Reconstructing words from right-bounded-block words





December 7, 2020 Marie Lejeune (FNRS grantee) joint work with P. Fleischmann, F. Manea, D. Nowotka, M. Rigo

The classical reconstruction problem

Let us consider finite words $u = u_1 \cdots u_n \in \mathcal{A}^*$.

A subword of u is a subsequence of the sequence of letters $(u_i)_{i=1}^n$, non necessarily contiguous.

aca is a subword of abcbaba

The classical reconstruction problem

Let us consider finite words $u = u_1 \cdots u_n \in \mathcal{A}^*$.

A subword of u is a subsequence of the sequence of letters $(u_i)_{i=1}^n$, non necessarily contiguous.

aca is a subword of abcbaba

The *binomial coefficient* $\binom{u}{v}$ denotes the number of times that v occurs as a subword in u.

We have

$$\binom{abcbaba}{aca} = 2.$$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$ Can you uniquely determine u?

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

Can you uniquely determine u?

u = a a a

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = bbb \stackrel{|}{a} a a \qquad \begin{pmatrix} u \\ ab \end{pmatrix} = 0$$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = abbb a a \begin{pmatrix} u \\ ab \end{pmatrix} \leq 3$$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = a \quad abbb \quad a \qquad \begin{pmatrix} u \\ ab \end{pmatrix} \leq 6$$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = a \quad a \quad a \quad b \qquad \begin{pmatrix} u \\ ab \end{pmatrix} = 3$$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = a a a b$$
 $\binom{u}{ab} = 3$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = b a a a b$$
 $\binom{u}{ab} \leq 6$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = ab \ a \ a b$$
 $\begin{pmatrix} u \\ ab \end{pmatrix} = 4$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = ab \ a \ a bb \qquad \begin{pmatrix} u \\ ab \end{pmatrix} = 7$$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = a ab a b$$
 $\begin{pmatrix} u \\ ab \end{pmatrix} = 5$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

$$u = a \ abb \ ab$$
 $\begin{pmatrix} u \\ ab \end{pmatrix} = 7$

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

Can you uniquely determine *u*?

$$u = a \ abb \ a b \qquad \begin{pmatrix} u \\ ab \end{pmatrix} = 7$$

Answer: NO. $u_1 = abaabb$ and $u_2 = aabbab$ are two words satisfying the conditions.

Let $\mathcal{A} = \{a, b\}$. Let $u \in \mathcal{A}^6$ be an unknown word of length 6 such that • $\binom{u}{a} = 3$ • $\binom{u}{ab} = 7$

Can you uniquely determine *u*?

$$u = a \ abb \ ab \ \begin{pmatrix} u \\ ab \end{pmatrix} = 7$$

Answer: NO. $u_1 = abaabb$ and $u_2 = aabbab$ are two words satisfying the conditions.

Add the following condition: $\binom{u}{aab} = 5$. Can you uniquely determine *u*? **YES**, u = aabbab.

Let *u* be a finite word and $k \in \mathbb{N}$. The *k*-deck of *u* is the multiset of subwords of *u* of length *k*. It is always of cardinality $\binom{|u|}{k}$.

Let *u* be a finite word and $k \in \mathbb{N}$. The *k*-deck of *u* is the multiset of subwords of *u* of length *k*. It is always of cardinality $\binom{|u|}{k}$.

The 3-deck of *aabbab* is

 $\{aaa, aab_5, aba_4, abb_6, bab_2, bba, bbb\}.$

Let *u* be a finite word and $k \in \mathbb{N}$. The *k*-deck of *u* is the multiset of subwords of *u* of length *k*. It is always of cardinality $\binom{|u|}{k}$.

The 3-deck of *aabbab* is

 $\{aaa, aab_5, aba_4, abb_6, bab_2, bba, bbb\}.$

Reconstruction problem

Let \mathcal{A} be an alphabet, and n an integer. What is the minimal k such that any word from \mathcal{A}^n can be uniquely determined from its k-deck?

Let \mathcal{A} be an alphabet, u and v two words. We denote by Q(u, v) the following question:

What is the value of the binomial coefficient $\binom{u}{v}$?

Let A be an alphabet, u and v two words. We denote by Q(u, v) the following question:

What is the value of the binomial coefficient $\binom{u}{v}$?

Our variant Let \mathcal{A} be an alphabet and $n \in \mathbb{N}$. What is the minimal number k such that any word u from \mathcal{A}^n can be uniquely determined by asking k questions $Q(u, v_1), \ldots, Q(u, v_k)$, sequentially? Let A be an alphabet, u and v two words. We denote by Q(u, v) the following question:

What is the value of the binomial coefficient $\binom{u}{v}$?

Our variant Let \mathcal{A} be an alphabet and $n \in \mathbb{N}$. What is the minimal number k such that any word u from \mathcal{A}^n can be uniquely determined by asking k questions $Q(u, v_1), \ldots, Q(u, v_k)$, sequentially?

By sequentially, we mean that, for all *i*, the answers to $Q(u, v_1), \ldots, Q(u, v_i)$ can influe the choice of v_{i+1} .

Reconstructing words from right-bounded-block words

Classical reconstruction problem: survey of the results

2) Binary case: the results



Extending to an arbitrary finite alphabet

• Square matrices from some of their minors [Manvel, Stockmeyer, 1971]

- Square matrices from some of their minors [Manvel, Stockmeyer, 1971]
- Graphs from some of their subgraphs? [Kelly and Ulam's conjecture, 1957; Harary's conjecture, 1963]

- Square matrices from some of their minors [Manvel, Stockmeyer, 1971]
- Graphs from some of their subgraphs? [Kelly and Ulam's conjecture, 1957; Harary's conjecture, 1963]
- Words from some of their subwords? [Kalashnik, 1973]

Reconstruction problem consists in determining the function f.

Reconstruction problem consists in determining the function f.

• Schützenberger, Simon (1975) showed that $f(n) \leq \lfloor n/2 \rfloor + 1$.

Reconstruction problem consists in determining the function f.

- Schützenberger, Simon (1975) showed that $f(n) \leq \lfloor n/2 \rfloor + 1$.
- Krasikov, Roditty (1997) improved it: $f(n) \leq \lfloor 16\sqrt{n}/7 \rfloor + 5$.

Reconstruction problem consists in determining the function f.

