The equality problem for rational series with multiplicities in the tropical semiring is undecidable

Daniel Krob

The equality problem for rational series with multiplicities in the tropical semiring is undecidable

Daniel KROB

LACIM and CNRS(Institut Blaise Pascal; LITP) [1]

0 Introduction

The tropical semiring is the semiring denoted by $\mathcal{M}$ which has support $\mathbb{N} \cup \{+\infty\}$ and operations $a \oplus b = \min\{a, b\}$ and $a \otimes b = a + b$. It was first introduced in the context of cost minimization in Operations Research. However it appeared that $\mathcal{M}$ plays in fact a central role in several decision problems concerning rational languages (see [15] for a survey of the tropical semiring theory and of its applications). For instance, I. Simon showed that the finite power property for recognizable languages can be reduced to the limitedness problem for the tropical semiring (cf [15]). In the same way, series with multiplicities in the tropical semiring can also be used in order to analyse the non-deterministic behaviour of finite usual automata (cf [17]).

One of the main open questions in the theory of the tropical semiring was to see if it is possible to decide whether two given $\mathcal{M}$-rational series are equal or not (cf [15, 16]). We offer here an answer to this problem since we show in this paper that the equality problem for $\mathcal{M}$-rational series over an alphabet with at least two letters is *undecidable*. One should notice that most people thought that a decision procedure existed (cf [15] for instance) and our result is indeed based on a rather surprising encoding of a 10th Hilbert problem.

It is also interesting to precise the structure of the proof of our undecidability result. Indeed it appears that we use as a main tool the tropical "ring" $\mathcal{Z} = (\mathbb{Z} \cup \{+\infty\}, \min, +)$ which is just the extension of $\mathcal{M}$ to arbitrary integers. The importance of $\mathcal{Z}$ comes from the equivalence with respect to decidability of the equality problems for $\mathcal{M}$ and $\mathcal{Z}$. According to this result, we can reduce our problem to showing that the equality

[1] Mailing adress : Université de Rouen; Faculté des Sciences (Informatique); 76134 Mont Saint-Aignan Cedex - FRANCE

problem for $\mathcal{Z}$-rational series over an alphabet with at least two letters is undecidable. To prove this last result, we show in fact that the decidability of the equality problem for $\mathcal{Z}$ is equivalent to the decidability of the local inequality problem for $\mathcal{Z}$. Using this equivalence and a reduction to a 10th Hilbert problem, we prove then the undecidability of the equality problem for $\mathcal{Z}$-rational series over an alphabet with at least two letters. Hence this allows us to obtain the undecidability of the same problem for $\mathcal{M}$-rational series. Moreover our methods give us also immediately other decidability and undecidability results for connected problems. In particular, we solve also another open question (cf [16]) by showing that the equality problem for rational series over an alphabet with at least two letters and with multiplicities in the semiring $\mathcal{N} = (\mathbb{N} \cup \{-\infty\}, \max, +)$ is undecidable. In the same way, we obtained also as an immediate byproduct of our results, a new proof of a difficult undecidability result of Ibarra (cf [9]).

Let us finally add that a first version of our result appeared in [11]. Unfortunately there was a large technical gap in the corresponding proof. This explains the purpose of this new paper where we give a correct proof, based however on the same idea than in [11].

1 Preliminaries

The *tropical "ring"* is the commutative semiring denoted by $\mathcal{Z}$ which has $\mathbb{Z} \cup \{+\infty\}$ as support, whose addition $\oplus$ is defined by $a \oplus b = \min\{a, b\}$ and whose product $\otimes$ is given by $a \otimes b = a + b$. The operations of $\mathbb{Z}$ are extended to $\mathcal{Z}$ in the usual natural way and the units for $\oplus$ and $\otimes$ are respectively $+\infty$ and 0. The *tropical semiring* is the subsemiring of $\mathcal{Z}$ denoted by $\mathcal{M}$ which has $\mathbb{N} \cup \{+\infty\}$ as support. Let us also introduce the "dual" semiring $\mathcal{N}$ of $\mathcal{M}$ which is the semiring whose support is $\mathbb{N} \cup \{-\infty\}$, whose addition $\oplus$ is given by $a \oplus b = \max\{a, b\}$ and whose product $\otimes$ is defined by $a \otimes b = a + b$. Finally let us consider the subsemiring $\mathcal{Z}^-$ of $\mathcal{Z}$ whose support is $\mathbb{Z}^- \cup \{+\infty\}$. Note that $\mathcal{Z}^-$ is clearly isomorphic to $\mathcal{N}$, an effective isomorphism beeing obtained by the mapping $x \to -x$ from $\mathcal{Z}^-$ into $\mathcal{N}$.

We refer to [2] for all generalities concerning series and rational series with multiplicities in an arbitrary semiring K. We will denote here by $K << A >>$ the K-algebra of series over A with multiplicities in K and by $K\mathrm{Rat}(A)$ the K-algebra of K-rational series. Let us also recall that a K-representation of order n of a free monoid A^* is just a monoid morphism from A^* into the monoid of square matrices of order n with entries in K. Then a K-automaton of order n is a triple (I, μ, T) where μ is a K-representation of order n of A^* and where I and T are respectively a row and a column vector of order n with entries in K (see [2] for more details). It should also be noted that we will use the equivalence between recognizable and rational series with multiplicities in a semiring throughout all this paper without mentioning it explicitely.

Let us now precise some notions concerning K-rational series that we will use in the sequel. First we will denote by $\underline{L}$ the characteristic series of any language $L \subset A^*$ which is the series of $K << A >>$ defined by

$$\forall\ w \in A^*,\ (\underline{L}|w) = \begin{cases} 1_K & \text{if } w \in L \\ 0_K & \text{if } w \notin L \end{cases}$$

Note that $\underline{L}$ is always a K-rational series when L is a rational language (cf [2]). We also denote here as usually by $S \odot T$ the Hadamard product of two series S, T which is the series defined by $(S \odot T|w) = (S|w)(T|w)$ for every $w \in A^*$. We recall that $S \odot T$ is a K-rational series when S and T are K-rational series (see [2] for more details). Finally the constant K-rational series whose every coefficient is equal to k, will be always denoted by k.

Let us also give some denotations concerning words. Thus let w be a word over an alphabet A and let a be a letter of A. Then $|w|$ will denote the length of w and $|w|_a$ will denote the number of a that occur in w.

Let us now recall the following result which is folklore (it is in fact a general property of positive semirings).

PROPOSITION 1.1 : Let S be a rational series of $\mathcal{Z}\mathrm{Rat}(A)$. Then the set

$$\{\ w \in A^*,\ (S|w) = +\infty\ \}$$

is a constructible rational language of $\mathrm{Rat}(A)$.

Proof : Let π be the morphism of semirings from $\mathcal{Z}$ into the boolean semiring $\mathcal{B}$ defined by $\pi(+\infty) = 0$ and $\pi(z) = 1$ for every $z \in \mathcal{Z} - \{+\infty\}$. We also denote by π its natural extension as an algebra morphism from $\mathcal{Z} << A >>$ into $\mathcal{B} << A >>$. Then we have

$$\{\ w \in A^*,\ (S|w) = +\infty\ \} = \{\ w \in A^*,\ (\pi(S)|w) = 0\ \} = A^* - \pi(S)$$

where we identified the $\mathcal{B}$-rational series $\pi(S)$ with its support which is clearly a rational language. Our result follows now since $\pi(S)$ is obviously constructible. ∎

Note : It follows also from proposition 1.1 that it is decidable whether a recognizable series of $\mathcal{Z}\mathrm{Rat}(A)$ is equal to $+\infty$ or whether it has a coefficient equal to $+\infty$.

Finally let us recall the following result of Adler (cf [1, 5]) :

THEOREM 1.2 : Every diophantine equation is equivalent to an equation of the form

$$P(x_1, \ldots, x_n) = 1$$

where P is an homogeneous polynomial of degree 4 of some $\mathbb{Z}$-algebra $\mathbb{Z}[x_1, \ldots, x_n]$.

It is easy to deduce from Adler's theorem the following undecidability result :

COROLLARY 1.3 : It is undecidable to decide whether there exists a n-uple of *strictly positive* integers $(x_1, \ldots, x_n)$ in $(\mathbb{N} - \{0\})^n$ such that $P(x_1, \ldots, x_n) = 1$ where P is an homogeneous polynomial of degree 4 of $\mathbb{Z}[x_1, \ldots, x_n]$.

