KYBERNETIKA — VOLUME 39 (2003), NUMBER 2, PAGES 177 - 192

SOME REPRESENTATIONS FOR SERIES ON IDEMPOTENT SEMIRINGS – or How to Go Beyond Recognizability Keeping Representability

INES KLIMANN

In this article, we compare different types of representations for series with coefficients in complete idempotent semirings. Each of these representations was introduced to solve a particular problem. We show how they are or are not included one in the other and we present a common generalization of them.

Keywords: idempotent semirings, recognizable series
AMS Subject Classification: 93B25, 13F25

1. INTRODUCTION

The aim of this article is to compare different types of representations for series on certain idempotent semirings (certain of them are available also without such hypotheses, for example, it is possible to define rational series with coefficients in a non-complete semiring).

After recalling the basic definitions in Section 2, we introduce in Section 3 the classical notions of recognizable and rational series [2]. In Section 4 we develop the notion of pseudo-recognizable series, introduced in [7] to solve certain inequations on series. Then we present the non linear representations which appear in a paper of J.-E. Pin and J. Sakarovitch [9] to solve the following classical problem of formal language theory: let $L_1, \dots, L_n$ be n languages recognized by the monoids $M_1, \dots M_n$ respectively; given an operation φ, how to build a monoid M, function of $M_1, \dots M_n$, which recognizes the language $(L_1, \dots, L_n)\varphi$? Finally, in the last section, we build a common generalization of multi-representations and non linear representations.

For each type of representation introduced, we show that, in the particular case of a finite semiring of coefficients, the corresponding notion of regularity is in fact the simple rationality.

2. BASES

In this section, we explore the basic properties of idempotent semirings and series.

If X is a set, we denote by $\mathcal{P}(X)$ the power set of X and by $\mathcal{P}_f(X)$ the set of finite subsets of X. The set of rational languages over an alphabet Σ is denoted by $\mathrm{Rat}(\Sigma)$.

If w is a word and a a letter, we denote by $|w|$ the length of w and by $|w|_a$ the number of occurrences of letter a in w.

2.1. Idempotent semirings

2.1.1. Definition and basic properties.

A *semiring* is a quintuple $(\mathcal{S}, +, *, 0, 1)$ with the following properties (see [5]):

- $(\mathcal{S}, +, 0)$ is a commutative monoid,
- $(\mathcal{S}, *, 1)$ is a monoid; as usual we denote by ab the product $a*b$ for all a, $b \in \mathcal{S}$,
- 0 is absorbing: $a0 = 0a = 0$ for all $a \in \mathcal{S}$,
- multiplication is distributive with respect to addition, i.e. $a(b+c) = ab + ac$ and $(b+c)a = ba + ca$, for all $a, b, c \in \mathcal{S}$.

A semiring is *commutative* if multiplication is commutative and is *idempotent* if addition is idempotent. In this paper, we always treat with idempotent semirings. We often write $(\mathcal{S}, +, *)$ or simply $\mathcal{S}$ for the semiring $(\mathcal{S}, +, *, 0, 1)$.

Examples 1.

- The boolean semiring $\mathbb{B} = \{0, 1\}$ is a finite commutative idempotent semiring.
- The tropical semiring $\mathbb{N}_{\min} = (\mathbb{N} \cup \{+\infty\}, \min, +, +\infty, 0)$ is an infinite commutative idempotent semiring. Several exotic semirings of this type can be defined, like $\mathbb{Z}_{\min} = (\mathbb{Z} \cup \{+\infty\}, \min, +, +\infty, 0)$.
- The set of (recognizable) languages over a fixed alphabet, with union for addition and concatenation for multiplication, is an idempotent semiring (with identity elements: for addition the empty set and for multiplication the singleton $\{1\}$ containing the empty word). It is infinite if the alphabet is non-empty and non commutative if the alphabet contains at least two letters.
- The power set $\mathcal{P}(M)$ of a monoid M, provided with union and multiplication, is an idempotent semiring. It is commutative if M is commutative. If P and Q are subsets of M, their product is the subset $PQ = \{pq \mid p \in P \text{ and } q \in Q\}$ of M.

Throughout this paper, $\mathcal{S}$ denotes an idempotent semiring and Σ a finite alphabet.

We consider the *natural order* over $\mathcal{S}$ given by: $a \leq b$ if and only if there exists $c \in \mathcal{S}$ such that $b = a + c$. It is well known and easy to see that $a \leq b$ if and only if $b = a + b$. It follows in particular that the least element of $\mathcal{S}$ is 0.

Note that for the tropical semiring, the natural order is exactly the inverse of the usual order on $\mathbb{N}$; in $\mathbb{N}$: $2 \leq 3$, but in $\mathbb{N}_{\min}$: $3 \leq 2$.

Multiplication is compatible with the order.

2.1.2. Supremum and infimum

If $\mathcal{T} \subseteq \mathcal{S}$ is non empty, the sum of its elements is its supremum. By analogy, if $\mathcal{T}$ is any subset of $\mathcal{S}$, we denote by $\sum_{x \in \mathcal{T}} x$ the supremum of $\mathcal{T}$, if it exists. This notation is justified since, in particular, the supremum of $\mathcal{T} \cup \mathcal{T}'$ is the sum of the suprema of $\mathcal{T}$ and $\mathcal{T}'$. If $\mathcal{T} \subseteq \mathcal{S}$ is non empty and has an infimum, we denote it by $\bigcap_{t \in \mathcal{T}} t$, or $a \cap b$ if $\mathcal{T} = \{a, b\}$.

Examples 2.

- The infimum of two (rational) languages is their intersection (note that intersection on languages preserve rationality).
- If $\mathcal{S} = \mathbb{N}_{\min}$, the infimum of two elements is their maximum in the usual order.