- Schützenberger, Simon (1975) showed that $f(n) \leq \lfloor n/2 \rfloor + 1$.
- Krasikov, Roditty (1997) improved it: $f(n) \leq \lfloor 16\sqrt{n}/7 \rfloor + 5$.
- Dudik, Schulmann (2003) gave a lower bound: $f(n) \ge 3^{(\sqrt{2/3}-o(1))\log_3^{1/2}n}.$

Reconstruction problem consists in determining the function f.

- Schützenberger, Simon (1975) showed that $f(n) \leq \lfloor n/2 \rfloor + 1$.
- Krasikov, Roditty (1997) improved it: $f(n) \leq \lfloor 16\sqrt{n}/7 \rfloor + 5$.
- Dudik, Schulmann (2003) gave a lower bound: $f(n) \ge 3^{(\sqrt{2/3} - o(1)) \log_3^{1/2} n}.$

...

Let $u \in \mathcal{A}^*$ and $k \in \mathbb{N}$ such that the k-deck of u is known. Therefore, $\binom{u}{v}$ is known for every $v \in \mathcal{A}^k$. Let $u \in \mathcal{A}^*$ and $k \in \mathbb{N}$ such that the k-deck of u is known. Therefore, $\binom{u}{v}$ is known for every $v \in \mathcal{A}^k$. But in fact, $\binom{u}{v}$ is known for every $v \in \mathcal{A}^{\leq k}$. Let $u \in \mathcal{A}^*$ and $k \in \mathbb{N}$ such that the k-deck of u is known. Therefore, $\binom{u}{v}$ is known for every $v \in \mathcal{A}^k$. But in fact, $\binom{u}{v}$ is known for every $v \in \mathcal{A}^{\leq k}$.

Proposition: The 1-deck of u is known from its k-deck.

Proof: We obviously have $\binom{u}{a^k} = \binom{\binom{u}{a}}{k}$, for any $a \in \mathcal{A}$.
Properties of the *k*-deck

Proposition: The (k-1)-deck of u is known from its k-deck.

Proof: Let $x \in \mathcal{A}^{k-1}$. For any $a \in \mathcal{A}$, we have

$$\binom{u}{x}\binom{u}{a} = \sum_{j=0}^{k} \binom{u}{x_1 \cdots x_{j-1} a x_j \cdots x_{k-1}} + \binom{u}{x}\binom{x}{a},$$

Properties of the *k*-deck

Proposition: The (k - 1)-deck of u is known from its k-deck.

Proof: Let $x \in \mathcal{A}^{k-1}$. For any $a \in \mathcal{A}$, we have

$$\binom{u}{x}\binom{u}{a} = \sum_{j=0}^{k} \binom{u}{x_1 \cdots x_{j-1} a x_j \cdots x_{k-1}} + \binom{u}{x}\binom{x}{a},$$

 $\binom{u}{abac}\binom{u}{a} = \binom{u}{aabac} + \binom{u}{aabac} + \binom{u}{abaac} + \binom{u}{abaac} + \binom{u}{abaac} + \binom{u}{abaca} + \binom{u}{abac} + \binom{u}{abaac} + \binom{u}{abaa$

Properties of the *k*-deck

Proposition: The (k - 1)-deck of u is known from its k-deck.

Proof: Let $x \in \mathcal{A}^{k-1}$. For any $a \in \mathcal{A}$, we have

$$\binom{u}{x}\binom{u}{a} = \sum_{j=0}^{k} \binom{u}{x_1 \cdots x_{j-1} a x_j \cdots x_{k-1}} + \binom{u}{x}\binom{x}{a},$$

 $\binom{u}{abac}\binom{u}{a} = \binom{u}{aabac} + \binom{u}{aabac} + \binom{u}{abaac} + \binom{u}{abaac} + \binom{u}{abaac} + \binom{u}{abaca} + \binom{u}{abac} + \binom{u}{abaac} + \binom{u}{abaa$

and thus

$$\binom{u}{x} = \frac{1}{\binom{u}{a} - \binom{x}{a}} \left[\sum_{j=0}^{k} \binom{u}{x_1 \cdots x_{j-1} a x_j \cdots x_{k-1}} \right]$$

Knowing the entire k-deck of u requires to ask $(\# A)^k$ questions Q(u, v) with $v \in A^k$.

- Knowing the entire k-deck of u requires to ask $(\# A)^k$ questions Q(u, v) with $v \in A^k$.
- Some binomial coefficients can be deducted from other ones. In fact, it suffices to have the answer to questions Q(u, v) with $v \in \mathcal{A}^{\leq k}$ a Lyndon word.

- Knowing the entire k-deck of u requires to ask $(\# A)^k$ questions Q(u, v) with $v \in A^k$.
- Some binomial coefficients can be deducted from other ones. In fact, it suffices to have the answer to questions Q(u, v) with $v \in \mathcal{A}^{\leq k}$ a Lyndon word.
- Recall: Let < be a total order on \mathcal{A} . A word u is Lyndon if for any factorization $u = x \cdot y \in \mathcal{A}^+ \times \mathcal{A}^+$, we have $xy <_{lex} yx$.

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}, y \in \mathcal{A}^{n_y}$$
, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

$$\{w = w_1 \cdots w_n : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\} \text{ a partition} \\ \text{of } [n] \text{ s.t. } w_{i_1} \cdots w_{i_{n_x}} = x \text{ and } w_{j_1} \cdots w_{j_{n_y}} = y\}.$$

$$ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}.$$

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Let
$$x \in \mathcal{A}^{n_x}$$
, $y \in \mathcal{A}^{n_y}$, $n = n_x + n_y$ and $[n] = \{1, \ldots, n\}$.

The *infiltration* $x \downarrow y$ is the multiset

$$\{w = w_1 \cdots w_{n'} : \exists I_x = \{i_1 < \ldots < i_{n_x}\}, I_y = \{j_1 < \ldots < j_{n_y}\}, n' \le n$$

s.t. $I_x \cup I_y = [n']$ and
 $w_{i_1} \cdots w_{i_{n_x}} = x, w_{j_1} \cdots w_{j_{n_y}} = y$, if well defined}.

Note that $x \sqcup y \subseteq x \downarrow y$.

Example:

Proposition (Reutenauer): If v is non-Lyndon, $\exists x, y \in A^+$ such that v = xy and every word in the multiset $x \sqcup y$ is $<_{lex}$ less than or equal to v.

Proposition (Reutenauer): If v is non-Lyndon, $\exists x, y \in A^+$ such that v = xy and every word in the multiset $x \sqcup y$ is $<_{lex}$ less than or equal to v.

For any multiset S, denote by S_v the multiplicity of v in S.

