Selective Clinical Estimation of Childhood Abdominal Pain based on Pruned Artificial Neural Networks

DIMITRIOS MANTZARIS[1], GEORGE ANASTASSOPOULOS[1],
ADAM ADAMOPOULOS[2] , IOANNIS STEPHANAKIS[3],
KATERINA KAMBOURI[4] and STEFANOS GARDIKIS[4]

[1] Medical Informatics Laboratory, Democritus University of Thrace, GR-68100,
Alexandroupolis, GREECE

[2] Medical Physics Laboratory, Democritus University of Thrace, GR-68100,
Alexandroupolis, HELLAS

[3] HellenicTelecom. Organization (OTE), GR-151 24,
Athens, GREECE

[4] Pediatric Surgery Department, Democritus University of Thrace, GR-68100,
Alexandroupolis, GREECE

Abstract: The abdominal pain is a very common disease in childhood. Pediatric surgeons have to estimate at least 15 clinical and laboratory factors in order to make a diagnosis and decide about performing a surgical operation of the abdomen. Artificial Neural Networks (ANNs) architectures can be implemented in order to face the problem of abdominal pain and to assist surgeons in appendicitis prediction, avoiding an unnecessary operative treatment of child. There is a redundancy in the diagnostic factors, so genetic algorithms were used for pruning ANNs architecture, minimizing the number of diagnostic factors used during the training phase and therefore minimizing the number of nodes in the ANN input and hidden layer, and also for minimizing the Mean Square Error (MSE) of the trained ANN at the testing phase. The obtained results may support that a number of the diagnostic factors that are recorded in patient's history may be omitted with no compromise to the fidelity of clinical evaluation.

Key-Words: - Artificial Neural Networks, Genetic Algorithms, Non – symbolic artificial intelligence methods, pruning, Abdominal pain in childhood.

1 Introduction

The increase of computer power and the desire of using non symbolic methods for problems' solution were conduced to growth of Artificial Intelligence (AI) engineering. Artificial Neural Networks (ANNs), Fuzzy Logic and Genetic Algorithms are fields of Computational Intelligence.

ANNs have been successfully applied in a wide range of problems due to their capability to correlate input data to corresponding output data. Some fields of ANNs implementations are bioengineering [1], defense, aerospace, telecommunications [2], robotics [3], image processing [3-5], applied mathematics [6], financial analysis [7], microelectronics [8], intrusion detection systems and other fields [9]. In addition, ANNs have been proven as a powerful tool in medical prognosis and diagnosis [1, 10-15].

In ANN design, one important point is the development of appropriate ANN's architecture for solving specific problem. The selection of befitting ANN architecture is very important, as an ANN with few neurons would be unable to learn, while a big one leads to inefficient generalization ability, presenting overtraining [14]. The problem of ANN's size is critical in cases of hardware implementation of ANNs. However, a remarkable point in the ANN field is the construction of a structural and operational simplified ANN without any loss in terms of performance and functional ability.

Early works for the investigation of ANN pruning used trial and error. However, in the last decade some more efficient methods have been developed which are divided in two main different approaches. These are the constructive and the pruning (destructive) algorithms [16, 17]. A constructive algorithm starts from a minimal ANN and step-by-step generate a more and more complex

ANN structure by adding nodes in all the layers of the ANN. On the opposite, the pruning (destructive) algorithms start from a maximal ANN and delete unnecessary nodes, or even layers and synaptic connections during training.

In another approach to solve the problem Genetic Algorithm (GA) search was used. As GA has been proved a powerful optimization tool in the case that the search space is of great complexity and large size, they can be utilized for the investigation of the optimal ANN architecture that solves a specific problem [16 – 20]. Also, GA search was used for the detection of the optimal ANN architecture even in cases that the training phase of the ANN is bypassed [21 – 26].

The present work attempts ANN pruning in order to search for, and detect the essential smallest input data set of diagnostic factors necessary for ANN training. This is done by selecting some diagnostic factors to train ANN and by avoiding others. The reduced data set that can be used for ANN training in such a way that the performance of the prognostic ability is further improved. The selection of some diagnostic factors and the elimination of some others, and subsequently the decrease of the number of inputs of ANN are evolved by performing Genetic Algorithm (GA) search. The GA used is two-objective, thus the GA has to search for diagnostic data sets that at the same time: (a) minimize the number of diagnostic factors used during the training phase and therefore minimize the number of nodes in the ANN input and hidden layer, and (b) minimize the Mean Square Error of the trained ANN at the testing phase.

