Lecture Notes in Computer Science 673

Gérard Cohen Teo Mora
Oscar Moreno (Eds.)

Applied Algebra, Algebraic Algorithms and Error-Correcting Codes

10th International Symposium, AAECC-10
San Juan de Puerto Rico, Puerto Rico, May 1993
Proceedings

Springer-Verlag

Lecture Notes in Computer Science 673
Edited by G. Goos and J. Hartmanis

Advisory Board: W. Brauer D. Gries J. Stoer

Volume Editors

Gérard Cohen
Ecole Nationale Supérieure des Télécommunications
46, rue Barrault, F-75634 Paris Cedex 13, France

Teo Mora
Università di Genova, Dipartimento di Matematica
Via L. B. Alberti 4, I-16132 Genova, Italy

Oscar Moreno
University of Puerto Rico, Department of Mathematics
Rio Pedras, Puerto Rico 00931

CR Subject Classification (1991): E.3-4, I.1, G.2, F.2

ISBN 3-540-56686-4 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-56686-4 Springer-Verlag New York Berlin Heidelberg

Typesetting: Camera ready by author/editor
Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
45/3140-543210 - Printed on acid-free paper

Preface

The AAECC Symposia Series was started ten years ago by Alain Poli (Toulouse), who organized, together with R. Desq, D. Lazard and P. Camion, the first conference in the series (Toulouse, June 1983) and was in charge of most of the following editions.

AAECC (the acronym has shifted its meaning over the years before stabilizing as "Applied Algebra, Algebraic Algorithms and Error Correcting Codes") aims to attract high-level research papers and to encourage cross-fertilization among different areas which share the use of algebraic methods and techniques for applications in the sciences of computing, communications, and engineering.

Algebra, in its broader sense, has always been viewed as a frame to describe in a formal setting both the properties of the objects giving mathematical models of reality and the rules under which they can be manipulated. Its importance for applications has grown in recent years with the introduction of technological areas (related to signal processing, error correcting codes, information processing, software engineering, etc.) in which the symbolic nature of the objects studied make the techniques of calculus and numerical analysis inapplicable. For these areas, algebra provides both a theoretical framework for the development of theories and algorithmic techniques for the concrete manipulation of objects.

While in principle covering any area related to applications of algebra to communication and computer sciences, by their previous history the AAECC Symposia are mainly devoted to research in coding theory and computer algebra.

The theory of error-correcting codes deals with the transmission of information in the presence of noise. Coding is the systematic use of redundancy in the formation of the messages to be sent so as to enable the recovery of the information present originally after it has been corrupted by (not too much) noise in the transmission over the channel. There has been a great deal of theoretical and applied work in this subject since the famous paper of Shannon in 1949. Applications of coding range from the lowly Hamming codes used in dynamic memories to the sophisticated Reed-Solomon codes used in compact disks and in many commercial and military systems. There are also convolutional codes widely used in satellite systems.

Computer algebra is devoted to the investigation of algorithms, computational methods, software systems and computer languages, oriented to scientific computations performed on exact and often symbolic data, by manipulating formal expressions by means of the algebraic rules they satisfy. It studies such problems from three different but confluent viewpoints: a) development and analysis of algebraic algorithms (both from the viewpoint of practical performance and of theoretical complexity); b) design and analysis of software systems for symbolic manipulation; c) applications of scientific and/or technological systems. It is important to stress that the mathematical theories to which computer algebra applies are not necessarily only the algebraic ones: polynomial equations, algebraic geometry, commutative algebra and group theory have a well-established research activity using symbolic computation techniques, but the same is equally

true for analytic theories, e.g. differential equations, as shown by a couple of papers in these proceedings. Computer algebra views algebra more as a method than as an object of research.

In the past, coding has interacted with group theory, combinatorics and finite geometries (the proof of the non-existence of a projective plane of order 10 by a coding approach is a recent example). More recently it has developed remarkable and unexpected connections with algebraic geometry and number theory (Goppa's algebraic geometric codes, Serre's improvement on Weil's bound for number of points of curves over finite fields, the p-adic Serre bound, improvements on Ax and Chevalley-Warning Theorems, etc.). This connection is creating links between the two major areas represented in AAECC, coding theory and computer algebra, e.g. by the use of Gröbner bases for decoding algebraic geometric codes or other algebraic codes.

Questions of complexity are naturally linked with the computational issues of both coding theory and computer algebra and represent an important share of the area which AAECC aims to cover; the same holds for cryptography where algebraic techniques are gaining relevance.

