

Proceedings of the IASTED International Conference



Artificial Intelligence and Soft Computing

May 27-30, 1998
Cancún, Mexico

Editor: M.H. Hamza

A Publication of
The International Association of Science and Technology
for Development - IASTED

ISBN: 0-88986-256-7
ISSN: 1482-7913

IASTED/ACTA Press
Anaheim ♦ Zürich

Proceedings of the IASTED International Conference

Artificial Intelligence and Soft Computing

**May 27-30, 1998
Cancún, Mexico**

Editor: M.H. Hamza

**A Publication of
The International Association of Science and Technology
for Development - IASTED**

**ISBN: 0-88986-256-7
ISSN: 1482-7913**

**IASTED/ACTA Press
Anaheim ♦ Calgary ♦ Zürich**

Table of Contents

Table of Contents	i	A Distributed Problem Solving Approach to Information Systems Modeling <i>M. Panti, A. Cucchiarelli, S. Valenti</i>	67
Expert Systems and Neural Networks	1	Generating and Understanding of Weak Information Structures by Humans <i>H. Maeda, T. Nishida</i>	74
ESDRD: An Intelligent Tool for Teaching Rheumatologic Skills <i>C. Rivas-Echeverria, F. Rivas-Echeverria</i>	3	Contextual Information Retrieval in Case-based Reasoning Systems: A Representational Scheme <i>A. Rodriguez</i>	78
Aspects of Temporal Usability via Constraint-based Reasoning: The use of Constraints as a Debugging Mechanism within General Descriptions of the User's Possible Actions <i>A. Ayes, G. Kelleher</i>	7	Performance of a Simple Cooperative Individual Situation Assessment (CISA) with Respect to Information Sharing Strategy Metrics <i>H. Hexmoor, E. Cuddihy, J. Llinas</i>	82
A Constructive Decision Boundary Modelling Algorithm <i>D. McLean, Z. Bandar, J.D. O'Shea</i>	11	On Improving Query Processing for Non-Recursive Queries in Indefinite Deductive Databases <i>H.D. Kim</i>	86
Flow Measurement Errors Estimation using Neural Networks <i>C. Leon de Mora, F. Biscarri Trivino, A. Menendez Martinez</i>	15	The Hybrid Method of TS and GAs for the Large-Scale Resource Allocation Problems <i>S.-A. Liu, X. Yin, B.-L. Zheng, M.-G. Wang</i>	90
Inductive Programming and a Multiversion Approach to Data Mining <i>D. Partridge, D. Tallis, W. Wang</i>	19	Approximate Reasoning in the Agricultural Context <i>M. Antonio de Oliveira, M. de Paiva Bastos Gottgroy</i>	94
Neural Control of a Non-Linear Plant: Fluid Mixer Case <i>R.M. de Brito, S. Baptistela, J. Barreto</i>	23	Linguistic Expression in Decision Making with Data Uncertainty <i>L.F. Sugianto, J. Kacprzyk</i>	98
Intelligent Systems	27	Validation of Algebraic Specifications <i>G.-H. Kwon, Y. Oh, Y. Chung</i>	102
A Learning Recommendation Agent in Virtual Environments <i>N. Zacharis, T. Panayiotopoulos</i>	29	Applications	107
Develop Fault Detection and Recovery System through Model Checking <i>W. Wen, J. Callahan</i>	22	Audio Word Spotting using Approximate String Match <i>W.Y. Wong, J. Robertson, C. Chung</i>	109
A Language and Architecture for Software Synthesis Control <i>R. Hewett</i>	37	Land Surface Classification of Landsat 5 TM Data using SOFM and GA <i>S.G. Lee, J.G. Han</i>	113
A Paradigm for Intelligent Decision and Control: Neural Nets, Fuzzy and Neuro-Fuzzy Implementations <i>R.V. Mayorga</i>	41	Optimal Planning of Energetic Operation using Genetic Algorithms <i>A.A.F.M. Carneiro, P.T. Leite, A.C.P.L.F. Carvalho</i> ..	117
A Personal Digital Assistance for Information Gathering <i>J.H. Lee, H.C. Kwon, K.J. Yoo, H.-S. Kim, B.N. Yoon</i>	45	Simulating of Ultrafast Nonlinear Soliton Logic for Optical Super Computing <i>V.N. Serkin, E.M. Schmidt, T. Belyaeva, O. Starostenko, G. Melo M.</i>	121
SOCRATES—An Integrated Intelligent System for Power System Control Center Operator Assistance and Training <i>Z.A. Vale, C. Ramos, A. Silva, L. Faria, J. Santos, M. Fernanda Fernandes, C. Rosado, A. Marques</i>	49	Acoustic Diagnosis of Electric Motors using Wavelet Transforms and Fuzzy Algorithms <i>G. Wirth, D.A. Mlynski</i>	125
GUI Design and Visual Programming <i>R.Y. Lee, S. Irish, N.C. Debnath</i>	52	Subsystem Testing in Automated Dynamic Code Verification <i>W.K. Chan, J. Lee, K. Simpson, J.-S. Sunwoo</i>	129
Information Systems	57	Prosody Modelling for Syllable-based Speech Synthesis <i>I. Kopecek, K. Pala</i>	134
Adaptive Retrieval of Reusable Software Components <i>A.A. Tzotsis, L. Jin</i>	59	Biologically Motivated Object Identification in a Multi-Context Scene <i>K. Ricanek, J.H. Kim, G.L. Lebbly</i>	138

