Doklady Mathematics, Vol. 69, No. 1, 2004, pp. 152–154. Translated from Doklady Akademii Nauk, Vol. 394, No. 4, 2004, pp. 462–464.

COMPUTER SCIENCE

Correct Sigma-Pi Extensions of One Admissible Class of Algorithms

Z. M. Shibzukhov

Presented by Academician Yu.I. Zhuravlev July 28, 2003

Received August 5, 2003

Special Sigma-Pi extensions of one class of algorithms are used here for effectively constructing correct recognition algorithms based on standard training data in the algebraic approach to recognition problems [1]. An admissible class of algorithms is defined that transform the set of original descriptions of analyzed objects into a set of sparse vectors (internal representation) satisfying one natural requirement, which can be viewed as a variant of logical consistency of internal representations of various objects and classes. Algorithms in a Sigma-Pi extension are multilinear forms of algorithms from an admissible class. For a standard recognition problem, it is shown that there exists a nonempty set of correct algorithms from the Sigma-Pi extension of a given finite set of algorithms from an admissible class if it satisfies a simple easily checkable condition. The corresponding algorithmic operators from the Sigma-Pi extension can be constructed following a simple constructive learning recurrence procedure, which converges in a finite number of steps.

ADMISSIBLE CLASSES OF ALGORITHMS AND DECISION RULES

Let $\mathfrak{S}$ be a class of analyzed objects. For each object $S \in \mathfrak{S}$, its standard description $\mathbf{x}(S) = (x_1(S), x_2(S), \ldots, x_n(S))$ is defined; here, x_i: $\mathfrak{S} \to \mathbf{D}_i$, where $\mathbf{D}_i$ is the set of all values of the ith component of the standard description. For each object $S \in \mathfrak{S}$, standard data $\mathbf{y}(S) = (y_1(S), y_2(S), \ldots, y_m(S))$ are also defined; here, y_j: $\mathfrak{S} \to \mathbf{E}_j$, where $\mathbf{E}_j$ is the set of all values of the jth component of the standard data.

Suppose that we are given a finite training set of analyzed objects $\mathbf{S} = \{S_k\} \subseteq \mathfrak{S}$, where $k = 1, 2, \ldots, N$; $\mathbf{X} = \{x_k\}$ is the corresponding set of standard descriptions, where $\mathbf{x}_k = \mathbf{x}(S_k)$; and $\mathbf{Y} = \{y_k\}$ is the corresponding set of standard data, where $\mathbf{y}_k = y(S_k)$. Consider the problem $\mathbf{Z}(\mathbf{X}, \mathbf{Y})$ of constructing an algorithm $\mathbf{A}$ such that $y(\mathbf{x}) \equiv \mathbf{A}(\mathbf{x})$ on $\mathbf{X}$. The algorithm $\mathbf{A}$ is then called correct for the problem $\mathbf{Z}(\mathbf{X}, \mathbf{Y})$.

Let $\mathfrak{A}$ be a class of algorithms Υ: $D_1 \times D_2 \times \ldots \times D_n \to \mathbb{K}^K$, where $\mathbb{K}$ is a ring without zero divisors. Let a finite subset of algorithms $\boldsymbol{\Upsilon} = \{\Upsilon_1, \Upsilon_2, \ldots, \Upsilon_K\} \subset \mathfrak{A}$ be given. We use the notation $\mathbf{U} = \boldsymbol{\Upsilon}(\mathbf{X}) = \{\mathbf{u}_k\}$, where $\mathbf{u}_k = \boldsymbol{\Upsilon}(\mathbf{x}_k)$.

For an arbitrary $u \in \mathbb{K}^K$, we define the function

$$\delta(u) = \begin{cases} 1, & \text{if } u \neq 0 \\ 0, & \text{if } u = 0. \end{cases}$$

For an arbitrary vector $\mathbf{u} \in \mathbb{K}^K$, we define the function $\delta(\mathbf{u}) = (\delta(u_1), \delta(u_2), \ldots, \delta(u_K))$. Admissible classes of algorithms and admissible sets of algorithms are defined below.

