Approximate comparison of distance automata*

Thomas Colcombet and Laure Daviaud

Université Sorbonne Paris Cité, Cnrs, Liafa

Abstract

Distance automata are automata weighted over the semiring $(\mathbb{N} \cup \{\infty\}, \min, +)$ (the tropical semiring). Such automata compute functions from words to $\mathbb{N} \cup \{\infty\}$ such as the number of occurrences of a given letter. It is known that testing $f \leqslant g$ is an undecidable problem for f, g computed by distance automata. The main contribution of this paper is to show that an approximation of this problem becomes decidable.

We present an algorithm which, given $\varepsilon > 0$ and two functions f, g computed by distance automata, answers "yes" if $f \leqslant (1-\varepsilon)g$, "no" if $f \nleqslant g$, and may answer "yes" or "no" in all other cases. This result highly refines previously known decidability results of the same type.

The core argument behind this quasi-decision procedure is an algorithm which is able to provide an approximated finite presentation to the closure under products of sets of matrices over the tropical semiring.

We also provide another theorem, of affine domination, which shows that previously known decision procedures for cost-automata have an improved precision when used over distance automata.

1998 ACM Subject Classification F.4.3 Formal Languages

Keywords and phrases Distance automata, tropical semiring, decidability, cost functions

Digital Object Identifier 10.4230/LIPIcs.STACS.2013.574

1 Introduction

One way to see language theory, and in particular the theory of regular languages, is as a toolbox of constructions and decision procedures allowing high level handling of languages. These high level operations can then be used as black-boxes in various decision procedures, such as in verification. Since the early times of automata theory, the need for the effective handling of functions rather than sets (as languages) was already apparent. Schützenberger proposed already in the sixties models of finite state machines used for computing functions. These are now known as weighted automata [11] and are the subject of much attention from the research community. In general, weighted automata are non-deterministic automata, weighted over some semiring $(S, \oplus, \otimes)$. The value computed by such an automaton over a given word is then the sum (for $\oplus$) over every run over this word of the product (for $\otimes$) of the weights along the run.

Several instances of this model are very relevant for modelling the behaviour of systems, and henceforth attract much attention. This is in particular the case of probabilistic automata (over the semiring $(\mathbb{R}^+, +, \times)$ with some additional constraints enforcing weights to remain in $[0, 1]$), and distance automata which are automata weighted over the semiring $(\mathbb{N} \cup \{\infty\}, \min, +)$. In such an automaton, each transition is labelled with a non-negative integer (usually 0 or 1), and the weight of a word is the minimum over all possible paths of

* The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°259454.

the sum of the weights. These automata naturally capture some optimisation problems since computing the value amounts to find the path of minimal weight.

The subject of this paper is to develop algorithmic tools for distance automata, and more precisely to develop the question of comparing distance automata. We know from the beginning that exact comparison is beyond reach.

▶ **Theorem 1** (Krob [7]). *The problem to determine, given two functions f, g computed by distance automata, whether $f = g$ or not is undecidable. The problem whether $f \leqslant g$ or not is also undecidable, even if g is deterministic.*

Despite this, some positive results exist but for a comparison relation less precise than inequality, namely *domination*. Given two functions $A^* \to \mathbb{N} \cup \{\infty\}$, *$f$ is dominated* by g (and we note $f \preccurlyeq g$) if there is a function $\alpha : \mathbb{N} \to \mathbb{N}$, extended with $\alpha(\infty) = \infty$, such that

$$f \leqslant \alpha \circ g \ .$$

Moreover, if α is a polynomial, we say that *f is polynomially dominated by g*. The following theorem shows the good properties of the domination relation.

▶ **Theorem 2** ([2] extending results and techniques from [4, 9, 13, 6, 1]). *Given two functions computed by distance automata, domination is decidable. Furthermore, if a function dominates another, then it polynomially dominates it*[1].

The motivation of this work is to improve Theorem 2 and to answer the following question:

> Is it possible to decide "approximations" of the inequality of functions computed by distance automata that are finer than domination ?

We answer positively this question in two ways. We first show:

▶ **Theorem 3** (affine domination). *Given two functions f and g computed by distance automata, if f is dominated by g then f is affinely dominated by g, i.e., $f \leqslant \alpha \circ g$ for some polynomial α of degree 1.*

A consequence of this theorem is that the decision procedure provided by Theorem 2 in fact decides the affine domination, which is finer than the polynomial domination[2].

Our second, and main contribution is an even more accurate decision-like procedure. One says that an algorithm, given two functions f and g and some real $\varepsilon > 0$, *ε-approximates the inequality* if:

- if $f \leqslant (1-\varepsilon)g$, the output is "yes",
- if $f \not\leqslant g$, the output is "no",
- otherwise the output can be either "yes" or "no".

Hence, if such an algorithm answers "yes", one has a guaranty that $f \leqslant g$. Conversely if f is ε-inferior to g (meaning $f \leqslant (1-\varepsilon)g$), one is sure that the algorithm answers "yes". Our second and main result reads as follows:

▶ **Theorem 4** (approximate comparison). *There is an EXPSPACE algorithm which ε-approximates the inequality of functions computed by distance automata.*

[1] Technically, this is not stated in [2] , but can be derived directly from the proofs which explicitly compute the function α using operations preserving polynomials.

[2] Theorem 2 holds for more general classes of automata, cost automata, for which affine domination does not hold. Affine domination is specific to distance automata.

This result is in fact a consequence of a theorem – called the core theorem below – stating that it is possible, given a set of matrices X in the tropical semiring, to approximate (in a suitable way) the set

$$\left\{\frac{1}{\ell}(M_1 \otimes \cdots \otimes M_\ell) \;:\; M_1, \ldots, M_\ell \in X\right\} ,$$

where $\otimes$ denotes the product of matrices. More precisely, the core theorem states that it is possible to approximate the upper envelope of the set of pairs

$$\{(M_1 \otimes \cdots \otimes M_\ell, \ell) \;:\; M_1, \ldots, M_\ell \in X\}$$

for a suitable notion of approximation. This core theorem, Theorem 6, will be described precisely in the first section of this paper.