Proof : It follows clearly from the undecidability of Hilbert's 10th problem (cf [5]) and from theorem 1.2 that it is undecidable to decide whether there exists a n-uple $(x_1, \ldots, x_n)$ of non-negative integers in $\mathbb{N}$ such that $P(x_1, \ldots, x_n) = 1$ when P is an homogeneous polynomial of degree 4 of $\mathbb{Z}[x_1, \ldots, x_n]$. Hence the problem of deciding whether there exists a n-uple $(x_1, \ldots, x_n)$ of integers in $\mathbb{Z}$ such that $P(x_1, \ldots, x_n) = 1$ is also undecidable under the same hypotheses.

Let now P be an homogeneous polynomial of degree 4 in $\mathbb{Z}[x_1, \ldots, x_n]$. Let us then introduce new variables $(y_{i,1}, y_{i,2})_{i=1,n}$ and let us consider the polynomial Q defined by

$$Q(y_{1,1}, y_{1,2}, \ldots, y_{n,1}, y_{n,2}) = P(y_{1,1} - y_{1,2}, \ldots, y_{n,1} - y_{n,2}) \ .$$

Q is clearly an homogeneous polynomial of degree 4. Moreover it is easy to see that every integer $x \in \mathbb{Z}$ can be written as the difference $y - z$ of two strictly positive integers. It follows immediately from this observation that the equation $P(x_1, \ldots, x_n) = 1$ has a solution in $\mathbb{Z}^n$ if and only if the equation $Q(y_{1,1}, \ldots, y_{n,2}) = 1$ has a solution in strictly positive integers.

Our corollary follows now immediately from this result and from above remarks. ∎

2 Some relations between decidability problems

Let K be a totally ordered semiring. We can extend the order of K to the K-algebra $K << A >>$ of series with multiplicities in K by defining

$$P \leq Q \quad \text{iff} \quad \forall\ w \in A^*,\ (P|w) \leq (Q|w)$$

for every series $P, Q \in K << A >>$. Let us then consider the four problems of equality, inequality, local inequality and local equality for K-rational series over A :

$$P, Q \in \mathrm{KRat}(A), \quad P = Q \ ? \qquad (Eq)$$

$$P, Q \in \mathrm{KRat}(A), \quad P \leq Q \ ? \qquad (Ineq)$$

$$P, Q \in \mathrm{KRat}(A), \quad \exists\ w \in A^*,\ (P|w) \leq (Q|w) \ ? \qquad (LocalIneq)$$

$$P, Q \in \mathrm{KRat}(A), \quad \exists\ w \in A^*,\ (P|w) = (Q|w) \ ? \qquad (LocalEq)$$

In general, these problems are not connected. [2] However it appears that the three first above problems are equivalent with respect to decidability when K is the tropical "ring" or semiring equiped with the total order induced by the usual order of $\mathbb{Z}$. [3]

PROPOSITION 2.1 : Let $K = \mathcal{Z}$ or $K = \mathcal{M}$. Then the three following assertions that deal with decidability problems for K-rational series, are equivalent :

1. The equality problem (Eq) is decidable.
2. The inequality problem $(Ineq)$ is decidable.
3. The local inequality problem $(LocalIneq)$ is decidable.

Proof : The fact that assertion 2 implies assertion 1 is immediate since we have

$$P = Q \iff \begin{cases} P \leq Q \\ Q \leq P \end{cases}$$

The fact that assertion 1 implies assertion 2 follows also immediately from the relation

[2] For instance, when $K = \mathbb{N}$, the equality problem is decidable and the inequality problem is undecidable (see [6]).

[3] Note that $\mathbb{N}$ or $\mathbb{Z}$ can also be equiped with the opposite total order which corresponds to the natural order in the sense of the theory of ordered semigroups. We will not use this natural order here since it is clearly equivalent to the previous one in our context of decidability questions.

$$P \leq Q \iff P = P \oplus Q = \min(P, Q)$$

Let us now show that assertion 3 implies assertion 2. Hence let P, Q be two K-rational series where $K = \mathcal{M}$ or $K = \mathcal{Z}$. Then the set $I = \{\ w \in A^*,\ (Q|w) = +\infty\ \}$ is rational and constructible according to proposition 1.1. Let us now consider the series $\overline{P}$ defined by

$$\overline{P} = (P \odot \underline{A^* - I}) \oplus \underline{I} = \begin{cases} 0 & \text{if } w \in I \\ (P|w) & \text{if } w \notin I \end{cases}$$

which is clearly K-rational. We can define in the same way the series $\overline{Q}$. Then we have

$$\begin{aligned} P \leq Q &\iff \overline{P} \leq \overline{Q} \\ &\iff \forall\ w \in A^*,\ (\overline{P}|w) \leq (\overline{Q}|w) \\ &\iff \forall\ w \in A^*,\ (\overline{P}|w) < (\overline{Q}|w) + 1 \end{aligned}$$

this last equivalence coming from the fact that $\overline{Q}$ has no value equal to $+\infty$. It follows immediately from these relations that we have

$$P \leq Q \iff \neg(\exists\ w \in A^*,\ (\overline{P}|w) \geq (\overline{Q} \odot 1|w))$$

Hence it follows clearly from this last equivalence that assertion 3 implies assertion 2.

Let us now show that assertion 2 implies assertion 3. Thus let P, Q be two K-rational series. According to proposition 1.1, the set $I = \{\ w \in A^*,\ (Q|w) = +\infty\ \}$ is an effective rational language. Hence we can decide whether I is empty or not. If I is non-empty, the local inequality problem has obviously a positive answer since every word $w \in I$ satisfies to $(LocalIneq)$. On the other hand, if I is empty, we have

$$\begin{aligned} (LocalIneq) &\iff \neg(\forall\ w \in A^*,\ (P|w) > (Q|w)) \\ &\iff \neg(\forall\ w \in A^*,\ (P|w) \geq (Q|w) + 1) \\ &\iff \neg(P \geq Q \odot 1) \end{aligned}$$

Note that the second above equivalence follows from the fact that $I = \emptyset$. Hence it follows immediately from these last equivalences that assertion 2 implies assertion 3. This ends the proof of our proposition. ■

PROPOSITION 2.2 : Let $K = \mathcal{Z}$ or $K = \mathcal{M}$. Then the decidability of the local equality problem $(LocalEq)$ for K-rational series implies the decidability of the local inequality problem $(LocalIneq)$ for K-rational series.

Proof : It is immediate since we clearly have

$$\exists\ w \in A^*,\ (P|w) \leq (Q|w) \iff \exists\ w \in A^*,\ (P|w) = (P \oplus Q|w) = \min((P|w), (Q|w))$$

Hence our proposition is proved. ■

3 Undecidability of the equality problem for $\mathcal{Z}$

This section is devoted to the proof of the undecidability of the equality problem for $\mathcal{Z}$-rational series over alphabets with at least two letters. This result implies in fact the undecidability of the same problem for $\mathcal{M}$-rational series as we will see later.

THEOREM 3.1 : Let A be any alphabet with at least 2 letters. Then the equality problem is undecidable for $\mathcal{Z}$-rational series over A.

Proof : We should first notice that the decidability of the equality problem for K-rational series on an arbitrary semiring K and over a k-letter alphabet A_k is equivalent to the decidability of the same problem for a l-letter alphabet A_l when $k, l \geq 2$.

Indeed, it suffices to use an adapted encoding of A_k^* over A_l^* in order to prove this result. For instance, let a, b be two distinct letters of A_l and let σ be the monoid morphism from A_k^* into A_l^* defined by $\sigma(a_i) = a^i b$ for every $a_i \in A_k = \{a_1, \ldots, a_k\}$. Then we can also denote by σ its extension as a K-algebra morphism from $K << A_k >>$ into $K << A_l >>$. It is easy to see that σ is injective and preserves rationality. Hence deciding whether two K-rational series E, F of $K << A_k >>$ are equal, is equivalent to deciding whether the two K-rational series $\sigma(E), \sigma(F)$ of $K << A_l >>$ are equal. This proves our claim.

