



The South African Institute of Computer Scientists  
and  
Information Technologists

**Proceedings**

**of the**

**1996 National Research and  
Development Conference**

**Industry meets Academia**

Interaction Conference Centre, University of Natal,  
Durban .  
26 & 27 September

**Edited by  
Vevek Ram**

©1996 Copyrights reside with the original authors who may be contacted directly

**ISBN 0-620-20568-7**

Cover printed by Natal Printers (Pty) Ltd, Pietermaritzburg

Copying by the Multicopy Centre, University of Natal, Pietermaritzburg

Binding by Library Technical Services, University of Natal, Pietermaritzburg

The views expressed in this book are those of the individual authors

## **FOREWORD**

This book is a collection of papers presented at the National Research and Development Conference of the Institute of Computer Scientists and Information Technologists, held on 26 & 27 September, at the Interaction Conference Centre, University of Natal, Durban. The Conference was organised by the Department of Computer Science and Information Systems of The University of Natal, Pietermaritzburg.

The papers contained herein range from serious technical research to work-in-progress reports of current research to industry and commercial practice and experience. It has been a difficult task maintaining an adequate and representative spread of interests and a high standard of scholarship at the same time. Nevertheless, the conference boasts a wide range of high quality papers. The program committee decided not only to accept papers that are publishable in their present form, but also papers which reflect this potential in order to encourage young researchers and to involve practitioners from commerce and industry.

The organisers would like to thank IBM South Africa for their generous sponsorship and all the members of the organising and program committees, and the referees for making the conference a success. The organisers are indebted to the Computer Society of South Africa (Natal Chapter) for promoting the conference among its members and also to the staff and management of the Interaction Conference Centre for their contribution to the success of the conference.

On behalf of the Organising Committee

Vevek Ram

Editor and Program Chair

Pietermaritzburg, September 1996

## **Organising Committee**

### **Conference General Chairs**

Mr Rob Dempster and Prof Peter Warren (UNP)

### **Organising Chair**

Dr Don Petkov (UNP)

### **Secretariat**

Mrs Jenny Wilson

### **Program Chair**

Prof Vevek Ram (UNP)

### **Program Committee**

Prof Peter Wentworth, Rhodes  
Dr Milan Hajek, UDW  
Prof Derek Smith, UCT  
Prof Anthony Krzesinski, Stellenbosch  
Dr Don Petkov, UNP  
Mr Rob Dempster, UNP  
Prof Peter Warren, UNP

# Table of Contents

Foreword	i
Organising Committee	ii
List of Contributors	vi
<b>Keynote Speaker</b>	
<i>The Role of Formalism in Engineering Interactive Systems</i>	1
M D Harrison and D J Duke	
<b>Plenary</b>	
<i>Industry-Academic-Government Cooperation to boost Technological Innovation and People Development in South Africa</i>	15
Tjaart J Van Der Walt	
<i>Checklist support for ISO 9001 audits of Software Quality Management Systems</i>	17
A J Walker	
<i>The IS Workers, they are a-changin'</i>	29
Derek Smith	
<b>Research</b>	
<i>Examination Timetabling</i>	35
E Parkinson and P R Warren	
<i>Generating Compilers from Formal Semantics</i>	43
H Venter	
<i>Efficient State-exploration</i>	63
J. Geldenhuys	
<i>A Validation Model of the VMTP Transport Level Protocol</i>	75
H.N. Roux and P.J.A. de Villiers	
<b>Intelligent Systems</b>	
<i>Automated Network Management using Artificial Intelligence</i>	87
M Wazzenboeck	
<i>A framework for executing multiple computational intelligent programs using a computational network</i>	89
H L Viktor and I Cloete	
<i>A Script-Based prototype for Dynamic Deadlock Avoidance</i>	95
C N Blewett and G J Erwin	
<i>Parallelism: an effective Genetic Programming implementation on low-powered Mathematica workstations</i>	107
H. Suleman and M. Hajek	
<i>Feature Extraction Preprocessors in Neural Networks for Image Recognition</i>	113
D Moodley and V Ram	

