

November 10, 2020

# The algebra of binary trees is affine complete

André Arnold, Patrick Cégielski, Serge Grigorieff and Irène Guessarian\*

**Abstract.** A function on an algebra is congruence preserving if, for any congruence, it maps pairs of congruent elements onto pairs of congruent elements. We show that on the algebra of binary trees whose leaves are labeled by letters of an alphabet containing at least three letters, a function is congruence preserving if and only if it is polynomial.

## 1. The algebra of binary trees

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 a 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 polynomial. In [1] we have proved that every unary CP function  $f : T(\Sigma) \rightarrow T(\Sigma)$  is polynomial, where  $T(\Sigma)$  is the algebra of full binary trees with leaves labelled by letters of an alphabet  $\Sigma$  having at least 3 letters. We here generalize this result proving that any CP function  $f : \mathcal{T}(\Sigma)^n \rightarrow \mathcal{T}(\Sigma)$  is polynomial, where  $\mathcal{T}(\Sigma)$  is the algebra of arbitrary binary trees (possibly non full) with labelled leaves. This generalization is twofold (1) any binary tree is allowed in  $\mathcal{T}(\Sigma)$ , and (2) CP functions of arbitrary arity are allowed. This exhibits an example of a non commutative and non associative affine complete algebra.

### 1.1. Trees, congruences

For an algebra  $\mathcal{A}$  with domain  $A$ , a *congruence*  $\sim$  on  $\mathcal{A}$  is an equivalence relation on  $A$  which is compatible with the operations of  $\mathcal{A}$ . We state the characterization of congruences by kernels of homomorphisms.

**Lemma 1.1.** *Let  $\mathcal{A} = \langle A, \star \rangle$ , be an algebra with a binary operations  $\star$ . An equivalence  $\sim$  on  $A$  is a congruence iff there exists an algebra  $\mathcal{B} = \langle B, * \rangle$  with a binary operation  $*$ , and there exists  $\theta : A \rightarrow B$  a homomorphism such that  $\sim$  coincides with the kernel congruence  $\ker(\theta)$  of  $\theta$  defined by  $x \sim_\theta y$  iff  $\theta(x) = \theta(y)$ .*

---

\* Corresponding author.

Let  $\Sigma$  be an alphabet not containing  $\{0, 1\}$ . We shall represent the algebra of binary trees over  $\Sigma$ , i.e., trees with leaves labeled by letters of  $\Sigma$ , as a set of words  $\mathcal{T}(\Sigma)$  on the alphabet  $\Sigma \cup \{0, 1\}$ , together with the binary product operation  $\star$ .

**Definition 1.2.** The algebra  $\mathcal{B} = \langle \mathcal{T}(\Sigma), \star \rangle$  of binary trees over  $\Sigma$  is defined as follows.

- A binary tree over  $\Sigma$  is a set of words  $t \subseteq \{0, 1\}^* \Sigma$  such that : For any  $ua, vb \in t$  if  $ua \neq vb$  then  $u$  is not a prefix of  $v$  and  $v$  is not a prefix of  $u$ . The carrier set  $\mathcal{T}(\Sigma)$  is the set of all binary trees. The empty set  $\emptyset$  is a binary tree denoted by  $\mathbf{0}$ .
- The binary product operation  $\star$  is defined by: for  $t, t' \in \mathcal{T}(\Sigma)$ ,  $t \star t' = 0.t \cup 1.t'$ . In particular,  $\mathbf{0} \star \mathbf{0} = \mathbf{0}$ .

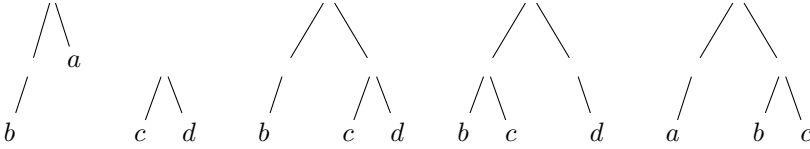


FIGURE 1. From left to right,  $t = \{00b, 1a\}$ ,  $\tau = \{0c, 1d\}$ ,  $\gamma_{a \rightarrow \tau}(t)$ ,  $t_1 = \{00b, 01c, 11d\}$ ,  $t_2 = \{00a, 10b, 11c\}$ . Trees  $\gamma_{a \rightarrow \tau}(t)$ ,  $t_1$ ,  $t_2$  have the same size 6, trees  $t_2$  and  $\gamma_{a \rightarrow \tau}(t)$  are similar (have the same skeleton)

When the alphabet  $\Sigma$  is clear, we will denote by  $\mathcal{T}$  the set of all binary trees. Trees are generated by  $\{\mathbf{0}\} \cup \Sigma$  and the operation  $\star$ .

