Derivatives of rational expressions with multiplicity

Sylvain Lombardy Jacques Sakarovitch *

Abstract

This paper introduces a generalization of the *partial derivatives* of rational expressions, due to Antimirov, to rational expressions with multiplicity. We define the derivation of a rational expression with multiplicity in such a way that the result is a *polynomial of expressions*. This amount to interpreting the addition symbol at the upper level in the semiring of coefficients.

Former result of Brzozoski and of Antimirov are then expressed in that framework that allows to deal with rational power series, and automata and expressions with multiplicity as well.

Résumé

Ce papier présente une généralisation des *dérivées partielles* d'expressions rationnelles, dues à Antimirov, aux expressions avec multiplicité. La derivation d'une expression rationnelle avec multiplicité est définie de sorte que la dérivée est un *polynôme d'expressions*. Ce qui revient à interpréter, le symbole d'addition, s'il est l'opérateur externe de l'expression, comme l'addition dans le semi-anneau des coefficient.

Les résultats antérieurs d'Antimirov et de Brzozowski, exprimés dans ce cadre, peuvent se généraliser aux séries rationnelles, aussi bien qu'aux automates et aux expressions avec multiplicité

1 Introduction

The purpose of this paper is to generalize the definition and constructions, due to V. Antimirov ([1]), of the so-called *partial derivatives* — and that we shall call here *derived terms* — from rational expressions and languages to rational expressions and languages *with multiplicity*.

In 1964, J. Brzozowski defined the *derivatives* of a rational expression ([3]). He showed that, *modulo* the axioms of associativity, commutativity, and idempotency of the addition (on the set of words) — the ACI-properties — the set of derivatives of a given expression is finite, yielding both a new proof for (one direction of) Kleene's Theorem and an algorithm turning an expression into a deterministic finite automaton.

This problem (of turning an expression into a finite automaton) has attracted much attention since the beginning of the theory ([7, 11]) and is an area of active research since then. In 1995, V. Antimirov made a fundamental contribution by defining the "*partial*

*Laboratoire Traitement et Communication de l'Information (C. N. R. S. URA 820), E. N. S. T., Paris.

derivatives" of an expression ([1]). Using his own words, "the idea behind [his] construction is that it allows to take into account the ACI-properties of only those occurrences of "+" in [rational] terms which appear at the very upper level". Roughly speaking, the derivation proposed by Antimirov has two effects. First, it performs the "normal" derivation and, second, it breaks the result into "parts", hence the name *partial derivatives*, such that this result is the sum of the parts. This construction, which still keeps "a touch of magic", has a number of outcomes: the number of partial derivatives is not only finite but "small": smaller than or equal to the number of letters in the expression; they are easier to compute than the classical derivatives and they yield a *non-deterministic* finite automaton with (almost) the same number of state; finally, the subset construction applied to that automaton gives back the deterministic one computed by Brzozowski's algorithm. The computation of Antimirov's automaton has been made really efficient in [5].

What is presented here is at the same time a *formalization* and a *generalization* of Antimirov's construction. We first define the rational expression with multiplicity in a semiring $\mathbb{K}$. For sake of simplicity in dealing with those expressions, we suppose that $\mathbb{K}$ is *commutative*, although most of our construction and formulae are independent of this assumption.

We then define the derivation of a $\mathbb{K}$-expression with respect to a letter and then to a word. The main feature of our definition, that indeed realizes Antimirov's main idea, is that the result of the derivation of a $\mathbb{K}$-expression is not a $\mathbb{K}$-expression anymore but a *polynomial* of certain $\mathbb{K}$-expressions (which we call *derived terms*) with coefficients in $\mathbb{K}$.

The generalization of Antimirov's results is then straightforward. The derived terms are exactly the partial derivatives, and they are the state of an automaton — an automaton with multiplicity, of course — which recognizes the series that is denoted by the expression. In particular, let us note that this technique overcomes the problem that the addition is not idempotent anymore, and that Brzozowski's theorem does not hold in that setting.

This is an example of the interest in taking multiplicities and generalization to series into account for a simplification and a better understanding of constructions and results on languages.

Let us note that Caron and Flouret [4] have generalized the Glushkov construction to expressions with multiplicity, yielding a $\mathbb{K}$-automaton that plays the same role with respect to ours as the so-called *position automaton* do with respect to Antimirov's automaton in the Boolean case.

In [9], Krob has considered the derivation of rational expressions with multiplicity (in a commutative semiring) with another problem in mind. He has characterized the set of axioms such that the derivatives of identities remain identities.

The paper is organized as follow: in the next sections we define rational expressions with multiplicity, the series they denote and describe a set of "trivial" identities necessary to allow reasonnable computation.

In section 3, after the definition of derivatives, we state the main theorem; it says that the derivative of an expression with respect to a word is a linear combination of expressions taken in a finite set independant of this word. This allows to build, in the following part, a finite automaton based on derivatives. The number of states of this automaton, aside the initial state, is equal to the *litteral length* of the expression.

We study then a phenomenon that can arise if the coefficients belong to a semiring which is not positive; in this case, the computation may generate some useless derived terms —called *shadow terms*. Finally, we give some variations on the definition of the derivative that can prove to be more suitable in some cases.

2 Rational expressions with multiplicity

Let A be a finite alphabet and let $\mathbb{K}$ be a semiring. The addition of $\mathbb{K}$, *associative* and *commutative*, is denoted by $\oplus$, its multiplication simply by concatenation.

Power series The semiring of formal power series over A^* with multiplicity in $\mathbb{K}$ is denoted by $\mathbb{K}\langle\langle A^*\rangle\rangle$. The inherited *commutative* addition in $\mathbb{K}\langle\langle A^*\rangle\rangle$ is denoted by $\oplus$, the (Cauchy) product by concatenation. For (rational) power series, their definitions, their notations, the related results, we refer to [2] which we basically follow.

But for one point: in [2], the semiring of coefficients is always equipped with the *discrete* topology; as it can easily be seen, this is an unnecessary assumption and $\mathbb{K}$ may well be equipped with *any* (metric or order) topology. For instance, in $\mathbb{Q}$, with the usual topology, $(1/2)^* = 1 + 1/2 + (1/2)^2 + \cdots = 2$ (*cf.* [10, 12]). The semiring $\mathbb{K}\langle\langle A^*\rangle\rangle$ is then supposed to be equipped with the product topology derived from the topology on $\mathbb{K}$.

The coefficient of a word f of A^* in a series s of $\mathbb{K}\langle\langle A^*\rangle\rangle$ is denoted by $<s, f>$. The *constant term* of s is the coefficient of the empty word 1_{A^*} in s and is denoted by $\mathsf{c}(s)$. A series is *proper* if $\mathsf{c}(s) = 0_{\mathbb{K}}$. The *proper part* of s is the proper series, denoted by s_p, which coincide with s on all words f different from 1_{A^*} and one can write $s = \mathsf{c}(s)\,1_{A^*} \oplus s_p$. The star of a proper series is always well defined. The star of a series which is not proper may be defined or not, and the decision between the both cases is given by the following proposition:

Lemma 1 [2, Exer I.3.4] *The star of a series s is defined iff the star of $\mathsf{c}(s)$ is defined (in $\mathbb{K}$) and it holds:*

$$s^* = (\mathsf{c}(s)^*\, s_p)^*\, \mathsf{c}(s)^* \quad . \qquad \blacksquare$$

Rational expressions From now on, and for a reason we shall explain later, *we suppose that the semiring $\mathbb{K}$ is commutative.*

The definition of *rational expressions over* $\mathbb{K}$ goes as the one of classical rational expressions: it amounts to the construction of a set of well-formed formulae.

Let $\{\mathsf{0}, \mathsf{1}, +, \cdot, *\}$ be a set of operations. Naturally, $+$ and $\cdot$ are binary, $*$ is unary, and $\mathsf{0}$ and $\mathsf{1}$ are "0-ary" operations, that is, they are constants. Moreover, for every k in $\mathbb{K}$, there is a unary operation, again denoted by k.

i) $\mathsf{0}$, $\mathsf{1}$, and a, for every a in A, are rational expressions (the *atomic* expressions or formulae).

ii) If E is a rational expression and k is in $\mathbb{K}$, then $(k\,\mathsf{E})$ is a rational expression.

iii) If E and F then $(\mathsf{E}+\mathsf{F})$, $(\mathsf{E}\cdot\mathsf{F})$, and (E^*) are rational expressions.