Algorithms, Logic and Soft Computing	143	Nested Joint Probability Model for Morphological Analysis	
Suitable Semantics for Logic Programming		<i>K. Fujimoto, N. Inui, Y. Kotani</i>	211
<i>J. Arrazola, M. Osorio</i>	145	A Fuzzy Automaton Approach to Dialog Systems	
On Monotonic Extensions of Boolean Functions		<i>O. Unold</i>	215
<i>W. Dosch</i>	149	Cased-based Reasoning for Intelligent Question Answering	
"Soft" Approaches to Analogical Alignment and Sub-Graph Isomorphism		<i>U. Loerch, H.W. Guesgen</i>	219
<i>T. Veale</i>	153	Object-Oriented Systems	223
Toward a Joint Utilisation of Algebraic Specifications and Program Synthesis		Contravariance-based Class Identification for Integration of Class Hierarchies	
<i>L. Pierre, P. Manoury</i>	157	<i>M. Sugimori, K. Mizumachi, Y. Oda, S. Miyake, K. Morikawa, M. Nagata</i>	225
A Method of Automatic Program Designing and Code Generation using Informal Procedure Call Sentences		An Object-Oriented Approach to Managing Heterogeneous Pathology Information Systems	
<i>M. Osogami, F. Nishida</i>	161	<i>J.C. Sargis, W.A. Gray</i>	232
Parallel Architecture Application for Two Control Algorithms		Frameworks for Reengineering Legacy Software System into Design Pattern and Software Architecture	
<i>J. Kucharski, A. Skorek</i>	165	<i>H.-K. Kim</i>	238
Toward Hybrid Translation: Focusing on the Pattern-based Compound Unit Information in Syntactic Analysis		Modeling Spatio-Temporal Relationships in Object-Oriented Applications	
<i>H. Jung, S.K. Choi, C.-M. Sim, S. Yuh, T. Kim, D.-I. Park</i>	169	<i>I. Besembel-Carrera, J. Montilva-Calderon</i>	242
Computing a Qualitative Representation for the Local Space		Requirements and Comparison of Object Evolution Mechanisms for Object-Oriented Systems	
<i>M.E. Jefferies, W.-K. Yeap 173</i>		<i>S. Hammoudi</i>	246
Linguistics	177	Object Migration Behavior Modeling with Petri Nets	
A Method for Resolving Cohesive Relations of Unknown Words in Text Structure		<i>H. Sato, T. Hayashi</i>	250
<i>Y. Sato, K. Araki, K. Tochinai</i>	179	Priority-Ceiling based Concurrency Control of Object-Oriented Database	
Evaluation of Prediction Method of Target Words using Inductive Learning		<i>K. Asai, J. Nishibayashi, K. Yoshihara, M. Hagata ..</i>	254
<i>H. Sasaoka, K. Araki, Y. Momouchi, K. Tochinai</i>	183	Integrating AI and Object Technology for Manufacturing	
The Combination Analysis of Adjacent Nouns in a Korean Compound Noun		<i>S.M. Chan, T.L. Lammers</i>	261
<i>Y.S. Chae</i>	187	Design and Implementation of a User-Friendly Simulation Modelling Tool	
Clustering Korean Nouns based on Syntactic Relations and Corpus Data		<i>Y.-H. Wang</i>	265
<i>H.-J. Kim, J.-D. Park, M.-G. Jang, P.M. Ryu, D.-I. Park</i>	191	Neural Networks	269
A Probabilistic Dependency Parser for Korean Sentence using Case Frame Information		Estimating Neural Network Complexity	
<i>S.H. Kim, J.H. Shin, H.R. Park, P.M. Ryu</i>	195	<i>D. Opitz</i>	271
The Construction of Predicate Subcategorization using Tree Tagged Corpus		An Artificial Neural Network that Models Human Decision Making	
<i>P.-M. Ryu, M.-G. Jang, J.-D. Park, D.-I. Park, S.-H. Kim</i>	199	<i>T.-S. Quah</i>	275
278-104-An Implementation of English-to-Korean Machine Translation System for HTML Documents		Neural Network System based on Associative Memory to Perform Trajectory Generation and Inverse Kinematics	
<i>C.-M. Sim, H. Jung, S. Yuh, T. Kim, D.-I. Park, H.-C. Kwon</i>	203	<i>M. Vieira, A.F.R. Araujo</i>	279
Esoro/KE: A Korean-English Machine Translation System		Developing Artificial Neural Networks for Autonomous Agents using Evolutionary Programming	
<i>S. Yuh, H. Jung, Y. Kim, S.-K. Choi, T. Kim, D.-I. Park, J. Seo</i>	207	<i>J.M. Barreto, M. Roisenberg, F.M. de Azevedo</i>	283
		An Universal Neural Network Representation for Fuzzy-Controllers	
		<i>St. Niendieck, A. Tenhagen, W.-M. Lippe</i>	287
		GeNeSys: A Generic Neural System	
		<i>R. ...-Galan, A. Garcia-Tejedor, E. ...-Martin, A. Garcia-Crespo</i>	289