An essential property of this algebra  $\mathcal{B}$  is that its elements are uniquely decomposable.

**Lemma 1.3 (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  $\mathcal{T}^2$  such that  $t = t_1 \star t_2$ .*

This property allows us to associate with each  $t \in \mathcal{T}$  its *size*  $|t|$  (number of nodes)

- $|\mathbf{0}| = 0$ , and for all  $a \in \Sigma$ ,  $|a| = 1$ ,
- if  $t \notin \{\mathbf{0}\} \cup \Sigma$  then  $t = t_1 \star t_2$ , and  $|t| = |t_1| + |t_2| + 1$ .

If  $|t| > 1$  then there exist  $t_1, t_2$  with  $|t_i| < |t|$  such that  $t = t_1 \star t_2$ . Trees  $t \star t'$ ,  $\mathbf{0} \star t'$ ,  $t \star \mathbf{0}$  are trees whose root has two sons, a single right son, a single left son, respectively. See Figure 1.

## 1.2. Homomorphisms, graftings

**Lemma 1.4.** *Let  $\mathcal{B} = \langle B, \star \rangle$  be an algebra with a binary operation  $\star$ . Every mapping  $h: \Sigma \rightarrow B$  can be uniquely extended to a homomorphism  $h: \mathcal{T} \rightarrow B$ .*

**Remark 1.5.** 1) Because of the universal property of Lemma 1.4, homomorphisms are (uniquely) defined by giving their values on  $\Sigma$ .

2) For every endomorphism,  $h(\mathbf{0}) = \mathbf{0}$ . Otherwise, as  $\mathbf{0} = \mathbf{0} \star \mathbf{0}$ ,  $h(\mathbf{0}) = h(\mathbf{0}) \star h(\mathbf{0})$ , if  $h(\mathbf{0}) = t$  with  $|t| \geq 1$  then  $t = t \star t$  implies  $|t| = 2|t| + 1$ , a contradiction.

**Definition 1.6.** For a given  $a \in \Sigma$ , let  $\nu_a$  be the endomorphism sending  $\Sigma$  onto  $a$ . If  $\nu_a(t) = \nu_a(t')$ , trees  $t$  and  $t'$  are said to be *similar* which is denoted by  $t \sim_s t'$ .

The congruence  $\sim_s$  does not depend on the choice of the letter  $a \in \Sigma$ . From an intuitive viewpoint,  $t \sim_s t'$  means that  $t$  and  $t'$  have the same skeleton, i.e., they are identical except for the leaf labels. See Figure 1. Other congruences fundamental for our proof are the kernels of the grafting endomorphisms, defined below.

**Definition 1.7 (Grafting).** Let  $a \in \Sigma$  and  $\tau \in \mathcal{T}$ . Then the grafting  $\gamma_{a \rightarrow \tau}: \mathcal{T} \rightarrow \mathcal{T}$  is the endomorphism defined by its restriction on  $\Sigma$

$$\gamma_{a \rightarrow \tau}(b) = \begin{cases} \tau & \text{if } b = a, \\ b & \text{if } b \neq a. \end{cases}$$

For any  $a \in \Sigma$ , and any  $\tau \in \mathcal{T}$ ,  $\gamma_{a \rightarrow \tau}$  is the endomorphism sending the letter  $a$  on  $\tau$  and each other letter on itself.

An endomorphism  $h$  of  $\langle \mathcal{T}(\Sigma), \star \rangle$  is *idempotent* if for every  $t \in \mathcal{T}$ ,  $h(h(t)) = h(t)$ . By Lemma 1.4  $h$  is idempotent iff for every  $a \in \Sigma$ ,  $h(h(a)) = h(a)$ . For instance if  $a$  does not occur in  $\tau$  then  $\gamma_{a \rightarrow \tau}$  is idempotent.

**Proposition 1.8.** Let  $\tau \in \mathcal{T}$  be of size  $|\tau| \neq 1$ , let  $t, t' \in \mathcal{T}$ , and let  $a_1 \neq a_2$  be two letters in  $\Sigma$ . If  $\gamma_{a_i \rightarrow \tau}(t) = \gamma_{a_i \rightarrow \tau}(t')$  for  $i = 1, 2$ , then  $t = t'$ .

*Proof.* By induction on  $\min(|t|, |t'|)$ .