We denote by $\mathbb{K}\,\mathsf{RatE}\,A^*$ the set of rational expressions over A with multiplicity in $\mathbb{K}$.

The complexity of a rational expression can be described by different parameters. The *litteral length*, denoted by $\ell(\mathsf{E})$, is the number of atomic formulae in E that are letters (*e.g.* $\ell((1/3\,(a^*){+}1/6\,(b^*))^*{+}(2\,1)) = 2$). Remark that $\ell((\mathsf{E}+\mathsf{F})) = \ell((\mathsf{E}\cdot\mathsf{F})) = \ell(\mathsf{E})+\ell(\mathsf{F})$ and that $\ell((\mathsf{E}^*)) = \ell((k\,\mathsf{E})) = \ell(\mathsf{E})$. This is the parameter on which the automaton we build in section 4 depends. Whereas, many of our proofs are actually by induction on the depth of an rational expression. The *depth*[1] of an expression is inductively defined by:

$$\begin{aligned} \mathsf{d}(\mathsf{0}) = \mathsf{d}(\mathsf{1}) &= 0, \\ \forall a\in A \qquad \mathsf{d}(\mathsf{a}) &= 0, \\ \mathsf{d}((k\,\mathsf{E})) = \mathsf{d}((\mathsf{E}^*)) &= 1+\mathsf{d}(\mathsf{E}), \\ \mathsf{d}((\mathsf{E}\cdot\mathsf{F})) = \mathsf{d}((\mathsf{E}+\mathsf{F})) &= 1+\max(\mathsf{d}(\mathsf{E}),\mathsf{d}(\mathsf{F})). \end{aligned}$$

The constant term of an expression E is defined by induction on the depth of the expression E:

$$\mathsf{c}((k\,\mathsf{E})) = k\ \mathsf{c}(\mathsf{E})\,, \quad \mathsf{c}((\mathsf{E}+\mathsf{F})) = \mathsf{c}(\mathsf{E})\oplus\mathsf{c}(\mathsf{F})\,, \quad \mathsf{c}((\mathsf{E}\cdot\mathsf{F})) = \mathsf{c}(\mathsf{E})\ \mathsf{c}(\mathsf{F})$$
$$\text{and} \quad \mathsf{c}((\mathsf{E}^*)) = \mathsf{c}(\mathsf{E})^* \quad \text{iff the latter is defined in } \mathbb{K}.$$

A rational expression E is a formula. It can either be *valid* and denotes a rational series, or not. We say that an expression is valid if $\mathsf{c}(\mathsf{E})$ is defined.

The series *denoted by a valid expression* E, which we note as $|\mathsf{E}|$, is defined by induction on the depth of the expression E as well:

$$|\mathsf{0}| = 0_{\mathbb{K}}\,, \quad |\mathsf{1}| = 1_{A^*}\,, \quad |a| = a\,, \quad \text{for every } a \text{ in } A\,, \quad |(k\,\mathsf{E})| = k\,|\mathsf{E}|\,,$$
$$|(\mathsf{E}+\mathsf{F})| = |\mathsf{E}|\oplus|\mathsf{F}|\,, \quad |(\mathsf{E}\cdot\mathsf{F})| = |\mathsf{E}|\,|\mathsf{F}|\,, \quad \text{and} \quad |(\mathsf{E}^*)| = |\mathsf{E}|^*.$$

Two valid expressions are *equivalent* if they denote the same series.

The definition of the constant term of an expression is consistent with the one of series denoted by an expression. Actually, $\mathsf{c}(\mathsf{E})$ and $|\mathsf{E}|$ are defined by the same induction. Moreover the constant term of an expression and the constant term of the series denoted by this expression are identical on the atomic expressions. Hence $\mathsf{c}(\mathsf{E})$ is equal to $\mathsf{c}(|\mathsf{E}|)$. The last equation is besides made consistent by Lemma 1: $|(\mathsf{E}^*)|$ is defined iff $\mathsf{c}(\mathsf{E})^*$ is defined thus iff (E^*) is valid.

Remark that Lemma 1 allows to define the star of a series with a constant term denoted by an expression E without requiring any information on the expression that denotes $|\mathsf{E}|_p$.

[1] We choose “depth ” rather than “height ” in order to avoid any confusion with the “star height ” of a rational expression (that we are dealing with in another paper).

Examples 1 : In the forthcoming examples, $A = \{a, b\}$ and $\mathbb{K} = \mathbb{Q}$.

i) $\mathsf{E}_1 = (((\frac{1}{6}\,a^*) + (\frac{1}{3}\,b^*))^*)$. Let $\mathsf{F}_1 = ((\frac{1}{6}\,a^*) + (\frac{1}{3}\,b^*))$. It comes $\mathsf{c}(\mathsf{F}_1) = \frac{1}{2}$, hence $\mathsf{c}(\mathsf{F}_1)^* = 2$. Thus, although $|\mathsf{F}_1|$ is not proper, the series denoted by E_1 is well-defined.

ii) $\mathsf{E}_2 = (a\,b\,a + (a\,(a - b\,a)))$.

For simplicity, we write $a\,b$ for $(a \cdot b)$, $a\,b\,a$ for $((a \cdot b) \cdot a)$ and $(a - b\,a)$ for $(a + (-1_{\mathbb{K}}\,(b \cdot a)))$.

iii) $\mathsf{E}_3 = 5\,((2\,a\,b) + ((3\,b) \cdot (4\,(a\,b)^*)))^*$ (This is the running example in [4].) □

Trivial identities The following identities trivially hold on rational expressions with multiplicity:

$$(k\,\mathsf{0}) \equiv \mathsf{0}\,, \quad (0_{\mathbb{K}}\,\mathsf{E}) \equiv \mathsf{0}\,, \quad (\mathsf{0} \cdot \mathsf{E}) \equiv (\mathsf{0} \cdot \mathsf{E})\mathsf{0}\,, \tag{1}$$

$$\mathsf{0} + \mathsf{E} \equiv \mathsf{E} + \mathsf{0} \equiv \mathsf{E}\,, \quad (1_{\mathbb{K}}\,\mathsf{E}) \equiv \mathsf{E}\,, \quad (k\,\mathsf{1}) \cdot \mathsf{E} \equiv \mathsf{E} \cdot (k\,\mathsf{1}) \equiv (k\,\mathsf{E}). \tag{2}$$

The first reason why the semiring $\mathbb{K}$ has been supposed to be *commutative* is to keep the definition of rational expressions with multiplicity simple enough, while the basic following property still holds:

Proposition 2 *A series of $\mathbb{K}\langle\langle A^* \rangle\rangle$ is $\mathbb{K}$-rational iff it is denoted by a rational expression with multiplicity in $\mathbb{K}$.* ■

One can see the trivial identities as rewriting rules and it should be clear that every rational expression is equivalent to an expression in "reduced form" which is unique and which can be computed in a time proportional to the length of the expression, provided the multiplication in $\mathbb{K}$ is seen as an operation with fixed cost.

Remark 1 Note that the trivial identities have nothing to do with the associativity and commutativity axioms for the $+$ operation: $(a+(b+c))$, $((a+b)+c)$ and $(a+(c+b))$ are three different expressions, nor with the associativity of "$\cdot$", in spite of the simplifications we have used in the examples.

3 Derivatives and derived terms

We now introduce polynomials of expressions and their derivatives. The set $\mathbb{K}\langle \mathbb{K}\,\mathsf{RatE}\,A^* \rangle$ of linear combinations of rational expressions, or *polynomials of expressions*, is a $\mathbb{K}$-semimodule; the addition is commutative and the multiplication by an element of $\mathbb{K}$ is distributive:

$$k\,\mathsf{E} \oplus k'\,\mathsf{F} = k'\,\mathsf{F} \oplus k\,\mathsf{E} \tag{3}$$

$$k\,\mathsf{E} \oplus k'\,\mathsf{E} = [k\,k']\mathsf{E} \tag{4}$$

We define a multiplication law on the monomials (*i.e.* on the elements of the base of the semimodule), which is generalized to polynomials by a distributivity axiom:

$$[k\,\mathsf{E}][k'\,\mathsf{F}] \equiv [k\,k']\,(\mathsf{E}\cdot\mathsf{F}) \tag{5}$$

$$\big([\mathsf{E}\oplus\mathsf{E}']\cdot\mathsf{F}\big) \equiv (\mathsf{E}\cdot\mathsf{F})\oplus(\mathsf{E}'\cdot\mathsf{F})\,, \qquad \big(\mathsf{E}\cdot[\mathsf{F}\oplus\mathsf{F}']\big) \equiv (\mathsf{E}\cdot\mathsf{F}')\oplus(\mathsf{E}\cdot\mathsf{F}') \tag{6}$$

In the following, $[k\,\mathsf{E}]$ or $k\,\mathsf{E}$ is a monomial whereas $(k\,\mathsf{E})$ is an expression.

