City Research Online

City St George's, University of London

Citation: Cowell, R. (2001). FINEX: Forensic Identification by Network Expert Systems (Statistical Research Paper No. 22). London, UK: Faculty of Actuarial Science & Insurance, City University London.

This is the unspecified version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/2366/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

FINEX: Forensic Identification by Network Expert Systems

by

Robert G. Cowell

Statistical Research Paper No. 22

March 2001

ISBN 1 901615 60X

"Any opinions expressed in this paper are my/our own and not necessarily those of my/our employer or anyone else I/we have discussed them with. You must not copy this paper or quote it without my/our permission".

FINEX: **F**orensic **I**dentification by **N**etwork **EX**pert systems.

Robert G. Cowell
Department of Actuarial Science and Statistics,
City University London, Northampton Square, London EC1V 0HB

DASS Statistical Research Report 22
March 20, 2001

Abstract

It has recently been shown how to formulate complex problems of forensic identification inference, such as occur in disputed paternity cases, in terms of Bayesian networks [1][2]. This makes it possible to use available software, such as the free program XBAIES [5] or commercial systems such as HUGIN [4], to analyse such problems efficiently and accurately. However, while specifying an appropriate network for a specific problem is not difficult, current Bayesian network software is ill-suited both to the task of editing the many local conditional probability tables — whose values are determined by the standard Mendelian model of genetic inheritance — and in combining the evidence from several genetic markers.

Here I introduce two software tools, called FINEX and FINEXGUI, that enable users to quickly specify and analyse such problems, and that free the user of the tedious task of editing conditional probabilities.

Keywords Bayesian networks, genetic markers, forensic identification.

1 Introduction

Consider the following simple case of disputed paternity. A mother **m** claims that the father of her daughter **c** is a certain male, but he denies this; we denote this putative father by **pf**. DNA evidence in the form of measurements on genetic markers is available on all three parties. The problem of determining whether the putative father, **pf** is the true father (labelled by **tf**) can be viewed as a problem of model selection. In one model **pf** is taken to be the true father, while in a second model some other person (who we shall assume in our example is drawn randomly from the general population) is taken to be the true father. For both models the joint probability distribution for the possible genetic make-ups of all parties *prior* to taking the DNA evidence into account is established. From these the likelihood of the available DNA evidence is evaluated for each model, and on the basis of their relative values a conclusion is drawn as to whether to accept or reject the mother's accusation that **pf** is the father of child **c**.

In the following sections I will describe how the problem described above may be analysed by using Bayesian networks, and describe a purpose written program to analyse this and similar problems. The plan of the paper is as follows. Section 2 gives a brief resumé of the ideas and basic probabilistic results about DNA and genetic inheritance required for this paper. Section 3 presents the above paternity example as a standard forensic diagram, and describes how it may be represented by a Bayesian network which may be used to find a likelihood ratio for paternity. Section 4.1 introduces a forensic indentification specification language using the above example. Section 4.2 describes the program FINEX, which takes a specification script as input, and performs the required probability calculations. FINEX differs from other specialist forensic software, such as **familias**[3], in being based upon general Bayesian network software and inference algorithms. In Section 4.3 the program FINEXGUI is introduced, which the user can use to create a specification script using simple graphical manipulations.

2 Some background information on genetics

Every cell of a person contains 46 chromosomes, which can be grouped into 23 distinguishable pairs of *homologous* chromosomes. For each homologous pair, one chromosome is inherited from the person's mother, the other from the person's father. However these inherited chromosomes are not copies of the parents' chromosomes. This is because sperm and egg cells are produced by a cell-division process called *meiosis*, during which each homologous pair of chromosomes randomly splits and recombines in a process called *crossover*, to create a new pair of chromosomes. One chromosome from each of these crossover pairs is represented in the gamete (sperm or egg), which thus has 23 unpaired chromosomes. Such cells are called *haploid*. When two haploid cells combine the resulting zygotic cell has 23 pairs of chromosomes; such cells are called *diploid*. Thus a child will inherit, for each of its father's 23 pairs of homologous chromosomes, only one of the two crossover chromosomes, each with equal probability; the same goes for the mother's 23 pairs of chromosomes. During cell division of a fertilized egg these chromosomes are replicated. Hence almost any cell in the body may be used for forensic identificiton. However, it is not possible to tell solely from an individual's chromosome which parent it came from.