## **Real-Time Systems**

- The real-time control system model - an Holistic Approach to System Design* 119  
T Considine
- Neural networks for process parameter identification and assisted controller tuning for control loops* 127  
M McLeod and VB Bajic
- Reference Model for the Process Control Domain of Application* 137  
N Dhevcharan, A L Steenkamp and V Ram

## **Database Systems**

- The Pearl Algorithm as a method to extract information out of a database* 145  
J W Kruger
- Theory meets Practice: Using Smith's Normalization in Complex Systems* 151  
A van der Merwe and W Labuschagne
- A Comparison on Transaction Management Schemes in Multidatabase Systems* 159  
K Renaud and P Kotze

## **Education**

- Computer-based applications for engineering education* 171  
A C Hansen and P W L Lyne
- Software Engineering Development Methodologies applied to Computer-Aided Instruction* 179  
R de Villiers and P Kotze
- COBIE: A Cobol Integrated Environment* 187  
N Pillay
- The Design and Usage of a new Southern African Information Systems Textbook* 195  
G J Erwin and C N Blewett
- Teaching a first course in Compilers with a simple Compiler Construction Toolkit* 211  
G Ganchev
- Teaching Turing Machines: Luxury or Necessity?* 219  
Y Velinov

## **Practice and Experience**

- Lessons learnt from using C++ and the Object Oriented Approach to Software Development* 227  
R Mazhindu-Shumba
- Parallel hierarchical algorithm for identification of large-scale industrial systems* 235  
B Jankovic and VB Bajic

## **Information Technology and Organizational Issues**

*A cultural perspective on IT/End user relationships* 243  
A C Leonard

*Information Security Management: The Second Generation* 257  
R Von Solms

*Project Management in Practice* 267  
M le Roux

*A Case-Study of Internet Publishing* 271  
A Morris

*The Role of IT in Business Process Reengineering* 285  
C Blewett, J Cansfield and L Gibson

## **Abstracts**

*On Total Systems Intervention as a Systemic Framework for the Organisation of the Model Base of a Decision Support Systems Generator* 299  
D Petkov and O Petkova

*Modular Neural Networks Subroutines for Knowledge Extraction* 300  
A Vahed and I Cloete

*Low-Cost Medical Records System: A Model* 301  
O A Daini and T Seipone

*A Methodology for Integrating Legacy Systems with the Client/Server Environment* 302  
M Redelinghuys and A L Steenkamp

*Information Systems Outsourcing and Organisational Structure* 303  
M Hart and Kvavatzandis

*The relational organisation model* 304  
B Laauwen

*The Practical Application of a New Class of Non-Linear Smoothers for Digital Image Processing* 305  
E Cloete

*A Technology Reference Model for Client/Server Software Development* 306  
R C Nienaber

*The Feasibility Problem in the Simplex Algorithm* 307  
T G Scott, J M Hattingh and T Steyn

**Author Index** 309

## List of Contributors

**Vladimir B Bajic**

Centre for Engineering Research,  
Technikon Natal,  
P O Box 953  
Durban 4000

**C N Blewett**

Department of Accounting  
University of Natal  
King George V Avenue  
Durban 4001

**Justin Cansfield**

Department of Accounting  
University of Natal  
King George V Avenue  
Durban 4001

**Tom Considine**

Apron Services (Pty) Ltd  
P O Johannesburg  
International Airport  
1600

**Eric Cloete**

School of Electrical Engineering  
Cape Technikon  
Box 652  
Cape Town

**I Cloete**

Computer Science Department  
University of Stellenbosch  
Stellenbosch  
7600

**O A Daini**

Department of Computer Science  
University of Botswana  
Gaborone  
Botswana

**Nirvani Devcharan**

Umgeni Water  
Box 9  
Pietermaritzburg  
3200

**P J A de Villiers**

Department of Computer Science  
University of Stellenbosch  
Stellenbosch  
7700

**Ruth de Villiers**

Department of Computer Science and  
Information Systems  
UNISA  
Box 392, Pretoria, 0001