Our undecidability proof is based on a reduction to Adler's restriction of Hilbert's 10th problem (see theorem 1.2 and corollary 1.3). [4] Let now $P(x)$ be an homogeneous polynomial of degree 4 in several indeterminates of $\mathbb{Z}[x]$ [5]. By distinguishing all variables, it is easily seen that the equation $P(x) = 1$ where the variables involved in x belong to $\mathbb{N} - \{0\}$, can be transformed in an equivalent way as a system of the form

$$\left\{ \begin{array}{l} \displaystyle\sum_{i=1}^{p} p_i \, x_1^{(i)} x_2^{(i)} x_3^{(i)} x_4^{(i)} = 1 \\ \forall \, (i,j,k,l) \in V, \; x_k^{(i)} = x_l^{(j)} \end{array} \right.$$

where V is some subset of $[1,p] \times [1,p] \times [1,4] \times [1,4]$, where $(p_i)_{i=1,\ldots,p}$ is a family of integers of $\mathbb{Z}$ and where all the variables $x_k^{(i)}$ take their values in $\mathbb{N} - \{0\}$. Moreover, introducing new variables $(y_1^{(i)}, y_2^{(i)}, y_3^{(i)})_{i=1,p}$ in order to encode partial products, it is easy to transform the previous system into the equivalent one

$$\left\{ \begin{array}{l} \displaystyle\sum_{i=1}^{p} p_i \, y_3^{(i)} = 1 \\ \forall \, (i,j,k,l) \in V, \; x_k^{(i)} = x_l^{(j)} \end{array} \right. \qquad \left\{ \begin{array}{l} \forall \, i \in [1,p], \; y_1^{(i)} = x_1^{(i)} x_2^{(i)} \\ \forall \, i \in [1,p], \; y_2^{(i)} = x_3^{(i)} x_4^{(i)} \\ \forall \, i \in [1,p], \; y_3^{(i)} = y_1^{(i)} y_2^{(i)} \end{array} \right.$$

where all considered variables have values in $\mathbb{N} - \{0\}$. But this last system can also be clearly transformed into the single equation

$$\begin{array}{l} - \, | \displaystyle\sum_{i=1}^{p} p_i \, y_3^{(i)} - 1 | - \sum_{i=1}^{p} | y_1^{(i)} - x_1^{(i)} x_2^{(i)} | - \sum_{i=1}^{p} | y_2^{(i)} - x_3^{(i)} x_4^{(i)} | \\ - \displaystyle\sum_{i=1}^{p} | y_3^{(i)} - y_1^{(i)} y_2^{(i)} | - \sum_{(i,j,k,l) \in V} | x_k^{(i)} - x_l^{(j)} | \; = \; 0 \qquad (HD) \end{array}$$

where all considered variables belong always to $\mathbb{N} - \{0\}$. Hence, according to corollary 1.3, it follows from our reduction process that it is undecidable to see whether an equation of the form (HD) has a solution in strictly positive integers.

[4] Note that the use of Adler's reduction of the 10th Hilbert problem is not essential in our proof. In fact, we introduced it here only for reducing the complexity of our encoding.

[5] Here x denotes of course a vector of variables $x = (x_1, \ldots, x_n)$.

Let now $A = \{a, b, c, d\}$ be a four-letter alphabet. According to proposition 2.1 and to our first remark, it suffices to show that the local inequality problem for $\mathcal{Z}$-rational series over A is undecidable in order to prove our theorem. We will show this fact by a suitable encoding of equation (HD) in terms of $\mathcal{Z}$-rational series. But let us now give some lemmas that will allow us to construct this encoding.

LEMMA 3.2 : Let $\epsilon \in \{-1, +1\}$. Then the series $Var1(\epsilon)$ defined by

$$(Var1(\epsilon)|w) = \begin{cases} \epsilon\, n_1 & \text{when } w = ab^{n_1}c \ldots ab^{n_k}c \in (ab^+c)^+ \\ 0 & \text{when } w \notin (ab^+c)^+ \end{cases}$$

is a $\mathcal{Z}$-rational series of $\mathcal{Z}\mathrm{Rat}(a, b, c)$.

Proof : Let us consider the $\mathcal{Z}$-representation μ of $\{a, b, c\}^*$ which is defined by

$$\mu(a) = \begin{pmatrix} 0 & +\infty \\ +\infty & 0 \end{pmatrix}, \quad \mu(b) = \begin{pmatrix} \epsilon & +\infty \\ +\infty & 0 \end{pmatrix} \quad \text{and} \quad \mu(c) = \begin{pmatrix} +\infty & 0 \\ +\infty & 0 \end{pmatrix}.$$

A simple computation allows us to see that we have

$$\mu(ab^{n_1}c\,ab^{n_2}c \ldots ab^{n_k}c) = \begin{pmatrix} +\infty & n_1 \\ +\infty & 0 \end{pmatrix}$$

for every non-empty vector $(n_i)_{i\in[1,k]} \in \mathbb{N}^k$. Let us now denote by $S1$ the $\mathcal{Z}$-rational series defined by $(S1|w) = \mu_{1,2}(w)$ for every $w \in \{a, b, c\}^*$. It is then easy to see that $Var1(\epsilon) = (S1 \odot \underline{L}) \oplus \underline{\{a, b, c\}^* - L}$ where L denotes the rational language $(ab^+c)^+$. Our lemma follows now immediately from this last formula. ∎

LEMMA 3.3 : Let $\epsilon \in \{-1, +1\}$. Then the series $VarA(\epsilon)$ defined by

$$(VarA(\epsilon)|w) = \begin{cases} \epsilon\,|w|_a = \epsilon\, k & \text{when } w = ab^{n_1}c \ldots ab^{n_k}c \in (ab^+c)^+ \\ 0 & \text{when } w \notin (ab^+c)^+ \end{cases}$$

is a $\mathcal{Z}$-rational series of $\mathcal{Z}\mathrm{Rat}(a, b, c)$.

Proof : Using the $\mathcal{Z}$-representation μ of order 1 of $\{a, b, c\}^*$ defined by

$$\mu(a) = (\,\epsilon\,) \quad \text{and} \quad \mu(b) = \mu(c) = (\,0\,),$$

it is not difficult to see that the series $S_a(\epsilon)$ defined by $(S_a(\epsilon)|w) = \epsilon|w|_a$ for every $w \in \{a, b, c\}^*$ is $\mathcal{Z}$-rational. But it is easily shown that $VarA(\epsilon) = (S_a(\epsilon) \odot \underline{L}) \oplus \underline{\{a, b, c\}^* - L}$ where L denotes the rational language $(ab^+c)^+$. The lemma follows now clearly. ∎

LEMMA 3.4 : Let $\epsilon \in \{-1, +1\}$. Then the series $VarB(\epsilon)$ defined by

$$(VarB(\epsilon)|w) = \begin{cases} \epsilon\,|w|_b = \epsilon\,(\sum_{i=1}^{k} n_i) & \text{when } w = ab^{n_1}c \ldots ab^{n_k}c \in (ab^+c)^+ \\ 0 & \text{when } w \notin (ab^+c)^+ \end{cases}$$

is a $\mathcal{Z}$-rational series of $\mathcal{Z}\mathrm{Rat}(a, b, c)$.

Proof : Let us consider the $\mathcal{Z}$-representation μ of $\{a, b, c\}^*$ defined by

$$\mu(a) = \mu(c) = (\,0\,) \qquad \text{and} \qquad \mu(b) = (\,\epsilon\,).$$

Using this representation, it is easy to show that the series

$$S_b(\epsilon) = \sum_{w \in \{a,b,c\}^*} \epsilon\,|w|_b\ w$$

is $\mathcal{Z}$-rational. But it is easily seen that $VarB(\epsilon) = (\underline{L} \odot S_b(\epsilon)) \oplus \underline{\{a,b,c\}^* - L}$ where L denotes the rational language $(ab^+c)^+$. Our lemma follows now immediately from this last formula. ■

LEMMA 3.5 : Let $k \in \mathbb{Z}$. Then the series $VarY(\alpha)$ defined by

$$(VarY(\alpha)|w) = \begin{cases} \alpha\, n & \text{when } w = a^n \in a^+ \\ 0 & \text{when } w \notin a^+ \end{cases}$$

is a $\mathcal{Z}$-rational series of $\mathcal{Z}\mathrm{Rat}(a,b,c)$.

Proof : Using the $\mathcal{Z}$-representation μ of order 1 of $\{a,b,c\}^*$ defined by

$$\mu(a) = (\,\alpha\,) \qquad \text{and} \qquad \mu(b) = \mu(c) = (\,0\,),$$

it can easily be shown that the series

$$S_a(\alpha) = \sum_{w \in \{a,b,c\}^*} \alpha\,|w|_a\ w$$

is $\mathcal{Z}$-rational. Let now consider the one-letter rational language $L = a^+$. An immediate computation shows then that $VarY(\alpha) = (\underline{L} \odot S_a(\alpha)) \oplus \underline{\{a,b,c\}^* - L}$ which is hence a $\mathcal{Z}$-rational series. ■

LEMMA 3.6 : Let $\epsilon \in \{-1,+1\}$. Then the series defined by

$$(MinMax(\epsilon)|w) = \begin{cases} \min{}_{i\in[1,k]}\,(\epsilon\, n_i) & \text{when } w = a\,b^{n_1}\,c\ \ldots\ a\,b^{n_k}\,c\ \in (ab^+c)^+ \\ 0 & \text{when } w \notin (ab^+c)^+ \end{cases}$$

is a $\mathcal{Z}$-rational series of $\mathcal{Z}\mathrm{Rat}(a,b,c)$.

