Affine completeness of some free binary algebras

Contents

1. Introduction

A function on an algebra is congruence preserving if, for any congruence, it maps pairs of congruent elements onto pairs of congruent elements.

A polynomial function on an algebra is any function defined by a term of the algebra using variables, constants and the operations of the algebra. Obviously, every polynomial function is congruence preserving. An algebra is said to be affine complete if every congruence preserving function is a polynomial function.

We proved in [3] that if Σ has at least three elements, then the free monoid Σ^* generated by Σ is affine complete. If Σ has just one letter a, then the free monoid a^* is isomorphic to $\langle \mathbb{N}, + \rangle$, and we proved in [2] that, e.g., $f\colon \mathbb{N} \to \mathbb{N}$ defined by $f(x) = \texttt{if } x == 0 \texttt{ then } 1 \texttt{ else } \lfloor ex! \rfloor$, where $e = 2.718\ldots$ is the Euler number, is congruence preserving but not polynomial. Thus $\langle \mathbb{N}, + \rangle$, or equivalently the free monoid a^* with concatenation, is not affine complete. Intuitively, this stems from the fact that the more generators Σ^* has, the more congruences it has too: thus $\mathbb{N}$ with just one generator, has very few congruences, hence many functions, including non polynomial ones, can preserve all congruences of $\mathbb{N}$. We also proved in [1] that, when Σ has three letters, in the algebra of full binary trees with leaves labelled by letters in Σ, every unary CP function is polynomial. These previous works left several open questions. What happens if Σ has one or two letters: for algebras of trees? for non unary CP functions on trees? for the free monoid generated by two letters? We answer these three questions in the present paper: these algebras are affine complete.

For full binary trees and at least three letters in Σ, the proof of [1] consisted in showing that CP functions which coincide on Σ are equal, and in building for any CP function f a polynomial P_f such that $f(a) = P_f(a)$ for $a \in \Sigma$, wherefrom we inferred that $f = P_f$ for any t. We now generalize this result in three ways: we consider arbitrary trees (with labelled leaves) where the empty tree is allowed, the alphabet Σ may have one or two letters instead of at least three, and CP functions of any arity are allowed. Our method mostly uses congruences $\sim_{u,v}$ which substitute for occurrences of a tree u a smaller tree v: in fact, we even restrict ourselves to congruences such that u belongs to a subset $\mathcal{T}$ which is chosen in a way ensuring that every congruence class has a unique smallest canonical representative. Using these congruences, we build, for each CP function f, and $\tau \in \mathcal{T}$, a polynomial P_τ such that, for trees $u_1, \ldots, u_n$ small enough, $f(u_1, \ldots, u_n) = P_\tau(u_1, \ldots, u_n)$. We finally show that polynomials which coincide on Σ coincide on the whole algebra, wherefrom we conclude that all the P_τ are equal and f is a polynomial.

The next question is: is $\{a, b\}^*$ equipped with concatenation affine complete? We show in the present paper that the answer is positive. The essential tool used in [3] was the notion of Restricted Congruence Preserving functions (RCP), i.e., functions preserving only the congruences defined by kernels of endomorphisms $\langle \Sigma^*, \cdot \rangle \to \langle \Sigma^*, \cdot \rangle$, which allowed to prove that RCP functions are polynomial, implying that a fortiori CP functions are polynomial. Unfortunately, the fundamental property $\mathcal{P}$ below, which was implicitly used when there are three letters, no longer holds where there are only two letters.

($\mathcal{P}$) Let $\gamma_{a,b}$ be the homomorphism substituting b for a, if $f\colon \Sigma \to \Sigma$ is such that for all $a, b \in \Sigma$, $\gamma_{a,b}(f(a)) = \gamma_{a,b}(f(b))$ then f is either a constant function, or the identity.

Let $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$. When $n = 2$, alas, property $(\mathcal{P})$ is no longer true and restricting ourselves to RCP functions cannot help in proving that CP functions are polynomial. For instance, the function $f\colon \Sigma^* \to \Sigma^*$ defined by $f(w) = \sigma_1^{|w|_{\sigma_1}} \cdots \sigma_n^{|w|_{\sigma_n}}$, where $|w|_\sigma$ denotes the number of occurrences of the letter σ in w, is clearly neither polynomial, nor CP (the congruence "to have the same first letter" is not preserved). Fortunately f is not RCP when $n \geq 3$, and thus is not a counter-example to the result stated in [3], but it is RCP when $n = 2$. Thus, for words in Σ^*, we here have to use a new method, which also works even when $|\Sigma| = 2$ and which is very similar to the method used for trees, even though the proofs are more complex to take into account the associativity of the product (usually called concatenation) of words.

Most of the proofs of intermediate Lemmas and Propositions are identical for trees and for words or have only minor differences. Important differences, related to the associativity or non associativity of the product in the corresponding algebras, are located in the the proofs of just two Assumptions, that we prove separately.

The paper is thus organized as follows. In section 2, we recall the basics about algebras, polynomials and congruence preserving functions. In Section 3 we prove that the relation between the length of the value of a function and the length of its arguments is affine for both CP functions and polynomials. In Section 4 we define the main kind of congruences we will use and we show how to compute canonical representatives for these congruences. In section 5, we define polynomials associated with a CP function and prove that CP functions are polynomial under two Assumptions given in the previous section. In Section 6 (resp. 7) we prove these two Assumptions for the algebra of trees (resp. the free monoid). Section 7 ends with an application of the result on lengths of Section 3 which immediately implies the affine completeness of the free commutative monoid.

2. Binary algebras

Let Σ be a nonempty finite alphabet, whose letters will be denoted by $a, b, c, d, \ldots$.

We consider an algebraic structure $\langle \mathcal{A}(\Sigma), \star, \mathbf{0} \rangle$, with $\mathbf{0} \notin \Sigma$, subsuming both the free monoid and the set of binary trees, satisfying the following axioms (Ax-1), (Ax-2), (Ax-3)

(Ax-1) $\Sigma \cup \{\mathbf{0}\} \subseteq \mathcal{A}(\Sigma)$,

(Ax-2) if $u \notin \Sigma \cup \{\mathbf{0}\}$ then $\exists v, w \in \mathcal{A}(\Sigma) : u = v \star w$.

(Ax-3) there exists a mapping $|\cdot|\colon \mathcal{A}(\Sigma) \to \mathbb{N}$ such that

- $|\mathbf{0}| = 0$,
- $|\sigma| = 1$, for all $\sigma \in \Sigma$,
- $|u \star v| = |u| + |v|$.

$|u|$ is said to be the <u>length</u> of u (it is equal to the number of occurrences of letters of Σ in u). We similarly define, for $\sigma \in \Sigma$ and $u \in \mathcal{A}(\Sigma)$, $|u|_\sigma$ which is the number occurrences of the letter σ in u.

The free monoid and the algebra of binary trees are examples of such an algebra. If $\mathcal{A}(\Sigma)$ is the set of words Σ^* on the alphabet Σ, $\star$ is the (associative) concatenation of words, and $\mathbf{0}$ is the empty word ε, we get the free monoid. If $\mathcal{A}(\Sigma)$ is the set of binary trees whose leaves are labelled by letters of Σ, $t \star t'$ is a tree consisting of a root whose left subtree is t and whose right subtree is t', and $\mathbf{0}$ is the empty tree then we get the algebra of binary trees. In the case of trees the operation $\star$ is not associative. The free commutative monoid $\langle \mathbb{N}^p, +, (0, \ldots, 0)\rangle$ is also a binary algebra satisfying (Ax-1), (Ax-2), (Ax-3).

For our proofs the main difference between trees and the other examples relates to point (Ax-2) above: the decomposition $u = v \star w$ is unique for trees and not for the other examples.