The series denoted by a polynomial of rational expressions is obtained by extending by linearity the interpretation defined on rational expressions.

Remark 2 If $\mathbb{K}$ is not commutative, the interpretation of the left handside and the right handside of the identity (5) may differ. This is the main reason for our assumption of commutativity. However this difficulty can be overcome, and it will be done in a forthcoming work.

The set of polynomials of rational expressions is not a semialgebra. Actually, the multiplication that we define is not associative:

$$[[\mathsf{E}]\,[\mathsf{F}]]\,[\mathsf{G}] = ((\mathsf{E}\cdot\mathsf{F})\cdot\mathsf{G}) \neq (\mathsf{E}\cdot(\mathsf{F}\cdot\mathsf{G})) = [\mathsf{E}]\,[[\mathsf{F}]\,[\mathsf{G}]]$$

However, it does not cause any problem in the framework of derivatives.

3.1 Derivatives

Definition 1 *& Notation* *Let* E *be in* $\mathbb{K}\,\mathsf{RatE}\,A^*$ *and let* a *be in* A. *The* derivative *of* E *with respect to* a, *denoted by* $\frac{\partial}{\partial a}\mathsf{E}$, *is a polynomial of rational expressions with coefficients in* $\mathbb{K}$, *defined inductively by the following formulae.*

$$\frac{\partial}{\partial a}\mathsf{0} = \frac{\partial}{\partial a}\mathsf{1} = \mathsf{0}$$

$$\forall a \in A \qquad \frac{\partial}{\partial a}a = \mathsf{1}\,, \quad \frac{\partial}{\partial a}b = \mathsf{0} \quad \forall b \in A\,,\ b \neq a$$

$$\frac{\partial}{\partial a}(k\,\mathsf{E}) = k\,\frac{\partial}{\partial a}\mathsf{E} \tag{7}$$

$$\frac{\partial}{\partial a}(\mathsf{E}+\mathsf{F}) = \frac{\partial}{\partial a}\mathsf{E}\oplus\frac{\partial}{\partial a}\mathsf{F} \tag{8}$$

$$\frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial}{\partial a}\mathsf{E}\right]\cdot\mathsf{F}\right)\oplus\mathsf{c}(\mathsf{E})\,\frac{\partial}{\partial a}\mathsf{F} \tag{9}$$

$$\frac{\partial}{\partial a}(\mathsf{E}^*) = \mathsf{c}(\mathsf{E})^*\,\left(\left[\frac{\partial}{\partial a}\mathsf{E}\right]\cdot(\mathsf{E}^*)\right) \tag{10}$$

The derivative of a polynomial of expressions is defined by linearity:

$$\frac{\partial}{\partial a}\left(\bigoplus_{i\in I}\mathsf{E}_i\right) = \bigoplus_{i\in I}\frac{\partial}{\partial a}\mathsf{E}_i \tag{11}$$

Implicitely, the (polynomial of) expressions are reduced by trivial identities (*e.g.* if $\mathsf{c}(\mathsf{E}) = 0_{\mathbb{K}}$, then $\frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F})$ is equal to $\big(\big[\frac{\partial}{\partial a}\,\mathsf{E}\big]\cdot\mathsf{F}\big)$ and not to $\big(\big[\frac{\partial}{\partial a}\,\mathsf{E}\big]\cdot\mathsf{F}\big) \oplus 0_{\mathbb{K}}\,\frac{\partial}{\partial a}\,\mathsf{F})$. If we compare the equations (7) to (10) to the classical ones, it is the replacement of a "+" by a "$\oplus$" in the right handside of (8) and (9) that realizes the generalization of the idea of Antimirov. Equations (7), (10) and (11) are the natural ones that are necessary for the generalization to expressions with multiplicity. Notice that equation (10) is defined only if E is a valid expression.

Once again, let us stress on the fact that with these conventions, a derivative of *an expression* with respect to a word is not an expression anymore, but *a polynomial of rational expressions* (with coefficients in $\mathbb{K}$).

Contrary to the Boolean case, the number of polynomials obtained by iterating the derivative process can be infinite. Theorem 1 will state that all these different polynomials are linear combination of a fixed finite number of expressions.

The derivative of an expression with respect to a *word* f is defined by induction on the length of f (by convention, the derivation with respect to the empty word is the identity):

$$\forall f \in A^*\,,\ \forall a \in A \qquad \frac{\partial}{\partial f a}\,\mathsf{E} = \frac{\partial}{\partial a}\left(\frac{\partial}{\partial f}\,\mathsf{E}\right) \tag{12}$$

The derivative of an expression with respect to a word corresponds to the *(left) quotient* of a series by a word. Recall that if s is a series in $\mathbb{K}\langle\langle A^*\rangle\rangle$, the left quotient of s by a word f in A^* is the series $f^{-1}s$ defined by

$$\forall g \in A^* \qquad <f^{-1}s, g> \;=\; <s, f\,g> \quad .$$

The link between derivative and quotient is explained in the following proposition:

Proposition 3 $\quad \forall f \in A^+ \qquad |\frac{\partial}{\partial f}(\mathsf{E})| = f^{-1}|\mathsf{E}|$

This result that can be proved directly will appear as a corollary of more precise properties of the derivative of a rational expression.

The following properties are easily established, by induction on the length of the word f:

Proposition 4 $\quad \forall f \in A^+ \qquad \frac{\partial}{\partial f}(\mathsf{E}+\mathsf{F}) = \frac{\partial}{\partial f}\,\mathsf{E} \oplus \frac{\partial}{\partial f}\,\mathsf{F}$ ■

Proposition 5 $\quad \forall f \in A^+ \qquad \frac{\partial}{\partial f}(\mathsf{E}\cdot\mathsf{F}) = \left[\frac{\partial}{\partial f}\,\mathsf{E}\right]\cdot\mathsf{F} \oplus \left[\bigoplus_{\substack{f=g\,h\\ g\in A^*,\,h\in A^+}} \mathsf{c}(\frac{\partial}{\partial g}\,\mathsf{E})\ \frac{\partial}{\partial h}\,\mathsf{F}\right]$ ■

Proposition 6 $\quad \forall f \in A^+$

$$\frac{\partial}{\partial f}(\mathsf{E}^*) = \bigoplus_{\substack{f=g_1 g_2\cdots g_n\\ g_1,g_2,\cdots,g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\,\mathsf{c}(\frac{\partial}{\partial g_1}\,\mathsf{E})\ \mathsf{c}(\mathsf{E})^*\,\mathsf{c}(\frac{\partial}{\partial g_2}\,\mathsf{E})\cdots\mathsf{c}(\mathsf{E})^*\,\mathsf{c}(\frac{\partial}{\partial g_{n\text{-}1}}\,\mathsf{E})\ \mathsf{c}(\mathsf{E})^*\left[\frac{\partial}{\partial g_n}\,\mathsf{E}\right]\cdot(\mathsf{E}^*)$$

■

Although this last equation can be simplified in the commutative case, we choose to write it in this form which is more relevant to the process of derivation.

