The Mathematics behind the Property of Associativity

Jean-Luc Marichal[1] and Bruno Teheux[1]

FSTC / Mathematics Research Unit
University of Luxembourg, Luxembourg, Luxembourg
{jean-luc.marichal,bruno.teheux@}@uni.lu

Abstract. The well-known equation of associativity for binary operations may be naturally generalized to variadic operations. In this talk, we illustrate different approaches that can be considered to study this extension of associativity, as well as some of its generalizations and variants, including barycentric associativity and preassociativity.

1 Introduction

Let X and Y be arbitrary nonempty sets. We regard tuples $\mathbf{x}$ in X^n as n-strings over X. We let $X^* = \bigcup_{n\geqslant 0} X^n$ be the set of all strings over X, with the convention that $X^0 = \{\varepsilon\}$, where ε is called the *empty string*. We denote the elements of X^* by bold roman letters $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$. If we want to stress that such an element is a letter of X, we use non-bold italic letters x, y, z, etc. The *length* of a string $\mathbf{x}$ is denoted by $|\mathbf{x}|$. We endow the set X^* with the concatenation operation, for which ε is the neutral element, i.e., $\varepsilon\mathbf{x} = \mathbf{x}\varepsilon = \mathbf{x}$ (in other words, we consider X^* as the free monoid generated by X). Moreover, for every string $\mathbf{x}$ and every integer $n \geqslant 0$, the power $\mathbf{x}^n$ stands for the string obtained by concatenating n copies of $\mathbf{x}$. In particular we have $\mathbf{x}^0 = \varepsilon$.

Let Y be a nonempty set. Recall that, for every integer $n \geq 0$, a function $F\colon X^n \to Y$ is said to be n*-ary*. Similarly, a function $F\colon X^* \to Y$ is said to have an *indefinite arity* or to be *variadic*. A variadic function $F\colon X^* \to Y$ is said to be a *variadic operation on* X (or an *operation* for short) if $\operatorname{ran}(F) \subseteq X \cup \{\varepsilon\}$. It is *standard* if $F(\mathbf{x}) = F(\varepsilon)$ if and only if $\mathbf{x} = \varepsilon$, and ε*-standard* if $\varepsilon \in Y$ and we have $F(\mathbf{x}) = \varepsilon$ if and only if $\mathbf{x} = \varepsilon$.

The main functional properties for variadic functions that we present and investigate in this talk are given in the following definition (see [2, 4, 6]).

Definition 1. *A function* $F\colon X^* \to X^*$ *is said to be* associative *if, for every* $\mathbf{x},\mathbf{y},\mathbf{z} \in X^*$*, we have*
$$F(\mathbf{x}\mathbf{y}\mathbf{z}) = F(\mathbf{x}F(\mathbf{y})\mathbf{z}).$$
It is said to be barycentrically associative *(or* B-associative*) if, for every* $\mathbf{x},\mathbf{y},\mathbf{z} \in X^*$ *, we have*
$$F(\mathbf{x}\mathbf{y}\mathbf{z}) = F(\mathbf{x}F(\mathbf{y})^{|\mathbf{y}|}\mathbf{z}).$$

A variadic function $F\colon X^* \to Y$ *is said to be* preassociative *if, for every* $\mathbf{x},\mathbf{y},\mathbf{y}',\mathbf{z} \in X^*$ *, we have*
$$F(\mathbf{y}) = F(\mathbf{y}') \implies F(\mathbf{x}\mathbf{y}\mathbf{z}) = F(\mathbf{x}\mathbf{y}'\mathbf{z}).$$

It is said to be barycentrically preassociative *(or* B-preassociative*) if for every* $\mathbf{x}, \mathbf{y}, \mathbf{y}', \mathbf{z} \in X^*$ *, we have*

$$\left.\begin{array}{c} F(\mathbf{y}) = F(\mathbf{y}') \\ |\mathbf{y}| = |\mathbf{y}'| \end{array}\right\} \implies F(\mathbf{xyz}) = F(\mathbf{xy}'\mathbf{z}).$$

The following results show that preassociativity is a weaker form of associativity, and that B-preassociativity is a weaker form of B-associativity.

Proposition 1 ([2]). *A function* $F\colon X^* \to X^*$ *is associative if and only if it is preassociative and satisfies* $F \circ F = F$.

Proposition 2 ([6]). *A function* $F\colon X^* \to X^*$ *is B-associative if and only if it is B-preassociative and satisfies* $F(\mathbf{x}) = F(F(\mathbf{x})^{|\mathbf{x}|})$ *for all* $\mathbf{x} \in X^*$.

