The Mathematics behind the Property of Associativity

An invitation to study the many variants of associativity

Bruno Teheux
and Jean-Luc Marichal

University of Luxembourg

Associativity for binary functions

$X, Y \equiv$ non-empty sets

$F\colon X \times X \to X$ is *associative* if

$$F(x, F(y,z)) = F(F(x,y),z)$$

Associativity enables us to define expressions like

$$F(x,y,z,t)$$

$$\begin{aligned} &= F(F(F(x,y),z),t) \\ &= F(x, F(F(y,z),t)) \\ &= \cdots \end{aligned}$$

Define $F\colon \bigcup_{n\geq 2} X^n \to X : \mathbf{x} \in X^n \mapsto F(x_1,\ldots,x_n)$

Notation

We regard n-tuples $\mathbf{x}$ in X^n as *n-strings* over X

0-string: ε
1-strings: $x, y, z, \ldots$
n-strings: $\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots$
$|\mathbf{x}| \;=\;$ length of $\mathbf{x}$

$$X^* := \bigcup_{n \geq 0} X^n$$

We endow X^* with concatenation (X^* is a free monoid)

Any $F : X^* \to Y$ is called a *variadic function*, and we set

$$F_n := F|_{X^n}.$$

We assume

$$\boxed{F(\mathbf{x}) \;=\; \varepsilon \quad \Longleftrightarrow \quad \mathbf{x} \;=\; \varepsilon}$$

Associativity for variadic operations

$F\colon X^* \to X \cup \{\varepsilon\}$ is called a *variadic operation*.

Definition. $F\colon X^* \to X \cup \{\varepsilon\}$ is *associative* if

$$F(\mathbf{xyz}) \;=\; F(\mathbf{x}F(\mathbf{y})\mathbf{z}) \qquad \forall\, \mathbf{xyz} \in X^*$$

Examples.

- the sum $x_1 + \cdots + x_n$,
- the minimum $x_1 \wedge \ldots \wedge x_n$,
- variadic extensions of binary associative functions.

F_1 may differ from the identity map!

Associativity for string functions

Definition. $F\colon X^* \to X^*$ is *associative* if

$$F(\mathbf{xyz}) \;=\; F(\mathbf{x}F(\mathbf{y})\mathbf{z}) \qquad \forall\ \mathbf{xyz} \in X^*$$

Associativity for string functions

$$F(\mathbf{xyz}) = F(\mathbf{x}F(\mathbf{y})\mathbf{z}) \qquad \forall\, \mathbf{xyz} \in X^*$$

Examples.

- sorting in alphabetical order
- letter removing, duplicate removing

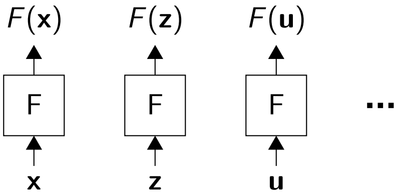
Associativity for string functions

$$F(\mathbf{xyz}) = F(\mathbf{x}F(\mathbf{y})\mathbf{z}) \qquad \forall\ \mathbf{xyz} \in X^*$$

Examples. [. . .] duplicate removing

INPUT: $\mathbf{xzu}\cdots$ in blocks of unknown length given at unknown time intervals.

OUTPUT: $F(\mathbf{xzu}\cdots)$



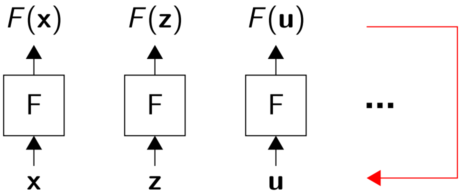
Associativity for string functions

$$F(\mathbf{xyz}) \;=\; F(\mathbf{x}F(\mathbf{y})\mathbf{z}) \qquad \forall\ \mathbf{xyz} \in X^*$$

Examples. [. . .] duplicate removing

INPUT: $\mathbf{xzu}\cdots$ in blocks of unknown length given at unknown time intervals.