Examples 1 (continued) **:**

i) $$\frac{\partial}{\partial a}\mathsf{E}_1 = \frac{\partial}{\partial a}(\mathsf{F_1}^*) = 2\,\frac{\partial}{\partial a}\left(\tfrac{1}{6}\,a^*\right)\cdot \mathsf{F_1}^* \oplus 2\,\frac{\partial}{\partial a}\left(\tfrac{1}{3}\,b^*\right)\cdot \mathsf{F_1}^* = \tfrac{1}{3}\,(a^*\cdot \mathsf{F_1}^*)$$

$$\frac{\partial}{\partial b}\mathsf{E}_1 = 2\,\frac{\partial}{\partial b}\left(\tfrac{1}{3}\,b^*\right)\cdot \mathsf{F_1}^* = \tfrac{2}{3}\,(b^*\cdot \mathsf{F_1}^*)$$

$$\begin{aligned}\frac{\partial}{\partial aa}\mathsf{E}_1 &= \tfrac{1}{3}\,\frac{\partial}{\partial a}(a^*\cdot \mathsf{F_1}^*) = \tfrac{1}{3}\left(\frac{\partial}{\partial a}\,a^*\right)\cdot \mathsf{F_1}^* \oplus \tfrac{1}{3}\,\mathsf{c}(a^*)\,\frac{\partial}{\partial a}(\mathsf{F_1}^*)\\ &= \tfrac{1}{3}\,(a^*\cdot \mathsf{F_1}^*)\oplus \tfrac{1}{9}\,(a^*\cdot \mathsf{F_1}^*) = \tfrac{4}{9}\,(a^*\cdot \mathsf{F_1}^*)\end{aligned}$$

$$\frac{\partial}{\partial ab}\mathsf{E}_1 = \tfrac{1}{3}\,\frac{\partial}{\partial b}(a^*\cdot \mathsf{F_1}^*) = \tfrac{1}{3}\left(\frac{\partial}{\partial b}\,a^*\right)\cdot \mathsf{F_1}^* \oplus \tfrac{1}{3}\,\mathsf{c}(a^*)\,\frac{\partial}{\partial b}(\mathsf{F_1}^*) = \tfrac{2}{9}\,(b^*\cdot \mathsf{F_1}^*)$$

$$\frac{\partial}{\partial ba}\mathsf{E}_1 = \tfrac{2}{3}\,\frac{\partial}{\partial a}(b^*\cdot \mathsf{F_1}^*) = \tfrac{2}{3}\left(\frac{\partial}{\partial a}\,b^*\right)\cdot \mathsf{F_1}^* \oplus \tfrac{2}{3}\,\mathsf{c}(b^*)\,\frac{\partial}{\partial a}(\mathsf{F_1}^*) = \tfrac{2}{9}\,(a^*\cdot \mathsf{F_1}^*)$$

$$\begin{aligned}\frac{\partial}{\partial bb}\mathsf{E}_1 &= \tfrac{2}{3}\,\frac{\partial}{\partial b}(b^*\cdot \mathsf{F_1}^*) = \tfrac{2}{3}\left(\frac{\partial}{\partial b}\,b^*\right)\cdot \mathsf{F_1}^* \oplus \tfrac{2}{3}\,\mathsf{c}(b^*)\,\frac{\partial}{\partial b}(\mathsf{F_1}^*)\\ &= \tfrac{2}{3}\,(b^*\cdot \mathsf{F_1}^*)\oplus \tfrac{4}{9}\,(b^*\cdot \mathsf{F_1}^*) = \tfrac{10}{9}\,(b^*\cdot \mathsf{F_1}^*)\end{aligned}$$

ii) $$\frac{\partial}{\partial a}\mathsf{E}_2 = b\,a \oplus (a - b\,a) \qquad\qquad \frac{\partial}{\partial b}\mathsf{E}_2 = 0$$

$$\frac{\partial}{\partial aa}\mathsf{E}_2 = \frac{\partial}{\partial a}\,b\,a \oplus \frac{\partial}{\partial a}(a - b\,a) = 1$$

$$\frac{\partial}{\partial ab}\mathsf{E}_2 = \frac{\partial}{\partial b}\,b\,a \oplus \frac{\partial}{\partial b}(a - b\,a) = a \oplus (-1_{\mathbb{K}})\,a = 0$$

□

3.2 Derived terms

We state now the main theorem that is a generalization of Antimirov's result.

Theorem 1 *Let* E *be in* $\mathbb{K}\mathsf{RatE}\,A^*$*. There exist an integer* m*,* $m \leqslant \ell(\mathsf{E})$*, and* m *rational expressions* K_1*,* K_2*, ... ,* K_m *such that for every word* f *in* A^+*, there exist* m *coefficients in* $\mathbb{K}$*,* $k_1^{(f)}$*,* $k_2^{(f)}$*, ... ,* $k_m^{(f)}$*, such that*

$$\frac{\partial}{\partial f}\mathsf{E} = \bigoplus_{i=1}^{i=m} k_i^{(f)}\,\mathsf{K}_i \quad .$$

The following statement proves to be very convenient in order to establish Theorem 1 and some further results of this paper.

Proposition 7 *Let* E *be in* $\mathbb{K}\mathsf{RatE}\,A^*$. *There exist an integer* n, $n \leqslant \ell(\mathsf{E})$, *and* n *rational expressions* $\mathsf{K}_1, \mathsf{K}_2, \ldots, \mathsf{K}_n$ *such that for every letter* a *in* A, *there exist* n *coefficients,* $k_1^{(a)}, k_2^{(a)}, \ldots, k_n^{(a)}$, *and* n^2 *coefficients* $\{z_{i,j}^{(a)}\}_{i,j\in[n]}$ *in* $\mathbb{K}$ *such that*

i) $\displaystyle \frac{\partial}{\partial a}\mathsf{E} = \bigoplus_{i\in[n]} k_i^{(a)}\,\mathsf{K}_i$;

ii) $\forall i \in [n]$ $\displaystyle \frac{\partial}{\partial a}\mathsf{K}_i = \bigoplus_{j\in[n]} z_{i,j}^{(a)}\,\mathsf{K}_j$.

We call the expressions K_i, the existence of which is asserted in the proposition, the *derived terms* of E. In the case where $\mathbb{K} = \mathbb{B}$, they are exactly what Antimirov called "*partial derivatives*" of E, with the explication that they are "*parts*" of the *derivatives* of E ([1]).[2]

Examples 1 (continued) **:** The derived terms of E_1 are $(a^* \cdot \mathsf{F}_1{}^*)$ and $(b^* \cdot \mathsf{F}_1{}^*)$. □

As we shall see in the proof of the Theorem 1, the expressions that appear in the theorem are derived terms. However, it may happen that the numbers m and n are different ($m \leqslant n$ always holds). The derived terms that *do not* appear in the Theorem 1 are called *shadow derived terms*. Their significance will be better understood in the next section. It is also easily seen that if $\mathbb{K} = \mathbb{B}$ or, more generally, if $\mathbb{K}$ is a *positive* semiring, there is no shadow term.[3]

Examples 1 (continued) **:** There is no shadow expression among the derived terms of E_1 whereas those of E_2 are $b\,a$, $(a - b\,a)$, a and $\mathsf{1}$: $a = \frac{\partial}{\partial b}(a - b\,a)$ is a shadow derived term of E_2. □

Proof of Proposition 7. By induction on the depth of the expression E (not on its *litteral* length). The statement obviously holds for $\mathsf{0}$ and $\mathsf{1}$ and for $\mathsf{E} = a$, $a \in A$. We then successively show:

a) If it is true for E, it is true for $(k\,\mathsf{E})$, $k \in \mathbb{K}$. Obvious from (7). The derived terms of $(k\,\mathsf{E})$ are the same as those of E.

b) If it is true for E and F, it is true for $(\mathsf{E} + \mathsf{F})$. It holds

$$\frac{\partial}{\partial a}(\mathsf{E}{+}\mathsf{F}) = \frac{\partial}{\partial a}\mathsf{E} \oplus \frac{\partial}{\partial a}\mathsf{F} = \bigoplus_{i\in[n]} k_i^{(a)}\,\mathsf{K}_i \oplus \bigoplus_{p\in[s]} l_p^{(a)}\,\mathsf{L}_p$$

with obvious notation. The set of derived terms of $(\mathsf{E} + \mathsf{F})$ is the union of those of E and F and this set clearly satisfies the proposition.

[2] We consider this wording as rather unfortunate, since the derivatives of a rational expression are already "partial" in essence as we have derivatives with respect to every letter of the alphabet, *i.e.* to every (non commutative) unknown. Moreover, "*partial derivative*" is an established expression in the area of mathematics, that has a definite relationship with the expression "*derivative*".