Proof : Let us consider the $\mathcal{Z}$-representation μ of $\{a,b,c\}^*$ defined by

$$\mu(a) = \begin{pmatrix} 0 & 0 & +\infty \\ +\infty & +\infty & +\infty \\ +\infty & +\infty & 0 \end{pmatrix}, \quad \mu(b) = \begin{pmatrix} 0 & +\infty & +\infty \\ +\infty & \epsilon & +\infty \\ +\infty & +\infty & 0 \end{pmatrix}$$

and by

$$\mu(c) = \begin{pmatrix} 0 & +\infty & +\infty \\ +\infty & +\infty & 0 \\ +\infty & +\infty & 0 \end{pmatrix}.$$

An easy computation allows to show that we have

$$\mu(ab^{n_1}c\,ab^{n_2}c\ \ldots\ ab^{n_k}c) = \begin{pmatrix} 0 & +\infty & \oplus_{i=1}^{k}\,\epsilon\, n_i \\ +\infty & +\infty & +\infty \\ +\infty & +\infty & 0 \end{pmatrix}$$

for every non-empty family $(n_i)_{i\in[1,k]}$ of positive integers. Let us now denote by S_ϵ the $\mathcal{Z}$-rational series which is equal to

$\mu_{1,3}(w)$ for every $w \in \{a,b,c\}^*$. It is then easy to see that $MinMax(\epsilon) = (S_\epsilon \odot \underline{L}) \oplus \underline{\{a,b,c\}^* - L}$ where L denotes the rational language $(ab^+c)^+$. Our lemma follows now immediately. ■

LEMMA 3.7 : Let M be a square matrix of order n and let $p \in \mathbb{N} - \{0\}$. Let us then denote by $E_{n,p}$ and $\mathcal{N}(M,p)$ the square matrices of order $n\,p$ defined by

$$E_{n,p} = \begin{array}{c} \\ n \\ n \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cccc} n & \ldots & n & n \end{array} \\ \begin{pmatrix} +\infty & \ldots & +\infty & Id_n \\ Id_n & \ldots & +\infty & +\infty \\ \vdots & \ddots & \vdots & \vdots \\ +\infty & \ldots & Id_n & +\infty \end{pmatrix} \end{array} \quad \text{and} \quad \mathcal{N}(M,p) = \begin{array}{c} \\ n \\ n \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cccc} n & n & \ldots & n \end{array} \\ \begin{pmatrix} M & +\infty & \ldots & +\infty \\ +\infty & Id_n & \ldots & +\infty \\ \vdots & \vdots & \ddots & \vdots \\ +\infty & +\infty & \ldots & Id_n \end{pmatrix} \end{array}$$

Let now $(N_i)_{i=1,\ldots,p}$ be a family of square matrices of order n. Then we have

$$\mathcal{N}(N_1,p)\,E_{n,p}\,\mathcal{N}(N_2,p)\,E_{n,p}\ \ldots\ \mathcal{N}(N_N,p)\,E_{n,p} \ = \ \begin{array}{c} \\ n \\ n \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cccc} n & n & \ldots & n \end{array} \\ \begin{pmatrix} N_1 & +\infty & \ldots & +\infty \\ +\infty & N_2 & \ldots & +\infty \\ \vdots & \vdots & \ddots & \vdots \\ +\infty & +\infty & \ldots & N_p \end{pmatrix} \end{array}$$

Proof : It is an easy verification that we leave to the reader. ■

Let us now consider an equation of the form (HD). Note first that we will use in the sequel of this proof all the notations used in the definition of this equation. Let us then introduce the rational language C over the alphabet $A = \{a,b,c,d\}$ defined by

$$C = (\ (ab^+c)^+\,d\,(ab^+c)^+\,d\,(ab^+c)^+\,d\,a^+d\ )^p = \prod_{i=1}^{p}\ (\ (ab^+c)^+\,d\,(ab^+c)^+\,d\,(ab^+c)^+\,d\,a^+d\ )$$

In the sequel, a generic word of C will be denoted by

$$\begin{aligned} \underline{w} = w_1^{(1)}(\underline{n(1,1)})\,d\,w_2^{(1)}(\underline{n(1,2)})\,d\,w_3^{(1)}(\underline{n(1,3)})\,d\,a^{y_3^{(1)}}\,d\ \ldots \\ \ldots\ w_1^{(p)}(\underline{n(p,1)})\,d\,w_2^{(p)}(\underline{n(p,2)})\,d\,w_3^{(p)}(\underline{n(p,3)})\,d\,a^{y_3^{(p)}}\,d \end{aligned}$$

where we have for every $i \in [1,p]$ and every $j \in \{1,2,3\}$

$$w_j^{(i)}(\underline{n(i,j)}) = ab^{n(i,j)_1}c\,ab^{n(i,j)_2}c\ \ldots\ ab^{n(i,j)_k}c$$

with $\underline{n(i,j)} = (n(i,j)_l)_{l\in[1,k]} \in (\mathbb{N}-\{0\})^k$. We will also especially be interested in the words of C which have the form

$$\begin{aligned} w(\underline{xy}) = (ab^{x_1^{(1)}}c)^{x_2^{(1)}}\,d\,(ab^{x_3^{(1)}}c)^{x_4^{(1)}}\,d\,(ab^{y_1^{(1)}}c)^{y_2^{(1)}}\,d\,a^{y_3^{(1)}}\,d\ \ldots \\ \ldots\ (ab^{x_1^{(p)}}c)^{x_2^{(p)}}\,d\,(ab^{x_3^{(p)}}c)^{x_4^{(p)}}\,d\,(ab^{y_1^{(p)}}c)^{y_2^{(p)}}\,d\,a^{y_3^{(p)}}\,d \end{aligned}$$

where $\underline{xy}$ denotes the vector $(x_1^{(i)},x_2^{(i)},x_3^{(i)},x_4^{(i)},y_1^{(i)},y_2^{(i)},y_3^{(i)})_{i=1,p}$ of $(\mathbb{N}-\{0\})^{7p}$, since this will be our basic encoding of the variables involved in equation (HD).

Let now (J,μ,F) be a $\mathcal{Z}$-automaton of order n that recognizes the series $Var1(-1)$ of lemma 3.2. We can then define a $\mathcal{Z}$-representation ν of A^* of order np as follows

$$\forall\ \alpha \in \{a,b,c\},\ \nu(\alpha) = \mathcal{N}(\mu(\alpha),p) \quad \text{and} \quad \nu(d) = E_{n,p}$$

where we took the notations of lemma 3.7. Then according to lemma 3.2 and to lemma 3.7, it is easy to see that we have

$$\begin{array}{c} {\scriptstyle I}\downarrow \\ \left(\ldots\ +\infty\ \ J\ \ +\infty\ \ \ldots \right) \end{array} \nu(w) \begin{pmatrix} \vdots \\ +\infty \\ F \\ +\infty \\ \vdots \end{pmatrix} \overset{I}{\leftarrow} = \begin{cases} -n(i,3)_1 & \text{when } w = \underline{w} \in C \\ 0 & \text{when } w \notin C \end{cases}$$

for every $i \in [1,p]$, where each symbol $+\infty$ denotes in fact a block of order n and where we set $I = (i-1)p+3$. Hence it follows that the series $SumP_{1,1}^{(i)}$ defined by $(SumP_{1,1}^{(i)}|w) = 0$ when $w \notin C$ and by $(SumP_{1,1}^{(i)}|w) = -n(i,3)_1$ when $w = \underline{w} \in C$ is a $\mathcal{Z}$-rational series for every $i \in [1,p]$. Arguing as above but using now lemma 3.4, it is also easy to see that the series $SumP_{1,2}^{(i)}$ defined by $(SumP_{1,2}^{(i)}|w) = 0$ when $w \notin C$ and by

$$(SumP_{1,2}^{(i)}|w) = |w_1^{(i)}(\underline{n(i,1)})|_b$$

when $w = \underline{w} \in C$ is a $\mathcal{Z}$-rational series for every $i \in [1,p]$. It follows that the series $SumP_1^{(i)} = SumP_{1,1}^{(i)} \odot SumP_{1,2}^{(i)}$ is a $\mathcal{Z}$-rational series for every $i \in [1,p]$ and we clearly have

$$(SumP_1^{(i)}|w) = \begin{cases} |w_1^{(i)}(\underline{n(i,1)})|_b - n(i,3)_1 & \text{when } w = \underline{w} \in C \\ 0 & \text{when } w \notin C \end{cases}$$

