Computing Equality-Free String Factorisations

Markus L. Schmid

Trier University, Germany

CiE 2015

Basic Concepts

A finite alphabet	$\Sigma = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$
strings	$w = \mathtt{daabaccabd}$
string factorisations	$(\mathtt{daa}, \mathtt{b}, \mathtt{acca}, \mathtt{bd})$ $\mathtt{daa} \cdot \mathtt{b} \cdot \mathtt{acca} \cdot \mathtt{bd}$

String factorisations

Let $p = (u_1, u_2, \ldots, u_k)$ be a factorisation.

- $\mathsf{sf}(p) = \{u_1, u_2, \ldots, u_k\}$ *set of factors*,
- $\mathsf{s}(p) = k$ *size*,
- $\mathsf{c}(p) = |\,\mathsf{sf}(p)|$ *cardinality*,
- $\mathsf{w}(p) = \max\{|u_i| \mid 1 \leq i \leq k\}$ *width*.

String factorisations

Let $p = (u_1, u_2, \ldots, u_k)$ be a factorisation.

- $\mathsf{sf}(p) = \{u_1, u_2, \ldots, u_k\}$ *set of factors*,
- $\mathsf{s}(p) = k$ *size*,
- $\mathsf{c}(p) = |\,\mathsf{sf}(p)|$ *cardinality*,
- $\mathsf{w}(p) = \max\{|u_i| \mid 1 \leq i \leq k\}$ *width*.

Central notion of this talk

A factorisation p is *equality-free* if $\mathsf{s}(p) = \mathsf{c}(p)$.
(p is *repetitive* $\Leftrightarrow$ p is not equality-free).

String factorisations

Let $p = (u_1, u_2, \ldots, u_k)$ be a factorisation.

- $\mathsf{sf}(p) = \{u_1, u_2, \ldots, u_k\}$ *set of factors*,
- $\mathsf{s}(p) = k$ *size*,
- $\mathsf{c}(p) = |\,\mathsf{sf}(p)|$ *cardinality*,
- $\mathsf{w}(p) = \max\{|u_i| \mid 1 \leq i \leq k\}$ *width*.

Central notion of this talk

A factorisation p is *equality-free* if $\mathsf{s}(p) = \mathsf{c}(p)$.
(p is *repetitive* $\Leftrightarrow$ p is not equality-free).

Example

$p = \mathtt{aab} \cdot \mathtt{ba} \cdot \mathtt{cba} \cdot \mathtt{aab} \cdot \mathtt{ba} \cdot \mathtt{aab}$.

- $\mathsf{sf}(p) = \{\mathtt{aab}, \mathtt{ba}, \mathtt{cba}\}$, $\mathsf{s}(p) = 6$, $\mathsf{c}(p) = 3$ and $\mathsf{w}(p) = 3$,
- p is not equality-free (i. e., p is repetitive).

Computing equality-free factorisations

Find equality-free factorisation with large size

Computing equality-free factorisations

Find equality-free factorisation with large size

`abbcbaabbc`

Computing equality-free factorisations

Find equality-free factorisation with large size

`abbc` · `ba` · `abbc`

Computing equality-free factorisations

Find equality-free factorisation with large size

$\texttt{ab} \cdot \texttt{bc} \cdot \texttt{ba} \cdot \texttt{ab} \cdot \texttt{bc}$

Computing equality-free factorisations

Find equality-free factorisation with large size

$$\texttt{ab} \cdot \texttt{bc} \cdot \texttt{ba} \cdot \texttt{abb} \cdot \texttt{c}$$

Computing equality-free factorisations

Find equality-free factorisation with large size

$$\texttt{ab} \cdot \texttt{bc} \cdot \texttt{ba} \cdot \texttt{a} \cdot \texttt{bb} \cdot \texttt{c}$$

Computing equality-free factorisations

Find equality-free factorisation with large size
Can we do better than 6?

$$\texttt{ab} \cdot \texttt{bc} \cdot \texttt{ba} \cdot \texttt{a} \cdot \texttt{bb} \cdot \texttt{c}$$

Computing equality-free factorisations

Find equality-free factorisation with large size

Can we do better than 6?

We need `a`, `b` and `c` as single factors!

`abbcbaabbc`

Computing equality-free factorisations

Find equality-free factorisation with large size

Can we do better than 6?

We need `a`, `b` and `c` as single factors!

$$\texttt{a} \cdot \texttt{b} \cdot \texttt{bcbaabb} \cdot \texttt{c}$$

Computing equality-free factorisations

Find equality-free factorisation with large size

Can we do better than 6?

We need `a`, `b` and `c` as single factors!

$$\texttt{a} \cdot \texttt{b} \cdot \texttt{bc} \cdot \texttt{ba} \cdot \texttt{abb} \cdot \texttt{c}$$

Computing equality-free factorisations

Find equality-free factorisation with large size

Can we do better than 6? No!