Definition 1. A finite set $\boldsymbol{\Upsilon}$ of algorithms is admissible for $\mathbf{X}$ and $\mathbf{Y}$ if $\mathbf{U} = \boldsymbol{\Upsilon}(\mathbf{X})$ and $\mathbf{Y}$ satisfies the admissibility condition

$$\delta(\mathbf{u}_{k'}) = \delta(\mathbf{u}_{k''}) \Rightarrow \mathbf{u}_{k'} = \mathbf{u}_{k''} \wedge \mathbf{y}_{k'} = \mathbf{y}_{k''}.$$

Let $\mathfrak{X}$ be a class of sets of standard descriptions $\mathbf{X} \subseteq \mathbf{D}_1 \times \mathbf{D}_2 \times \ldots \times \mathbf{D}_n$, $\mathbf{Y}$ be a class of sets of standard data $\mathbf{Y} \subseteq \mathbf{E}_1 \times \mathbf{E}_2 \times \ldots \times \mathbf{E}_m$. Denote by $\mathfrak{Z}(\mathfrak{X}, \mathfrak{Y})$ the class of problems $\mathbf{Z}(\mathbf{X}, \mathbf{Y})$, where $\mathbf{X} \in \mathfrak{X}$ and $\mathbf{Y} \in \mathfrak{Y}$.

Definition 2. $\mathfrak{A}$ is an admissible class of algorithms for $\mathfrak{X}$ and $\mathfrak{Y}$ if, for any pair of finite sets of standard descriptions $\mathbf{X} \in \mathfrak{X}$ and standard data $\mathbf{Y} \in \mathfrak{Y}$, there exists a finite admissible subset of algorithms $\boldsymbol{\Upsilon} \subset \mathfrak{A}$.

Let $\mathbf{dr}$: $\mathbb{K}^m \to \mathbf{E}_1 \times \mathbf{E}_2 \times \ldots \times \mathbf{E}_m$ be a decision rule.

Definition 3. A decision rule $\mathbf{dr}$ is admissible for $\mathbf{Y}$ if, for any vector $s \in \mathbb{K}^m$, any values $0 \neq p_t \in \mathbb{K}$ $(1 \leq t \leq l)$, and any element $\mathbf{y} \in \mathbf{Y}$, there exist vector weights $\mathbf{w}_t \in \mathbb{K}^m$ $(t = 1, 2, \ldots, m)$ that solve the equation

$$\mathbf{dr}(\mathbf{s} + p_1\mathbf{w}_1 + \ldots + p_l\mathbf{w}_l) = \mathbf{y}.$$

Let us define the class of decision rules $\mathfrak{R}$ that are admissible for $\mathfrak{Y}$.

Research Institute of Applied Mathematics and Automation, Kabardino-Balkar Scientific Center, Russian Academy of Sciences, ul. Shortanova 89a, Nalchik, 360000 Russia
e-mail: szport@fromru.com

Definition 4. A class of decision rules $\mathfrak{R}$ is admissible for $\mathfrak{Y}$ if, for any $\mathbf{Y} \in \mathfrak{Y}$, there exists a decision rule $\mathbf{dr} \in \mathfrak{R}$ that is admissible for $\mathbf{Y}$.

ORDERED SPARSE DESCRIPTIONS OF OBJECTS

A vector $\mathbf{u} = (u_1, u_2, \ldots, u_K)$ is sparse if there is an index $1 \le i \le K$ such that $u_i = 0$. Let $\mathbf{I} = \{i_k\}$ be a sequence of multi-indices, $\mathbf{i}_k \subseteq \{1, 2, \ldots, K\}$. For definiteness, we assume that the indices in $\mathbf{i}_k$ are arranged in increasing order.