Using the same method, we can also construct for every $i \in [1,p]$ a $\mathcal{Z}$-rational series $SumN_1^{(i)}$ such that

$$(SumN_1^{(i)}|w) = \begin{cases} -|w_1^{(i)}(\underline{n(i,1)})|_b + n(i,3)_1 & \text{when } w = \underline{w} \in C \\ 0 & \text{when } w \notin C \end{cases}$$

It follows that the series $Sum_1^{(i)} = SumN_1^{(i)} \oplus SumP_1^{(i)}$ is $\mathcal{Z}$-rational for every $i \in [1,p]$ and we clearly have

$$(Sum_1^{(i)}|w) = \begin{cases} -\ |\ |w_1^{(i)}(\underline{n(i,1)})|_b - n(i,3)_1\ | & \text{when } w = \underline{w} \in C \\ 0 & \text{when } w \notin C \end{cases}$$

Therefore the series $Sum_1 = Sum_1^{(1)} \odot \ldots \odot Sum_1^{(p)}$ is $\mathcal{Z}$-rational and we have

$$(Sum_1|w) = -\sum_{i=1}^{p} |\ |w_1^{(i)}(\underline{n(i,1)})|_b - n(i,3)_1\ |$$

when $w = \underline{w} \in C$ and $(Sum_1|w) = 0$ when $w \notin C$.

Using now lemmas 3.3, 3.4, 3.5 and 3.7, the same method allows us to show that the series Sum_2 and Sum_3 defined respectively by

$$(Sum_2|w) = -\sum_{i=1}^{p} |\ |w_2^{(i)}(\underline{n(i,2)})|_b - |w_3^{(i)}(\underline{n(i,3)})|_a\ |$$

and

$$(Sum_3|w) = - \sum_{i=1}^{p} \mid |w_3^{(i)}(\underline{n(i,3)})|_b - y_3^{(i)} \mid$$

when $w = \underline{w} \in C$ and by $(Sum_2|w) = (Sum_3|w) = 0$ when $w \notin C$ are $\mathcal{Z}$-rational series.

Arguing in the same way, it can also be shown that there exists a $\mathcal{Z}$-rational series Sum_4 such that $(Sum_4|w) = 0$ when $w \notin C$ and

$$(Sum_4|w) = - \sum_{(i,j,k,l) \in V} \mid X_k^{(i)} - X_l^{(j)} \mid$$

when $w = \underline{w} \in C$, where $X_k^{(i)}$ is equal to

$$X_k^{(1)} = n(k,1)_1 \ , \quad X_k^{(2)} = |w_1^{(k)}(\underline{n(k,1)})|_a$$
$$X_k^{(3)} = n(k,2)_1 \ , \quad X_k^{(4)} = |w_2^{(k)}(\underline{n(k,2)})|_a$$

according to the case considered. Using now lemmas 3.5 and 3.7, the above method allows also to construct a $\mathcal{Z}$-rational series Sum_5 such that $(Sum_5|w) = 0$ when $w \notin C$ and

$$(Sum_5|w) = - \mid \sum_{i=1}^{p} p_i \, y_3^{(i)} - 1 \mid$$

when $w = \underline{w} \in C$.

Using finally lemmas 3.6 and 3.7 and the previous method, we can easily construct a $\mathcal{Z}$-rational series Sum_6 such that $(Sum_6|w) = -1$ when $w \notin C$ and such that

$$(Sum_6|w) = \sum_{i=1}^{p} \sum_{j=1}^{3} \ (\ \min(n(i,j)_k)_k - \max(n(i,j)_k \,)_k \ )$$

when $w = \underline{w} \in C$.

Let us now consider the $\mathcal{Z}$-rational series HD defined by

$$HD = Sum_1 \odot Sum_2 \odot Sum_3 \odot Sum_4 \odot Sum_5 \odot Sum_6$$

We clearly have $(HD|w) = -1$ when $w \notin C$. On the other hand, $(HD|w)$ is obviously a sum of negative integers when $w \in C$. Hence the general term of the series HD is always negative or equal to zero. It follows immediately that $(HD|w) \geq 0$ if and only if $(HD|w) = 0$ and this last condition is satisfied if and only if $w \in C$ and $(Sum_k|w) = 0$ for every $k \in [1,6]$. Note now that the condition $w \in C$ and $(Sum_6|w) = 0$ is clearly equivalent to the fact that $w = w(\underline{xy})$ for some vector $\underline{xy} \in (\mathbb{N} - \{0\})^{7p}$. Hence $(HD|w)$ can be equal to zero only if w has this form. But we clearly have

$$\begin{aligned}(HD(w(\underline{xy})) \ = \ & - \mid \sum_{i=1}^{p} p_i \, y_3^{(i)} - 1 \mid - \sum_{i=1}^{p} \mid y_1^{(i)} - x_1^{(i)} x_2^{(i)} \mid - \sum_{i=1}^{p} \mid y_2^{(i)} - x_3^{(i)} x_4^{(i)} \mid \\ & - \sum_{i=1}^{p} \mid y_3^{(i)} - y_1^{(i)} y_2^{(i)} \mid - \sum_{(i,j,k,l) \in V} \mid x_k^{(i)} - x_l^{(j)} \mid\end{aligned}$$

for every $\underline{xy} \in (\mathbb{N} - \{0\})^{7p}$. It follows now immediately from our study that the diophantine equation (HD) has a solution in positive integers if and only if there exists a word $w \in A^*$ such that $(HD|w) \geq 0$. Hence it follows from a previous remark that the local inequality

problem for $\mathcal{Z}$-rational series over A is undecidable. Thus, according to our reduction work, this ends our proof. ■

Notes : 1) Using the same kind of ideas than in the above proof, it can be shown that any diophantine equation of degree k can be directly encoded as a local inequality problem for $\mathcal{Z}$-rational series over an alphabet with 4 letters.

2) The above proof shows that there exists a $\mathcal{Z}$-rational series HD which has a non-recursive 0-support. [6] This is completely different from the situation that occurs in $\mathcal{M}$ since m-supports of $\mathcal{M}$-rational series are always rational languages (see [12] for more details).

3) In the above proof, we often used rational series S in $\mathcal{Z}\mathrm{Rat}(A)$ such that $-S$ remains also in $\mathcal{Z}\mathrm{Rat}(A)$. It should be noticed that this property does not hold in general (see [10] or [12]). In fact, it can be shown that the equality problem is decidable when restricted to $\mathcal{Z}$-rational series S such that $-S$ is also $\mathcal{Z}$-rational (see [12]).

As an immediate corollary of the previous theorem, we obtain according to propositions 2.1 and 2.2 :

COROLLARY 3.8 : Let A be an alphabet with at least 2 letters. Then the equality, inequality, local equality and local inequality problems are all undecidable questions for $\mathcal{Z}$-rational series over A.

Note : The proof of theorem 3.1 shows in fact than an inequality problem in $\mathcal{Z}\mathrm{Rat}(A)$ of the form $S \leq -1$ is already undecidable when $|A| \geq 2$.

4 Undecidability of the equality problem for $\mathcal{M}$

4.1 Reduction of decidability problems

In this section, we show that the decidability for $\mathcal{M}$ (resp. $\mathcal{N}$) of any problem considered in section 2 is equivalent to the decidability of the same problem for $\mathcal{Z}$. Let us now first prove this equivalence for $\mathcal{M}$ and $\mathcal{Z}$.

[6] For every element $m \in \mathcal{Z}$, the m-support of a series $S \in \mathcal{Z} \ll A \gg$ is the language that consists in the words w such that $(S|w) = m$.

THEOREM 4.1 : Let A be an arbitrary alphabet. Then the equality problem, the inequality problem, the local equality problem or the local inequality problem for $\mathcal{M}$-rational series over A is decidable if and only if the same problem is decidable for $\mathcal{Z}$-rational series over A.

Proof : Since all the proofs are the same, we shall only show here the equivalence between the decidability of the equality problems for $\mathcal{M}$ and $\mathcal{Z}$. Clearly we just have then to prove that the decidability of the equality problem in $\mathcal{M}$ implies the decidability of the same problem in $\mathcal{Z}$.

