PhD Defense

Distance and Conjugacy of Word Transducers

Saina Sunny
saina19231102@iitgoa.ac.in

Roll No: 19231102

Advisor: Dr. Amaldev Manuel

*School of Mathematics and Computer Science
Indian Institute of Technology Goa, India

July 25, 2025

Preliminaries

1. Distance on Words

Metric on Words

- A metric on words is a distance between words that satisfies
  - $d(u, v) = 0 \iff u = v$.
  - $d(u, v) = d(v, u)$.
  - $d(u, v) \leq d(u, w) + d(w, v)$.

- An edit distance between two words is the minimum number of given edit operations – such as insertions, deletions, substitutions – required to rewrite one word to another.

- Example, $d(hello, yellow) = 2$.

$$\cancel{h}yellow$$

$$yellow$$

Common Edit Distances

Edit Distance	Permissible Operations
Hamming	letter-to-letter substitutions
Conjugacy	left and right cyclic shifts
Transposition	swapping adjacent letters
Longest Common Subsequence	insertions and deletions
Levenshtein	insertions, deletions, and substitutions
Damerau-Levenshtein	insertions, deletions, substitutions and adjacent transpositions

Table 1: Edit Distances.

Preliminaries

2. Word Transducers

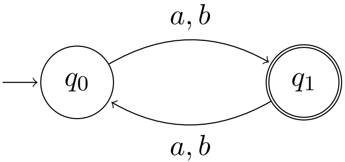
Finite State Word Transducer

- Machine that reads an input word and produces output word(s) using finite memory.
- Examples: spell checkers, grammatical tools.

Automaton vs. Transducer

Automaton

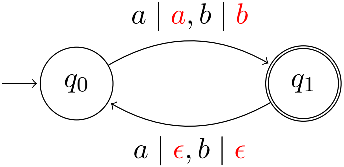
- Accepts a set of words.



Accepts odd length words.

Transducer

- Defines a relation over input-output words.

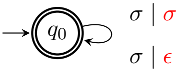


Outputs letters at odd positions.

$$aba \rightarrow aa$$

Rational Relations

- Rational relations are relations defined by transducers.

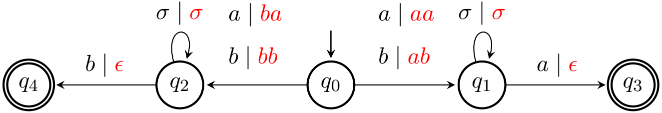


defines the relation $\{(u, v) \mid v \text{ is a subword of } u\}$.

$\sigma \in \{a, b\}$

Rational Functions

- Rational functions are functions defined by transducers.

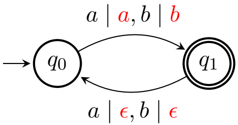


defines $\mathsf{f}_{\mathsf{last}} : \mathsf{u}\sigma \mapsto \sigma\mathsf{u}$

$$\sigma \in \{a, b\}$$

Sequential Functions

- Sequential functions are functions defined by input-deterministic transducers.



- The function $\mathsf{f}_{\mathsf{last}} : \mathsf{u}\sigma \to \sigma\mathsf{u}$ is not sequential.

Recap: (Sub)classes of Transducers

$$\text{sequential functions} \subsetneq \text{rational functions} \subsetneq \text{rational relations}$$

- Rational relations - relations defined by transducers.
- Rational functions - functions defined by transducers.
- Sequential functions - functions defined by input-deterministic transducers.

Distance and Conjugacy of Word Transducers

1. Compare word transducers by defining and computing a metric over transducers.
2. Study a combinatorial property: conjugacy of transducers.
3. Study the approximate variants of some classical problems on transducers.

1. Comparing Word Transducers

Joint work with

Dr. C. Aiswarya (Chennai Mathematical Institute)

Dr. Amaldev Manuel (Indian Institute of Technology Goa)

How to Compare Transducers?

- Equivalence - whether two transducers define the same relation (or function).
  - Rational relations - Undecidable [Fischer and Rosenberg, 1968].
  - Rational functions - PSPACE-complete [Gurari and Ibarra, 1983].
  - Sequential functions - PTIME [Gurari and Ibarra, 1983].
- We quantitatively compare transducers by computing their distance.