OUTPUT: $F(\mathbf{xzu}\cdots)$



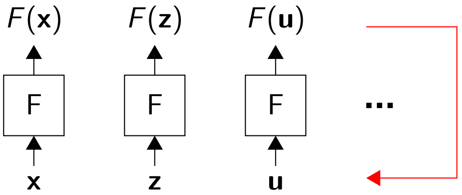
Associativity for string functions

$$F(\mathbf{xyz}) = F(\mathbf{x}F(\mathbf{y})\mathbf{z}) \qquad \forall\ \mathbf{xyz} \in X^*$$

Examples. [...] duplicate removing

INPUT: $\mathbf{xzu}\cdots$ in blocks of unknown length given at unknown time intervals.

OUTPUT: $F(\mathbf{xzu}\cdots)$



"Highly" distributed algorithms

Associativity for string functions

$$F(\mathbf{xyz}) \;=\; F(\mathbf{x}F(\mathbf{y})\mathbf{z}) \qquad \forall\ \mathbf{xyz} \in X^*$$

Proposition.

(1) If $F, G\colon X^* \to X^*$ are associative, then

$$F \;=\; G \qquad \Longleftrightarrow \qquad (F_1 = G_1 \;\text{ and }\; F_2 = G_2)$$

(2) $G\colon X^2 \to X$ is associative if and only if it admits a variadic associative extension $F\colon X^* \to X \cup \{\varepsilon\}$ (*i.e.*, $F_2 = G$).

Preassociative variadic functions

Definition. We say that $F\colon X^* \to Y$ is *preassociative* if

$$F(\mathbf{y}) \ = \ F(\mathbf{y}') \quad \Rightarrow \quad F(\mathbf{x}\mathbf{y}\mathbf{z}) \ = \ F(\mathbf{x}\mathbf{y}'\mathbf{z})$$

Examples. $F_n(\mathbf{x}) = x_1^2 + \cdots + x_n^2 \quad (X = Y = \mathbb{R})$
$F_n(\mathbf{x}) = |\mathbf{x}| \quad (X \text{ arbitrary}, \ Y = \mathbb{N})$

Slogan. Preassociativity is a *composition-free* version of associativity.

Fact. For $F\colon X^* \to Y$

$$F \text{ is preassociative} \qquad \Longleftrightarrow \qquad \ker(F) \text{ is a congruence on } X^*$$

Associative and preassociative functions

Proposition. Let $F\colon X^* \to X^*$.

$$F \text{ is associative}$$
$$\iff$$
$$F \text{ is preassociative} \quad \text{and} \quad F \circ F = F.$$

Proposition Let $F\colon X^* \to \mathrm{ran}(F)$ be preassociative and

$$g\colon \mathrm{ran}(F) \to Z$$

If g is one-to-one or constant, then $g \circ F$ is preassociative.

Problem. Let $F\colon X^* \to Y$ be preassociative. For which g is $g \circ F$ preassociative?

Hard! Characterize $[\ker(F))$ in the congruence lattice of X^*.

Associative and preassociative functions

Theorem. (AC) Let $F\colon X^* \to Y$. The following conditions are equivalent.

(i) F is preassociative.

(ii) $F = f \circ H$ where

$H\colon X^* \to X^*$ is associative and $f\colon \mathrm{ran}(H) \to Y$ is one-to-one.

Associative and preassociative functions

Theorem. (AC) Let $F\colon X^* \to Y$. The following conditions are equivalent.

(i) F is preassociative.

(ii) $F = f \circ H$ where

$H\colon X^* \to X^*$ is associative and $f\colon \mathrm{ran}(H) \to Y$ is one-to-one.

Proof.

$$X^* \underset{g}{\overset{F}{\rightleftarrows}} \mathrm{ran}(F)$$

Define

$$g(F(\mathbf{x})) \in \mathbf{x}/\ker(F),$$

$$H := g \circ F,$$

then

$$F = F \circ H.$$

Factorizations lead to axiomatizations of function classes

A three step technique:

(Binary) Start with a class associative functions $F : X^2 \to X$,

(Source) Axiomatize all their associative extensions $F : X^* \to X \cup \{\varepsilon\}$,

(Target) Use factorization theorem to weaken this axiomatization to capture preassociativity.

The methodology will be used for other factorization results.