In Section 2 we present some classical definitions and formally state our core theorem. Section 3 is devoted to the proof of the core theorem. Section 4 applies the core theorem for answering our original motivation, and shows the decidability of the approximate comparison between distance automata. We prove on the way our result of affine domination, Theorem 3. Section 5 concludes the paper.

2 Description of the core theorem

In this first section, we introduce the basic definitions, and define sufficient material for stating our core theorem 6. Its proof is the subject of Section 3 and its application to the comparison of distance automata is the subject of Section 4. We first introduce some classical algebraic definitions in Section 2.1, and finally state our core theorem in Section 2.2.

2.1 Classical definitions

A *semigroup* $(S, \cdot)$ is a set S equipped with an associative binary operation "$\cdot$". If the product has furthermore a neutral element, it is called a *monoid*. The monoid is said *commutative* when $\cdot$ is commutative. An *idempotent* in a monoid is an element e such that $e \cdot e = e$. Given a subset A of a semigroup, $\langle A \rangle$ denotes the closure of A under product, *i.e.,* the least sub-semigroup that contains A. Given two subsets X, Y of a semigroup, $X \cdot Y$ denotes the set $\{a \cdot b \;:\; a \in X,\; b \in Y\}$.

A *semiring* is a set S equipped with two binary operations $\oplus$ and $\otimes$ such that $(S, \oplus)$ is a commutative monoid of neutral element 0, $(S, \otimes)$ is a monoid of neutral element 1, 0 is absorbing for $\otimes$ (i.e., $x \otimes 0 = 0 \otimes x = 0$) and $\otimes$ distributes over $\oplus$. We will consider three semirings: $(\mathbb{R}^+ \cup \{\infty\}, \min, +)$, denoted $\overline{\mathbb{R}^+}$, its restriction to $\mathbb{N} \cup \{\infty\}$, denoted $\overline{\mathbb{N}}$, and its restriction to $\{0, \infty\}$ denoted $\mathbb{B}$. The third, finite semiring is called the *Boolean semiring*, since if we identify 0 with "true" and ∞ with "false", then $\oplus$ is the disjunction and $\otimes$ the conjunction. Remark that in the three cases, the "0" is ∞, and the "1" is 0.

Let S be $\overline{\mathbb{R}^+}$, $\overline{\mathbb{N}}$ or $\mathbb{B}$. The set of matrices with m rows and n columns over S is denoted $\mathcal{M}_{m,n}(S)$. For $M \in \mathcal{M}_{m,n}(S)$, we denote by $\widetilde{M}$ the matrix over $\mathbb{B}$ in which all entries of M different from ∞ are changed into 0. We define the multiplication $A \otimes B$ of two matrices A, B (provided the number n of columns of A equals the number of rows of B) as usual by:

$$(A \otimes B)_{i,j} = \bigoplus_{0<k\leqslant n} (A_{i,k} \otimes B_{k,j}) = \min_{0<k\leqslant n} (A_{i,k} + B_{k,j}) \;.$$

For a positive integer k, we also use the notation $M^k = \underbrace{M \otimes \cdots \otimes M}_{k \text{ times}}$.

For $\lambda \in S$, we denote by λA the matrix such that $(\lambda A)_{i,j} = \lambda A_{i,j}$, with the convention $\lambda\infty = \infty$ (the standard product is used here, not the one of the semiring). We denote also by $B + \lambda$ the matrix such that $(B + \lambda)_{i,j} = B_{i,j} + \lambda$. Finally, we write $A \leqslant B$ if for all i, j, $A_{i,j} \leqslant B_{i,j}$.

2.2 Weighted matrices and the core theorem

In this section we state our core approximation result, Theorem 6. This theorem states that given a set of weighted matrices, it is possible to compute a finite presentation of its closure under product up to some approximation. Hence we have to introduce weighted matrices, the approximation, and what are finite presentations before disclosing the statement. This requires some specific definitions that we present beforehand. We fix now a positive integer n, and all matrices implicitly belong to $\mathcal{M}_{n,n}(\overline{\mathbb{R}^+})$.

As already mentioned in the introduction, our goal is to approximate a set of pairs (M, ℓ) where M is a matrix and ℓ is a positive integer. We call such pairs weighted matrices. A *weighted matrix* is an ordered pair (M, ℓ) where $M \in \mathcal{M}_{n,n}(\overline{\mathbb{R}^+})$ and ℓ is a positive integer. The positive integer ℓ is called the *weight* of the weighted matrix. The set of weighted matrices is denoted by $\mathcal{W}_{n,n}$. Weighted matrices have a semigroup structure $(\mathcal{W}_{n,n}, \otimes)$, where $(M, \ell) \otimes (M', \ell')$ stands for $(M \otimes M', \ell + \ell')$. Given A, B subsets of $\mathcal{W}_{n,n}$, one denotes by $A \otimes B$ the set $\{M \otimes N \ : \ M \in A,\ N \in B\}$, and by $\langle A \rangle$ the closure under $\otimes$ of A. With this terminology, our goal is, given a finite set of weighted matrices X, to approximate $\langle X \rangle$. (Intuitively, if (M, ℓ) is a weighted matrix, M represents the behaviour of a distance automaton that computes a function f, over a word w, while ℓ stands for the lengths of w. So, weighted matrices let us compare $f(w)$ with $|w|$ which is exactly what we want. The operation $\otimes$ between two weighted matrices matches with the concatenation of words, i.e. the product in the tropical semiring for matrices and the sum for lengths.)

