Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems 299 (2016) 89–97

FUZZY sets and systems

www.elsevier.com/locate/fss

Relaxations of associativity and preassociativity for variadic functions

Miguel Couceiro [a], Jean-Luc Marichal [b,*], Bruno Teheux [b]

[a] *LORIA, (CNRS – Inria Nancy Grand Est – Université de Lorraine), BP239, 54506 Vandoeuvre-lès-Nancy, France*
[b] *Mathematics Research Unit, FSTC, University of Luxembourg, 6, rue Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg*

Received 5 August 2015; accepted 4 November 2015

Available online 18 November 2015

Abstract

In this paper we consider two properties of variadic functions, namely associativity and preassociativity, that are pertaining to several data and language processing tasks. We propose parameterized relaxations of these properties and provide their descriptions in terms of factorization results. We also give an example where these parameterized notions give rise to natural hierarchies of functions and indicate their potential use in measuring the degrees of associativeness and preassociativeness. We illustrate these results by several examples and constructions and discuss some open problems that lead to further directions of research.

Keywords: Associativity; Preassociativity; Variadic function; String function; Functional equation; Axiomatization

1. Introduction

Let X be an arbitrary nonempty set, called the *alphabet*, and its elements are called *letters*. The symbol X^* stands for the set $\bigcup_{n\geqslant 0} X^n$ of all tuples on X, and its elements are called *strings*, where the empty string ε is such that $X^0=\{\varepsilon\}$. 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 $x, y, z, \ldots$. We assume that X^* is endowed with the concatenation operation (the empty string ε being the neutral element) for which we adopt the juxtaposition notation. For instance, if $\mathbf{x}\in X^m$ and $y\in X$, then $\mathbf{x}y\varepsilon=\mathbf{x}y\in X^{m+1}$. 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 $\mathbf{x}^0=\varepsilon$. The *length* of a string $\mathbf{x}$ is denoted by $|\mathbf{x}|$. In particular, we have $|\varepsilon|=0$.

Let Y be a nonempty set. Recall that a function $F\colon X^*\to Y$ is said to be *variadic* and that, for every integer $n\geqslant 0$, a function $F\colon X^n\to Y$ is said to be *n-ary*. A unary operation on X^* is a particular variadic function $F\colon X^*\to X^*$ called a *string function* over the alphabet X.

* Corresponding author.
E-mail addresses: miguel.couceiro@inria.fr (M. Couceiro), jean-luc.marichal@uni.lu (J.-L. Marichal), bruno.teheux@uni.lu (B. Teheux).

http://dx.doi.org/10.1016/j.fss.2015.11.003
0165-0114/

Definition 1.1. A function $F\colon X^* \to X^*$ is said to be *associative* [3] if for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^*$, we have

$$F(\mathbf{xyz}) = F(\mathbf{x}F(\mathbf{y})\mathbf{z}). \tag{1}$$

A function $F\colon X^* \to Y$ is said to be *preassociative* [5,6] if for any $\mathbf{x}, \mathbf{y}, \mathbf{y}', \mathbf{z} \in X^*$, we have

$$F(\mathbf{y}) = F(\mathbf{y}') \quad \Rightarrow \quad F(\mathbf{xyz}) = F(\mathbf{xy}'\mathbf{z}). \tag{2}$$

Associative string functions and preassociative variadic functions as well as some of their variants have been studied in [3–8]. For instance, it has been shown [3] that a function $F\colon X^* \to X^*$ is associative if and only if it is preassociative and satisfies the condition $F = F \circ F$. Also, under the Axiom of Choice, a function $F\colon X^* \to Y$ is preassociative if and only if it can be written as a composition of the form $F = f \circ H$, where $H\colon X^* \to X^*$ is associative and $f\colon \mathrm{ran}(H) \to Y$ is one-to-one.

It is noteworthy that several data processing tasks correspond to associative and preassociative functions. For instance, the function which corresponds to sorting the letters of every string in alphabetical order is associative. Similarly, the function that transforms a string of letters into upper case is also associative. Another natural example of a preassociative function is the mapping that outputs the length of strings.

