

## Data Mining in Durham

*Data Mining is a term inspired by the exploits of the gold diggers of 1849. Based on the expectation that 'there's gold in them thar hills'...*

The Autumn NCAF Meeting 2002 will be held in the University of Durham, hosted by Kathryn Thornton from the Computer Science Department. The special theme for this meeting is Data Mining.

Data Mining is a term inspired by the exploits of the gold diggers of 1849. Based on the expectation that 'there's gold in them thar hills', Data Mining (DM) refers to the use of a variety of techniques which are applied to data with the aim of discovering valuable items of information or decision making knowledge. These can subsequently be used in areas such as decision support, forecasting, prediction and estimation. The base data often contains large volumes of records which, in its raw form, is of low value and little use can be made of it directly. It is the 'hidden' information brought out by the mining techniques that produces the business advantages. The Computer Science Department of Durham is currently engaged in investigating the application of data mining in the areas of intelligent decision support and functional programming.

The University itself has an interesting history dating back to the efforts of William van Mildert, last Prince Bishop of Durham, and of the Dean and Chapter. This led, in 1832, to the passage through Parliament of an 'Act to enable the Dean and Chapter of Durham to appropriate part of the property of their church to the establishment of a University in connection therewith'. Temporary accommodation was provided in the house known as Archdeacon's Inn on Palace Green and the first students came into residence in 1833. On 1 June 1837 a Royal Charter was issued recognising and confirming the constitution of the University. Seven days later the first Durham degrees were conferred under the authority of this Charter. An Order of the Queen in Council of 8 August 1837 appropriated Durham Castle, previously a palace of the Palatinate See, to the uses of the University. One of the objectives of the founders was to establish in the North of England 'an Institution which should secure to its inhabitants the advantage of a sound yet not expensive academical education'.

The new University in 1832 was Collegiate, although initially there was only one college, now University College situated in Durham Castle. In 1846 this was followed by Hatfield Hall, where expenses were reduced by providing all meals in common at a fixed charge and by letting the rooms furnished. Unattached, later known as non-collegiate students, were first admitted in 1871. They themselves established a St. Cuthbert's Society in 1888. In 1947 St. Cuthbert's Society became the recognised designation of the non-collegiate students. Its current relationship with Council was adopted in 1948. Bede College, established independently for men in 1839, took University degree students from 1892. In 1975 it was merged with its women's counterpart, St. Hild's College, which had been founded independently in 1858 and brought into connection with the University in 1896. Two private halls, St. Chad's and St. John's founded in 1904 and 1909 took the style and title of an Independent College in 1919.

The current University attractions include:

The Oriental Museum: from eighth-century grave figures to contemporary New Year street lions – discover ancient Egypt, Japan and China in the heart of Durham.

Old Fulling Mill Museum: one of the most photographed museums in the world because of its position underneath the cathedral, it covers the history of the area.

Botanic Garden: an 18-acre garden with trees, shrubs and flowers from all over the world, sculptures, and a reptile and exotic plant house.

Mixing old and new, this promises to be another highly enjoyable NCAF meeting. Join us in Durham, and who knows what nuggets you will find!

**Rick Adderley**  
Editor

### INSIDE

#### Inside Pages

- Review of Sheffield meeting
- Neural Networks in the Automotive Industry
- Puzzle Corner No. 21

#### Back Page

- News from BCS SGES
- Diary Dates

---

# Sheffield Review

*NCAF  
wouldn't be  
complete without  
Puzzle Corner.  
Quite how  
Fenella took a  
problem involving  
bolts in buckets  
and made the  
pantomime  
consist of a blind-  
folded Marta  
Milo extracting  
pens from the  
trouser pocket of  
Andy Wright,  
I don't know.*

The first day's talks at the Sheffield meeting proved to be demanding and stimulating. First up was Joab Winkler (Sheffield University), whose talk might have been titled 'Everything you should know about linear algebra but didn't know to ask'. We are all familiar with the need to consider the conditioning of matrices when solving linear equations, but Joab showed us how the problem can be tackled in a more sophisticated way by considering Picard ratios. If these decay to zero, a solution can be found using the pseudo-inverse; if they grow, then small singular values can be truncated to give a well-posed solution; if they have constant magnitude, then a smoothness regulariser should be imposed.