We describe now the notion of approximation that we use. Given some $\varepsilon > 0$ and two weighted matrices (M, ℓ) and (M', ℓ'), one writes

$$(M, \ell) \preccurlyeq_\varepsilon (M', \ell') \qquad \text{if} \qquad \ell \geqslant \ell', \widetilde{M} = \widetilde{M'} \text{ and } M \leqslant M' + \varepsilon\ell \ .$$

Remark that in particular, this implies $\frac{1}{\ell}M \leqslant \frac{1}{\ell'}M' + \varepsilon$, which is the intention behind this definition. The definition of $\preccurlyeq_\varepsilon$ is more constraining: this is mandatory for having better properties with respect to the product of matrices, such as in Lemma 5 below. This definition extends to sets of weighted matrices as follows. Given two such sets X, X', $X \preccurlyeq_\varepsilon X'$ if for all $(M, \ell) \in X$, there exists $(M', \ell') \in X'$ such that $(M, \ell) \preccurlyeq_\varepsilon (M', \ell')$. One writes $X \approx_\varepsilon X'$ if $X \preccurlyeq_\varepsilon X'$ and $X' \preccurlyeq_\varepsilon X$ (and says X is ε-equivalent to X').

The following lemma establishes some simple properties of the $\preccurlyeq_\varepsilon$ relations (as a consequence, the same properties hold for $\approx_\varepsilon$).

▶ **Lemma 5.** *Given $X, X', Y, Y', Z \subseteq \mathcal{W}_{n,n}$ and $\varepsilon, \eta > 0$,*
- *if $X \preccurlyeq_\varepsilon Y$ and $Y \preccurlyeq_\eta Z$ then $X \preccurlyeq_{\varepsilon+\eta} Z$,*
- *if $X \preccurlyeq_\varepsilon X'$ and $Y \preccurlyeq_\varepsilon Y'$ then $X \otimes Y \preccurlyeq_\varepsilon X' \otimes Y'$,*
- *if $X \preccurlyeq_\varepsilon X'$ then $\langle X \rangle \preccurlyeq_\varepsilon \langle X' \rangle$.*

Proof. *First item.* If $(M, \ell) \preccurlyeq_\varepsilon (M', \ell') \preccurlyeq_\eta (M'', \ell'')$, then $\ell \geqslant \ell' \geqslant \ell''$, $\widetilde{M} = \widetilde{M'} = \widetilde{M''}$ and $M \leqslant M' + \varepsilon\ell \leqslant M'' + \eta\ell' + \varepsilon\ell \leqslant M'' + (\varepsilon + \eta)\ell$. This easily extends to sets of weighted matrices.

Second item. Assume $(M, \ell) \preccurlyeq_\varepsilon (M', \ell')$ and $(N, t) \preccurlyeq_\varepsilon (N', t')$. Then, $\ell + \ell' \geqslant t + t'$, $\widetilde{M \otimes N} = \widetilde{M' \otimes N'}$ and $M \otimes N \leqslant (M' + \varepsilon\ell) \otimes (N' + \varepsilon t) \leqslant M' \otimes N' + \varepsilon(\ell + t)$. This naturally extends to sets of weighted matrices.

Third item. By induction, applying the second item. ◀

The last ingredient required is to describe how to represent (infinite) sets of weighted matrices. Call a set of weighted matrices $W \subseteq \mathcal{W}_{n,n}$ *finitely presented* if it is a finite union of singleton sets, and of sets of the form $\{(kM, k) : k \geqslant \ell\}$ where $M \in \mathcal{M}_{n,n}(\overline{\mathbb{R}^+})$ and ℓ is a positive integer. Our algorithm manipulates finitely presented sets of weighted matrices.

The core technical contribution of this paper can now be stated, as follows.

▶ **Theorem 6** (core theorem). *Given $X \subseteq \mathcal{W}_{n,n}$ finitely presented and $\varepsilon > 0$, one can compute effectively $Y \subseteq \mathcal{W}_{n,n}$ finitely presented such that:*

$$Y \approx_\varepsilon \langle X \rangle \ .$$

A sketch of the proof of this result will be the subject of Section 3. The application of this theorem to the comparison of distance automata is presented in Section 4. The two sections are independent.

3 Proof of the core theorem

In this section we describe the key arguments involved in the proof of Theorem 6. It is the combination of several arguments. The first one is the use of the forest factorisation theorem of Simon.

3.1 The main induction: the forest factorization theorem of Simon

The forest factorization theorem of Simon [12] is a powerful combinatorial tool for understanding the structure of finite semigroups. In this short abstract, we will not describe the original statement of this theorem, in terms of trees of factorisations, but rather a direct consequence of it which is central in our proof.

▶ **Theorem 7** (equivalent to the forest factorization theorem [12][3]). *Given a semigroup morphism φ from $(S, \otimes)$ (possibly infinite) to a finite semigroup $(T, \cdot)$, and some $X \subseteq S$, set $X_0 = X$ and for all $k \geqslant 0$,*

$$X_{k+1} = X_k \cup X_k \otimes X_k \cup \bigcup_{e \text{ is idempotent } \in T} \langle X_k \cap \varphi^{-1}(e) \rangle \ ,$$

then $\langle X \rangle = X_N$ for $N = 3|T| - 1$.

This proposition teaches us that, for computing the closure under product in the semigroup S, it is sufficient to know how to compute (a) the union of sets, (b) the product of sets, and (c) the restriction of a set to the inverse image of an idempotent by φ, and (d) the closure under product of sets of elements that all have the same idempotent image under φ. Of course, this proposition is interesting when the semigroup T is cleverly chosen.

In our case, we are going to use the above proposition with $(S, \otimes) = \mathcal{W}_{n,n}$, $(T, \cdot) = \mathcal{M}_{n,n}(\mathbb{B})$, and φ the morphism which maps (M, ℓ) to $\widetilde{M}$. Our algorithm will compute, given a finitely presented set of weighted matrices X, an approximation of $\langle X \rangle$ following the same inductive construction as in the forest factorisation theorem. This is justified by the two following lemmas, for which we provide the sketch of a proof.

[3] Modern proofs of this theorem can be found in [8, 3], in particular with the exact bound of $N = 3|T| - 1$ (Simon's original proof only provides $N = 9|T|$).

▶ **Lemma 8.** *For all $\varepsilon > 0$ and all finitely presented sets $X, Y \subseteq \mathcal{W}_{n,n}$ there exists effectively a finitely presented set* $\texttt{product}(\varepsilon, X, Y) \subseteq \mathcal{W}_{n,n}$ *such that*

$$\texttt{product}(\varepsilon, X, Y) \approx_\varepsilon X \otimes Y \ .$$

Given a set X of weighted matrices, let us set $\widetilde{X} = \{\widetilde{M} \mid (M, \ell) \in X\}$.

