of *baabaaabaab*:



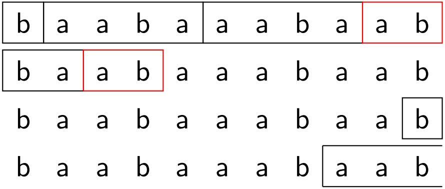
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



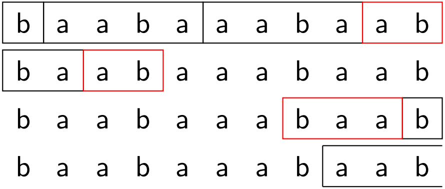
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



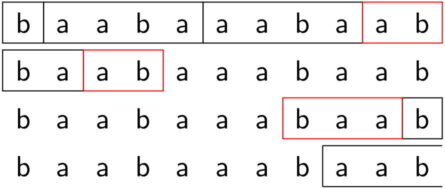
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



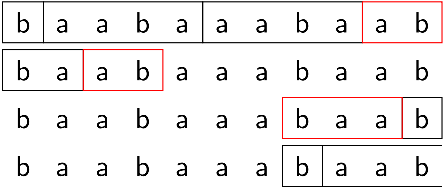
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



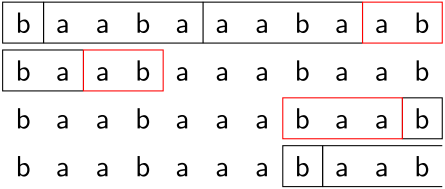
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



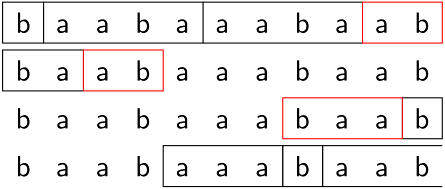
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



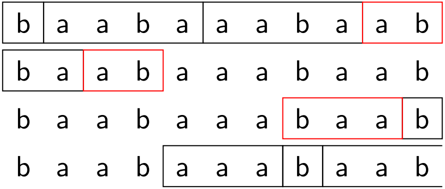
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



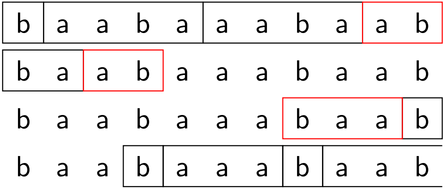
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



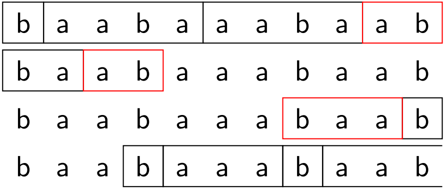
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_{\triangleleft} = \{baa\}$.

Possible factorizations of *baabaaabaab*:



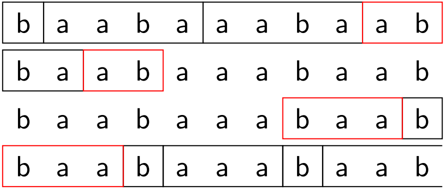
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of $baabaaabaab$:



Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:

b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b
b	a	a	b	a	a	a	b	a	a	b

Proposition

$M \subseteq Z_\triangleleft S^{-1} \cup Z_\triangleleft X^* P$.

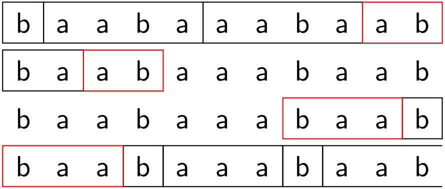
Structure of M

The words in Z stops the factorizations on $Fact(X^*) = SX^*P$.

Example for $X = \{aaa, aaba, b, ba\}$.

$P = \{\varepsilon, a, aa, aab\}$, $S = \{\varepsilon, a, aa\}$, $Z = \{ab, aabb\}$ and $Z_\triangleleft = \{baa\}$.

Possible factorizations of *baabaaabaab*:



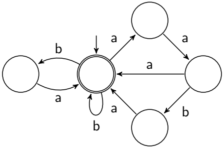
Proposition

An uncompletable word w is minimal if and only if
$w \in (P^{-1}Z \cup SX^*Z) \cap (Z_\triangleleft S^{-1} \cup Z_\triangleleft X^*P)$.