The proposed method was used for determination of the minimum and therefore essential diagnosis factors for abdominal pain in childhood clinical evaluation. The abdominal pain diagnosis except of the traditional methods (clinical, laboratory, and medical imaging) could also be supported by fuzzy logic techniques, numerically scoring systems, etc [27].

The obtained results proved that the medical diagnostic factors recorder by the pediatric surgeons present a high level of redundancy and overlapping. Therefore some of these factors proved essential for clinical evaluation and prognosis, whereas, some other factors could be omitted and totally neglected during clinical evaluation with no loss in prognosis ability. From a technical point of view, the detection of the essential diagnostic factors could support the utilization of smaller and simpler data sets for ANN training and therefore could be useful for the design and use of ANN with simpler architecture and improved performance.

2 Abdominal Pain Disease and its Diagnostic Factors

Appendicitis is the most common surgical condition of the abdomen. Diagnosis and treatment have certainly improved during the last years but appendicitis still continues to cause significant morbidity and remains, although rarely, a cause of death. The mortality rate for children with appendicitis ranges from 0.1-1% [28].

Appendicitis occurs with a male-to-female ratio of 3:2 with a peak incidence between the ages 12 and 18 years. The mean age of a patient is 6-10 years. The lifetime risk is 8.6% for boys and 6.7% for girls [28]. The role of race, ethnicity, health insurance, education, access to healthcare, and economic status on the development and treatment of appendicitis are widely debated [29]. A genetic predisposition holds in some cases, particularly for children who develop appendicitis before the age of 6 years. Although the disorder is uncommon in infants and elderly, these groups have a disproportionate number of compilations because of delays in diagnoses and the presence of comorbid conditions.

Many terms have been used to describe the varying stages of appendicitis, including acute focal appendicitis, acute supurative appendicitis, gangrenous appendicitis and perforated appendicitis. These distinctions are vague, and only the clinically relevant distinction of perforated (gangrenous appendicitis includes into this entity as dead intestine functionally acts as a perforation) versus non-perforated appendicitis (acute focal and supurative appendicitis) should be made.

Acute Appendicitis (AA) has a variability of clinical presentation. Clinical experience and technology advances in diagnostic methods are not foolproof. A typical scenario of acute appendicitis manifestation is the follow: The child describes some mild gastrointestinal symptoms before the onset of pain, such as indigestion or "gastritis". Typical in an early stage the pain is non specific in the epigastric or umbilical region and after few hours (4-6 hrs) it becomes localized at the lower right quadrant (LRQ) of the abdomen over the appendix. Anorexia, nausea and vomiting regularly follow the onset of pain within few hours. Localized tenderness in the LRQ (McBurney’s point) is essential for the diagnosis. Other signs suggesting acute appendicitis include: psoas muscle sign, obturator muscle sign, Rosving’s sign and rebound tenderness. As the disease progresses to perforation, peritonitis ensues. Perforation may result in temporary relief of symptoms as the pain of the distended appendix is relieved. Initially, peritonitis is reflected as local muscular rigidity. This progresses from simple voluntary guarding to generalized rigidity of the abdomen. The appendix commonly, but by no means always, ruptures about 24 to 48 hours after the onset of symptoms.

Leucocytosis (11,000 to 16,000/ mm3), with an increase in neutrophil count, has been considered to be significance in patients with AA. Greater specifity and sensitivity for diagnosis of AA is a neutrophil-lymphocyte ratio greater than 3.5. Positive values for C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) are useful, but negative values do not rule out the disorder. Combinations of all these tests may be the most helpful, and it found that normal total leukocyte count, neutrophil percentage, and normal CRP level correctly ruled out AA with 100% accuracy. Urinalysis is useful for detecting urinary tract disease; normal findings on urinalysis are of limited diagnostic value for appendicitis [30].

The appendicitis diagnosis is based on 15 factors presented in Table 1. This study attempts to distinguish a subset of appendicitis medical factors, as inputs of a pruned ANN, eliminating redundant factors without any compromise in terms of appendicitis prognosis.