▶ **Lemma 9.** *For all $\varepsilon > 0$ and all finitely presented set $X \subseteq \mathcal{W}_{n,n}$ such that $\widetilde{X} = \{e\}$ for an idempotent e, there exists effectively a finitely presented set* $\texttt{idempotent}(\varepsilon, X) \subseteq \mathcal{W}_{n,n}$ *such that*

$$\texttt{idempotent}(\varepsilon, X) \approx_\varepsilon \langle X \rangle \ .$$

Assuming that Lemmas 8 and 9 hold, it is easy to provide an algorithm which, given $X \subseteq \mathcal{W}_{n,n}$ finitely presented, computes $X' \subseteq \mathcal{W}_{n,n}$ finitely presented such that $X' \approx_\varepsilon \langle X \rangle$. The principle of the algorithm is to implement Theorem 7, using finitely presented sets that approximate the X_k's.

- Set $Y_0 = X$ and $N = 3(2^{n^2}) - 1 = 3|\mathcal{M}_{n,n}(\mathbb{B})| - 1$.
- For all $0 \leqslant k \leqslant N$, set $\varepsilon(k) = \frac{\varepsilon}{2^{N-k}}$ and

$$Y_{k+1} = Y_k \cup \texttt{product}(\varepsilon(k), Y_k, Y_k) \cup \bigcup_{e \otimes e = e \in \mathcal{M}_{n,n}(\mathbb{B})} \texttt{idempotent}(\varepsilon(k), Y_k \cap \varphi^{-1}(e)) \ .$$

- output Y_N

Correction can be justified as follows: one proves by induction that $Y_k \approx_{\varepsilon(k)} X_k$ for all $k = 0, \ldots, N$ where X_k is defined as in Theorem 7 (with $S = \mathcal{W}_{n,n}$, $T = \mathcal{M}_{n,n}(\mathbb{B})$ and $\varphi(M, \ell) = \widetilde{M}$). For $k = 0$, one has $X_k = X = Y_k$. Let $k \geqslant 0$, suppose that $Y_k \approx_{\varepsilon(k)} X_k$, then by Lemma 8, Lemma 5 and the induction hypothesis,

$$\texttt{product}(\varepsilon(k), Y_k, Y_k) \approx_{\varepsilon(k)} Y_k \otimes Y_k \approx_{\varepsilon(k)} X_k \otimes X_k \ .$$

Finally, by Lemma 5, $\texttt{product}(\varepsilon(k), Y_k, Y_k) \approx_{2\varepsilon(k)} X_k \otimes X_k$. Similarly, by Lemma 9, for all idempotent e, $\texttt{idempotent}(\varepsilon(k), Y_k \cap \varphi^{-1}(e)) \approx_{2\varepsilon(k)} \langle X_k \cap \varphi^{-1}(e)\rangle$. Thus $Y_{k+1} \approx_{\varepsilon(k+1)} X_{k+1}$.

Hence, what remains to be done is to establish Lemmas 8 and 9.

3.2 Approximate products of sets

In this part, we give the main ideas of the proof of Lemma 8. It shows explicit examples of the approximation arguments that are used in a more advanced way for proving Lemma 9.

Proof of Lemma 8. Since the finitely presented sets of weighted matrices are closed under union, it is sufficient to prove Lemma 8 for the atomic blocks of the finite presentation. Namely, it is sufficient to consider the case $X = \{(M, x)\}$ or $X = \{(\ell M, \ell) \mid \ell \geqslant x\}$ together with $Y = \{(N, y)\}$ or $Y = \{(\ell N, \ell) \mid \ell \geqslant y\}$. This results in four possibilities, among which only three remain up to symmetry: (a) $X = \{(M, x)\}$ and $Y = \{(N, y)\}$, (b) $X = \{(M, x)\}$ and $Y = \{(\ell N, \ell) \mid \ell \geqslant y\}$, and finally (c) $X = \{(\ell M, \ell) \mid \ell \geqslant x\}$ and $Y = \{(\ell N, \ell) \mid \ell \geqslant y\}$.

Let us explain the most interesting case, case (c). Let a be the maximum absolute value of a non-infinite entry of M or N. Choose some z such that $2ax \leqslant \varepsilon z$ and $2ay \leqslant \varepsilon z$, and let Z be the set $Z_1 \cup Z_2$ defined by:

$$Z_1 = \{(x'M \otimes y'N, x' + y') \mid x' + y' < z\} \ ,$$

and $$Z_2 = \{(\ell(\lambda M \otimes (1-\lambda)N), \ell) \mid \ell \geqslant z,\ \lambda \in [0,1]\} \ .$$

The set Z_1 is finite, and merely lists all weighted matrices of weight less than z in $X \otimes Y$. The set Z_2 (which is not finitely presented) takes all barycentres of M and N, and produces corresponding weighted matrices for all possible weights greater or equal to z. We need to prove two things. First that $Z \approx_{\frac{\varepsilon}{2}} X \otimes Y$, and second that one can further approximate Z_2 by a finitely presented set $Z_3 \approx_{\frac{\varepsilon}{2}} Z_2$. By Lemma 5 we can then conclude that $X \otimes Y \approx_\varepsilon Z_1 \cup Z_3$, and that $Z_1 \cup Z_3$ is finitely presented and computable from X and Y.