An example based on Aczélian semigroups

Theorem (Aczél 1949). $H\colon \mathbb{R}^2 \to \mathbb{R}$ is

- continuous
- one-to-one in each argument
- associative

if and only if

$$H(xy) \;=\; \varphi^{-1}(\varphi(x) + \varphi(y))$$

where $\varphi\colon \mathbb{R} \to \mathbb{R}$ is continuous and strictly monotone.

Source class of associative variadic operations

$$H_n(\mathbf{x}) \;=\; \varphi^{-1}(\varphi(x_1) + \cdots + \varphi(x_n))$$

An example based on Aczélian semigroups

Target axiomatization theorem

Let $F\colon \mathbb{R}^* \to \mathbb{R} \cup \{\varepsilon\}$. The following assertions are equivalent:

(i) F is preassociative and

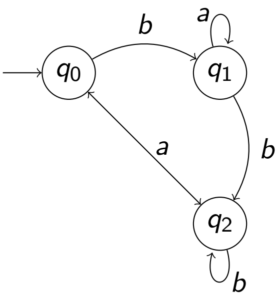
- $\mathrm{ran}(F_1) = \mathrm{ran}(F)$,
- F_1 and F_2 are continuous,
- F_1 and F_2 one-to-one in each argument,

(ii) we have

$$F_n(\mathbf{x}) \;=\; \psi\big(\varphi(x_1) + \cdots + \varphi(x_n)\big)$$

where $\varphi, \psi\colon \mathbb{R} \to \mathbb{R}$ are continuous and strictly monotone.

Transition systems



For instance,

$$\delta(q_0, ababb) = q_2$$

A *transition system* over X:

$$\mathcal{A} = (Q, q_0, \delta)$$

where $q_0 \in Q$ is the *initial state* and

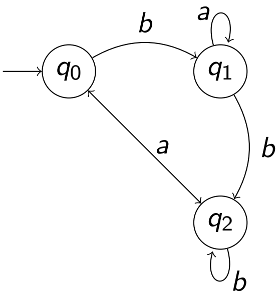
$$\delta \colon Q \times X \to Q$$

is the *transition function.*

The map δ is extended to $Q \times X^*$ by

$$\delta(q, \varepsilon) := q,$$
$$\delta(q, \mathbf{x}y) := \delta(\delta(q, \mathbf{x}), y)$$

Transition systems



A *transition system* over X:

$$\mathcal{A} = (Q, q_0, \delta)$$

where $q_0 \in Q$ is the *initial state* and

$$\delta \colon Q \times X \to Q$$

is the *transition function.*

The map δ is extended to $Q \times X^*$ by

$$\delta(q, \varepsilon) := q,$$
$$\delta(q, \mathbf{x}y) := \delta(\delta(q, \mathbf{x}), y)$$

For instance,

$$\delta(q_0, ababb) = q_2$$

Definition.

$$F_{\mathcal{A}}(\mathbf{x}) := \delta(q_0, \mathbf{x})$$

Preassociativity and transition systems

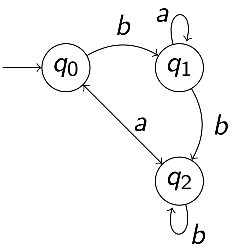
$$F_{\mathcal{A}}(\mathbf{x}) := \delta(q_0, \mathbf{x})$$

Fact. If $\mathcal{A}$ is transition system,

- $F_{\mathcal{A}}$ is "half"-preassociative:

$$F_{\mathcal{A}}(\mathbf{x}) = F_{\mathcal{A}}(\mathbf{y}) \implies F_{\mathcal{A}}(\mathbf{xz}) = F_{\mathcal{A}}(\mathbf{yz})$$

- $F_{\mathcal{A}}$ may not be preassociative:



$$F_{\mathcal{A}}(b) = q_1 = F_{\mathcal{A}}(ba)$$
$$F_{\mathcal{A}}(bb) = q_2 \neq q_0 = F_{\mathcal{A}}(bba)$$