Recall that an ordered set is *complete* if each of its subsets has a supremum. A semiring $\mathcal{S}$ is *complete* if it is complete as an ordered set and satisfies the following distributivity conditions:

$$\text{for all } \mathcal{T} \subseteq \mathcal{S}, s \in \mathcal{S} : \left(\sum_{t \in \mathcal{T}} t\right) s = \sum_{t \in \mathcal{T}} (ts) \text{ and } s \left(\sum_{t \in \mathcal{T}} t\right) = \sum_{t \in \mathcal{T}} (st).$$

Examples 3.

- If Ξ is an alphabet, $(\mathcal{P}(\Xi^*), \cup, \cdot)$ is a complete semiring.
- The tropical semiring is complete.

If we now assume that $\mathcal{S}$ is a complete idempotent semiring, then every $\mathcal{T} \subseteq \mathcal{S}$ has an infimum: the sum of all x such that $x \leq t$ for each $t \in \mathcal{T}$. It follows directly that the operation of infimum is idempotent and compatible with the order.

Proposition 1. In a complete idempotent semiring $\mathcal{S}$, the following distributivity property holds. Let Y be a subset of $\mathcal{S}$ and let $x \in \mathcal{S}$. Then:

$$x \left(\bigcap_{y \in Y} y\right) \leq \bigcap_{y \in Y} (xy) \quad \text{and} \quad \left(\bigcap_{y \in Y} y\right) x \leq \bigcap_{y \in Y} (yx).$$

From now on, we assume that $\mathcal{S}$ is complete and moreover that the operations $+$ and $\cap$ supply $\mathcal{S}$ with a structure of a distributive lattice, i.e. $+$ distributes over $\cap$ and $\cap$ over $+$. This property is used in Section 4 to define pseudo-recognizable series.

Example 4.

- If $\mathcal{S}$ is the set of (recognizable) languages on Σ, union and intersection distribute one with respect to the other.
- If $\mathcal{S} = \mathbb{N}_{\min}$, minimum and maximum distribute one with respect to the other.

We define a new operation which is not classical in series theory, but corresponds to the notion of residuation [3]. If a and y are elements of a semiring $\mathcal{S}$, the *cut* $a\backslash y$ of y by a is the element

$$a\backslash y = \sup\{x \mid ax \leq y\}.$$

This element is well-defined since the semiring is idempotent and complete.

If a, x and y are elements of $\mathcal{S}$, then $ax \leq y$ if and only if $x \leq a\backslash y$. As a direct consequence, for each $x \in \mathcal{S}$, $1\backslash x = x$.

If the semiring of coefficients is the tropical semiring, the cut is the quasi-difference [4]: $a\backslash x = \max_{\mathbb{Z}}(0_{\mathbb{Z}}, x - a)$ (here the difference is the classical difference on $\mathbb{Z}$).

Product has precedence with respect to cut, and cut with respect to sum and infimum: $a\backslash xy = a\backslash(xy)$, $a\backslash x + y = (a\backslash x) + y$ and $a\backslash x \cap y = (a\backslash x) \cap y$.

2.2. Formal series

We consider the set $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ of formal series on Σ, with coefficients in $\mathcal{S}$. A typical element of $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ is written $A = \sum_{w\in\Sigma^*}(A, w)w$, with $(A, w) \in \mathcal{S}$.

The *constant coefficient* of A is $(A, 1)$; A is *proper* if its constant coefficient is zero. We identify an element a of $\mathcal{S}$ with the constant series, also denoted by a, defined by $(a, 1) = a$ and $(a, w) = 0$ for every non-empty word w. In the same way, we identify a word $w \in \Sigma^*$ with the series also denoted by w and defined by $(w, u) = 0$ if $u \neq w$ and $(w, w) = 1$. Our notation is taken from [2].

The *support* of a series A is the language $\operatorname{supp} A = \{w \in \Sigma^* \mid (A, w) \neq 0\}$. If $s \in \mathcal{S}$, the *s-support* of a series A is the language $A^{-1}s = \{w \mid (A, w) = s\}$.

A series is said to be a *language* if its coefficients belong to $\{0, 1\}$, a *polynomial* if its support is finite. The set of polynomials is denoted by $\mathcal{S}\langle\Sigma\rangle$.

Operations on $\mathcal{S}$ are extended to the set of formal series by letting

$$(S + T, w) = (S, w) + (T, w) \text{ and } (ST, w) = \sum_{uv=w} (S, u)(T, v).$$

These operations provide $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ with a semiring structure. The natural order over $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ is then exactly the extension of the order of $\mathcal{S}$: $X \leq Y$ if and only if for all $w \in \Sigma^*$, $(X, w) \leq (Y, w)$. Since $\mathcal{S}$ is idempotent and complete, $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ is a complete idempotent semiring. In particular, the infimum $A \cap B$ of two elements $A, B \in \mathcal{S}\langle\langle\Sigma\rangle\rangle$ is given by $(A \cap B, w) = (A, w) \cap (B, w)$ for each $w \in \Sigma^*$. The distributivity of the lattice $(\mathcal{S}\langle\langle\Sigma\rangle\rangle, +, \cap)$ is inherited from the distributivity of $(\mathcal{S}, +, \cap)$.

If S is a series and u a word, we denote by $u^{-1}S$ the series whose coefficient on a word v is equal to (S, uv).

Remark 2. $\mathbb{B}\langle\langle\Sigma\rangle\rangle$ can be identified with $\mathcal{P}(\Sigma^*)$: $(A, w) = 1$ if and only if $w \in A$. The order previously introduced corresponds to inclusion, the infimum on elements of $\mathcal{S}$ to conjunction and the infimum on series to intersection of languages.

If a series S is proper, the family $(S^n)_{n\geq 0}$ is locally finite and hence summable. The *star* of S is the sum of this family: $S^* = \sum_{n\geq 0} S^n$.

A series S of $\mathcal{P}(\Xi^*)\langle\langle\Sigma\rangle\rangle$, where Σ and Ξ are alphabets, is called a *transduction*.

Remark 3. If S belongs to $\mathrm{Rat}(\Xi)\langle\langle\Sigma\rangle\rangle$, S is a rational series if and only if it is a rational transduction in the classical sense of the term (i. e. its graph is a rational relation over Σ and Ξ) [1, Proposition III.7.3].

