

Proceedings

**1997 IEEE
INTERNATIONAL SYMPOSIUM
ON INFORMATION THEORY**

Maritim Hotel and Congress Center
Ulm, Germany

June 29 - July 4, 1997



Sponsored by the Information Theory Society
of the Institute of Electrical and
Electronics Engineers

**Proceedings
1997 IEEE
International Symposium on
Information Theory**

**Maritim Hotel and Congress Center
Ulm, Germany**

June 29 - July 4, 1997

*Sponsored by The Information Theory Society of
The Institute of Electrical and Electronics Engineers*

Proceedings 1997 IEEE International Symposium on Information Theory

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331. All rights reserved. Copyright ©1997 by the Institute of Electrical and Electronics Engineers, Inc.

IEEE Catalog Number:	97CH36074	
ISBN:	0-7803-3956-8	Softbound Edition
	0-7803-3957-6	Casebound Edition
	0-7803-3958-4	Microfiche Edition
Library of Congress:	72-179437	

Additional copies of this publication are available from

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08854-1331
1-800-678-IEEE

A Hierarchy of Hyperbolic Macrodynamic Equations as a Model for Network Training

Roderick V. Nicholas Melnik¹, Email: melnik@usq.edu.au

Dept. of Mathematics and Computing, University of Southern Queensland, Toowoomba 4350, QLD, Australia

Abstract — The author proposes mathematical models of hyperbolic type for training of neural networks, and its computational implementation using the Markov Chain approximation method.

I. INTRODUCTION

Let a mapping $x_t : \mathcal{R} \rightarrow \Sigma$ be a sigmoidal function (activator) coupled to a neural network defined by its neurons

$$x_t \circ \mu : T \otimes \Sigma \otimes U_T \rightarrow \mathcal{R}, \quad (1)$$

where $\mu : T \otimes \Sigma \otimes U_T \rightarrow \mathcal{R}$ will be referred to as the decision making function. In (1) we assume that the network is trained by a dynamic system with the state space Σ during time defined by a set T , and U_T is a set of all permissible training strategies. Let H be a mapping $X_T \rightarrow R^M$, and $H_T : \mathcal{N} \rightarrow \mathcal{N}$ is a T -computable function. Then H can be in principle arbitrarily well approximated by a network implementing H_T due to the Gödel numeration procedure. On the one hand if X_T is a compact Borel set and $H \in L^1(X_T)$ then for any arbitrary small $\epsilon > 0$ there exists a feedforward neural network (FNN) \tilde{H} such that $\|H - \tilde{H}\|_{L^1(X_T)} < \epsilon$. On the other hand there exists a block feedback network (BFN) \tilde{H}_T such that for any input $n \in M \subset \mathcal{N}$ a finite number of network steps produces $H(n)$ [2]. The main purpose of this paper is to show how both of the above approaches are connected by a training algorithm that can change weight values and the network structure.

II. NETWORKS TRAINING AS A PROBLEM IN OPTIMAL CONTROL

Let us choose as a function H the Hamiltonian of the dynamic system associated with the network defined above. Then the process of approximation of the function H can be seen as a construction of a training strategy for a network \tilde{H}_n^ϵ such that $\|H - \tilde{H}_n^\epsilon\|_{L^1(X_T)} \rightarrow \min$, where $H \in L^1(X_T)$, ϵ is an arbitrary small positive number, and n is the number of sigmoidal nodes of the network. Though we do not know *a-priori* the character of dependency of the network on parameters which determine the function μ , the author observes that the problem of constructing the mapping μ is intrinsically connected with the definition of dynamic rules in *singular stochastic control problems*. Mathematically speaking, we expect that if $\epsilon \rightarrow 0^+$ and $n \rightarrow \infty$ then $\tilde{H}_n^\epsilon \rightarrow H$, but to prove this conjecture rigorously we need a connection rule between ϵ , n , and the topology specified in X_T . Such a connection defines the system stability. Similarly, BFN can be applied to the approximation of H on an arbitrary set X_T , but questions of network dimension and architecture should be addressed [2]. Using this analogy it is proposed to analyse the network performance using a family of Discrete Markovian Decision processes. In the constructed hyperbolic mathematical model the training process