We need `a`, `b` and `c` as single factors!

$$\texttt{a} \cdot \texttt{b} \cdot \texttt{bc} \cdot \texttt{ba} \cdot \texttt{abb} \cdot \texttt{c}$$

Computing equality-free factorisations

Find equality-free factorisation with small width

`aabbccaabbcc`

Computing equality-free factorisations

Find equality-free factorisation with small width

$\mathtt{a} \cdot \mathtt{abbccaabbcc}$

Computing equality-free factorisations

Find equality-free factorisation with small width

$\mathtt{a} \cdot \mathtt{ab} \cdot \mathtt{bccaabbcc}$

Computing equality-free factorisations

Find equality-free factorisation with small width

$\mathtt{a} \cdot \mathtt{ab} \cdot \mathtt{b} \cdot \mathtt{ccaabbcc}$

Computing equality-free factorisations

Find equality-free factorisation with small width

$$\texttt{a} \cdot \texttt{ab} \cdot \texttt{b} \cdot \texttt{c} \cdot \texttt{caabbcc}$$

Computing equality-free factorisations

Find equality-free factorisation with small width

$$\mathtt{a} \cdot \mathtt{ab} \cdot \mathtt{b} \cdot \mathtt{c} \cdot \mathtt{ca} \cdot \mathtt{abbcc}$$

Computing equality-free factorisations

Find equality-free factorisation with small width

$$\texttt{a} \cdot \texttt{ab} \cdot \texttt{b} \cdot \texttt{c} \cdot \texttt{ca} \cdot \texttt{abb} \cdot \texttt{cc}$$

Computing equality-free factorisations

Find equality-free factorisation with small width
Can we do better than 3?

$$\texttt{a} \cdot \texttt{ab} \cdot \texttt{b} \cdot \texttt{c} \cdot \texttt{ca} \cdot \texttt{abb} \cdot \texttt{cc}$$

Computing equality-free factorisations

Find equality-free factorisation with small width

Can we do better than 3? Yes!

$$\texttt{aa} \cdot \texttt{b} \cdot \texttt{bc} \cdot \texttt{ca} \cdot \texttt{a} \cdot \texttt{bb} \cdot \texttt{cc}$$

Computing equality-free factorisations

Find equality-free factorisation with small width

`aabbccaabbccaabbcc`

Computing equality-free factorisations

Find equality-free factorisation with small width

$\texttt{aab} \cdot \texttt{bcc} \cdot \texttt{aa} \cdot \texttt{bbc} \cdot \texttt{caa} \cdot \texttt{bb} \cdot \texttt{cc}$

Computing equality-free factorisations

Find equality-free factorisation with small width

$\texttt{aab} \cdot \texttt{bcc} \cdot \texttt{aa} \cdot \texttt{bbc} \cdot \texttt{caa} \cdot \texttt{bb} \cdot \texttt{cc}$

Computing equality-free factorisations

Given a string w and $m \in \mathbb{N}$

- $\exists$ equality-free factorisation p of w with $\mathsf{s}(p) \geq m$? EF-s
- $\exists$ equality-free factorisation p of w with $\mathsf{w}(p) \leq m$? EF-w

Computing repetitive factorisations

What is a good measure of repetitiveness?

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

$\mathsf{c}(p) = \mathsf{s}(p) \Rightarrow$ equality-free $\Rightarrow$ not repetitive at all,
$\mathsf{c}(p) = 1 \Rightarrow p = u \cdot u \cdots u \Rightarrow$ very repetitive.

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

$\mathsf{c}(p) = \mathsf{s}(p) \Rightarrow$ equality-free $\Rightarrow$ not repetitive at all,
$\mathsf{c}(p) = 1 \Rightarrow p = u \cdot u \cdots u \Rightarrow$ very repetitive.

$\mathsf{c}(p) \leq 3, \mathsf{s}(p) \geq 5$

`aabcacaaabaab`

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

$\mathsf{c}(p) = \mathsf{s}(p) \Rightarrow$ equality-free $\Rightarrow$ not repetitive at all,
$\mathsf{c}(p) = 1 \Rightarrow p = u \cdot u \cdots u \Rightarrow$ very repetitive.

$\mathsf{c}(p) \leq 3, \mathsf{s}(p) \geq 5$

$$\mathtt{a} \cdot \mathtt{a} \cdot \mathtt{b} \cdot \mathtt{c} \cdot \mathtt{a} \cdot \mathtt{c} \cdot \mathtt{a} \cdot \mathtt{a} \cdot \mathtt{a} \cdot \mathtt{b} \cdot \mathtt{a} \cdot \mathtt{a} \cdot \mathtt{b}$$

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

$\mathsf{c}(p) = \mathsf{s}(p) \Rightarrow$ equality-free $\Rightarrow$ not repetitive at all,
$\mathsf{c}(p) = 1 \Rightarrow p = u \cdot u \cdots u \Rightarrow$ very repetitive.

$\mathsf{c}(p) \leq 2, \mathsf{s}(p) \geq 5$

$$\mathtt{a} \cdot \mathtt{a} \cdot \mathtt{b} \cdot \mathtt{c} \cdot \mathtt{a} \cdot \mathtt{c} \cdot \mathtt{a} \cdot \mathtt{a} \cdot \mathtt{a} \cdot \mathtt{b} \cdot \mathtt{a} \cdot \mathtt{a} \cdot \mathtt{b}$$

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

$\mathsf{c}(p) = \mathsf{s}(p) \Rightarrow$ equality-free $\Rightarrow$ not repetitive at all,
$\mathsf{c}(p) = 1 \Rightarrow p = u \cdot u \cdots u \Rightarrow$ very repetitive.

$\mathsf{c}(p) \leq 2, \mathsf{s}(p) \geq 5$

$$\mathtt{aab} \cdot \mathtt{ca} \cdot \mathtt{ca} \cdot \mathtt{aab} \cdot \mathtt{aab}$$

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

$\mathsf{c}(p) = \mathsf{s}(p) \Rightarrow$ equality-free $\Rightarrow$ not repetitive at all,
$\mathsf{c}(p) = 1 \Rightarrow p = u \cdot u \cdots u \Rightarrow$ very repetitive.