Outline

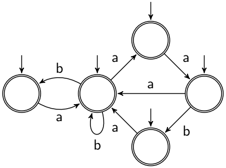
1. Background
2. Theoretical results
3. Computation of a minimal uncompletable word
4. Experimental search using a genetic algorithm
5. Results

Usual algorithm



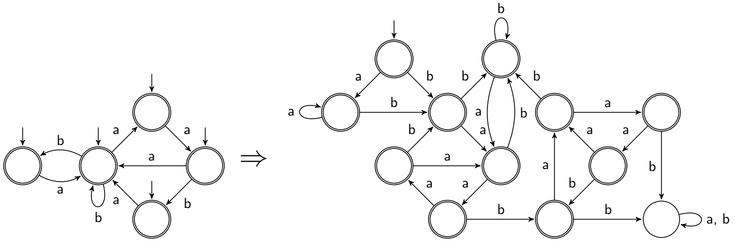
- Build a semi-flower automaton for X.

Usual algorithm



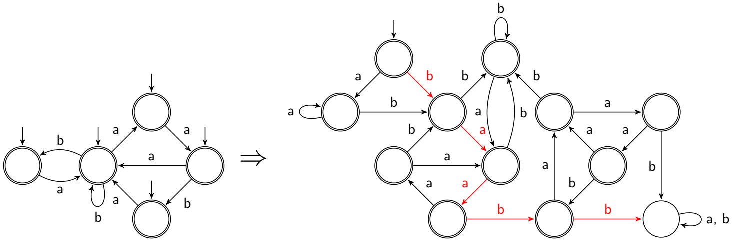
- Build a semi-flower automaton for X.
- Make all states initial and accepting.

Usual algorithm



- Build a semi-flower automaton for X.
- Make all states initial and accepting.
- Determinize it.

Usual algorithm



- Build a semi-flower automaton for X.
- Make all states initial and accepting.
- Determinize it.
- Find a shortest path to the sink state (if it exists).

Cost and improvements

- Exponential blow-up (determinization).
- Build directly a smaller automaton recognizing $Fact(X^*)$.
- Incremental build during the BFS.
- Detection of non-optimal paths.
- Still exponential but effectively more efficient.

Res automaton

The automaton $Res(X)$, $X \subseteq \Sigma^{\leq k}$:

- States are subsets of $\Sigma^{\leq k}$,
- Initial state is $X \cup S \setminus \{\varepsilon\}$,
- Transition function

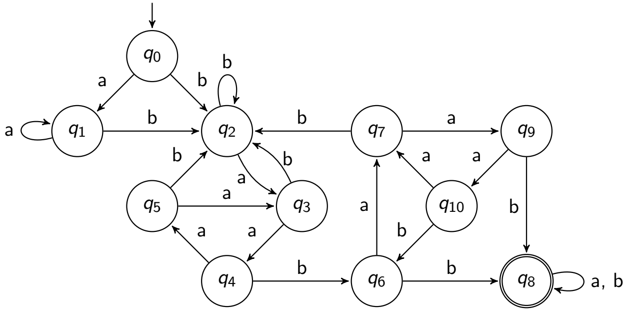
$$res(Y, \sigma) = \begin{cases} \sigma^{-1} Y & \text{if } \sigma \notin Y \\ (\sigma^{-1} Y \setminus \{\varepsilon\}) \cup X & \text{otherwise} \end{cases}$$

Intuitively, a state gathers the suffixes of X which can achieve a factorization on $SX*$.

There exists a morphism from the previous automaton to this one.

