Applications & Science of Computational Intelligence III

Vol. 4055

PROCEEDINGS OF SPIE
SPIE—The International Society for Optical Engineering

Applications and Science of Computational Intelligence III

Kevin L. Priddy
Paul E. Keller
David B. Fogel
Chairs/Editors

24–27 April 2000
Orlando, Florida

Sponsored and Published by
SPIE—The International Society for Optical Engineering

SPIE PROCEEDINGS SERIES

Volume 4055

SPIE is an international technical society dedicated to advancing engineering and scientific applications of optical, photonic, imaging, electronic, and optoelectronic technologies.

The papers appearing in this book compose the proceedings of the technical conference cited on the cover and title page of this volume. They reflect the authors' opinions and are published as presented, in the interests of timely dissemination. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by SPIE. Papers were selected by the conference program committee to be presented in oral or poster format, and were subject to review by volume editors or program committees.

Please use the following format to cite material from this book:

Author(s), "Title of paper," in *Applications and Science of Computational Intelligence III*, Kevin L. Priddy, Paul E. Keller, David B. Fogel, Editors, Proceedings of SPIE Vol. 4055, page numbers (2000).

ISSN 0277-786X
ISBN 0-8194-3681-X

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 360/676-3290 (Pacific Time) • Fax 360/647-1445

Printed in the United States of America.

Conference Committee

Conference Chairs

Kevin L. Priddy, Battelle Memorial Institute
Paul E. Keller, Pacific Northwest National Laboratory
David B. Fogel, Natural Selection, Inc.

Program Committee

Stanley C. Ahalt, The Ohio State University
Peter J. Angeline, Natural Selection, Inc.
Gianfranco Basti, Pontifical Lateran University (Italy)
Joydeep Ghosh, University of Texas at Austin
Charles W. Glover, Oak Ridge National Laboratory
Clifford G. Y. Lau, Office of Naval Research
Karl Mathia, Equipe Technologies, Inc.
Antonio Luigi Perrone, Pontifical Lateran University (Italy)
Steven K. Rogers, Qualia Computing, Inc.
Stephan Rudolph, Universität Stuttgart (Germany)

Session Chairs

1 Theoretical Foundations
Mark E. Oxley, Air Force Institute of Technology

2 Algorithms and Architectures I
Antonio Luigi Perrone, Pontifical Lateran University (Italy)

3 Algorithms and Architectures II
Gianfranco Basti, Pontifical Lateran University (Italy)

4 Evolutionary Computation I
David B. Fogel, Natural Selection, Inc.

5 Evolutionary Computation II
David B. Fogel, Natural Selection, Inc.

6 Hardware
Kevin L. Priddy, Battelle Memorial Institute

7 Neural Networks in Modeling, Identification, and Control
Stephan Rudolph, Universität Stuttgart (Germany)

8 Biology
Krzysztof J. Cios, University of Toledo

9 Applications I
Stanley C. Ahalt, The Ohio State University

10 Applications II
Charles W. Glover, Oak Ridge National Laboratory

11 Signal Processing
William S. Hortos, Florida Institute of Technology

12 Image Processing
Paul E. Keller, Pacific Northwest National Laboratory

13 Intelligent Symbolic Computing
Eugene Santos, Jr., University of Connecticut

14 Grand Challenges
Kevin L. Priddy, Battelle Memorial Institute

Introduction

Welcome to SPIE's third annual conference on the Applications and Science of Computational Intelligence. This year's conference covers theory, algorithms, hardware, and many other unique applications of computational intelligence.

We have provided two panel sessions on Thursday to pique your interest in other methods of computing and in the challenges we all face as we move forward in implementing our technology in the products of tomorrow.

I wish to take this opportunity to thank my co-chairs, Dr. Paul Keller and Dr. David Fogel for their support, as well as the program committee, each of the participants at the conference, and the staff at SPIE.

Please take the time to confer with the authors and with the program committee about your views on the conference.

I look forward to seeing you at the conference.

Kevin L. Priddy

Contents

SESSION 1

Theoretical Foundations

Topological-based Capability Measures of Artificial Neural Network Architectures

Mark E. Oxley[a] and Martha Alvey Carter[b]

[a]Department of Mathematics and Statistics
Air Force Institute of Technology
2950 P Street
Wright-Patterson AFB, OH 45433-7765

[b]National Air Intelligence Center
4115 Hebble Creek Rd.
Wright-Patterson AFB, OH 45433-5622

ABSTRACT

Current measures of an artificial neural network (ANN) capability are based on the V-C dimension and its variations. These measures may be underestimating the actual ANN's capabilities and hence overestimating the required number of examples for learning. This is caused by relying on a single invariant description of the problem set, which, in this case is cardinality, and requiring worst case geometric arrangements and colorings.