Algorithms, Logic and Soft Computing	143	Nested Joint Probability Model for Morphological Analysis	
Suitable Semantics for Logic Programming		<i>K. Fujimoto, N. Inui, Y. Kotani</i>	211
<i>J. Arrazola, M. Osorio</i>	145	A Fuzzy Automaton Approach to Dialog Systems	
On Monotonic Extensions of Boolean Functions		<i>O. Unold</i>	215
<i>W. Dosch</i>	149	Cased-based Reasoning for Intelligent Question Answering	
"Soft" Approaches to Analogical Alignment and Sub-Graph Isomorphism		<i>U. Loerch, H.W. Guesgen</i>	219
<i>T. Veale</i>	153	Object-Oriented Systems	223
Toward a Joint Utilisation of Algebraic Specifications and Program Synthesis		Contravariance-based Class Identification for Integration of Class Hierarchies	
<i>L. Pierre, P. Manoury</i>	157	<i>M. Sugimori, K. Mizumachi, Y. Oda, S. Miyake, K. Morikawa, M. Nagata</i>	225
A Method of Automatic Program Designing and Code Generation using Informal Procedure Call Sentences		An Object-Oriented Approach to Managing Heterogeneous Pathology Information Systems	
<i>M. Osogami, F. Nishida</i>	161	<i>J.C. Sargis, W.A. Gray</i>	232
Parallel Architecture Application for Two Control Algorithms		Frameworks for Reengineering Legacy Software System into Design Pattern and Software Architecture	
<i>J. Kucharski, A. Skorek</i>	165	<i>H.-K. Kim</i>	238
Toward Hybrid Translation: Focusing on the Pattern-based Compound Unit Information in Syntactic Analysis		Modeling Spatio-Temporal Relationships in Object-Oriented Applications	
<i>H. Jung, S.K. Choi, C.-M. Sim, S. Yuh, T. Kim, D.-I. Park</i>	169	<i>I. Besembel-Carrera, J. Montilva-Calderon</i>	242
Computing a Qualitative Representation for the Local Space		Requirements and Comparison of Object Evolution Mechanisms for Object-Oriented Systems	
<i>M.E. Jefferies, W.-K. Yeap</i>	173	<i>S. Hammoudi</i>	246
Linguistics	177	Object Migration Behavior Modeling with Petri Nets	
A Method for Resolving Cohesive Relations of Unknown Words in Text Structure		<i>H. Sato, T. Hayashi</i>	250
<i>Y. Sato, K. Araki, K. Tochinai</i>	179	Priority-Ceiling based Concurrency Control of Object-Oriented Database	
Evaluation of Prediction Method of Target Words using Inductive Learning		<i>K. Asai, J. Nishibayashi, K. Yoshihara, M. Hagata ..</i>	254
<i>H. Sasaoka, K. Araki, Y. Momouchi, K. Tochinai</i>	183	Integrating AI and Object Technology for Manufacturing	
The Combination Analysis of Adjacent Nouns in a Korean Compound Noun		<i>S.M. Chan, T.L. Lammers</i>	261
<i>Y.S. Chae</i>	187	Design and Implementation of a User-Friendly Simulation Modelling Tool	
Clustering Korean Nouns based on Syntactic Relations and Corpus Data		<i>Y.-H. Wang</i>	265
<i>H.-J. Kim, J.-D. Park, M.-G. Jang, P.M. Ryu, D.-I. Park</i>	191	Neural Networks	269
A Probabilistic Dependency Parser for Korean Sentence using Case Frame Information		Estimating Neural Network Complexity	
<i>S.H. Kim, J.H. Shin, H.R. Park, P.M. Ryu</i>	195	<i>D. Opitz</i>	271
The Construction of Predicate Subcategorization using Tree Tagged Corpus		An Artificial Neural Network that Models Human Decision Making	
<i>P.-M. Ryu, M.-G. Jang, J.-D. Park, D.-I. Park, S.-H. Kim</i>	199	<i>T.-S. Quah</i>	275
278-104-An Implementation of English-to-Korean Machine Translation System for HTML Documents		Neural Network System based on Associative Memory to Perform Trajectory Generation and Inverse Kinematics	
<i>C.-M. Sim, H. Jung, S. Yuh, T. Kim, D.-I. Park, H.-C. Kwon</i>	203	<i>M. Vieira, A.F.R. Araujo</i>	279
Esoro/KE: A Korean-English Machine Translation System		Developing Artificial Neural Networks for Autonomous Agents using Evolutionary Programming	
<i>S. Yuh, H. Jung, Y. Kim, S.-K. Choi, T. Kim, D.-I. Park, J. Seo</i>	207	<i>J.M. Barreto, M. Roisenberg, F.M. de Azevedo</i>	283
		An Universal Neural Network Representation for Fuzzy-Controllers	
		<i>St. Niendieck, A. Tenhagen, W.-M. Lippe</i>	287
		GeNeSys: A Generic Neural System	
		<i>R. Rodriguez-Galan, A. Garcia-Tejedor, E. Rodriguez-Martin, A. Garcia-Crespo</i>	289