Fact 2.1 (Unicity of decomposition). *If t is a tree not in $\{\mathbf{0}\} \cup \Sigma$ then there exists a unique ordered pair $\langle t_1, t_2\rangle \neq \langle \mathbf{0}, \mathbf{0}\rangle$ in $\mathcal{A}^2$ such that $t = t_1 \star t_2$.*

An element of $\mathcal{A}$ (a word or a tree) will be called an *object*.

2.1. Polynomials

We denote by $\mathcal{A}$ the set $\mathcal{A}(\Sigma)$. We also consider the infinite set of variables $X = \{x_i \mid i \geq 1\}$, disjoint from Σ. We denote by $\mathcal{A}_n$, the set $\mathcal{A}(\Sigma \cup \{x_1, \ldots, x_n\})$. Note that $\mathcal{A} = \mathcal{A}_0$ and that $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$.

Definition 2.2. A n-ary polynomial with variables $\{x_1, \ldots, x_n\}$ is an element P of $\mathcal{A}_n$. The multidegree of P is the n-tuple $\langle k_1, \ldots, k_n\rangle$ where $k_i = |P|_{x_i}$. With every such polynomial P we associate a n-ary polynomial function $\tilde{P} \colon \mathcal{A}^n \to \mathcal{A}$ defined by:
for any $\vec{u} = \langle u_1, \ldots, u_i, \ldots, u_n\rangle \in \mathcal{A}^n$,

$$\tilde{P}(\vec{u}) = \begin{cases} P & \text{if } P = \mathbf{0} \text{ or } P \in \Sigma \\ u_i & \text{if } P = x_i \\ \widetilde{P_1}(\vec{u}) \star \widetilde{P_2}(\vec{u}) & \text{if } P = P_1 \star P_2 \end{cases}$$

Note. In the case of words we have to prove that the value of $\widetilde{P}$ is independent of its decomposition $P = P_1 \star P_2$. This is due to the fact that $\widetilde{P}(\vec{u})$ can be seen as a homomorphic image of P by an homomorphism from $\mathcal{A}_n$ to $\mathcal{A}$.

From now on we simply write P instead of $\tilde{P}$ for denoting the function associated with the polynomial P.

2.2. Sub-objects

Let $\mathcal{A}_{1,1}$ be the set of degree 1 unary polynomials with variable y, i.e., elements $P \in \mathcal{A}(\Sigma \cup \{y\})$ such that $|P|_y = 1$, or objects of $\mathcal{A}(\Sigma \cup \{y\})$ with exactly one occurrence of y.

Definition 2.3. An element u of $\mathcal{A}$ is a sub-object of an element $t \in \mathcal{A}$, if there exists an occurrence of u inside t, formally: if there exists a polynomial $P \in \mathcal{A}_{1,1}$ such that $P(u) = t$.

In the case of words (resp. trees), sub-objects are factors (resp. subtrees).

Definition 2.4. A *sub-polynomial* Q of a polynomial $P \in \mathcal{A}^n$ is a sub-object of P.

2.3. Congruence preserving functions

Definition 2.5. A congruence on $\langle \mathcal{A}, \star, \mathbf{0} \rangle$ is an equivalence relation $\sim$ compatible with $\star$, i.e., $s_1 \sim s'_1$ and $s_2 \sim s'_2$ imply $s_1 \star s_2 \sim s'_1 \star s'_2$.

Definition 2.6. A function $f \colon \mathcal{A}^n \to \mathcal{A}$ is congruence preserving (abbreviated into CP) on $\langle \mathcal{A}, \star, \mathbf{0} \rangle$ if, for all congruences $\sim$ on $\langle \mathcal{A}, \star, \mathbf{0} \rangle$, for all $t_1, \ldots, t_n$, $t'_1, \ldots, t'_n$ in $\mathcal{A}$, $t_i \sim t'_i$ for all $i = 1, \ldots, n$, implies $f(t_1, \ldots, t_n) \sim f(t'_1, \ldots, t'_n)$.

Obviously, every polynomial function is CP. Our goal is to prove the converse, namely

Theorem 2.7. *Assume $|\Sigma| \geq 2$ for words and $|\Sigma| \geq 1$ for trees. If $f \colon \mathcal{A}(\Sigma)^n \to \mathcal{A}(\Sigma)$ is CP then there exists a polynomial P_f such that $f = \widetilde{P_f}$.*

This is the main result of the paper, which will be proven in Sections 5, 6 and 7.

3. Length condition

For polynomials, as a consequence of (Ax-3), we get:

Fact 3.1. *If $P \in \mathcal{A}_n$ is a polynomial of multidegree $\langle k_1, \ldots, k_n \rangle$ then*
$|P(u_1, \ldots, u_n)| = |P(\mathbf{0}, \ldots, \mathbf{0})| + \sum_{i=1}^{n} k_i.|u_i|$.

A necessary condition for a function $f \colon \mathcal{A}^n \to \mathcal{A}$ to be polynomial is that f has in someway a multidegree $\langle k_1, \ldots, k_n \rangle$, playing the rôle of the multidegree of polynomials, i.e., such that $|f(u_1, \ldots, u_n)| = |f(\mathbf{0}, \ldots, \mathbf{0})| + \sum_{i=1}^{n} k_i.|u_i|$. For words when $|\Sigma| \geq 3$, the existence of such a multidegree is proved in [3]. We here generalise this proof so that it also applies to trees and to smaller alphabets.

Lemma 3.2. *Let $f \colon \mathcal{A}(\Sigma)^n \to \mathcal{A}(\Sigma)$ be a n-ary CP function.*
(1) There exist functions $\lambda, \lambda_i \colon \mathbb{N}^n \to \mathbb{N}$ such that $|f(u_1, \ldots, u_n)| = \lambda(|u_1|, \ldots, |u_n|)$ and $|f(u_1, \ldots, u_n)|_i = \lambda_i(|u_1|_i, \ldots, |u_n|_i)$, for $i = 1, 2$.
(2) $\lambda(p_1 + q_1, \ldots, p_n + q_n) = \lambda_1(p_1, \ldots, p_n) + \lambda_2(q_1, \ldots, q_n)$.

Proof. For an object $u \in \mathcal{A}$, denote by $|u|_1 = |u|_a$ the number of occurrences of the letter a in u, and let $|u|_2 = |u| - |u|_1$. Formally, $|\varepsilon|_1 = 0$, $|a|_1 = 1$, $|\sigma|_1 = 0$ for $\sigma \neq a$, and $|t \star t'|_1 = |t|_1 + |t'|_1$.

(1) As the relation $|u| = |v|$ is a congruence and f is CP, $|u_i| = |v_i|$ for $i = 1, \ldots, n$ implies $|f(u_1, \ldots, u_n)| = |f(v_1, \ldots, v_n)|$ hence $|f(u_1, \ldots, u_n)|$ depends only on the lengths $|u_1|, \ldots, |u_n|$, and λ is well defined. Similarly for λ_i, $i = 1, 2$ as $|u|_i = |v|_i$ is also a congruence.

(2) Consider objects u_i with $|u_i|_1 = p_i$ and $|u_i|_2 = q_i$ (see Figure 1). On the one hand, $|f(u_1, \ldots, u_n)| = \lambda(|u_1|, \ldots, |u_n|) = \lambda(p_1 + q_1, \ldots, p_n + q_n)$, $|f(u_1, \ldots, u_n)|_1 = \lambda_1(p_1, \ldots, p_n)$ and $|f(u_1, \ldots, u_n)|_2 = \lambda_2(q_1, \ldots, q_n)$. On the other hand, $|f(u_1, \ldots, u_n)| = |f(u_1, \ldots, u_n)|_1 + |f(u_1, \ldots, u_n)|_2$, hence (2). $\square$

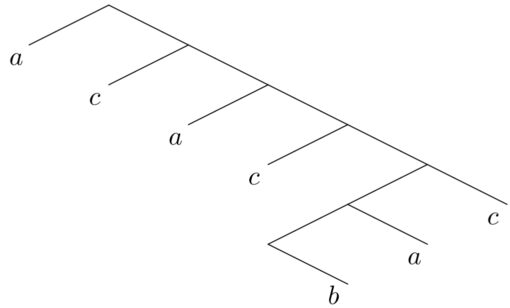


FIGURE 1. A tree u_i with $p_i = |u_i|_1 = 3$ and $q_i = |u_i|_2 = 4$.