*Basis Case 0:* If  $\min(|t|, |t'|) = 0$  then one of  $t, t'$  is  $\mathbf{0}$ , say  $t = \mathbf{0}$ . If  $t' \neq \mathbf{0}$  then  $t'$  contains at least one occurrence of some letter  $b$ . As  $\gamma_{a_i \rightarrow \tau}(t') = \gamma_{a_i \rightarrow \tau}(t) = \gamma_{a_i \rightarrow \tau}(\mathbf{0}) = \mathbf{0}$  we have  $\gamma_{a_i \rightarrow \tau}(t') = \mathbf{0}$ , which implies (because  $t' \neq \mathbf{0}$  was supposed) that  $\tau = \mathbf{0}$ . Then  $\gamma_{a_i \rightarrow \tau}(t') = \mathbf{0}$  implies that all leaves of  $t'$  are equal to both  $a_1$  and  $a_2$ , a contradiction. Hence  $t' = \mathbf{0}$  and  $t = t'$ .

*Basis Case 1:* If  $\min(|t|, |t'|) = 1$  then one of  $t, t'$  is a letter, say  $t = b$ , and there is one  $i$ , say  $i = 1$ , such that  $a_1 \neq b$ , thus  $b = \gamma_{a_1 \rightarrow \tau}(t) = \gamma_{a_1 \rightarrow \tau}(t')$ .

- If  $t'$  is a letter  $c \neq b$ , then  $\gamma_{a_1 \rightarrow \tau}(c) = b$ . If  $c = a_1$  then  $\gamma_{a_1 \rightarrow \tau}(c) = \tau$  whose size is not 1, a contradiction. Hence  $c \neq a_1$  and  $\gamma_{a_1 \rightarrow \tau}(c) = c \neq b$ , a contradiction. Hence  $t' = t = b$ .
- If  $|t'| > 1$  then  $t' = t'_1 \star t'_2$ , and  $\gamma_{a_1 \rightarrow \tau}(t') = \gamma_{a_1 \rightarrow \tau}(t'_1) \star \gamma_{a_1 \rightarrow \tau}(t'_2)$  which can be only of size 0 or  $\geq 2$ , contradicting  $\gamma_{a_1 \rightarrow \tau}(t') = b$ . this case is excluded.

*Induction:* If  $\min(|t|, |t'|) > 1$  then  $t = t_1 \star t_2$  and  $t' = t'_1 \star t'_2$  with  $\min(|t_i|, |t'_i|) < \min(|t|, |t'|)$ , for  $i = 1, 2$ . By Lemma 1.3,  $\gamma_{a_j \rightarrow \tau}(t_1) \star \gamma_{a_j \rightarrow \tau}(t_2) = \gamma_{a_j \rightarrow \tau}(t'_1) \star \gamma_{a_j \rightarrow \tau}(t'_2)$  implies  $\gamma_{a_j \rightarrow \tau}(t_i) = \gamma_{a_j \rightarrow \tau}(t'_i)$ , for  $j = 1, 2$ . By the induction hypothesis  $t_i = t'_i$ , hence  $t = t'$ .  $\square$

**Proposition 1.9.** *Let  $a \in \Sigma$ , with  $|\Sigma| \geq 3$ ,  $t, t' \in \mathcal{T}$  such that  $t \sim_s t'$ .*

(1) *If, for some  $\tau \in \mathcal{T}$  of size  $|\tau| \neq 1$ ,  $\gamma_{a \rightarrow \tau}(t) = \gamma_{a \rightarrow \tau}(t')$ , then  $t = t'$ .*

(2) *If, for some  $a \in \Sigma$  and for all  $b \neq a$ ,  $b \in \Sigma$ ,  $\gamma_{a \rightarrow b}(t) = \gamma_{a \rightarrow b}(t')$ , then  $t = t'$ .*

*Proof.* Both (1) and (2) are proved by induction on  $|t| = |t'|$ , and in both cases, the result obviously holds if  $t = t' = \mathbf{0}$ .

*Basis:* If  $|t| = |t'| = 1$ .

(1) We assume that  $t = b \neq c = t'$ .

(i) If  $a \notin \{b, c\}$  then  $\gamma_{a \rightarrow \tau}(t) = b \neq c = \gamma_{a \rightarrow \tau}(t')$ , a contradiction.

(ii) Otherwise,  $a \in \{b, c\}$ , e.g.,  $a = b = t$ , then  $\gamma_{a \rightarrow \tau}(t) = \gamma_{a \rightarrow \tau}(a) = \tau$  and  $\gamma_{a \rightarrow \tau}(t') = \gamma_{a \rightarrow \tau}(c) = c$ , hence  $\tau = c$ , which contradicts  $|\tau| \neq 1$ .