Let $u \in \mathcal{A}^*$, v be non-Lyndon and x, y as in the previous proposition. We have

$$\binom{u}{v} = \frac{1}{(x \sqcup y)_{v}} \left[\binom{u}{x} \binom{u}{y} - \sum_{w \in \mathcal{A}^{+} \setminus \{v\}} (x \downarrow y)_{w} \binom{u}{w} \right]$$

Note that in the previous formula, $x, y, w \prec v$. $x \prec v \Leftrightarrow |x| < |v|$ or |x| = |v| and $x <_{lex} v$

Proposition (Reutenauer): If v is non-Lyndon, $\exists x, y \in A^+$ such that v = xy and every word in the multiset $x \sqcup y$ is $<_{lex}$ less than or equal to v.

For any multiset S, denote by S_v the multiplicity of v in S.

Let $u \in \mathcal{A}^*$, v be non-Lyndon and x, y as in the previous proposition. We have

$$\binom{u}{v} = \frac{1}{(x \sqcup y)_{v}} \left[\binom{u}{x} \binom{u}{y} - \sum_{w \in \mathcal{A}^{+} \setminus \{v\}} (x \downarrow y)_{w} \binom{u}{w} \right]$$

Note that in the previous formula, $x, y, w \prec v$. $x \prec v \Leftrightarrow |x| < |v| \text{ or } |x| = |v| \text{ and } x <_{lex} v$

We can apply the formula recursively on all non-Lyndon words.

Why Lyndon words are enough: an example

Let $u \in \mathcal{A}^*$ and v = abaab. Words x = ab and y = aab are such that xy = v and $w \in x \sqcup y \Rightarrow w \preceq v$.

 $ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}$ $ab \downarrow aab = \{abaab, aabab_3, aaabb_6, aabb_4, aaab_3, abab, aab\}$
Why Lyndon words are enough: an example

Let $u \in \mathcal{A}^*$ and v = abaab. Words x = ab and y = aab are such that xy = v and $w \in x \sqcup y \Rightarrow w \preceq v$.

 $ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}$ $ab \downarrow aab = \{abaab, aabab_3, aaabb_6, aabb_4, aaab_3, abab, aab\}$

$$\binom{u}{v} = \frac{1}{(x \sqcup y)_v} \left[\binom{u}{x} \binom{u}{y} - \sum_{w \in \mathcal{A}^+ \setminus \{v\}} (x \downarrow y)_w \binom{u}{w} \right]$$

$$\begin{pmatrix} u \\ abaab \end{pmatrix} = \frac{1}{1} \left[\begin{pmatrix} u \\ ab \end{pmatrix} \begin{pmatrix} u \\ aab \end{pmatrix} - \left[3 \begin{pmatrix} u \\ aabab \end{pmatrix} + 6 \begin{pmatrix} u \\ aaabb \end{pmatrix} + 4 \begin{pmatrix} u \\ aabb \end{pmatrix} + 3 \begin{pmatrix} u \\ aaab \end{pmatrix} + \begin{pmatrix} u \\ abab \end{pmatrix} + \begin{pmatrix} u \\ aab \end{pmatrix} \right] \right]$$

Why Lyndon words are enough: an example

Let $u \in \mathcal{A}^*$ and v = abaab. Words x = ab and y = aab are such that xy = v and $w \in x \sqcup y \Rightarrow w \preceq v$.

 $ab \sqcup aab = \{abaab, aabab_3, aaabb_6\}$ $ab \downarrow aab = \{abaab, aabab_3, aaabb_6, aabb_4, aaab_3, abab, aab\}$

$$\binom{u}{v} = \frac{1}{(x \sqcup y)_v} \left[\binom{u}{x} \binom{u}{y} - \sum_{w \in \mathcal{A}^+ \setminus \{v\}} (x \downarrow y)_w \binom{u}{w} \right]$$

$$\begin{pmatrix} u \\ abaab \end{pmatrix} = \frac{1}{1} \left[\begin{pmatrix} u \\ ab \end{pmatrix} \begin{pmatrix} u \\ aab \end{pmatrix} - \left[3 \begin{pmatrix} u \\ aabab \end{pmatrix} + 6 \begin{pmatrix} u \\ aaabb \end{pmatrix} + 4 \begin{pmatrix} u \\ aabb \end{pmatrix} + 3 \begin{pmatrix} u \\ aaab \end{pmatrix} + \begin{pmatrix} u \\ abab \end{pmatrix} + \begin{pmatrix} u \\ aab \end{pmatrix} \right] \right]$$

• Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^k \Rightarrow$ knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$. There are $(\# \mathcal{A})^k$ such words.

- Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^k \Rightarrow$ knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$. There are $(\# \mathcal{A})^k$ such words.
- Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$ and v Lyndon \Rightarrow knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$. There are $\sum_{i=1}^{k} \frac{1}{i} \sum_{d|i} \mu(d) (\# \mathcal{A})^{\frac{i}{d}}$ such words.

- Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^k \Rightarrow$ knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$. There are $(\# \mathcal{A})^k$ such words.
- Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$ and v Lyndon \Rightarrow knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$. There are $\sum_{i=1}^{k} \frac{1}{i} \sum_{d|i} \mu(d) (\# \mathcal{A})^{\frac{i}{d}}$ such words.
- Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^k$ and v Lyndon $\not\Rightarrow$ knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$.

- Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^k \Rightarrow$ knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$. There are $(\# \mathcal{A})^k$ such words.
- Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$ and v Lyndon \Rightarrow knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$. There are $\sum_{i=1}^{k} \frac{1}{i} \sum_{d|i} \mu(d) (\# \mathcal{A})^{\frac{i}{d}}$ such words.
- Knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^k$ and v Lyndon $\not\Rightarrow$ knowing $\binom{u}{v}$ for all $v \in \mathcal{A}^{\leq k}$.

Counterexample: $u_1 = babaa, u_2 = bbaaa.$ We have $\binom{u_1}{aab} = \binom{u_2}{aab}, \binom{u_1}{abb} = \binom{u_2}{abb}, \binom{u_1}{ab} \neq \binom{u_2}{ab}.$ • We want to reduce the number of computed $\binom{u}{v}$. So, knowing the whole *k*-deck of *u* is too much.

• We want to reduce the number of computed $\binom{u}{v}$. So, knowing the whole *k*-deck of *u* is too much.