Res automaton

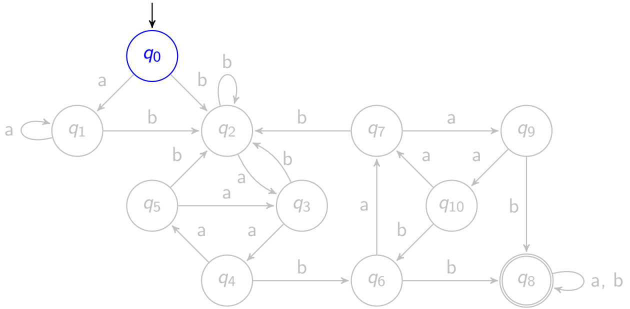
Example for $X = \{aaa, aaba, b, ba\}$.



$$q_0 = X \cup \{a, aa\}$$
$$q_1 = X \cup \{a, aa, aba\}$$
$$q_2 = X \cup \{a\}$$
$$q_3 = X \cup \{aa, aba\}$$
$$q_4 = \{a, aa, aba, ba\}$$
$$q_5 = X \cup \{a, ba\}$$
$$q_6 = \{a\}$$
$$q_7 = X$$
$$q_8 = \varnothing$$
$$q_9 = \{aa, aba\}$$
$$q_{10} = \{a, ba\}$$

Res automaton

Example for $X = \{aaa, aaba, b, ba\}$.



$q_0 = X \cup \{a, aa\}$ $\quad$ $q_4 = \{a, aa, aba, ba\}$ $\quad$ $q_8 = \varnothing$

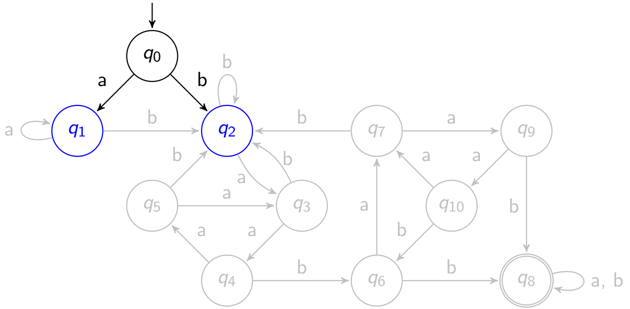
$q_1 = X \cup \{a, aa, aba\}$ $\quad$ $q_5 = X \cup \{a, ba\}$ $\quad$ $q_9 = \{aa, aba\}$

$q_2 = X \cup \{a\}$ $\quad$ $q_6 = \{a\}$ $\quad$ $q_{10} = \{a, ba\}$

$q_3 = X \cup \{aa, aba\}$ $\quad$ $q_7 = X$

Res automaton

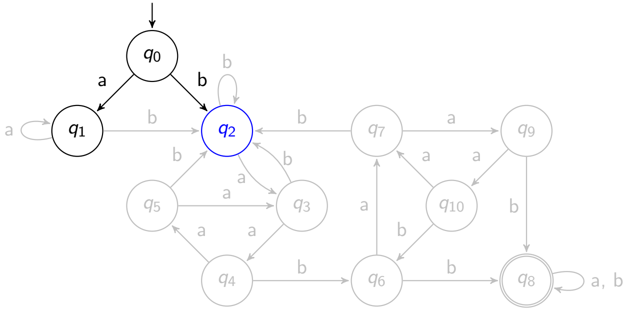
Example for $X = \{aaa, aaba, b, ba\}$.



$$q_0 = X \cup \{a, aa\} \qquad q_4 = \{a, aa, aba, ba\} \qquad q_8 = \varnothing$$
$$q_1 = X \cup \{a, aa, aba\} \qquad q_5 = X \cup \{a, ba\} \qquad q_9 = \{aa, aba\}$$
$$q_2 = X \cup \{a\} \qquad q_6 = \{a\} \qquad q_{10} = \{a, ba\}$$
$$q_3 = X \cup \{aa, aba\} \qquad q_7 = X$$

Res automaton

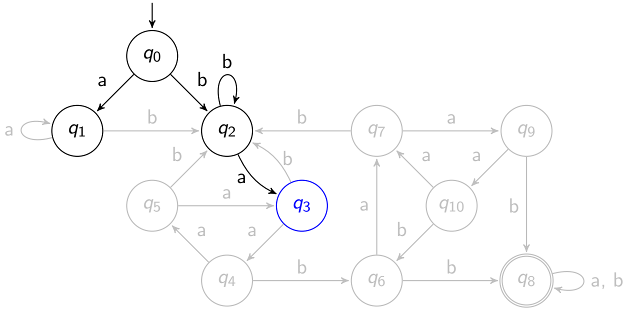
Example for $X = \{aaa, aaba, b, ba\}$.