**G J Erwin**

Business Information Systems  
University of Durban-Westville  
Private Bag X54001  
Durban 4000

**G Ganchev**

Computer Science Department  
University of Botswana  
PBag 0022  
Gaborone, Botswana

**J Geldenhuys**

Department of Computer Science  
University of Stellenbosch  
Stellenbosch  
7700

**Louise Gibson**

BIS, Dept Accounting & Finance  
University of Durban  
Pvt Bag X10  
Dalbridge 4014

**Mike Hart**

Department of Information Systems  
University of Cape Town  
Rondebosch  
7700

**M. Hajek**

Department of Computer Science  
University of Durban-Westville  
Pvt Bag X54001  
Durban 4000

**A C Hansen**

Dept of Agricultural Engineering  
University of Natal  
Private Bag X01  
Scottsville 3209

**J M Hattingh**

Department of Computer Science  
Potchefstroom University for CHE  
Potchefstroom 2520



**Boris Jankovic**  
Centre for Engineering Research  
Technikon Natal  
P O Box 953,  
Durban 4000

**Paula Kotze**  
Department of Computer Science and  
Information Systems  
UNISA  
Box 392  
Pretoria, 0001

**J W Kruger**  
Vista University  
Soweto Campus  
Box 359  
Westhoven 2124

**A C Leonard**  
Dept of Informatics  
University of Pretoria  
Pretoria  
2000

**Ben Laauwen**  
Laauwen and Associates  
P O Box 13773  
Sinoville  
0129

**Mari Le Roux**  
Information technology, development: project  
leader  
Telkom IT 1015  
Box 2753  
Pretoria 0001

**P W L Lyne**  
Dept of Agricultural Engineering  
University of Natal  
Private Bag X01  
Scottsville 3209

**Rose Mazhindu-Shumba**  
Computer Science Department  
University of Zimbabwe  
Box MP167  
Harare, Zimbabwe

**Meredith McLeod**  
Centre for Engineering Research,  
Technikon Natal,  
P O Box 953  
Durban 4000

**D Moodley**  
Computer Management Systems  
Box 451  
Umhlanga Rocks  
4320

**Andrew Morris**  
P O Box 34200  
Rhodes Gift  
7707

**R C Nienaber**  
Technikon Pretoria  
Dept of Information Technology  
Private Bag X680  
Pretoria 0001

**E Parkinson**  
Department of Computer Science  
University of Port Elizabeth  
Box 1600  
Port Elizabeth 6000

**Don Petkov**  
Department of Computer Science and  
Information Systems  
University of Natal  
PBag x01  
Scottsville 3209

**Olga Petkov**  
Technikon Natal  
Box 11078  
Dorpspruit 3206  
Pietermaritzburg

**N Pillay**  
Technikon Natal  
Box 11078  
Dorpspruit 3206  
Pietermaritzburg

**V Ram**

Department of Computer Science and  
Information Systems  
University of Natal  
PBag x01  
Scottsville 3209

**Melinda Redelinghuys**

Department of Computer Science and  
Information Systems  
UNISA  
Box 392  
Pretoria, 0001

**Karen Renaud**

Computer Science and Information Systems  
UNISA  
Box 392  
Pretoria, 0001

**H N Roux**

Department of Computer Science  
University of Stellenbosch  
Stellenbosch  
7700

**T G Scott**

Department of Computer Science  
Potchefstroom University for CHE  
Potchefstroom  
2520

**T Seipone**

Department of Computer Science  
University of Botswana  
Gaborone  
Botswana

**Derek Smith**

Department of Information Systems  
University of Cape Town  
Rondebosch  
7700

**Anette L Steenkamp**

Department of Computer Science and  
Information Systems  
UNISA  
Box 392  
Pretoria, 0001

**T Steyn**

Department of Computer Science  
Potchefstroom University for CHE  
Potchefstroom 2520

**H. Suleman**

Department of Computer Science  
University of Durban-Westville  
Pvt Bag X54001  
Durban 4000

**A Vahed**

Department of Computer Science  
University of Western Cape  
Private Bag X17  
Bellville 7530