3. SOME CLASSICAL REPRESENTATIONS

3.1. Rational series

Like for languages, *rational operations* (sum, product and star of proper series) allow to define a particular subset of $\mathcal{S}\langle\langle\Sigma\rangle\rangle$; the set of *rational series* on Σ with coefficients in $\mathcal{S}$ is the rational closure of $\mathcal{S}\langle\Sigma\rangle$ in $\mathcal{S}\langle\langle\Sigma\rangle\rangle$.

Examples 5.

- If $\mathcal{S} = \mathbb{B}$, a series is rational if and only if its support is a rational language.
- Let Σ be the two-letter-alphabet $\{a, b\}$.
  - Let $\mathcal{S} = \mathbb{N}_{\min}$. Define the series R by $1_{\mathbb{N}}a + 1_{\mathbb{N}_{\min}}b$ and the series Q_a by R^*. The series Q_a is rational, its coefficient on a word w is the number of occurrences of a in w.
  - Let Ξ be the one-letter-alphabet $\{c\}$ and $\mathcal{S} = \mathcal{P}(\Xi^*)$ be the set of words on Ξ. Define the series T by $\{c\}a + \{1\}b$ and the series S_a by T^*. The series S_a is rational, its coefficient on a word w is $\{c^{|w|_a}\}$.

3.2. Recognizable series

Let $\mathcal{S}^{n\times m}$ denote the set of (n, m)-matrices with entries in $\mathcal{S}$.

A series $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ is *recognizable* [2, Chapter 1] if and only if there exist an integer $n \geq 1$, a morphism of monoids $\mu : \Sigma^* \to \mathcal{S}^{n\times n}$ and two vectors $\lambda \in \mathcal{S}^{1\times n}$ and $\gamma \in \mathcal{S}^{n\times 1}$ such that, for all words w, $(S, w) = \lambda\mu(w)\gamma$. The triple (λ, μ, γ) is called a *linear representation* of S and n is its dimension.

Examples 6. Let Σ be the two-letter-alphabet $\{a, b\}$.

- Let $\mathcal{S} = \mathbb{N}_{\min}$. Define the morphism $\mu : \Sigma^* \to \mathbb{N}_{\min}^{1\times 1} : a \mapsto 1_{\mathbb{N}}, b \mapsto 1_{\mathbb{N}_{\min}}$ and the vectors $\lambda = \gamma = 1_{\mathbb{N}_{\min}}$. This linear representation recognizes the series Q_a introduced in Example 5, which is then also recognizable.

- Let Ξ be the one-letter-alphabet $\{c\}$ and $\mathcal{S} = \mathcal{P}(\Xi^*)$ be the set of words on Ξ. Define the morphism $\mu : \Sigma^* \to \mathcal{P}(\Xi^*)^{1\times 1} : a \mapsto \{c\}, b \mapsto \{1\}$ and the vectors $\lambda = \gamma = \{1\}$. This linear representation recognizes the series R_a introduced in Example 5, which is then also recognizable.

A morphism $\mu : \Sigma^* \to \mathcal{S}^{n\times n}$ being fixed, we denote by $S(\lambda, \gamma)$ the series S recognized by (λ, μ, γ).

The Kleene–Schützenberger theorem is the cornerstone of the theory of formal series:

Theorem 4. (Kleene–Schützenberger, [2]) A formal series is rational if and only if it is recognizable.

Here is an algebraic characterization of recognizability. A submodule of $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ is *stable* if for each element T of this submodule and each word u, the series $u^{-1}T$ still belongs to the submodule.

Proposition 5. [2] A series $S \in \mathcal{S}\langle\langle\Sigma\rangle\rangle$ is recognizable iff there exists a stable, finite generated left submodule of $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ containing S.

3.3. Hadamard product of recognizable series

In this subsection, we prove that the Hadamard product of two recognizable series with coefficients in a finite semiring $\mathcal{S}$ is recognizable, whether the semiring is commutative or not.

The *Hadamard product* of two series S and T is the series $S \odot T$ such that for all words w, $(S \odot T, w) = (S, w)(T, w)$.

Lemma 6. If $\mathcal{S}$ is finite, $S \in \mathcal{S}\langle\langle\Sigma\rangle\rangle$ is rational if and only if all its s-supports are rational ($s \in \mathcal{S}$).

Proof. Let S be a rational series. According to J. Berstel and C. Reutenauer [2, Proposition III.2.2], the s-supports of S are rational. Conversely, if the s-supports of S are rational, according to [2, Proposition III.2.1], for each $s \in \mathcal{S}$, the series $\sum_{w\in S^{-1}s} w$ is rational. The result follows straightforwardly since
$S = \sum_{s\in\mathcal{S}} s\left(\sum_{w\in S^{-1}s} w\right)$. □

Proposition 7. If $\mathcal{S}$ is finite, the Hadamard product of two recognizable series is recognizable.

Proof. Let S and T be two recognizable series and r in $\mathcal{S}$. The r-support of $S \odot T$ is rational. Indeed, the set $\Theta = \{(s,t) \in \mathcal{S} \times \mathcal{S} \mid st = r\}$ is finite because $\mathcal{S}$ is finite. Furthermore, for each word w, w belongs to $(S \odot T)^{-1}r$ if and only if there exist (s,t) in Θ such that w belongs to $S^{-1}s \cap T^{-1}t$.

And so
$$(S \odot T)^{-1}r = \bigcup_{(s,t)\in\Theta} (S^{-1}s \cap T^{-1}t).$$

Now, the languages $S^{-1}s$ and $T^{-1}t$ are recognizable by Lemma 6. Since the set Θ is finite, the r-supports of $S \odot T$ are recognizable. It follows by Lemma 6 that the series $S \odot T$ is recognizable. □

4. MULTI–REPRESENTATIONS AND PSEUDO–RECOGNIZABILITY

4.1. Definitions

In this section, we extend the notion of linear representation introduced in Section 3.2 to that of multi-representation. We defined this new notion in [7], in analogy with multi-automata [6].