(2) We assume that  $t = b \neq c = t'$ .

(i) The case  $a \notin \{b, c\}$  yields a contradiction as in case (1).

(ii) Otherwise, e.g.,  $a = b$ , there exists  $d \notin \{a, c\}$ , and we get  $\gamma_{a \rightarrow d}(t) = \gamma_{a \rightarrow d}(a) = d$  and  $\gamma_{a \rightarrow d}(t') = \gamma_{a \rightarrow d}(c) = c$ , contradicting  $\gamma_{a \rightarrow d}(t) = \gamma_{a \rightarrow d}(t')$ .

*Induction:* As in Proposition 1.8 in both cases: since  $t$  and  $t'$  are similar,  $t = t_1 \star t_2$  and  $t' = t'_1 \star t'_2$  with  $t_i$  similar to  $t'_i$  and  $|t_i| < |t'_i|$ .  $\square$

### 1.3. Congruence preserving functions on trees

**Definition 1.10.** A function  $f: \mathcal{T}^n \rightarrow \mathcal{T}$  is *congruence preserving* (abbreviated into CP) if for all congruences  $\sim$  on  $\mathcal{T}$ , for all  $t_1, \dots, t_n, t'_1, \dots, t'_n$  in  $\mathcal{T}$ ,  $t_i \sim t'_i$  for all  $i = 1, \dots, n$ , implies  $f(t_1, \dots, t_n) \sim f(t'_1, \dots, t'_n)$ .

**Remark 1.11.** (1) It follows from Lemma 1.1 that CP functions are characterized by the fact that for all homomorphisms  $h$ ,  $h(t_i) = h(t'_i)$  for all  $i = 1, \dots, n$ , implies  $h(f(t_1, \dots, t_n)) = h(f(t'_1, \dots, t'_n))$ .

(2) If  $f$  is CP and endomorphism  $h$  is idempotent then  $h(f(t_1, \dots, t_n)) = h(f(h(t_1), \dots, h(t_n)))$ . Indeed, let  $\sim_h$  be the congruence associated with  $h$ , for  $i = 1, \dots, n$ , we have  $h(t_i) = h(h(t_i))$ , hence  $t_i \sim_h h(t_i)$  whence the result.

We will show that congruence preserving functions on the algebra  $\langle \mathcal{T}(\Sigma), \star \rangle$  are polynomial. Let us first formally define polynomials on trees.

**Definition 1.12.** Let  $x_1, \dots, x_n \notin \Sigma$  be called *variables*. A *polynomial*  $P(x_1, \dots, x_n)$  is a tree on the alphabet  $\Sigma \cup \{x_1, \dots, x_n\}$ .

With every polynomial  $P(x_1, \dots, x_n)$  we will associate a *polynomial function*  $\tilde{P}: \mathcal{T}^n \rightarrow \mathcal{T}$  defined by: for any  $\vec{u} = \langle t_1, \dots, t_i, \dots, t_n \rangle \in \mathcal{T}^n$ ,

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

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

**Theorem 1.13.** *Let  $|\Sigma| \geq 3$ . If  $g: \mathcal{T}^n \rightarrow \mathcal{T}$  is CP then there exists a polynomial  $P_g$  such that  $g = \widetilde{P_g}$ .*

## 2. Equality of CP functions

**Notation 2.1.** For any  $f : \mathcal{T}^n \rightarrow \mathcal{T}$ , we denote by  $f|_{\Sigma^n}$  its restriction to  $\Sigma^n$ .

In this section we prove that if  $f$  and  $g$  are two CP functions, then  $f|_{\Sigma^n} = g|_{\Sigma^n}$  implies  $f = g$ , provided that  $\Sigma$  contains at least three letters.

**Lemma 2.2.** *Suppose  $\Sigma$  has at least three letters. If  $f$  and  $g$  are CP functions on  $\mathcal{T}$  such that for all  $a \in \Sigma$ ,  $f(a) = g(a)$  then for all  $t \in \mathcal{T}$ ,  $f(t)$  and  $g(t)$  are similar.*