[3] And this is the reason why this notion of shadow term does not appear in the work of Antimirov.

c) If it is true for E and F, it is true for $(\mathsf{E}\cdot\mathsf{F})$. It holds

$$\frac{\partial}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial}{\partial a}\mathsf{E}\right]\cdot\mathsf{F}\right) \oplus \mathsf{c}(\mathsf{E})\ \frac{\partial}{\partial a}\mathsf{F} = \bigoplus_{i\in[n]} k_i^{(a)}\ (\mathsf{K}_i\cdot\mathsf{F}) \oplus \bigoplus_{p\in[s]} \left(\mathsf{c}(\mathsf{E})\, l_p^{(a)}\right)\ \mathsf{L}_p$$

The set of derived terms of $(\mathsf{E}\cdot\mathsf{F})$ is the union of the set $\{(\mathsf{K}_i\cdot\mathsf{F})\}_{i\in[n]}$ and of $\{\mathsf{L}_p\}_{p\in[s]}$ and one verifies that, for every i in $[n]$ and every a in A, it holds:

$$\frac{\partial}{\partial a}(\mathsf{K}_i\cdot\mathsf{F}) = \bigoplus_{j\in[n]} z_{i,j}^{(a)}\ (\mathsf{K}_j\cdot\mathsf{F}) \oplus \bigoplus_{p\in[s]} \left(\mathsf{c}(\mathsf{K}_i)\, l_p^{(a)}\right)\ \mathsf{L}_p$$

d) If it is true for E, it is true for (E^*) . It holds

$$\frac{\partial}{\partial a}(\mathsf{E}^*) = \mathsf{c}(\mathsf{E})^* \left(\left[\frac{\partial}{\partial a}\mathsf{E}\right]\cdot(\mathsf{E}^*)\right) = \bigoplus_{i\in[n]} \left(\mathsf{c}(\mathsf{E})^*\, k_i^{(a)}\right)\ (\mathsf{K}_i\cdot\mathsf{E}^*)$$

The set of derived terms of (E^*) is $\{(\mathsf{K}_i\cdot\mathsf{E}^*)\}_{i\in[n]}$. And one verifies that, for every i in $[n]$ and every a in A, it holds:

$$\frac{\partial}{\partial a}(\mathsf{K}_i\cdot\mathsf{E}^*) = \bigoplus_{j\in[n]} z_{i,j}^{(a)}\ (\mathsf{K}_j\cdot\mathsf{E}^*) \oplus \bigoplus_{j\in[n]} \left(\mathsf{c}(\mathsf{K}_i)\ \mathsf{c}(\mathsf{E})^*\, k_j^{(a)}\right)\ (\mathsf{K}_j\cdot\mathsf{E}^*) \qquad \blacksquare$$

Proof of Theorem 1. By induction on the length of f. If f is a letter, the theorem is equivalent to the equality i) of the Proposition 7. For every f in $\mathbb{A}^+$ and every a in A, it holds:

$$\begin{aligned}\frac{\partial}{\partial fa}\mathsf{E} &= \frac{\partial}{\partial a}\left(\frac{\partial}{\partial f}\mathsf{E}\right) = \bigoplus_{i\in[n]} k_i^{(f)}\ \frac{\partial}{\partial a}\mathsf{K}_i \\ &= \bigoplus_{i\in[n]} k_i^{(f)}\left[\bigoplus_{j\in[n]} z_{i,j}^{(a)}\ \mathsf{K}_j\right] = \bigoplus_{j\in[n]}\left[\bigoplus_{i\in[n]}\left(k_i^{(f)}\ z_{i,j}^{(a)}\right)\right]\mathsf{K}_j \end{aligned} \tag{13}$$

One observes here the first benefit of Proposition 7: the proof of Theorem 1 goes by induction on the length of f and not on the depth of E and the formulae involved keep much shorter than they would with a direct proof. $\blacksquare$

Remark 3 Indeed, Proposition 7 is not only a convenient statement to prove Theorem 1; it describes the most natural algorithm in order to compute the derived terms of E. The derived terms are computed by successive derivations, until their set does not grow anymore.

The following makes the link between the coefficient of f in the series denoted by E and the derivative of E by f.

Theorem 2 *Let* E *be in* $\mathbb{K}\,\mathsf{RatE}\,A^*$*, let* $\mathsf{K}_1, \mathsf{K}_2, \ldots, \mathsf{K}_m$ *be its derived terms and, for every word* f *in* A^+*, let* $k_1^{(f)}, k_2^{(f)}, \ldots, k_m^{(f)}$ *be the coefficients defined in Theorem 1. It then holds:*

$$<|\mathsf{E}|, f> \; = \mathsf{c}(\frac{\partial}{\partial f}\,\mathsf{E}) = \bigoplus_{i=1}^{i=m} k_i^{(f)}\;\mathsf{c}(\mathsf{K}_i) \tag{14}$$

Proof. The proof goes by induction on the depth of the expression and makes use of propositions 4, 5 and 6.

The result is true for $\mathsf{0}$ and $\mathsf{1}$: the derivation with respect to any word f in A^+ is null and the coefficient of f in the series 0 and 1_{A^*} is actually null.

Likewise, if the expression is reduced to a letter a in A, except if $f = a$. In this case, the derived expression is $\mathsf{1}$ and the theorem holds.

For every k in $\mathbb{K}$, and for every rational expressions E and F, for which the theorem holds,

$$
\begin{aligned}
<|(k\,\mathsf{E})|, f> \; &= k \; <|\mathsf{E}|, f> \; = k\;\mathsf{c}(\frac{\partial}{\partial f}\,\mathsf{E}) = \mathsf{c}(\frac{\partial}{\partial f}(k\,\mathsf{E}))\\
<|(\mathsf{E}+\mathsf{F})|, f> \; &= \; <|\mathsf{E}|, f> \;\oplus\; <|\mathsf{F}|, f>\\
&= \mathsf{c}(\frac{\partial}{\partial f}\,\mathsf{E}) \oplus \mathsf{c}(\frac{\partial}{\partial f}\,\mathsf{F})\\
&= \mathsf{c}(\frac{\partial}{\partial f}(\mathsf{E}+\mathsf{F}))\\
<|(\mathsf{E}\cdot\mathsf{F})|, f> \; &= \; <|(\mathsf{E})|, f> \;\; <|(\mathsf{F})|, 1_{A^*}> \;\oplus \bigoplus_{\substack{f=g\,h\\ g\in A^*,\, h\in A^+}} <|(\mathsf{E})|, g> \;\; <|(\mathsf{F})|, h>\\
&= \mathsf{c}(\frac{\partial}{\partial f}\,E)\,\mathsf{c}(\mathsf{F}) \oplus \bigoplus_{\substack{f=g\,h\\ g\in A^*,\, h\in A^+}} \mathsf{c}(\frac{\partial}{\partial g}\,\mathsf{E})\,\mathsf{c}(\frac{\partial}{\partial h}\,\mathsf{F})\\
&= \mathsf{c}\left(\left[\frac{\partial}{\partial f}\,E\right]\cdot\mathsf{F} \oplus \bigoplus_{\substack{f=g\,h\\ g\in A^*,\, h\in A^+}} \mathsf{c}(\frac{\partial}{\partial g}\,\mathsf{E})\,\frac{\partial}{\partial h}\,\mathsf{F}\right)\\
&= \mathsf{c}(\frac{\partial}{\partial f}(\mathsf{E}\cdot\mathsf{F}))
\end{aligned}
$$

All these equations, as well as the second equality of the theorem, directly follow from the linearity of $\mathsf{c}(\mathsf{E})$. To prove the result on E^*, in order to avoid an infinite sum, we use the

Lemma 1 before applying the same arguments.

$$\begin{aligned}
<|(\mathsf{E}^*)|, f> \;&=\; <(\mathsf{c}(\mathsf{E})^*|\mathsf{E}|_p)^*\ \mathsf{c}(\mathsf{E})^*, f> \;=\; <(\mathsf{c}(\mathsf{E})^*|\mathsf{E}|_p)^*, f>\ \mathsf{c}(\mathsf{E})^* \\
&= \bigoplus_{\substack{f=g_1g_2\cdots g_n\\ g_1,g_2,\cdots,g_n\in A^+}} <\mathsf{c}(\mathsf{E})^*|\mathsf{E}|_p, g_1> \ldots <\mathsf{c}(\mathsf{E})^*|\mathsf{E}|_p, g_n>\ \mathsf{c}(\mathsf{E})^* \\
&= \bigoplus_{\substack{f=g_1g_2\cdots g_n\\ g_1,g_2,\cdots,g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\ <|\mathsf{E}|_p, g_1> \ldots \mathsf{c}(\mathsf{E})^*\ <|\mathsf{E}|_p, g_n>\ \mathsf{c}(\mathsf{E})^*
\end{aligned}$$

As the g_i are all different from 1_{A^*}, $<|\mathsf{E}|_p, g_i> \;=\; <|\mathsf{E}|, g_i>$.