If c is an element of $\mathcal{S}$, we define the series $c\backslash S$ by $(c\backslash S, w) = c\backslash(S,w)$, for each word w. A morphism μ being fixed, for each triple $(c,\lambda,\gamma) \in \mathcal{S} \times \mathcal{S}^{1\times n} \times \mathcal{S}^{n\times 1}$, we denote by $S(c,\lambda,\gamma)$ the series $c\backslash S(\lambda,\mu)$. Note that $S(1,\lambda,\gamma) = S(\lambda,\gamma)$.

In some cases, such a series is recognizable. We fix μ. Let $S = S(c,\lambda,\gamma)$ and $T = S(\lambda,\gamma)$: $c\backslash T = S$.

Proposition 8. If $\mathcal{S}$ is finite, $S(c,\lambda,\gamma)$ is rational.

Proof. We claim that $c\backslash T = \sum_{s\in\mathcal{S}}(c\backslash s)T^{-1}s$. Indeed, for each word $w \in \Sigma^*$ and every $s \in \mathcal{S}$, we have

$$\begin{cases} (T^{-1}s, w) = 1 & \text{if } (T,w) = s \\ (T^{-1}s, w) = 0 & \text{if } (T,w) \neq s. \end{cases}$$

So

$$\begin{aligned} \sum_{s\in\mathcal{S}}(c\backslash s)(T^{-1}s,w) &= \sum_{s\in\mathcal{S}}\sum_{w\in T^{-1}s} c\backslash s \\ &= \sum_{s\in\mathcal{S}} c\backslash(T,w) = \sum_{s\in\mathcal{S}}(c\backslash T,w). \end{aligned}$$

Hence S is a finite sum of recognizable series, since, by Lemma 6, the s-supports of T are rational languages. Thus, S is recognizable. □

In the commutative case, the Hankel matrix of a series can be used. The *Hankel matrix* of a series S on a commutative semiring is the infinite $(\Sigma^* \times \Sigma^*)$-matrix $H(S)$ defined by $H(S)_{u,v} = (S, uv)$.

Let us recall the key result about Hankel matrices.

Theorem 9. [10, Corollary II.3.2], [2, Thm. II.1.2] Let $\mathcal{S}$ be a commutative semiring. A series with coefficients in $\mathcal{S}$ is recognizable if and only if its Hankel matrix has only a finite number of independent columns.

Corollary 10. If the semiring of coefficients $\mathcal{S}$ is commutative, the series $S(c, \lambda, \gamma)$ is rational.

Proof. The following relation holds, for any words u, v:

$$H(S)_{u,v} = S(uv) = c\backslash T(uv) = c\backslash H(T)_{u,v}.$$

If $H(T)$ has a finite number of independent columns, so has $H(S)$. □

Let us define multi-representations and pseudo-recognizability.

Let $n \geq 1$ be an integer, μ be a morphism of monoids $\Sigma^* \to \mathcal{S}^{n\times n}$ and Φ be a positive boolean formula on $\mathcal{S} \times \mathcal{S}^{1\times n} \times \mathcal{S}^{n\times 1}$. It is convenient to call *atom* an element of $\mathcal{S} \times \mathcal{S}^{1\times n} \times \mathcal{S}^{n\times 1}$.

The series $S(\Phi)$ is the image of Φ by the morphism from the free distributive lattice over $\mathcal{S} \times \mathcal{S}^{1\times n} \times \mathcal{S}^{n\times 1}$ into the distributive lattice $(\mathcal{S}\langle\langle\Sigma\rangle\rangle, +, \cap)$ obtained by mapping the atom (c, λ, γ) to $S(c, \lambda, \gamma)$.

The pair (μ, Φ) is by definition a *multi-representation* of S. We call μ the *base* and Φ the *acceptance formula* of the multi-representation (μ, Φ). The series S is said to be *recognized* by (μ, Φ).

A series is *pseudo-recognizable* if it has a multi-representation. As it is shown in the next examples, pseudo-recognizable series are not necessarily recognizable.

Examples 7.

- Let $\Sigma = \{a, b\}$ and $\Xi = \{c\}$. Consider series with coefficients in the commutative semiring $\mathcal{P}(\Xi^*)$.

  Let S and T be the series defined as follows: $(S, w) = \{c^{|w|_a}\}$ and $(T, w) = \{c^{|w|_b}\}$, for all $w \in \Sigma^*$.

  Both series are recognizable (see Example 5 and Theorem 4) and admit a linear representation with the same base. Indeed, let $\mu : \Sigma^* \to \mathcal{P}(\Xi^*)^{2\times 2}$ be the morphism defined by

  $$\mu(a) = \begin{pmatrix} \{c\} & \emptyset \\ \emptyset & \{1\} \end{pmatrix} \text{ and } \mu(b) = \begin{pmatrix} \{1\} & \emptyset \\ \emptyset & \{c\} \end{pmatrix},$$

  and let $\lambda_S = (\{1\}\ \emptyset)$, $\gamma_S = \binom{\{1\}}{\emptyset}$, $\lambda_T = (\emptyset\ \{1\})$ and $\gamma_T = \binom{\emptyset}{\{1\}}$. Then $(S, w) = \lambda_S \mu(w) \gamma_S$ and $(T, w) = \lambda_T \mu(w) \gamma_T$.

Now, the infimum of these series is pseudo-recognizable, recognized by (μ, Φ), where $\Phi = (1, \lambda_S, \gamma_S) \wedge (1, \lambda_T, \gamma_T)$. The coefficient of a word w in $S \cap T$ is the intersection of (S, w) and (T, w) (see Example 2). So we have:

$$(S \cap T, w) = \begin{cases} \emptyset & \text{if } |w|_a \neq |w|_b, \\ c^{|w|/2} & \text{if } |w|_a = |w|_b \left(= \frac{|w|}{2}\right). \end{cases}$$

Note that $S \cap T$ has rational coefficients since they are singletons. Hence, by Remark 3, $S \cap T$ is rational if and only if it is a rational transduction.