*Proof.* We have to show that  $\nu_a(f(t)) = \nu_a(g(t))$  for all  $t$ . As  $\nu_a$  is idempotent and  $f$  is CP, by Remark 1.11 (2),  $\nu_a(f(t)) = \nu_a(f(\nu_a(t)))$ , and similarly for  $g$ . Hence it suffices to prove  $f(\nu_a(t)) = g(\nu_a(t))$ . Let  $b_1, b_2 \in \Sigma$  such that  $a, b_1, b_2$  are pairwise distinct. As  $\gamma_{b_i \rightarrow \nu_a(t)}$  is idempotent, by Remark 1.11 (2), we have  $\gamma_{b_i \rightarrow \nu_a(t)}(f(b_i)) = \gamma_{b_i \rightarrow \nu_a(t)}(f(\nu_a(t)))$ . The same holds for  $g$ , i.e.,  $\gamma_{b_i \rightarrow \nu_a(t)}(g(b_i)) = \gamma_{b_i \rightarrow \nu_a(t)}(g(\nu_a(t)))$ . From  $f(b_i) = g(b_i)$ , we deduce that  $\gamma_{b_i \rightarrow \nu_a(t)}(f(\nu_a(t))) = \gamma_{b_i \rightarrow \nu_a(t)}(g(\nu_a(t)))$ . This equality holds for  $i = 1, 2$ , thus Proposition 1.8 implies  $f(\nu_a(t)) = g(\nu_a(t))$ .  $\square$

The following proposition shows that  $f$  is completely determined by its values on  $\Sigma$ .

**Proposition 2.3.** *Suppose  $\Sigma$  has at least three letters. If  $f$  and  $g$  are CP functions on  $\mathcal{T}$  such that for all  $a \in \Sigma$ ,  $f(a) = g(a)$  then for all  $t \in \mathcal{T}$ ,  $f(t) = g(t)$ .*

*Proof.* Let  $a$  be a letter that occurs in  $t$ . For any other letter  $b$ , the endomorphisms  $\gamma_{a \rightarrow b}$  and  $\gamma_{a \rightarrow t_b}$  are idempotent, where  $t_b = \gamma_{a \rightarrow b}(t)$ . Thus by Remark 1.11 (2),  $\gamma_{a \rightarrow t_b}(f(a)) = \gamma_{a \rightarrow t_b}(f(t_b))$ , and  $\gamma_{a \rightarrow t_b}(g(a)) = \gamma_{a \rightarrow t_b}(g(t_b))$ . As  $f(a) = g(a)$  we have  $\gamma_{a \rightarrow t_b}(f(t_b)) = \gamma_{a \rightarrow t_b}(g(t_b))$ . By Lemma 2.2,  $f(t_b)$  and  $g(t_b)$  are similar, and by Proposition 1.9 (1)  $f(t_b) = g(t_b)$ .

On the other hand, as  $f$  and  $g$  are CP and  $t \sim_{\gamma_{a \rightarrow b}} t_b$ , we get  $\gamma_{a \rightarrow b}(f(t)) = \gamma_{a \rightarrow b}(f(t_b))$  and  $\gamma_{a \rightarrow b}(g(t)) = \gamma_{a \rightarrow b}(g(t_b))$ , hence  $\gamma_{a \rightarrow b}(f(t)) = \gamma_{a \rightarrow b}(g(t))$ . As this is true for all  $b \neq a$ , we have by Proposition 1.9 (2),  $f(t) = g(t)$ .  $\square$

Proposition 2.3 now can be generalized.

**Notation 2.4.** For any function  $f : \mathcal{T}^{n+1} \rightarrow \mathcal{T}$ , any  $t \in \mathcal{T}$ , and  $\vec{u} = \langle t_1, \dots, t_n \rangle$ , we define

- (1) the  $n$ -ary function  $f_{\dots, t}$  by  $f_{\dots, t}(\vec{u}) = f_{\dots, t}(\vec{u}) = f(\vec{u}, t)$ .
- (2) the unary function  $f_{\vec{u}, \cdot}$  by  $f_{\vec{u}, \cdot}(t) = f(\vec{u}, t)$ .

**Proposition 2.5.** *Let  $f$  and  $g$  be  $n$ -ary CP functions on  $\mathcal{T}$  such that for all  $a_1, \dots, a_n \in \Sigma$ ,  $f(a_1, \dots, a_n) = g(a_1, \dots, a_n)$  then for all  $t_1, \dots, t_n \in \mathcal{T}$ ,  $f(t_1, \dots, t_n) = g(t_1, \dots, t_n)$ .*