Our main guest speaker was Zoubin Ghahramani from the Gatsby Institute. He showed how to carry out practical Bayesian inference for a large class of time series models including hidden Markov models (HMM), Kalman filters, and many extensions and combinations of these. By using a variational approach to fit a simpler distribution to an intractable joint distribution it is possible to perform the necessary integrals analytically. What makes the technique attractive is the fact that a simple extension of the EM algorithm can be used to train the models, so one gets (approximate) Bayesianism for the price of maximum likelihood. He illustrated his talk by fitting a large HMM to data consisting of sequences and showing that the model identified correctly the minimum number of states to do the job.

Kernel methods are very popular at the moment (at least with the academic community!) but one of the difficulties of applying them is how to choose an appropriate kernel function for a particular problem. Jaz Kandola (Royal Holloway College) showed how a kernel could be improved by increasing its alignment with the problem at hand; this non-parametric technique can also incorporate regularisation using support vector machine-like algorithms. Ralf Herbrich (Microsoft Research) showed us that there is still life in the good old perceptron. He has developed a special learning method that is adapted to the case where the two classes are very unbalanced. In an application to document categorisation based on Reuters news feeds (where for any given topic, only a very small proportion of the documents will be relevant) this technique significantly outperformed the support vector machine, and is much easier to implement and faster to train to boot! The final technical talk of the day was given by Manfred Opper (Aston University) who showed how methods derived from statistical physics can be used to approximate the averages needed in Bayesian inference. The advantage of this approach over the variational techniques described by Zoubin is that they work for a much wider class of models.

NCAF wouldn't be complete without Puzzle Corner. Quite how Fenella took a problem involving bolts in buckets and made the pantomime consist of a blind-folded Marta Milo extracting pens from the trouser pocket of Andy

Wright, I don't know. The remainder of the act was not suitable for a family readership, so we shall pass swiftly (unlike the coach driver) to the social event, which showed just how many NCAF members would rather drive diggers than do their day job.

## Scottish Universities INCITE

The second day had more of an application orientation. Kevin Swingler (Stirling University) talked about INCITE, an organisation set up by three Scottish universities to help university researchers exploit their results by arranging licensing. Graeme Manson (Sheffield University) showed how neural networks can be used to detect damage to aircraft wings: a novelty detection approach was used for each fault, and then the values were passed to an MLP to provide an overall classification. Lionel Tarassenko (Oxford University) brought NCAF up to date with progress on his Software Monitor, a device to provide multi-parameter patient-centred monitoring for hospital wards. He has used Neuroscale to visualise changes in patient state over time. By mapping large numbers of patients, it is possible to define a probability model of normal health and thus to set off an alarm when the probability of the measurements drops too low.

Combining classifiers to improve performance has a long history in neural computing. Bernard Buxton (UCL) showed us how genetic programming can create (non-linear) combinations that improve the area under the ROC curve, sometimes by surprisingly substantial amounts. Randa Herzallah took some well-known methods for modelling output noise (such as mixture density networks) and showed how they could be used to improve dramatically on standard inverse controllers for non-linear systems. Jim Stone (Sheffield University) proposed a different way of viewing blind source separation. By trying to maximise the temporal predictability of the separated signals (rather than their independence) it is possible to provide excellent separation regardless of the source distributions. The same approach can also be applied to extracting depth from stereo data and edge orientation from images. The final talk, from Marta Milo (Sheffield University), gave a broad overview of some applications of bioinformatics to gene expression, and a taste of where pattern analysis can be applied.

All in all, despite the indifferent weather, a very enjoyable and informative two days.

I would also like to take this opportunity to thank all those who helped to make it such an enjoyable meeting, in particular the organisers (Joab Winkler, Mahesan Niranjan, Mandy Bradley and Mark Cheeseman) the sponsors of our technical programme (Sheffield University) and everyone who attended and contributed so enthusiastically.