**A Van der Merwe**

Computer science and Informations Systems  
UNISA  
P O Box 392  
Pretoria,0001

**Tjaart J Van Der Walt**

Foundation for Research and Development  
Box 2600  
Pretoria, 0001

**K Vavatzandis**

Department of Information Systems  
University of Cape Town  
Rondebosch  
7700

**Y Velinov**

Dept Computer Science  
University of Natal  
Private Bag X01  
Scottsville 3209

**H Venter**

Department of Computer Science  
University of Port Elizabeth  
Box 1600  
Port Elizabeth 6000

**H L Viktor**

Computer Science Department  
University of Stellenbosch  
Stellenbosch  
7600

**R Von Solms**

Department of Information Technology  
Port Elizabeth Technikon  
Private Bag X6011  
Port Elizabeth 6000

**A J Walker**  
Software Engineering Applications  
Laboratory  
Electrical Engineering  
University of Witwatersrand  
Johannesburg

**Max Watzenboeck**  
University of Botswana  
Private Bag 0022  
Gaborone  
Botswana

**P Warren**  
Computer Science Department  
University of Natal  
P/Bag X01  
Scottsville 3209

# THE PEARL ALGORITHM AS A METHOD TO EXTRACT INFORMATION OUT OF A DATABASE.

J. W. Kruger  
KRGER-JW@SOREX.VISTA.AC.ZA  
Vista University - Soweto Campus  
P. O. Box 359, Westhoven 2142  
Tel: (011) 938-1701 x 238(W), (011) 477-2078 (H)

## Abstract

Big databases have been built up over time. The information in these databases can be converted by the Pearl algorithm [ Pearl 1988 ] from data to information. When a promotional drive is initiated, do we know who the potential customers are? The Pearl algorithm gives the belief that an individual will purchase. This means that the database can be used to select the individuals that are most likely to purchase. The breakeven point, where the cost of the promotional contact and the expected return are the same, can be calculated. The Pearl algorithm uses Bayes' probabilities to propagate belief through a tree.

## The Correlation between variables

The Pearson correlation between the numeric different fields on a flatfile (or relational) database can be calculated. Non-numeric fields, like yes/no answers can be converted to numeric by allocating a one to yes and a zero to no. The Pearl algorithm was developed for binary fields, and the conversion, mentioned here, give good results. In my personal experience market researchers without a statistical background tend to have too many non-numeric fields in their questionnaires. These fields are sometimes difficult to analyse and the data is then merely stored for possible future reference.

## Star decomposability

If we have three variables (Nodes A, B and C), then they are said to be star decomposable if a latent structure [ Lazarfeld 1966 ] exists that is the common cause to the three variables. The latent structure can be found as a hidden node. If the hidden variable is called W, then the correlation to the hidden variable can be determined (See figure 1).

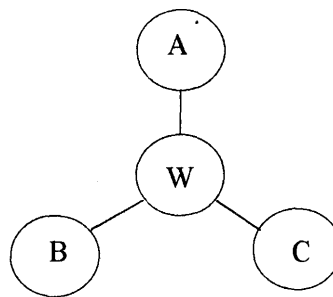


Figure 1. A star formation:

The latent structure, indicated by node W, links the leaf nodes A, B and C.

The correlation to W can be calculated from:

$$r_{AB} = r_{AW} \cdot r_{WB}$$

$$r_{AC} = r_{AW} \cdot r_{WC}$$

$$r_{BC} = r_{BW} \cdot r_{WC}$$

By solving:

$$r_{AW} = \sqrt{\frac{r_{AB} \cdot r_{AC}}{r_{BC}}}$$

The correlation to the hidden node, W, can be calculated.

## Causal Structure

To add a node, D, to this structure, one of the following correlation equations must hold:

1.  $r_{AD} \cdot r_{BC} = r_{AB} \cdot r_{CD}$
2.  $r_{BD} \cdot r_{AC} = r_{AB} \cdot r_{CD}$
3.  $r_{AD} \cdot r_{BC} = r_{AC} \cdot r_{BD}$

In equation 1, the node D must link to the arc between B and W.