Definition 5. A sequence $\mathbf{U}$ is ordered relative to zeros (briefly, zero-ordered) with respect to $\mathbf{I}$ if for any pair $j < k$, there exists an index $i \in \mathbf{i}_k$ such that $u_{ji} = 0$ and $u_{ki} \neq 0$.

If $\mathbf{i}_k = \{1, 2, \ldots, K\}$ for all k, then $\{\mathbf{u}_k\}$ is said to be ordered relative to zeros. For example, any sequence $\{\mathbf{u}_k\} \subseteq \{0, 1\}^n$ not containing identical vectors can be ordered relative to zeros. Let $\{\mathbf{c}_k\}$ be an arbitrary sequence of vectors from $\mathbb{K}^K$ that contains no sparse vectors, and let $\{\mathbf{u}_k\}$ be ordered relative to zeros. Then the sequence $\{\mathbf{c}_k \odot \mathbf{u}_k\}$, where $\odot$ denotes the componentwise multiplication of vectors, is ordered relative to zeros.

Let $\mathbf{U}$ and $\mathbf{Y}$ be sets satisfying the admissibility condition. Then $\mathbf{U} = \{\mathbf{u}_k\}$ consists of sparse vectors (except for possibly one vector) and is ordered relative to zeros.

For an arbitrary $\mathbf{U}$ that is ordered relative to zeros, we denote by $\mathfrak{I}(\mathbf{U})$ the set of all sequences of multi-indices $\mathbf{I}$ such that $\mathbf{U}$ is zero-ordered with respect to $\mathbf{I}$.

ALGORITHMIC Sigma-Pi OPERATORS AND Sigma-Pi EXTENSIONS

Let $\mathbb{K}$ be a ring with no zero divisors, for example, $\mathbb{Z}_p$ (p is a prime number), $\mathbb{GF}(p^m)$, $\mathbb{R}$, or $\mathbb{C}$.

We define Sigma-Pi functions having the representation

$$\mathrm{sp}(\mathbf{u}) = \theta(\mathbf{u}) + \sum_{k=1}^{N} w_k \mathrm{p}(\mathbf{u}, \mathbf{i}_k),$$

where $\theta(\mathbf{u})$ is an arbitrary function from $\mathbb{K}^K$ to $\mathbb{K}$, $w_k \in \mathbb{K}$,

$$\mathrm{p}(\mathbf{u}, \mathbf{i}_k) = \prod_{i \in \mathbf{i}_k} u_i,$$

and $\mathrm{p}(\mathbf{u}, \varnothing) \equiv 1$. The class $\mathfrak{SP}$ consists of algorithmic Sigma-Pi operators that can be represented as a composition of the form

$$\mathbf{spo} = \mathbf{dr} \circ \mathbf{sp}(\mathbf{u}),$$

$$\mathbf{sp}(\mathbf{u}) = (\mathrm{sp}_1(\mathbf{u}), \mathrm{sp}_2(\mathbf{u}), \ldots, \mathrm{sp}_m(\mathbf{u})),$$

where $\mathbf{dr} \in \mathfrak{R}$ and $\mathrm{sp}_1(\mathbf{u}), \mathrm{sp}_2(\mathbf{u}), \ldots, \mathrm{sp}_m(\mathbf{u})$ are Sigma-Pi functions.

Definition 6. The Sigma-Pi extension $\mathfrak{SP}(\mathbf{\Upsilon})$ of a set of algorithms $\mathbf{\Upsilon} = \{\Upsilon_1, \Upsilon_2, \ldots, \Upsilon_K\}$ is the following set of algorithms:

$$\mathfrak{SP}(\mathbf{\Upsilon}) = \{\mathbf{spo} \circ \Upsilon\colon \mathbf{spo} \in \mathfrak{SP}\}.$$

Theorem. *Let* $\mathbf{\Upsilon} \in \mathfrak{A}$ *be a finite set of algorithms that is admissible for* $\mathbf{X}$ *and* $\mathbf{Y}$, *and let* $\mathbf{dr} \in \mathfrak{R}$ *be a decision rule that is admissible for* $\mathbf{Y}$. *Then one can construct an algorithmic* Sigma-Pi *operator* $\mathbf{spo} \in \mathfrak{SP}(\mathbf{\Upsilon})$ *such that* $\mathbf{spo} \circ \mathbf{\Upsilon}$ *is correct for* $\mathbf{X}$ *and* $\mathbf{Y}$.