This study used a data set consisted of 516 cases, whereof 422 (81.78%) normal and 94 (18.22%) underwent operative treatment. The patients' records are obtained from Pediatric Surgery Clinical Information System of the University Hospital of Alexandroupolis, Greece.

3 Methods

The proposed method for elimination of abdominal pain in childhood diagnostic factors is based on GA search for the optimal combination of the essential diagnostic factors necessary for ANN training.

Among the great variety of ANN topologies and architectures and the related learning algorithms that have been proposed for ANN training, in the present work there were be used feed-forward ANN and the error-backpropagation learning algorithm [31].

GAs are adaptive, robust, efficient, stochastic and global search algorithms based on the mechanics of natural selection and natural genetics [31-33]. They are especially useful for complex optimization problems where the number of parameters is large and the analytical solutions are difficult to obtain.

Since the main issue in the present work is the detection and elimination of redundant and overlapping diagnostic factors and data that are kept in the patient's history, a GA search was performed in order to attain the most suitable subset of diagnostic factors that can be used for ANN training.

	Diagnostic Factors
1	Sex
2	Age
3	Religion
4	Demographic data
5	Duration of Pain
6	Vomitus
7	Diarrhea
8	Anorexia
9	Tenderness
10	Rebound
11	Leucocytosis
12	Neutrophilia
13	Urinalysis
14	Temperature
15	Constipation

Table 1. Abdominal pain clinical and laboratorial factors.

As it was shown in Table 1, the number of diagnostic factors is $N = 15$, so the possible combinations of input data subsets are given by $N!/k!(N-k)!$, for $k = 1 .. 15$, as presented in Fig. 1.

The total number of combinations is 32767, so exhaustive search is time-consuming and therefore the use of advanced optimization searching techniques, such as GA search, is required.

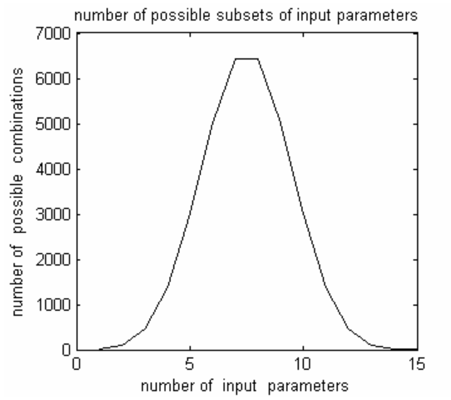

Fig. 1 Number of possible diagnostic factors combinations

For that purpose the population of the GA consisted of binary strings of 15 bits. Therefore, the chromosome's length of each GA individual is equal to the total number of the abdominal pain diagnostic factors presented in Table 1. In order the GA to perform the two-objective optimization task that is desirable, concerning the minimization of the input data set that is used for ANN training and at the same time the minimization of the MSE generated from the ANN during the testing phase the following fitness function was used:

$$f = MSE + I/N \tag{1}$$

where, I is the number of input nodes of the ANN and N is the total number of diagnostic factors in the original, full-sized training data set.

All the well-known binary GA operators for crossover and mutation can be used, that is the single-point crossover, two-point crossover, uniform crossover, scattered crossover, as well as, gaussian mutation and uniform mutation).

For each experiment, the GA was left to run for a sufficient number of generations and the fitness values of the GA individuals as well as the average fitness value of each generation were recorded. In addition, the diagnostic factors that were considered for ANN training were kept.

For each GA execution, the selected (pruned) input data set was divided in a training group of patterns (consisting of the 80% of the total cases) and a testing group of patterns (consisting of the rest 20% of the total cases).

4. Results

In order to obtain results for the ANN performance in training and testing for the full-sized data set that was consisted from all the 15 recorded diagnostic factors presented in Table 1, as a first step, an extensive investigation was performed, that included 10 computer experiments of ANN training and testing for a variety of values of the training epochs parameter, ranging from 50 up to 5000. The obtained results are presented in Table 2, where for each value of the number of training epochs the mean value and the standard deviation of MSE over 10 computer experiments are recorded.