Let us prove that $Z \approx_{\frac{\varepsilon}{2}} X \otimes Y$. Remark first that $X \otimes Y \subseteq Z$. For the converse direction, consider $(W, \ell) \in Z$. Clearly, if $\ell < z$, then $(W, \ell) \in Z_1 \subseteq X \otimes Y$. Otherwise, $W = (\lambda \ell M) \otimes ((1 - \lambda)\ell N)$. It is sufficient for us to find $x' \geqslant x$ and $y' \geqslant y$ such that $x' + y' = \ell$, and $\left|\lambda - \frac{x'}{\ell}\right| \leqslant \frac{\varepsilon}{2a}$: indeed, assuming the existence of such x', y', the matrix $W' = (x'M \otimes y'N, \ell)$ is such that $(W, \ell) \approx_{\frac{\varepsilon}{2}} (W', \ell)$, and furthermore $(W', \ell) \in X \otimes Y$. For proving the existence of such x', y', consider the evolution of the value $\frac{x'}{\ell}$ when x' ranges from x to $\ell - y$. Since $\ell \geqslant z$, $\frac{x}{\ell} \leqslant \frac{\varepsilon}{2a}$, and similarly $\frac{\ell - y}{\ell} \geqslant 1 - \frac{\varepsilon}{2a}$. Furthermore when x' increases of 1, the quantity $\frac{x'}{\ell}$ increases of at most $\frac{1}{z} \leqslant \frac{\varepsilon}{2a}$. As a consequence, $\frac{x'}{\ell}$ gets to be $\frac{\varepsilon}{2a}$-close of any $\lambda \in [0, 1]$ when x' ranges from x to $\ell - y$. Consider x' witnessing this fact and set $y' = \ell - x'$. The pair x', y' satisfies the requirement.

One now needs defining a set $Z_3 \approx_{\frac{\varepsilon}{2}} Z_2$ which is finitely presented. The set Z_3 is defined as the set Z_2, but for the fact that λ is discretized by steps of $\frac{\varepsilon}{4a}$. This can be written as:

$$Z_3 = \bigcup_{\lambda \in \left([0,1] \cap \frac{\varepsilon}{4a}\mathbb{N}\right)} \{(\ell(\lambda M \otimes (1 - \lambda)N), \ell) \mid \ell \geqslant z\} \ .$$

Clearly, this set is finitely presented. It is also simple to prove that $Z_3 \approx_{\frac{\varepsilon}{2}} Z_2$. ◀

4 Comparing distance automata

In this section, we consider the problem of comparing the functions computed by distance automata. In particular, we establish Theorem 3, and we reduce Theorem 4 to our core theorem, Theorem 6. We start by describing distance automata, and their relationship with matrices over the tropical semiring (Section 4.1).

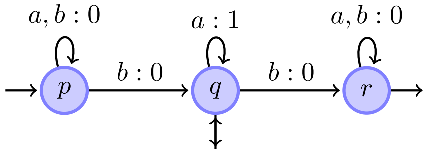
4.1 Distance automata

An *alphabet* is a finite set of symbols. The set of *words* over an alphabet $\mathbb{A}$ is denoted $\mathbb{A}^*$. A *distance automaton* is a tuple $(\mathbb{A}, Q, I, F, T)$, where Q is a finite set of *states* (that we can assume to be $\{1, \ldots, n\}$) where I (resp. T) is a row-vector (*resp.* column-vector) indexed by Q, and F is a morphism from words to $\mathcal{M}_{n,n}(\overline{\mathbb{N}})$. The function f computed by a distance automaton $(\mathbb{A}, Q, I, F, T)$ over an input word u is:

$$\begin{aligned} f : \mathbb{A}^* &\to \overline{\mathbb{N}} \\ u &\mapsto I \otimes F(u) \otimes T \ . \end{aligned}$$

We assume from now on that the initial and final vectors I, T of distance automata only range over $\{0, \infty\}$. The theorems are equally true without this assumption, but this simplifies slightly the proof. In practice the theorems without this restriction can be obtained by simple reductions to this case.

We have defined so far distance automata in terms of matrices. One can see this object in a more "automaton" form as follows. There is a transition labelled (a, x) from state p to state q if $x < \infty$ and $x = F(a)_{p,q}$. A state p is *initial* if $I_{1,p} = 0$. It is *final* if $T_{i,1} = 0$. An example of distance automaton is as follows:



One can redefine the function computed by a distance automaton as follows. A *run* of an automaton over a word $a_1 \dots a_k$ is a sequence $p_0, \dots, p_k$ of states. The *weight of a run* is the sum of the weights of its transitions, *i.e.*, $F(a_1)_{p_0,p_1} + \cdots + F(a_k)_{p_{k-1},p_k}$. Remark that if there is some non-existing transition in this sequence, say from p_{i-1} to p_i, this means that $F(a_i)_{p_{i-1},p_i} = \infty$, and as a consequence the run has an infinite weight. A run is *accepting* if p_0 is initial and p_k is final. One defines the function accepted by the automaton as:

$$
\begin{aligned}
&f : \mathbb{A}^* \to \overline{\mathbb{N}} \\
&\qquad u \mapsto \inf\{\mathrm{weight}(\rho) \ : \ \rho \text{ accepting run over } u\}\ .
\end{aligned}
$$

This definition is equivalent to the matrix version presented above.

For instance, the function computed by the above automaton associates to each word $u = a^{n_0}ba^{n_1} \dots ba^{n_k}$ the value $\min(n_0, \dots, n_k)$.

4.2 Superior limits

In this section, we present Theorem 10 which allows us to compute the superior limit of some infinite set of matrices.

In order to define the superior limit of a set of matrices, a topology is required. The matrices over $\overline{\mathbb{N}}$ are equipped with the following metric. When two matrices are distinct, their distance is $1/n$ where n is the maximal positive integer such that the entries that carry values at most n are the same in both matrices. If no such integer exists, the distance is 1.

Given $X \subseteq \mathcal{M}_{n,n}(\overline{\mathbb{N}})$, a matrix N (which may not be in X) belongs to the *superior limit* of X if:

- N is the limit of some sequence of matrices from X,
- there exists no $M \in X$ such that $M > N$.

Let us denote $\limsup(X)$ the set of matrices in the superior limit of S.