$\mathsf{c}(p) \leq 2, \mathsf{s}(p) > 5$?

$$\mathtt{aab} \cdot \mathtt{ca} \cdot \mathtt{ca} \cdot \mathtt{aab} \cdot \mathtt{aab}$$

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

$\mathsf{c}(p) = \mathsf{s}(p) \Rightarrow$ equality-free $\Rightarrow$ not repetitive at all,
$\mathsf{c}(p) = 1 \Rightarrow p = u \cdot u \cdots u \Rightarrow$ very repetitive.

$\mathsf{c}(p) \leq 2, \mathsf{w}(p) \leq 2$?

$$\texttt{aab} \cdot \texttt{ca} \cdot \texttt{ca} \cdot \texttt{aab} \cdot \texttt{aab}$$

Computing repetitive factorisations

What is a good measure of repetitiveness? The cardinality $\mathsf{c}(p)$!

$\mathsf{c}(p) = \mathsf{s}(p) \Rightarrow$ equality-free $\Rightarrow$ not repetitive at all,
$\mathsf{c}(p) = 1 \Rightarrow p = u \cdot u \cdots u \Rightarrow$ very repetitive.

$\mathsf{c}(p) \leq 2, \mathsf{w}(p) \leq 2$?

$$\texttt{aab} \cdot \texttt{ca} \cdot \texttt{ca} \cdot \texttt{aab} \cdot \texttt{aab}$$

Computing repetitive factorisations

Given a string w and $m, k \in \mathbb{N}$

- $\exists$ factorisation p of w with $\mathsf{c}(p) \leq k$, $\mathsf{s}(p) \geq m$? RF-s
- $\exists$ factorisation p of w with $\mathsf{c}(p) \leq k$, $\mathsf{w}(p) \leq m$? RF-w

Motivation: equality-free factorisations with small width

Collision-aware oligo design for gene synthesis

Goal: Construct long DNA strands.
Problem: Only very short pieces of DNA can be reliably constructed.
Solution: Find short pieces of DNA that will self-assemble.

Motivation: equality-free factorisations with small width

Collision-aware oligo design for gene synthesis

Goal: Construct long DNA strands.
Problem: Only very short pieces of DNA can be reliably constructed.
Solution: Find short pieces of DNA that will self-assemble.

$\Rightarrow$

Find a factorisation p of the DNA strand with

- $\mathsf{w}(p)$ is small,
- no factor is the complement of another,
- ...
- ...

Motivation: equality-free factorisations with large size

Pattern matching with variables

Given a string α with variables and a string w, can we uniformly replace the variables in α such that we obtain w?

If α is "simple enough", then this can be decided in poly-time.

Motivation: equality-free factorisations with large size

Pattern matching with variables

Given a string α with variables and a string w, can we uniformly replace the variables in α such that we obtain w?

If α is "simple enough", then this can be decided in poly-time.

Injective pattern matching with variables

Given a string α with variables and a string w, can we uniformly replace the variables in α such that we obtain w and different variables must be replaced by different strings?

Motivation: equality-free factorisations with large size

Pattern matching with variables

Given a string α with variables and a string w, can we uniformly replace the variables in α such that we obtain w?

If α is "simple enough", then this can be decided in poly-time.

Injective pattern matching with variables

Given a string α with variables and a string w, can we uniformly replace the variables in α such that we obtain w and different variables must be replaced by different strings?

For the "simple" patterns $x_1 x_2 \dots x_n$ this is equivalent to finding equality-free factorisations with size n.

Motivation: repetitive factorisations

Let p be a factorisation with $\mathsf{sf}(p) = \{u_1, u_2, \ldots, u_k\}$, i. e., $p = u_{j_1} \cdot u_{j_2} \cdot \ldots \cdot u_{j_n}$, $1 \leq j_i \leq k$, $1 \leq i \leq k$.

Motivation: repetitive factorisations

Let p be a factorisation with $\mathsf{sf}(p) = \{u_1, u_2, \ldots, u_k\}$, i. e., $p = u_{j_1} \cdot u_{j_2} \cdot \ldots \cdot u_{j_n}$, $1 \leq j_i \leq k$, $1 \leq i \leq k$.

The corresponding word can be represented by $j_1 j_2 \ldots j_n$ and $\mathsf{sf}(p)$

Complexity

Theorem (Condon, Maňuch, Thachuk, 2008)

Computing EF-w *is* NP*-complete (even if* $m \leq 2$ *or* $|\Sigma| \leq 2$*).*

Theorem (Fernau, Manea, Mercaş, S., 2015)

EF-s *is* NP*-complete.*

Complexity

Theorem (Condon, Maňuch, Thachuk, 2008)

Computing EF-w *is* NP*-complete (even if* $m \leq 2$ *or* $|\Sigma| \leq 2$*).*

Theorem (Fernau, Manea, Mercaş, S., 2015)

EF-s *is* NP*-complete.*

Contribution of this paper

Revisit the complexity of these problems (and RF-s, RF-w), also from the parameterised point of view.

Parameterised Complexity

Parameterised problem K:
instances are of the form (x, k), where k is the *parameter*

Parameterised Complexity

Parameterised problem K:
instances are of the form (x, k), where k is the *parameter*

K is *fixed-parameter tractable* (in FPT) $\iff$
K can be solved in $\mathcal{O}(f(k) \times p(|x|))$ (for recursive f and polynomial p).