Preassociativity and transition systems

$$F_{\mathcal{A}}(\mathbf{x}) := \delta(q_0, \mathbf{x})$$

Definition. A transition system is *preassociative* if it satisfies

$$\delta(q_0, \mathbf{x}) = \delta(q_0, \mathbf{y}) \quad \Longrightarrow \quad \delta(q_0, z\mathbf{x}) = \delta(q_0, z\mathbf{y})$$

Lemma.

$$\mathcal{A} \text{ preassociative} \quad \Longleftrightarrow \quad F_{\mathcal{A}} \text{ preassociative}$$

Preassociativity and transition systems

$$F_{\mathcal{A}}(\mathbf{x}) := \delta(q_0, \mathbf{x})$$

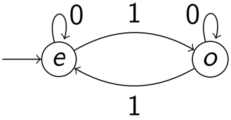
Definition. A transition system is *preassociative* if it satisfies

$$\delta(q_0, \mathbf{x}) = \delta(q_0, \mathbf{y}) \quad \Longrightarrow \quad \delta(q_0, z\mathbf{x}) = \delta(q_0, z\mathbf{y})$$

Lemma.

$$\mathcal{A} \text{ preassociative} \quad \Longleftrightarrow \quad F_{\mathcal{A}} \text{ preassociative}$$

Example. $X = \{0, 1\}$



$$F_{\mathcal{A}}(\mathbf{x}) = e \iff \#\{i \mid x_i = 1\} \text{ is even,}$$
$$F_{\mathcal{A}}(\mathbf{x}) = o \iff \#\{i \mid x_i = 1\} \text{ is odd.}$$

Preassociativity and transition systems

X, Q finite.

Definition. For an onto $F\colon X^* \to Q$, set

$$q_0 := F(\varepsilon),$$
$$\delta(q, z) := \{F(\mathbf{x}z) \mid q = F(\mathbf{x})\},$$
$$\boxed{\mathcal{A}^F := (Q, q_0, \delta)}$$

Generally, $\mathcal{A}^F$ is a non-deterministic transition system.

Lemma.

F is preassociative $\iff$ $\mathcal{A}^F$ is deterministic and preassociative

A criterion for preassociativity

F is preassociative $\iff$ $\mathcal{A}^F$ is deterministic and preassociative

For any state q of $\mathcal{A} = (Q, q_0, \delta)$, any $L \subseteq 2^{X^*}$ and $z \in X$, set

$$L^{\mathcal{A}}(q) := \{\mathbf{x} \in X^* \mid \delta(q_0, \mathbf{x}) = q\}$$
$$z.L := \{z\mathbf{x} \mid \mathbf{x} \in L\}$$

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

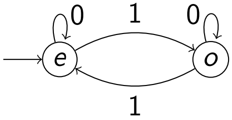
(i) $\mathcal{A}$ is preassociative,

(ii) for all $z \in X$ and $q \in Q$,

$$z.L^{\mathcal{A}}(q) \subseteq L^{\mathcal{A}}(q'), \qquad \text{for some } \; q' \in Q.$$

$$z.L^{\mathcal{A}}(q) \subseteq L^{\mathcal{A}}(q'), \qquad \text{for some} \quad q' \in Q.$$

Example. $X = \{0, 1\}$



$$L^{\mathcal{A}}(e) = \{\mathbf{x} \mid \mathbf{x} \text{ contains an even number of } 1\}$$
$$L^{\mathcal{A}}(o) = \{\mathbf{x} \mid \mathbf{x} \text{ contains an odd number of } 1\}$$

$$0.L^{\mathcal{A}}(o) \subseteq L^{\mathcal{A}}(o) \qquad 0.L^{\mathcal{A}}(e) \subseteq L^{\mathcal{A}}(e)$$
$$1.L^{\mathcal{A}}(o) \subseteq L^{\mathcal{A}}(e) \qquad 1.L^{\mathcal{A}}(e) \subseteq L^{\mathcal{A}}(o)$$

Associative length-based functions

Definition. $F\colon X^* \to X^*$ is *length-based* if

$$F \;=\; \phi \circ |\cdot| \qquad \text{for some } \phi\colon \mathbb{N} \to X^*.$$

Proposition. Let $F\colon X^* \to X^*$ be a length-based function. The following conditions are equivalent.