Proposition 3.3. *For any n-ary CP function $f\colon \mathcal{A}(\Sigma)^n \to \mathcal{A}(\Sigma)$, with $|\Sigma| \geq 2$, there exists a n-tuple $\langle k_1, \ldots, k_n\rangle$ of natural numbers, called the* <u>*multidegree*</u> *of f, such that $|f(u_1, \ldots, u_n)| = |f(\mathbf{0}, \ldots, \mathbf{0})| + \sum_{i=1}^{n} k_i.|u_i|$.*

Proof. Let $\vec{e_i} = \langle \overbrace{0, \ldots, 0}^{(i-1)\text{ times}}, 1, 0, \ldots, 0\rangle$, $\vec{0} = \langle 0, \ldots, 0\rangle$, and apply Lemma 3.2. We have for any $m_1, \ldots, m_i, \ldots, m_n$,

$$\lambda(m_1, \ldots, m_i + 1, \ldots, m_n) = \lambda_1(m_1, \ldots, m_i, \ldots, m_n) + \lambda_2(\vec{e_i}),$$
$$\lambda(m_1, \ldots, m_i, \ldots, m_n) = \lambda_1(m_1, \ldots, m_i, \ldots, m_n) + \lambda_2(\vec{0}).$$

Subtracting

$$\lambda(m_1, \ldots, m_i + 1, \ldots, m_n) - \lambda(m_1, \ldots, m_i, \ldots, m_n) = \lambda_2(\vec{e_i}) - \lambda_2(\vec{0}).$$

Setting $k_i = \lambda_2(\vec{e_i}) - \lambda_2(\vec{0})$, we get

$$\lambda(m_1, \ldots, m_i, \ldots, m_n) - \lambda(m_1, \ldots, m_i - 1, \ldots, m_n) = k_i$$
$$\vdots$$
$$\lambda(m_1, \ldots, 1, \ldots, m_n) - \lambda(m_1, \ldots, 0, \ldots, m_n) = k_i$$

Summing up $\lambda(m_1, \ldots, m_i, \ldots, m_n) - \lambda(m_1, \ldots, 0, \ldots, m_n) = k_i m_i$

Iterating for all i, $\lambda(m_1, \ldots, m_n) - \lambda(\vec{0}) = k_1 m_1 + \cdots + k_n m_n$. □

Proposition 3.3 holds both for words and trees. However, for trees the following better result holds even when $|\Sigma| = 1$.

Proposition 3.4. *In the algebra of trees, for any n-ary CP function $f\colon \mathcal{A}(\Sigma)^n \to \mathcal{A}(\Sigma)$, there exists a n-tuple $\langle k_1, \ldots, k_n\rangle$ of natural numbers, called the* <u>*multidegree*</u> *of f, such that $|f(u_1, \ldots, u_n)| = |f(\mathbf{0}, \ldots, \mathbf{0})| + \sum_{i=1}^{n} k_i.|u_i|$.*

Proof. For a tree $u \notin \Sigma$, $|u|_1$ (resp. $|u|_2$) is the number of left (resp. right) leaves, so that $|u| = |u|_1 + |u|_2$ for $u \notin \Sigma$. On Figure 1 $|u_i|_1 = 4$ and $|u_i|_2 = 3$. Formally, $|\mathbf{0}| = |\mathbf{0}|_1 = |\mathbf{0}|_2 = 0$. For $u = t \star t' \notin \Sigma$ we have

$$|u|_1 = |t'|_1 + \begin{cases} 1 & \text{if } t \in \Sigma, \\ |t|_1 & \text{if } t \notin \Sigma. \end{cases} \quad \text{and } |u|_2 = |t|_2 + \begin{cases} 1 & \text{if } t' \in \Sigma, \\ |t'|_2 & \text{if } t' \notin \Sigma. \end{cases}$$

We already know that the relation $\sim$ defined by $u \sim v$ iff $|u| = |v|$ is a congruence. For $j = 1, 2$, the relation $\sim_j$ defined by $u \sim_j v$ iff either $u = v \in \Sigma$ or $u, v \notin \Sigma$ and $|u|_j = |v|_j$ is a congruence. Hence if $f = \mathcal{A}^n \to \mathcal{A}$ is CP then for all $u_1, \ldots, u_n, v_1, \ldots, v_n \notin \Sigma$ such that $\forall i = 1, \ldots, n, |u_i|_j = |v_i|_j$ and $f(u_1, \ldots, u_n), f(v_1, \ldots, v_n) \notin \Sigma$, we have $|f(u_1, \ldots, u_n)|_j = |f(v_1, \ldots, v_n)|_j$. Without loss of generality, we may assume that for all $u_1, \ldots, u_n$, $f(u_1, \ldots, u_n)$ is not in Σ. This holds because $g(u_1, \ldots, u_n) = \mathbf{0} \star f(u_1, \ldots, u_n)$ is CP and $|g(u_1, \ldots, u_n)| = |f(u_1, \ldots, u_n)|$.

For $u \notin \Sigma$, $|u| = |u|_1 + |u|_2$. Exactly as in Proposition 3.3 we show that for any $m_1, \ldots, m_i, \ldots, m_n$, $\lambda(m_1, \ldots, m_n) - \lambda(\vec{0}) = k_1 m_1 + \cdots + k_n m_n$. It follows that for all $u_1, \ldots, u_n \notin \Sigma$, $|f(u_1, \ldots, u_n)| = |f(\mathbf{0}, \ldots, \mathbf{0})| + \sum_{i=1}^{n} k_i.|u_i|$.

Finally, as for all $u \in \mathcal{A}$, $u \star \mathbf{0} \notin \Sigma$ and $|u \star \mathbf{0}| = |u|$, we have: $|f(u_1, \ldots, u_n)| = |f(u_1 \star \mathbf{0}, \ldots, u_n \star \mathbf{0})| = |f(\mathbf{0}, \ldots, \mathbf{0})| + \sum_{i=1}^{n} k_i.|u_i \star \mathbf{0}| = |f(\mathbf{0}, \ldots, \mathbf{0})| + \sum_{i=1}^{n} k_i.|u_i|$. $\square$

4. The toolbox

4.1. Congruent substitutions

If f is CP then $f(u) \sim f(v)$ as soon as $u \sim v$. This is why we introduce specific congruences $\sim_{u,v}$ such that $u \sim_{u,v} v$, so that if for some polynomial Q, (which is also CP), we know that for some u, $f(u) = Q(u)$, then we know that for all v, $f(v) \sim_{u,v} Q(v)$. Thus it is important to describe the congruence classes of such congruences.

Definition 4.1. For u, v a couple of objects in $\mathcal{A}$ the relation $\sim_{u,v}$ is the equivalence relation generated by the set of pairs $\{\langle P(u), P(v)\rangle \mid P \in \mathcal{A}_{1,1}\}$. $\sim_{u,v}$ is clearly a congruence on $\langle \mathcal{A}, \star, \mathbf{0}\rangle$.

Given such a congruence, we can consider the quotient algebra. It may happen that each congruence class has a simple canonical representative. For instance, the canonical representative could be the shortest object in the congruence class, provided it is unique. However unicity of the shortest representative certainly does not hold for the congruences $\sim_{u,v}$ when $|u| = |v|$. It also happens that unicity does not hold even when $|u| > |v|$ (Remark 4.2).

Remark 4.2. Even if $|u| > |v|$, there might be several shortest congruent elements. For instance in the case of words, $ab \sim_{aa,b} aaa \sim_{aa,b} ba$, hence ab and ba are two shortest elements congruent to aaa.

Definition 4.3. For a given element τ of $\mathcal{A}$, an element $t \in \mathcal{A}$ is τ-reducible, if τ is a sub-object of t. We denote by Θ_τ the set of all τ-irreducible objects in $\mathcal{A}$.