The corresponding results that were obtained by executing the GA are presented in Table 3. For the same values of training epochs as in Table 2 there were recorded the fitness of the optimal individual found (in column 2), the MSE of the optimal individual (in column 3), the number of independent diagnostic factors that were used as an input for ANN training (in column 4), and explicitly, which diagnostic factors were used (in column 5).

As it is shown in column 4 of Table 3 the GA managed to converge to ANN that used 1 to 4 at most over the 15 diagnostic factors, or in other words the 6.67% up to 26.67% of the diagnostic inputs available. At the same time, the ANN that were trained with the pruned input data sets exhibited improved diagnostic ability, since their MSE at the testing phase is significantly decreased. This is more clearly shown in Figure 2, that the mean MSE of the full-trained ANN, as well as, the MSE of the ANN of the genetically evolved and pruned input data sets are plotted against the number of training epochs.

Training epochs	**mean MSE**	**Standard deviation of MSE**
50	0.02411304	0.00799345
100	0.02090588	0.01027768
200	0.02336139	0.00717854
300	0.02687388	0.01195818
400	0.03071900	0.01089655
500	0.02632005	0.00464043
1000	0.03141748	0.00889751
2000	0.03271966	0.01139088
5000	0.04471958	0.00959487

Table 2. Mean MSE and standard deviation of MSE on the test set over 10 full-sized experiments

As it is shown in Figure 2, for all the cases of the various numbers of training epochs the genetically pruned input data sets managed to train the ANN so that they exhibited smaller MSE that the full-trained ANN. The observed decrease in the genetically trained ANN is of the order of 15% up to 72%, depending on the number of training epochs. However, most important is the fact that in column 5 of Table 3 are presented explicitly which diagnostic factors were used as inputs for pruned ANN training. By looking over these results it is clearly that some parameters such as 3 (Religion), 9 (Tenderness), 11 (Leucocytosis) and 12 (Neutrophilia), seem to be more effective on ANN training and testing.

5. Discussion

Despite the fact that abdominal pain is one of the most frequently diseases in childhood, with mortality ranging from 0.1% to 1%, no investigation has been done on the effectiveness and weightiness of the various diagnostic factors in the clinical evaluation of the patients. A method for doing that is proposed in the present work. The method presented here utilized a specific GA to evolve the subsets of clinical data that are used as inputs for ANN training and testing. After adequate steps of genetic evolution, the GA converged to diagnostic factors subsets that were consisted from 1 up to 4 over a total of 15 diagnostic factors that are presented in Table 1. As it was found, the ANN that were trained using the pruned input data subsets presented in Table 3 overperformed the full-trained ANN in terms of the generated MSE on the test phase. Therefore the proposed methodology unveiled that by using specifically selected diagnostic factors instead of all of them, or in other words by using selective (pruned) training instead of full training of the ANN resulted to increase ANN performance and prognostic ability while at the same time the training procedure was considerably speed up. This stands because as it is indicated in Fig. 2, the MSE is decreased up to 72% while the size of the training input data is decreased up to 73.33 %. From this point of view, the obtained results may support that a number of the recorder diagnostic factors that are recorded in patient's history may be omitted with no compromise to the fidelity of clinical evaluation.

Training epochs	Fitness function	MSE on test set pruned inputs	Number of diagnostic parameters used	diagnostic parameters used
50	0.19844195	0.01493072	2	3 12
100	0.14202634	0.01702634	2	11 12
200	0.19833713	0.01083713	2	3 13
300	0.19489748	0.00739748	2	13 14
400	0.20321350	0.01571350	2	9 14
500	0.13497545	0.00997545	1	13
1000	0.32613259	0.01363259	4	3 7 10 12
2000	0.13946570	0.01446570	1	12
5000	0.20042594	0.01292594	2	9 11

Table 3. MSE of neural network trained with pruned sets of diagnostic input parameters