$$\begin{aligned}
<|(\mathsf{E}^*)|, f> \;&=\; \bigoplus_{\substack{f=g_1g_2\cdots g_n\\ g_1,g_2,\cdots,g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\ \mathsf{c}(\frac{\partial}{\partial g_1}\mathsf{E})\ldots \mathsf{c}(\mathsf{E})^*\ \mathsf{c}(\frac{\partial}{\partial g_n}\mathsf{E})\ \mathsf{c}(\mathsf{E})^* \\
&= \mathsf{c}\left(\bigoplus_{\substack{f=g_1g_2\cdots g_n\\ g_1,g_2,\cdots,g_n\in A^+}} \mathsf{c}(\mathsf{E})^*\ \mathsf{c}(\frac{\partial}{\partial g_1}\mathsf{E})\ldots \mathsf{c}(\mathsf{E})^*\ \left[\frac{\partial}{\partial g_n}\mathsf{E}\right]\cdot(\mathsf{E}^*)\right) \\
&= \mathsf{c}(\frac{\partial}{\partial f}\mathsf{E}^*) \qquad \blacksquare
\end{aligned}$$

Proof of Proposition 3. The definition of the derivative of an expression E implies that $\frac{\partial}{\partial fg}\mathsf{E} = \frac{\partial}{\partial g}\left[\frac{\partial}{\partial f}\mathsf{E}\right]$ for every pair of words f and g in A^*. Thus, for all f and g in A^*,

$$<f^{-1}|\mathsf{E}|, g> \;=\; <|\mathsf{E}|, fg> \;= \mathsf{c}(\frac{\partial}{\partial fg}\mathsf{E}) = \mathsf{c}\left(\frac{\partial}{\partial g}\left[\frac{\partial}{\partial f}\mathsf{E}\right]\right) \;=\; <|\frac{\partial}{\partial f}\mathsf{E}|, g> \qquad \blacksquare$$

Remark 4 The derivative and the left quotient are right actions of A^* on the set of polynomials of rational expressions and the set of rational series respectively. Theorem 1 says that the orbit of a rational expression with multiplicity under the action of A^* belongs to a finitely generated $\mathbb{K}$-semimodule. The function which maps a polynomial of expressions P onto the rational power series $|P|$ is a morphism of actions. Therefore, Theorem 1 implies that the orbit of a rational series under the action of A^* belongs to a finitely generated $\mathbb{K}$-semimodule as well, and provides a new proof for this classical result [2].

4 The automaton of derived terms

To any rational expression with multiplicity E in $\mathbb{K}\,\mathsf{RatE}\,A^*$, we associate a $\mathbb{K}$-automaton (*i.e.* an automaton over the alphabet A with multiplicity in $\mathbb{K}$) in the following way. Let $P = \{\mathsf{K}_1, \mathsf{K}_2, \ldots, \mathsf{K}_n\}$ be the set of the derived terms of E. Let $\mathsf{K}_0 = \mathsf{E}$ and let P_{E} be the union of P and K_0. The *automaton of derived terms of* E is the $\mathbb{K}$-automaton $\mathcal{A}_{\mathsf{E}} = \langle P_{\mathsf{E}}, A, Z, I, T \rangle$ defined by:

$$I_{\mathsf{K}_i} = \begin{cases} 1_{\mathbb{K}} & \text{if } \mathsf{K}_i = \mathsf{K}_0 \\ 0_{\mathbb{K}} & \text{otherwise} \end{cases}, \qquad Z_{\mathsf{K}_i,\mathsf{K}_j} = \bigoplus_{a \in A} z_{i,j}^{(a)}\, a\,, \qquad T_{\mathsf{K}_j} = \mathsf{c}(\mathsf{K}_j),$$

where the $z_{i,j}^{(a)}$ have been defined at Proposition 7.

Theorem 3 *Let* E *be in* $\mathbb{K}\mathsf{RatE}\, A^*$*. The series realized by the automaton of derived terms of* E *is equal to the series denoted by* E*:*

$$|\mathcal{A}_{\mathsf{E}}| = |\mathsf{E}|$$

Proof. The definition of the $\mathbb{K}$-automaton $\mathcal{A}_{\mathsf{E}}$ is indeed equivalent to the definition of a "$\mathbb{K}$-*representation*" (I, ζ, T). This is a part of the proof of the so-called Kleene-Schützenberger Theorem (*cf.* [2]). In the representation (I, ζ, T), I and T are the two vectors of dimension n (where $n = \mathsf{Card}(P_{\mathsf{E}})$) with entries in $\mathbb{K}$ defined above and $\zeta \colon A^* \to \mathbb{K}^{n \times n}$ is the morphism from A^* into the monoid of $n \times n$-matrices with entries in $\mathbb{K}$ defined by

$$\forall a \in A\,,\ \forall i, j \in [n] \qquad (a)\zeta_{i,j} = z_{i,j}^{(a)}.$$

The *series realized* by the representation (I, ζ, T) (and thus by the automaton $\mathcal{A}_{\mathsf{E}}$) is, by definition:

$$|\mathcal{A}_{\mathsf{E}}| = \bigoplus_{f \in A^*} (I \cdot (f)\zeta \cdot T)\, f.$$

Now, (13) directly shows, by induction on the length of f, that

$$\forall f \in A^+\,,\ \forall i \in [n] \qquad (I \cdot (f)\zeta)_i = k_i^{(f)}, \tag{15}$$

and then, by (14),

$$\forall f \in A^+ \qquad <|\mathcal{A}_{\mathsf{E}}|, f> \;= \bigoplus_{i[n]} k_i^{(f)}\ \mathsf{c}(\mathsf{K}_i) = \;<|\mathsf{E}|, f>. \tag{16}$$

The theorem is then trivially verified, since the coefficient of the empty word is

$$<|\mathcal{A}_{\mathsf{E}}|, 1_{A^*}> \;= \mathsf{c}(\mathsf{K}_0) = \mathsf{c}(\mathsf{E}). \qquad \blacksquare$$

Examples 1 (continued) **:** Figure 1 shows the automaton of derived terms of E_1. □

Proposition 8 *Let* E *be in* $\mathbb{K}\mathsf{RatE}\, A^*$ *and let* $\mathcal{A}_{\mathsf{E}}$ *be its automaton of derived terms. The* $\mathbb{K}$*-automaton obtained from* $\mathcal{A}_{\mathsf{E}}$ *by erasing the states that correspond to shadow derived terms and by trimming the result realizes the same series* $|\mathcal{A}_{\mathsf{E}}| = |\mathsf{E}|$.

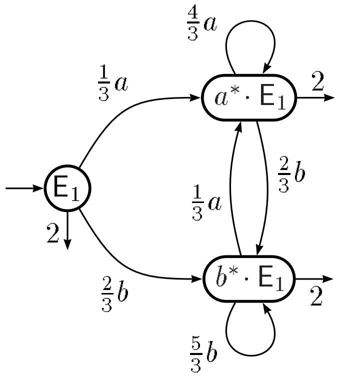


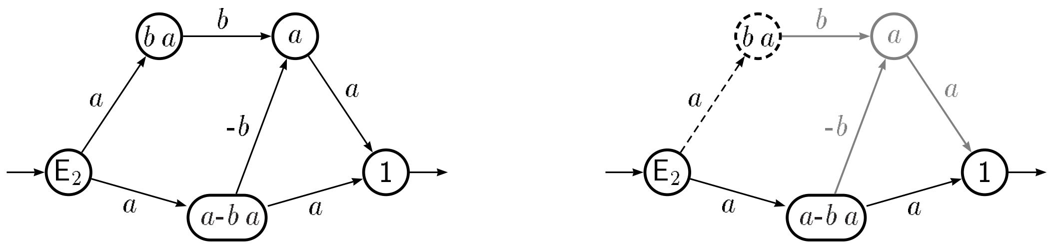
Figure 1: The automaton of derived terms of E_1

Proof. If K_i is a shadow derived term, then, by (15):

$$k_i^{(f)} = (I \cdot (f)\zeta)_i = 0_{\mathbb{K}}$$

for every f in A^* and this state can be erased in the representation (I, ζ, T) without changing the realized series. The further trimming of the corresponding automaton does not change the realized series either. ■

Example 1 (continued) : Figure 2 a) shows the automaton of derived terms of E_2; b) shows the effect of, first, the suppression of the state a and then the trimming that erases the state $b\,a$. □



(a) The automaton of derived terms of E_2 (b) The automaton of non-shadow derived terms of E_2

Figure 2: An example with shadow derived terms.