Let then R and S be two $\mathcal{Z}$-rational series over the alphabet A. According to the Kleene-Schützenberger theorem, R and S are $\mathcal{Z}$-recognizable series. Let us now consider two $\mathcal{Z}$-automata (I, μ, T) and (J, ν, F) of order m and n recognizing respectively R and S. Let us then consider for every $k \in \mathbb{Z}$ the new vectors $I(k), J(k), T(k), F(k)$ and the new $\mathcal{Z}$-representations μ_k and ν_k of A^* defined by

$$\forall\, a \in A, \quad \mu_k(a) = (\, \mu(a)_{i,j} + k \,)_{1 \le i,j \le m} \ , \ \nu_k(a) = (\, \nu(a)_{i,j} + k \,)_{1 \le i,j \le n}$$

$$I(k) = (\, I_i + k \,)_{i=1,\ldots,m} \ , \ J(k) = (\, J_i + k \,)_{i=1,\ldots,n}$$

$$T(k) = (\, T_i + k \,)_{i=1,\ldots,m} \ , \ F(k) = (\, F_i + k \,)_{i=1,\ldots,n}$$

It is easy to see that we have

$$I(k)\, \mu_k(w)\, T(k) = I\, \mu(w)\, T + 2k + k|w|$$

$$J(k)\, \nu_k(w)\, F(k) = J\, \nu(w)\, F + 2k + k|w|$$

for every word $w \in A^*$. Let now R_k and S_k be the two series recognized respectively by the automata $(I(k), \mu_k, T(k))$ and $(J(k), \nu_k, F(k))$. It follows immediately from the above computations that we have $S = T$ iff $S_k = T_k$ for any fixed $k \in \mathbb{Z}$. But it is easy to see that T_k and S_k are $\mathcal{M}$-recognizable series when k is greater than

$$-\min_{i,j} (\ \mu(a)_{i,j}, \nu(a)_{i,j}, I_i, J_i, T_i, F_i\ ) \ \in \mathbb{Z} \cup \{-\infty\}.$$

Hence we showed that the equality of two $\mathcal{Z}$-recognizable series is equivalent to the equality of two $\mathcal{M}$-recognizable series. This ends proving that the decidability of the equality problem for $\mathcal{M}$-rational series implies the decidability of the same problem for $\mathcal{Z}$-rational series. Our theorem is then proved. ∎

Using the same method as in the previous theorem, we can also get the following result that shows our equivalence result for $\mathcal{Z}$ and $\mathcal{N}$.

THEOREM 4.2 : Let A be an arbitrary alphabet. Then the equality problem, the inequality problem, the local equality problem or the local inequality problem

for $\mathcal{N}$-rational series over A is decidable if and only if the same problem is decidable for $\mathcal{Z}$-rational series over A.

Proof : Since all the proofs are similar, we shall also only show here the equivalence between the equality problems for $\mathcal{N}$ and $\mathcal{Z}$. It suffices then clearly to prove that the decidability of the equality problem for $\mathcal{N}$ implies the decidability of the same problem

for $\mathcal{Z}$. But since $\mathcal{N}$ is effectively isomorphic to $\mathcal{Z}^-$ (cf section 1), we have just to prove that the decidability of the equality problem for $\mathcal{Z}^-$ implies the decidability of the same problem for $\mathcal{Z}$.

Let us now take the same notations that in the proof of theorem 4.1. It is easy to see that S_k and R_k are $\mathcal{Z}^-$-recognizable series when k is less than

$$M = -\max_{i,j} \left(\mu(a)_{i,j}, \nu(a)_{i,j}, I_i, J_i, T_i, F_i \right) \in \mathbb{Z}$$

where the above maximum is only taken over the values which are not equal to $+\infty$ (when every value involved in the above maximum is equal to $+\infty$, we set $M = +\infty$). Hence, arguing as in the proof of the previous theorem, it follows that the equality of two $\mathcal{Z}$-rational series is equivalent to the equality of two $\mathcal{Z}^-$-rational series. This ends therefore the proof of our theorem. ∎

4.2 Undecidability of the equality problem for $\mathcal{M}$ and $\mathcal{N}$

As immediate corollaries of the previous results, we get the following undecidability results. Observe that they solve in particular the open problems we speak of in our introduction.

COROLLARY 4.3 : Let A be an alphabet with at least two letters. Then the equality problem, the inequality problem, the local equality problem and the local inequality problem are all undecidable problems for $\mathcal{M}$ or $\mathcal{N}$-rational series over A.

Notes : 1) Our undecidability result implies in particular that no "equality theorem" in the sense of Eilenberg (cf [6] th. VI.8.1) can hold in $\mathcal{M}$. However it is interesting to notice that this can also be directly proved. Indeed, if we consider the following family of matrices

$$M_n = \begin{pmatrix} 1 & n \\ 1 & 0 \end{pmatrix}$$

which is indexed by $n \in \mathbb{N}$, it is easy to prove that we have

$$S_n = \sum_{k=0}^{+\infty} (M_n)_{1,1}^k \, a^k = \sum_{k=0}^{n} k \, a^k + \sum_{k=n+1}^{+\infty} (n+1) \, a^k$$

It follows from this computation that the two distinct series S_n and S_{n+1} (both associated to a $\mathcal{M}$-representation of order 2) coincide up to the order n. Hence we obtain effectively that no "equality theorem" (cf [6] th. VI.8.1) is possible for the tropical semiring, even in fact with one-letter series. [7]

2) (Simon, [18]) I. Simon introduced in [17] a hierarchy, denoted $(H_i(A))_{i \in \mathbb{N}}$, of $\mathcal{M}$-rational series which is strict when $|A| \geq 2$. The series of $H_0(A)$ are exactly the limited series and it is easy to prove that the equality problem for such series is decidable. On the other hand, using the same argument than in the proof of theorem 4.1, we can easily show that the equality problem in $\mathcal{M}\mathrm{Rat}(A)$ and in $H_1(A)$ are equivalent with respect to

[7] I. Simon ([18]) asked whether it is possible to construct the same kind of counterexample with both bounded coefficients and bounded number of states. Such an example is in fact impossible to find as we will see in section 4.3.

decidability. It follows therefore that the equality problem for $\mathcal{M}$-rational series in $H_1(A)$ is undecidable when $|A| \geq 2$. Hence the limit between decidability and undecidability for the equality problem for $\mathcal{M}$ occurs when one passes from 0 to 1 in Simon's hierarchy.

3) It is also interesting to observe that all the results and methods obtained and developped before, gave us that the problem

$$P, Q \in \mathcal{M}\mathrm{Rat}(A),\ P \leq Q,\ \exists\ w \in A^*,\ (P|w) = (Q|w)$$

is undecidable when A has at least two letters. This result should be put in parallel with the same result for $\mathbb{N}$-rational series (see [6]) which is comparatively much more easier to obtain. Note also that the proof of this last result is based on an encoding of the Post correspondance problem, which is possible since $M_n(\mathbb{N})$ contains non-trivial free submonoids when $n \geq 2$. This is not the case neither for $M_n(\mathcal{M})$, nor for $M_n(\mathcal{Z})$ by an easy growth argument. Hence it seems difficult to adapt for $\mathcal{M}$-rational series the undecidability proof of the above result that works for $\mathbb{N}$-rational series. It was this situation that suggested to several authors (see [15] for instance) that the equality problem for $\mathcal{M}$ was decidable.

Moreover we can also obtain as an application of our methods the following decidability results which make complete our study of the decidability of the four problems considered in section 2 for $\mathcal{Z}$, $\mathcal{M}$ and $\mathcal{N}$.

COROLLARY 4.4 : The equality problem, the inequality problem, the local equality problem and the local inequality problem are decidable for $\mathcal{M}$, $\mathcal{N}$ or $\mathcal{Z}$-rational series over a *one-letter* alphabet $A = \{a\}$.

Proof : According to theorems 4.1, 4.2 and to propositions 2.1 and 2.2, it suffices to show that the local equality problem is decidable for $\mathcal{M}$-rational series over a one-letter alphabet in order to prove our corollary. Using now a classical embedding due to C. Choffrut of $\mathcal{M}$ into the semiring $\mathrm{Rat}(b^*)$ of rational languages over a one-letter alphabet $\{b\}$ (cf [3] or [4]), this last problem appears in fact as an intersection problem for special kinds of rational languages of $\mathrm{Rat}(a^* \times b^*)$. The decidability of our result follows now from the decidability of the intersection problem for $\mathrm{Rat}(a^* \times b^*)$ (see [7] for instance). This ends our proof. ■

Note : Using a fine study of the iterated power of a square matrix with entries in $\mathcal{Z}$ or a study of one-letter $\mathcal{Z}$-rational expressions, it can also be shown directly that the equality problem for one-letter $\mathcal{Z}$-rational series is decidable.