In equation 2, the node D must link to the arc between A and W.

In equation 3, the node D must link to the arc between C and W.

By continuing with this reasoning a causal structure that is an acyclical graph (Tree / belief network) can be built up.

## Joint-Occurrence probabilities

For the time being, assume that we are working with binary variables and that W is the central node between the leaf nodes:  $X_1$ ,  $X_2$  and  $X_3$ :

Define the seven joint-occurrence probabilities as [Lazarsfeld 1966]:

$$p_i = p(x_i = 1)$$

$$p_{ij} = p(x_i = 1, x_j = 1)$$

$$p_{ijk} = p(x_i = 1, x_j = 1, x_k = 1)$$

The standard deviation of a Bernoulli variable is given by:  $\sigma_i = [p_i (1 - p_i)]^{1/2}$

and the correlation coefficients:  $\rho_{ij} = (p_{ij} - p_i p_j) / \sigma_i \sigma_j$

## Link Matrix

Define  $f_i = p(x_i = 1 \mid w = 1)$  and  $g_i = p(x_i = 1 \mid w = 0)$ . Now we can solve the elements of the link matrices [Bhat 1984] or transition probability matrix.

The prior probabilities of the node W, multiplied by the transition matrix, gives the prior probabilities for the node  $X_i$ .

$$\text{Let } S_i = \pm [(p_{ij} - p_i p_j) (p_{ik} - p_i p_k) / (p_{jk} - p_j p_k)]^{1/2}$$

$$\mu_i = (p_i p_{ijk} - p_{ij} p_{ik}) / (p_{jk} - p_j p_k)$$

$$K = S_i / p_i - p_i / S_i + \mu_i / (S_i p_i)$$

and  $\alpha = t^2 / (1 + t^2)$ , where t is the solution to  $t^2 + Kt - 1 = 0$

$$\text{then } f_i = p_i + S_i [(1 - \alpha) / \alpha]$$

$$\text{and } g_i = p_i - S_i [(1 - \alpha) / \alpha]$$

Anybody interested in the manipulations can look at the theorem by Lazarfeld [ Lazarfeld 1966, Pearl 1988 ].

### Non-binary variables

The Pearl algorithm was developed for binary variables. For variables in more states or for continuous variables the causal structure can still be found.

To generalise, a method to calculate the transition matrices must be found. In general we must solve  $3n(n - 1)$  unknown parameters (elements) to find the transition matrices, to connect three leaf nodes to the belief network (The tree). A generalisation to a continuous state space scenario is needed.

A heuristic to solve the elements of the link matrices with more than two states has been developed . [ Kruger 1996a ]

### Belief Propagation

In a single chain structure (see Figure 2):

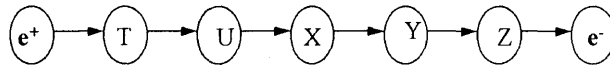


Figure 2. A causal chain with evidential data at its head ( $e^+$ ) and tail ( $e^-$ )

To propagate the prior probability down away from the root, the prior probability vector must be pre-multiplied to the transition matrix.

To propagate the likelihood towards the root, the likelihood must be post-multiplied to the transition matrix.

### For the tree however:

Consider a typical node  $x$ , with  $m$  children  $y_1, y_2, \dots, y_m$ , and parent  $u$ :

### Belief updating in a tree:

As with the chain structure above; Using the causal support,  $\pi(x)$ , and diagnostic support,  $\lambda(x)$ , the belief distribution of  $x$  is:

$$BEL(x) = \alpha P(e^+ | x) P(x | e^-) = \alpha \lambda(x) \pi(x)$$

$\alpha$  is a normalizing constant to make  $\sum BEL(x) = 1$ .

Note: The multiplication of these vectors must not be confused with the scalar product of vectors. By multiplication we mean that the corresponding co-ordinates are multiplied, giving a vector of products, each co-ordinate corresponding to the belief that the state is true.