Throughout this note, we focus on the associativity and preassociativity properties, leaving the discussion on the properties of B-associative and B-preassociative functions to the oral presentation.

2 Factorization of preassociative functions

Recall that an equivalence relation θ on X^* is called a *congruence* if it satisfies

$$(\mathbf{x}_1 \theta \mathbf{y}_1 \quad \& \quad \mathbf{x}_2 \theta \mathbf{y}_2) \implies \mathbf{x}_1\mathbf{x}_2 \theta \mathbf{y}_1\mathbf{y}_2.$$

The definition of preassociativity and B-preassociativity can be restated as follows. As usual, for any function $F\colon Z \to Y$, we denote by $\ker(F)$ the equivalence relation defined by $\ker(F) = \{(x, y) \in Z^2 \mid F(x) = F(y)\}$.

Lemma 1. *A variadic function* $F\colon X^* \to Y$ *is preassociative if and only if* $\ker(F)$ *is a congruence on* X^*.

If $F\colon Z \to Y$ and if $g\colon Y \to Y'$ is an injective function, then $\ker(F) = \ker(g \circ F)$. Hence, we have the following easy corollary.

Corollary 1 ([3,4]). *Let* $F\colon X^* \to Y$ *be a variadic function. If* F *is preassociative and* $g\colon \mathrm{ran}(F) \to Y'$ *is constant or one-to-one, then* $g \circ F$ *is preassociative.*

In general, given a preassociative function $F\colon X^* \to Y$, characterizing the maps $g\colon Y \to Y'$ such that $g \circ F$ is preassociative is a difficult task since it amounts to characterizing the congruences above $\ker(F)$ on X^*.

It is easily seen that the only one-to-one associative function $F\colon X^* \to X^*$ is the identity. The next result shows that an associative function which is non-injective is in some sense highly non-injective.

Proposition 3 ([2]). *Let* $F\colon X^* \to X^*$ *be an associative function different from the identity. Then there is an infinite sequence of associative functions* $(F^m\colon X^* \to X^*)_{m \geq 1}$ *such that* $\ker(\mathrm{id}) \subsetneq \cdots \subsetneq \ker(F^2) \subsetneq \ker(F^1) \subseteq \ker(F)$.

By carefully choosing g in Corollary 1 we can give the following characterization of preassociative functions (see [2, 4, 6]).

Proposition 4 (Factorization of preassociative functions). *Let $F\colon X^* \to Y$ be a function. The following conditions are equivalent.*

(i) *F is preassociative.*
(ii) *There exists an associative function $H\colon X^* \to X^*$ and a one-to-one function $f\colon \mathrm{ran}(H) \to Y$ such that $F = f \circ H$.*

For any variadic function $F\colon X^* \to Y$ and any integer $n \geq 0$, we denote by F_n the *n-ary part* of F, *i.e.*, the restriction $F|_{X^n}$ of F to the set X^n. We also let $X^+ = X^* \backslash \{\varepsilon\}$ and denote the restriction $F|_{X^+}$ of F to X^+ by F^+.

Corollary 2. *Let $F\colon X^* \to Y$ be a standard function. The following conditions are equivalent.*

(i) *F is preassociative and satisfies $\mathrm{ran}(F_1) = \mathrm{ran}(F^+)$.*
(ii) *There exists an associatve ε-standard operation $H\colon X^* \to X \cup \{\varepsilon\}$ and a one-to-one function $f\colon \mathrm{ran}(H^+) \to Y$ such that $F^+ = f \circ H^+$.*

Corollary 2 enables us to produce axiomatizations of classes of preassociative functions from known axiomatizations of classes of associative functions. Let us illustrate this observation on an example. Further examples can be found in [5]. Let us recall an axiomatization of the Aczélian semigroups.

Proposition 5 ([1]). *Let $\mathbb{I}$ be a nontrivial real interval, possibly unbounded. An operation $H\colon \mathbb{I}^2 \to \mathbb{I}$ is continuous, one-to-one in each argument, and associative if and only if there exists a continuous and strictly monotone function $\phi\colon \mathbb{I} \to \mathbb{J}$ such that*

$$H(x,y) = \phi^{-1}(\phi(x) + \phi(y)),$$

where $\mathbb{J}$ is a real interval of the form $]-\infty,b[$, $]-\infty,b]$, $]a,+\infty[$, $[a,+\infty[$ or $\mathbb{R} =]-\infty,+\infty[$ ($b \leq 0 \leq a$). For such an operation H, the interval I is necessarily open at least on one end. Moreover, ϕ can be chosen to be strictly increasing.

Corollary 2 leads to the following characterization result.