But the inverse of a rational transduction is a rational transduction (it is easy to see using transducers) and a rational transduction preserves rational languages [1, Corollary III.4.2]. Since the support of a transduction $\Sigma^* \to \Xi^*$ is the inverse image[1] of Ξ^*, the series $S \cap T$ is pseudo-recognizable and not recognizable (here $\mathrm{supp}(S \cap T) = \{w \in \Sigma^* \mid |w|_a = |w|_b\}$, which is not recognizable [2, Example III.3.1]).

- Consider series on $\Sigma = \{a, b\}$ with coefficients in $\mathbb{N}_{\min}$. We take series S and T as follows: $(S, w) = |w|_a$ and $(T, w) = |w|_b$, for all $w \in \Sigma^*$. Both series are recognizable. We can obtain some multi-representations for them from the previous example. The infimum of these series is pseudo-recognizable. Now, we saw in Example 2 that the infimum on $\mathbb{N}_{\min}$ is the maximum on $\mathbb{N}$, so $(S \cap T, w) = |w|_a \cap |w|_b = \max(|w|_a, |w|_b)$. But this series is not recognizable on $\mathbb{Z}_{\min}$ [8], and hence cannot be recognizable on $\mathbb{N}_{\min}$.

Remark 11. As shown in [7], if the semiring of coefficients is finite, then a formal series is pseudo-recognizable if and only if it is recognizable.

Note that, in general, there exist series which are not pseudo-recognizable. For example, if the semiring of coefficients is countable, there is a countable number of pseudo-recognizable series, but the set of series is not countable.

Example 8. It is easy to see, as a special case of Remark 11, that the pseudo-recognizable series with coefficients in $\mathbb{B}$ are recognizable. Indeed, these series can be identified with their supports and the infimum of two languages is their intersection. In particular, the series $\sum_{|w|_a = |w|_b} w$ is not pseudo-recognizable in $\mathbb{B}\langle\langle\Sigma\rangle\rangle$ [2, Example III.3.1].

Proposition 12. The set of pseudo-recognizable series is closed under (finite) addition and (finite) infimum.

We denote by $\mathcal{S}_{\mathrm{PsRec}}\langle\langle\Sigma\rangle\rangle$ the set of pseudo-recognizable series on Σ, with coefficients in $\mathcal{S}$.

[1]The inverse image of a language L by a transduction $S : \Sigma^* \to \Xi^*$ is the set $S^{-1}(L) = \{w \in \Sigma^* \mid (S, w) \cap L \neq \emptyset\}$, so $S^{-1}(\Xi^*) = \{w \in \Sigma^* \mid (S, w) \cap \Xi^* \neq \emptyset\} = \{w \in \Sigma^* \mid (S, w) \neq \emptyset\} = \mathrm{supp}\, S$.

4.2. Why using multi-representations?

Multi-representations can be used to solve certain equations on formal series. Consider the following problem: let A, B and K be recognizable series on an idempotent semiring $\mathcal{S}$, what can be said about the supremal series $X \leq K$ such that $AX \leq X + B$? It is shown in [7] that this series exists, is recognizable if $\mathcal{S}$ is finite (with a constructive proof), and is pseudo-recognizable if $\mathcal{S} = \mathbb{N}_{\min}$ and A is a language.

5. REPRESENTATION IN THE PIN–SAKAROVITCH WAY

5.1. Non linear representations

This subsection is tightly inspired by a paper of J.-E. Pin and J. Sakarovitch [9].

Let $(M, 1_M)$ and $(N, 1_N)$ be monoids. The *free product* (or *coproduct*) $M * N$ of monoids M and N is the quotient $(M \cup N)^*/\mathcal{R}$, where $\mathcal{R}$ is the set of relations:

$$\mathcal{R} = \{m\cdot m' = mm', n\cdot n' = nn',\ m\cdot 1_N = m, 1_N\cdot m = m,\\ n\cdot 1_M = n,\ 1_M\cdot n = n \,|\, m, m' \in M,\ n, n' \in N\}.$$

We can identify $M * N$ to the elements of the form $m_0 n_1 m_1 \cdots n_r m_r$ with $m_0, \ldots, m_r \in M$ and $n_1, \ldots, n_r \in N$ with the product

$$(m_0 n_1 m_1 \cdots n_r m_r)(m'_0 n'_1 m'_1 \cdots n'_{r'} m'_{r'})\\ = m_0 n_1 m_1 \cdots n_r (m_r m'_0) n'_1 m'_1 \cdots n'_{r'} m'_{r'}.$$

This operation provides $M * N$ with a monoid structure. The set $\mathcal{P}(M * N)$ has a semiring structure inherited from the monoid structure of $M * N$ (see Example 1).

Let $\mathcal{S}$ be a complete idempotent semiring and $\Omega = \{\omega_1, \ldots, \omega_r\}$ be an alphabet. We denote by $\mathcal{S} * \Omega^*$ the free product of $\mathcal{S}$ and Ω^* provided with their multiplicative structure. If $\tau \in \mathcal{S} * \Omega^*$, we call *specialization of τ in the r-tuple* $(s_1, \ldots, s_r) \in \mathcal{S}^r$ the image of τ by the morphism from $\mathcal{S} * \Omega^*$ into $(\mathcal{S}, \cdot)$ which associates s_i to ω_i, $1 \leq i \leq r$. This element is denoted by $\tau(s_1, \ldots, s_r)$.