Stable Adaptive Neural Network Controller with Learned Envelope Functions for Vibration Suppression <i>D. Martin</i>	293	Knowledge Systems	365
A Neural Solution for the Assignment Task <i>D. Benic, N. Sakic</i>	298	A Case-based Reasoning System involving a Quantized Method for Gait Disorder Diagnosis <i>O.A. Kuchar, Y.L. Lazukova, D. Riordan, J.L. Leahey</i>	367
A Biologically Fine-Grained Artificial Neural Network: Towards a Hybrid Model <i>J.L. Garcia Rosa, E. Francozo</i>	302	Expert System Approach to Diagnosing and Destroying Unknown Computer Viruses <i>N. Thanh Thuy, T.M. Nhat Quang</i>	371
Roles of Visual Nonlinear Processing in Neural Networks <i>N. Ishii, M. Nakamura, K. Yamauchi, H. Sasaki</i>	306	The Machine Intelligence based on MetaLingua <i>A.Sh. Abdoullaev</i>	375
Feature Selection and Classification for Diagnosing Breast Cancer <i>R. Nezafat, A. Tabesh, S. Akhavan, C. Lucas, M.A. Zia</i>	310	A Grammatical Model for Knowledge Representation in Artificial Neural Networks <i>P. Maani, N. Rashnavadi, A.R. Mirzai</i>	380
Neural Networks Applications	315	A Way of Multi-database Mining <i>N. Zhong, S. Yamashita</i>	384
Prediction of Survival in Trauma Patients using a Probabilistic Neural Network <i>W.J. Ketcherside, J.K. Blundell, V. Morse, J.P. Place</i>	317	Approaches to Validating the Discovered Knowledge from Databases <i>M.M. Owrang O.</i>	388
Pathological Discrimination of Larynx through Neural Network <i>M. de Oliveira Rosa, J.C. Pereira, S.A. Arujo, M. Grellet</i>	321	A Memory Model for Storing and Retrieving Mechanical Design <i>Y. Dandekar, I. Zeid</i>	394
An ANN based Music Score Recognizer <i>W.S. Wijesema, A.D.R. Sasanka, K. Hettiarachchi, S.D.I. De Silva, T.C.J. De Silva</i>	325	The Knowledge-based Simulation applied to the Waiting Phenomenons <i>E. Duran, R. Costaguta</i>	398
Adaptive Control of Biped Robot Locomotion using Multiple Models <i>R.K. Speer, W.E. Moore</i>	329	Algorithms for Modelling the Student's Solution Strategy in the DIATS <i>O.R. Joshua, D.H. Scuse</i>	402
Neural Network Architecture applied to Transmission Line Protection <i>D.V. Coury, M. Oleskovicz</i>	333	Learning	409
Financial Prediction using DOMM Manual Model Selection <i>M. Zhang, S. Xu</i>	337	Improving Representation Spaces using Genetic Algorithms <i>H. Vafaie, J. Bala</i>	411
Parallel Power System Analysis using Heterogeneous Neural Networks <i>V. Salami, M.J. Cooke, G.L. Lebbby</i>	341	Case-based Reasoning (CBR) and Machine Learning within the Aviation Domain <i>I. Dattani, R.V. Magaldi</i>	415
Improvement of Pattern Recognition in the Field of Mapping using Neural Networks <i>K. Hadjar, M. Khalfallah</i>	345	A Reactive Rule-based System for Trajectory Planning of a Mobile Robot with Low Sensory Resources <i>G. de A. Barreto, A.F.R. Araujo, M. de O. Rosa</i>	419
Pattern Classification using Parallel Neural Systems <i>M.C. King, G.L. Lebbby</i>	350	An Empirical Evaluation of Machine Learning Techniques for Automated Information Filtering <i>D. Opitz, R. Potluri</i>	423
Pattern Classification using a Minimum Volume Ellipsoidal Neural Network <i>G.L. Lebbby</i>	354	The Fusion of Supervised and Unsupervised Approaches to Image Segmentation <i>P. Pachowicz, S.W. Baik, J. Bala, A. Hadjarian</i>	427
A New Power Swing Blocking Principle for Distance Protection based on Neural Networks <i>E. Vazquez, O. Chacon, H.J. Altuve, M. Ramirez</i>	359	Improving the Performance of Differential Competitive Learning Model in Clustering Tasks <i>G. de A. Barreto, A.F.R. Araujo</i>	431
Power System Protection using Pattern Recognition Classifier <i>M. Ahmed, G.L. Lebbby</i>	363	A Reinforcement Learning based on Potential Field Methods to navigate in Initially Unknown Environments <i>A.F.R. Araujo, A.P.S. Braga</i>	435
		A Comparison of Functional Networks and Neural Networks <i>E. Castillo, J.M. Gutierrez</i>	439

The Generalized Autodistributy Functional Networks: Applications to Statistical Problems	
<i>E. Castillo, A. Cobo, E. Pruneda, J.M. Sarabia</i>	443
Universal Hierarchical Encoding Algorithm for Data Compression	
<i>J. Chung, Z. Bandar</i>	447
Fuzzy Systems	451
Fuzzy Classifier System: An Application for Fault Tolerance in Industrial Processes	
<i>M. Cerrada L., J. Aguilar C.</i>	453
High Level Object Reuse in Data Mining	
<i>S.H. Rubin</i>	457
Study on Price Forecasting and Decision-making Model System for Iron and Steel Enterprise	
<i>S.-A. Liu, Q. Wang, B.-L. Zheng, M.G. Wang</i>	461
Group Decision Making for the Selection Among CIM Systems using Fuzzy Concept	
<i>C. Kahraman, E. Tolga</i>	465
A New Approach to Learning Fuzzy If-Then Rules by Neural Networks	
<i>A. Sanei, E. Eslami, A.R. Mirzai</i>	469
Dealing with Impreciseness in Architectural Synthesis	
<i>C. Chantrapornchai, S. Tongsima, E.H.-M. Sha, S.X. Hu</i>	473
Adaptive Fuzzy Control System Design through TD Learning	
<i>C.-C. Wong, Y.-W. Ling</i>	477
Artificial Odor Discrimination System using Fuzzy Learning Vector Quantization Neural Network	
<i>B. Kusumoputro, H. Budiarto</i>	481
A Genetically Optimised Fuzzy Interface Algorithm	
<i>K.A. Crockett, Z. Bandar, A. Al-Attar</i>	485
A New Approach to Classical Backpropagation Algorithm for Neuro-Fuzzy-GA Systems Learning	
<i>L.M. Brasil, F.M. de Azevedo, J.M. Barreto, M. Noirhomme-Fraiture</i>	489
A Gradient Descent Learning Rule for Fuzzy Neural Networks	
<i>A. Klingebiel, A. Tenhagen, W.-M. Lippe</i>	493
Predicting Unemployment Rates using Fuzzy Time Series Model and Neural Network	
<i>C.-C. Chiu</i>	496
Author Index	501

IMPROVING THE PERFORMANCE OF DIFFERENTIAL COMPETITIVE LEARNING MODEL IN CLUSTERING TASKS

Guilherme de A. Barreto

Aluizio F. R. Araújo

University of São Paulo (USP), Department of Electrical Engineering

Av. Dr. Carlos Botelho, 1465

13560-970 São Carlos - SP, Brazil.