Each chromosome consists of a sequence of large molecules called *neucleotides*, there being four neucleotides in all. (These are adenine, cytosine, guanine and thymine, referred to by the letters A,C,G and T respectively). Genes are particular subsequences in the paired chromosomes; any specific gene is associated with a specific chromosome at a particular position called a *locus* (hence there are two genes at a locus of a chromosome pair). A gene used in forensic identification at a particular locus is called a *marker*. A marker may have one of a number of alternative chemical compositions, called *alleles*. The number of possible alleles will vary according to the particular marker. An unordered pair of alleles is associated with each marker measured by a forensic laboratory, an allele from each of the two homologous chromosomes associated with the marker. One of these is inherited from the father, and is called the *paternal gene*, the other is inherited from the mother and is called the *maternal gene*. This pair is called the *genotype*. (I shall not go into the details of how the laboratory measurements of genotypes are performed.) Mendel's first law dictates the form of the probabilities of individual genes being passed on from parent to offspring.

Table 1 shows the alleles for the marker HBGG, which is routinely used in forensic identification, and their (possible) frequency of occurence in a (specific) population. Such frequencies — typically obtained from a forensic analysis of many individuals — together with the Mendelian law of gene inheritance, define all of the probabilities that will enter into the genetic inheritance models considered later.

3 Forensic diagrams

The disputed paternity scenario outlined in Section 1 is illustrated diagrammatically in Figure 1. In such diagrams males are represented by rectangles, females by circles. Time order is top to

Table 1: Alleles and illustrative population frequencies of the marker HBGG.

Allele	A	B	C
Freq. in Pop.	0.53	0.45	0.02

bottom, so that nodes on a lower level in the diagram represent later generations. The shaded nodes indicate that forensic measurements are available on those individuals. The question that the diagram asks, is whether **pf** if the true father **tf** of the child **c**.

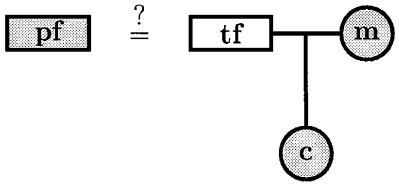

Figure 1: Diagram of a simple paternity dispute

Following [1][2] we model this diagram of disputed paternity by the paternity network Figure 2, which describes the inhertance pathways of alleles. Genotype nodes have labels ending with **gt** while paternal and maternal gene nodes have labels ending with **pg** and **mg** respectively. Finally, the figure includes a *query node*, **tf=pf?**, whose function we shall describe later.

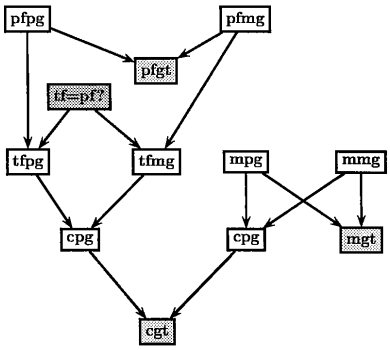

Figure 2: Paternity network corresponding to Figure 1.

Let us look at the semantics of Figure 2, which is an example of a Bayesian network. First note that each genotype node has no graphical child in the graph.[1] Furthermore each genotype

[1]Unfortunately, the terms *parent* and *child* have a specific meaning for Bayesian networks which are apt to

node has two graphical parents, one a paternal gene and one a maternal gene — recall that the genotype is an unordered pair of maternal and paternal genes. The conditional probability table associated with a genotype node will consist of ones and zeroes, indicating the logical possibility or impossibility of a given genotype being a combination of the given maternal and paternal genes.