Parameterised Complexity

Parameterised problem K:
instances are of the form (x, k), where k is the *parameter*

K is *fixed-parameter tractable* (in FPT) $\iff$
K can be solved in $\mathcal{O}(f(k) \times p(|x|))$ (for recursive f and polynomial p).

K is NP-hard even if $k \leq c$ for constant $c \Rightarrow K \notin \text{FPT}$ (unless P = NP).

Equality-Free Factor Cover (EFFC)

Equality-free factor cover

Given a string w and a set F of strings,
$\exists$ equality-free factorisation p of w with $\mathsf{sf}(p) \subseteq F$? EFFC

Equality-Free Factor Cover (EFFC)

Equality-free factor cover

Given a string w and a set F of strings,
$\exists$ equality-free factorisation p of w with $\mathsf{sf}(p) \subseteq F$? EFFC

Theorem

EFFC *is* NP*-complete (even for fixed* Σ *with* $|\Sigma| = 2$*).*

Equality-Free Factor Cover (EFFC)

Equality-free factor cover

Given a string w and a set F of strings,
$\exists$ equality-free factorisation p of w with $\mathsf{sf}(p) \subseteq F$? EFFC

Theorem

EFFC *is* NP*-complete (even for fixed* Σ *with* $|\Sigma| = 2$*).*

Proof Sketch

Let $w \in \Sigma^*$, $F = \{v \mid w = uvu', |v| \leq m\}$.

Equality-Free Factor Cover (EFFC)

Equality-free factor cover

Given a string w and a set F of strings,
$\exists$ equality-free factorisation p of w with $\mathsf{sf}(p) \subseteq F$? EFFC

Theorem

EFFC *is* NP*-complete (even for fixed* Σ *with* $|\Sigma| = 2$*).*

Proof Sketch

Let $w \in \Sigma^*$, $F = \{v \mid w = uvu', |v| \leq m\}$.
w has equality-free factorisation p with $\mathsf{w}(p) \leq m$ $\iff$
w has equality-free factorisation p' with $\mathsf{sf}(p') \subseteq F$.

Equality-Free Factor Cover (EFFC)

Theorem

EFFC *can be solved in time* $\mathcal{O}(|w|^{|F|+1})$.

Equality-Free Factor Cover (EFFC)

Theorem

EFFC *can be solved in time* $\mathcal{O}(|w|^{|F|+1})$.

Proof Sketch

Let $w \in \Sigma^*$ and let p be an equality-free factorisation for w with $\mathsf{sf}(p) \subseteq F$.

Equality-Free Factor Cover (EFFC)

Theorem

EFFC *can be solved in time* $\mathcal{O}(|w|^{|F|+1})$.

Proof Sketch

Let $w \in \Sigma^*$ and let p be an equality-free factorisation for w with $\mathsf{sf}(p) \subseteq F$.

- $\mathsf{s}(p) \leq |w|$
- $\mathsf{s}(p) \leq |F|$

Equality-Free Factor Cover (EFFC)

Theorem

EFFC *can be solved in time* $\mathcal{O}(|w|^{|F|+1})$.

Proof Sketch

Let $w \in \Sigma^*$ and let p be an equality-free factorisation for w with $\mathsf{sf}(p) \subseteq F$.

- $\mathsf{s}(p) \leq |w|$
- $\mathsf{s}(p) \leq |F|$

Enumerate all equality-free factorisations with $\mathsf{sf}(p) \subseteq F$ and $\mathsf{s}(p) \leq \min\{|w|, |F|\}$.

Equality-Free Factor Cover (EFFC)

Theorem

The Problem EFFC *can be solved in time* $\mathcal{O}(|w| \times (2^{|F|} - 1) \times |F|!)$.

Equality-Free Factor Cover (EFFC)

Theorem

The Problem EFFC *can be solved in time* $\mathcal{O}(|w| \times (2^{|F|} - 1) \times |F|!)$.

Proof Sketch

$w \in \Sigma^*$, $F = \{u_1, u_2, \ldots, u_\ell\}$
$\Gamma = \{1, 2, \ldots, \ell\}$, $h : \Gamma^* \to \Sigma^*$, $h(i) = u_i$, $i \in \Gamma$

Equality-Free Factor Cover (EFFC)

Theorem

The Problem EFFC *can be solved in time* $\mathcal{O}(|w| \times (2^{|F|} - 1) \times |F|!)$.

Proof Sketch

$w \in \Sigma^*$, $F = \{u_1, u_2, \ldots, u_\ell\}$
$\Gamma = \{1, 2, \ldots, \ell\}$, $h : \Gamma^* \to \Sigma^*$, $h(i) = u_i$, $i \in \Gamma$

w has equality-free factorisation p with $\mathsf{sf}(p) \subseteq F \iff$
$\exists v \in \Gamma^*$ with $|v|_i \leq 1, i \in \Gamma$, $h(v) = w$.

Equality-Free Factor Cover (EFFC)

Theorem

The Problem EFFC *can be solved in time* $\mathcal{O}(|w| \times (2^{|F|} - 1) \times |F|!)$.

Proof Sketch

$w \in \Sigma^*$, $F = \{u_1, u_2, \ldots, u_\ell\}$
$\Gamma = \{1, 2, \ldots, \ell\}$, $h : \Gamma^* \to \Sigma^*$, $h(i) = u_i$, $i \in \Gamma$

w has equality-free factorisation p with $\mathsf{sf}(p) \subseteq F \iff$
$\exists v \in \Gamma^*$ with $|v|_i \leq 1, i \in \Gamma$, $h(v) = w$.