Theorem 1 ([5]). *Let $\mathbb{I}$ be a nontrivial real interval, possibly unbounded. A standard function $F\colon \mathbb{I}^* \to \mathbb{R}$ is preassociative and satisfies $\mathrm{ran}(F^+) = \mathrm{ran}(F_1)$, and F_1 and F_2 are continuous and one-to-one in each argument if and only if there exist continuous and strictly monotone functions $\phi\colon \mathbb{I} \to \mathbb{J}$ and $\psi\colon \mathbb{J} \to \mathbb{R}$ such that*

$$F_n(\mathbf{x}) = \psi\left(\sum_{i=1}^{n} \phi(x_i)\right), \qquad n \geq 1,$$

where $\mathbb{J}$ is a real interval of one of the forms $]-\infty,b[$, $]-\infty,b]$, $]a,+\infty$, $[a,+\infty[$ or $\mathbb{R} =]-\infty,\infty[$ ($b \leq 0 \leq a$). For such a function F, we have $\psi : F_1 \circ \phi^{-1}$ and $\mathbb{I}$ is necessarily open at least on one end. Moreover, ϕ can be chosen to be strictly increasing.

3 Preassociativity and transition systems

As illustrated below, preassociative functions $F\colon X^* \to Y$ can be characterized as functions that can be computed through some special transition systems. In this note, a *transition system* is a triple $\mathcal{A} = (Q, q_0, \delta)$ where Q is a set of *states*, $q_0 \in Q$ is an *initial state* and $\delta : Q \times X \to Q$ is a transition map. As usual, the map δ is extended to $Q \times X^*$ by setting for any $q \in Q$,

$$\delta(q, \varepsilon) = q$$

$$\delta(q, \mathbf{x}y) = \delta(\delta(q, \mathbf{x}), y), \quad y \in X, \mathbf{x} \in X^*.$$

We say that a function $F\colon X^* \to Y$ is *right-preassociative* if it satisfies $F(\mathbf{x}) = F(\mathbf{y}) \implies F(\mathbf{xz}) = F(\mathbf{yz})$ for every $\mathbf{xyz} \in X^*$.

Definition 2. *Assume that $F : X^* \to Y$ is an onto right-preassociative function. We define the transitions system $\mathcal{A}(F) = (Y, q_0, \delta)$ by*

$$q_0 = F(\varepsilon) \quad \textit{and} \quad \delta(F(\mathbf{x}), z) = F(\mathbf{x}z).$$

We call $\mathcal{A}(F)$ the transitions system associated *with F.*

The language of transition systems give an elegant way to characterize preassociativity. Indeed, transition systems that arise from preassociative functions can be characterized in the following way. For any transition system $\mathcal{A} = (Q, q_0, \delta)$ and any $q \in Q$, let $L^{\mathcal{A}}(q) = \{\mathbf{x} \in X^* \mid \delta(q_0, \mathbf{x}) = q\}$ and $\mathcal{L}^{\mathcal{A}} = \{L^{\mathcal{A}}(q) \mid q \in Q\}$.

Theorem 2. *Let $\mathcal{A} = (Q, q_0, \delta)$ be a transition system. The following conditions are equivalent.*

(i) There is a preassociative function $F\colon X^* \to Q$ such that $\mathcal{A} = \mathcal{A}(F)$.
(ii) For every $z \in X$, the map defined on $\mathcal{L}^{\mathcal{A}}$ by $L \mapsto zL = \{z\mathbf{x} \mid \mathbf{x} \in L\}$ is valued in $\{2^L \mid L \in \mathcal{L}^{\mathcal{A}}\}$.

Acknowledgement. This research is supported by the internal research project F1R-MTH-PUL-15MRO3 of the University of Luxembourg.

References

1. J. Aczél. Sur les opérations définies pour nombres réels. *Bull. Soc. Math. France*, 76:59-64, 1949.
2. E. Lehtonen, J.-L. Marichal, B. Teheux. Associative string functions. *Asian-European Journal of Mathematics*, 7(4):1450059 (18 pages), 2014.
3. J.-L. Marichal and B. Teheux. Associative and preassociative functions. *Semigroup Forum*, 89(2):431–442, 2014.
4. J.-L. Marichal and B. Teheux. Associative and preassociative functions (improved version). Working paper (arXiv:1309.7303v3).
5. J.-L. Marichal and B. Teheux. Preassociative aggregation functions. *Fuzzy Sets and Systems*, 268:15–26, 2015.
6. J.-L. Marichal and B. Teheux. Barycentrically associative and preassociative functions. *Acta Mathematica Hungarica*, 145(2):468–488, 2015.