{gbarreto, aluizioa}@sel.eesc.sc.usp.br

ABSTRACT

In this paper we propose two neural algorithms that can be considered a simplification and a generalization of the Differential Competitive Learning (DCL) neural network, respectively. Firstly, we suggest some simplifications for the original DCL model to eliminate redundant aspects of the competition mechanism. We get rid of the lateral connections arguing that it is possible because the winning neuron is chosen based solely on metrical similarity measures and the lateral feedback weights play no effective role. The activation rule is made simpler requiring less computational effort. In the second model, we show how to combine lateral connections with metrical relations on the activation and the learning rules of DCL to effectively estimate cluster centroids. This model is also less sensitive to weight initialization. A number of simulations are carried out to compare the presented models in unsupervised clustering tasks.

Keywords: Neural networks, unsupervised learning, differential competitive learning, inhibition, clustering.

1. INTRODUCTION

Unsupervised artificial neural networks (UANN) models have offered new approaches to the solution of many pattern recognition tasks. Statistical pattern classification, cluster detection, vector quantization and probability density function estimation are some areas in which UANN have successfully been used [1]. In this paper, we are particularly concerned with cluster detection via centroid estimation.

Clustering a set of p patterns comprises finding m disjoint partitions so that the members of each partition are more similar to each other than to the remaining patterns. Unsupervised learning through competitive neural networks can be used for clustering by discovering the salient statistical features in the input [2], [3].

In competitive learning systems, a set of neurons compete among themselves for the right to respond to an input pattern. The winners of the competition are allowed to modify their weight vectors to become more similar to

the input patterns. In effect, competitive UANN models modify their weight vectors in an adaptive fashion in order to estimate the center of gravity of the clusters embedded in the input distribution [4]. Centroid estimation is useful in a number of applications, such as speech and image compression [5].

The goal is to minimize the mean square error E in finding m cluster centroids (w_1, w_2, \dots, w_m) for p patterns (x_1, x_2, \dots, x_p) [5], [6]:

$$E = \sum_{j=1}^m \left\{ \left(\frac{1}{p} \right) \sum_{i=1}^p M_{ji} \|x_i - w_j\|^2 \right\} \quad (1)$$

in which $M_{ji} = 1$ if x_i belongs to cluster m and $M_{ji} = 0$ otherwise.

In this paper, we introduce two competitive models for clustering based on Differential Competitive Learning (DCL) [7], [8]. The first model derives directly from DCL requiring less computational efforts. The second is less sensitive to weight initialization than the others two. In addition, both models can detect subtle aspects of the input pattern distribution dividing the input patterns in clusters with low clustering error.

The paper is organized as follows. In Section 2 we present the original DCL model and discuss some features of the model. In Section 3 a simplified version of the DCL is presented based on the discussion in Section 2. The previous two models have some limitations and in Section 4 we present a novel neural network model with the aim of solving some of the problems. Then, in Section 5 we carry out a number of simulations to illustrate the performance of the models in simple clustering tasks. Finally, in Section 6, we discuss the results of the simulations and the issues for further work.

2. DIFFERENTIAL COMPETITIVE LEARNING

Kong and Kosko [7] and Kosko [8], [9] have proposed the DCL model as a new unsupervised learning paradigm for adaptive vector quantization (AVQ). It has been used

in a number of applications such as density estimation and phoneme recognition [7] and printed circuit board inspection [10].

The DCL model comprises two layers of neurons (Figure 1). The input layer has n neurons and the output layer has m neurons. The input neurons propagate the observed inputs, through feedforward weight connections, to the output layer. In both layers, the neurons may have linear or non-linear (sigmoid) activations. Each neuron within the competition layer excites itself and inhibits the others through feedback inhibitory connections.

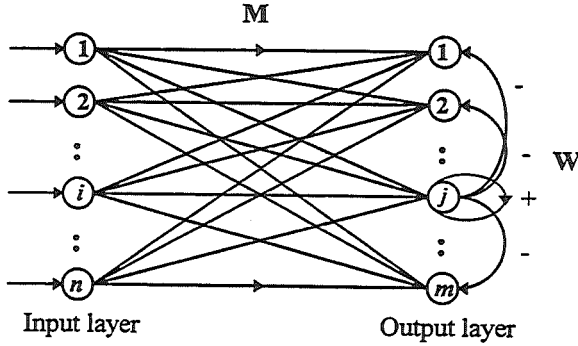


FIGURE 1. TOPOLOGY OF THE LATERALLY INHIBITORY DIFFERENTIAL COMPETITIVE LEARNING MODEL.