Next consider those **pg** nodes which have graphical parents (excluding for now the **tfpg** and **tfmg** nodes). Again there are exactly two graphical parents: one, the father's maternal gene, the other the father's paternal gene. Similarly those **mg** nodes which have parents in the graph also have exactly two graphical parents: the mother's maternal and paternal genes. The conditional probability tables associated with these nodes are readily deduced from the law of Medelian inheritance, and will have entries consisting only of the values 0, 0.5 and 1.

Now consider those **mg** and **pg** nodes which do not have graphical parents; in reality the individuals represented by such nodes will have real parents, grandparents, etc. For the purposes of the paternity network such individuals are ignored (unless there is DNA information about them or their other descendants). For these nodes the associated unconditional probabilities are simply the allele frequencies as for example in Table 1.

The query node **tf=pf?** has two states, *yes* and *no*. This node has no parents, and its unconditional probability table takes values $(0.5, 0.5)$. Now consider the node **tfpg**, which has the graphical parents **pfpg** and **tf=pf?**. For the state *yes* of **tf=pf?**, the true father's paternal gene will be identical to the putative father's paternal gene. This is represented by the conditional probability table $P(\mathbf{tfpg} \mid \mathbf{pfpg}, (\mathbf{tf} = \mathbf{pf?}) = \textit{yes})$ taking values of zero unless the state of **tfpg** is the same as that of **pfpg**, in which case the value is unity. However, for the state *no* of **tf=pf?**, the putative father's paternal gene is irrelevant to the true father's paternal gene, in which case the conditional probability table $P(\mathbf{tfpg} \mid \mathbf{pfpg}, (\mathbf{tf} = \mathbf{pf?}) = \textit{no})$ is taken to be that of the distribution of genes in the population. Similar remarks apply to the node **tfmg**.

Having set up the Bayesian network model for a particular marker, we may now enter and propagate the evidence known about the genotype nodes **cgt**, **mfgt**, and **pfgt**, and then read of the posterior probabilities on the node **tf=pf?**. These give us the likelihoods of the two paternal hypotheses because of the uniform prior place on this node. If we repeat this for a number of different markers which are on distinct homologous chromosomes, (we will require for each such marker a different Bayesian network having the same structure, however the states of the nodes and their associated probability tables will depend upon the marker), then because of the independence in the inheritance of these chromosomes, the individual likelihoods may be multiplied together to give a joint likelihood over all markers. Finally, an overall likelihood ratio of the two hypotheses may be calculated, upon which a judgement about the disputed paternity

conflict with the more natural familial meanings of the terms for the individuals that we are modelling. To avoid confusion, when referring to a parent or child in the Bayesian network sense, I will use the terms *graphical parent* and *graphical child.* respectively

may be made.

4 A specification language

4.1 The simple paternity example

Using the simple paternity example I shall now illustrate a specification language for forensic identification which can be interpreted by the program FINEX to calculate the required likelihoods. Let us suppose that the following observations on the marker FESFPS have been collected: for the mother **m**, alleles 10 and 12; for the putative father **pf**, alleles 10 and 12; and for the child **c**, alleles 12 and 12. With this data, Figure 3 shows a specification of the paternity problem intended for the FINEX interpretor.

```
genepool gp < "genepool.data" ;
female  m  < gp ;
male    tf < gp ;
male    pf ?< gp : tf ;
union   u  < m + tf ;
female  c  < u ;

data    pf <  "FESFPS"    "10"   "12" ;
data    m  <   "FESFPS"   "10"   "12" ;
data    c  <   "FESFPS"   "12"   "12" ;

likelihood ;
```

Figure 3: Specification of simple paternity problem.

Each command ends with a semicolon. The first line declares a variable **gp** to be of type `genepool` — currently FINEX expects data on population allele frequencies to be stored in textfiles — so the first line says to read the data held in the file `genepool.data` into the variable **gp**. The '<' symbol is an *assignment* operator. Note that in the previous diagrams the background population genepool was implicit; here the specification language for FINEX requires that it be defined explicitly. This has the bonus that several different ethnic groups may be modelled explicitly.