4.3 $\mathcal{M}$-automata with constrained entries

We devote this section to the study of the equality problem for $\mathcal{M}$-automata whose entries are supposed to belong to some fixed set. Let us therefore give the following definition.

DEFINITION 4.1 : Let S be a subset of $\mathbb{Z}$. A $\mathcal{Z}$-automaton $\mathcal{A} = (I, \mu, T)$ is then said to be a S-automaton iff every non-infinite entry of I, $(\mu(a))_{a \in A}$ and T belongs to S.

Note first that the equality problem is decidable for S-automata when $|S| = 1$ as claims the following result.

PROPOSITION 4.5 : Let $S = \{k\}$ be a one-element subset of $\mathbb{Z}$. Then the equality problem for S-automata is decidable.

Proof : Using the same method than in the proof of theorem 4.1, it can be easily seen that the equality problem for $\{k\}$-automata is equivalent to the same problem for $\{0\}$-automata. But $\{0, +\infty\}$ is a subsemiring of $\mathcal{M}$ which is isomorphic to the boolean semiring. Hence it follows from these results that the equality problem for $\{k\}$-automata is equivalent to the equality problem for usual boolean automata. Our proposition follows then immediately. ■

On the other hand, the proof of theorem 3.1 shows that the equality problem is undecidable for $\{-1, 0, 1\}$-automata. Using the same kind of method than in the proof of theorem 4.1, it follows easily that the equality problem is undecidable for $\{0, 1, 2\}$-automata. Therefore the question remains to see where is the limit between undecidability and decidability when we constraint the entries of a $\mathcal{M}$-automaton to belong to some given set.

In this direction, let us now have a look on the class of $\{0, 1\}$-automata for which we shall also prove that the equality problem is undecidable. In fact, we will show that every $\mathcal{M}$-automaton can be simulated in a "rational way" by a $\{0, 1\}$-automaton. This last result will allow us to reduce the equality problem for general $\mathcal{M}$-automata to the same problem for $\{0, 1\}$-automata. However let us first introduce a useful denotation.

Notation : Let A be an alphabet, let t be a letter that does not belong to A, let n be some integer and let σ_n be the substitution of A^* that maps every letter $a \in A$ onto a^n. Then we will denote by $T(n)$ the rational language over $A \cup \{t\}$ defined by

$$T(n) = \sigma_n(A^*)\, t^n$$

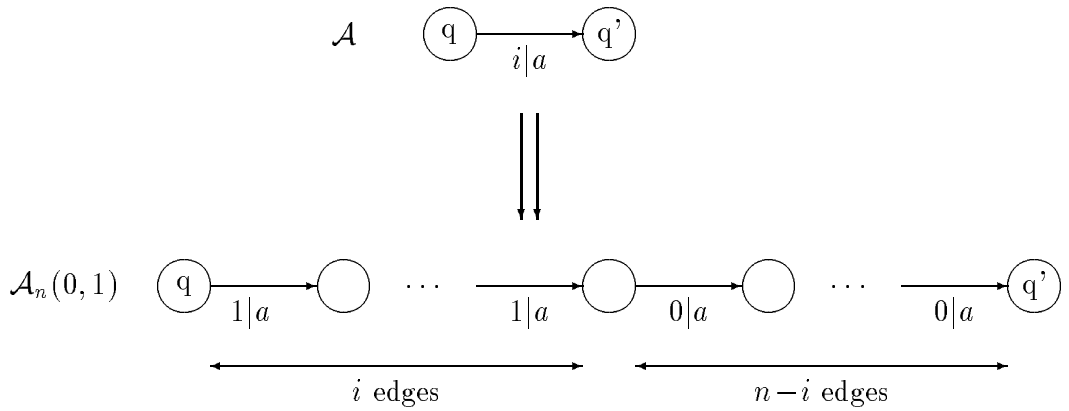
PROPOSITION 4.6 : Let A be an alphabet, let $\mathcal{A} = (I, \mu, T)$ be a $\mathcal{M}$-automaton over A and let N be the maximal value of the non-infinite elements involved in the entries of I, $(\mu(a))_{a \in A}$ and T. Then, for every $n \geq N$, there exists a $\{0, 1\}$-automaton $\mathcal{A}_n(0, 1)$ over $A \cup \{t\}$ such that

$$(\mathcal{A}_n(0,1)|w) = \begin{cases} (\mathcal{A}|u) & \text{when } w = \sigma_n(u)\, t^n \in T(n) \text{ with } u \in A^* \\ +\infty & \text{when } w \notin T(n) \end{cases}$$

Proof : According to a classical result (see theorem 7.2 of [13] for instance), one can always suppose that the initial vector I of $\mathcal{A}$ is equal to $I = (0\ +\infty\ \ldots\ +\infty)$. Let us now consider some integer $n \geq N$ where N is the maximal value of the non-infinite entries of I, $(\mu(a))_{a \in A}$ and T. We can then construct the $\{0, 1\}$-automaton $\mathcal{A}_n(0, 1)$ obtained from $\mathcal{A}$ as follows :

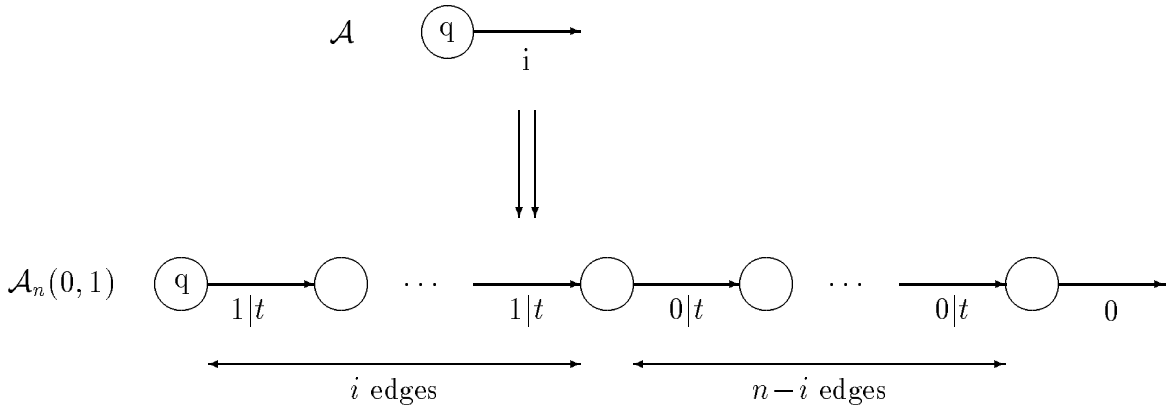
- for every pair (q, q') of states in $\mathcal{A}$ and for every letter $a \in A$ such that $\mu(a)_{q,q'} \neq +\infty$, we replace the edge of $\mathcal{A}$ going from q to q' and labeled by a by n new edges (creating therefore the new $n - 1$ necessary states) all labeled by a and equiped with the

following costs : the first $\mu(a)_{q,q'}$ new edges have cost 1 and the last $n - \mu(a)_{q,q'}$ new edges have cost 0.



The previous picture explains the operation that we make on every edge of $\mathcal{A}$ in order to obtain the automaton $\mathcal{A}_n(0,1)$. Note that all the infinite costs on the edges of $\mathcal{A}$ do not change in $\mathcal{A}_n(0,1)$.

- the value of the entries of the initial and final vectors of $\mathcal{A}_n(0,1)$ corresponding to any new created state in the previous operation is set to $+\infty$.
- the values of the entries of the initial vector of $\mathcal{A}_n(0,1)$ corresponding to the states of $\mathcal{A}$ are not changed.
- for every state q of $\mathcal{A}$ such that $T_q \neq +\infty$, the new value of the final vector of $\mathcal{A}_n(0,1)$ on q is set to $+\infty$. Then $n-1$ new states and n corresponding new edges all labelled with the extra letter t are created, the first new edge beeing issued from q. These new edges are equiped with the following costs : the T_q first new edges have cost 1 and the $n - T_q$ last new edges have cost 0. Moreover the entry of the final vector of $\mathcal{A}_n(0,1)$ corresponding to the last of these new created states is set to 0.