The DCL algorithm is summarized below:

- (1) Initialize the excitatory weight connections with m samples drawn from the input distribution and initialize the activations of the output neurons:

$$\mathbf{m}_j(0) = \mathbf{x}_j \text{ and } y_j(0) = 0, \quad j = 1, \dots, m$$

- (2) Present a randomly chosen sample $\mathbf{x}(t)$ to the net.
- (3) Find the closest ("winner") weight vector $\mathbf{m}_k(t)$ according to the Euclidean norm:

$$\|\mathbf{m}_k(t) - \mathbf{x}(t)\| = \min_j \|\mathbf{m}_j(t) - \mathbf{x}(t)\| \quad (2)$$

- (4) Update the activations of the output neurons through:

$$y_j(t+1) = y_j(t) + \sum_{i=1}^n S_i(x_i) m_{ij}(t) + \sum_{r=1}^m S_r(y_r) w_{rj}(t) \quad (3)$$

in which $S_i(x_i)$ and $S_r(y_r)$ are linear or sigmoid-type transfer functions. The inhibitory weight, w_{rj} , connects the output neuron r to the output neuron j .

- (5) Update the winning weight vector $\mathbf{m}_k(t)$ through:

$$\mathbf{m}_k(t+1) = \mathbf{m}_k(t) + \eta(t) \text{sgn}[\Delta y_k(t)] [\mathbf{x}(t) - \mathbf{m}_k(t)] \quad (4)$$

in which $0 < \eta(t) \ll 1$ is the learning rate and $\text{sgn}[\Delta y_k(t)] = \text{sgn}[y_k(t+1) - y_k(t)]$.

- (6) Repeat steps 2-5 continuously for the number of times t_{\max} specified by the user.

The learning process takes place only if the neuron activations change. If $\Delta y_k(t) > 0$, the winning weight vector $\mathbf{m}_k(t)$ moves towards the direction of the input $\mathbf{x}(t)$. If $\Delta y_k(t) < 0$, the winning weight vector is moved in the opposite direction, away from the input. This speed-velocity information provides unsupervised learning reinforcement during learning [7].

It is worthwhile to note that the winning neuron is chosen based on Euclidean distance measures. Another way to find the winner is via lateral feedback weights [1] by choosing the neuron with highest response for a given input pattern. The use of both mechanism is redundant and brings no special advantage to the model. In the next section, we suggest some simplifications to the DCL model in order to eliminate the redundancy in the competition mechanism.

3. THE SIMPLIFIED MODEL

The competition mechanism in DCL determines the winning neuron based on metrical similarity. Hence we argue that DCL does not effectively use its inhibitory connections. Hence, we get rid of the lateral connections introducing the Simplified Differential Competitive Learning (SDCL) neural algorithm. As a result, the activation rule (3) is made simpler and requires less computational efforts than DCL:

$$y_j(t+1) = y_j(t) + \sum_{i=1}^n S_i(x_i) m_{ij}(t) \quad (5)$$

The computational savings in using (5) instead of (3) are of particular interest when the $S_i(y_i)$ sigmoid are sigmoid functions. Sigmoids are computationally much heavier than the distance computations [11]. Moreover, simplicity and locality are the major constraints in designing a neural system. These are the obvious requirements if the neural system has to be hard-wired in analog or hybrid VLSI architectures. The absence of lateral connections simplifies VLSI implementation of artificial neural models substantially. The activation rule and the neural network operation require less computational effort and use only the information available locally at the neuron level.

The original DCL and its simplified version modify their feedforward excitatory weights according to the competitive learning rule [2]. Winner-take-all (WTA) behavior has the advantage of simplicity, but also has some drawbacks including underutilized output neurons (dead units), which reduces the representational ability

competitive neural networks [1], [3], [4]. In the next section, we propose a neural algorithm in order to show how Euclidean distance can be used together with lateral inhibition to correctly estimate cluster centroids.

4. THE GENERALIZED MODEL

This new model will be called Generalized DCL (GDCL). GDCL has the same topology as DCL. The inputs are non-linearly transformed according to:

$$a_j(t) = \exp\left(\frac{\|\mathbf{x}(t) - \mathbf{m}_j(t)\|^2}{2\rho^2}\right) \quad (5)$$

in which $\mathbf{x}(t)$ is the input vector, $\mathbf{m}_j(t)$ is the weight vector associated with the j th output and ρ is the selectivity parameter. This model has no "winner", that is, all output neurons are updated at each presentation of an input pattern. It is worth noting that simultaneous updating of all the neurons should accelerate the convergence of the network and helps to avoid dead units [1], [6].

Because of the new strategy for weight modification, the simple competitive learning rule can be no longer used. Hence, some kind of correlation-based learning [3], [4] rule must be used. The weights are adjusted through:

$$\mathbf{m}_j(t+1) = \mathbf{m}_j(t) + \eta^*(t)y(t)[\mathbf{x}(t) - \mathbf{m}_j(t)] \quad (6)$$

in which $\mathbf{m}_j(t)$ is the feedforward weight connection between input i and output neuron j , $\mathbf{x}(t)$ is the input vector, $y_j(t)$ is the output transfer function defined as:

$$y_j(t) = g\left(\sum_{r=1}^n w_{rj} a_r(t)\right) = g\left(\alpha a_j(t) - \beta \sum_{r \neq j} w_{rj} a_r(t)\right) \quad (7)$$

in which w_{rj} is connection from the output unit r to unit j ($w_{jj} = \alpha$ and $w_{rj} = -\beta$). The transfer function $g(v) = 1$ if $v > 1$, $g(v) = 0$ if $v < 0$ and $g(v) = v$ otherwise. The signals $y_j(t)$ are normalized so that $\sum_j y_j(t) = 1$. The extended learning rate $\eta^*(t)$ is defined as $\eta^*(t) = \eta(t) \text{sgn}[y(t+1) - y(t)]$.