The second line declares **mf** to be of type **female**, and that she inherits her alleles from **gp**, the population genepool. The third line in a similar manner declares the true father **tf** to be of type **male**, again drawn from **gp**. The fourth line declares **pf**, the putative father, to be of type **male**, but with his genes drawn either from the general population genepool or from **tf**. The question mark flags **pf** to be used for generating the query node, the composite symbol '? <' being called a *query-assignment*. There may be several terms on the right of the query-assignment, each separated by a colon, but there can be only one person declared using a query-assignment statement.

Now the child needs to be declared; before doing this another data type is first introduced, called a **union**. (A union corresponds to a *marriage node* in genetics.) A union represents all the possible children that a couple can have (so a union is equivalent to a subpopulation

genepool), hence on the right the two individuals making up the couple are specified: must be one male and the other female, but they may be specified in any order. After declaring the union **u**, the child **c** is declared to be of type **female**, drawing her genes from the union **u**.

An important constraint is that a variable occurring on the right-hand-side of an assignment or query-assignment statement must have been declared in some previous line.

The next three lines specify the data available on the individuals; note that data on several markers may be given. The final line `likelihood` is a command to evaluate the likelihoods of the various queries based upon the specified data. An alternative form of this command is to write `likelihood` followed by a subset of names of observed markers. Individual likelihoods based upon these markers, and overall likelihood-ratios, are then evaluated.

The above example shows all of the datatypes of the specification language. An interesting aspect is that the specification of even this simple problem is not unique. Figure 4 shows an alternative specification.

```
genepool gp < "genepool.data" ;
female  m  < gp ;
male    pf < gp ;
male    tf ?< gp : pf ;
union   u  < m + tf ;
female  c  < u ;

data  pf <  "FESFPS"   "10"   "12" ;
data  m  <  "FESFPS"   "10"   "12" ;
data  c  <  "FESFPS"   "12"   "12" ;

likelihood ;
```

Figure 4: Alternative specification of the simple paternity problem

These two specifications differ in the questions that are being posed. The specification of Figure 3 asks of **pf**: Is **pf** the true father, or someone else from the general population (excluding **c** and **m**)? The alternative specification in Figure 4 asks the different question: Is **tf**, the true father, the same person as the putative father, **pf**, or someone else from the general population? The first form asks a question about **pf**, the second asks a logically equivalent question about **tf** — for the problem of determining if **c** is the child of **pf** by **m** the questions are the same. Yet another way of posing the paternity question is to place the emphasis of question onto the child **c**, as shown in Figure 5, where now we replace the node **tf** by **af**, meaning *alternative male*. The question then posed is not one of identity but one of the child's origins: Is **c** a child **c1** by **af** or a child **c2** by **pf** ? (This specification also describes the following scenario: A woman **m** has two daughters, **c1** by **af** and **c2** by **pf**. Both daughters have disappeared, but a body **c** is found which is known to be one of the two daughters. DNA evidence is available on **m**, **pf** and **c**. Which of the two daughters is **c**?)

Corresponding to each of these three representations will be three different Bayesian networks, but they all give the same likelihoods regarding the paternity of **c**. (Incidentally, the Bayesian network in Figure 2 corresponds to the specification in Figure 4.)

```
genepool gp < "genepool.data" ;
female  m  < gp ;
male    pf < gp ;
male    af < gp;
union   u1 < m + af ;
union   u2 < m + pf ;
female  c1 < u1 ;
female  c2 < u2 ;
female  c ?< c1 : c2 ;

data   pf <  "FESFPS"   "10"   "12" ;
data   m  <  "FESFPS"   "10"   "12" ;
data   c  <  "FESFPS"   "12"   "12" ;

likelihood ;
```

Figure 5: A third specification of the simple paternity problem

4.2 FINEX

FINEX is an interpretor for processing commands in the specification language, and is built upon the software that drives XBAIES [5]. It can be run interactively at the command line, with the user typing in commands. However, it is more convenient to prepare a text file containing the commands, and to have FINEX read this file through standard input redirection.