depends on Markov Chain parameters with possible discontinuities that are dependent on values of the sigmoidal function x_t . If $H \in L^1(X_T)$ the process of training is associated with the generalized energy equation [1]:

$$(1 + v_1) \left[\frac{\partial \mu}{\partial x} + \frac{1}{v_1} \left(\frac{\partial \mu}{\partial t} + f_0 \right) \right] = 0, \quad (2)$$

where v_1 is the velocity of information transmission between neurons, and f_0 is a training goal defined by *a-priori* knowledge on X_T and the function H . Algorithmically the function μ is required to be adapted to such *a-priori* given information.

III. NETWORKS AND MARKOV CHAINS

Many mathematical problems in optimal control and information theory is reducible to the modelling of the training process for networks. The main results of this paper is based on the idea of the Markovian character of the process of training (x_t, μ) . The author develops some previously obtained results for the Generalized Dynamic System (GDS) evolution [1]. Using conditional probabilities of the Markov Chain jumps the notion of the GDS velocity function between two macroscopic events is introduced as a measure of changes which take place on the microscopic level with respect to the macroscopic behaviour of the system. It is proposed to view a class of neural networks with built-in training ability as Infinite Length Perturbed Markov Chains (ILPMC). The practical implementation of training requires an approximation of the network activator by a *piecewise-deterministic stochastic process*. Such non-diffusion stochastic models have been previously studied by mathematicians in the theory of DMDP and recently attracted the attention of theoretical physicists. Since in practice information on the function H and topology of X_T are often subject to *a-priori* assumptions, the author proposes an approximation of the training process by an appropriately constructed Markov Chain. Stability of such approximation has been derived when the decision making function is modeled by (2).

IV. CONCLUSIONS.

This study proposed a hyperbolic mathematical model for the process of network training. An analogy of this process with the solution of optimal control problems has been revealed and new results in this direction have been obtained. A connection of trained neural networks with Perturbed Markov Chains allows us to derive computational models for the process of training. Finally, stability of the process has been discussed.

REFERENCES

- [1] V.N. Melnik, "Nonconservation law equation in mathematical modelling: aspects of approximation," *Proc. of the International Conference AEMC'96, Sydney*, pp. 423-430, 1996.
- [2] S. Santini and A. Del Bimbo, "Properties of block feedback neural networks," *Neural Networks*, Vol.8, No.4, pp. 579-596, 1995.

¹This work was supported by a grant from the CIAM, School of Mathematics at the University of South Australia

Table of Contents

Plenary Session I

The Way to the Proof of Fermat's Theorem — Frey, G.	1
--	---

Shannon Lecture I

Information Theoretic Methods in Probability and Statistics — Csiszár, I.	2
--	---

Plenary Session II

Practical Implementations of Cryptography — Landrock, P.	3
---	---

Shannon Lecture II

Back from Infinity: A Constrained Resources Approach to Information Theory — Ziv, J.	4
---	---

Plenary Session III

Source Coding in Multimedia — Rabiner, L.R.	5
--	---

Session MOAM1 Iterative Decoding

A General Algorithm for Distributing Information in a Graph — Aji, S.M. and McEliece, R.J.	6
Soft-Input Soft-Output Building Blocks to Construct and Iteratively Decode Code Networks — Benedetto, S. and Montorsi, G.	7
Serial Concatenated Trellis Coded Modulation with Iterative Decoding — Benedetto, S., Divsalar, D., Montorsi, G., and Pollara, F.	8
On the Performance of Recursive Decoding Schemes — Soljanin, E. and Urbanke, R.	9
Hybrid Concatenated Codes and Iterative Decoding — Divsalar, D. and Pollara, F.	10
Simultaneous Decoding and Phase Synchronization Using Iterative "Turbo" Decoding — Wiberg, N.	11
Approaching Shannon's Capacity Limit by 0.27 dB Using Hamming Codes in a 'Turbo'-Decoding Scheme — Nickl, H., Hagenauer, J., and Burkert, F.	12