**Ian Nabney**  
**Aston University**

---

# Neural Networks in the Automotive Industry

How many adverts for cars these days are emphasising the benefits of their high tech, 'intelligent' content. It's not that long ago that the automotive industry was regarded as conservative towards new technology, but attitudes are changing. Technologies such as Artificial Neural Networks (ANNs) are in some of today's production vehicles providing functionality as diverse as airbag deployment and engine misfire detection.

As society becomes more environmentally aware legislation is being introduced in Europe and the US which will reduce the permitted levels of vehicle emissions. The motivation for building better engines is no longer simply market driven, it is now forced by the need for automotive manufacturers to comply with legislation. This push for automotive manufacturers to produce leaner and cleaner engines is causing them to consider new approaches.

## Active Valve Train

The most recent response to the pressures of emission legislation involves a complete change to the valve train system, i.e. the operation of inlet and exhaust valves which allow fuel and air into the combustion chamber and allow exhaust gases out. A fully flexible valve train system is seen by the automotive industry as a key driver in the future of engine performance and fuel economy. However, little in the way of flexibility is possible with the traditional approach of a rotating camshaft because the valve operating profile is fixed.

Prestige car manufacturer Lotus have been developing their approach to flexible valve trains, the Active Valve Train (AVT) system, for several years. The Lotus AVT system allows for each valve to be operated individually with varying lift profiles and dispenses with camshafts completely to allow optimisation of the gas exchange process across the whole engine speed and load range.

Lotus use an electro-hydraulic actuator to operate inlet and exhaust valves. Due to the demands for 'real-time' operation of this highly non-linear system, the control of AVT is a difficult task. The non-linear aspects of the system make it particularly difficult to define an appropriate

analytical solution, but ANNs are ideally suited to this type of demanding application.

## The Neural Networks Approach

The neural networks approach to applications such as AVT is, in comparison to analytical methods, an easy way to achieve the desired solution. Fundamentally, all that is required is for the network to be trained on data characteristics of the operating envelope of the actuator. Once data has been gathered it is used to train a network so that it learns the transfer function of the actuator. Clearly, software simulations of networks are inadequate for such a demanding application so a VindAX processor produced by AXEON Ltd was used.

The trained VindAX was used to control an AVT system where it was shown to perform well. The basic ANNs approach to this problem was achieved in a much shorter timescale and used much less resources than the traditional analytical approach. Furthermore, the resultant ANNs based system has been found to have much greater flexibility than could be achieved with a traditional approach.

## Conclusions

Due, in part, to legislative pressures and limitations of standard microprocessor technologies ANNs are now being used for demanding 'real-time' non-linear applications requiring an embedded solution. While the automotive manufacturers have previously been reluctant to embrace such technologies they are now learning that there are many benefits in doing so, including reduced time lines and resources to generate a solution. ANNs are no panacea, but when used appropriately with the right applications they frequently result in improved solutions over traditional analytical methods.

Further details of VindAX and AVT can be found at <http://www.axeon.com/> and <http://www.lotuscars.co.uk>

**Neil Lightowler**  
**AXEON Ltd**

## PUZZLE CORNER

### Number 21

"He's only got a few minutes of oxygen left," said Carl, the assistant manager. The manager of the MegaCorp Bank had already been locked in the safe for 20 minutes, and he was the only one who knew the combination to the lock. "So how do we open the lock?" asked Lisa.

"It's fully electronic," Carl replied. "It has a circular memory of M bits. You clear the memory and key in the number, N, of the bit you want to set. You press the CYCLE button and the lock continues upwards in memory setting every Nth bit. Because it's circular it doesn't stop when it gets to the end, it just carries on until it attempts to set a bit which is already set. At that point, if all the bits in memory are set, the lock doesn't open. If any bits are still clear the lock opens."

"So N=1 is always a non-starter," said Lisa. "What's the value of M?" she asked.

"We don't know the number in full," answered Carl, "but we know how to calculate it. This is a Sophie Germain Inc Model 25 lock, so M is 4 to the power 25 plus 25 to the power 4. We've already tried N=3, 5 and 7 without success and we're running out of time to try many more. Each attempt takes about a minute for the lock to fully cycle."