In this paper we introduce and study certain relaxations of associativity and preassociativity. Let $\mathcal{A}$ denote the class of associative string functions on X^* and let $\mathcal{P}$ denote the class of preassociative variadic functions on X^*. For a fixed nonempty subset D of X^*, define the following classes of functions:

$$\mathcal{A}_D = \{F\colon X^* \to \mathrm{ran}(F) \mid F(D) \subseteq X^* \text{ and (1) holds for all } \mathbf{x}, \mathbf{y}, \mathbf{z} \in X^* \text{ such that } \mathbf{y} \in D\},$$

$$\mathcal{A}'_D = \{F\colon X^* \to \mathrm{ran}(F) \mid F(D) \subseteq X^* \text{ and (1) holds for all } \mathbf{x}, \mathbf{y}, \mathbf{z} \in X^* \text{ such that } F(\mathbf{y}) \in F(D)\},$$

$$\mathcal{P}_D = \{F\colon X^* \to \mathrm{ran}(F) \mid \text{(2) holds for all } \mathbf{x}, \mathbf{y}, \mathbf{y}', \mathbf{z} \in X^* \text{ such that } \mathbf{y}, \mathbf{y}' \in D\},$$

$$\mathcal{P}'_D = \{F\colon X^* \to \mathrm{ran}(F) \mid \text{(2) holds for all } \mathbf{x}, \mathbf{y}, \mathbf{y}', \mathbf{z} \in X^* \text{ such that } \mathbf{y} \in D\}.$$

It is clear that $\mathcal{A}_{X^*} = \mathcal{A}'_{X^*} = \mathcal{A}$ and $\mathcal{P}_{X^*} = \mathcal{P}'_{X^*} = \mathcal{P}$. When $D \subsetneq X^*$, these classes of functions correspond to relaxations of associativity and preassociativity for which we have $\mathcal{A}'_D \subseteq \mathcal{A}_D$ and $\mathcal{P}'_D \subseteq \mathcal{P}_D$. For instance, functions $F\colon X^* \to X^*$ that are in $\mathcal{A}'_D$ are characterized by the fact that for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^*$ the value $F(\mathbf{xyz})$ can be replaced with $F(\mathbf{x}F(\mathbf{y})\mathbf{z})$ whenever $F(\mathbf{y}) = F(\mathbf{y}')$ for some $\mathbf{y}' \in D$.

Certain of these relaxations are particularly natural. For instance, consider the subset

$$D = \{x^n \mid x \in X, \ n \in \mathbb{N}\},$$

where $\mathbb{N}$ denotes the set of nonnegative integers. Any function $F\colon X^* \to X^*$ in $\mathcal{A}_D$ has the property that the value $F(\mathbf{xyz})$ can be replaced with $F(\mathbf{x}F(\mathbf{y})\mathbf{z})$ whenever $\mathbf{y}$ is a repeated letter. Further examples include:

- $D = \{\mathbf{x} \in X^* \mid |\mathbf{x}| \leqslant m\}$ for some integer $m \geqslant 0$,
- $D = \{\mathbf{x} \in X^* \mid |\mathbf{x}| \geqslant m\}$ for some integer $m \geqslant 0$,
- $D = X^*\mathbf{w}X^* = \{\mathbf{xwx}' \mid \mathbf{x}, \mathbf{x}' \in X^*\}$ for a given $\mathbf{w} \in X^*$,
- $D = \{x \in \mathbb{R} \mid x \leqslant s\}$ for some threshold $s \in \mathbb{R}$ (observe that $D \subsetneq \mathbb{R} \subsetneq \mathbb{R}^*$).

The function classes defined above can be motivated by indexation techniques in natural language processing (NLP) as they include noteworthy examples such as the Soundex encoding and its variants (see, e.g., [1,2]).

Example 1.2. Let $X = \{a, b, \ldots, z\}$, let $\mathbf{w} \in X^*$, and let $D = X^*\mathbf{w}X^*$. Consider also $F\colon X^* \to X^*$ defined by $F(\mathbf{x}) = \mathbf{w}$ if $\mathbf{x} \in X^*\mathbf{w}X^*$, and $F(\mathbf{x}) = \varepsilon$, otherwise. It is easy to see that F is in $\mathcal{A}'_D$. However, it is not in $\mathcal{A}$ unless $|\mathbf{w}| \leqslant 1$. For example, if $\mathbf{w} = ab$, then $F(ab) = ab \neq \varepsilon = F(F(a)b)$.

Fact 1.3. *For any nonempty subsets D_1 and D_2 of X^* such that $D_1 \subseteq D_2$, the following inclusions hold:*

$$\mathcal{A}_{D_2} \subseteq \mathcal{A}_{D_1}, \qquad \mathcal{A}'_{D_2} \subseteq \mathcal{A}'_{D_1},$$

$$\mathcal{P}_{D_2} \subseteq \mathcal{P}_{D_1}, \qquad \mathcal{P}'_{D_2} \subseteq \mathcal{P}'_{D_1}.$$