In Figure 2, Q_τ is τ-reducible, Q and P_τ are τ-irreducible, and in Figure 3, t'' is τ-irreducible.

We now extend Definition 4.3 of τ-irreducible objects in $\mathcal{A}$ to polynomials in $\mathcal{A}_n$.

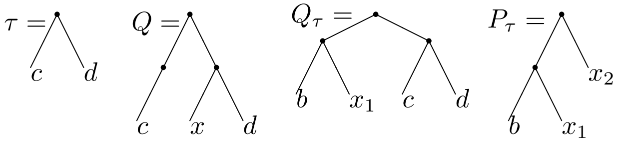


FIGURE 2. From left to right: tree $\tau = c \star d$, a τ-irreducible polynomial Q with variable x, a τ-reducible polynomial Q_τ with variable x_1 together with its associated τ-irreducible polynomial $P_\tau = Red^*_{\tau,x_2}(Q_\tau)$.

Definition 4.4. Let $\tau \in \mathcal{A}$. A polynomial $P \in \mathcal{A}_n$ is said to be τ-irreducible if any sub-object v of P which is in $\mathcal{A}$ is τ-irreducible.

Intuitively, the constant sub-objects ("coefficients") of P are τ-irreducible. In Figure 2, Q_τ is the only τ-reducible polynomial.

4.2. Canonical representatives

In fact it is possible to define and to "compute" a canonical representative t' of t for $\sim_{\tau,v}$ if $|\tau| > |v|$. To this end we stepwise replace every occurrence of τ inside t by v. To make this process deterministic we define the *reduct* $Red_{\tau,v}(t)$ obtained by replacing by v the "leftmost" occurrence of τ inside a τ-reducible object t.

Definition 4.5. (Definition of $Red_{\tau,v}(t)$.)
Case of trees If $t = \tau$ then $Red_{\tau,v}(t) = v$. Oherwise, since $t \neq \tau$ is τ-reducible, $|t| > |\tau| \geq 1$, hence, by (Ax-2), $t = t_1 \star t_2$, and at least one t_i is τ-reducible. Either $t_1 \in \mathcal{A}$ is τ-reducible, and then $Red_{\tau,v}(t) = Red_{\tau,v}(t_1) \star t_2$, or t_1 is τ-irreducible, then t_2 is τ-reducible and $Red_{\tau,v}(t) = t_1 \star Red_{\tau,v}(t_2)$. Figure 3 illustrates this reduction process.
Case of words Since τ is a factor of t, there exists a shortest prefix t' of t such that $t = t'\tau t''$. Then $Red_{\tau,v}(t) = t'vt''$.

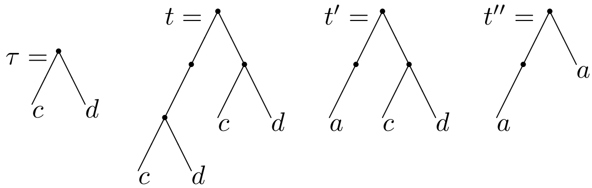


FIGURE 3. From left to right, $\tau = c\star d$, $t = ((c\star d)\star\mathbf{0})\star(c\star d)$, $t' = (a\star\mathbf{0})\star(c\star d)) = Red_{\tau,a}(t)$, $t'' = Red_{\tau,a}(t') = (a\star\mathbf{0})\star a)$.

We iterate this partial reduction function to get a mapping $Red^*_{\tau,v} : \mathcal{A} \to \Theta_\tau$ inductively defined by:

$$Red^*_{\tau,v}(t) = \begin{cases} t & \text{if } t \in \Theta_\tau \\ Red^*_{\tau,v}(Red_{\tau,v}(t)) & \text{if } t \notin \Theta_\tau. \end{cases}$$

Proposition 4.6. $Red^*_{\tau,v}(u \star w) = Red^*_{\tau,v}(Red^*_{\tau,v}(u) \star w)$.

Proof. By definition, $Red^*_{\tau,v}(t) = Red^k_{\tau,v}(t)$, where k is the least integer such that $Red^k_{\tau,v}(t)$ is τ-irreducible. If $Red^*_{\tau,v}(u \star w) = Red^p_{\tau,v}(u \star w)$ and $Red^*_{\tau,v}(u) = Red^q_{u,v}(u)$, necessarily $q \leq p$ and we have by induction on $i = 0, \ldots, q$, $Red^p_{\tau,v}(u \star w) = Red^{p-i}_{\tau,v}(Red^i_{\tau,v}(u) \star w)$ hence the result for $i = q$. □

Although $Red^*_{\tau,v}(t)$ is a canonical representative of the congruence class of t modulo $\sim_{\tau,v}$, it is not necessarily the only object of the equivalence class of t having minimal length, as shown in Remark 4.2.

To prevent such situations, we will first define for each algebra a suitably chosen subset $\mathcal{T}$ of the algebra ensuring that for each $\tau \in \mathcal{T}$, there exists a unique canonical representative of shortest length in the class of $\sim_{\tau,v}$ for each $v \in \mathcal{A}$ such that $|v| < |\tau|$ (Proposition 4.8). This set $\mathcal{T}$ has to satisfy the following assumption.

Assumption 4.7. $\forall \tau \in \mathcal{T},\ v \in \mathcal{A},\ P \in \mathcal{A}_{1,1},\ \ Red^*_{\tau,v}(P(\tau)) = Red^*_{\tau,v}(P(v))$.

Proposition 6.3 (resp. 7.1) shows that this assumption holds for the set $\mathcal{T}$ of trees defined by (6.1) in Section 6 (resp. the set $\mathcal{T}$ of words defined by (7.1) in Section 7).

Provided the truth of this assumption, we get:

Proposition 4.8. *(Existence of a canonical representative) Let $\tau \in \mathcal{T}$, and $v \in \mathcal{A}$ with $|\tau| > |v|$. For any $t, t' \in \mathcal{A}$, $t \sim_{\tau,v} t'$ iff $Red^*_{\tau,v}(t) = Red^*_{\tau,v}(t')$.*

Proof. By the definition of $Red^*_{\tau,v}$, for all t, t', $t \sim_{\tau,v} Red^*_{\tau,v}(t)$, and $t' \sim_{\tau,v} Red^*_{\tau,v}(t')$. Hence $Red^*_{\tau,v}(t) = Red^*_{\tau,v}(t')$ implies $t \sim_{\tau,v} t'$ by transitivity.

Conversely, if $t \sim_{\tau,v} t'$ then there exist $t_1 = t,\ t_2, \ldots, t_n = t'$, and $P_i \in \mathcal{A}_{1,1}$ (see Definition 4.1) such that for each $i = 1, \ldots, n-1$, $t_i = P_i(\tau)$ and $t_{i+1} = P_i(v)$ (or vice-versa). By Assumption 4.7, $Red^*_{\tau,v}(t_i) = Red^*_{\tau,v}(t_{i+1})$, hence $Red^*_{\tau,v}(t) = Red^*_{\tau,v}(t')$. □

Proposition 4.9. *Let $\tau \in \mathcal{T}$, t and t' be two objects such that $|v| < |\tau|$, $t \sim_{\tau,v} t'$, and $|t| < |\tau|$. Then $t = t'$ if and only if $|t| = |t'|$.*

Proof. If $t = t'$ then obviously $|t| = |t'|$. Since $t \sim_{\tau,v} t'$, by Proposition 4.8, $Red^*_{\tau,v}(t) = Red^*_{\tau,v}(t')$. But $|t'| = |t| < |\tau|$ implies that both t' and t are τ-irreducible, hence $t = Red^*_{\tau,v}(t) = Red^*_{\tau,v}(t') = t'$. □

4.3. Strong irreducibility

By Propositions 4.8 and 4.9, we get that if $|t| < |\tau|$ and $|Red^*_{\tau,v}(t')| > |\tau|$ then $t \not\sim_{\tau,u} t'$. To prove that if $|t'| > |\tau|$ then $|Red^*_{\tau,v}(t')| > |\tau|$, it is enough to prove that if t' contains a sub-object w of length $n \geq |\tau|$ then w is a sub-object of $Red^*_{\tau,v}(t')$. This leads to the following definition.