Distance between Functions

- Let d be a metric on words. We can lift it to word-to-word functions.

$$d(f,g) = \begin{cases} \sup\{\, d(f(w), g(w)) \mid w \in dom(f)\} & \text{if } dom(f) = dom(g) \\ \infty & \text{otherwise} \end{cases}$$

Distance between Functions

- Let d be a metric on words. We can lift it to word-to-word functions.

$$d(f,g) = \begin{cases} \sup\{\, d(f(w), g(w)) \mid w \in dom(f) \} & \text{if } dom(f) = dom(g) \\ \infty & \text{otherwise} \end{cases}$$

Examples

- Consider functions $\mathsf{f}_{\mathsf{last}} : \mathsf{u}\sigma \mapsto \sigma\mathsf{u}$ and $\mathsf{f}_{\mathsf{id}} : \mathsf{u}\sigma \mapsto \mathsf{u}\sigma$

$$d(f_{\mathsf{last}}, f_{\mathsf{id}}) = 2 \text{ (w.r.t. Levenshtein distance).}$$

$$d(f_{\mathsf{last}}, f_{\mathsf{id}}) = \infty \text{ (w.r.t. Hamming distance).}$$

Distance between Functions

- Let d be a distance on words. We can lift it to word-to-word functions.

$$d(f,g) = \begin{cases} \sup\{\, d(f(w), g(w)) \mid w \in dom(f) \} & \text{if } dom(f) = dom(g) \\ \infty & \text{otherwise} \end{cases}$$

- f and g are close if their distance $d(f,g)$ is finite.

Distance between Functions: Results

Theorem

The distance between rational functions w.r.t. metrics given in Table 1 are computable.

Distance between Functions: Results

Theorem

The distance between rational functions w.r.t. metrics given in Table 1 are computable.

Edit Distance	Permissible Operations
Hamming	letter-to-letter substitutions
Conjugacy	left and right cyclic shifts
Transposition	swapping adjacent letters
Longest Common Subsequence	insertions and deletions
Levenshtein	insertions, deletions, and substitutions
Damerau-Levenshtein	insertions, deletions, substitutions and adjacent transpositions

Distance between Functions: Results

Theorem

The distance between rational functions w.r.t. metrics given in Table 1 are computable.

Problem	Input	Question
Closeness problem	functions f, g	$d(f, g) < \infty$?
k-Closeness problem	integer k, functions f, g	$d(f, g) \leq k$?

Theorem

The closeness and k-closeness problems for rational functions w.r.t. metrics given in Table 1 are decidable.

Diameter of a Relation

- The *diameter* of a relation R with respect to a metric d is the supremum of the distance of the every pair in R.

$$\mathsf{dia}_d(R) = \sup \{\, d(u,v) \mid (u,v) \in R \,\}\,.$$

Diameter of a Relation

- The *diameter* of a relation R with respect to a metric d is the supremum of the distance of the every pair in R.

$$\mathsf{dia}_d(R) = \sup \{ d(u,v) \mid (u,v) \in R \} .$$

Examples

The diameter of $\{((ab)^n, (ba)^n) \mid n \geq 0\}$ w.r.t. Levenshtein distance is 2.

$$abab \cdots ab$$

$$baba \cdots ba$$

Diameter of a Relation

- The *diameter* of a relation R with respect to a metric d is the supremum of the distance of the every pair in R.

$$\mathsf{dia}_d(R) = \sup\{\, d(u,v) \mid (u,v) \in R \,\}\,.$$

Examples

The diameter of $\{((ab)^n, (ba)^n) \mid n \geq 0\}$ w.r.t. Levenshtein distance is 2.

$$\cancel{a}bab \cdots ab\textcolor{blue}{a}$$

$$baba \cdots ba$$

Diameter of a Relation : Results

Theorem

The distance problem between rational functions is mutually reducible to the diameter problem of a rational relation.

1. Distance to Diameter:
   - Let f and g be two rational functions with identical domain.
   - Construct a rational relation $R = \{(f(w), g(w)) \mid w \in dom(f)\}$.
   - For any metric d, $d(f, g) = \mathsf{dia}_d(R)$.
2. Diameter to Distance: using a theorem by [Nivat, 1968].

Diameter of a Relation : Boundedness Questions

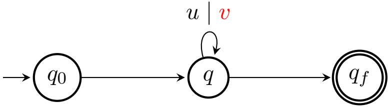
Problem	Input	Question
Bounded diameter problem	rational relation R	$\mathsf{dia}_d(R) < \infty$?
k-Bounded diameter problem	integer k, rational relation R	$\mathsf{dia}_d(R) \leq k$?

Bounded Diameter

- Assume that the rational relation is given by a transducer $\mathcal{T}$.
- If the diameter is finite, can you say something about the loops of $\mathcal{T}$?