In order to avoid different neurons to respond to the same input pattern, we suggest the use of fixed lateral inhibitory connections. It is worth emphasizing that here the lateral connections play an effective role. This is not the case for DCL. Computer simulations illustrate that the combination of these mechanism is able to improve the performance of DCL models in clustering tasks.

5. SIMULATION AND RESULTS

We simulated the models, written in ANSI C, in a SUN workstation ULTRA-1. For all tests $n = 2$, $m = 4$. In the first experiment, the pattern set consists of 2000 two-dimensional Gaussian-distributed pattern vectors with

standard deviation $\sigma = 0.09$, and with centroids or modes at $(1/3, 1/3)$, $(1/3, 2/3)$, $(2/3, 1/3)$ and $(2/3, 2/3)$.

Figure 1a shows the evolution of training for the SDCL model where $\eta(t) = 0.01$ and the initial weight vectors were $\{(0.48, 0.48), (0.52, 0.48), (0.52, 0.52), (0.48, 0.52)\}$. The SDCL weights estimated the centroids of the distributions correctly. The final weights after 5000 iterations were $\{(0.329, 0.332), (0.673, 0.328), (0.662, 0.662), (0.322, 0.659)\}$.

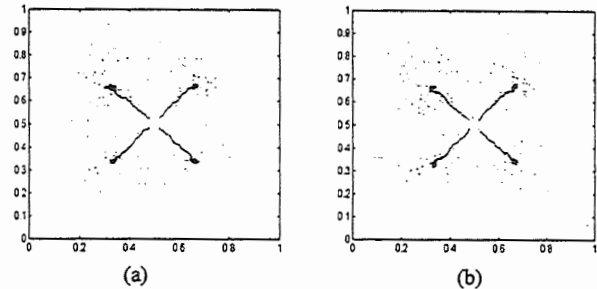


FIGURE 1. CENTROID CONVERGENCE OF SYNAPTIC VECTORS FOR: (a) SDCL AND (b) GDCL. THE INPUTS ARE LINEARLY TRANSUCED.

We repeat the previous test to evaluate the ability of the GDCL model in estimating centroids. The same training pattern set and initial weights of the previous experiment are used. For this model, $\eta(t) = 0.01$ and $\rho = 0.025$, $\beta = -2$ and $\alpha = 1$. Figure 1b shows the results of the simulation. After 5000 iterations the final weights were $\{(0.323, 0.330), (0.660, 0.337), (0.324, 0.660), (0.672, 0.666)\}$.

Figure 2 shows the trajectories of one of the SDCL and GDCL synaptic vectors reaching the centroid at $(2/3, 1/3)$.

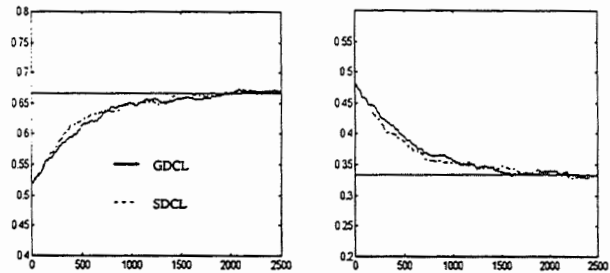


FIGURE 2. TRAJECTORIES OF ONE OF THE SDCL AND GDCL WEIGHT VECTORS REACHING THE CENTROID $(2/3, 1/3)$.

Both models have fast convergence, reaching the centroids after 1600 training iterations. The original DCL was unable to converge to the centroids for this pattern set.

In the third experiment, we evaluate the performance of the models on a simple clustering task [6]. For this test, the pattern set comprises 24 two-dimensional vectors:

$$\begin{aligned} \mathbf{v}_1 &= (0.18, 0.28) & \mathbf{v}_7 &= (0.23, 0.28) & \mathbf{v}_{13} &= (0.27, 0.20) & \mathbf{v}_{19} &= (0.45, 0.24) \\ \mathbf{v}_2 &= (0.19, 0.26) & \mathbf{v}_8 &= (0.24, 0.22) & \mathbf{v}_{14} &= (0.27, 0.18) & \mathbf{v}_{20} &= (0.47, 0.26) \\ \mathbf{v}_3 &= (0.20, 0.28) & \mathbf{v}_9 &= (0.23, 0.20) & \mathbf{v}_{15} &= (0.27, 0.16) & \mathbf{v}_{21} &= (0.47, 0.24) \\ \mathbf{v}_4 &= (0.18, 0.30) & \mathbf{v}_{10} &= (0.23, 0.18) & \mathbf{v}_{16} &= (0.34, 0.22) & \mathbf{v}_{22} &= (0.47, 0.22) \\ \mathbf{v}_5 &= (0.22, 0.26) & \mathbf{v}_{11} &= (0.25, 0.20) & \mathbf{v}_{17} &= (0.43, 0.26) & \mathbf{v}_{23} &= (0.47, 0.20) \\ \mathbf{v}_6 &= (0.22, 0.24) & \mathbf{v}_{12} &= (0.25, 0.20) & \mathbf{v}_{18} &= (0.44, 0.28) & \mathbf{v}_{24} &= (0.49, 0.22) \end{aligned}$$

In the trained networks, the patterns are grouped in classes according to the response of the output neurons.

The weight vector of each output neuron represents a cluster centroid. Each input pattern belongs to the cluster identified by the output neuron with the highest response.