If w is of length < k, then

$$\begin{pmatrix} u \\ w \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} u \\ v \end{pmatrix} = 0 \quad \forall v \text{ having } w \text{ as subword}$$

and a whole part of the k-deck of u is known.

• We want to reduce the number of computed $\binom{u}{v}$. So, knowing the whole *k*-deck of *u* is too much.

If w is of length < k, then

$$\begin{pmatrix} u \\ w \end{pmatrix} = 0 \Rightarrow \begin{pmatrix} u \\ v \end{pmatrix} = 0 \quad \forall v \text{ having } w \text{ as subword}$$

and a whole part of the k-deck of u is known.

• We don't want to restrict to words v having all the same length.

Reconstructing words from right-bounded-block words

1 Classical reconstruction problem: survey of the results





Extending to an arbitrary finite alphabet

•
$$Q(u, b)? = 6. \rightarrow {\binom{u}{a}} = 4.$$

 $\exists s_1, \dots, s_5 \in \mathbb{N}_0 \text{ s.t. } u = b^{s_1} a b^{s_2} a b^{s_3} a b^{s_4} a b^{s_5}$
and $s_1 + s_2 + s_3 + s_4 + s_5 = 6$

•
$$Q(u, b)? = 6. \rightarrow {\binom{u}{a}} = 4.$$

 $\exists s_1, \dots, s_5 \in \mathbb{N}_0 \text{ s.t. } u = b^{s_1} a b^{s_2} a b^{s_3} a b^{s_4} a b^{s_5}$
and $s_1 + s_2 + s_3 + s_4 + s_5 = 6$

• Q(u, ab)? = 4. $s_2 + 2s_3 + 3s_4 + 4s_5 = 4.$ $(s_1, s_2, s_3, s_4, s_5) \in \{(5, 0, 0, 0, 1), (4, 1, 0, 1, 0), (4, 0, 2, 0, 0), (3, 2, 1, 0, 0), (2, 4, 0, 0, 0)\}.$

•
$$Q(u, b)? = 6. \rightarrow {\binom{u}{a}} = 4.$$

 $\exists s_1, \dots, s_5 \in \mathbb{N}_0 \text{ s.t. } u = b^{s_1} a b^{s_2} a b^{s_3} a b^{s_4} a b^{s_5}$
and $s_1 + s_2 + s_3 + s_4 + s_5 = 6$

• Q(u, ab)? = 4. $s_2 + 2s_3 + 3s_4 + 4s_5 = 4.$

 $(s_1, s_2, s_3, s_4, s_5) \in \{(5, 0, 0, 0, 1), (4, 1, 0, 1, 0), (4, 0, 2, 0, 0), (3, 2, 1, 0, 0), (2, 4, 0, 0, 0)\}.$

• $Q(u, a^2b)? = 2.$

$$s_3 + 3s_4 + 6s_5 = 2.$$

The unique solution is: (4, 0, 2, 0, 0) and u = bbbbaabbaa.

First question: Q(u, b)? Assume $\binom{u}{b} \geq \frac{|u|}{2}$.

$$u = b^{s_1} a b^{s_2} a \cdots a b^{s_{|u|_a+1}}$$

First question: Q(u, b)? Assume $\binom{u}{b} \ge \frac{|u|}{2}$.

$$u = b^{s_1} a b^{s_2} a \cdots a b^{s_{|u|_a+1}}$$

Ask $Q(u, ab), Q(u, a^2b), \ldots, Q(u, a^mb)$ until u is determined. In all cases: $m \le |u|_a$.

First question: Q(u, b)? Assume $\binom{u}{b} \ge \frac{|u|}{2}$.

$$u = b^{s_1} a b^{s_2} a \cdots a b^{s_{|u|_a+1}}$$

Ask $Q(u, ab), Q(u, a^2b), \ldots, Q(u, a^mb)$ until u is determined. In all cases: $m \le |u|_a$.

$$\begin{cases} s_1 + s_2 + \dots + s_{|u|_a+1} = \binom{u}{b} \\ s_2 + 2s_3 + \dots + |u|_a s_{|u|_a+1} = \binom{u}{ab} \\ s_3 + 3s_4 + \dots + \binom{|u|_a - 1}{2} s_{|u|_a+1} = \binom{u}{a^2b} \\ \vdots \\ s_{|u|_a+1} = \binom{u}{a^{|u|_ab}} \end{cases}$$

Coefficients of the *i*-th equation are the first coefficients of the *i*-th column of the Pascal triangle.

First question: Q(u, b)? Assume $\binom{u}{b} \leq \frac{|u|}{2}$.

$$u = a^{s_1} b a^{s_2} b \cdots b a^{s_{|u|_b+1}}$$

Ask $Q(u, ba), Q(u, b^2a), \ldots, Q(u, b^ma)$ until u is determined. In all cases: $m \le |u|_b$.

$$\begin{pmatrix} s_1 & +s_2 & +\dots & +s_{|u|_b+1} & = \binom{u}{a} \\ s_2 & +2s_3 & +\dots & +|u|_b s_{|u|_b+1} & = \binom{u}{ba} \\ s_3 & +3s_4 + \dots & +\binom{|u|_b-1}{2} s_{|u|_b+1} & = \binom{u}{b^2a} \\ \vdots & & \\ s_{|u|_b+1} & = \binom{u}{b^{|u|_ba}}. \end{cases}$$

Coefficients of the *i*-th equation are the first coefficients of the *i*-th column of the Pascal triangle.

The number of questions is at most

$$\begin{cases} |u|_{a} + 1 & \text{if } |u|_{a} \leq \frac{|u|}{2} \\ |u|_{b} + 1 & \text{if } |u|_{b} \leq \frac{|u|}{2} \end{cases}$$

Hence

$$\left\lfloor \frac{|u|}{2} \right\rfloor + 1$$

questions are enough.

Reconstructing words from right-bounded-block words

1 Classical reconstruction problem: survey of the results

2 Binary case: the results



Extending to an arbitrary finite alphabet

The idea: projections on binary alphabets

Let $\mathcal{A} = \{a_1, \ldots, a_q\}$ and $u \in \mathcal{A}^*$. Idea: use the algorithm on binary alphabets. Let $\mathcal{A} = \{a_1, \ldots, a_q\}$ and $u \in \mathcal{A}^*$. Idea: use the algorithm on binary alphabets.