The corresponding algorithmic Sigma-Pi operator can be constructed by using a recurrence learning procedure based on $\mathbf{U}$ and $\mathbf{Y}$, whose elements are previously reordered so that $\mathbf{U}$ becomes zero-ordered with respect to some sequence of multi-indices $\mathbf{I} \in \mathfrak{I}(\mathbf{U})$ [2, 3]. In the general case, one can construct a set of such Sigma-Pi operators by using various $\mathbf{I}$. To reduce the complexity of algorithmic operators, it is possible to use a procedure for constructing minimal (with respect to $\subset$) sequences of multi-indices $\mathbf{I} \in \mathfrak{I}(\mathbf{U})$ [2, 4], where $\mathbf{I}' \subset \mathbf{I}''$ if $\mathbf{i}'_k \subseteq \mathbf{i}''_k$ for any k and there is at least one k such that $\mathbf{i}'_k \subset \mathbf{i}''_k$.

Definition 7. The Sigma-Pi extension $\mathfrak{SP}(\mathfrak{A})$ of a class $\mathfrak{A}$ of admissible algorithms is the following class of algorithms:

$$\mathfrak{SP}(\mathfrak{A}) = \{\mathbf{spo} \circ \Upsilon\colon \mathbf{spo} \in \mathfrak{SP}, \mathbf{\Upsilon} \subset \mathfrak{A}\}.$$

Theorem. *If a class* $\mathfrak{A}$ *of algorithms is admissible for* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *and if a class* $\mathfrak{R}$ *of decision rules is admissible for* $\mathfrak{Y}$, *then the Sigma-*Pi *extension* $\mathfrak{SP}(\mathfrak{A})$ *is correct for the class of problems* $\mathfrak{Z}(\mathfrak{X}, \mathfrak{Y})$.

ALGORITHMIC SIGMA^2-PI OPERATORS

Let $\mathbf{U}$ be ordered relative to zeros, and let $I_k(\mathbf{U})$ be a set of multi-indices $\mathbf{i}$ such that, for any $j < k$, there exists i such that $u_{ji} = 0$ and $u_{ki} \neq 0$. For $I_k \subseteq I_k(\mathbf{U})$, we define functions of the form

$$\mathrm{spn}_k(\mathbf{u}, I_k) = f \circ \mathrm{sp}_k(\mathbf{u}, I_k),$$

$$\mathrm{sp}_k(\mathbf{u}, I_k) = \sum_{\mathbf{i} \in I_k} c_{\mathbf{i}} \cdot \mathrm{pn}(\mathbf{u}, \mathbf{i}),$$

where f is an arbitrary function from $\mathbb{K}$ to $\mathbb{K}$ such that $f(s) = 0 \Leftrightarrow s = 0$, $c_i \in \mathbb{K}$, $\mathrm{sp}_k(\mathbf{u}_k, I_k) \neq 0$, and

$$\mathrm{pn}(\mathbf{u}, \mathbf{i}) = g_{\mathbf{i}} \circ \mathrm{p}(\mathbf{u}, \mathbf{i}),$$

where g_i are arbitrary functions from $\mathbb{K}$ to $\mathbb{K}$ such that

$g_i(s) = 0 \Leftrightarrow s = 0$. Consider functions of the form

$$\mathrm{Sp}(\mathbf{u}) = \theta(\mathbf{u}) + \sum_{k=1}^{N} w_k \mathrm{spn}_k(\mathbf{u}, I_k).$$