There are at most $(2^{|F|} - 1) \times |F|!$ such words v.

Factor Cover (FC)

Theorem

FC *can be solved in time* $\mathcal{O}(|F| \times |w|^2)$.

Factor Cover (FC)

Theorem

FC *can be solved in time* $\mathcal{O}(|F| \times |w|^2)$.

Proof Sketch

Dynamic programming + KMP.

Factor Cover (FC)

Theorem

FC *can be solved in time* $\mathcal{O}(|F| \times |w|^2)$.

Proof Sketch

Dynamic programming + KMP.

Remark: We shall need this algorithm later for computing repetitive factorisations with large size or small width.

Max/Min Equality-Free Fact. Size/Width

Theorem (Condon, Maňuch, Thachuk, 2008)

EF-w *is* NP*-complete (even if* $m \leq 2$ *or* $|\Sigma| \leq 2$*).*

Max/Min Equality-Free Fact. Size/Width

Theorem (Condon, Maňuch, Thachuk, 2008)

EF-w *is* NP-*complete (even if* $m \leq 2$ *or* $|\Sigma| \leq 2$*).*

Theorem

EF-w *can be solved in time* $\mathcal{O}(m^{m^2 \times |\Sigma|^m + 2} \times |\Sigma|^m)$.

Max/Min Equality-Free Fact. Size/Width

Theorem (Condon, Maňuch, Thachuk, 2008)

$\mathsf{EF\text{-}w}$ *is* NP-*complete (even if* $m \leq 2$ *or* $|\Sigma| \leq 2$*).*

Theorem

$\mathsf{EF\text{-}w}$ *can be solved in time* $\mathcal{O}(m^{m^2 \times |\Sigma|^m + 2} \times |\Sigma|^m)$.

Proof Sketch

Let p be equality-free factorisation of w with $\mathsf{w}(p) \leq m$.

Max/Min Equality-Free Fact. Size/Width

Theorem (Condon, Maňuch, Thachuk, 2008)

EF-w *is* NP*-complete (even if* $m \leq 2$ *or* $|\Sigma| \leq 2$*).*

Theorem

EF-w *can be solved in time* $\mathcal{O}(m^{m^2 \times |\Sigma|^m + 2} \times |\Sigma|^m)$.

Proof Sketch

Let p be equality-free factorisation of w with $\mathsf{w}(p) \leq m$.

$\Rightarrow \mathsf{s}(p) \leq m \times |\Sigma|^m \Rightarrow |w| \leq m^2 \times |\Sigma|^m$.

Max/Min Equality-Free Fact. Size/Width

Theorem (Condon, Maňuch, Thachuk, 2008)

EF-w *is* NP-*complete (even if* $m \leq 2$ *or* $|\Sigma| \leq 2$*).*

Theorem

EF-w *can be solved in time* $\mathcal{O}(m^{m^2 \times |\Sigma|^m + 2} \times |\Sigma|^m)$.

Proof Sketch

Let p be equality-free factorisation of w with $\mathsf{w}(p) \leq m$.

$\Rightarrow \mathsf{s}(p) \leq m \times |\Sigma|^m \Rightarrow |w| \leq m^2 \times |\Sigma|^m$.

Check $|w| \leq m^2 \times |\Sigma|^m$, if yes, enumerate all factorisations with width of at most m.

Max/Min Equality-Free Fact. Size/Width

Dichotomy for EF-w w.r.t. parameters m and $|\Sigma|$:

- $m \leq c$ and $|\Sigma|$ unbounded: NP-complete if and only if $c \geq 2$.

Max/Min Equality-Free Fact. Size/Width

Dichotomy for EF-w w.r.t. parameters m and $|\Sigma|$:

- $m \leq c$ and $|\Sigma|$ unbounded: NP-complete if and only if $c \geq 2$.
- $|\Sigma| \leq c$ and m unbounded: NP-complete if and only if $c \geq 2$.

Max/Min Equality-Free Fact. Size/Width

Dichotomy for EF-w w.r.t. parameters m and $|\Sigma|$:

- $m \leq c$ and $|\Sigma|$ unbounded: NP-complete if and only if $c \geq 2$.
- $|\Sigma| \leq c$ and m unbounded: NP-complete if and only if $c \geq 2$.
- $|\Sigma| \leq c$ and $m \leq c'$: poly-time.

Max/Min Equality-Free Fact. Size/Width

Dichotomy for EF-w w.r.t. parameters m and $|\Sigma|$:

- $m \leq c$ and $|\Sigma|$ unbounded: NP-complete if and only if $c \geq 2$.
- $|\Sigma| \leq c$ and m unbounded: NP-complete if and only if $c \geq 2$.
- $|\Sigma| \leq c$ and $m \leq c'$: poly-time.

What about equality-free factorisations with large size (EF-s)??

Max/Min Equality-Free Fact. Size/Width

Open Problem

Is EF-s NP-complete for fixed alphabets?

Reminder: In the real world, there are only fixed alphabets!

Max/Min Equality-Free Fact. Size/Width

Open Problem

Is EF-s NP-complete for fixed alphabets?

Reminder: In the real world, there are only fixed alphabets!

At least, poly-time (fpt) if m is bounded:

Theorem

EF-s *can be solved in time* $\mathcal{O}((\frac{m^2+m}{2} - 1)^m)$.

Max/Min Equality-Free Fact. Size/Width

Open Problem

Is EF-s NP-complete for fixed alphabets?