Definition 4.10. Let $\tau \in \mathcal{A}$, an object w is said to be $\underline{\text{strongly } \tau\text{-irreducible}}$ if $|w| \geq |\tau|$ and if whenever w is a sub-object of some $t \in \mathcal{A}$, w also is a sub-object of $Red^*_{\tau,v}(t)$ for any v such that $|v| < |\tau|$.

We finally state the following assumption on $\mathcal{T}$, the truth of which is proven in Proposition 6.4 (resp. 7.3) for trees (resp. for words).

Assumption 4.11. *For all $\tau \in \mathcal{T}$ and for all τ-irreducible unary polynomials P of degree k such that $|\tau| \geq 2k+4$, we have the following property:*

If for all $u \in \mathcal{A}$ such that $|u| \leq 1$, $P(u)$ is τ-reducible, then there exists $\theta \in \mathcal{A}$ of length 1 and a strongly τ-irreducible sub-object w of $P(\theta)$ of length not less than $|\tau|$ (i.e., $|w| \geq |\tau|$).

5. Proof of the main Theorem

From now on, we postulate the existence of a set $\mathcal{T}$ which satisfies Assumptions 4.7 and 4.11.

5.1. The induction hypothesis

The polynomiality of CP functions will be proved by induction on their arity. The basic step of this induction is obvious and common to all algebras we consider: a function of arity 0 is a constant, which is a polynomial function.

For the inductive step, note that if $n \geq 0$ and f is a $(n+1)$-ary CP function of multidegree $\langle k_1, \ldots, k_n, k_{n+1} \rangle$, then for all t, f_t defined by $f_t(u_1, \ldots, u_n) = f(u_1, \ldots, u_n, t)$ is CP with multidegree $\langle k_1, \ldots, k_n \rangle$, hence the induction hypothesis:

Fact 5.1. **Induction hypothesis.** *For any $t \in \mathcal{A}$, there exists a polynomial Q_t of multidegree $\langle k_1, \ldots, k_n \rangle$ such that:*

$$\forall u_1, \ldots, u_n \in \mathcal{A}, \quad Q_t(u_1, \ldots, u_n) = f(u_1, \ldots, u_n, t).$$

Definition 5.2. The *polynomial P_τ associated with f and $\tau \in \mathcal{T}$* is the unique τ-irreducible polynomial of multidegree $\langle k_1, \ldots, k_n, m \rangle$ such that

$$\forall u_1, \ldots, u_n \in \mathcal{A}, \ P_\tau(u_1, \ldots, u_n, \tau) = Q_\tau(u_1, \ldots, u_n) = f(u_1, \ldots, u_n, \tau).$$

It is also defined by $P_\tau = Red^*_{\tau, x_{n+1}}(Q_\tau)$, considering P_τ and Q_τ as objects in $\mathcal{A}(\Sigma \cup \{x_1, \ldots, x_n, x_{n+1}\})$.

Figure 2 illustrates this definition in the algebra of binary trees.

5.2. Partial polynomiality of CP functions

Assuming the hypothesis stated in Fact 5.1, we can proceed and prove

Proposition 5.3. *Let $\tau \in \mathcal{T}$. If $|u| < |\tau|$ and if $|f(u_1, \ldots, u_n, u)| < |\tau|$ then*

- $f(u_1, \ldots, u_n, u) = Red^*_{\tau,u}(P_\tau(u_1, \ldots, u_n, u))$
- *either $m = k_{n+1}$ and $f(u_1, \ldots, u_n, u) = P_\tau(u_1, \ldots, u_n, u)$, or $m < k_{n+1}$ and $P_\tau(u_1, \ldots, u_n, u)$ is τ-reducible.*

Proof. Obviously, $f(u_1, \ldots, u_n, u) \sim_{\tau,u} f(u_1, \ldots, u_n, \tau) = P_\tau(u_1, \ldots, u_n, \tau) \sim_{\tau,u} P_\tau(u_1, \ldots, u_n, u)$. As $|f(u_1, \ldots, u_n, u)| < |\tau|$, $f(u_1, \ldots, u_n, u)$ is τ-irreducible. Thus, by Assumption 4.7, $f(u_1, \ldots, u_n, u) = Red^*_{\tau,u}(P_\tau(u_1, \ldots, u_n, u))$. Let $d = |f(u_1, \ldots, u_n, \tau)| = |P_\tau(u_1, \ldots, u_n, \tau)|$. Then $|f(u_1, \ldots, u_n, u)| = d - k_{n+1}(|\tau| - |u|)$ and $|P_\tau(u_1, \ldots, u_n, u)| = d - m(|\tau| - |u|)$.

By Proposition 4.9, $P_\tau(u_1,\dots,u_n,u) = f(u_1,\dots,u_n,u)$ if and only if $|P_\tau(u_1,\dots,u_n,u)| = |f(u_1,\dots,u_n,u)|$ if and only if $m = k_{n+1}$.

Since $f(u_1,\dots,u_n,u) = Red^*_{\tau,u}(P_\tau(u_1,\dots,u_n,u))$, if $f(u_1,\dots,u_n,u) \neq P_\tau(u_1,\dots,u_n,u)$ then $P_\tau(u_1,\dots,u_n,u)$ is not τ-irreducible.
Hence $d - m(|\tau| - |u|) = |P_\tau(u_1,\dots,u_n,u)| \geq |\tau| > |f(u_1,\dots,u_n,u)| = d - k_{n+1}(|\tau| - |u|)$, which implies $m < k_{n+1}$. □

An immediate consequence of Proposition 5.3 is:

Proposition 5.4. *Let $\tau \in \mathcal{T}$, let $\langle k_1,\dots,k_n,m\rangle$ be the multidegree of P_τ. Then*

(1) either $m = k_{n+1}$ and for all $u \in \mathcal{A}$ such that $|u| \leq |\tau|$, and for all $u_1,\dots,u_n \in \mathcal{A}$ such that $|f(u_1,\dots,u_n,u)| < |\tau|$, we have $P_\tau(u_1,\dots,u_n,u) = f(u_1,\dots,u_n,u)$,

(2) or $m < k_{n+1}$ and for all $u \in \mathcal{A}$ such that $|u| \leq |\tau|$, and for all $u_1,\dots,u_n \in \mathcal{A}$ such that $|f(u_1,\dots,u_n,u)| < |\tau|$, $P_\tau(u_1,\dots,u_n,u)$ is τ-reducible.

5.3. Polynomiality of CP functions

We first prove that for almost all τ we are in case (1) of Proposition 5.4.

Proposition 5.5. *Let $\langle k_1,\dots,k_n,k_{n+1}\rangle$ be the multidegree of f, let $k = k_1 + \cdots + k_n + k_{n+1}$, and let $\tau \in \mathcal{T}$ be such that $\tau \geq 2k+4$. For all $u \in \mathcal{A}$ such that $|u| < |\tau|$ and for all $u_1,\dots,u_n \in \mathcal{A}$ such that $|f(u_1,\dots,u_n,u)| < |\tau|$, we have $P_\tau(u_1,\dots,u_n,u) = f(u_1,\dots,u_n,u)$.*

Proof. By Proposition 5.4 it is enough to prove that $m < k_{n+1}$ is impossible.

Let P_τ be the τ-irreducible polynomial associated with τ of multidegree $\langle k_1,\dots,k_n,m\rangle$ and let us assume that $m < k_{n+1}$. Then, by Proposition 5.4, we have: for all $u \in \mathcal{A}$ such that $|u| \leq |\tau|$ and $|f(u,\dots,u,u)| < |\tau|$, the object $P_\tau(u,\dots,u,u)$ is τ-reducible.