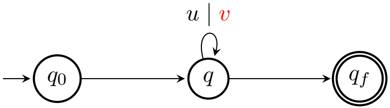
Bounded Diameter

- Assume that the rational relation is given by a transducer $\mathcal{T}$.
- If the diameter is finite, can you say something about the loops of $\mathcal{T}$?
- Let (u, v) be an input-output pair along a loop in $\mathcal{T}$.
    - $|u| = |v|$.



Bounded Diameter

- Assume that the rational relation is given by a transducer $\mathcal{T}$.
- If the diameter is finite, can you say something about the loops of $\mathcal{T}$?
- Let (u, v) be an input-output pair along a loop in $\mathcal{T}$.
  - $|u| = |v|$.
  - $u = xy$ and $v = yx$ (u and v are conjugate) [using Fine-Wilf theorem]



Bounded Diameter

- Assume that the rational relation is given by a transducer $\mathcal{T}$.
- If the diameter is finite, then every input-output pair generated along any loop of $\mathcal{T}$ must be conjugate.

Lemma

The diameter of a rational relation given by a transducer $\mathcal{T}$ w.r.t. Levenshtein family of distances is finite iff the input-output pairs produced in the loops of $\mathcal{T}$ are conjugates.*

$$\cancel{x}yxy\cdots xy{\color{blue}x}$$

$$yxyx\cdots yx$$

*Levenshtein family — Levenshtein, longest common subsequence, Damerau-Levenshtein distance

2. Conjugacy of a Rational Relation

Joint work with

Dr. C. Aiswarya (Chennai Mathematical Institute)

Dr. Amaldev Manuel (Indian Institute of Technology Goa)

Conjugate Words

- Two words u and v are conjugate
  - if there exists words x and y such that $u = xy$ and $v = yx$.

$$\begin{array}{cc} listen & was \\ enlist & saw \end{array}$$

Conjugate Words

- Two words u and v are conjugate
  - if there exists words x and y such that $u = xy$ and $v = yx$.

$$\begin{array}{cc} listen & was \\ enlist & saw \end{array}$$

  - or, iff there exists a word z such that $uz = zv$ [Lyndon and Schützenberger, 1962].

$$listen\textcolor{red}{list} = \textcolor{red}{list}enlist$$

Conjugate Words

- Two words u and v are conjugate
  - if there exists words x and y such that $u = xy$ and $v = yx$.

$$\begin{array}{cc} listen & was \\ enlist & saw \end{array}$$

  - or, iff there exists a word z such that $uz = zv$ [Lyndon and Schützenberger, 1962].

$$listen\color{red}{list} = \color{red}{list}enlist$$

  - or, iff there exists a word z such that $zu = vz$.

$$\color{red}{en}listen = enlist\color{red}{en}$$

Conjugacy of Relations

- A relation is conjugate if every pair of words in the relation are conjugate.

Examples

The relation $\{((ab)^n, (ba)^n) \mid n \geq 0\}$ is conjugate.

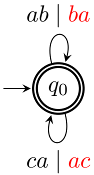
$$a/bab \cdots ab/$$

$$/baba \cdots b/a$$

Conjugacy of Relations : Problem Statement

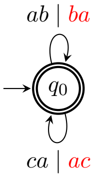
- A relation is conjugate if every pair of words in the relation are conjugate.
- Given a rational relation, is it conjugate?
- Checking if it contains at least one pair of conjugate words is undecidable [Finkel et al., 2023].

Conjugacy of Relations : Challenge

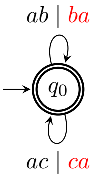


$(abca, baac)$ is not conjugate.

Conjugacy of Relations : Challenge

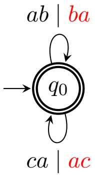


$(abca, baac)$ is not conjugate.

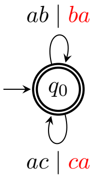


defines a conjugate relation.

Conjugacy of Relations : Challenge



$(abca, baac)$ is not conjugate.



defines a conjugate relation. $\boxed{uz = zv}$

$a/b\cdots ab/ \qquad a/c\cdots ac/$

$/ba\cdots b/a \qquad /ca\cdots c/a$

Conjugacy of Relations : Results

- A word z is a *common witness* of a set of pairs G if

$$\forall (u,v) \in G, uz = zv \;\; \textit{or} \;\; \forall (u,v) \in G, zu = vz.$$

Theorem

Let G be an arbitrary set of pairs of words. TFAE.