Reminder: In the real world, there are only fixed alphabets!

At least, poly-time (fpt) if m is bounded:

Theorem

EF-s *can be solved in time* $\mathcal{O}((\frac{m^2+m}{2} - 1)^m)$.

Proof Sketch

$|w| \geq \Sigma_{i=1}^{m} i = \frac{m^2+m}{2} \Rightarrow$ split w into factors of different lengths.

Max/Min Equality-Free Fact. Size/Width

Open Problem

Is EF-s NP-complete for fixed alphabets?

Reminder: In the real world, there are only fixed alphabets!

At least, poly-time (fpt) if m is bounded:

Theorem

EF-s *can be solved in time* $\mathcal{O}((\frac{m^2+m}{2} - 1)^m)$.

Proof Sketch

$|w| \geq \Sigma_{i=1}^{m} i = \frac{m^2+m}{2} \Rightarrow$ split w into factors of different lengths.
$|w| \leq \frac{m^2+m}{2} - 1 \Rightarrow$ enumerate all factorisations of size m.

Max/Min Repetitive Factorisation Size/Width

We have three parameters:

- $|\Sigma|$,
- m (size/width bound),
- k (bound on the cardinality).

Max/Min Repetitive Factorisation Size/Width

We have three parameters:

- $|\Sigma|$,
- m (size/width bound),
- k (bound on the cardinality).

Open Problem

Is RF-s NP-complete?

Max/Min Repetitive Factorisation Size/Width

We have three parameters:

- $|\Sigma|$,
- m (size/width bound),
- k (bound on the cardinality).

Open Problem

Is RF-s NP-complete?

However, if $|\Sigma|$, m or k is a constant, then we can solve it in poly-time.

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-s *can be solved in time*

- $\mathcal{O}(k^2 \times |w|^{2k+3})$,
- $\mathcal{O}(|\Sigma|^2 \times |w|^{2|\Sigma|+1})$,
- $\mathcal{O}(m^2 \times |w|^{2m+1})$.

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-s *can be solved in time*

- $\mathcal{O}(k^2 \times |w|^{2k+3})$,
- $\mathcal{O}(|\Sigma|^2 \times |w|^{2|\Sigma|+1})$,
- $\mathcal{O}(m^2 \times |w|^{2m+1})$.

Proof Sketch

Let $F_w = \{u \mid u \text{ is a factor of } w\}$.

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-s *can be solved in time*

- $\mathcal{O}(k^2 \times |w|^{2k+3})$,
- $\mathcal{O}(|\Sigma|^2 \times |w|^{2|\Sigma|+1})$,
- $\mathcal{O}(m^2 \times |w|^{2m+1})$.

Proof Sketch

Let $F_w = \{u \mid u \text{ is a factor of } w\}$.
Problem FC: Does w have a factorisation p with $\mathsf{sf}(p) \subseteq F$ for given F?
Solve FC on every $F \subseteq F_w$ with $|F| \leq k$.

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-s *can be solved in time*

- $\mathcal{O}(k^2 \times |w|^{2k+3})$,
- $\mathcal{O}(|\Sigma|^2 \times |w|^{2|\Sigma|+1})$,
- $\mathcal{O}(m^2 \times |w|^{2m+1})$.

Proof Sketch

Let $F_w = \{u \mid u \text{ is a factor of } w\}$.
Problem FC: Does w have a factorisation p with $\mathsf{sf}(p) \subseteq F$ for given F?
Solve FC on every $F \subseteq F_w$ with $|F| \leq k$.

$k \geq |\Sigma| \Rightarrow$ split w into factors of size 1.
$k \geq m \Rightarrow$ any factorisation of size m is fine.

Max/Min Repetitive Factorisation Size/Width

We know more about computing repetitive factorisations with small width (RF-w)!

Max/Min Repetitive Factorisation Size/Width

We know more about computing repetitive factorisations with small width (RF-w)!

If $|\Sigma|$ or k is a constant, then we can solve it in poly-time.

Theorem

RF-w *can be solved in time*

- $\mathcal{O}(k^2 \times m^k \times |w|^{k+3})$,
- $\mathcal{O}(|\Sigma|^2 \times m^{(|\Sigma|-1)} \times |w|^{|\Sigma|+2})$.

Max/Min Repetitive Factorisation Size/Width

We know more about computing repetitive factorisations with small width (RF-w)!

If $|\Sigma|$ or k is a constant, then we can solve it in poly-time.

Theorem

RF-w *can be solved in time*

- $\mathcal{O}(k^2 \times m^k \times |w|^{k+3})$,
- $\mathcal{O}(|\Sigma|^2 \times m^{(|\Sigma|-1)} \times |w|^{|\Sigma|+2})$.

Proof Sketch

Analogous to the proofs for RF-s.

Max/Min Repetitive Factorisation Size/Width

We know more about computing repetitive factorisations with small width (RF-w)!

If $|\Sigma|$ or k is a constant, then we can solve it in poly-time.

Theorem

RF-w *can be solved in time*

- $\mathcal{O}(k^2 \times m^k \times |w|^{k+3})$,
- $\mathcal{O}(|\Sigma|^2 \times m^{(|\Sigma|-1)} \times |w|^{|\Sigma|+2})$.

Proof Sketch

Analogous to the proofs for RF-s.

However, k cannot be bounded by m (the width bound), only by $\lceil \frac{|w|}{m} \rceil$.

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-w *is* NP*-complete even if* $m \leq 2$.