Session MOAM2 Combinatorics and Sequences

Modularity Properties of Graph Entropy — Körner, J. and Simonyi, G.	13
Correlation Attacks on Stream Ciphers Using the Lempel-Ziv Algorithm — Penzhorn, W.T.	14
Pseudoredundant Measurements in Nonlinear Mappings — Arndt, C. and Loffeld, O.	15
On Higher Order Autocorrelation Properties of Golay Complementary Sequences — Braun, V.	16
Design of Sequences of I.I.D. p -ary Random Variables — Kohda, T. and Tsuneda, A.	17
Multiple Packing of Euclidean Sphere — Blinovskiy, V.	18
Enumeration of Rare Events, Pyramid and Ball Cubic Lattices — Krichevskii, R.E. and Potapov, V.N. ...	19
On Efficient Inversion in Tower Fields of Characteristic Two — Fan, J.L. and Paar, C.	20
Data-Punctuated Token Trees — Berlekamp, E. and Hekuis, P.	21

Session MOAM3 Multi-User Channels

Arbitrarily Varying Multiple-Access Channels Part I. Ericson's Symmetrizability is Adequate, Gallager's Conjecture is True. — Ahlswede, R. and Cai, N.	22
Arbitrarily Varying Multiple-Access Channels Part II. Correlated Sender's Side Information, Correlated Messages, and Ambiguous Transmission — Ahlswede, R. and Cai, N.	23
Approximating the Output of the Multiple Access Channel — Steinberg, Y.	24
A New Perspective to AWGN Multi-Access Channels — Shu, L.	25
Generalized Time Sharing for Multiple Access Channels — Rimoldi, B.	26
Optimal Power Allocation over Parallel Gaussian Broadcast Channels — Tse, D.N.	27

Session WEAM4 Soft Decoding

Asymptotic Error Performance for Reliability-Based Decoding of Binary Linear Codes — Fossorier, M.P.C. and Lin, S.	300
Not all Nearest Neighbors are Equal in Bounded Distance Decoding — Amrani, O. and Be'ery, Y.	301
An Efficient Information Set Algorithm for Soft-Decision Decoding of Linear Block Codes — de Barros, D.J. and Dorsch, B.G.	302
Soft Decision Decoding of BCH Codes Using Error Magnitudes — Reid III, W.J., Joiner, L.L., and Komo, J.J.	303
Soft-Decision Decoding Metrics for DAPSK — Chen, S.-I. and Fuja, T.	304
A New Evaluation Method of a Block Error Probability of a Block Modulation Code over an AWGN Channel — Tomita, K., Takata, T., and Kasami, T.	305
Generalized Minimum-Distance Decoding in Euclidean-Space: Performance Analysis — Agrawal, D. and Vardy, A.	306
Trellis-Based Decoders for Reed-Muller Type-II Hybrid ARQ Protocols — Martin, I., Honary, B., Markarian, G., and Wicker, S.	307
Probability Density Function of the Line-of-Sight Angle Error — Beth, Th., Grassl, M., Lazic, D.E., and Lukhaub, H.	308

Session WEAM5 Source Coding and Prediction

Redundancy of the Lempel-Ziv Codes — Savari, S.A.	309
On Pattern Frequency Occurrences in a Markovian Sequence — Régnier, M. and Szpankowski, W.	310
Average Profile of the Lempel-Ziv Scheme for a Markovian Source — Tang, J. and Szpankowski, W.	311
On the Distribution of Recurrence Times and the Exact Asymptotics of Lempel-Ziv Coding — Kontoyiannis, I.	312
Universal Coding and Universal Prediction — Suzuki, J.	313
Minimum Description Length Prediction of Next Symbol — Krichewski, R.E.	314
Asymptotic Minimax Regret for Data Compression, Gambling and Prediction — Xie, Q. and Barron, A.R.	315
Recurrence Times and Pointwise Lower Bounds in Data Compression — Algaet, P.	316