▶ **Theorem 10** (consequence of [5]). *Given a set $X \subseteq \mathcal{M}_{n,n}(\overline{\mathbb{N}})$, the set $\limsup(X)$ is finite. Furthermore, there is a PSPACE algorithm which, given a morphism F from $\mathbb{A}^*$ to $\mathcal{M}_{n,n}(\overline{\mathbb{N}})$, and a non-deterministic automaton for a language $L \subseteq \mathbb{A}^*$, enumerates $\limsup(F(L))$.*

The first part of the statement is a consequence of Higman's lemma. The second part relies on a result of Hashiguchi [5] (improved by Leung and Podolskiy [10]) which implies that the non-infinite entries in the matrices in $\limsup(F(L))$ are at most exponential. This is crucial for representing matrices in polynomial space, and hence exploring the state space in PSPACE.

4.3 A first reduction: the theorem of affine domination

Our goal in this section is to establish the theorem of affine domination (Theorem 3). This will be the opportunity to introduce some notations used in the subsequent section.

Let us fix ourselves two distance automata over the same alphabet $\mathbb{A}$. The first one, $\mathcal{A}_f = (\mathbb{A}, Q_f, F, I_f, T_f)$ calculates a function f. The second one, $\mathcal{A}_g = (\mathbb{A}, Q_g, G, I_g, T_g)$ calculates a function g.

Define $R_{p,0,q} \subseteq \mathbb{A}^*$ to be the set of words over which there is a run of $\mathcal{A}_g$ of weight 0 from state p to state q. Let ℓ be a non-null weight occurring in some transition of $\mathcal{A}_g$, and p, q be states in Q_g. Define $R_{p,\ell,q} \subseteq \mathbb{A}^*$ to contain the words over which there is a run of $\mathcal{A}_g$ from state p to state q which uses one transition of weight ℓ, and otherwise only transitions of weight 0. We will reuse this languages in the next section.

Proof of theorem 3. Let K be the largest number that occurs in one of $\limsup(F(R_{p,\ell,q}))$ for some states p, q and weight of a transition ℓ (such a number exists since by Theorem 10 it is the maximum of finitely many numbers). Given a matrix M, call an *m-expansion* of M a matrix $M' \geqslant M$ such that for all i, j, $M_{i,j} > K$ implies $M'_{i,j} \geqslant m$. We first show a claim concerning expansions.

Claim. For all $M \in F(R_{p,\ell,q})$ and for all m there exists an m-expansion $M' \in F(R_{p,\ell,q})$ of M.
Indeed, by definition of the superior limit, there is some $L \in \limsup(F(R_{p,\ell,q}))$ such that $L \geqslant M$. Furthermore, by choice of K, whenever $M_{i,j} > K$, $L_{i,j} = \infty$. Finally, still by definition of the superior limit, L is the limit of a sequence of matrices in $F(R_{p,\ell,q})$. Hence, for all m, there exists a matrix M' in this sequence which is sufficiently close to L that it is an m-expansion of M. This proves the claim.

Let us turn now to the core of the proof. Our goal is to prove that if f is dominated by g, (*i.e.,* there exists $\alpha : \mathbb{N} \to \mathbb{N}$ extended with $\alpha(\infty) = \infty$ such that $f \leqslant \alpha \circ g$), then $f \leqslant K(1 + g)$. The proof is by contraposition. Thus, assume $f \not\leqslant K(1 + g)$. This means $f(u) > Kg(u) + K$ for some word u. We have to prove that f is not dominated by g.

The first case is $g(u) = 0$. This means that $u \in R_{p,0,q}$ with p initial and q final. Using the above claim, one can choose for all m a word $v^{(m)} \in R_{p,0,q}$ such that $F(v^{(m)})$ is an m-expansion of $F(u)$. Since $f(u) > K$, this means that for all initial state r and all final state s of $\mathcal{A}_f$, $F(u)_{r,s} > K$. This means that for all such r, s, $F(v^{(m)})_{r,s} \geqslant m$. It follows that $f(v^{(m)}) \geqslant m$. Hence over the sequence $(v^{(m)})_m$, g is bounded and f tends to infinity. This forbids the existence of a function α such that $f \leqslant \alpha \circ g$, f is not dominated by g.

Assuming $g(u) \neq 0$, the argument is similar. Remark first that $g(u)$ is finite since $f(u) > Kg(u) + K$. This means one can find $p_0, \ldots, p_k$ with p_0 initial, p_k final, and such that:

$$u = u_1 \ldots u_k, \quad u_1 \in R_{p_0,\ell_1,p_1}, \ldots, u_k \in R_{p_{k-1},\ell_k,p_k} \ ,$$

where $\ell_1, \ldots, \ell_k$ are all non-null and of sum $g(u)$. By the above claim, for all $i = 1 \ldots k$, and all m, one can select $v_i^{(m)}$ in R_{p_{i-1},ℓ_i,p_i} such that $F(v_i^{(m)})$ is an m-expansion of $F(u_i)$. Consider now the word $v^{(m)} = v_1^{(m)} \ldots v_k^{(m)}$. Clearly $g(v^{(m)}) = g(u)$. For the sake of contradiction, assume now that $f(v^{(m)}) < m$ for some m. This means that there exists $q_0, \ldots, q_k$ such that q_0 is initial, q_k is final, and $F(v_i^{(m)})_{q_{i-1},q_i} < m$ for all $i = 1 \ldots k$. Since $F(v_i^{(m)})$ is an m -expansion of $F(u_i)$, this implies $F(u_i)_{q_{i-1},q_i} \leqslant K$. It follows that $f(u) \leqslant Kk \leqslant Kg(u)$. A contradiction. Hence $f(v^{(m)}) \geqslant m$. Thus, g is bounded over $(v^{(m)})_m$ while f is not. As a consequence, f is not dominated by g. ◀

4.4 The reduction construction

We reuse definitions and notations of automata $\mathcal{A}_f$ and $\mathcal{A}_g$ given in the preceding section. In particular, we use the sets $R_{p,\ell,q}$ again.

Our goal is to construct a finite set of weighted matrices X that captures the relationship between f and g. The key ideas behind this reduction are the following. Each matrix (M, ℓ) in X corresponds to a set of runs of g, that start in a given state p and end in a given state

q, and use exactly one transition of non-null weight ℓ. The corresponding matrix M is in charge of (a) simulating the behaviour of F over some word corresponding to such a run (there may be infinitely many such runs, but only the finitely many matrices of the superior limit need to be considered), and (b) keeping information concerning the first and last state of the run of $\mathcal{A}_g$ for being able to check that pieces of run of g are correctly concatenated.