In general, FINEX takes the specification input and, for each marker that data is available on and a likelihood is required for: (i) constructs an appropriate Bayesian network; (ii) enters and propagates the marker evidence; and (iii) evaluates the likelihood for that marker. Overall likelihood-ratios are also returned (calculated in the case of data on several markers that they are on distinct homologous chromosomes).

The marker FESFPS has seven possible allele values, labelled by integers from 8 to 14. However in the simple example only the alleles 10 and 12 were observed. Thus FINEX implements the procedure described in [2], whereby to increase the efficiency of the likelihood calculations for each marker all of the unobserved alleles are combined into one unobserved state **x**. In the example this leads to a reduction of the largest conditional probability table from a potential size of $7^3 = 343$ to one of just 27 entries.

The output on passing FINEX the specification in Figure 3, using allele frequencies given by: $p(10) = 0.288$, $p(12) = 0.258$, $p(x) = 0.454$, is shown in Figure 6. From this we see that the likelihood ratio of **pf** being the true father to **pf** being a person drawn at random from the general population is almost two, based upon this single maker DNA evidence.

It is worth emphasizing that a user only needs to specify the population allele frequencies; the user does not need to construct and edit potentially large conditional probabilities tables — FINEX handles their construction automatically. To reduce the possibility of user error in typing in the specification commands, either interactively or when creating a script file, the program FINEXGUI was written which allows a user to diagrammatically build up the specification.

```
Marker : FESFPS
"pfQN"
%   pfQN: pfQN

    0.340369  % pfQN=gp_?   |
    0.659631  % pfQN=tf_?   |

    prob(evid) = 0.00418498

likelihood ratio for the various queries
gp_?    0.516
tf_?    1.93798
```

Figure 6: FINEX output based upon the input Figure 2

4.3 FINEXGUI

FINEXGUI is a graphical interface program which allows users to construct a diagrammatic representation of commands which would appear in a specification script, and indeed it generates a specification file from the diagram for interpretation by FINEX. Figure 7 shows in diagrammatic form the script of Figure 3. It is actually very close to the diagram that a user would construct using FINEXGUI. The only difference is in the query-assignment, which in Figure 7 is illustrated by the curved line with the question mark nearby crossing the directed edges from **gp** and **tf** to **pf**. In FINEXGUI, a specific query-person node is formed, which is displayed differently from the nodes representing other people. Almost all input is mouse driven, using pull down menus, with data entered on a person by means of a mouse selectable popup-listbox.

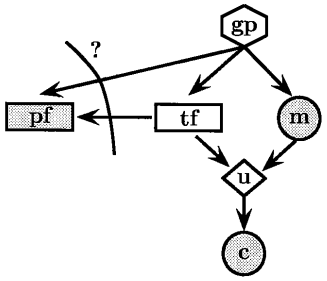

Figure 7: FINEX network corresponding to Figures 1 and 3.

Figure 7 can be viewed as an elaboration of Figure 1, with explicit inclusion of the population genepool and a greater explicitness of the identity question being posed.

5 Summary

I have introduced two programs, FINEX and FINEXGUI and an associated specification language for use in problems of forensic identification using genetic markers. The programs automate a large amount of the specification of the Bayesian networks described in [2], thus saving a considerable amount of time and eliminating the chances of human error arising in editing conditional probability tables using currently available Bayesian network software. At present the program runs on MS Windows 9x and NT operating systems, though porting to other platforms is underway. It is planned to combine the FINEX and FINEXGUI programs into one, and to extend the functionality in other ways, such as dealing with genetic mutation. (There is some limited functionality for modelling mutation at the moment.) It should be stressed that the specification language is under active development, so that future versions could look quite different from the format presented in this paper.