We now consider the τ-irreducible unary polynomial P'_τ of degree $M = k_1 + \cdots + k_n + m < k$, obtained by substituting x_1 for any variable x_i in P_τ. Since $P'_\tau(u)$ is τ-reducible for all u such that $|u| \leq 1 < |\tau|$, by Assumption 4.11 there exist θ of length 1 and a strongly τ-irreducible sub-object w of $P'_\tau(\theta) = P_\tau(\theta,\dots,\theta,\theta)$ of length not less than τ. By Proposition 5.3, w is a sub-object of $Red^*_{\tau,\theta}(P_\tau(\theta,\dots,\theta,\theta)) = f(\theta,\dots,\theta,\theta)$. Hence $|w| \leq |f(\theta,\dots,\theta,\theta)| < |\tau| \leq |w|$, a contradiction. □

Let τ_1 and τ_2 be such that $|\tau_i| > |f(a,\dots,a)|$. Then, by Proposition 5.5, we have :
For all $u_1,u_2,\dots,u_n,u$ such that $|u|$ and $|f(u_1,\dots,u_n)|$ are less that $|\tau_1|$ and $|\tau_2|$ then

$$P_{\tau_1}(u_1,\dots,u_n,u) = \quad f(u_1,\dots,u_n,u) \quad = P_{\tau_2}(u_1,\dots,u_n,u). \tag{5.1}$$

We first prove that $P_{\tau_1} = P_{\tau_2}$ as a consequence of the next Proposition by observing that equation (5.1) holds for all u_i, u of length 1.

Proposition 5.6. *Let P, Q be polynomials of multidegree $\langle k_1,\dots,k_n\rangle$.*
If, for all $u_1,u_2,\dots,u_n$ of length 1, $P(u_1,\dots,u_n) = Q(u_1,\dots,u_n)$ then $P = Q$.

Proof. For a polynomial P in the algebra of trees, we define $s(P)$ to be the number of symbols of $\Sigma \cup \{\star\} \cup \{x_1, \dots, x_n\}$ occurring in P. Formally $s(\mathbf{0}) = 0$, $s(a) = 1$ for $a \in \Sigma \cup \{x_1, \dots, x_n\}$, and $s(u \star v) = 1 + s(u) + s(v)$. For P in the algebra of words, we set $s(P) = |P|$.

In both cases there exists at least two distinct objects of length 1: either two distinct letters a, b, or the trees $a \star \mathbf{0}$ and $\mathbf{0} \star a$.

The proof is by induction on $s(P)$.

Basis.

(1) If $s(P) = s(Q) = 0$ then $P = \mathbf{0} = Q$.

(2) If $s(P) = s(Q) = 1$ then $P, Q \in \Sigma \cup \{x_1, \dots, x_n\}$. If P and Q are both constants, the result follows from equality $P(u, \dots, u) = Q(u, \dots, u)$. If $P = x_i$ and $Q = x_j$ with $i \neq j$, the hypothesis $P(u_1, \dots, u_n) = Q(u_1, \dots, u_n)$ leads to a contradiction, as soon as $u_i \neq u_j$, hence $i = j$. If P is a constant u and Q is a variable x_i, we have $u = P(u', \dots, u') = Q(u', \dots, u') = u'$, a contradiction when $u \neq u'$.

Inductive step. If $s(P) > 1$ then $P = P_1 \star P_2$ and $Q = Q_1 \star Q_2$, (taking $|P_1| = |Q_1| = 1$ in case of words). For any $u_1, u_2, \dots, u_n$ of length 1, we have $Q(u_1, \dots, u_n) = P(u_1, \dots, u_n) = P_1(u_1, \dots, u_n) \star P_2(u_1, \dots, u_n) = Q_1(u_1, \dots, u_n) \star Q_2(u_1, \dots, u_n)$ which implies $P_i(u_1, \dots, u_n) = Q_i(u_1, \dots, u_n)$, hence, by the induction hypothesis, $P_1 = Q_1$ and $P_2 = Q_2$, and thus $P = Q$. $\square$

Theorem 5.7. *Let f be a CP function of multidegree $\langle k_1, \dots, k_n, k_{n+1} \rangle$. There exists a polynomial P_f of multidegree $\langle k_1, \dots, k_n, k_{n+1} \rangle$ such for all $u_1, \dots, u_n$, $u \in \mathcal{A}$, $P_f(u_1, \dots, u_n, u) = f(u_1, \dots, u_n, u)$.*

Proof. By Propositions 5.5 and 5.6 there exists a unique polynomial P_f such that for all τ of length greater than $|f(a, a, \dots, a)|$, $P_\tau = P_f$. For any $u_1, \dots, u_n, u$ there exists τ such that $|\tau| > \max(|u|, |f(u_1, \dots, u_n, u)|)$. By Proposition 5.5, $f(u_1, \dots, u_n, u) = P_\tau(u_1, \dots, u_n, u) = P_f(u_1, \dots, u_n, u)$. $\square$

6. The case of trees

We here consider the algebra of binary trees with labelled leaves. For this algebra of trees we set

$$\mathcal{T} \quad = \quad \{\, \tau \in \mathcal{A} \mid |\tau| \geq 2 \,\} \tag{6.1}$$

Proposition 6.1. *If a tree w is τ-irreducible, then it is strongly τ-irreducible.*

Proof. By definition of $Red^*_{\tau,v}$, it is enough to show that if w is a subtreee of t then it is a subtree of $Red_{\tau,v}(t)$. The proof is by induction on $|t|$ such that w is a subtree of t. If t is τ-irreducible then $Red_{\tau,v}(t) = t$ and the result is proved. Otherwise, $t = t_1 \star t_2$, with w subtree of some t_i, and $Red_{\tau,v}(t) = Red_{\tau,v}(t_1) \star t_2$ or $Red_{\tau,v}(t) = t_1 \star Red_{\tau,v}(t_2)$. In both cases, w is a subtree of $Red_{\tau,v}(t)$. $\square$

6.1. Canonical representative

For trees, we can improve Proposition 4.6.

Proposition 6.2. $Red^*_{\tau,v}(u \star w) = Red^*_{\tau,v}(Red^*_{\tau,v}(u) \star Red^*_{\tau,v}(w))$.

Proof. By taking Proposition 4.6 into account, we just have to prove that $Red^*_{\tau,v}(u \star w) = Red^*_{\tau,v}(u \star Red^*_{\tau,v}(w))$ when u is τ-irreducible. This a consequence of the definition of the leftmost reduction for trees: $Red_{\tau,v}(u \star w) = u \star Red_{\tau,v}(w)$. □

We now prove that Assumption 4.7 holds for our algebra of binary trees.

Proposition 6.3. $\forall P \in \mathcal{A}_{1,1} \qquad Red^*_{\tau,v}(P(\tau)) = Red^*_{\tau,v}(P(v))$.

Proof. The proof is by induction on $|P|$. If $P = y$ then $Red^*_{\tau,v}(\tau) = Red^*_{\tau,v}(v) = v$.

If $P = P_1 \star P_2$ then by Proposition 6.2,

$$Red^*_{\tau,v}(P(\tau)) = Red^*_{\tau,v}(Red^*_{\tau,v}(P_1(\tau)) \star Red^*_{\tau,v}(P_2(\tau))), \text{and}$$
$$Red^*_{\tau,v}(P(v)) = Red^*_{\tau,v}(Red^*_{\tau,v}(P_1(v)) \star Red^*_{\tau,v}(P_2(v))).$$

Then, by the induction hypothesis, $Red^*_{\tau,v}(P_i(v)) = Red^*_{\tau,v}(P_i(\tau))$, for $i = 1, 2$, and thus $Red^*_{\tau,v}(P(v)) = Red^*_{\tau,v}(P(\tau))$. □

6.2. Strongly irreducible trees

The following Proposition assures that Assumption 4.7 holds for trees.