1. *G^* is conjugate.*
2. *G^* has a common witness z.*
3. *G has a common witness z.*
4. *Roots of G has a common witness z.*

G^* consists of all pairs obtained by pointwise concatenation of finite pairs in G.

Conjugacy of Relations : Results

Theorem

Let

$$G = (u_0, v_0){G_1}^*(u_1, v_1)$$

where G_1 *is an arbitrary sets of pairs of words, and* $(u_0, v_0), (u_1, v_1)$ *are arbitrary pairs of words. TFAE.*

1. G *is conjugate.*
2. $G_1 \cup \{(u_1 u_0, v_1 v_0)\}$ *has a common witness.*
3. G *has a common witness.*

Conjugacy of Relations : Results

Theorem

Let

$$G = (u_0, v_0){G_1}^*(u_1, v_1) \cdots (u_{k-1}, v_{k-1}){G_k}^*(u_k, v_k)$$

where $k > 0$, $G_1, \ldots, G_k$ are arbitrary sets of pairs of words, and $(u_0, v_0), \ldots, (u_k, v_k)$ are arbitrary pairs of words. TFAE.

1. *G is conjugate.*
2. *Each singleton redux of G (where all but one Kleene star is substituted with (ϵ, ϵ)) has a common witness z.*
3. *G has a common witness z.*

Conjugacy of Relations : Results

Theorem

Checking if a rational relation is conjugate is decidable.

Deciding Conjugacy of a Rational Relation

- Every rational relation can be expressed as a rational expression over pairs of words (using operations union, product and Kleene star).

$$((a, aa) + (b, b))^* \text{ represents } \{(u, v) \mid v \text{ is obtained from } u \text{ by duplicating } a\text{'s}\}.$$

- An expression is conjugate if the relation it represents is conjugate.

Deciding Conjugacy of a Rational Relation

1. Assume that the rational relation is given as a rational expression E over pairs.
2. Convert the rational expression to sum of sumfree expressions.

   (Since every rational expression is equivalent to a sum of sumfree expressions.)
3. Check the conjugacy of each sumfree expression

$$(u_0, v_0){E_1}^*(u_1, v_1)E_2^* \cdots {E_k}^*(u_k, v_k)$$

   by computing a common witness for it.

   (Since the union operation preserves conjugacy, E is conjugate if each of its summands is conjugate.)

Conjugacy of Relations : Results

- The common witnesses of a sumfree expression can be computed inductively in $\mathcal{O}(h \cdot m^2)$ where h is the star height of the expression and m is the length of the expression.

Theorem

Checking if a rational relation given by a rational expression is conjugate is decidable in exponential time.

3. Approximate Problems on Transducers

Joint work with

Prof. Emmanuel Filiot (Université libre de Bruxelles)

Dr. Ismaël Jecker (Université de Franche-Comté)

Dr. Khushraj Madnani (Max Planck Institute for Software Systems)

Recap: (Sub)classes of Transducers

$$\text{sequential functions} \subsetneq \text{rational functions} \subsetneq \text{rational relations}$$

- Rational relations - relations defined by transducers.
- Rational functions - functions defined by transducers.
- Sequential functions - functions defined by input-deterministic transducers.

Class Membership Problems

$$\text{sequential function} \xleftarrow{\text{determinisation}} \text{rational function} \xleftarrow{\text{functionality}} \text{rational relation}$$

Problem	Input	Question
Functionality	rational relation R	Is R a function?
Determinisation	rational function f	Is f sequential?

Class Membership Problems

$$\text{sequential function} \xleftarrow{\mathsf{determinisation}} \text{rational function} \xleftarrow{\mathsf{functionality}} \text{rational relation}$$

Problem	Result
Functionality	P [Choffrut, 1977, Weber and Klemm, 1995]
Determinisation	P [Schützenberger, 1975, Gurari and Ibarra, 1983]

- We study approximate versions of these problems.

Approximate Class Membership Problems

$$\text{sequential function} \xleftarrow{\text{apx. determinisation}} \text{rational function} \xleftarrow{\text{apx. functionality}} \text{rational relation}$$

Problem	Input	Question
Apx. Functionality	rational relation R	Is R close to a function?
Apx. Determinisation	rational function f	Is f close to a sequential function?

Approximate Class Membership Problems

$$\text{sequential function} \xleftarrow{\text{apx. determinisation}} \text{rational function} \xleftarrow{\text{apx. functionality}} \text{rational relation}$$

Problem	Input	Question
Apx. Functionality	rational relation R	$\exists$ rational function f s.t. $d(R, f) < \infty$?
Apx. Determinisation	rational function f	$\exists$ sequential function g s.t. $d(f, g) < \infty$?