A capability measure allows aligning the measure with desired characteristics of the problem sets. The mathematical framework has been established in which to express other desired invariant descriptors of a capability measure. New invariants are defined on the problem space that yield new capability measures of ANNs that are based on topological properties. A specific example of an invariant is given which is based on topological properties of the problem set and yields a new measure of ANN architecture.

Keywords: Artificial neural network architecture, Capability measure, Invariants

1. INTRODUCTION

Let A be an ANN architecture. We want to measure the capability of A. Typically, the capability is defined as the ability of A to successfully classify data, though there are other capabilities one might want to measure. We want to assign a positive real number to the architecture A that corresponds to the measure of A, $\mu(A)$. Given another architecture B, we would like to know if B is better, worse or the same in its capability compared to A. We will make this comparison based upon the measures of the architectures, that is, A is more capable than B if $\mu(A) \geq \mu(B)$. The purpose of this paper is to determine a mapping μ so that we can compare architectures.

2. BACKGROUND

2.1. Measure Theory

A measure is a set function, that is, its input is a set and its output is a positive real number. There are generalizations of measures based upon the type of output. For example, if the output is also negative the measure is called a signed measure. If the output is a complex number then the measure is called a complex measure. Sometimes there are several measures of interest so one may concatenate them into a

E-mail: Mark.Oxley@afit.af.mil, mac05@naic.wpafb.af.mil

In *Applications and Science of Computational Intelligence III*, Kevin L. Priddy, Paul E. Keller, David B. Fogel, Editors, Proceedings of SPIE Vol. 4055 (2000) • 0277-786X/00/$15.00

vector. This yields a vector measure. The fundamental characteristics of a measure are: (1) their output is a scalar that has an ordering (partial or total), and (2) there is one output that corresponds to the input (that is, a function). Recall that the real numbers are totally ordered.

Definition 1. (*Measure*)(See Halmos.[4]) A measure m is a real-valued function that is defined on a σ-algebra of sets $\mathcal{S}$, that is, $m : \mathcal{S} \to \mathbb{R}$, that satisfies the properties:

1. $m(\emptyset) = 0$ ($\emptyset$ is the empty set);
2. $m(S) \geq 0$ for all $S \in \mathcal{S}$;
3. $m(\bigcup_n S_n) = \sum_n m(S_n)$ for all countable disjoint collection $\{S_n\} \subset \mathcal{S}$.

Recall the definition of a σ-algebra is a Boolean algebra with the added property that for each countable collection $\{S_n\} \subset \mathcal{S}$, then

$$\bigcup_n S_n \in \mathcal{S}.$$

A Boolean algebra of sets is a collection of subsets of some *universe* set. Let $\mathcal{Y}$ denote a collection of subsets of $\mathbb{R}^d$. An example of $\mathcal{Y}$ is the power set $\mathcal{P}(\mathbb{R}^d)$ which is a σ-algebra. Other examples of interest are topological spaces.

Definition 2. (*Topological Spaces*)(See Royden.[8]) A topological space $(U, \mathcal{T})$ is a non-empty set U (the universe set) of objects together with a family of subsets $\mathcal{T}$ (the topology) (We say the set $O \in \mathcal{T}$ is an *open* set.) possessing the following properties:

1. $U \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$;
2. for every $O_1, O_2 \in \mathcal{T}$ then $O_1 \cap O_2 \in \mathcal{T}$ (finite intersection);
3. for every subcollection $\{O_\alpha\} \subset \mathcal{T}$ then $\cup_\alpha O_\alpha \in \mathcal{T}$ (infinite unions).

If a σ-algebra also satisfies property (3), then the σ-algebra is also a topological space.

Suppose A has d inputs and 1 output, then A corresponds to a family of functions F. Assume the d inputs are real-valued and the single output is also real-valued. Thus, for an instantiation of a weight vector $w \in W$ (W is the set of possible weight vectors for the architecture), there is a function that corresponds to that fixed network (not necessarily one-to-one; see Carter.[1]) Therefore, we can identify the architecture A with the family of functions

$$F = \{f_w : \mathbb{R}^d \to \mathbb{R} : w \in W\}.$$

For each $f \in F$ we define the set $s_f = \{x \in \mathbb{R}^d : f(x) > 0\}$. The family of these sets is denoted by

$$S = \{s_f \in \mathcal{P}(\mathbb{R}^d) : f \in F\}.$$

Hence, we can identify the family S with architecture A. That is, there exists a mapping ϕ that maps A to S, so $\phi(A) = S$.