Acknowledgments

I would like to thank the European Science Foundation HSSS initiative, and the Danish Research Councils through their PIFT programme and the DINA project, for financial support in attending workshops. I would also like to thank A. P. Dawid, S. L. Lauritzen, J. Mortera, V. L. Pascali and N. A. Sheehan for introducing me to concepts and problems of statistical genetics and forensic identification.

References

[1] Dawid, A. P., van Boxel, D., Mortera, J. and Pascali, V. L. Inference about disputed paternity from an incomplete pedigree using a probabilistic expert system. *Bull. Int. Statist. Inst. Contributed Papers, Book 1*, **58**, pp. 241–242. (1999).

[2] Dawid, A. P., Mortera, J., Pascali, V. L. and van Boxel, D. Probabilistic expert systems for forensic inference from genetic markers. *Research Report 215.* Department of Statistical Science, University College London. (2000)

[3] Egeland, T., Mostad, P. F., Mevåg, B. and Stenersen, M. Beyond traditional paternity and identification cases. Selecting the most probable pedigree. *Forensic Science International*, **110**, pp. 47–59. (2000).

[4] HUGIN is a commercial package available at *http://www.hugin.dk.*

[5] XBAIES, a general shell for building Bayesian networks models, is freely available at *http://www.staff.city.ac.uk/~rgc* for educational use.

City University London

DEPARTMENT OF ACTUARIAL SCIENCE AND STATISTICS

Actuarial Research Papers since 2000

---

122. Booth P.M. and Cooper D.R. The Tax Treatment of Pensions. April 2000. 36 pages. ISBN 1 901615 42 1

123. Walsh D.E.P. and Rickayzen B.D. A Model for Projecting the number of People who will require Long-Term Care in the Future. Part I: Data Considerations. July 2000. 37 pages. ISBN 1 901615 43 X

124. Rickayzen B.D. and Walsh D.E.P. A Model for Projecting the number of People who will require Long-Term Care in the Future. Part II: The Multiple State Model. July 2000. 27 pages. ISBN 1 901615 44 8

125. Walsh D.E.P. and Rickayzen B.D. A Model for Projecting the number of People who will require Long-Term Care in the Future. Part III: The Projected Numbers and The Funnel of Doubt. July 2000. 61 pages. ISBN 1 901615 45 6

126. Cooper D.R. Security for the Members of Defined Benefit Pension Schemes. July 2000. 23 pages. ISBN 1 901615 45 4

127. Renshaw A.E. and Haberman S. Modelling for mortality reduction factors. July 2000. 32 pages. ISBN 1 901615 47 2

128. Ballotta L. and Kyprianou A.E A note on the -quantile option. September 2000. ISBN 1 901615 49 9

129. Spreeuw J. Convex order and multistate life insurance contracts. December 2000. ISBN 1 901615 50 2

130. Spreeuw J. The Probationary Period as a Screening Device. December 2000. ISBN 1 901615 51 0

131. Owadally M.I. and Haberman S. Asset Valuation and the Dynamics of Pension Funding with Random Investment Returns. December 2000. ISBN 1 901615 52 9

132. Owadally M.I. and Haberman S. Asset Valuation and Amortization of Asset Gains and Losses in Defined Benefit Pension Plans. December 2000. ISBN 1 901615 53 7

133. Owadally M.I. and Haberman S. Efficient Amortization of Actuarial Gains/Losses and Optimal Funding in Pension Plans. December 2000. ISBN 1 901615 54 5

134. Ballotta L. -quantile Option in a Jump-Diffusion Economy. December 2000. ISBN 1 901615 55 3

135. Renshaw A. E. and Haberman S. On the Forecasting of Mortality Reduction Factors. February 2001. ISBN 1 901615 56 1

136. Haberman S., Butt Z. & Rickayzen B. D. Multiple State Models, Simulation and Insurer Insolvency. February 2001. 27 pages. ISBN 1 901615 57 X