*Proof.* Induction on  $n$ . For  $n = 1$  the result was proved in Proposition 2.3. Assume by induction that the result holds for all  $k \leq n$ . By the hypothesis, for all  $a_1, \dots, a_n, a \in \Sigma$ , we have  $f(a_1, \dots, a_n, a) = g(a_1, \dots, a_n, a)$ , i.e.,  $f_{\dots, a}(a_1, \dots, a_n) = g_{\dots, a}(a_1, \dots, a_n)$ . By the induction applied to  $f_{\dots, a}$ , for

all  $\vec{u} \in \mathcal{T}^n$ ,  $f_{\dots, a}(\vec{u}) = g_{\dots, a}(\vec{u})$ , or equivalently  $f_{\vec{u}, \cdot}(a) = g_{\vec{u}, \cdot}(a)$ . Let  $\vec{u} \in \Sigma^n$ . As  $f_{\vec{u}, \cdot}(a) = g_{\vec{u}, \cdot}(a)$ , applying now the induction hypothesis to  $f_{\vec{u}, \cdot}$  yields  $f_{\vec{u}, \cdot}(t) = g_{\vec{u}, \cdot}(t)$  for all  $t$ , hence  $f(\vec{u}, t) = g(\vec{u}, t)$ .  $\square$

### 3. The algebra of binary trees is affine complete

To prove that any CP function is polynomial, as a consequence of Proposition 2.5 and of the fact that a polynomial function is CP, it is enough to show that the restriction  $f|_{\Sigma^n}$  of  $f: \mathcal{T}^n \rightarrow \mathcal{T}$  to  $\Sigma^n$  is equal to the restriction  $\tilde{P}|_{\Sigma^n}$  of a  $n$ -ary polynomial function. For such restricted functions we introduce a weakened version WCP of the CP condition, namely,

**Definition 3.1.** Function  $g: \mathcal{T}^n \rightarrow \mathcal{T}$  is said to be WCP iff for any mapping  $h: \Sigma \rightarrow \Sigma$ ,  $\forall \vec{u}, \vec{v} \in \Sigma^n$ ,  $h(\vec{u}) = h(\vec{v}) \implies h(g(\vec{u})) = h(g(\vec{v}))$ , where for  $\vec{u} = \langle u_1, \dots, u_n \rangle$ ,  $h(\vec{u})$  denotes  $\langle h(u_1), \dots, h(u_n) \rangle$ .

Every CP function is clearly WCP.

**Lemma 3.2.** *If  $g$  is WCP then for all  $\vec{u}, \vec{v} \in \Sigma^n$ ,  $g(\vec{u})$  and  $g(\vec{v})$  are similar.*

*Proof.* As  $\nu_a(\vec{u}) = \nu_a(\vec{v}) = \langle a, \dots, a \rangle$  for all  $\vec{u}, \vec{v} \in \Sigma^n$ ,  $\nu_a(g(\vec{u})) = \nu_a(g(\vec{v})) = \nu_a(g(a, \dots, a))$ .  $\square$

We often use a different form of the condition WCP, which deals only with alphabetic graftings.

**Proposition 3.3.** *A fonction  $g$  is WCP if and only if*

*(GCP) ( $G$  for graftings) for all  $a_1, a_2, \dots, a_n \in \Sigma$ ,  $i \in \{1, \dots, n\}$  and  $b_i \in \Sigma$ ,  $\gamma_{a_i \rightarrow b_i}(g(a_1, \dots, a_n)) = \gamma_{a_i \rightarrow b_i}(g(a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n))$ .*

*Proof.* Since  $\gamma_{a_i \rightarrow b_i}(a_1, \dots, a_n) = \gamma_{a_i \rightarrow b_i}(a_1, \dots, a_{i-1}, b_i, a_{i+1}, \dots, a_n)$ , clearly WCP implies GCP. The proof of the converse is by induction on  $n$ . It is obviously true for  $n = 0$ .

Otherwise, let  $\vec{u}, \vec{v} \in \Sigma^n$  such that  $h(\vec{u}) = h(\vec{v})$ , and let  $a, b \in \Sigma$  such that  $h(a) = h(b)$ . By (GCP), we have  $\gamma_{a \rightarrow b}(g(\vec{u}, a)) = \gamma_{a \rightarrow b}(g(\vec{u}, b))$ , hence  $h(\gamma_{a \rightarrow b}(g(\vec{u}, a))) = h(\gamma_{a \rightarrow b}(g(\vec{u}, b)))$ .

But  $h(\gamma_{a \rightarrow b}(c)) = \begin{cases} h(c) & \text{if } c \neq a \\ h(b) = h(a) & \text{if } c = a \end{cases}$  hence  $h \circ \gamma_{a \rightarrow b} = h$ . Therefore  $h(g(\vec{u}, a)) = h(g(\vec{u}, b))$ , and by the induction applied to  $g_{\dots, b}$ ,  $h(g(\vec{u}, a)) = h(g(\vec{u}, b)) = h(g(\vec{v}, b))$ .  $\square$

Let us first study unary WCP functions whose restriction to  $\Sigma$  is contained in  $\Sigma$ .