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-w *is* NP*-complete even if* $m \leq 2$.

Hitting Set (HS)
Instance: $U = \{x_1, \ldots, x_\ell\}$, $S_1, \ldots, S_n \subseteq U$ and $q \in \mathbb{N}$.
Question: $\exists\, T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n$?

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-w *is* NP-*complete even if* $m \leq 2$.

Hitting Set (HS)
Instance: $U = \{x_1, \ldots, x_\ell\}$, $S_1, \ldots, S_n \subseteq U$ and $q \in \mathbb{N}$.
Question: $\exists\, T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n$?

HS instance: $(U, S_1, \ldots, S_n, q)$ with $S_i = \{y_{i,1}, y_{i,2}, \ldots, y_{i,r}\}$, $1 \leq i \leq n$.

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-w *is* NP*-complete even if* $m \leq 2$.

Hitting Set (HS)
Instance: $U = \{x_1, \ldots, x_\ell\}$, $S_1, \ldots, S_n \subseteq U$ and $q \in \mathbb{N}$.
Question: $\exists\, T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n$?

HS instance: $(U, S_1, \ldots, S_n, q)$ with $S_i = \{y_{i,1}, y_{i,2}, \ldots, y_{i,r}\}$, $1 \leq i \leq n$.

RF-w instance:
$\Sigma = U \cup \{\$_{i,j} \mid 1 \leq i \leq n, 1 \leq j \leq r-1\} \cup \{¢\}$,

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-w *is* NP*-complete even if* $m \leq 2$.

Hitting Set (HS)
Instance: $U = \{x_1, \dots, x_\ell\}$, $S_1, \dots, S_n \subseteq U$ and $q \in \mathbb{N}$.
Question: $\exists\, T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n$?

HS instance: $(U, S_1, \dots, S_n, q)$ with $S_i = \{y_{i,1}, y_{i,2}, \dots, y_{i,r}\}$, $1 \leq i \leq n$.

RF-w instance:
$\Sigma = U \cup \{\$_{i,j} \mid 1 \leq i \leq n, 1 \leq j \leq r-1\} \cup \{¢\}$,
$w = ¢¢\, v_1\, ¢\, v_2\, ¢ \dots ¢\, v_n\, ¢$,

Max/Min Repetitive Factorisation Size/Width

Theorem

RF-w *is* NP-*complete even if* $m \leq 2$.

Hitting Set (HS)
Instance: $U = \{x_1, \ldots, x_\ell\}$, $S_1, \ldots, S_n \subseteq U$ and $q \in \mathbb{N}$.
Question: $\exists\ T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n$?

HS instance: $(U, S_1, \ldots, S_n, q)$ with $S_i = \{y_{i,1}, y_{i,2}, \ldots, y_{i,r}\}$, $1 \leq i \leq n$.

RF-w instance:
$\Sigma = U \cup \{\$_{i,j} \mid 1 \leq i \leq n, 1 \leq j \leq r-1\} \cup \{¢\}$,
$w = ¢¢\, v_1\, ¢\, v_2\, ¢ \ldots ¢\, v_n\, ¢$,
$v_i = y_{i,1}\$_{i,1}y_{i,2}\$_{i,2} \ldots \$_{i,r-1}y_{i,r}$, $1 \leq i \leq n$.

Max/Min Repetitive Factorisation Size/Width

$T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n \Rightarrow$

Max/Min Repetitive Factorisation Size/Width

$T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n \Rightarrow$

$w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

Max/Min Repetitive Factorisation Size/Width

$T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n \Rightarrow$

$w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,
$v_i = y_{i,1}\$_{i,1} \cdot y_{i,2}\$_{i,2} \cdot \ldots \cdot y_{i,j_i-1}\$_{i,j_i-1} \cdot y_{i,j_i} \cdot \$_{i,j_i} y_{i,j_i+1} \cdot \ldots \cdot \$_{i,r-1} y_{i,r}$,
i, $1 \leq i \leq n$, $y_{i,j_i} \in T$,

Max/Min Repetitive Factorisation Size/Width

$T \subseteq U$ with $|T| \leq q$ and $T \cap S_i \neq \emptyset$, $1 \leq i \leq n \Rightarrow$

$w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,
$v_i = y_{i,1}\$_{i,1} \cdot y_{i,2}\$_{i,2} \cdot \ldots \cdot y_{i,j_i-1}\$_{i,j_i-1} \cdot y_{i,j_i} \cdot \$_{i,j_i} y_{i,j_i+1} \cdot \ldots \cdot \$_{i,r-1} y_{i,r}$,
i, $1 \leq i \leq n$, $y_{i,j_i} \in T$,

has width 2 and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x¢$ and $¢x$, $x \in U \cup \{¢\}$, into $x \cdot ¢$ and $¢ \cdot x$, respectively.
$\Rightarrow w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x¢$ and $¢x$, $x \in U \cup \{¢\}$, into $x \cdot ¢$ and $¢ \cdot x$, respectively.
$\Rightarrow w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

$\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$ is maintained.

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x¢$ and $¢x$, $x \in U \cup \{¢\}$, into $x \cdot ¢$ and $¢ \cdot x$, respectively.
$\Rightarrow w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

$\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$ is maintained.

$v_i = y_{i,1}\$_{i,1}y_{i,2}\$_{i,2}y_{i,3}\$_{i,3} \ldots \$_{i,r-1}y_{i,r}$ has odd length $\Rightarrow$

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x¢$ and $¢x$, $x \in U \cup \{¢\}$, into $x \cdot ¢$ and $¢ \cdot x$, respectively.
$\Rightarrow w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

$\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$ is maintained.