(i) F is associative

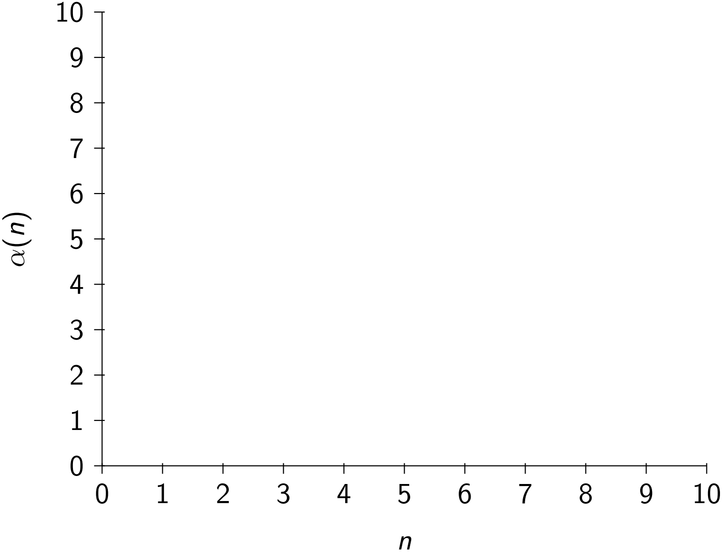
(ii)

$$|F(\mathbf{x})| \;=\; \alpha(|\mathbf{x}|)$$

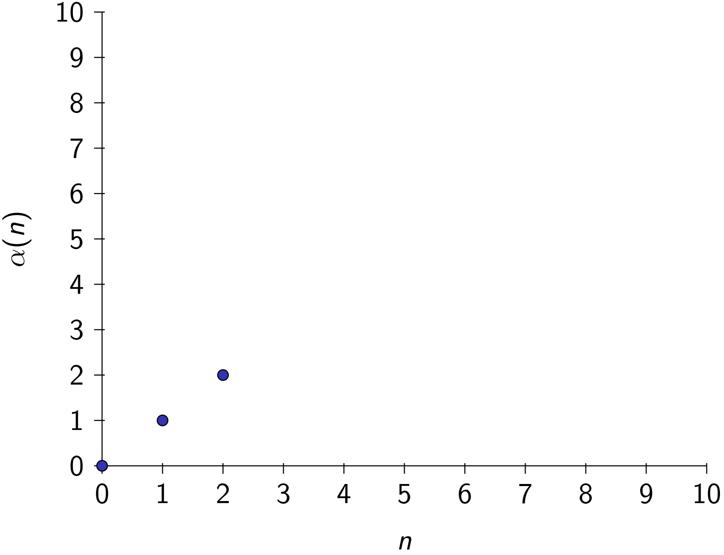
where $\alpha\colon \mathbb{N} \to \mathbb{N}$ satisfies

$$\alpha(n+k) \;=\; \alpha(\alpha(n)+k), \qquad \forall n, k \in \mathbb{N}$$

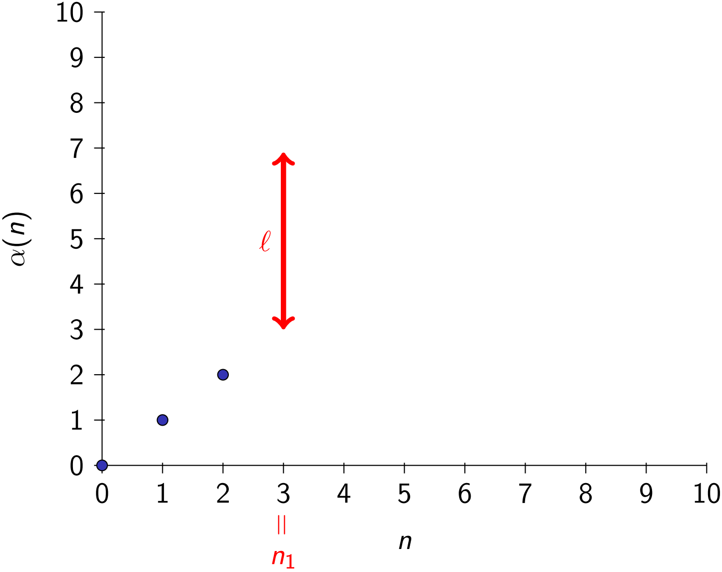
$$\alpha(n+k) = \alpha(\alpha(n)+k), \qquad \forall n, k \in \mathbb{N}$$