Let $\{a, b\} \subset A$. Denote by $\pi_{a,b}(u)$ the projection of u on the binary alphabet $\{a, b\}$:

$$\pi_{a,b}: \begin{cases} a \mapsto a \\ b \mapsto b \\ c \mapsto \varepsilon \text{ if } c \in \mathcal{A} \setminus \{a, b\} \end{cases}$$

Let $\mathcal{A} = \{a_1, \ldots, a_q\}$ and $u \in \mathcal{A}^*$. Idea: use the algorithm on binary alphabets.

Let $\{a, b\} \subset A$. Denote by $\pi_{a,b}(u)$ the projection of u on the binary alphabet $\{a, b\}$:

$$\pi_{a,b}: \begin{cases} a \mapsto a \\ b \mapsto b \\ c \mapsto \varepsilon \text{ if } c \in \mathcal{A} \setminus \{a, b\} \end{cases}$$

1 Try to reconstruct $\pi_{a,b}(u)$ for every subalphabet $\{a, b\} \subset \mathcal{A}$ of size 2.

Let $\mathcal{A} = \{a_1, \ldots, a_q\}$ and $u \in \mathcal{A}^*$. Idea: use the algorithm on binary alphabets.

Let $\{a, b\} \subset A$. Denote by $\pi_{a,b}(u)$ the projection of u on the binary alphabet $\{a, b\}$:

$$\pi_{a,b}: \begin{cases} a \mapsto a \\ b \mapsto b \\ c \mapsto \varepsilon \text{ if } c \in \mathcal{A} \setminus \{a, b\} \end{cases}$$

Ory to reconstruct π_{a,b}(u) for every subalphabet {a, b} ⊂ A of size 2.
Ory Combine all projections {π_{a,b}(u)} to reconstruct u.

Determine the 1-deck: Q(u, a)? = 3, Q(u, b)? = 1. $\rightarrow |u|_n = 2$.

Determine the 1-deck: Q(u, a)? = 3, Q(u, b)? = 1. $\rightarrow |u|_n = 2$.

• Consider $\{b, n\}$. Take b < n. Q(u, bn)? = 2

 $\pi_{b,n}(u) = n^{s_1} b n^{s_2}$ and $s_1 + s_2 = 2, s_2 = 2$

Therefore $\pi_{b,n}(u) = bnn$.

Determine the 1-deck: Q(u, a)? = 3, Q(u, b)? = 1. $\rightarrow |u|_n = 2$.

• Consider $\{b, n\}$. Take b < n. Q(u, bn)? = 2

$$\pi_{b,n}(u)=n^{s_1}bn^{s_2}$$
 and $s_1+s_2=2,s_2=2$

Therefore $\pi_{b,n}(u) = bnn$.

2 Consider the subalphabet $\{b, a\}$. Take b < a. Q(u, ba)? = 3

$$\pi_{a,b}(u) = a^{t_1} b a^{t_2}$$
 and $t_1 + t_2 = 3, t_2 = 3$

Therefore $\pi_{a,b}(u) = baaa$.

Determine the 1-deck: Q(u, a)? = 3, Q(u, b)? = 1. $\rightarrow |u|_n = 2$.

• Consider $\{b, n\}$. Take b < n. Q(u, bn)? = 2

$$\pi_{b,n}(u)=n^{s_1}bn^{s_2}$$
 and $s_1+s_2=2,s_2=2$

Therefore $\pi_{b,n}(u) = bnn$.

2 Consider the subalphabet $\{b, a\}$. Take b < a. Q(u, ba)? = 3

$$\pi_{a,b}(u) = a^{t_1} b a^{t_2}$$
 and $t_1 + t_2 = 3, t_2 = 3$

Therefore $\pi_{a,b}(u) = baaa$.

Solution So

 $\pi_{n,a}(u) = a^{p_1} n a^{p_2} n a^{p_3}$ and $p_1 + p_2 + p_3 = 3, p_2 + 2p_3 = 3, p_3 = 1$

Therefore $\pi_{n,a}(u) = anana$.

Determine the 1-deck: Q(u, a)? = 3, Q(u, b)? = 1. $\rightarrow |u|_n = 2$.

• Consider $\{b, n\}$. Take b < n. Q(u, bn)? = 2

$$\pi_{b,n}(u)=n^{s_1}bn^{s_2}$$
 and $s_1+s_2=2,s_2=2$

Therefore $\pi_{b,n}(u) = bnn$.

2 Consider the subalphabet $\{b, a\}$. Take b < a. Q(u, ba)? = 3

$$\pi_{a,b}(u) = a^{t_1} b a^{t_2}$$
 and $t_1 + t_2 = 3, t_2 = 3$

Therefore $\pi_{a,b}(u) = baaa$.

Solution So

$$\pi_{n,a}(u) = a^{p_1} n a^{p_2} n a^{p_3}$$
 and $p_1 + p_2 + p_3 = 3, p_2 + 2p_3 = 3, p_3 = 1$

Therefore $\pi_{n,a}(u) = anana$.

We get u = banana. Marie Lejeune (Liège University) Let $\mathcal{A} = \{a_1, \ldots, a_q\}$.

- Is *u* always uniquely determined from
 {*π*_{ai,aj}(*u*) : {*a*_i, *a*_j} ⊂ {*a*₁,...,*a*_q}}? How to reconstruct it?
- Compare the maximal number of questions with the bound of the classical reconstruction problem.

K-markings

Let $K = (k_a)_{a \in \mathcal{A}}$ be a sequence of natural numbers. Let $u^{(1)}, \ldots, u^{(\ell)}$ be words of \mathcal{A}^* .

A *K*-marking of $u^{(1)}, \ldots, u^{(\ell)}$ is a mapping

$$\psi: \{(j,i): j \in [\ell], i \in [|u^{(j)}|]\} \to \mathbb{N}$$

such that, $\forall j \in [\ell], \forall i, m \in [|u^{(j)}|], a \in \mathcal{A}$, there holds

• if
$$u_i^{(j)} = a$$
 then $\psi(j, i) \le k_a$,

• if
$$i < m$$
 and $u_i^{(j)} = u_m^{(j)}$ then $\psi(j, i) < \psi(j, m)$.

K-markings

Let $K = (k_a)_{a \in \mathcal{A}}$ be a sequence of natural numbers. Let $u^{(1)}, \ldots, u^{(\ell)}$ be words of \mathcal{A}^* .

A *K*-marking of $u^{(1)}, \ldots, u^{(\ell)}$ is a mapping

$$\psi:\{(j,i):j\in [\ell],i\in [|u^{(j)}|]\}\to \mathbb{N}$$

such that, $\forall j \in [\ell], \forall i, m \in [|u^{(j)}|], a \in \mathcal{A}$, there holds

• if
$$u_i^{(j)} = a$$
 then $\psi(j,i) \leq k_a$,

• if
$$i < m$$
 and $u_i^{(j)} = u_m^{(j)}$ then $\psi(j, i) < \psi(j, m)$.