**Proposition 3.4.** *Assume  $|\Sigma| \geq 3$ . Let  $f: \mathcal{T} \rightarrow \mathcal{T}$  be WCP and such that  $f(\Sigma) \subseteq \Sigma$ . Then  $f|_{\Sigma^n}$  is (1) either a constant function (2) or the identity.*

*Proof.* If  $f$  is not the identity there exist  $a, b$ , with  $a \neq b$  and  $f(a) = b$ . As  $\gamma_{a \rightarrow b}(f(b)) = \gamma_{a \rightarrow b}(f(a)) = \gamma_{a \rightarrow b}(a) = b$ , we get  $f(b) \in \{a, b\}$ .

For  $c \notin \{a, b\}$ ,  $\gamma_{a \rightarrow c}(f(c)) = \gamma_{a \rightarrow c}(f(a)) = b$  implies  $f(c) = b$ . It remains to prove that  $f(b) = b$ . From  $\gamma_{b \rightarrow c}(f(b)) = \gamma_{b \rightarrow c}(f(c)) = b$ , we deduce that

$f(b) \in \{c, b\}$ , hence  $f(b) \in \{a, b\} \cap \{c, b\} = b$  which concludes the proof for the case  $t \in \Sigma$ .  $\square$

We now will generalize Proposition 3.4 by Proposition 3.5 (replacing a unary  $f$  by a  $n$ -ary  $g$ ).

**Proposition 3.5.** *Assume  $|\Sigma| \geq 3$ . If  $g: \mathcal{T}^n \rightarrow \mathcal{T}$  is WCP and such that  $g(\Sigma^n) \subseteq \Sigma$ , then  $g|_{\Sigma^n}$  is (1) either a constant function (2) or a projection  $\pi_i^n$ .*

*Proof.* The proof is by induction on  $n$ . It is true for  $n = 1$ . If  $g$  is of arity  $n + 1$  then, by induction hypothesis, for any  $a \in \Sigma$ , the function  $g_{\dots, a}$  of arity  $n$  is either a constant or a projection  $\pi_i^n$ . We first show that these functions are all equal to a given  $\pi_i^n$  or all equal to a same constant.

Let us assume that  $g_{\dots, a} = \pi_i^n$ . Let  $\vec{u} = \langle a, \dots, a, c, a, \dots, a \rangle$  and  $\vec{v} = \langle a, \dots, a, d, a, \dots, a \rangle$  with  $a, c, d$  pairwise disjoint, so that for any  $b$ ,  $\gamma_{a \rightarrow b}(g(\vec{u}, a)) = c$  and  $\gamma_{a \rightarrow b}(g(\vec{v}, a)) = d$ . It follows from the GCP condition that  $\gamma_{a \rightarrow b}(g(\vec{u}, a)) = \gamma_{a \rightarrow b}(g(\vec{u}, b)) = c$  and  $\gamma_{a \rightarrow b}(g(\vec{v}, a)) = \gamma_{a \rightarrow b}(g(\vec{v}, b)) = d$ , which is impossible if  $g_{\dots, b}$  is either a constant or a projection  $\pi_j^n$  with  $j \neq i$ . Hence all  $g_{\dots, a}$  are equal to  $\pi_i^n$  implying  $g = \pi_i^{n+1}$ .

Assume now all the  $g_{\dots, a}$  are constant. For every  $\vec{u}, \vec{v}, a$ , we have  $g(\vec{u}, a) = g(\vec{v}, a)$ . We choose an arbitrary  $\vec{u} \in \Sigma^n$  which will be fixed. By the induction hypothesis  $g_{\vec{u}, \cdot}$  is either (1) the identity, or (2) a constant. In case (1) for all  $\vec{v}, a$ ,  $g(\vec{u}, a) = g(\vec{v}, a) = a$  and  $g = \pi_{n+1}^{n+1}$ . In case (2) for all  $\vec{v}, a, b$ ,  $g(\vec{u}, a) = g(\vec{v}, b) = a$  and  $g$  is a constant.  $\square$

As CP functions are WCP, for  $g$  a CP function such that for some  $a_1, \dots, a_n \in \Sigma$ ,  $g(a_1, \dots, a_n) \in \Sigma$ , we have shown that there exists a polynomial  $\widetilde{P}_g$ , which is either a constant  $a \in \Sigma$  or an  $x_i$ , such that  $g = \widetilde{P}_g$ . We will now extend to the case when  $g(a_1, \dots, a_n) \notin \Sigma$ .