If σ belongs to $\mathcal{P}(\mathcal{S}*\Omega^*)$, we call *specialization of σ in the r-tuple* $(s_1, \ldots, s_r) \in \mathcal{S}^r$ the image of σ by the morphism of complete idempotent semirings, from $\mathcal{P}(\mathcal{S} * \Omega^*)$ into $\mathcal{S}$, which associates s_i to ω_i, $1 \leq i \leq r$. More concretely, it is the element

$$\sigma(s_1, \ldots, s_r) = \sum_{\tau \in \sigma} \tau(s_1, \ldots, s_r).$$

A formal series S is *representable* if and only if there exist an integer $n \geq 1$, a morphism $\mu : \Sigma^* \to \mathcal{S}^{n\times n}$ and an element $\sigma \in \mathcal{P}(\mathcal{S} * \Omega^*)$, where Ω is a n^2-letters alphabet, $\Omega = \{\varepsilon_{11}, \ldots, \varepsilon_{ij}, \ldots, \varepsilon_{nn}\}$, such that for all words w

$$(S, w) = \sigma\Big(\big(\mu(w)\big)_{11}, \ldots, \big(\mu(w)\big)_{ij}, \ldots, \big(\mu(w)\big)_{nn}\Big).$$

The couple (μ, σ) is a *non linear representation* of the series S.

A series is *representable by a singleton* if it has a non linear representation (μ, σ) where σ is a singleton of $\mathcal{P}(\mathcal{S} * \Omega^*)$.

We denote by $\mathcal{S}_{\text{Rep}}\langle\langle\Sigma\rangle\rangle$ the set of all representable series on Σ, with coefficients in $\mathcal{S}$.

Example 9. A recognizable series is representable.

Example 10. Consider the following transduction from $\Sigma = \{a\}$ into $\Xi = \{b\}$:

$$(S, a^n) = \sum_{m \in M} b^{mn}, \text{ where } M \text{ is any set of integers.}$$

This transduction is representable in $\mathcal{P}(\Xi^*)\langle\langle\Sigma\rangle\rangle$. Indeed, let us consider the one-letter-alphabet $\Omega = \{\omega\}$, the element $\sigma = \sum_{m \in M} \omega^m$ of $\mathcal{P}\big(\mathcal{P}(\Xi^*) * \Omega^*\big)$ and the morphism $\mu : \Sigma^* \to \mathcal{P}(\Xi^*) : a \mapsto \{b\}$. We have $(S, w) = \sigma\big(\mu(w)\big)$.

Proposition 13. The series

$$(S, a^n) = \sum_{m \in M} b^{mn}, \text{ where } M \text{ is any infinite set of integers,}$$

is not recognizable.

Proof. By contradiction: assume that S is recognizable and M is infinite.

We denote by u_q the word a^q (for q in $\mathbb{N}$) and by B_q the coefficient of u_q in S, i.e.

$$B_q = \{b^{mq} \mid m \in M\}.$$

Note that words in B_q have as a length a multiple of q.

Since the series S is recognizable, according to Proposition 5, it belongs to a stable, finite generated left submodule of $\mathcal{S}\langle\langle\Sigma\rangle\rangle$, say $\mathcal{M}$. We denote by $(S_i)_{i \in I}$ a finite generating family of $\mathcal{M}$.

The submodule $\mathcal{M}$ is stable and so for each positive integer q, the series $u_q^{-1}S$ belongs to $\mathcal{M}$. Hence, there exist families $(\alpha_{q,i})_{i \in I}$ of elements of $\mathcal{P}(\Xi^*)$ such that

$$\forall\, q \in \mathbb{N},\ u_q^{-1}S = \sum_{i \in I} \alpha_{q,i} S_i.$$

That is, for the coefficient of a^n:

$$\forall\, q \in \mathbb{N},\ \forall\, n \in \mathbb{N},\ \sum_{i \in I} \alpha_{q,i}(S_i, a^n) = (u_q^{-1}S, a^n) = (S, a^{n+q}) = B_{n+q}. \tag{1}$$

Let us have a look to the set I.

For $i \in I$, we set

$$Q_i = \{q \mid \alpha_{q,i} \neq \emptyset\} \text{ and } N_i = \{n \mid (S_i, a^n) \neq \emptyset\}.$$

If Q_i is finite, we denote by $\bar{q}_i$ its maximal element and in the same way, if N_i is finite, we denote by $\bar{n}_i$ its maximal element. We set

$$\bar{q} = \max\{\bar{q}_i, i \in I \mid Q_i \text{ finite}\} \text{ and } \bar{n} = \max\{\bar{n}_i, i \in I \mid N_i \text{ finite}\},$$

with the convention $\max(\emptyset) = 0$.

Note that if an element i of I is such that Q_i is finite and $q > \bar{q}$, then $\alpha_{q,i} = \emptyset$. In the same way, if N_i is finite and $n > \bar{n}$, then $(S_i, a^n) = \emptyset$.

We denote by J the following subset of I:

$$J = \{i \in I \mid Q_i \text{ and } N_i \text{ are infinite}\}.$$

According to the above remark, we have for $q > \bar{q}$ and $n > \bar{n}$:

$$\sum_{i \in I} \alpha_{q,i}(S_i, a^n) = \sum_{i \in J} \alpha_{q,i}(S_i, a^n).$$

Equation (1) now becomes:

$$\forall\, q > \bar{q},\, \forall\, n > \bar{n},\ \sum_{i \in J} \alpha_{q,i}(S_i, a^n) = B_{n+q}. \tag{2}$$

Let us fix some i in J: Q_i is infinite by definition of J.

For all q in Q_i, let v_q be a word of $\alpha_{q,i}$.

From Equation (2),

$$v_q(S_i, a^n) \subseteq B_{n+q}.$$

In particular, the length of any word of $v_q(S_i, a^n)$ is a multiple of $n + q$. If (S_i, a^n) is not the empty language, let x and y be two of its words.

For each element q of Q_i, $n + q$ divides $|v_q x|$ and $|v_q y|$, hence $n + q$ divides $||x| - |y||$. Since the set Q_i is infinite, $||x| - |y||$ can be divided by an infinity of integers, and so it is 0. Hence x and y have the same length. As we are dealing with a one-letter-alphabet, x and y are the same word.

We conclude that the set (S_i, a^n) is either empty or a singleton for $i \in J$. Using similar arguments, we can prove that the set $\alpha_{q,i}$ is either empty or a singleton for $i \in J$. The left term of Equation (2) is a finite sum of empty sets and singleton sets and the right one is infinite for $n + q > 0$, hence we have a contradiction. □

We saw in Example 7 that rational transduction preserve rational languages [1, Corollary III.4.2]. Here is a generalization to representable transductions.