$$
\begin{array}{lll}
q_0 = X \cup \{a, aa\} & q_4 = \{a, aa, aba, ba\} & q_8 = \varnothing \\
q_1 = X \cup \{a, aa, aba\} & q_5 = X \cup \{a, ba\} & q_9 = \{aa, aba\} \\
q_2 = X \cup \{a\} & q_6 = \{a\} & q_{10} = \{a, ba\} \\
q_3 = X \cup \{aa, aba\} & q_7 = X &
\end{array}
$$

Res automaton

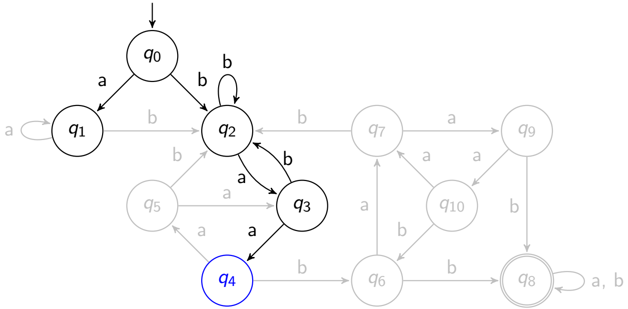
Example for $X = \{aaa, aaba, b, ba\}$.



$q_0 = X \cup \{a, aa\}$
$q_1 = X \cup \{a, aa, aba\}$
$q_2 = X \cup \{a\}$
$q_3 = X \cup \{aa, aba\}$
$q_4 = \{a, aa, aba, ba\}$
$q_5 = X \cup \{a, ba\}$
$q_6 = \{a\}$
$q_7 = X$
$q_8 = \varnothing$
$q_9 = \{aa, aba\}$
$q_{10} = \{a, ba\}$

Res automaton

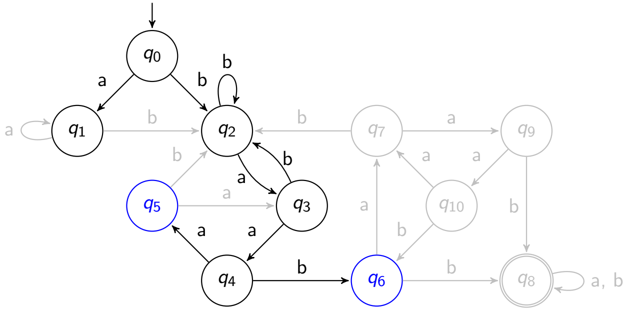
Example for $X = \{aaa, aaba, b, ba\}$.



$$q_0 = X \cup \{a, aa\} \qquad q_4 = \{a, aa, aba, ba\} \qquad q_8 = \varnothing$$
$$q_1 = X \cup \{a, aa, aba\} \qquad q_5 = X \cup \{a, ba\} \qquad q_9 = \{aa, aba\}$$
$$q_2 = X \cup \{a\} \qquad q_6 = \{a\} \qquad q_{10} = \{a, ba\}$$
$$q_3 = X \cup \{aa, aba\} \qquad q_7 = X$$

Res automaton

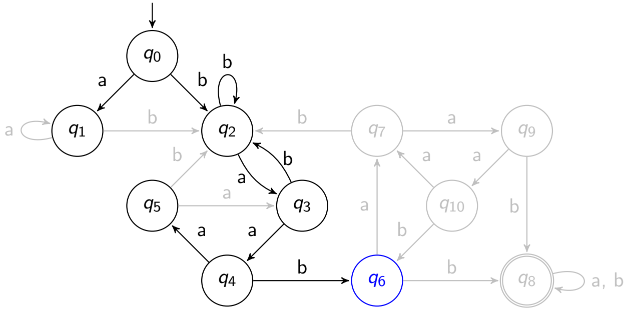
Example for $X = \{aaa, aaba, b, ba\}$.