Let $\mathcal{A}$ denote a collection of ANN architectures with d inputs and 1 output (both real-valued), so $A \in \mathcal{A}$ is a specific architecture. Let $\mathcal{S}$ denote the corresponding collection of families of sets, so $S \in \mathcal{S}$. From our previous discussion, there exists a mapping ϕ that maps $\mathcal{A}$ to $\mathcal{S}$, that is, $\phi(\mathcal{A}) = \mathcal{S}$. So, the measure of A, denoted by $\mu(A)$, will be defined in terms of a measure of S. That is, we seek a measure m, defined on S, so that $m(S)$ is a positive real number. We define $\mu = m \circ \phi$ since

$$[m \circ \phi](A) = m(\phi(A)) = m(S)$$

then

$$\mu(A) = m(\phi(A)).$$

A word of warning is appropriate now. Given $\mathcal{A}$ we form $\mathcal{S}$, but $\mathcal{S}$ may not be a σ-algebra. This motivates the following theorems.

Theorem 1. Let $\mathcal{S}$ be a collection of sets. There exists a σ-algebra Σ that contains $\mathcal{S}$.

Proof. Let $\mathcal{S}$ be the generator of a σ-algebra, say Σ, then Σ contains $\mathcal{S}$.

Definition 3. Let $\mathcal{S}$ be a collection of sets. Let $\Sigma(\mathcal{S})$ denote the set of σ-algebras that contains $\mathcal{S}$. That is,

$$\Sigma(\mathcal{S}) = \{\Sigma : \Sigma \text{ is a } \sigma\text{-algebra, } \Sigma \supseteq \mathcal{S}\}.$$

Theorem 2. Let $\mathcal{S}$ be a collection of sets. There exists a smallest σ-algebra that contains $\mathcal{S}$, denoted $\sigma(\mathcal{S})$. That is, if $\Sigma_1 \in \Sigma(\mathcal{S})$ and $\Sigma_1 \subset \sigma(\mathcal{S})$, then, in fact, $\Sigma_1 = \sigma(\mathcal{S})$.

Proof. $(\Sigma(\mathcal{S}), \subseteq)$ forms a partially ordered set. Applying Zorn's Lemma[5] gives the results.

The capability of the family S is based on its ability to implement dichotomies of data in $\mathbb{R}^d$. We review this to establish some notation.

2.2. Classification Problem

Let $X \subset \mathbb{R}^d$ be a non-empty set. The dichotomy of X is the partition of its elements into two disjoint subsets $X^+$and X^-, such that $X^+ \cup X^- = X$ and $X^+ \cap X^- = \emptyset$. A dichotomy is denoted by the ordered pair (X^+, X^-) and is also referred to as a signed set. A signed set $\mathbf{X} = (X^+, X^-)$ is said to be implemented by a function $f \in F$ if

$$f(x) > 0 \quad \text{for all} \quad x \in X^+$$

$$f(x) < 0 \quad \text{for all} \quad x \in X^-.$$

It will be convenient to say that the family F implements the signed set $\mathbf{X}$ if there exists some $f \in F$ such that f implements that signed set $\mathbf{X}$. The (two-class) classification problem is posed as: given a signed set $\mathbf{X}$ find a function $f \in F$ which implements it. Typically, the set X is finite, that is, the cardinality of X, denoted by card(X), is a (finite) positive integer. Because of the relationship between the family F and the family S, we will say the family S implements the signed set $\mathbf{X}$ if there exists a set $s \in S$ such that

$$s \cap X^+ = X^+ \quad \text{and} \quad \tilde{s} \cap X^- = X^-$$

where $\tilde{s}$ denotes the set complement of s with respect to $\mathbb{R}^d$. The equivalent statements are,

$$\tilde{s} \cap X^+ = \emptyset \quad \text{and} \quad s \cap X^- = \emptyset.$$

The measure of S is based on some quantifier Q acting on the collection of signed sets that S can implement. We are interested in finite data, so let $\mathcal{X}$ denote the collection of finite signed sets in $\mathbb{R}^d$, that is,

$$\mathcal{X} = \{\mathbf{X} = (X^+, X^-) : \text{card}(X^+) + \text{card}(X^-) < \infty\}.$$