One also needs to define the part of the matrix in charge of controlling the validity of the run of $\mathcal{A}_g$. The construction behind Lemma 11 below is the one of a deterministic automaton, that reads words over the alphabet Q_g^2, and accepts a word $(p_1, q_1)\dots(p_k, q_k)$ if, either p_1 is not initial, or q_k is not final, or if $q_{i-1} \neq p_i$ for some i. One can verify that this language is accepted by a deterministic and complete automaton of states $Q_g \uplus \{i, \bot\}$. The unique initial state is i, and, when reading the word $(p_1, q_1)\dots(p_k, q_k)$, the automaton reaches state $\bot$ if p_1 is not initial or $q_{i-1} \neq p_i$ for some i, otherwise it reaches state q_k. The final states are the ones not in T_g plus $\bot$ plus possibly i if there are no states that are both initial and final in g. Translated in matrix form, this yields Lemma 11.

▶ **Lemma 11.** *There are $(|Q_g| + 2, |Q_g| + 2)$-matrices $(C^{p,q})_{p,q\in Q_g}$ over $\mathbb{B}$ and vectors I_C and T_C such that for all $p_1, q_1, \dots, p_k, q_k \in Q_g$,*

$$I_C \otimes C^{p_1,q_1} \otimes \cdots \otimes C^{p_k,q_k} \otimes T_C = \begin{cases} \infty & \textit{if } p_1 \in I_g,\ q_1 = p_2,\ \dots,\ q_{k-1} = p_k \textit{ and } q_k \in T_g, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. This is implemented in matrix form as follows. For each p, q where $p, q \in Q_g$, set the matrix $C^{p,q}$ that has indices in $Q_g \cup \{i, \bot\}$, to be such that:

$$(C^{p,q})_{p',q'} = \begin{cases} 0 & \text{if } p' = i,\ p \in I_g \text{ and } q' = q, \\ 0 & \text{if } p' = i,\ p \notin I_g \text{ and } q' = \bot, \\ 0 & \text{if } p' = p \text{ and } q' = q, \\ 0 & \text{if } p' \neq i \text{ and } p' \neq p \text{ and } q' = \bot, \\ \infty & \text{otherwise.} \end{cases}$$

Define furthermore I_C be the vector with all entries ∞ but i which is 0, and let T_C be the vector with all entries equal to 0 except T_g and i if there is a state both initial and final in $\mathcal{A}_g$. ◀

We can now construct the set X as follows:

$$X = \left\{ \left(\begin{pmatrix} M & \infty \\ \infty & C^{p,q} \end{pmatrix}, \ell \right) \ :\ M \in \limsup(F(R_{p,\ell,q})) \right\} \tag{1}$$

and the vectors

$$I = (I_f\ I_C) \qquad \text{and} \qquad T = \begin{pmatrix} T_f \\ T_C \end{pmatrix} . \tag{2}$$

The following lemma shows the validity of the construction, and more particularly how it relates the comparison of distance automata to the computation of the closure of a set of weighted matrices.

▶ **Lemma 12.** *For all $\beta > 0$, $f \leqslant \beta g$ if and only if for all $(W, \ell) \in \langle X \rangle$, $I \otimes W \otimes T \leqslant \beta\ell$.*

Proof. Assume first $f \not\leqslant \beta g$, which means $f(u) > \beta g(u)$ for some u. Then clearly, $g(u)$ is finite and hence, there is an accepting run ρ of g over u. This means that one can find $p_0, \ldots, p_k$ with p_0 initial, p_k final, such that:

$$u \in R_{p_0,\ell_1,p_1} R_{p_1,\ell_2,p_2} \ldots R_{p_{k-1},\ell_k,p_k} \ ,$$

where $\ell_1, \ldots, \ell_k$ are all non-null and of sum $\ell = g(u)$. For all $i = 1 \ldots k$, set M_i to be some matrix in $\limsup(F(R_{p_{i-1},\ell_i,p_i}))$ such that $F(u_i) \leqslant M_i$. Let also C_i be C^{p_{i-1},p_i}. Clearly, the weighted matrix

$$(W_i, \ell_i) \qquad \text{with} \quad W_i = \begin{pmatrix} M_i & \infty \\ \infty & C_i \end{pmatrix}$$

belongs to X. Hence (W, ℓ) belongs to $\langle X \rangle$, where $W = W_1 \otimes \cdots \otimes W_k$. We then have $I \otimes W \otimes T = \min(x_f, x_C)$ with

$$x_f = I_f \otimes M_1 \otimes \cdots \otimes M_k \otimes T_f \qquad \text{and} \qquad x_C = I_C \otimes C_1 \otimes \cdots \otimes C_k \otimes T_C \ .$$

By choice of the M_i's, $x_f \geqslant I_f \otimes F(u) \otimes T_f = f(u)$. Furthermore, by Lemma 11, $x_C = \infty$. It follows that $I \otimes W \otimes T \geqslant f(u) > \beta g(u) = \beta \ell$.

Assume now that $f \leqslant \beta g$. Consider some $(W, \ell) \in \langle X \rangle$, it is obtained as $(W, \ell) = (W_1, \ell_1) \otimes \cdots \otimes (W_k, \ell_k)$ with $(W_i, \ell_i) \in X$ for all i. By definition of X, each of the W_i's can be written, for some $p_i, q_i \in Q_g$, as

$$W_i = \begin{pmatrix} M_i & \infty \\ \infty & C^{p_i,q_i} \end{pmatrix} \qquad \text{with} \quad M_i \in \limsup F(R_{p_i,\ell_i,q_i}).$$

Once more, one has $I \otimes W \otimes T = \min(x_f, x_C)$ with

$$x_f = I_f \otimes M_1 \otimes \cdots \otimes M_k \otimes T_f \qquad \text{and} \qquad x_C = I_C \otimes C_1 \otimes \cdots \otimes C_k \otimes T_C \ .$$

Remark first that if $x_C = 0$, clearly, $I \otimes W \otimes T = 0 \leqslant \beta \ell$. Hence, let us assume that $x_C = \infty$. This means by Lemma 11 that p_1 is initial, q_k is final, and $p_i = q_{i-1}$ for all $i = 2 \ldots k$. One needs to prove $x_f \leqslant \beta \ell$.