Apx Class Membership: Results

Theorem

The approximate functionality problem for rational relations w.r.t. a metrics given in Table 1 are decidable.

Theorem

The approximate determinisation problem for rational functions w.r.t. Levenshtein family of distances are decidable.

Approximate Determinisation: Example

$$\text{sequential function} \xleftarrow{\text{approx. determinisation}} \text{rational function}$$

Examples

- The function $\mathsf{f}_{\mathsf{last}} : \mathsf{u}\sigma \to \sigma\mathsf{u}$ is approx-determinisable w.r.t. Levenshtein.
- The function $\mathsf{f}_{\mathsf{id}} : \mathsf{u}\sigma \to \mathsf{u}\sigma$ is sequential and $d(f_{\mathsf{last}}, f_{\mathsf{id}})$ is finite.

Approximate Determinisation: Example

$$\text{sequential function} \xleftarrow{\text{approx. determinisation}} \text{rational function}$$

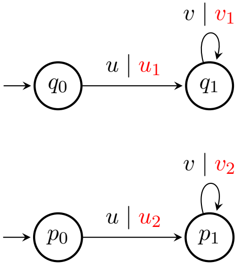
Examples

- The function $\mathsf{f}_{\mathsf{last}} : \mathsf{u}\sigma \to \sigma\mathsf{u}$ is approx-determinisable w.r.t. Levenshtein.
- The function $\mathsf{f}_{\mathsf{id}} : \mathsf{u}\sigma \to \mathsf{u}\sigma$ is sequential and $d(f_{\mathsf{last}}, f_{\mathsf{id}})$ is finite.

- Exact deterministation was characterised using a structural property of a transducer called *twinning property* [Choffrut, 1977].

Approximate Twinning Property (ATP)

- A transducer $\mathcal{T}$ satisfies *approximate* twinning iff for all situations



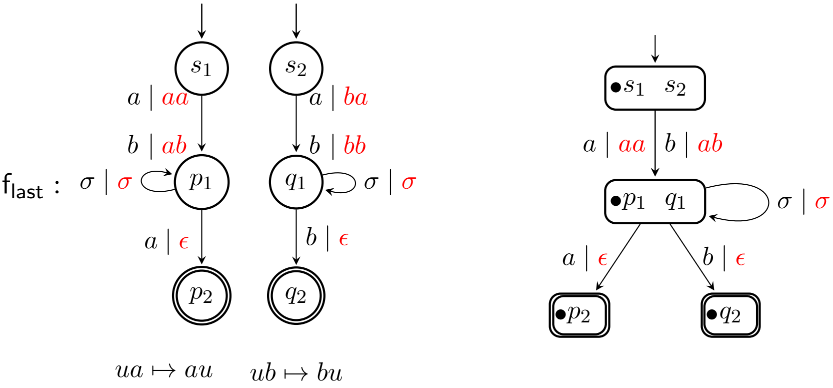
v_1 and v_2 are conjugates

i.e., $\exists$ words x, y s.t. $v_1 = xy$ and $v_2 = yx$

Approximate Determinisation: Characterisation

- **ATP** is sufficient for certain subclassses of rational functions to be approx. determinisable.
  1. union of input-deterministic transducers
  2. "concatenation" of input-deterministic transducers

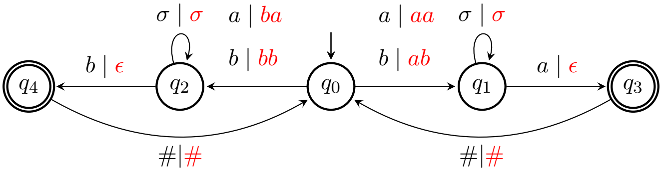
Approximate Determinisation: Construction



- Construction: extend automata subset construction. On any input, choose the output of the transducer with the smallest index.

Approximate Determinisation: Characterisation

- **ATP** is not sufficient for rational functions to be approximately determinisable w.r.t. Levenshtein family of distances.



$\mathsf{f}^*_{\mathsf{last}} : \mathsf{u}_1\# \cdots \mathsf{u}_n\# \mapsto \mathsf{f}_{\mathsf{last}}(\mathsf{u}_1)\# \cdots \mathsf{f}_{\mathsf{last}}(\mathsf{u}_n)\#$ is not approx. determinisable w.r.t. Levenshtein.