Session WEAM6 Network Access

Reliable Message Delivery via Unreliable Networks — Chan, A.H. and Chan, V.W.S.	317
Can Spatial Separation Increase Throughput of Conflict Resolution Algorithms? — Sayrafian-Pour, K. and Ephremides, A.	318
Iterative Multiuser Detection of a Multiple Access FSK System — Arnold, D.M., Schlegel, C., and Perez, L.C.	319
The Synchronization of Short Packets — Collins, O.M.	320
The Role of Information in Random Accessing — Lober, U.	321
A Hierarchy of Hyperbolic Macrodynamical Equations as a Model for Network Training — Melnik, R.V.N.	322
Evaluation of Petri Nets Compressibility — de Assis, F.M. and Barros, T.C.	323
Multicasting of Hierarchically Compressed Information in a Network — Kautia, A.	324

Chairmen

Daniel Costello, Bernhard Dorsch, Han Vinck

Finance:	Martin Bossert
Registration:	Martin Bossert, Han Vinck
Publications and Publicity:	Werner Henkel, Klaus Huber
Local Arrangements:	Jürgen Lindner
International Advisory Committee:	Rolf Johannesson

Acknowledgements

Grateful acknowledgement is given to the following organizations for their financial support of the 1997 IEEE ISIT:

AEG Mobile Communication, Ulm
Alcatel SEL, Stuttgart
City of Ulm
Daimler Benz Research Institute, Ulm
Daimler Benz Aerospace, Ulm
Deutsche Telekom AG, Ulm
E Plus Mobilfunk GmbH, Düsseldorf
Mannesmann Mobilfunk GmbH, Stuttgart
NASA Ames Research Center, Maffett Field, CA, USA
Siemens AG, Munich
Siemens AG, Ulm
The National Science Foundation, USA
Ulmer Universitätsgesellschaft
University of Ulm

Program Committee

Joachim Hagenauer (Chairman)

Henk van Tilborg (Chairman)

Rudolph Ahlswede

John Anderson

Ezio Biglieri

Martin Bossert

Pierre Chevillat

Daniel Costello

Ilya Dumer

Tony Ephremides

David Forney

Costas Georgiades

Christoph Günther

Te Sun Han

Kees Schouhamer Immink

Rolf Johannesson

Ralf Kötter

Torleiv Kløve

Shu Lin

Jim Massey

Ueli Maurer

Edward van der Meulen

Heinrich Meyr

Jim Modestino

Ben Smeets

Aimo Tietäväinen

Gottfried Ungerböck

Frans Willems

Ephraim Zehavi

Schriftliches Grußwort von Herrn Minister

Klaus von Trotha

für den Tagungsband „I S I T '97“ in Ulm

It is my pleasure to welcome the IEEE Information Theory Society to Baden-Württemberg for its International Symposium on Information Theory. It shows that the German scientists working in this field have major contributions to this growing area of science. The presence of so many well known researchers in the field will definitely stimulate our students to start or continue their studies in this direction of Information Sciences on a high level. The IEEE German Chapter on Information Theory has about 200 members, highlighting the importance of the Information Theory Society for the local scientific community. Germany has a great tradition in research, and Information theory plays an important role in the fields of Mathematics and Communications. Several publications by Claude Shannon have appeared in German technical journals. A well-known meeting point for Information theorists is the Oberwolfach conference center in the beautiful Black Forest, not far from Stuttgart. I am sure that many new contacts will be made at this Symposium, not only between scientists, but also between science and industry. The choice of the city of Ulm as venue is an excellent one from several points of view.

Ulm is a city dedicated to science with many research establishments and a young university. It has been chosen to be the city where research in multimedia applications will be concentrated in the near future. Ulm is in the middle of a number of very exciting attractions. I hope that you will take the opportunity to visit a few of these and that your stay will be a profitable one, both from a scientific as well as from a cultural point of view. I conclude using the famous Ulmer saying that it is „In Ulm, um Ulm, und um Ulm herum“, where the action is!



Klaus von Trotha

Minister für Wissenschaft, Forschung und Kunst
des Landes Baden-Württemberg