Theorem 14. [9, Corollary 5.3] Let Ξ be an alphabet. The inverse image of a recognizable language on Ξ by a representable transduction (from Σ^* into Ξ^*) is a recognizable language of Σ^*, and hence a rational language of Σ^*.

5.2. Representable vs. pseudo-recognizable

5.2.1. General case

Theorem 14 says, in particular, that the inverse image of Ξ^* by a representable transduction, is a rational set of Σ^*. But the support of the transduction proposed in Example 7 is the language $\{w \in \Sigma^* \mid |w|_a = |w|_b\}$ which is not rational (see Example 8). So

$$\mathcal{P}(\Xi^*)_{\mathrm{PsRec}}\langle\langle\Sigma\rangle\rangle \not\subseteq \mathcal{P}(\Xi^*)_{\mathrm{Rep}}\langle\langle\Sigma\rangle\rangle.$$

Proposition 15. The series

$$(S, a^n) = \sum_{m \in M} b^{mn}, \text{ where } M \text{ is any infinite set of integers,}$$

is not pseudo-recognizable.

P r o o f. The proof of this proposition is quite similar to the proof of Proposition 13.

We denote by $\mathcal{S}$ the semiring of coefficients, that is $\mathcal{P}(\{b\}^*)$ and, as in the proof of Proposition 13, by u_q the word a^q and by B_q its coefficient in S: $B_q = \{b^{mq} \mid m \in M\}$.

Assume S is pseudo-recognizable. Let (μ, Φ) be a multi-representation of S. We can put Φ in normal form. In other words, we write Φ as

$$\Phi = \bigvee_{E \in \mathcal{P}'} \left(\bigwedge_{(c,\lambda,\gamma) \in E} (c, \lambda, \gamma) \right),$$

where $\mathcal{P}'$ is a finite subset of $\mathcal{P}_f(\mathcal{S} \times \mathcal{S}^{1\times n} \times \mathcal{S}^{n\times 1})$.

By Corollary 10, we know that left-cut has no influence on rationality, because the semiring $\mathcal{S}$ is commutative (see Example 1). So we can write:

$$\Phi = \bigvee_{E \in \mathcal{P}} \left(\bigwedge_{(\lambda,\gamma) \in E} (1, \lambda, \gamma) \right),$$

where $\mathcal{P}$ is a finite subset of $\mathcal{P}_f(\mathcal{S}^{1\times n} \times \mathcal{S}^{n\times 1})$.

Hence the series S can be written as:

$$S = \sum_{E \in \mathcal{P}} \left(\bigcap_{(\lambda,\gamma) \in E} S(1, \lambda, \gamma) \right), \tag{3}$$

where the sum is done over a finite set.

Let us consider a particular E in $\mathcal{P}$. From Equation (3), we obtain:

$$\bigcap_{(\lambda,\gamma) \in E} S(1, \lambda, \gamma) \leq S. \tag{4}$$

To simplify notation, we assume that E has two elements. Equation (4) becomes $T \cap U \leq S$, where T and U are rational series.

By the algebraic characterization of Proposition 5, we know that T (resp. U) belongs to a stable, finite generated left submodule of $\mathcal{S}\langle\langle\Sigma\rangle\rangle$, say $\mathcal{M}$ (resp. $\mathcal{N}$). We denote by $(T_i)_{i\in I}$ (resp. $(U_j)_{j\in J}$) a finite generating family of $\mathcal{M}$ (resp. $\mathcal{N}$).

The submodules $\mathcal{M}$ (resp. $\mathcal{N}$) is stable, and so for all positive integer q, the series $u_q^{-1}T$ (resp. $u_q^{-1}U$) belongs to $\mathcal{M}$ (resp. $\mathcal{N}$). Hence, there exist families $(\alpha_{q,i})_{i\in I}$ (resp. $(\beta_{q,j})_{j\in J}$)) of elements of $\mathcal{S}$ such that

$$\forall\, q \in \mathbb{N},\ u_q^{-1}T = \sum_{i\in I} \alpha_{q,i}T_i \text{ and } u_q^{-1}U = \sum_{j\in J} \beta_{q,j}U_j.$$

That is, for the coefficient of a^n:

$$\forall\, q \in \mathbb{N},\ \forall\, n \in \mathbb{N},\ \left(\sum_{i\in I} \alpha_{q,i}(T_i, a^n)\right) = (u_q^{-1}T, a^n) = (T, a^{n+q}),$$

$$\text{and } \left(\sum_{j\in J} \beta_{q,j}(U_j, a^n)\right) = (u_q^{-1}U, a^n) = (U, a^{n+q}).$$

And so for all integers q and n, we have:

$$\left(\sum_{i\in I} \alpha_{q,i}(T_i, a^n)\right) \cap \left(\sum_{j\in J} \beta_{q,j}(U_j, a^n)\right) = (T, a^{n+q}) \cap (U, a^{n+q}) \subseteq B_{n+q}. \tag{5}$$

For $i \in I$, we set

$$QN_i = \{(q,n) \mid (\alpha_{q,i}(T_i, a^n)) \cap (U, a^{n+q}) = \emptyset,\ i \in I\}$$

If QN_i is finite, we denote by $\bar{q}_i$ the maximal q that belongs to it and by $\bar{n}_i$ the maximal n. We set:

$$\bar{q} = \max\{\bar{q}_i, i \in I \mid QN_i \text{ finite}\} \text{ and } \bar{n} = \max\{\bar{n}_i, i \in I \mid QN_i \text{ finite}\}.$$

With a similar argument as in the proof of Proposition 13, we can prove that for $q > \bar{q}$ and $n > \bar{n}$, the pieces of the sets $\alpha_{q,i}$ and (T_i, a^n) that enter into Equation (5) are either empty sets or singleton ones. So the left member of Equation (5) is a finite set.