The previous picture explains the operation that we make on every final state of $\mathcal{A}$ in order to obtain the automaton $\mathcal{A}_n(0,1)$. Note that every state whose final vector entry is $+\infty$ in $\mathcal{A}$ has the same final infinite value in $\mathcal{A}_n(0,1)$.

- the entries of the initial and final vectors of $\mathcal{A}_n(0,1)$ corresponding to a new created state which is not the last one in the previous operation, is set to $+\infty$.

It follows now easily from our construction that

$$(\mathcal{A}_n(0,1)|w) = \begin{cases} (\mathcal{A}|u) & \text{if } w = \sigma_n(u)\,t^n \text{ for some } u \in A^* \\ +\infty & \text{if } w \notin \sigma_n(A^*)\,t^n \end{cases}$$

This ends therefore the proof of our proposition. ■

Note : Observe that the automaton $\mathcal{A}_n(0,1)$ constructed in the proof of the previous proposition is in fact a *distance automaton* (cf [8] or [19]), i.e. a $\{0,1\}$-automaton such that all entries of the initial and final vectors consist only of 0 or $+\infty$.

As an immediate consequence of the previous proposition, we get :

COROLLARY 4.7 : Let A be an alphabet with at least two letters. Then the equality problem for $\{0,1\}$-automata over A is undecidable.

Proof : Let $\mathcal{A}$ and $\mathcal{B}$ be two $\mathcal{M}$-automata. Let then N be some integer in $\mathbb{N}$ such that every non-infinite element involved in the entries of the initial vector, the final vector or the transition matrices of $\mathcal{A}$ and $\mathcal{B}$ is less than N. Then, if $\mathcal{A}_N(0,1)$ and $\mathcal{B}_N(0,1)$ denote the $\{0,1\}$-automata associated to $\mathcal{A}$ and $\mathcal{B}$ by proposition 4.6, the equivalence of $\mathcal{A}$ and $\mathcal{B}$ is clearly equivalent to the equivalence of $\mathcal{A}_N(0,1)$ and of $\mathcal{B}_N(0,1)$. Our result follows now immediately from corollary 4.3. ■

Note : It follows immediately from the note following proposition 4.6 and from the proof of the previous corollary that the equality problem for distance automata is also undecidable.

Using the previous result, we can easily deduce that the equality problem for S-automata is undecidable when $|S| = 2$:

COROLLARY 4.8 : Let A be an alphabet with at least two letters and let k, l be two different integers in $\mathbb{Z}$. Then the equality problem for $\{k,l\}$-automata over A is undecidable.

Proof : We can always suppose that $k < l$. Note first that using the same trick than the one involved in the proof of theorem 4.1, the equality problem for $\{k,l\}$-automata can be easily reduced to an equality problem for $\{0, l-k\}$-automata. Hence we can suppose that $k = 0$ and argue only with $\{0,l\}$-automata in order to prove our corollary.

Therefore let now $\mathcal{A}$ and $\mathcal{B}$ be two $\{0,l\}$-automata. Let then $\mathcal{A}(0,1)$ and $\mathcal{B}(0,1)$ be the two $\{0,1\}$-automata obtained respectively from $\mathcal{A}$ and $\mathcal{B}$ by replacing by 1 every entry equal to l in the initial vector, the final vector and the transition matrices of $\mathcal{A}$ and $\mathcal{B}$. An easy computation shows then that we have

$$(\mathcal{A}|w) = l\,(\mathcal{A}(0,1)|w) \qquad \text{and} \qquad (\mathcal{B}|w) = l\,(\mathcal{B}(0,1)|w)$$

for every $w \in A^*$. It follows immediately that the equivalence of $\mathcal{A}$ and $\mathcal{B}$ is clearly equivalent to the equivalence of the two $\{0,1\}$-automata $\mathcal{A}(0,1)$ and $\mathcal{B}(0,1)$. Our corollary follows now from the previous corollary 4.7. ■

Note : This last result solves our initial problem : the equality problem is decidable for S-automata when $|S| = 1$, but becomes undecidable when $|S| \geq 2$.

4.4 Other connected undecidability results

Let us denote by $Rat(b)$ the subsemiring of rational subsets over the one-letter alphabet $\{b\}$ equiped with union and intersection as sum and product. Let us also denote by $Fin(b)$ the subsemiring of $Rat(b)$ whose support is the family of finite subsets of $\{b\}^*$.

Then $\mathcal{M}$ can be identified with the subsemiring of $Fin(b)$ that consists in the sets of the form $(1 + b)^m$ together with empty set. [8] It follows now immediately from this embedding and from corollary 4.3 that the following result (due originally to Ibarra (cf [9]) and also obtained with another method by Lisovik (cf [14])) holds :

COROLLARY 4.9 : Let A be an alphabet with at least two letters. Then the equality problem is an undecidable problem for $Rat(b)$ or $Fin(b)$-rational series over A.

Note : The original proof of Ibarra that corresponded to the undecidability of the equality problem from $Fin(b)$ was very technical and used directly an encoding of a Turing machine halting problem.

Acknowledgements

I want to thank here professor C. Choffrut who discovered a gap in the first "proof" of the undecidability result (cf [11]), professor S. Grigorieff who indicated me the wonderful reference [5] and professor I. Simon for the personal papers he communicated me and for all the e-conversations we had that allowed to improve the final version of this paper. Special thanks are also due to professor A. Weber : section 4.3 is in fact a consequence of several discussions we had together. Finally I must also thank the "LAboratoire de Combinatoire et d'Informatique Mathématique" (LACIM; Université du Québec à Montréal) for its kind support during the preparation of the first version of this paper.

[8] This is the embedding due to C. Choffrut (cf [3] or [4]) we used already before.

References

[1] ADLER A., *A reduction of homogeneous diophantine problems*, Journ. of London Math. Soc. (2), **3**, pp. 446-448, 1971

[2] BERSTEL J., REUTENAUER C., *Rational series and their languages*, Springer Verlag, 1986

[3] CHOFFRUT C., *Séries rationnelles d'image finie*, Technical Report 79-6, LITP, Paris, 1979

[4] CHOFFRUT C., *Rational relations and rational series*, Theor. Comput. Sci., **98**, pp. 5-13, 1992

[5] DAVIS M., MATIJASEVIC Y., ROBINSON J., *Hilbert's tenth problem, diophantine equations : positive aspects of a negative solution*, Proceedings of Symposia in Pure Mathematics, Vol. 28, pp. 323-378, 1976

[6] EILENBERG S., *Automata, languages and machines*, Vol. A, Academic Press, 1974

[7] GIBBONS A., RYTTER W., *On the decidability of some problems about rational subsets of free partially commutative monoids*, Theor. Comp. Sci., **48**, pp. 329-337, 1986

[8] HASHIGUSHI K., *Improved limitedness theorem on finite automata with distance functions*, Theor. Comput. Sci., **72**, p. 27-38, 1990

[9] IBARRA O.H., *The unsolvability of the equivalence problem for ϵ-free NGSM's with unary input (output) alphabet and applications*, SIAM J. Comput., **7**, 4, p. 524-532, 1978

[10] KOYABASHI N., [in "LATIN'92 Proceedings"], Lect. Notes in Comput. Sci., **583**, Springer Verlag, 1992

[11] KROB D., *The equality problem for rational series with multiplicities in the tropical semiring is undecidable*, [in "Automata, Languages and Programming", ICALP'92 Proceedings], Lect. Notes in Comput. Sci., **623**, p. 101-112, Springer Verlag, 1992

[12] KROB D., *Some consequences of a Fatou property of the tropical semiring*, LITP Technical Report 92-64, Paris, 1992

[13] KUICH W., SALOMAA A., *Semirings, automata, languages*, Springer Verlag, 1986

[14] LISOVIK L.P., *The identity problem for regular events over the direct product of free and cyclic semigroups*, Dok. Akad. Nauk. Ukraiinskoj RSR, ser. A 6, p. 410-413, 1979 (in ukrainien)

[15] SIMON I., *Recognizable sets with multiplicities in the tropical semiring*, [in "MFCS'88 Proceedings"], Lect. Notes in Comput. Sci., **324**, p. 107-120, Springer Verlag, 1988

[16] SIMON I., *Some open problems for automata with multiplicities*, Private communication

[17] SIMON I., *The nondeterministic complexity of a finite automaton*, [in "Mots", M. Lothaire], pp. 384-400, Hermès, 1990

[18] SIMON I., Personal communication

[19] WEBER A., *Distance automata having large distance or finite ambiguity*, [in "MFCS'90 Proceedings"], Lect. Notes in Comput. Sci., **452**, p. 508-515, Springer Verlag, 1990