Finally let us mention the area of sequence design or spread spectrum multiple access, represented here by an invited contribution: originally developed in the Second World War for communications in a hostile enviroment where the enemy tries to jam one's message, it now includes non-military applications such as mobile radio, cellular telephony, and wireless computer communications.

Except for AAECC 1 (*Discrete Mathematics*, **56**,1985) and AAECC 7 (*Discrete Applied Mathematics*, **33**,1991), the proceedings of all the symposia are published in Springer *Lecture Notes in Computer Sciences*, Vols. 228, 229, 307, 356, 357, 508, 539.

It is a policy of AAECC to maintain a high scientific standard, comparable to that of a journal, and at the same time a fast publication of the proceedings. This is made possible only thanks to the cooperation of a large body of referees.

We aimed to have each submission evaluated by at least three referees, and we failed only in 9 cases. We had 147 independent reports from 105 referees on the 47 submissions. Of these, 6 were withdrawn during the procedure, 12 were rejected, 7 accepted for oral presentation only, 22 accepted for oral presentation and inclusion in the proceedings. The proceedings also contain six invited contributions; a seventh, by G. Lachaud, was not received in time for inclusion in the proceedings.

The conference was organized by the University of Puerto Rico and sponsored by the Army Research Office Cornell MSI project and by the NSF EPSCoR of Puerto Rico project.

We express our thanks to the staff of the Gauss Laboratory of the University of Puerto Rico and especially to Tita Santos, for handling the local organization, and to the Springer-Verlag staff and especially to A. Hofmann for their help in the preparation of these proceedings.

February 1993 G. Cohen, T. Mora, O. Moreno

Conference Board

Gerard Cohen (Paris), Teo Mora (Genova), Oscar Moreno (Puerto Rico)

Conference Committee

T. Beth (Karlsruhe), J. Calmet (Karlsruhe), G. Cohen (Paris), M. Giusti (Palaiseau), J. Heintz (Buenos Aires), H. Imai (Yokohama), H. Janwa (Bombay), R. Kohno (Yokohama), H. F. Mattson (Syracuse), A. Miola (Roma), T. Mora (Genova), O. Moreno (Puerto Rico), A. Poli (Toulouse), T. R. N. Rao (Lafayette, LA), S. Sakata (Toyohashi)

Referees

G. Attardi, T. Beth, E. Biglieri, D. Bini, M. Blaum, M. Bronstein, D. Bruschi, G. Butler, P. Camion, J.F. Canny, H. Chabanne, I. Chakravarti, A. Chan, P. Charpin, M. Clausen, A. Cohen, G. Cohen, D. Coppersmith, J. Davenport, J.L. Dornstetter, L.A. Dunning, A. Duval, D. Duval, T. Etzion, H.J. Fell, G.L. Feng, R. Froeberg, G. Gallo, Z. Ge, W. Geiselmann, M. Giusti, S. Goldwasser, D. Gordon, M. Goresky, R. Grossman, S. Harari, J. Heintz, T. Helleseth, H. Imai, H. Janwa, J. Justesen, M. Kalkbrener, E. Kaltofen, A. Kerber, N. Koblitz, R. Kohno, G. Lachaud, W. Lassner, D. Lazard, D. Lebrigand, S. Litsyn, A. Logar, I. Luengo, H.S. Maini, H.F. Mattson, H.M. Möller, T. Mora, R. Morelos-Zaragoza, O. Moreno, J. Moulin-Ollagnier, D. Mundici, H. Niederreiter, A.M. Odlyzko, F. Ollivier, G. Persiano, K.T. Phelps, V. Pless, A. Poli, C. Pomerance, T.R.N. Rao, T. Recio, J.J. Risler, L. Robbiano, F. Rodier, M.F. Roy, S. Sakata, R. Safavi-Naini, B.D. Saunders, W. Schmidt, R. Schoof, C. Scovel, J. Seberry, W. Seiler, N. Sendrier, A. Sgarro, K. Shirayanagi, M. Singer, P. Sole, H.J. Stetter, D.R. Stinson, B. Sturmfels, M. Sweedler, M. Szegedy, A. Tietavainen, H.C.A. van Tilborg, J. van Tilburg, L. Tolhuizen, C. Traverso, K. Tzeng, U. Vaccaro, F. Vatan, S. Vladut, J. Wolfmann, H. Yoshida, G. Zemor.