The effectively computation of the shadow terms depends on the semiring of coefficients. In many cases —and in particular in all classical cases— this computation does not bring any problem. On the one hand, if the semiring is positive, there is no shadow term. For instance, Boolean semiring, sub-semirings of $\mathbb{R}_+$, $(\mathsf{max}, +)$-semirings, $(\mathcal{P}(A^*), \cup, .)$,*etc.* are positive[4]. On the other hand, K_i is a shadow term if and only if the series realized by the automaton $\mathcal{A}_i = \langle\, P_{\mathsf{E}}, A, Z, I, \{\mathsf{K}_i\}\,\rangle$ is equal to zero. This can be easily decided if the

[4]Notice that not all idempotent semirings are positive: some may have zero divisors.

semiring of coefficients is a sub-semiring of a field (*cf.* [6], Equality theorem). Examples of such semirings are $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{Z}[X]$, *etc.*

5 Variations

The definition of derivations and of derived terms has been chosen such a way they can stand as a perfect generalization of those defined by Antimirov and, in particular, in order to get the same bound on the number of states of the automaton of derived terms. But as there are many different semirings, and for other developments, it may be interesting to choose slightly different definitions for the derivatives.

Here are two of them, the first one leads to an automaton that may be smaller than the automaton obtained by the basic definitions, the second one may yield a bigger automaton whose structure is closer to the expression.

A convenient way to describe the modifications in the definition of derivatives is to take another convention for the derivative with respect to the empty word. We recall that until now, we suppose that the derivative with respect to the empty word is the identity on the set of rational expressions with multiplicity: $\frac{\partial}{\partial 1_{A^*}} \mathsf{E} = \mathsf{E}$. We shall see that this single modification implies deep differences between automata of derived terms.

5.1 Building the derivatives with unitary expressions

In the previous part we define the automaton of derived terms in which every state is characterized by an expression. So E and $(k\,\mathsf{E})$, which are of course distinct expressions for k different from $1_{\mathbb{K}}$, label two different states. We can define derivatives such that such a pair of states will finally be merged in the automaton of derived terms.

With this intention, we assume that there exists another identity on rational expressions:

$$(k\,(k'\,\mathsf{E})) \equiv ([k\,k']\,\mathsf{E}) \tag{17}$$

A consequence of the identity (17) is that any expression E in $\mathbb{K}\mathsf{RatE}\,A^*$ may be written in a *unique* way as $\mathsf{E} = k\,\mathsf{M}$ where M is what we call a *unitary expression* or, more precisely *the unitary expression in* E if we refer to E. It is thus a letter or an expression of the form: $(\mathsf{E}+\mathsf{F})$, $(\mathsf{E}\cdot\mathsf{F})$ or (E^*) — but not of the form $(k\,\mathsf{E})$ — *e.g.* E_1 and E_3 (*cf.* Example 1) are unitary expressions, $((2\,a\,b) + ((3\,b)\cdot(4\,(a\,b)^*)))^*$ is the unitary expression in E_2.

Remark 5 The "extraction" of a unitary expression is not a process of factorization: $((2\,a) + (4\,b))$ is unitary while $(a + (2\,b))$ is the unitary expression in $(2\,(a + (2\,b)))$.

We define the derivative with unitary expressions whose result is a polynomial of unitary expressions. For this purpose, we define $\frac{\partial'}{\partial 1_{A^*}}$ which maps every expression $\mathsf{E} = (k\,\mathsf{M})$ in $\mathbb{K}\mathsf{RatE}\,A^*$ (where M is the unitary expression in E) into the monomial $k\,\mathsf{M}$.

We can then define the derivative with unitary expressions of E in $\mathbb{K}\,\mathsf{RatE}\,A^*$ with respect to a as

$$\frac{\partial'}{\partial a}\,\mathsf{E} = \frac{\partial'}{\partial 1_{A^*}}(\frac{\partial}{\partial a}\,\mathsf{E}). \tag{18}$$

The aim of this method is that the coefficients pop up as soon as possible. Applying this, we give a modification to the equation (9):

$$\frac{\partial'}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial'}{\partial a}\,\mathsf{E}\right]\cdot\left[\frac{\partial'}{\partial 1_{A^*}}\,\mathsf{F}\right]\right) \oplus \mathsf{c}(\mathsf{E})\,\frac{\partial'}{\partial a}\,\mathsf{F} \tag{19}$$

This modification is consistent with the equation (9) since $\frac{\partial}{\partial 1_{A^*}}\,\mathsf{F} = \mathsf{F}$, for every F in $\mathbb{K}\,\mathsf{RatE}\,A^*$.

Let E be an expression in $\mathbb{K}\,\mathsf{RatE}\,A^*$. Let $P' = \{\mathsf{K}'_1, \mathsf{K}'_2, \ldots, \mathsf{K}'_n\}$ be the set of unitary derived terms and $k_0\,\mathsf{K}'_0 = \frac{\partial'}{\partial 1_{A^*}}\,\mathsf{E}$. Let $\{z'^{(a)}_{i,j} \mid i,j \in [n], a \in A\}$ be elements of $\mathbb{K}$ such that

$$\frac{\partial'}{\partial a}\,\mathsf{K}'_i = \bigoplus_{j\in[n]} z'^{(a)}_{i,j}\,\mathsf{K}'_j. \tag{20}$$

The *automaton of unitary derived terms of* E is the $\mathbb{K}$-automaton $\mathcal{A}'_{\mathsf{E}} = \langle\, P' \cup \mathsf{K}'_0, A, Z', I', T' \,\rangle$ defined by:

$$I'_{\mathsf{K}'_i} = \begin{cases} k_0 & \text{if} \quad \mathsf{K}'_i = \mathsf{K}'_0 \\ 0_{\mathbb{K}} & \text{otherwise} \end{cases}, \qquad Z'_{\mathsf{K}'_i,\mathsf{K}'_j} = \bigoplus_{a\in A} z'^{(a)}_{i,j}\,a\,, \qquad T'_{\mathsf{K}'_j} = \mathsf{c}(\mathsf{K}'_j).$$

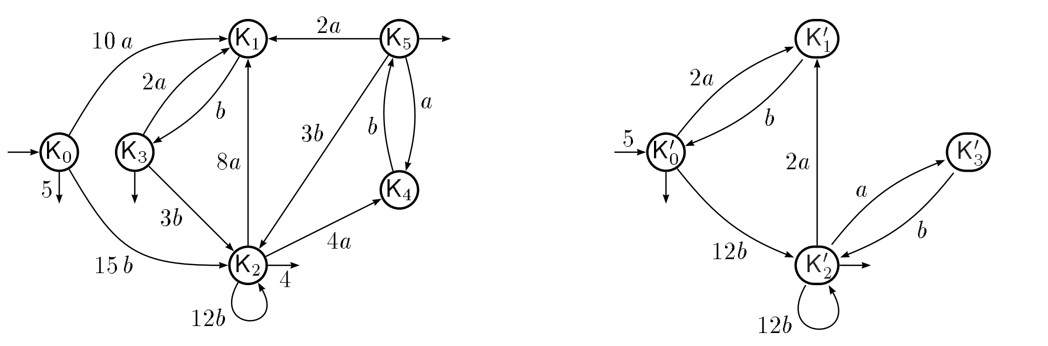
The proof that the number of derived terms obtained by this method is still smaller than the length of the expression and that the unitary derived terms automaton of E realizes the series $|\mathsf{E}|$ is straightforward from the proof in the "classic" case. Their number are even smaller or equal than the number of the derived terms.