Assume for the sake of contradiction that $x_f > \beta \ell$. By continuity of the product, and using the definition of the superior limit, there exist words $u_1, \ldots, u_k$ such that for all $i = 1 \ldots k$, $u_i \in R_{p_i,\ell_i,q_i}$, and $I_f \otimes F(u_1) \otimes \cdots \otimes F(u_k) \otimes T_f > \beta \ell$. Furthermore, by definition of the sets R_{p_i,ℓ_i,q_i}, the fact that p_1 is initial, that q_k is final, and that $q_{i-1} = p_i$ for all $i = 2 \ldots k$, it follows that $g(u_1 \ldots u_k) = \ell$. It follows that $f(u_1 \ldots u_k) > \beta g(u_1 \ldots u_k)$. A contradiction. ◀

We are now ready to establish the main theorem of the paper.

Proof of Theorem 4. Let us consider two functions f and g computed by distance automata and some $\varepsilon > 0$. The algorithm works as follows. It computes the set X of weighted matrices as defined in this section (1), as well as the corresponding vectors I, T (2). Using Theorem 6, it computes a finitely presented set Y of weighted matrices such that $Y \approx_{\frac{\varepsilon}{2}} \langle X \rangle$. Then it tests the existence in Y of a weighted matrix (M, ℓ) such that $I \otimes \frac{1}{\ell} M \otimes T > 1 - \frac{\varepsilon}{2}$. This is easy to do for finitely presented sets. If such a weighted matrix exists, the algorithm answers "no". It answers "yes" otherwise. Let us show the correctness of this approach.

- Assume $f \leqslant (1-\varepsilon) g$, and that, for the sake of contradiction, the algorithm answers "no". This means that $I \otimes \frac{1}{\ell} M \otimes T > 1 - \frac{\varepsilon}{2}$ for some weighted matrix $(M, \ell) \in Y$. Furthermore, there exists $(M', \ell') \in \langle X \rangle$ such that $(M, \ell) \preccurlyeq_{\frac{\varepsilon}{2}} (M', \ell')$. This implies $\frac{1}{\ell} M \leqslant \frac{1}{\ell'} M' + \frac{\varepsilon}{2}$. It follows that $I \otimes M' \otimes T > (1-\varepsilon) \ell'$. This contradicts Lemma 12.

- Assume $f \not\preccurlyeq g$, then by Lemma 12, there exists a matrix $M \in \langle X \rangle$ such that $I \otimes \frac{1}{\ell} M \otimes T > 1$. Furthermore, there exists $M' \in Y$ such that $(M, \ell) \preccurlyeq_{\frac{\varepsilon}{2}} (M', \ell')$. This implies $\frac{1}{\ell} M \leqslant \frac{1}{\ell'} M' + \frac{\varepsilon}{2}$, and hence $I \otimes \frac{1}{\ell'} M' \otimes T > 1 - \frac{\varepsilon}{2}$. The algorithm answers "no".

◀

5 Conclusion and further remarks

In this paper, we provided an algorithm for deciding the approximate comparison of distance automata. This algorithm involves the computation of the closure under product of sets of weighted matrices, a result of independent interest.

The main open question is the complexity of the problem. It is clear that the problem is at least PSPACE hard. A correct implementation of the arguments in this paper shows that EXSPACE is an upper bound. We do not know the exact complexity.

Acknowledgments

We would like to thank Jean Mairesse for numerous enlightening discussions as well as the anonymous referees for their helpful comments.

References

1. Mikolaj Bojańczyk and Thomas Colcombet. Bounds in ω-regularity. In *LICS 06*, pages 285–296, 2006.
2. Thomas Colcombet. The theory of stabilisation monoids and regular cost functions. In *Automata, languages and programming. Part II*, volume 5556 of *Lecture Notes in Comput. Sci.*, pages 139–150. Springer, Berlin, 2009.
3. Thomas Colcombet. Green's relations and their use in automata theory. In Adrian Horia Dediu, Shunsuke Inenaga, and Carlos Martín-Vide, editors, *LATA*, volume 6638 of *Lecture Notes in Computer Science*, pages 1–21. Springer, 2011. Invited lecture.
4. Kosaburo Hashiguchi. Limitedness theorem on finite automata with distance functions. *J. Comput. Syst. Sci.*, 24(2):233–244, 1982.
5. Kosaburo Hashiguchi. New upper bounds to the limitedness of distance automata. *Theor. Comput. Sci.*, 233(1–2):19–32, 2000.
6. Daniel Kirsten. Distance desert automata and the star height problem. *RAIRO*, 3(39):455–509, 2005.
7. Daniel Krob. The equality problem for rational series with multiplicities in the tropical semiring is undecidable. *Internat. J. Algebra Comput.*, 4(3):405–425, 1994.
8. Manfred Kufleitner. A proof of the factorization forest theorem. Technical Report Nr. 2007/05, Universität Stuttgart, Germany, 2007.
9. Hing Leung. On the topological structure of a finitely generated semigroup of matrices. *Semigroup Forum*, 37:273–287, 1988.
10. Hing Leung and Viktor Podolskiy. The limitedness problem on distance automata: Hashiguchi's method revisited. *Theoretical Computer Science*, 310(1-3):147–158, 2004.
11. Marcel-Paul Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.
12. Imre Simon. Piecewise testable events. In H. Brackage, editor, *Proc. 2nd GI Conf.*, volume 33, pages 214–222. Springer, 1975.
13. Imre Simon. On semigroups of matrices over the tropical semiring. *RAIRO ITA*, 28(3-4):277–294, 1994.