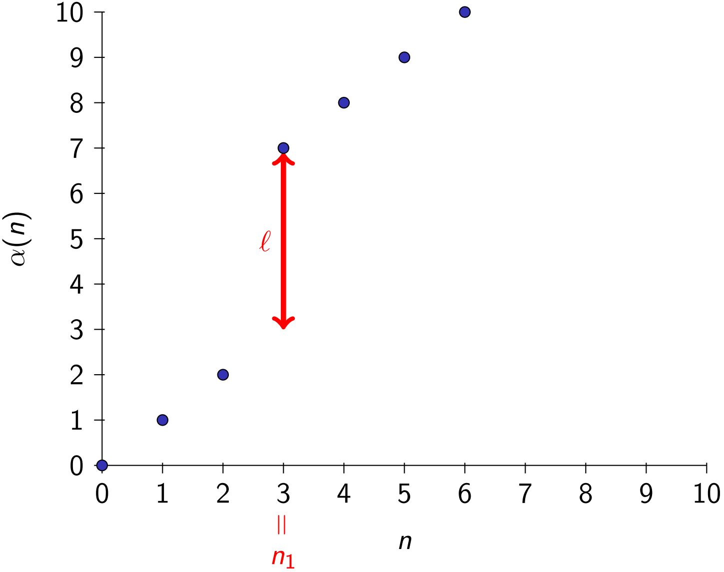
$$\alpha(n+k) = \alpha(\alpha(n)+k), \qquad \forall n, k \in \mathbb{N}$$



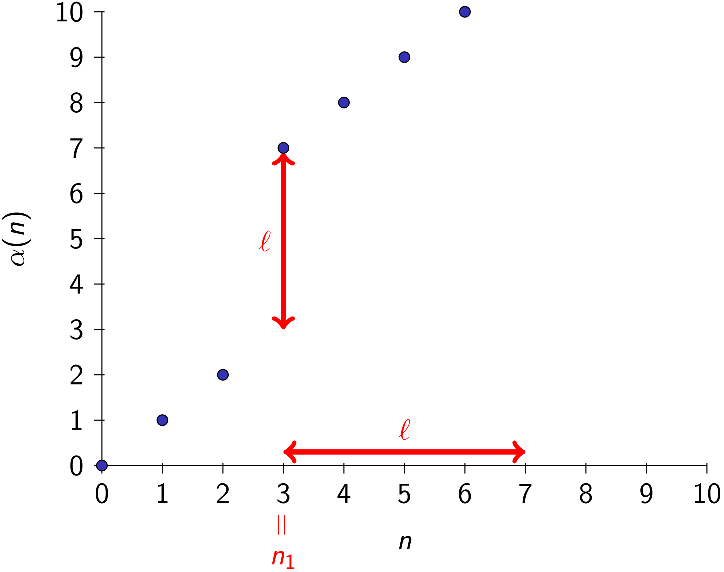
$$\alpha(n+k) = \alpha(\alpha(n)+k), \qquad \forall n,k \in \mathbb{N}$$