Example:

$$\mathcal{A} = \{a, b, c\}, \ K = (2, 3, 2)$$

and $u^{(1)} = bcab, u^{(2)} = aba$

j	1				2		
i	1	2	3	4	1	2	3
$u_i^{(j)}$	b	с	а	b	а	b	а

Topological sorting of a directed graph

Let G = (V, E) be a directed graph. A *topological sorting* of G is a linear ordering $v_1 < \ldots < v_n$ of V such that every edge in G is of the type (v_i, v_j) with i < j.



Topological sorting of a directed graph

Let G = (V, E) be a directed graph. A *topological sorting* of G is a linear ordering $v_1 < \ldots < v_n$ of V such that every edge in G is of the type (v_i, v_j) with i < j.



Folklore: Any directed graph has a topological sorting if and only if it is acyclic.

Topological sorting of a directed graph

Let G = (V, E) be a directed graph. A *topological sorting* of G is a linear ordering $v_1 < \ldots < v_n$ of V such that every edge in G is of the type (v_i, v_j) with i < j.



Folklore: Any directed graph has a topological sorting if and only if it is acyclic.

The topological sorting of an acyclic graph is unique if and only if there is a path going through every vertex.

Graph G_{ψ} associated to a K-marking ψ

Let ψ be a *K*-marking of $u^{(1)}, \ldots, u^{(\ell)}$. Define the graph $G_{\psi} = (V_{\psi}, E_{\psi})$ such that:

$$egin{aligned} V_\psi &= \{(a)_i: a \in \mathcal{A}, 1 \leq i \leq k_a\} \ E_\psi &= \{((a)_i, (a)_{i+1})\} \cup \{((u_i^{(j)})_{\psi(j,i)}, (u_{i+1}^{(j)})_{\psi(j,i+1)})\} \end{aligned}$$
Let ψ be a *K*-marking of $u^{(1)}, \ldots, u^{(\ell)}$. Define the graph $G_{\psi} = (V_{\psi}, E_{\psi})$ such that:

$$egin{aligned} &V_\psi = \{(a)_i: a \in \mathcal{A}, 1 \leq i \leq k_a \} \ &E_\psi = \{((a)_i, (a)_{i+1})\} \cup \{((u_i^{(j)})_{\psi(j,i)}, (u_{i+1}^{(j)})_{\psi(j,i+1)})\} \end{aligned}$$



Let ψ be a *K*-marking of $u^{(1)}, \ldots, u^{(\ell)}$. Define the graph $G_{\psi} = (V_{\psi}, E_{\psi})$ such that:

$$egin{aligned} &V_\psi = \{(a)_i: a \in \mathcal{A}, 1 \leq i \leq k_a \} \ &E_\psi = \{((a)_i, (a)_{i+1})\} \cup \{((u_i^{(j)})_{\psi(j,i)}, (u_{i+1}^{(j)})_{\psi(j,i+1)})\} \end{aligned}$$



Let ψ be a *K*-marking of $u^{(1)}, \ldots, u^{(\ell)}$. Define the graph $G_{\psi} = (V_{\psi}, E_{\psi})$ such that:

$$egin{aligned} &V_\psi = \{(a)_i: a \in \mathcal{A}, 1 \leq i \leq k_a \} \ &E_\psi = \{((a)_i, (a)_{i+1})\} \cup \{((u_i^{(j)})_{\psi(j,i)}, (u_{i+1}^{(j)})_{\psi(j,i+1)})\} \end{aligned}$$



Let ψ be a *K*-marking of $u^{(1)}, \ldots, u^{(\ell)}$. Define the graph $G_{\psi} = (V_{\psi}, E_{\psi})$ such that:

$$egin{aligned} V_\psi &= \{(a)_i: a \in \mathcal{A}, 1 \leq i \leq k_a \} \ E_\psi &= \{((a)_i, (a)_{i+1})\} \cup \{((u_i^{(j)})_{\psi(j,i)}, (u_{i+1}^{(j)})_{\psi(j,i+1)})\} \end{aligned}$$



Let ψ be a *K*-marking of $u^{(1)}, \ldots, u^{(\ell)}$. Define the graph $G_{\psi} = (V_{\psi}, E_{\psi})$ such that:

$$egin{aligned} &V_\psi = \{(a)_i: a \in \mathcal{A}, 1 \leq i \leq k_a \} \ &E_\psi = \{((a)_i, (a)_{i+1})\} \cup \{((u_i^{(j)})_{\psi(j,i)}, (u_{i+1}^{(j)})_{\psi(j,i+1)})\} \end{aligned}$$

Example:



 ψ is a K-marking $\Rightarrow G_{\psi}$ admits a topological sorting

Theorem: Let $u^{(1)}, \ldots, u^{(\ell)} \in \mathcal{A}^*$. Let $K = (k_a)_{a \in \mathcal{A}}$. There exists a word $u \in \mathcal{A}^*$ such that $u^{(j)}$ is a subword of u for all $j \in [\ell]$ and $|u|_a = k_a$ for all $a \in \mathcal{A}$ \Leftrightarrow

```
Theorem: Let u^{(1)}, \ldots, u^{(\ell)} \in \mathcal{A}^*. Let \mathcal{K} = (k_a)_{a \in \mathcal{A}}.
There exists a word u \in \mathcal{A}^* such that u^{(j)} is a subword of u for all j \in [\ell]
and |u|_a = k_a for all a \in \mathcal{A}
\Leftrightarrow
there exists a \mathcal{K}-marking \psi of the words u^{(1)}, \ldots, u^{(\ell)} and a topological
sorting of G_{\psi}.
```

```
Theorem: Let u^{(1)}, \ldots, u^{(\ell)} \in \mathcal{A}^*. Let K = (k_a)_{a \in \mathcal{A}}.
There exists a word u \in \mathcal{A}^* such that u^{(j)} is a subword of u for all j \in [\ell]
and |u|_a = k_a for all a \in \mathcal{A}
\Leftrightarrow
there exists a K-marking \psi of the words u^{(1)}, \ldots, u^{(\ell)} and a topological
sorting of G_{\eta}.
Moreover, if the values of K are minimal.
```

Values of K are minimal means that

$$\forall a \in \mathcal{A}, k_a = \max_{j \in [\ell]} |u^{(j)}|_a$$

```
Theorem: Let u^{(1)}, \ldots, u^{(\ell)} \in \mathcal{A}^*. Let K = (k_a)_{a \in \mathcal{A}}.
There exists a word u \in \mathcal{A}^* such that u^{(j)} is a subword of u for all j \in [\ell]
and |u|_a = k_a for all a \in \mathcal{A}
\Leftrightarrow
there exists a K-marking \psi of the words u^{(1)}, \ldots, u^{(\ell)} and a topological
sorting of G_{\psi}.
```

Moreover, if the values of K are *minimal*, if there is a unique K-marking ψ such that G_{ψ} admits a topological sorting and if this topological sorting is unique, then u is unique.