For this experiment the learning rates are decreased according to: $\eta(t) = \eta_0(1 - t/t_{max})$ in which $\eta_0 = 1$ and $t_{max} = 2500$. The initial weights were chosen at random from the pattern set and $\beta = -0.9$ and $\alpha = 1.3$. We compare SDCL and GDCL with the WTA network [2]. The clusters found by SDCL and GDCL are $\{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$, $\{v_8, v_9, v_{10}, v_{11}, v_{12}, v_{13}, v_{14}, v_{15}\}$, $\{v_{16}\}$ and $\{v_{17}, v_{18}, v_{19}, v_{20}, v_{21}, v_{22}, v_{23}, v_{24}\}$ which can be visualized by the Voronoi tessellation [3] shown in Figure 3. GDCL has achieved this result in 40 out of 55 attempts, SDCL needs 29 out of 55 and WTA in only 13 out of 55 trials. In most cases, WTA has tried to shift the weight vectors to oversampled regions (Figure 4) finding the following clusters: $\{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$, $\{v_8, v_9, v_{10}, v_{11}, v_{12}, v_{13}, v_{14}, v_{15}\}$, $\{v_{16}, v_{17}, v_{18}\}$ and $\{v_{19}, v_{20}, v_{21}, v_{22}, v_{23}, v_{24}\}$.

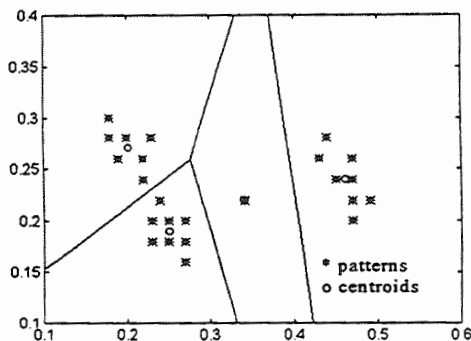


FIGURE 3. VORONOI DIAGRAM REPRESENTING CLUSTERING RESULTS FOR SDCL AND GDCL.

The result shown in Figure 3 suggests that SDCL and GDCL have paid a close attention to pattern v_{16} which is an outlier and assigned a weight vector to represent that.

It is important to note that the original DCL model was unable to cluster the input patterns for the training trials. In fact, this simulation suggests that the negative feedback through lateral connections in DCL generates some kind of instability.

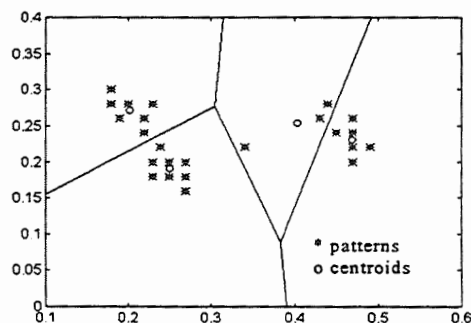


FIGURE 4. VORONOI DIAGRAM REPRESENTING CLUSTERING RESULTS FOR WTA.

This experiment shows that GDCL is less sensitive to weight initialization than SDCL. Both, GDCL and SDCL converge faster than WTA. For the clusters shown in

Figure 3 the final clustering error calculated through $\sum_{i=1}^n \|x_i - c_i\|^2$ is 0.00070 and for those in Figure 4 the error is 0.00087.

6 CONCLUSION AND FURTHER WORK

We have suggested a set of modifications to the original DCL model introducing the SDCL and GDCL models in order to enhance the performance in clustering tasks. In SDCL, we assessed the role played by inhibitory lateral weights in the competition mechanism. We got rid of the inhibition arguing that the use of both Euclidean distance relations and lateral connection is redundant to determining the winning neuron. In GDCL, we showed how to use Euclidean distances and inhibition effectively updating all the output neurons. The simulations illustrate that SDCL and GDCL are more stable than the original DCL always converging to cluster centroids. Moreover, the GDCL model is less sensitive to weight initialization than SDCL and WTA. SDCL and GDCL also achieve lower clustering error than WTA for certain input patterns.

Further work must be developed in order to simulate the SDCL and GDCL models on real world applications such as speech and image processing.

7 REFERENCES

- [1] S.J. Kia and G.G. Coghill, Unsupervised clustering as centroid estimation using dynamic competitive learning, *Biological Cybernetics*, 67, 1992, 433-443.
- [2] D.E. Rumelhart and D. Zipser, Feature discovery by competitive learning, *Cognitive Sci.*, 9, 1985, 75-112.
- [3] J. Hertz, A. Krogh and R.G. Palmer, *Introduction to the Theory of Neural Computation*. (Redwood City, CA: Addison-Wesley Publishing Company, 1991).
- [4] M.H. Hassoun, *Fundamentals of Artificial Neural Networks*. (A Bradford Book: MIT Press, 1995).
- [5] S.C. Ahalt, A.K. Krishnamurthy, P. Chen and D.E. Neuhoff, Competitive learning algorithms for vector quantization, *Neural Networks*, 3, 1990, 277-290.
- [6] S.J. Kia and G.G. Coghill, High performance clustering using dynamic competitive learning, *Electronics Letters*, 28(18), 1992, 1753-1755.
- [7] S.G. Kong and B. Kosko, Differential competitive learning for centroid estimation and phoneme recognition, *IEEE Trans. Neural Networks*, 2(1), 1991, 118-124.
- [8] B. Kosko, *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. New Jersey: Prentice Hall, 1992).
- [9] B. Kosko, Unsupervised learning in noise, *IEEE Transactions on Neural Networks*, 1(1), 1990, 44-57.
- [10] Madhav Moganti, PCB inspection using differential competitive learning and Fuzzy associative memory, *Master thesis*, University of Missouri-Rolla, May 94, 1994.
- [11] T. Kohonen, G. Barna and R. Chrisley, Statistical pattern recognition with neural networks: Benchmarking model, *Proceedings of the IEEE International Conference on Neural Networks*, San Diego, 1988.