The difference comes in the way that  $\lambda(x)$  is calculated.

$$\lambda(x) = P(e^- | x) = P(\text{evidence from the children} | x) = \prod_{i=1}^m \lambda(y_i \text{ of } x). \text{ because } x \text{ separates its children and the siblings are conditionally independent.}$$

$$\pi(x) = \sum_u P(x | u) \pi_x(u) = \pi(u) P(x | u)$$

### Bottom-up propagation:

Node x uses the  $\lambda(x)$  message from the children to compute a new message  $\lambda(x \text{ of } u)$  to send to its parent u.

$$\lambda(x \text{ of } u) = \lambda(x) P(x | u)$$

### Top-down propagation:

The new  $\pi$  message sent to by node x to its j-th child  $y_j$  is:

$$\pi(y_j) = \alpha \pi(x) \prod_{k \neq j} \lambda(y_k \text{ of } x)$$

but,

$$BEL(x) = \alpha \lambda(x) \pi(x), \text{ therefore } \pi(y_j) = \alpha BEL(x) / \lambda(y_j).$$

Readers interested in Belief propagation in more general networks are referred to Pearl [1996]. As the Pearl algorithm produces a causal tree, belief propagation in more general networks are not included here.

From the above it is evident that we only need the prior probabilities of the root and the likelihood functions of the leaf nodes to calculate the belief of all the nodes.

The prior probabilities for the leaf nodes can be used as likelihoods. The prior probability for the root must still be found. The easiest is to make the dependent variable the root. This means that certain transition matrices will have to be inverted. Of the dependent variable is the root, then the prior probabilities can be found empirically.

## Application

In a database, the variable, of whether the client will purchase or not, can be found from previous experience (data). Last time a promotion went out, who purchased. It can also be found from a pilot study. The causal structure with other fields in the database can be found by the Pearl algorithm.

Read the clients in the database sequentially. Instantiate the known fields for each client on the database.

Now these likelihoods and the prior probability for the root can be propagated through the belief network, to get the belief of all the undefined variables in the belief network.

Kruger [Kruger 1996a] gives a method to evaluate the different possible belief networks created. Kruger also compares the accuracy of belief networks on data, used to create the networks, with applying the algorithm to other data.

The belief of a state like purchase can be multiplied by the expected monetary gain of a purchase, to get the expected gain of contacting this client. If the gain is more than the cost, then it will be wise to go ahead. If the gain is less than the cost, then of course no contact must be made.

With this method qualitatively seemingly unrelated fields can assist in the highly competitive field of retail sales. Fields giving biographical data like gender, home language, marital status can now be used to predict sales.

Personell Psychologists use biographical information to decide on the the selection of applicants. They create a weighted biographical information index, test the validity of this index and then base their decisions on this. The Pearl algorithm is a much more refined method to obtain a valid selection criterium out of a database [ McCormich 1992, Kruger 1996b ].

### **Acknowledgement**

I want to thank Prof. David Lubinsky from the University of the Witwatersrand for the help he gave me in understanding the Pearl algorithm and Bayes' methods.

### **References**

1. Bhat, U. N., "Elements of Applied Stochastic Processes", John Wiley & Sons, Inc., 1984
2. (a) Kruger, J. W., "Generalizing the number of states in Bayesian Belief Propagation, as applied to Portfolio Management", Masters Research Report, University of the Witwatersrand, 1996.
3. (b) Kruger, J. W., "Reliability and Validity of selection criteria into courses and of tests and exams.", Proceedings of the South African Computer Lecturers Association Conference at Pilansberg, 1996.
4. Lazarfeld, P.F., "Latent structure analysis, in measurement and prediction", eds. S. A. Stoufer, L. Guttman, E.A. Suchman, P.F.Lazarfeld, S.A.Star and J.A.Classen, Wiley, New York, 1966
5. McCormick E. J., and Ilgen D, "Industrial and Organisational Psychology" 8th edition, Scotprint Ltd, Musselburgh, ISBN 0-415-09452-6, 1992
6. Pearl, J., "Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference" Morgan Kaufmann, San Mateo, California 1988