137. Khorasanee M.Z. A Cash-Flow Approach to Pension Funding. September 2001. 34 pages. ISBN 1 901615 58 8

Statistical Research Papers

1. Sebastiani P. Some Results on the Derivatives of Matrix Functions. December 1995. 17 Pages. ISBN 1 874 770 83 2

2. Dawid A.P. and Sebastiani P. Coherent Criteria for Optimal Experimental Design. March 1996. 35 Pages. ISBN 1 874 770 86 7

3. Sebastiani P. and Wynn H.P. Maximum Entropy Sampling and Optimal Bayesian Experimental Design. March 1996. 22 Pages. ISBN 1 874 770 87 5

4. Sebastiani P. and Settimi R. A Note on D-optimal Designs for a Logistic Regression Model. May 1996. 12 Pages. ISBN 1 874 770 92 1

5. Sebastiani P. and Settimi R. First-order Optimal Designs for Non Linear Models. August 1996. 28 Pages. ISBN 1 874 770 95 6

6. Newby M. A Business Process Approach to Maintenance: Measurement, Decision and Control. September 1996. 12 Pages. ISBN 1 874 770 96 4

7. Newby M. Moments and Generating Functions for the Absorption Distribution and its Negative Binomial Analogue. September 1996. 16 Pages. ISBN 1 874 770 97 2

8. Cowell R.G. Mixture Reduction via Predictive Scores. November 1996. 17 Pages. ISBN 1 874 770 98 0

9. Sebastiani P. and Ramoni M. Robust Parameter Learning in Bayesian Networks with Missing Data. March 1997. 9 Pages. ISBN 1 901615 00 6

10. Newby M.J. and Coolen F.P.A. Guidelines for Corrective Replacement Based on Low Stochastic Structure Assumptions. March 1997. 9 Pages. ISBN 1 901615 01 4.

11. Newby M.J. Approximations for the Absorption Distribution and its Negative Binomial Analogue. March 1997. 6 Pages. ISBN 1 901615 02 2

12. Ramoni M. and Sebastiani P. The Use of Exogenous Knowledge to Learn Bayesian Networks from Incomplete Databases. June 1997. 11 Pages. ISBN 1 901615 10 3

13. Ramoni M. and Sebastiani P. Learning Bayesian Networks from Incomplete Databases. June 1997. 14 Pages. ISBN 1 901615 11 1

14. Sebastiani P. and Wynn H.P. Risk Based Optimal Designs. June 1997. 10 Pages. ISBN 1 901615 13 8

15. Cowell R. Sampling without Replacement in Junction Trees. June 1997. 10 Pages. ISBN 1 901615 14 6

16. Dagg R.A. and Newby M.J. Optimal Overhaul Intervals with Imperfect Inspection and Repair. July 1997. 11 Pages. ISBN 1 901615 15 4

17. Sebastiani P. and Wynn H.P. Bayesian Experimental Design and Shannon Information. October 1997. 11 Pages. ISBN 1 901615 17 0

18. Wolstenholme L.C. A Characterisation of Phase Type Distributions. November 1997. 11 Pages. ISBN 1 901615 18 9

19. Wolstenholme L.C. A Comparison of Models for Probability of Detection (POD) Curves. December 1997. 23 Pages. ISBN 1 901615 21 9

20. Cowell R.G. Parameter Learning from Incomplete Data Using Maximum Entropy I: Principles. February 1999. 19 Pages. ISBN 1 901615 37 5

21. Cowell R.G. Parameter Learning from Incomplete Data Using Maximum Entropy II: Application to Bayesian Networks. November 1999. 12 Pages ISBN 1 901615 40 5

22. Cowell R.G. FINEX : Forensic Identification by Network Expert Systems. March 2001. 10 pages. ISBN 1 901615 60X

23. Cowell R.G. When Learning Bayesian Networks from Data, using Conditional Independence Tests is Equivalant to a Scoring Metric. March 2001. 11 pages. ISBN 1 901615 61 8