Values of K are minimal means that

$$\forall a \in \mathcal{A}, k_a = \max_{j \in [\ell]} |u^{(j)}|_a$$

Illustrating the theorem



No valid topological sorting G_{ψ} for this K-marking ψ of $u^{(1)}, u^{(2)}$.

Illustrating the theorem

$$A = \{a, b, c\}, K = (2, 3, 2)$$

j	1				2		
i	1	2	3	4	1	2	3
$u_i^{(j)}$	b	с	а	b	а	b	а
$\psi'(j,i)$	1	2	1	3	1	2	2



Another K-marking ψ' of $u^{(1)}, u^{(2)}$. $G_{\psi'}$ admits a topological sorting.

There exists $u \in \{a, b, c\}^*$ having $u^{(1)}$ and $u^{(2)}$ as subwords, and such that $|u|_a = 2$, $|u|_b = 3$, $|u|_c = 2$.

 $u \in \{bcababc, bcabacb, bcbcaba, cbbcaba, \ldots\}$

Since $u^{(1)} = bcab$ and $u^{(2)} = aba$, the minimal K is (2, 2, 1).

j	1				2		
i	1	2	3	4	1	2	3
$u_i^{(j)}$	b	с	а	b	а	Ь	а
$\psi(j,i)$	1	1	1 or 2	2	1	1 or 2	2

There are 4 possible *K*-markings.

Since $u^{(1)} = bcab$ and $u^{(2)} = aba$, the minimal K is (2, 2, 1).

j	1				2		
i	1	2	3	4	1	2	3
$u_i^{(j)}$	b	С	а	b	а	b	а
$\psi_1(j,i)$	1	1	1	2	1	1	2

There are 4 possible K-markings.



 G_{ψ_1} does not have a topological sorting.

Since $u^{(1)} = bcab$ and $u^{(2)} = aba$, the minimal K is (2, 2, 1).

j	1				2		
i	1	2	3	4	1	2	3
$u_i^{(j)}$	b	с	а	b	а	b	а
$\psi_2(j,i)$	1	1	1	2	1	2	2

There are 4 possible K-markings.



Unique topological sorting of G_{ψ_2} . Reconstructing u = bcaba, having $u^{(1)}$ and $u^{(2)}$ as subwords.

Since $u^{(1)} = bcab$ and $u^{(2)} = aba$, the minimal K is (2, 2, 1).

j	1				2		
i	1	2	3	4	1	2	3
$u_i^{(j)}$	b	С	а	b	а	b	а
$\psi_{3}(j,i)$	1	1	2	2	1	1	2

There are 4 possible K-markings.



Unique topological sorting of G_{ψ_3} . Reconstructing u = abcab, having $u^{(1)}$ and $u^{(2)}$ as subwords.

Since $u^{(1)} = bcab$ and $u^{(2)} = aba$, the minimal K is (2, 2, 1).

j	1				2		
i	1	2	3	4	1	2	3
$u_i^{(j)}$	b	с	а	b	а	b	а
$\psi_4(j,i)$	1	1	2	2	1	2	2

There are 4 possible K-markings.



 G_{ψ_4} does not have a topological sorting.

Since $u^{(1)} = bcab$ and $u^{(2)} = aba$, the minimal K is (2, 2, 1).

j	1				2		
i	1	2	3	4	1	2	3
$u_i^{(j)}$	b	с	а	b	а	Ь	а
$\psi(j,i)$	1	1	1 or 2	2	1	1 or 2	2

Since there exist two K-markings ψ_2 and ψ_3 admitting a topological sorting of their associated graph, the word u is not unique.

Let \mathcal{A} be an alphabet of size q. Assume that $u^{(1)}, \ldots, u^{(\ell)}$ are all the projections of an unknown word $u \in \mathcal{A}^*$ over binary subalphabets.

Let \mathcal{A} be an alphabet of size q. Assume that $u^{(1)}, \ldots, u^{(\ell)}$ are all the projections of an unknown word $u \in \mathcal{A}^*$ over binary subalphabets.

Take $K = (k_a)_{a \in \mathcal{A}}$ where $k_a = |u^{(j)}|_a$ if $u^{(j)}$ is the projection of u over a subalphabet containing a. Then K is minimal.

Let \mathcal{A} be an alphabet of size q. Assume that $u^{(1)}, \ldots, u^{(\ell)}$ are all the projections of an unknown word $u \in \mathcal{A}^*$ over binary subalphabets.

Take $K = (k_a)_{a \in \mathcal{A}}$ where $k_a = |u^{(j)}|_a$ if $u^{(j)}$ is the projection of u over a subalphabet containing a. Then K is minimal.

• The K-marking
$$\psi$$
 of $u^{(1)}, \ldots, u^{(\ell)}$ is unique:

Let \mathcal{A} be an alphabet of size q. Assume that $u^{(1)}, \ldots, u^{(\ell)}$ are all the projections of an unknown word $u \in \mathcal{A}^*$ over binary subalphabets.

Take $K = (k_a)_{a \in A}$ where $k_a = |u^{(j)}|_a$ if $u^{(j)}$ is the projection of u over a subalphabet containing a. Then K is minimal.

 The K-marking \u03c6 of u⁽¹⁾,..., u^(\u03c6) is unique: If the m-th occurrence of letter a in u^(j) appears in position i, then \u03c6(i,j) = m.

Let \mathcal{A} be an alphabet of size q. Assume that $u^{(1)}, \ldots, u^{(\ell)}$ are all the projections of an unknown word $u \in \mathcal{A}^*$ over binary subalphabets.

Take $K = (k_a)_{a \in A}$ where $k_a = |u^{(j)}|_a$ if $u^{(j)}$ is the projection of u over a subalphabet containing a. Then K is minimal.