For each set $s \in S$ define the family of finite sets that s can implement, to be

$$i(s) = \{\mathbf{X} \in \mathcal{X} : s \text{ implements } \mathbf{X}\}.$$

Define the union over all these sets to be

$$\mathcal{I}(S) = \bigcup_{s \in S} i(s) = \{\mathbf{X} \in \mathcal{X} : s \text{ implements } \mathbf{X} \text{ for some } s \in S\}.$$

Now define the measure m to be in terms of Q acting on $\mathcal{I}(S)$ by

$$m(S) = Q(\mathcal{I}(S)).$$

2.3. Examples

We give some examples of quantifiers Q. Most of the interesting cases involve a set function q acting on sets in $\mathcal{X}$, such that $q(\mathbf{X}) \geq 0$ for each $\mathbf{X} \in \mathcal{X}$. Note that q is almost a measure (called a semi-measure), but may not satisfy the property (3) in the definition of a measure. Define Q in terms of q by

$$Q(\mathcal{C}) = \sup\{q(\mathbf{X}) : \mathbf{X} \in \mathcal{C}\}$$

for a subcollection of signed sets $\mathcal{C}$. Once we specific q and $\mathcal{C}$, then the quantifier Q follows, and thus, the measure m.

Example 1. Let $q_1(\mathbf{X}) = \text{card}(\mathbf{X}) = \text{card}(X^+) + \text{card}(X^-) = \text{card}(X)$ and $\mathcal{C}_1 = \mathcal{I}(S)$, then m_1 is

$$m_1(S) = Q_1(\mathcal{I}(S)) = \sup\{\text{card}(\mathbf{X}) : \mathbf{X} \in \mathcal{I}(S)\}.$$

This m is related to the V-C dimension, but not the V-C dimension.

Example 2. Let $q_2(\mathbf{X}) = \text{card}(\mathbf{X})$ and $\mathcal{C}_2 = \{\mathbf{X} \in \mathcal{I}(S) : S \text{ implements every } \mathbf{Y} \in \mathcal{P}(\mathbf{X})\}$ then m_2 is

$$m_2(S) = Q_2(\mathcal{C}_2) = \sup\{\text{card}(\mathbf{X}) : \mathbf{X} \in \mathcal{C}_2\}.$$

Then m_2 is the V-C dimension.[9,10]

Example 3. Let $q_3(\mathbf{X}) = GC(\mathbf{X})$, the geometric complexity of the signed set $\mathbf{X}$ (See Carter[1] and Carter and Oxley.[3]) and let $\mathcal{C}_3 = \mathcal{I}(S)$, then m_3 is

$$m_3(S) = Q_3(\mathcal{I}(S)) = \sup\{GC(\mathbf{X}) : \mathbf{X} \in \mathcal{I}(S)\}.$$

This is the Ox-Cart dimension.[1,3]

Other choices for q exist. The one that several researchers have sought is the following.

Example 4. Let $q_4(\mathbf{X})$ be given by

$$q_4(\mathbf{X}) = \text{card}(\{\text{hyperlanes that separates } \mathbf{X}\}).$$

But, this mapping is difficult to write out. Take $\mathcal{C}_4 = \mathcal{I}(S)$ then it leads to the measure

$$m_4(S) = Q_4(\mathcal{I}(S)) = \sup\{q_4(\mathbf{X}) : \mathbf{X} \in \mathcal{I}(S)\}.$$

The discovery of other measures of this form relies on the mapping q. We investigate this further. There is another property that q should satisfy and that is an invariant property. We begin with a discussion on the theory of invariants.

2.4. Invariants Theory

Because, an ANN capability measure should be about signed sets, what is sought is a set of invariants and a family of operators that are defined on signed sets. Hence, consider the following definition for a set of invariants on signed sets, $\mathcal{X}$.

Definition 4. We say the quantifier q is *invariant with respect to the family* $\mathcal{G}$ of operators (that act on $\mathcal{X}$) if

$$q(g(\mathbf{X})) = q(\mathbf{X}) \text{ for all } \mathbf{X} \in \mathcal{X} \text{ for all } g \in \mathcal{G}.$$

The family $\mathcal{G}$ usually is a group of operators (where composition is the group's binary operation.)

2.4.1. Operators defined on signed sets

The generalized framework for determining ANN capabilities will be centered on invariants. The following operators defined on $\mathcal{X}$ will help formalize the invariants desired. Specifically, it is desired that the mappings that characterize signed sets will be invariant to dilation, translation, or rotation of the signed sets.