Proposition 6.4. *For all $\tau \in \mathcal{T}$ and for all τ-irreducible unary polynomials P the following property holds.*

If for all $u \in \mathcal{A}$ such that $|u| \leq 1$, $P(u)$ is τ-reducible, then there exists $\theta \in \mathcal{A}$ of length 1 and a strongly τ-irreducible subtree w of $P(\theta)$ of length not less than $|\tau|$ (i.e., $|w| \geq |\tau|$).

Proof. Let $\tau \in \mathcal{T}$, which has length at least 2. Let P be a non constant τ-irreducible polynomial such that for all $u \in \mathcal{A}$ with length $|u| \leq 1$, $P(u)$ is τ-reducible. Let $\sigma \in \Sigma$, and let $t = \sigma \star \mathbf{0}$ and $t' = \mathbf{0} \star \sigma$, $t \neq t'$.

As $P(t)$ is τ-reducible, it must contain τ. But since P is τ-irreducible, there exists a non constant sub-polynomial Q of P such that $Q(t) = \tau$. Then $|Q(t)| = |Q(t')| = |\tau|$ and, as Q is non-constant, $Q(t') \neq \tau$. It follows that $Q(t')$ is τ-irreducible, hence strongly τ-irreducible by Proposition 6.1. We set $\theta = t'$ and $w = Q(t')$. □

7. The case of words

For words, proving Assumptions 4.7 and 4.11 requires more work because unicity of the decomposition fails in the free monoid.
As shown in Remark 4.2, Assumption 4.7 does not hold for any word τ. Indeed, Assumption 4.7 fails as soon as τ self-overlaps, i.e., when there exists a word t which is a both a strict prefix and a strict suffix of τ. For instance, if $\tau = aba$, $ab \sim_{aba,\varepsilon} ababa \sim_{aba,\varepsilon} ba$, while $Red_{aba,\varepsilon}(ab) = ab \neq ba =$

$Red_{aba,\varepsilon}(ba)$. Obviously, words such that a^nb^n do not self-overlap and thus satisfy Assumption 4.7. But we also need that these words satisfy Assumption 4.11. The condition that τ is not self-overlapping is not sufficient to satisfy Assumption 4.11. For instance, let $\tau = aabb$ and $P = aax_1bb$, which is τ-irreducible. The factors of length ≥ 4 of $P(a) = aaabb$ and $P(b) = aabbb$ are $aaabb$, $aabbb$, $aabb$, $aaab$, $abbb$. None of them is strongly τ-irreducible: $aaabb$, $aabbb$, $aabb$ are τ-reducible, and $aaab$, $abbb$ satisfy one of the forbidden property (1) or (2) of Proposition 7.2. We thus have to introduce a stronger constraint to define a suitable $\mathcal{T}$, which turns out to be

$$\mathcal{T} \quad = \quad \{a^nbab^n \mid n > 1\} \tag{7.1}$$

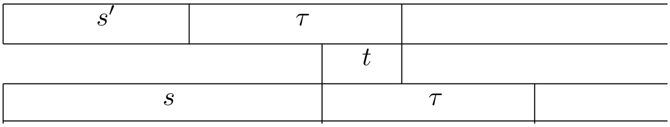
7.1. Canonical representative

Proposition 7.1. *For all P in $\mathcal{A}_{1,1}$* $\qquad Red^*_{\tau,v}(P(\tau)) = Red^*_{\tau,v}(P(v)).$

Proof. The proof is by induction on $|P|$.

Basis. If $P = y$ then $Red^*_{\tau,v}(\tau) = Red^*_{\tau,v}(v) = v$.

Induction. Let $P = uyw$ and let $s = Red^*_{\tau,v}(u) \in \Theta_\tau$. By Proposition 4.6, $Red^*_{\tau,v}(P(\tau)) = Red^*_{\tau,v}(s\tau w)$ and $Red^*_{\tau,v}(P(v)) = Red^*_{\tau,v}(svw)$. Thus, to prove the result it is enough to show that $Red_{\tau,v}(s\tau w) = svw$, i.e., that the shortest prefix $s\tau$ of $s\tau w$ is $s\tau$. Let us assume that there exists s' such that $s'\tau$ is a strict prefix of $s\tau$. Since since $s \in \Theta_\tau$, $s'\tau$ is not a prefix of s.



It follows that there exists a nonempty word t, with $0 < |t| < |\tau|$, which is both a suffix and a prefix of $\tau = a^nbab^n$, such that $s'\tau = st$.

The first letter of t has to be a and its last letter b. Therefore a^nb is a prefix of t and ab^n is a suffix of t, hence $t = a^nbab^n$, contradicting $|t| < |\tau|$. □

7.2. Strongly irreducible words

We state a sufficient condition for a word $w \in \mathcal{A}$ to be strongly τ-irreducible.

Proposition 7.2. *A nonempty word w is strongly τ-irreducible if it is τ-irreducible and it has the additional properties that τ and w do not overlap, i.e., there do not exist words u, t', t such that $t \notin \{\varepsilon, \tau\}$ and*

(1) either $w = ut$ and $\tau = tt'$,
(2) or $\tau = t't$ and $w = tu$.

Proof. It is enough to show that if a factor w of t satisfies the above hypothesis, then w is a factor of $Red_{\tau,v}(t)$ when $|v| < |\tau|$.

Let $t = w'\tau w''$ with w' τ-irreducible. Then $Red_{\tau,v}(t) = w'vw''$. As w is τ-irreducible and w and τ do not overlap, if w is a factor of t, it is a factor of w' or a factor of w'', hence a factor of $Red_{\tau,v}(t) = w'vw''$. □

The following proposition implies Assumption 4.11.

Proposition 7.3. *For all $\tau = a^n bab^n \in \mathcal{T}$ and for all τ-irreducible unary polynomials P of degree k such that $|\tau| \geq 2k+4$, the following property holds.*

If $P(\varepsilon)$ is τ-reducible, then there exists $\theta \in \{a,b\}$ and a strongly τ-irreducible sub-object w of $P(\theta)$ of length greater than $|\tau|$ (i.e., $|w| > |\tau|$).

Proof. Let $\tau = a^n bab^n \in \mathcal{T}$ and let P be a τ- irreducible polynomial of degree k such that $P(\varepsilon), P(a)$, and $P(b)$ are τ-reducible. Note that since $|\tau| = 2n+2$ the condition $|\tau| \geq 2k+4$ is equivalent to $n-1 > k$.

Since τ is a factor of $P(\varepsilon)$ there exists a factor Q of P such that $Q(\varepsilon) = \tau$, i.e.,

$$Q = ax^{p_1}ax^{p_2}a \cdots ax^{p_n}bx^m ax^{q_1}bx^{q_2}b \cdots x^{q_n}b$$

with $k = p+m+q < n-1$, where $p = p_1+p_2+\cdots+p_n$ and $q = q_1+q_2+\cdots+q_n$.

We show that at least one of the words $Q(a)$ or $Q(b)$ is strongly τ-irreducible.

We first show that if $Q(a) = a^{n+p}ba^{1+m+q_1}ba^{q_2}b \cdots a^{q_n}b$ is not strongly τ-irreducible, then $m = q = 0$.

If $Q(a)$ is not strongly τ-irreducible, then it is either τ-reducible and we are in case (i) below, or it is τ-irreducible and then we are in one of cases (ii) or (iii) below.

(i) $Q(a)$ is τ-reducible, i.e., $\exists u, v$ such that: $Q(a) = u\tau v$, or
(ii) $Q(a) = ut$ and $\tau = tv$, with $v \neq \varepsilon \neq t$ (Proposition 7.2 (1)), or
(iii) $Q(a) = tv$ and $\tau = ut$, with $u \neq \varepsilon \neq t$ (Proposition 7.2 (2)).

For both Cases (ii) and (iii), as both $Q(a)$ and τ start with a and end with b, the first letter of t is a and its last letter is b.