$$q_0 = X \cup \{a, aa\} \qquad q_4 = \{a, aa, aba, ba\} \qquad q_8 = \varnothing$$
$$q_1 = X \cup \{a, aa, aba\} \qquad q_5 = X \cup \{a, ba\} \qquad q_9 = \{aa, aba\}$$
$$q_2 = X \cup \{a\} \qquad q_6 = \{a\} \qquad q_{10} = \{a, ba\}$$
$$q_3 = X \cup \{aa, aba\} \qquad q_7 = X$$

Res automaton

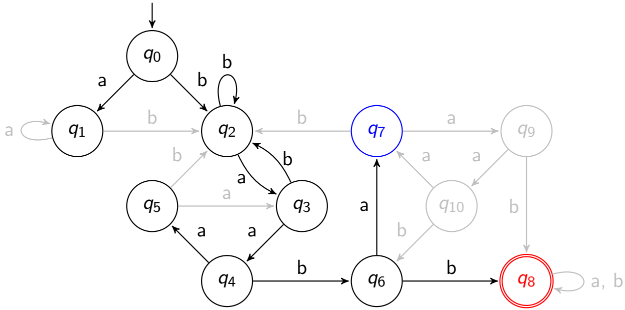
Example for $X = \{aaa, aaba, b, ba\}$.



$$q_0 = X \cup \{a, aa\} \qquad q_4 = \{a, aa, aba, ba\} \qquad q_8 = \varnothing$$
$$q_1 = X \cup \{a, aa, aba\} \qquad q_5 = X \cup \{a, ba\} \qquad q_9 = \{aa, aba\}$$
$$q_2 = X \cup \{a\} \qquad q_6 = \{a\} \qquad q_{10} = \{a, ba\}$$
$$q_3 = X \cup \{aa, aba\} \qquad q_7 = X$$

Res automaton

Example for $X = \{aaa, aaba, b, ba\}$.



$q_0 = X \cup \{a, aa\}$ $\quad q_4 = \{a, aa, aba, ba\}$ $\quad q_8 = \emptyset$

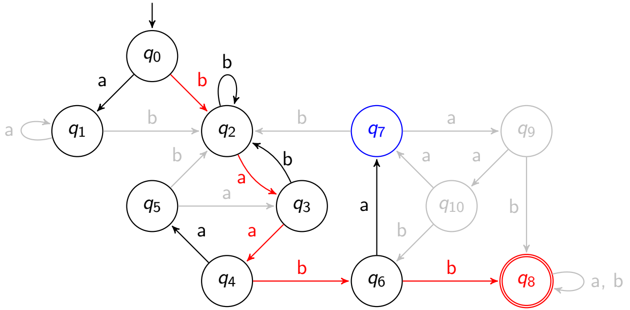
$q_1 = X \cup \{a, aa, aba\}$ $\quad q_5 = X \cup \{a, ba\}$ $\quad q_9 = \{aa, aba\}$

$q_2 = X \cup \{a\}$ $\quad q_6 = \{a\}$ $\quad q_{10} = \{a, ba\}$

$q_3 = X \cup \{aa, aba\}$ $\quad q_7 = X$

Res automaton

Example for $X = \{aaa, aaba, b, ba\}$.



$q_0 = X \cup \{a, aa\}$

$q_1 = X \cup \{a, aa, aba\}$

$q_2 = X \cup \{a\}$

$q_3 = X \cup \{aa, aba\}$

$q_4 = \{a, aa, aba, ba\}$

$q_5 = X \cup \{a, ba\}$

$q_6 = \{a\}$

$q_7 = X$

$q_8 = \varnothing$

$q_9 = \{aa, aba\}$

$q_{10} = \{a, ba\}$

Outline

1 Background

2 Theoretical results

3 Computation of a minimal uncompletable word

4 Experimental search using a genetic algorithm

5 Results

The search

- Our goal is to explore the lattice $(\mathcal{P}(\Sigma^{\leq k}), \subseteq)$.
- More precisely, the border between non-complete and complete set
  - $X \subseteq Y$ and X complete implies Y complete.
  - $X \subseteq Y$ and Y not complete implies X not complete and $uwl(X) \leq uwl(Y)$.
- The value $UWL(k)$ can be found on the border.
- Use of a SAT-solver to obtain sets to process.