- The K-marking \u03c6 of u⁽¹⁾,..., u^(\u03c6) is unique: If the m-th occurrence of letter a in u^(j) appears in position i, then \u03c6(i,j) = m.
- 2 It can be proven that G_{ψ} is acyclic and has a path going through every vertex, hence the topological sorting exists and is unique.

Let \mathcal{A} be an alphabet of size q. Assume that $u^{(1)}, \ldots, u^{(\ell)}$ are all the projections of an unknown word $u \in \mathcal{A}^*$ over binary subalphabets.

Take $K = (k_a)_{a \in A}$ where $k_a = |u^{(j)}|_a$ if $u^{(j)}$ is the projection of u over a subalphabet containing a. Then K is minimal.

- The K-marking \u03c6 of u⁽¹⁾,..., u^(\u03c6) is unique: If the m-th occurrence of letter a in u^(j) appears in position i, then \u03c6(i,j) = m.
- 2 It can be proven that G_{ψ} is acyclic and has a path going through every vertex, hence the topological sorting exists and is unique.

Taking the letters of vertices of G_{ψ} following the topological sorting gives the unique word u.

Example: reconstructing "banana" from its binary projections.

$$u^{(1)} = baaa, \ u^{(2)} = bnn, \ u^{(3)} = anana.$$

Then the minimal K is $(k_a, k_b, k_n) = (3, 1, 2)$.



Example: reconstructing "banana" from its binary projections.

$$u^{(1)} = baaa, \ u^{(2)} = bnn, \ u^{(3)} = anana.$$

Then the minimal K is $(k_a, k_b, k_n) = (3, 1, 2)$.



Reconstructing a word from its binary projections can be done in linear time w.r.t. the total length of the projections.

Let $u \in \{a_1, \ldots, a_q\}^n$.

• First, ask q-1 questions to determine the 1-deck of u.

- First, ask q-1 questions to determine the 1-deck of u.
- Let σ be a permutation of $\{1, \ldots, q\}$ such that $|u|_{a_{\sigma(1)}} \leq |u|_{a_{\sigma(2)}} \leq \ldots \leq |u|_{a_{\sigma(q)}}$

- First, ask q-1 questions to determine the 1-deck of u.
- Let σ be a permutation of $\{1, \ldots, q\}$ such that $|u|_{a_{\sigma(1)}} \leq |u|_{a_{\sigma(2)}} \leq \ldots \leq |u|_{a_{\sigma(q)}}$
- Consider all subalphabets $\{a_{\sigma(i)}, a_{\sigma(j)}\}$. Assume i < j. Ask questions $Q(u, a_{\sigma(i)}a_{\sigma(j)}), Q(u, a_{\sigma(i)}^2a_{\sigma(j)}), \ldots, Q(u, a_{\sigma(i)}^{|u|_{a_{\sigma(i)}}}a_{\sigma(j)})$. There are $|u|_{a_{\sigma(i)}}$ such questions.

- First, ask q-1 questions to determine the 1-deck of u.
- Let σ be a permutation of $\{1, \ldots, q\}$ such that $|u|_{a_{\sigma(1)}} \leq |u|_{a_{\sigma(2)}} \leq \ldots \leq |u|_{a_{\sigma(q)}}$
- Consider all subalphabets $\{a_{\sigma(i)}, a_{\sigma(j)}\}$. Assume i < j. Ask questions $Q(u, a_{\sigma(i)}a_{\sigma(j)}), Q(u, a_{\sigma(i)}^2 a_{\sigma(j)}), \dots, Q(u, a_{\sigma(i)}^{|u|_{a_{\sigma(i)}}} a_{\sigma(j)})$. There are $|u|_{a_{\sigma(i)}}$ such questions.
- All projections $\pi_{a_{\sigma(i)},a_{\sigma(j)}}(u)$ are uniquely determined; deduce u.

Let $u \in \{a_1, \ldots, a_q\}^n$.

- First, ask q-1 questions to determine the 1-deck of u.
- Let σ be a permutation of $\{1, \ldots, q\}$ such that $|u|_{a_{\sigma(1)}} \leq |u|_{a_{\sigma(2)}} \leq \ldots \leq |u|_{a_{\sigma(q)}}$
- Consider all subalphabets $\{a_{\sigma(i)}, a_{\sigma(j)}\}$. Assume i < j. Ask questions $Q(u, a_{\sigma(i)}a_{\sigma(j)}), Q(u, a_{\sigma(i)}^2a_{\sigma(j)}), \ldots, Q(u, a_{\sigma(i)}^{|u|_{a_{\sigma(i)}}}a_{\sigma(j)})$. There are $|u|_{a_{\sigma(i)}}$ such questions.
- All projections $\pi_{a_{\sigma(i)},a_{\sigma(j)}}(u)$ are uniquely determined; deduce u.

Number of asked questions:

$$(q-1)+\sum_{i=1}^q |u|_{a_{\sigma(i)}}(q-i)$$

Comparing to the classical reconstruction problem

Let $\mathcal{A} = \{a_1, \dots, a_q\}$. Recall: in the classical reconstruction problem, knowing the $(\lfloor \frac{16}{7}\sqrt{n} \rfloor + 5)$ -deck of u suffices.

Comparing to the classical reconstruction problem

Let $\mathcal{A} = \{a_1, \ldots, a_q\}$. Recall: in the classical reconstruction problem, knowing the $(\lfloor \frac{16}{7}\sqrt{n} \rfloor + 5)$ -deck of u suffices. Using Lyndon words, that leads to

$$\sum_{i=1}^{\lfloor \frac{16}{7}\sqrt{n}\rfloor+5}\frac{1}{i}\sum_{d\mid i}\mu(d)q^{\frac{i}{d}}$$

questions.

Comparing to the classical reconstruction problem

Let $\mathcal{A} = \{a_1, \ldots, a_q\}$. Recall: in the classical reconstruction problem, knowing the $(\lfloor \frac{16}{7}\sqrt{n} \rfloor + 5)$ -deck of u suffices. Using Lyndon words, that leads to

$$\sum_{i=1}^{\lfloor \frac{16}{7}\sqrt{n}\rfloor+5}\frac{1}{i}\sum_{d\mid i}\mu(d)q^{\frac{i}{d}}$$

questions. This bound is strictly greater than

$$(q-1) + \sum_{i=1}^{q} |u|_{s_{\sigma(i)}}(q-i)$$

for every q and $n \ge q - 1$.

Thank you!