Approximate Determinisation: Characterisation

- For rational functions to be approximately determinisable w.r.t. Levenshtein, **ATP** + twinning property must hold within SCCs of the transducer (**STP**).

Lemma

*A rational function given by a transducer $\mathcal{T}$ is approximately determinisable w.r.t. Levenshtein family of distances iff $\mathcal{T}$ satisfies **ATP** and **STP**.*

- Both **ATP** and **STP** are decidable properties for transducers.

Conclusion

Summary I : Comparing Word Transducers

Problem	Input	Question
Distance	rational functions f, g	$d(f, g)$?
Closeness	rational functions f, g	$d(f, g) < \infty$?
k-Closeness	integer k, rational functions f, g	$d(f, g) \leq k$?
Diameter	rational relation R	$\mathsf{dia}_d(R)$?
Bounded diameter	rational relation R	$\mathsf{dia}_d(R) < \infty$?
k-Bounded diameter	integer k, rational relation R	$\mathsf{dia}_d(R) \leq k$?
Index	relation R, S	$\mathsf{Index}(R, S)$?
Bounded index	relation R, S	$\mathsf{Index}(R, S) < \infty$?
k-Bounded index	integer k, relation R, S	$\mathsf{Index}(R, S) \leq k$?

Summary II : Conjugacy of a Rational Relation

Theorem

Checking if a rational relation is conjugate is decidable.

Towards this, we give characterisations for the conjugacy of sets of word pairs.

Summary III : Approximate Problems on Transducers

$$\text{sequential function} \xleftarrow{\text{apx. determinisation}} \text{rational function} \xleftarrow{\text{apx. functionality}} \text{rational relation}$$

Theorem

The approximate determinisation problem for rational functions w.r.t. Levenshtein family of distances are decidable.

Theorem

The approximate functionality problem for rational relations w.r.t. a metrics given in Table 1 are decidable.

Future Work

- Computing edit distance between regular functions (defined by two-way transducers).
- Deciding conjugacy of two-way transducers.
- Deciding "upto distance k" variant of approximate class membership problems on transducers.
- Study the approximate variants of other classical problems on transducers, for instance approximate synthesis/sequential uniformisation.

List of Publications

1. C. Aiswarya, Amaldev Manuel, and Saina Sunny. Edit Distance of Finite State Transducers. In 51st International Colloquium on Automata, Languages, and Programming (ICALP 2024). Volume 297 of LIPIcs, pp. 125:1-125:20.
   doi.org/10.4230/LIPIcs.ICALP.2024.125
2. C.Aiswarya, Amaldev Manuel and Saina Sunny. Deciding Conjugacy of a Rational Relation (Extended Abstract). In 28th International Conference on Developments in Language Theory (DLT 2024). LNCS, Vol 14791, Springer, Cham.
   doi.org/10.1007/978-3-031-66159-4_4
3. Emmanuel Filiot, Ismaël Jecker, Khushraj Madnani and Saina Sunny. Approximate Problems for Finite Transducers. In 52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025). Volume 334 of LIPIcs, pp. 155:1-155:19.
   doi.org/10.4230/LIPIcs.ICALP.2025.155

Thank you :-)

References I

Choffrut, C. (1977).
Une caractérisation des fonctions séquentielles et des fonctions sous-séquentielles en tant que relations rationnelles.
Theoretical Computer Science, 5(3):325–337.

Finkel, O., Halava, V., Harju, T., and Sahla, E. (2023).
On bi-infinite and conjugate post correspondence problems.
RAIRO Theoretical Informatics and Applications, 57:7.

Fischer, P. C. and Rosenberg, A. L. (1968).
Multitape one-way nonwriting automata.
Journal of Computer and System Sciences, 2(1):88–101.

Gurari, E. M. and Ibarra, O. H. (1983).
A note on finite-valued and finitely ambiguous transducers.
Mathematical systems theory, 16(1):61–66.

References II

Lyndon, R. and Schützenberger, M. (1962).
The equation $a^M = b^N c^P$ in a free group.
Michigan Math. J, 9(4):289–298.

Nivat, M. (1968).
Transduction des langages de Chomsky.
PhD thesis, Annales de l'Institut Fourier.

Schützenberger, M. P. (1975).
Sur les relations rationnelles.
In *Automata Theory and Formal Languages, 2nd GI Conference, 1975*, volume 33 of *Lecture Notes in Computer Science*, pages 209–213. Springer.

Weber, A. and Klemm, R. (1995).
Economy of description for single-valued transducers.
Information and Computation, 118(2):327–340.