$v_i = y_{i,1}\$_{i,1}y_{i,2}\$_{i,2}y_{i,3}\$_{i,3} \ldots \$_{i,r-1}y_{i,r}$ has odd length $\Rightarrow$
$v_i = \ldots \cdot \$_{i,j-1} \cdot y_{i,j} \cdot \$_{i,j} \ldots$ (call $y_{i,j}$ *isolated*)

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x¢$ and $¢x$, $x \in U \cup \{¢\}$, into $x \cdot ¢$ and $¢ \cdot x$, respectively.
$\Rightarrow w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

$\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$ is maintained.

$v_i = y_{i,1}\$_{i,1}y_{i,2}\$_{i,2}y_{i,3}\$_{i,3}\ldots\$_{i,r-1}y_{i,r}$ has odd length $\Rightarrow$
$v_i = \ldots \cdot \$_{i,j-1} \cdot y_{i,j} \cdot \$_{i,j} \ldots$ (call $y_{i,j}$ *isolated*)

Let T be the set of all isolated elements.

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x¢$ and $¢x$, $x \in U \cup \{¢\}$, into $x \cdot ¢$ and $¢ \cdot x$, respectively.
$\Rightarrow w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

$\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$ is maintained.

$v_i = y_{i,1}\$_{i,1}y_{i,2}\$_{i,2}y_{i,3}\$_{i,3}\ldots\$_{i,r-1}y_{i,r}$ has odd length $\Rightarrow$
$v_i = \ldots \cdot \$_{i,j-1} \cdot y_{i,j} \cdot \$_{i,j} \ldots$ (call $y_{i,j}$ *isolated*)

Let T be the set of all isolated elements.

$T \cap S_i \neq \emptyset$, $1 \leq i \leq n$.

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x¢$ and $¢x$, $x \in U \cup \{¢\}$, into $x \cdot ¢$ and $¢ \cdot x$, respectively.
$\Rightarrow w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

$\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$ is maintained.

$v_i = y_{i,1}\$_{i,1}y_{i,2}\$_{i,2}y_{i,3}\$_{i,3} \ldots \$_{i,r-1}y_{i,r}$ has odd length $\Rightarrow$
$v_i = \ldots \cdot \$_{i,j-1} \cdot y_{i,j} \cdot \$_{i,j} \ldots$ (call $y_{i,j}$ *isolated*)

Let T be the set of all isolated elements.

$T \cap S_i \neq \emptyset$, $1 \leq i \leq n$.

$\mathsf{sf}(p) = \{¢\} \cup \{\text{all } n(r-1) \text{ factors with } \$_{i,j}\} \cup T$.

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x¢$ and $¢x$, $x \in U \cup \{¢\}$, into $x \cdot ¢$ and $¢ \cdot x$, respectively.
$\Rightarrow w = ¢ \cdot ¢ \cdot v_1 \cdot ¢ \cdot v_2 \cdot ¢ \cdot \ldots \cdot ¢ \cdot v_n \cdot ¢$,

$\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$ is maintained.

$v_i = y_{i,1}\$_{i,1}y_{i,2}\$_{i,2}y_{i,3}\$_{i,3} \ldots \$_{i,r-1}y_{i,r}$ has odd length $\Rightarrow$
$v_i = \ldots \cdot \$_{i,j-1} \cdot y_{i,j} \cdot \$_{i,j} \ldots$ (call $y_{i,j}$ *isolated*)

Let T be the set of all isolated elements.

$T \cap S_i \neq \emptyset$, $1 \leq i \leq n$.

$\mathsf{sf}(p) = \{¢\} \cup \{\text{all } n(r-1) \text{ factors with } \$_{i,j}\} \cup T$.
$\mathsf{c}(p) = |\,\mathsf{sf}(p)| = 1 + n(r-1) + |T| \leq 1 + n(r-1) + q \Rightarrow$

Max/Min Repetitive Factorisation Size/Width

Let p be a factorisation of w with $\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$.

Split every $x\cent$ and $\cent x$, $x \in U \cup \{\cent\}$, into $x \cdot \cent$ and $\cent \cdot x$, respectively.
$\Rightarrow w = \cent \cdot \cent \cdot v_1 \cdot \cent \cdot v_2 \cdot \cent \cdot \ldots \cdot \cent \cdot v_n \cdot \cent$,

$\mathsf{w}(p) \leq 2$ and $\mathsf{c}(p) \leq 1 + q + n(r-1)$ is maintained.

$v_i = y_{i,1}\$_{i,1}y_{i,2}\$_{i,2}y_{i,3}\$_{i,3} \ldots \$_{i,r-1}y_{i,r}$ has odd length $\Rightarrow$
$v_i = \ldots \cdot \$_{i,j-1} \cdot y_{i,j} \cdot \$_{i,j} \ldots$ (call $y_{i,j}$ *isolated*)

Let T be the set of all isolated elements.

$T \cap S_i \neq \emptyset$, $1 \leq i \leq n$.

$\mathsf{sf}(p) = \{\cent\} \cup \{\text{all } n(r-1) \text{ factors with } \$_{i,j}\} \cup T$.
$\mathsf{c}(p) = |\,\mathsf{sf}(p)| = 1 + n(r-1) + |T| \leq 1 + n(r-1) + q \Rightarrow$
$|T| \leq q$.

Thank you very much for your attention.