The same conclusion can be drawn for all sets E which appear in Equation (4). Since the sum in Equation (3) is finite and the set B_{n+q} is infinite for q and n large enough, we have a contradiction. □

Corollary 16.

$$\mathcal{P}(\Xi^*)_{\mathrm{Rep}}\langle\langle\Sigma\rangle\rangle \not\subseteq \mathcal{P}(\Xi^*)_{\mathrm{PsRec}}\langle\langle\Sigma\rangle\rangle.$$

Hence there is no inclusion relation between the set $\mathcal{P}(\Xi^*)_{\mathrm{Rep}}\langle\langle\Sigma\rangle\rangle$ and the set $\mathcal{P}(\Xi^*)_{\mathrm{PsRec}}\langle\langle\Sigma\rangle\rangle$.

5.2.2. A particular case: the coefficients belong to a finite semiring

In this section, we show that a series on a finite semiring $\mathcal{S}$ is representable if and only if it is rational, so if and only if it is pseudo-recognizable, according to Remark 11.

Let $n \in \mathbb{N}$ and $\Omega = \{\omega_{11}, \ldots, \omega_{nn}\}$ be an alphabet. For all $\sigma \in \mathcal{P}(\mathcal{S} * \Omega^*)$, we call *specialization* of σ the subset of all specializations of σ in all the n^2-tuple of $\mathcal{S}$. We denote it by $\operatorname{Spec}\sigma$: $\operatorname{Spec}\sigma = \{\sigma(s_{11}, \ldots, s_{nn}) \mid (s_{11}, \ldots, s_{nn}) \in \mathcal{S}^{n^2}\}$.

Proposition 17. If $\mathcal{S}$ is finite, the set of specializations of all the elements of $\mathcal{S} * \Omega^*$ is finite.

Corollary 18. A representable series (on a finite semiring) is necessarily representable by an element of $\mathcal{P}_f(\mathcal{S} * \Omega^*)$.

Now, it is sufficient to prove that a series representable by a singleton is recognizable. Since the set of representable series is stable under addition, we will be able to conclude.

Lemma 19. A series representable by a singleton is recognizable.

Proof. Such a series is the Hadamard product of series recognized by singletons of the form $\{s_1 \omega_{ij} s_2\}$, $s_1, s_2 \in \mathcal{S}$, $\omega_{ij} \in \Omega$, which represent clearly recognizable series. Since $\mathcal{S}$ is finite, such a series is recognizable by Proposition 7. □

As a result, we obtain the following theorem:

Theorem 20. On a finite semiring, a series is representable if and only if it is recognizable.

6. A MORE GENERAL REPRESENTATION

In this section, we suggest a common generalization of both notions of pseudo-recognizability and representability. We work with series on Σ, with coefficients in a complete idempotent semiring $\mathcal{S}$.

We consider such a representation: let $n \geq 1$ be an integer, $\mu : \Sigma^* \to \mathcal{S}^{n \times n}$ be a morphism and Φ be a positive boolean formula on $\mathcal{S} \times \mathcal{P}(\mathcal{S} * \Omega^*)$, where $\Omega = \{\omega_{11}, \ldots, \omega_{nn}\}$ is an alphabet. We call *atoms* the elements $(c, \sigma) \in \mathcal{S} \times \mathcal{P}(\mathcal{S} * \Omega^*)$. The series recognized by such a representation, which we will call *pseudo-representable*, is the image of Φ by the morphism from the free distributive lattice on $\mathcal{S} \times \mathcal{P}(\mathcal{S} * \Omega^*)$ in the distributive lattice $\mathcal{S}\langle\langle\Sigma\rangle\rangle$ which associates to the atom (c, σ) the series whose value on a word w is $c \backslash \sigma\big((\mu(w))_{11}, \ldots, (\mu(w))_{nn}\big)$.

To obtain a representable series, it is sufficient to reduce Φ to an atom $(1, \sigma)$. To obtain a pseudo-recognizable series, it is sufficient for each σ which appears in the acceptance formula to be of the form $\{s_1 \omega_{ij} s_2\}$, $s_1, s_2 \in \mathcal{S}$, $\omega_{ij} \in \Omega$. Thus this is really a generalization of both notions. Furthermore, we have the following theorem.

Theorem 21. If $\mathcal{S}$ is finite, the pseudo-representable series are recognizable.

7. CONCLUSION

These representations allow us to go beyond representability. It could be interesting to see if, for recognizable series, they provide smaller representations than linear ones.

(Received December 9, 2002.)

REFERENCES

[1] J. Berstel: Transductions and Context–Free Languages. Teubner, Stuttgart 1979.
[2] J. Berstel and C. Reutenauer: Les séries rationnelles et leurs langages. Masson, Paris 1984. English translation: Rational Series and Their Languages, Springer–Verlag, Berlin 1988.
[3] T. S. Blyth and M. F. Janowitz: Residuation Theory. Pergamon Press, Oxford 1972.
[4] S. Eilenberg: Automata, Languages and Machines, vol. A. Academic Press, New York 1974.
[5] J. Gunawardena: An introduction to idempotency, in idempotency. Chapter 1 (J. Gunawardena, ed.), Cambridge University Press, Cambridge 1998.
[6] I. Klimann: New types of automata to solve fixed point problems. Theoret. Comput. Sci. *259* (2001), 1–2, 183–197.
[7] I. Klimann: A solution to the problem of (A,B)-invariance for series. Theoret. Comput. Sci. *293* (2003), 1, 115–139.
[8] N. Kobayashi: The closure under division and a characterization of the recognizable $\mathcal{Z}$-subsets. RAIRO Inform. Théor. Appl. *30* (1996), 3, 209–230.
[9] J.-E. Pin and J. Sakarovitch: Une application de la représentation matricielle des transductions. Theoret. Comp. Sci. *35* (1985), 271–293.
[10] A. Salomaa and M. Soittola: Automata–Theoretical Aspects of Formal Power Series. Springer–Verlag, Berlin 1978.

Dr. Ines Klimann, LIAFA – Université Paris 7 – Case 7014, 2, place Jussieu F-75251 Paris Cedex 5. France.
e-mail: klimann@liafa.jussieu.fr