Case(i) If τ is a factor of $Q(a)$ then bab^n is a factor of $Q(a)$. The only factor of $Q(a)$ starting and ending with b, ending with b, and containing $(n+1)$ b's is $ba^{1+m'+m_1}ba^{m_2}b \cdots a^{m_n}b$, which implies $m' + m_b = 0$.

Case(ii) Assume now $\exists u, v, t$ with $Q(a) = ut$ and $\tau = tv$, with $v \neq \varepsilon$. As t is a prefix of τ, we have $t = a^n b$ or $t = a^n bab^{n'}$ with $0 < n' < n$. Since t is a suffix of $Q(a)$, in all cases, $a^n b$ is a factor of $Q(a)$. As for all i $q_i \leq q < n-1$ and, since $1+m+q_1 \leq 1+p+m+q < 1+(n-1) = n$, the unique suffix of $Q(a)$ starting with $a^n b$ is $t = a^n ba^{1+m+q_1}ba^{q_2}b \cdots a^{q_n}b$. Since t is a prefix of τ, we have $n+1+m+q = |t|_a \leq |\tau|_a = n+1$, which implies $m = q = 0$.

Case(iii) Assume now $\exists u, v, t$ with $Q(a) = tv$ and $\tau = ut$, with $u \neq \varepsilon$. Since t is a suffix of τ, then either $t = ab^n$ or $t = a^{n'}bab^n$ with $0 < n' < n$. Since t is a prefix of $Q(a)$, $a^{n+p}b$ is also a prefix of t. Both cases are impossible since $n + p > n' \geq 1$.

Hence if $Q(a)$ is not strongly τ-irreducible, $m = q = 0$.

By a symmetrical reasoning on $Q(b) = ab^{p_1}ab^{p_2} \cdots ab^{p_n+m+q}ab^{q_n+n}$ we get that if $Q(b)$ is not strongly τ-irreducible, then $p = m = 0$.

Finally, if both $Q(a)$ and $Q(b)$ are not strongly τ-irreducible then $p = m = q = 0$, hence τ is a factor of P, contradicting the τ-irreducibility of P. Thus, either $Q(a)$ or $Q(b)$ is strongly τ-irreducible. Then choose $\theta \in \{a,b\}$ such that $w = Q(\theta)$ is strongly τ-irreducible. $\square$

Hence, Theorem 2.7 holds and if $|\Sigma| \geq 2$ then Σ^* is affine complete. Our proof method can be extended to the free commutative monoid with p generators when $p \geq 2$ as shown in the next subsection.

7.3. Application to free commutative monoids

Note that the free commutative monoid with p generators is isomorphic to $\mathbb{N}^p$. We now prove a variant of Proposition 3.3 which immediately implies that the commutative binary algebra $\langle \mathbb{N}^p, +, \vec{0} \rangle$ is affine complete, thus giving a very simple proof of already known results [5, 7].

For $u = \langle \ell_1, \ldots, \ell_p \rangle \in \mathbb{N}^p$ let $|u| = \ell_1 + \cdots + \ell_p$ and $|u|_j = \ell_j$ for $i = 1, \ldots, p$.

Proposition 7.4. *For any n-ary CP function $f \colon \mathcal{A}(\mathbb{N}^p)^n \to \mathbb{N}^p)$, with $p \geq 2$, there exists a n-tuple $\langle k_1, \ldots, k_n \rangle$ of natural numbers, called the multidegree of f, such that*

(i) $|f(u_1, \ldots, u_n)| = |f(\mathbf{0}, \ldots, \mathbf{0})| + \sum_{i=1}^n k_i.|u_i|$, and

(ii) for all $j = 1, \ldots, p$, $|f(u_1, \ldots, u_n)|_j = |f(\mathbf{0}, \ldots, \mathbf{0})|_j + \sum_{i=1}^n k_i.|u_i|_j$

Proof. The proof is almost identical to the proof of Proposition 3.3. We stress here the differences. For an object $u = \langle \ell_1, \ldots, \ell_p \rangle \in \mathbb{N}^p$, and an arbitrary element $j \in \langle 1, \ldots, p \rangle$, let us denote: $|u| = \ell_1 + \cdot + \ell_p$, $|u|_1 = \ell_j$, and $|u|_2 = |u| - |u|_1$. There exist λ, λ_1 such that $\lambda(m_1, \ldots, m_n)$ is the common value of all $|f(u_1, \ldots, u_n)|$ and $\lambda_1(m_1, \ldots, m_n)$ is the common value of all $|f(u_1, \ldots, u_n)|_1 = \ell_j$ for an arbitrary $j \in \{1, \ldots, p\}$. Lemma 3.2 and (i) are then proved as in Proposition 3.3. Moreover

$$\begin{aligned}
\lambda(m_1, \ldots, m_n) &= \lambda_1(m_1, \ldots, m_n) + \lambda_2(0, \ldots, 0) \\
&= \lambda_1(m_1, \ldots, m_n) - \lambda_1(0, \ldots, 0) + \lambda_1(0, \ldots, 0) + \lambda_2(0, \ldots, 0) \\
&= \lambda_1(m_1, \ldots, m_n) - \lambda_1(0, \ldots, 0) + \lambda(0, \ldots, 0) \text{ [Lemma 3.2 2]}
\end{aligned}$$

Hence $\lambda(m_1, \ldots, m_n) - \lambda(0, \ldots, 0) = \lambda_1(m_1, \ldots, m_n) - \lambda_1(0, \ldots, 0)$ which, as λ_1 can be any arbitrarily chosen λ_j, immediately implies (ii). □

Corollary 7.5. *The commutative algebra $\langle \mathbb{N}^p, +, 0 \rangle$ is affine complete.*

Proof. Proposition 7.4 (ii) means that the jth component $|f(x_1, \ldots, x_n)|_j$ of $f(x_1, \ldots, x_n)$ is of the form $c_j + \sum_{i=1}^n k_i.|x_i|_j$, for all $j = 1, \ldots, p$. Hence $f(x_1, \ldots, x_n) = c + \sum_{i=1}^n k_i.x_i$ is indeed a polynomial. □

8. Conclusion

It is known that, when the alphabet has just one letter, the free monoid is not affine complete [2]. It is also known that, when the alphabet has at least two letters, the free commutative monoid is affine complete since it is isomorphic to a free module or a vector space of dimension at least 2, known to be affine complete [5, 7].

We here prove that the (non commutative) free monoid Σ^* is affine complete as soon as its alphabet has at least two letters (generalizing [3] where the result was proved for $|\Sigma| \geq 3$).

We also prove that the algebra of binary trees with labelled leaves is affine complete for every nonempty finite alphabet Σ, i.e., not assuming that $|\Sigma| \geq 2$. This difference with the case of the free monoid might seem surprising. However since its product is not associative, the algebra of trees has more structure, hence more congruences, and thus less CP functions, than the free monoid.

References

[1] Arnold A., Cégielski P., Grigorieff S., Guessarian I.: Affine completeness of the algebra of full binary trees. Algebra Universalis, Springer Verlag, **81**, `https://doi.org/10.1007/s00012-020-00690-6(2020)81:55`

[2] Cégielski, P., Grigorieff, S., Guessarian, I.: Newton representation of functions over natural integers having integral difference ratios. International Journal of Number Theory, 11 (7), 2019–2139 (2015)

[3] Cégielski P., Grigorieff S., Guessarian I.: Congruence preserving functions on free monoids. Algebra Universalis, Springer Verlag, 78 (3), 389–406 (2017)

[4] Kaarli K., Pixley A.F.: Polynomial Completeness in Algebraic Systems. Chapman & Hall/CRC (2001)

[5] Nöbauer W.: Affinvollständige Moduln. Mathematische Nachrichten, **86**, 85–96 (1978)

[6] Ploščica M., Haviar M.: Congruence-preserving functions on distributive lattices. Algebra Universalis, **59**, 179–196 (2008)

[7] Werner H.: Produkte von KongruenzenKlassengeometrien universeller Algebren. Math. Z. **121**, 111–140 (1971)