Contents

Sequence Based Methods for Data Transmission and Source Compression

A. R. Calderbank[1], P. C. Fishburn[1] and A. Rabinovich[2]

[1] Mathematical Sciences Research Center,
AT&T Bell Laboratories, Murray Hill, NJ 07974
[2] Statistics Department, Stanford University, Palo Alto, CA 94305

Abstract. In the last 10 years the invention of trellis coded modulation has revolutionized communication over bandlimited channels and is starting to be used in magnetic storage. Part of the reason is that sophisticated signal processing systems involving finite state machines can now be fabricated inexpensively. This paper discusses new developments in the performance analysis of finite state machines.
This is the extended abstract of an invited lecture to be given at the 10th International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes, Puerto Rico, May 10–14, 1993.

1 Introduction

This paper surveys recent work on sequence based methods for data transmission and source compression. We shall focus on new methods for analyzing the expected and worst-case performance of finite state machines. We suppose that every state transition in the finite state machine is associated with a symbol or set of symbols. The output of a finite state machine is then a set of symbol sequences or codewords. This set can be searched efficiently to find the optimum codeword with respect to any nonnegative measure that can be calculated on a symbol by symbol basis. The search algorithm is dynamic programming, that is to say the Viterbi algorithm.

The most familiar application of the Viterbi algorithm is in the decoding of channel outputs that have been corrupted by noise. A more recent application is the trellis coded quantization work of Marcellin and Fischer [14] where the measure is mean squared error (mse). We shall begin by describing a new graphical method for analyzing the covering properties of binary convolutional codes. This may be viewed as trellis coded quantization of a binary source, since the squared Euclidean distance $d^2(x, y)$ between two binary vectors x, y is just the Hamming distance $d_H(x, y)$. For complete details see [1, 2]. These are the first papers to define covering radius of a convolutional code and to describe a procedure for calculating this quantity.

The evolution of the Viterbi algorithm is determined by vectors of path metrics. The set of possible vectors forms the decoder state space. Bounds on the differences between path metrics are of practical importance in digital implementations of the Viterbi algorithm. Section 3 describes an example taken from

magnetic recording where it was possible to completely determine the decoder state space.

Section 4 considers the problem of finding the closest convolutional codeword to a sequence of source samples drawn from a uniform source on $[0, 1]$. The mean squared error per dimension can be interpreted as the second moment of a Voronoi region of an infinite lattice. This quantity is of importance in data transmission and vector quantization.

2 Graphical Analysis of the Covering Properties of Convolutional Codes

In this section we consider quantization of equiprobable binary data using a decoder for a binary convolutional code. Given an arbitrarily long binary sequence we wish to calculate the expected and worst-case Hamming distortion per dimension. This normalization gives the fraction of bits that we need to change in order to reach a codeword in the convolutional code.

The key observation is that a convolutional code with 2^ν states gives 2^ν approximations to a given source sequence and that these approximations do not differ very much. If we subtract the smallest path metric from the others then we obtain a vector with bounded entries. These vectors determine the evolution of the Viterbi algorithm. Hence the possible one-step trajectories of the Viterbi algorithm determine a finite directed graph on these vectors.

We shall briefly describe the new graphical method by means of a simple but representative example, the rate $1/2$ convolutional code C with generator matrix $[1 + D^2, 1 + D + D^2]$. The encoder state diagram is shown below in Fig. 1.

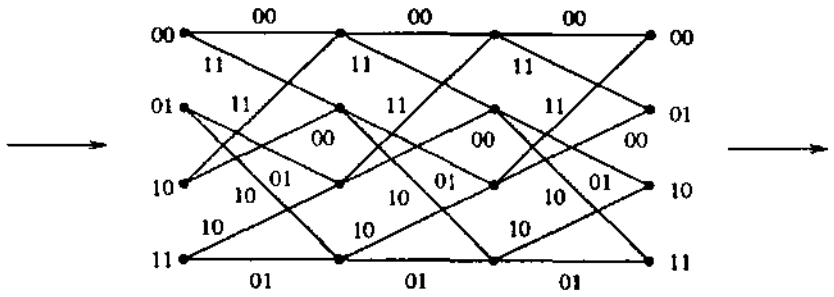

Fig. 1. Trellis diagram for the convolutional code with generator matrix $[1 + D^2, 1 + D + D^2]$.

The decoder has a copy of the trellis shown in Fig. 1. In every signaling interval, the decoder calculates and stores that path terminating in a given state that is closest to the binary data sequence. The decoder also calculates the *path metric* which measures distance from the data sequence to the codeword corresponding to the most likely path.