Define the class $\mathfrak{S}^2\mathfrak{P}$ of algorithmic Sigma^2-Pi-operators of the form

$$\mathbf{Spo}(\mathbf{u}) = \mathbf{dr} \circ \mathbf{Sp}(\mathbf{u}),$$
$$\mathbf{Sp}(\mathbf{u}) = (\mathrm{Sp}_1(\mathbf{u}), \mathrm{Sp}_2(\mathbf{u}), \ldots, \mathrm{Sp}_K(\mathbf{u})).$$

Definition 8. The Sigma^2-Pi extension $\mathfrak{S}^2\mathfrak{P}(\boldsymbol{\Upsilon})$ of a given set $\boldsymbol{\Upsilon}$ of algorithms is the following set of algorithms:

$$\mathfrak{S}^2\mathfrak{P}(\boldsymbol{\Upsilon}) = \{\mathbf{Spo} \circ \Upsilon \colon \mathbf{Spo} \in \mathfrak{S}^2\mathfrak{P}\}.$$

Theorem. *Let* $\boldsymbol{\Upsilon} \subset \mathfrak{A}$ *be a finite set of algorithms that is admissible for* $\mathbf{X}$ *and* $\mathbf{Y}$, *and let* $\mathbf{dr} \in \mathfrak{R}$ *be a decision rule that is admissible for* $\mathbf{Y}$. *Then one can construct an algorithmic Sigma*^{2}*-Pi operator* $\mathbf{Spo} \in \mathfrak{S}^2\mathfrak{P}(\boldsymbol{\Upsilon})$ *such that* $\mathbf{Sp} \circ \boldsymbol{\Upsilon}$ *is correct for* $\mathbf{X}$ *and* $\mathbf{Y}$.

The corresponding generalized Sigma-Pi operator can be constructed by using a similar constructive learning procedure based on $\mathbf{U}$ *and* $\mathbf{Y}$, *whose elements are previously reordered so that* $\mathbf{U}$ *becomes ordered relative to zeros. In the general case, it is possible to construct a set of such generalized Sigma*^{2}*-Pi operators by using various sets* $\{I_k\}$.

Definition 9. The Sigma^2-Pi extension $\mathfrak{S}^2\mathfrak{P}(\mathfrak{A})$ of a class $\mathfrak{A}$ of admissible algorithms is the following class of algorithms:

$$\mathfrak{S}^2\mathfrak{P}(\mathfrak{A}) = \{\mathbf{spo} \circ \Upsilon \colon \mathbf{spo} \in \mathfrak{S}^2\mathfrak{P}, \boldsymbol{\Upsilon} \subset \mathfrak{A}\}.$$

Theorem. *If a class* $\mathfrak{A}$ *of algorithms is admissible for* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *and if a class* $\mathfrak{R}$ *of decision rules is admissible for* $\mathfrak{Y}$, *then the Sigma*^{2}*-Pi extension* $\mathfrak{S}^2\mathfrak{P}(\mathfrak{A})$ *is correct for the class of problems* $\mathfrak{Z}(\mathfrak{X}, \mathfrak{Y})$.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 01-01-00142), the Sixth Contest of Scientific Projects of Young Scientists of the Russian Academy of Sciences (project no. 111), and state contract no. 10002-251/OMN-2/024-115/120503-062.

REFERENCES

1. Yu. I. Zhuravlev, *Selected Scientific Works* (Magistr, Moscow, 1998) [in Russian].
2. Z. M. Shibzukhov, Neĭrokomp. Razrabotka Primen., No. 5, 50 (2002).
3. Z. M. Shibzukhov, Dokl. Akad. Nauk **388**, 174 (2003) [Dokl. Math. **67**, 134 (2003)].
4. Z. M. Shibzukhov, Zh. Vychisl. Mat. Mat. Fiz **43**, 1260 (2003) [Comp. Math. Math. Phys. **43**, 1209 (2003)].