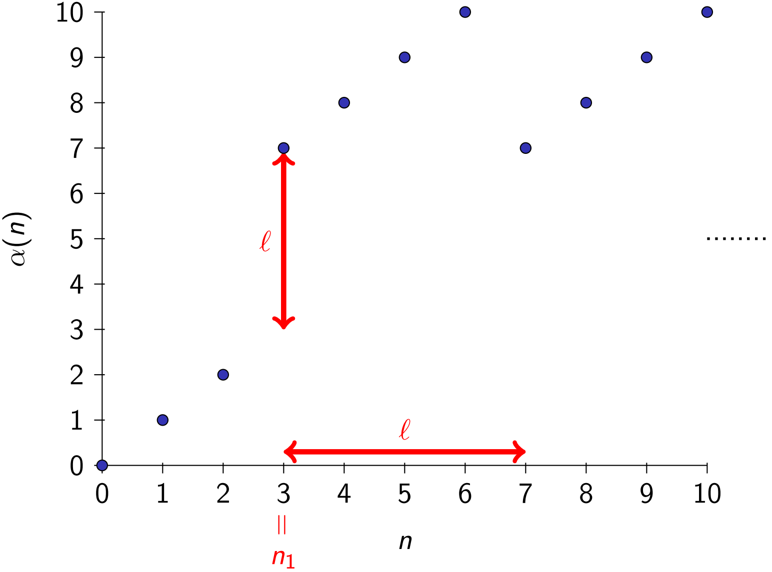
$$\alpha(n+k) \;=\; \alpha(\alpha(n)+k), \qquad \forall n,k \in \mathbb{N}$$



$$\alpha(n+k) \;=\; \alpha(\alpha(n)+k), \qquad \forall n,k \in \mathbb{N}$$



$$\alpha(n+k) \;=\; \alpha(\alpha(n)+k), \qquad \forall n, k \in \mathbb{N}$$



B-associativity and its variants

B-associative functions

Definition. A function $F\colon X^* \to X \cup \{\varepsilon\}$ is *B-associative* if

$$F(\mathbf{x}F(\mathbf{y})^{|\mathbf{y}|}\mathbf{z}) = F(\mathbf{xyz}), \qquad \forall \mathbf{xyz} \in X^*.$$

The function value does not change when replacing every letter of a substring of consecutive letters by the value of the function on this substring.

Example. {Arithmetic, geometric, harmonic} means!

Schimmack (1909), Kolmogoroff (1930), Nagumo (1930).

B-associative functions

The function value does not change when replacing every letter of a substring of consecutive letters by the value of the function on this substring.

B-associative functions

The function value does not change when replacing every letter of a substring ~~of consecutive letters~~ by the value of the function on this substring.

Strongly B-associative functions

Definition. A function $F\colon X^* \to X \cup \{\varepsilon\}$ is *strongly B-associative* if

> *The function value does not change when replacing every letter of a substring ~~of consecutive letters~~ by the value of the function on this substring.*

For instance,

$$\begin{aligned} F(x_1x_2x_3x_4x_5) &= F(F(x_1x_3)x_2F(x_1x_3)x_4x_5), \\ &= F(F(x_1x_3)x_2F(x_1x_3)F(x_4x_5)F(x_4x_5)). \end{aligned}$$

Strongly B-associative functions

Fact.

$$\text{Strongly B-associative} \quad \left\{ \begin{array}{c} \Longrightarrow \\ \not\Longleftarrow \end{array} \right\} \text{B-associative}$$

Example.

$$F(\mathbf{x}) = \sum_{i=1}^{n} \frac{2^{i-1}}{2^n - 1} x_i \qquad \text{is (not strongly) B-associative}$$

Strongly B-associative functions

Fact.

$$\text{Strongly B-associative} \quad \left\{ \begin{array}{c} \Longrightarrow \\ \not\Longleftarrow \end{array} \right\} \quad \text{B-associative}$$

Example.

$$F(\mathbf{x}) = \sum_{i=1}^{n} \frac{2^{i-1}}{2^n - 1} x_i \quad \text{is (not strongly) B-associative}$$

Proposition. The following conditions are equivalent.

(i) F is strongly B-associative

(ii)

$$F(\mathbf{xyz}) \;=\; F(F(\mathbf{xz})^{|\mathbf{x}|}\mathbf{y}F(\mathbf{xz})^{|\mathbf{z}|}) \qquad \forall\ \mathbf{xyz} \in X^*$$

Strong B-associativity and symmetry

Fact.

$$\text{B-associative} \; + \; \text{symmetric} \quad \left\{ \begin{array}{c} \Longrightarrow \\ \not\Longleftarrow \end{array} \right\} \text{strongly B-associative}$$

Example.

$$F(\mathbf{x}) = x_1 \qquad \text{is strongly B-associative but not symmetric}$$

Proposition. If $F\colon X^* \to X \cup \{\varepsilon\}$ is strongly B-associative, then $\mathbf{y} \mapsto F(x\mathbf{y}z)$ is symmetric for every $xz \in X^2$.