Examples 1 (continued) : Let $\mathsf{F}_3 = ((2ab)+((3b)\cdot(4(ab)^*)))^*$; $\mathsf{E}_3 = (5\,\mathsf{F}_3)$. We compute the derived terms (left column) and the unitary derived terms (right column):

$$
\begin{aligned}
\mathsf{K}_0 &= \mathsf{E}_3 \\
\frac{\partial}{\partial a}\mathsf{K}_0 &= 10\,(b\cdot\mathsf{F}_3) = 10\,\mathsf{K}_1 \\
\frac{\partial}{\partial b}\mathsf{K}_0 &= 15\,((4\,(ab)^*)\cdot\mathsf{F}_3) = 15\,\mathsf{K}_2 \\
\frac{\partial}{\partial b}\mathsf{K}_1 &= \mathsf{F}_3 = \mathsf{K}_3 \\
\frac{\partial}{\partial a}\mathsf{K}_2 &= 4\,((b\cdot(ab)^*)\cdot\mathsf{F}_3)\oplus 8\,(b\cdot\mathsf{F}_3) \\
&= 4\,\mathsf{K}_4 \oplus 8\,\mathsf{K}_1 \\
\frac{\partial}{\partial b}\mathsf{K}_2 &= 12\,((4\,(ab)^*)\cdot\mathsf{F}_3) = 12\,\mathsf{K}_2 \\
\frac{\partial}{\partial a}\mathsf{K}_3 &= 2\,(b\cdot\mathsf{F}_3) = 2\,\mathsf{K}_1 \\
\frac{\partial}{\partial b}\mathsf{K}_3 &= 3\,\mathsf{K}_2 \\
\frac{\partial}{\partial b}\mathsf{K}_4 &= ((ab)^*\cdot\mathsf{F}_3) = \mathsf{K}_5 \\
\frac{\partial}{\partial a}\mathsf{K}_5 &= \mathsf{K}_4 \oplus 2\,\mathsf{K}_1 \\
\frac{\partial}{\partial b}\mathsf{K}_5 &= 3\,\mathsf{K}_2 \\
\mathsf{c}(\mathsf{K}_0) &= 5, \mathsf{c}(\mathsf{K}_1) = \mathsf{c}(\mathsf{K}_4) = 0 \\
\mathsf{c}(\mathsf{K}_2) &= 4, \mathsf{c}(\mathsf{K}_3) = \mathsf{c}(\mathsf{K}_5) = 1
\end{aligned}
$$

$$
\begin{aligned}
\mathsf{K}'_0 &= \mathsf{F}_3 \qquad k_0 = 5 \\
\frac{\partial\text{'}}{\partial a}\mathsf{K}'_0 &= 2\,(b\cdot\mathsf{F}_3) = 2\,\mathsf{K}'_1 \\
\frac{\partial\text{'}}{\partial b}\mathsf{K}'_0 &= 12\,((ab)^*\cdot\mathsf{F}_3) = 12\,\mathsf{K}'_2 \\
\frac{\partial\text{'}}{\partial b}\mathsf{K}'_1 &= \mathsf{F}_3 = \mathsf{K}'_0 \\
\frac{\partial\text{'}}{\partial a}\mathsf{K}'_2 &= ((b\cdot(ab)^*)\cdot\mathsf{F}_3)\oplus 2\,(b\cdot\mathsf{F}_3) \\
&= \mathsf{K}'_3 \oplus 2\,\mathsf{K}'_1 \\
\frac{\partial\text{'}}{\partial b}\mathsf{K}'_2 &= 12\,((ab)^*\cdot\mathsf{F}_3) = 12\,\mathsf{K}'_2 \\
\frac{\partial\text{'}}{\partial b}\mathsf{K}'_3 &= ((ab)^*\cdot\mathsf{F}_3) = \mathsf{K}'_2 \\
\mathsf{c}(\mathsf{K}'_0) &= \mathsf{c}(\mathsf{K}'_2) = 1 \\
\mathsf{c}(\mathsf{K}'_1) &= \mathsf{c}(\mathsf{K}'_3) = 0
\end{aligned}
$$

Figure 3 a) shows the automaton of derived terms of E_3 (which is isomorphic to the Glushkov automaton computed in [4]); b) shows the automaton of unitary derived terms of E_3. □



(a) The automaton of derived terms of E_3 (b) The automaton of unitary derived terms

Figure 3: Two $\mathbb{K}$-automata for E_3

5.2 Breaking the derivatives

Another modification on the derivatives is to consider that the "+" operation of the rational expressions is weak and that the derivative with respect to the empty word breaks it. The breaking derivative of an expression with respect to the empty word is defined as the identity on atomic expressions and on the rational operations except for "+":

$$\frac{\partial''}{\partial 1_{A^*}}(\mathsf{E}+\mathsf{F}) = \frac{\partial''}{\partial 1_{A^*}}\,\mathsf{E} \oplus \frac{\partial''}{\partial 1_{A^*}}\,\mathsf{F} \tag{21}$$

The breaking derivative of an expression E with respect to a letter a is $\frac{\partial''}{\partial a}\,\mathsf{E} = \frac{\partial''}{\partial 1_{A^*}}(\frac{\partial}{\partial a}\,\mathsf{E})$.

The derivative of the "$\cdot$" is improved in the same way as for unitary derivatives:

$$\frac{\partial''}{\partial a}(\mathsf{E}\cdot\mathsf{F}) = \left(\left[\frac{\partial''}{\partial a}\,\mathsf{E}\right]\cdot\left[\frac{\partial''}{\partial 1_{A^*}}\,\mathsf{F}\right]\right) \oplus \mathsf{c}(\mathsf{E})\,\frac{\partial''}{\partial a}\,\mathsf{F} \tag{22}$$

Let E be an expression in $\mathbb{K}\mathsf{RatE}\,A^*$. Let $P'' = \{\mathsf{K}''_1, \mathsf{K}''_2, \ldots, \mathsf{K}''_n\}$ be the set of broken derived terms and P''_I the set of terms of $\frac{\partial''}{\partial 1_{A^*}}\,\mathsf{E}$. Let $\{z''^{(a)}_{i,j} \mid i,j \in [n], a \in A\}$ be elements of $\mathbb{K}$ such that

$$\frac{\partial''}{\partial a}\,\mathsf{K}''_i = \bigoplus_{j\in[n]} z''^{(a)}_{i,j}\,\mathsf{K}''_j. \tag{23}$$

The *automaton of broken derived terms of* E is the $\mathbb{K}$-automaton $\mathcal{A}''_{\mathsf{E}} = \langle\, P'' \cup P''_I, A, Z'', I'', T''\,\rangle$ defined by:

$$I''_{\mathsf{K}'_i} = \begin{cases} 1 & \text{if } \mathsf{K}''_i \in P''_I \\ 0_{\mathbb{K}} & \text{otherwise} \end{cases}, \qquad Z''_{\mathsf{K}''_i,\mathsf{K}''_j} = \bigoplus_{a\in A} z''^{(a)}_{i,j}\,a\,, \qquad T''_{\mathsf{K}'_j} = \mathsf{c}(\mathsf{K}''_j).$$

With this convention, the bound on the size of the automaton of derived terms does not hold anymore, but in some cases, the result of the algorithm may seem "closer" to the expression.

Example 2 : Let $\mathsf{E}_5 = a^* + b^*$. Figure 4 a) shows the automaton of derived terms of E_5, b) shows the automaton of its broken derived terms. □

(a) The automaton of derived terms of E_5 (b) The automaton of broken derived terms of E_5

Figure 4: Two $\mathbb{K}$-automata for E_5

References

[1] V. ANTIMIROV, Partial derivatives of regular expressions and finite automaton constructions. *Theoret. Comput. Sci. 155* (1996), 291–319.

[2] J. BERSTEL AND CH. REUTENAUER, *Les séries rationnelles et leurs langages.* Masson, 1984. Translation: *Rational Series and their Languages.* Springer, 1986.

[3] J. A. BRZOZOWSKI, Derivatives of regular expressions. *J. Assoc. Comput. Mach. 11* (1964), 481–494.

[4] P. CARON, AND M. FLOURET, Glushkov construction for multiplicities. *Pre-Proceedings of CIAA'00*, M. Daley, M. Eramian and S. Yu, eds, Univ. of Western Ontario, (2000), 52–61.

[5] J.-M. CHAMPARNAUD AND D. ZIADI, New finite automaton constructions based on canonical derivatives. *Pre-Proceedings of CIAA'00*, M. Daley, M. Eramian and S. Yu, eds, Univ. of Western Ontario, (2000), 36–43.

[6] S. EILENBERG, *Automata, Languages and Machines, volume A.* Academic Press, 1974.

[7] V. GLUSHKOV, The abstract theory of automata. *Russian Mathematical Surveys 16* (1961), 1–53.

[8] T. HARJU, AND J. KARHUMÄKI, The equivalence problem of multitape finite automata. *Theoret. Comput. Sci. 78* (1991), 347–355.

[9] D. KROB, Differentiation of K-rational expressions. *Int. J. of Algebra and Computation 2* (1992), 57–87.

[10] W. KUICH, AND A. SALOMAA, *Semirings, Automata, Languages.* Springer, 1986.

[11] R. MCNAUGHTON, AND H. YAMADA, Regular expressions and state graphs for automata. *IRE Trans. on electronic computers 9* (1960), 39–47.

[12] J. SAKAROVITCH, *Eléments de théorie des automates*, in preparation.