Why a SAT-solver?

- Easy encoding of sets by their indicator functions.
- Set properties are easy to implement.
- Possibility to dynamically add a huge number of clauses.
- Produces good candidates to test.
  - No superset of a known complete set and no subset of a known non-complete set.
  - At least one word of length k.
  - Z and $Z_{\triangleleft}$ are both non-empty.
  - No symetric of a known set.
  - Closed by concatenation on $\Sigma^{\leq k}$.
- Detects the exhaustion of the search space (the underlying formula become unsatisfiable).

Genetic algorithm

- Asking the SAT-solver a solution gives an assignation with the minimal number of variables set to true, *i.e.* a new set of minimal cardinality.
- Basically, we could visit the lattice by layers of sets of same cardinality. This is an inneficient strategy.
- To avoid that, the SAT-solver is driven by a genetic algorithm.

Genetic algorithm

- Genetic algorithm are biology inspired probabilistic algorithm used in optimization problems.
- A population of individuals evolves step by step simulating the survival of the fittest, reproduction and genetic phenoma as crossovers and mutations.
- In our background,
  - Individuals are subsets of $\Sigma^{\leq k}$.
  - The fitness function is *uwl*.
  - No crossover.
  - Mutations are extensions of sets with new words (if possible).

Outline

1. Background
2. Theoretical results
3. Computation of a minimal uncompletable word
4. Experimental search using a genetic algorithm
5. Results

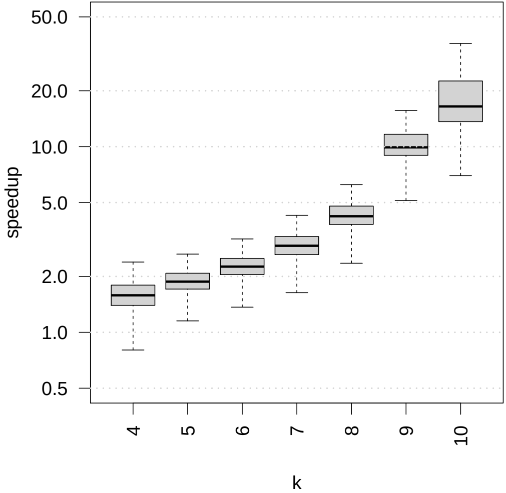
Performances

Speedup of residuals over flower-automata based algorithm on random instances.

k \ Density	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	1.13	1.14	1.11	1.35	1.15	1.21	1.33	1.52	1.51
3	1.04	**0.96**	1.26	1.23	1.26	1.53	1.50	1.62	1.74
4	**0.99**	1.06	1.00	**0.82**	1.03	1.41	1.98	1.89	2.03
5	**0.63**	**0.85**	1.09	1.27	1.33	1.69	1.34	2.27	2.29
6	**0.73**	1.04	1.30	1.55	1.96	2.05	3.22	3.34	4.58
7	1.02	1.13	1.54	1.85	2.85	3.78	5.02	6.31	8.36
8	**0.81**	1.24	1.72	2.36	3.41	4.58	6.89	9.70	14.66
9	**0.83**	1.50	2.40	3.12	4.56	7.35	10.38	16.74	24.32
10	1.03	2.10	3.76	5.29	7.33	11.37	18.33	24.67	42.89

Performances

Speedup of residuals over flower-automata based algorithm on a trace of the search.



Results

- For a binary alphabet,

k	1	2	3	4	5
$UWL(k)$	1	5	13	31	68

k	6	7	8	9	10
$UWL(k)$	≥ 153	≥ 279	≥ 524	≥ 553	≥ 570

- $UWL(k)$ is not a polynomial of degree less than 5.
- Best solutions need small words.
- For the best solutions, Z is a singleton and the word $z \in Z$ has length k.

Conclusion

Conjecture

If X is a maximal ($w.r.t.$ the inclusion) non-complete set, then Z is a singleton $\{z\}$ and z has length k.

- Computation of more values and lower bounds of UWL.
- Structure of the maximal non-complete sets.
- The exact role of the set Z.
- Search for alphabet of size greater than 2.

Thank you for your attention !