Definition 5. Let $\gamma \in \mathbb{R}^+$. Define $D_\gamma : \mathcal{P}(\mathbb{R}^d) \to \mathcal{P}(\mathbb{R}^d)$ as $D_\gamma(X) = \{\gamma x \in \mathbb{R}^d : x \in X\}$ for all $X \in \mathcal{P}(\mathbb{R}^d)$. Then, the dilation operator defined on signed sets $\mathbf{D}_\gamma : \mathcal{X} \to \mathcal{X}$ is defined as $\mathbf{D}_\gamma(\mathbf{X}) \doteq (D_\gamma(X^+), D_\gamma(X^-))$, for all $\mathbf{X} \in \mathcal{X}$.

Definition 6. Let $x_0 \in \mathbb{R}^d$. Define $T_{x_0} : \mathcal{P}(\mathbb{R}^d) \to \mathcal{P}(\mathbb{R}^d)$ as $T_{x_0}(X) = \{x_0 + x \in \mathbb{R}^d : x \in X\}$ for all $X \in \mathcal{P}(\mathbb{R}^d)$. Then, the translation operator defined on signed sets $\mathbf{T}_{x_0} : \mathcal{X} \to \mathcal{X}$, is defined as $\mathbf{T}_{x_0}(\mathbf{X}) \doteq (T_{x_0}(X^+), T_{x_0}(X^-))$, for all $\mathbf{X} \in \mathcal{X}$.

Definition 7. Let $\lambda \in \mathbb{R}^{d-1}$. Define $R_\lambda : \mathcal{P}(\mathbb{R}^d) \to \mathcal{P}(\mathbb{R}^d)$ as $R_\lambda(X) = \{r_\lambda(x) \in \mathbb{R}^d : x \in X\}$, for all $X \in \mathcal{P}(\mathbb{R}^d)$ where $r_\lambda : \mathbb{R}^d \to \mathbb{R}^d$ is a vector rotation operator that can be represented by an orthogonal matrix multiplication with the angle $\lambda = (\lambda_1, \lambda_2, ...\lambda_{d-1})$. Then, the rotation operator defined on signed sets $\mathbf{R}_\lambda : \mathcal{X} \to \mathcal{X}$ is defined as $\mathbf{R}_\lambda(\mathbf{X}) \doteq (R_\lambda(X^+), R_\lambda(X^-))$, for all $\mathbf{X} \in \mathcal{X}$.

Note that both $\mathbf{D}_\gamma$ and $\mathbf{R}_\lambda$ are linear operators. However, $\mathbf{T}_{x_0}$ is an affine operator.

Theorem 3. The quantifier, $q = \text{card}(\cdot)$, defined on $\mathcal{X}$ is invariant with respect to the group

$$\mathcal{G} = \{\mathbf{R}_\lambda : \lambda \in \mathbb{R}^{d-1}\} \cup \{\mathbf{D}_\gamma : \gamma \in \mathbb{R}^+\} \cup \{\mathbf{T}_{x_0} : x_0 \in \mathbb{R}^d\}.$$

Proof. We need to prove that $\text{card}(g(\cdot)) = \text{card}(\cdot)$ for all $g \in \mathcal{G}$. Let $g = T_{x_0}$ for some $x_0 \in \mathbb{R}^d$, and let $\mathbf{X} \in \mathcal{X}$, then

$$\begin{aligned} \text{card}(g(\mathbf{X})) &= \text{card}(\mathbf{T}_{x_0}(\mathbf{X})) \\ &= \text{card}(\mathbf{X}) \end{aligned}$$

Hence, $\text{card}(\mathbf{T}_{x_0}(\cdot)) = \text{card}(\cdot)$ for all $x_0 \in \mathbb{R}^d$. Similarly, $\text{card}(\mathbf{R}_\lambda(\cdot)) = \text{card}(\cdot)$ for all $\lambda \in \mathbb{R}^{d-1}$ and for all $\gamma \in \mathbb{R}^+$, $\text{card}(\mathbf{D}_\gamma(\cdot)) = \text{card}(\cdot)$. Therefore, cardinality is an invariant with respect to $\mathcal{G}$.

2.5. Topological-based Quantifiers

Cardinality is a quantifier that does not consider the geometric arrangement of the signed set. The geometric complexity[1,3] is a quantifier that considers the geometric arrangement, but unfortunately it is not continuous. To appreciate this statement, recall the distance between two sets X and Y in $\mathbb{R}^d$ is defined by

$$d(X,Y) = \max_{x \in X} \min_{y \in Y} \|x - y\|$$