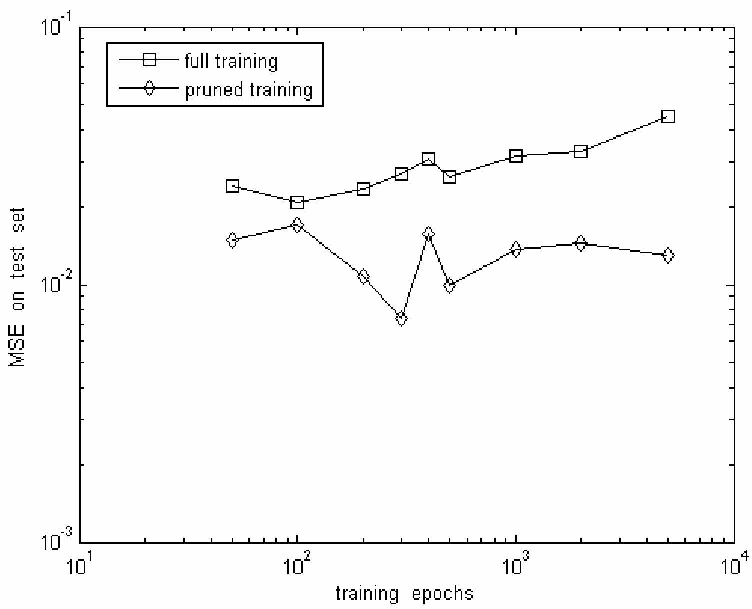

Fig. 2 Comparison of MSE on the test set among ANN with full training and ANN with pruned training.

References

[1] R. Nayak, L. Jain and B. Ting, Artificial Neural Networks in Biomedical Engineering: A Review, *Proc. 1st Asian-Pacific Congr. Computational Mechanics*, 2001, pp. 887-892.

[2] J.J. Murillo-Fuentes, and F.J. González-Serrano, Applying GCMAC to Predistortion in GSM Base Stations, *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'99),* Vol. V, 1999, pp. 2575-2578.

[3] L. Carnimeo, Target Recognition in Mobile Robot Vision Systems via Neural Networks with Local Interconnections, *Proceedings of the 3rd WSEAS Int.Conf. on Neural Network and Applications*, 2002, pp.107-111.

[4] M. Hitam, Y. Muslan, M. Deris, and M. Saman, Image Texture Classification Using Projection and Neural Network, *WSEAS Transactions on Systems*, Vol. 3, No. 4, 2004, pp. 1825- 1830.

[5] H. Rowley, S. Baluja, and T. Kanade, Neural Network-Based Face Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 1, 1998, pp. 23 – 38.

[6] M. Rao, and K. Ravindran, An Evolutionary Approach for Solving Differential Equations, *WSEAS Transactions on Circuits*, Vol. 1, Issue 1, 2002, pp. 100 – 105.

[7] G.M. Papadourakis, G. Spanoudakis, and A. Gotsiass, Application of Neural Networks in Short-Term Stock Price Forecasting, *Proceedings of 1st International Workshop Neural Networks in the Capital Markets*, 1993.

[8] V. Varshavsky, V. Marakhovsky, and I. Levin, Artificial Neurons Based on CMOS β-Driven Threshold Elements with Functional Inputs, *WSEAS Transactions on Systems*, Issue 2, Volume 3, 2004, pp. 442 – 448.

[9] D. Coury, and E. Segatto, Pattern Recognition to Distinguish Magnetizing Inrush from Internal Faults in Power Transformers, *WSEAS Transactions on Systems*, Vol. 3, 2004, pp. 1258-1264.

[10] M. Keogan, J. Lo, K. Freed, V. Raptopoulos, S. Blake, I. Kamel, K. Weisinger, M. Rosen, and R. Nelson, Outcome Analysis of Patients with Acute Pancreatitis by Using an Artificial Neural Network, *Academic Radiology*, Vol. 9, No. 4, 2002, pp. 410-419.

[11] K. Papik, B. Molnar, R. Schaefer, Z. Dombovari, Z. Tulassay, and J. Feher, Application of Neural Networks in Medicine - A Review, *Med Sci Monit*, 1998, pp. 538-546.

[12] T. Anagnostou, M. Remzi, and B. Djavan, Artificial Neural Networks for Decision-Making in Urologic Oncology, *Reviews in Urology*, Vol. 5, No. 1, 2003, pp.15-21.

[13] D. Tasoulis, P. Spyridonos, N. Pavlidis, D. Cavouras, P. Ravazoula, G. Nikiforidis, and M. Vrahatis, Urinary Bladder Tumor Grade Diagnosis Using On-line Trained Neural Networks, *Proc. Knowl. Based Intell. Inform. Eng. Syst., Springer-Verlag*, 2003, pp. 199-206.