A composition-free version of strong B-associativity

Definition. $F\colon X^* \to Y$ is *strongly B-preassociative* if

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

Example. The length function $F\colon X^* \to \mathbb{R}\colon \mathbf{x} \mapsto |\mathbf{x}|$ is strongly B-preassociative.

Strongly B-associative and B-preassociative functions

Proposition. Let $F\colon X^* \to X \cup \{\varepsilon\}$. The following conditions are equivalent.

(i) F is strongly B-associative.

(ii) F is strongly B-preassociative and satisfies $F(F(\mathbf{x})^{|\mathbf{x}|}) = F(\mathbf{x})$.

Strongly B-associative and B-preassociative functions

Proposition. Let $F\colon X^* \to X \cup \{\varepsilon\}$. The following conditions are equivalent.

(i) F is strongly B-associative.

(ii) F is strongly B-preassociative and satisfies $F(F(\mathbf{x})^{|\mathbf{x}|}) = F(\mathbf{x})$.

Theorem. (AC) Let $F\colon X^* \to Y$. The following conditions are equivalent.

(i) F is strongly B-preassociative and $\mathrm{ran}(F_n) = \{F(x^n) \mid x \in X\}$ for all n;

(ii) $F_n = f_n \circ H_n$ for every $n \geq 1$ where

- $H\colon X^* \to X \cup \{\varepsilon\}$ is strongly B-associative,
- $f_n\colon \mathrm{ran}(H_n) \to Y$ is one-to-one.

Strongly B-preassociative and associative functions

$H\colon X^* \to X^*$ is *length-preserving* if $|H(\mathbf{x})| = |\mathbf{x}|$ for all $\mathbf{x} \in X^*$.

Theorem. (AC) Let $F\colon X^* \to Y$. The following conditions are equivalent.

(i) F is strongly B-preassociative.

(ii) $F_n = f_n \circ H_n$ for every $n \geq 1$ where

- $H\colon X^* \to X^*$ is
  associative
  length-preserving
  strongly B-preassociative,
- $f_n\colon \mathrm{ran}(H_n) \to Y$ is one-to-one.

From the factorization theorem to axiomatizations of function classes

(Source) Start with a class of strongly B-associative functions which is axiomatized,

(Target) Use factorization theorem to weaken this axiomatization to capture strongly B-preassociativity.

An example based on quasi-arithmetic means

$\mathbb{I} \equiv$ non-trivial real interval.

Definition. $F\colon \mathbb{I}^* \to \mathbb{R}$ is a *quasi-arithmetic pre-mean function* if

$$F(\mathbf{x}) = f_n\Big(\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big), \qquad n \geq 1, \mathbf{x} \in X^n.$$

where f, f_n are

continous and strictly increasing

If $f_n = f^{-1}$ for every $n \geq 1$ then F is a *quasi-arithmetic mean.*

Example. The product function is a quasi-arithmetic pre-mean function over $\mathbb{I} =]0, +\infty[$ (take $f_n(x) = \exp(nx)$ and $f(x) = \ln(x)$) which is not a quasi-arithmetic mean function.

Characterization of quasi-arithmetic mean functions

Theorem (Kolmogoroff - Nagumo). Let $F\colon \mathbb{I}^* \to \mathbb{I}$. The following conditions are equivalent.

(i) F is a quasi-arithmetic mean function.

(ii) F is B-associative, and for every $n \geq 1$, F_n is

symmetric,

continuous,

strictly increasing in each argument,

reflexive.

Theorem. B-associativity and symmetry can be replaced by strong B-associativity. Moreover, reflexivity can be removed.

Characterization of quasi-arithmetic pre-mean functions

(Source) Quasi-arithmetic mean functions.

Theorem. (Target) Let $F\colon \mathbb{I}^* \to \mathbb{R}$. The following conditions are equivalent.

(i) F is a quasi-arithmetic pre-mean function

(ii) F is strongly B-preassociative, and for every $n \geq 1$, F_n is

symmetric,

continuous,

strictly increasing in each argument.