**Proposition 3.6.** *Assume that  $|\Sigma| \geq 3$ . Let  $g: \mathcal{T}^n \rightarrow \mathcal{T}$  be WCP. Then there exists a polynomial  $P_g$  such that  $g|_{\Sigma^n} = \widetilde{P}_g|_{\Sigma^n}$ .*

*Proof.* Let  $\sigma(g)$  be the common size of all the  $g(\vec{u})$ ,  $\vec{u} \in \Sigma^n$ . The proof is by induction on  $\sigma(g)$ .

*Basis:* If  $\sigma(g) = 0$  then  $g|_{\Sigma^n} = \tilde{P}|_{\Sigma^n} = \mathbf{0}$ . If  $\sigma(g) = 1$  then  $g(a_1, \dots, a_n) \in \Sigma$  and the result is proved in Proposition 3.5.

*Induction:* If  $\sigma(g) > 1$  there exists two functions  $g_i: \mathcal{T}^n \rightarrow \mathcal{T}$  for  $i = 1, 2$ , such that for all  $\vec{u} \in \Sigma^n$ ,  $g(\vec{u}) = g_1(\vec{u}) \star g_2(\vec{u})$ , with  $|\sigma(g_i)| < |\sigma(g)|$ . It remains to show that the  $g_1$  and  $g_2$  are WCP. Let  $\vec{u}, \vec{v} \in \Sigma^n$  such that  $h(\vec{u}) = h(\vec{v})$  for some mapping  $h: \Sigma \rightarrow \Sigma$ . Extend  $h$  as an endomorphism  $\mathcal{T} \rightarrow \mathcal{T}$  by Lemma 1.4, then  $h(g(\vec{u})) = h(g_1(\vec{u}) \star g_2(\vec{u})) = h(g_1(\vec{u})) \star h(g_2(\vec{u}))$ . Similarly,  $h(g(\vec{v})) = h(g_1(\vec{v})) \star h(g_2(\vec{v}))$ . As  $g$  is WCP and  $h(\vec{u}) = h(\vec{v})$ , hence  $h(g(\vec{u})) = h(g(\vec{v}))$ . Then by Lemma 1.3 (unique decomposition) we get  $h(g_i(\vec{u})) = h(g_i(\vec{v}))$  for  $i = 1, 2$ . As this is true for any  $h$ ,  $g_1$  and  $g_2$  are

WCP. By the induction hypothesis there exists  $P_i$  such  $\tilde{P}_i|_{\Sigma^n} = g_i|_{\Sigma^n}$ , hence  $g|_{\Sigma^n} = \widetilde{P_1|_{\Sigma^n}} \star \widetilde{P_2|_{\Sigma^n}} = \widetilde{P_1 \star P_2|_{\Sigma^n}}$ .  $\square$

**Theorem 3.7.** *If  $f: \mathcal{T}^n \rightarrow \mathcal{T}$  is CP then there exists a polynomial  $P$  such that  $f = \tilde{P}$ .*

*Proof.* Since  $f$  is CP,  $f$  also is WCP. By the previous proposition, there exists  $P$  such that  $f|_{\Sigma^n} = \tilde{P}|_{\Sigma^n}$ , and by Proposition 2.5,  $f = \tilde{P}$ .  $\square$

## 4. Conclusion

We proved that, when  $\Sigma$  has at least three letters, the algebra of arbitrary binary trees with leaves labeled by letters of  $\Sigma$  is an affine complete algebra (non commutative and non associative). Extending this result when  $\Sigma$  has at most two letters yields open problems.

## 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)
- [2] Kaarli K., Pixley A.F.: *Polynomial Completeness in Algebraic Systems*. Chapman & Hall/CRC (2001)
- [3] Ploščica M., Havíar M.: Congruence-preserving functions on distributive lattices. *Algebra Universalis*, **59**, 179–196 (2008)

André Arnold

LABRI, UMR 5800, Université Bordeaux, France.

e-mail: [andre.arnold@club-internet.fr](mailto:andre.arnold@club-internet.fr)

Patrick Cégielski

LACL, EA 4219, Université Paris XII – IUT de Sénart-Fontainebleau, France.

e-mail: [cegielski@u-pec.fr](mailto:cegielski@u-pec.fr)

Serge Grigorieff

IRIF, UMR 8243, CNRS & Université Paris-Diderot, France.

e-mail: [seg@irif.fr](mailto:seg@irif.fr)

Irène Guessarian\*

IRIF, UMR 8243, CNRS & Sorbonne Université, France.

e-mail: [ig@irif.fr](mailto:ig@irif.fr)