[14] D. Mantzaris, G. Anastassopoulos, A. Tsalkidis, and A. Adamopoulos, Intelligent Prediction of Vesicoureteral Reflux Disease, *WSEAS Transactions on Systems*, Vol. 4, Issue 9, 2005, pp. 1440-1449.

[15] M. Munley, J. Lo, G. Sibley, G. Bentel, M. Anscher and L. Marks, A Neural Network To Predict Symptomatic Lung Injury, *Phys. Med. Biol.*, Vol. 44, 1999, pp. 2241–2249.

[16] X. Yao X. and Y. Liu, A New Evolutionary System for Evolving Artificial Neural Networks, *IEEE Transactions on Neural Networks*, Vol. 8, No. 3, 1997.

[17] X. Yao, Evolving Artificial Neural Networks, *Proceedings of the IEEE*, Vol. 87, No. 9, 1999, pp. 1423-1447.

[18] D. Whitley and C. Bogart, The evolution of connectivity: Pruning neural networks using genetic algorithms, *International Joint Conference on Neural Networks*, Washington D.C., 1. Hillsdale, NJ: Lawpence Erlbaum, 1990, pp. 134-137.

[19] B. Happel et al., Design and evolution of modular neural network architectures, *Neural Networks*, Vol. 7, 1994, pp. 985 – 1004,.

[20] S. A. Billings and G. L. Zheng, Radial basis function network configuration using genetic algorithms, *Neural Networks*, Vol. 8, 1995, pp. 877-890,.

[21] L. Davis, Mapping classifier systems into neural networks. *Proceedings of the 1988 Conference on Neural Information Processing Systems*, Morgan Kaufmann, 1988.

[22] D. Whitley, Applying genetic algorithms to neural network problems, *International Neural Networks Society*, 1988, pp. 230.

[23] D. Montana and L. Davis, *Training feedforward neural networks using genetic algorithms,* BBN Systems and Technologies, Cambridge, MA 1989.

[24] A. Adamopoulos, E. Georgopoulos, G. Manioudakis and S. Likothanassis, An Evolutionary Method for System Structure Identification Using Neural Networks, *Neural Computation '98*, 1998.

[25] E.F. Georgopoulos, S.D. Likothanassis, A.V. Adamopoulos, Evolving artificial neural networks using genetic algorithms, *Neural Network World*, Vol.4, 2000, pp.565-574.

[26] N. García-Pedrajas, C. Hervás-Martínez, and J. Muñoz-Pérez, COVNET: A Cooperative Coevolutionary Model for Evolving Artificial Neural Networks, *IEEE Transactions on Neural Networks*, Vol. 14, No. 3, 2003, pp. 575-596.

[27] M. Blazadonakis, V. Moustakis, and G. Charissis, Deep Assessment of Machine Learning Techniques Using Patient Treatment in Acute Abdominal Pain in Children, *Artificial Intelligence in Medicine*, 1996, Vol. 8, pp. 527- 542.

[28] D.G. Addiss, N. Shaffer, B.S. Fowler, and R.V. Tauxe, The Epidemiology of Appendicitis and Appendectomy in the United States, *Am J Epidemiol*, Vol. 132, 1990, pp. 910-925.

[29] B.C. Nwomeh, D.J. Chisolm, D.A. Caniano, and K.J. Kelleher Racial and Socioeconomic Disparity in Perforated Appendicitis Among Children: Where Is The Problem?, *Pediatrics*, Vol. 117, 2006, pp. 870-875.

[30] S. Gardikis, S. Touloupidis, G. Dimitriadis, C. Limas, S. Antypas, T. Dolatzas, A. Polychronidis, C. Simopoulos, Urological Symptoms of Acute Appendicitis in Childhood and Early Adolescence, *Int Urol Nephrol.*, Vol. 34, 2002, pp. 189-192.

[31] D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence. Piscataway*, NJ: IEEE Press, 1995.

[32] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning. Reading*, MA: Addison-Wesley, 1989.

[33] M. Mitchell, *An Introduction to Genetic Algorithms*, Cambridge, MA: MIT Press, 1996.