"Don't worry," said Lisa, "I'll have the combination for you in a few moments!"

**How did Lisa know she could find the combination so quickly, what was the number she entered, and what is the smallest number which also works?**

The answer will be given at the next NCAF meeting 10-11 September 2002, University of Durham.

**Fenella the Rottweiler**



## Chairman

Graham Hesketh  
Rolls-Royce plc

## Secretary

Mark Cheeseman  
Rolls-Royce plc

## Treasurer

Dr Andrew Starr  
University of Manchester

## Editor of Neural Computing & Applications Journal

Professor John MacIntyre  
University of Sunderland

## Managing Editor of Networks

Inspector Rick Adderley  
West Midlands Police

## Publicity Officer

Richard Ellis  
Stratum Management Ltd

Mandy Bradley  
Technical Forecasts Ltd

Dr Richard Everson  
University of Exeter

Dr Neil Lightowler  
AXEON Ltd

Dr Ian Nabney  
Aston University

Dr Kathryn Thornton  
University of Durham

Please contact NCAF  
through Mark Cheeseman  
Secretary – NCAF  
PO Box 5944  
Derby DE24 8ZD U.K.  
Tel: +44 (0) 1332 246989  
Fax: +44 (0) 1332 247129  
e-mail: [enquiries@ncaf.org.uk](mailto:enquiries@ncaf.org.uk)  
<http://ncaf.org.uk>

# News from BCS SGES

## Looking Back a Few Months and Forward a Few Years!

SGES activities have a distinct conference flavour at the moment.

ES2001 was held in Cambridge before Christmas and was a great success. Some forty papers were presented in the Application and Technical streams covering a huge range of AI domains, techniques and applications. There was plenty there for those with an interest in natural computing techniques. The first day of the conference was given over to tutorials, including a fascinating introduction to artificial immune systems from Dr Jon Timmis of the University of Kent at Canterbury, and an excellent down to earth, practical tutorial on building better neural networks by Dr Antony Browne of London Guildhall University. The natural computing thread continued in the main conference with a range of both application and technical papers on neural networks, GAs. Keynote speeches were given by Prof. Derek Sleeman of the University of Aberdeen on the re-use of knowledge and Rudolph Berger who gave a fascinating overview of the innovative work of Media Lab Europe. All this in the historic surrounding of Peterhouse College, Cambridge.

So that's looking back a few months, but what about looking forward? Well many committee members are hard at work on ES2002, to be held in Cambridge from 10–12 December. We've had approximately a hundred papers submitted, and these are currently being reviewed by our large team of referees. With this bumper crop of papers, a wide range of tutorials and workshops, the competition for Progress Towards Machine Intelligence, the usual social events and Gala Dinner, it looks like being an excellent and enjoyable conference. Full details are at [www.bcs-sges.org/es2002](http://www.bcs-sges.org/es2002) – see you there!

The keen-eyed will note when they visit the website that the Group has taken a leaf out of NCAF's book, and that we are in the process of changing our name to SGAI, the 'BCS Specialist Group on Artificial Intelligence'. More details and a full story in the next issue of *Networks*.

But looking even further ahead we are pleased to announce that a bid by SGES and the University of Edinburgh to host the 2005 International Joint Conference on Artificial Intelligence has been accepted by the IJCAI Trustees. The biennial IJCAI conferences are the premier world forum for presenting advances in Artificial Intelligence, and it is a great achievement to have been able to bring the conference back to Britain after an absence of over 30 years, thanks largely to the efforts of Dr. Rob Milne, SGES Treasurer, who has led the preparations for our bid over the last few years. The conference will be held in Edinburgh, and is timed to finish just before the Edinburgh Festival. It may seem a long way off, but with an expected attendance of some 2000+ delegates from all parts of the world, spread across dozens of streams, there is a lot of work to be done and planning has begun. It is hoped that NCAF will be able to play a full role in supporting IJCAI 2005, and it is certain that there will be plenty there to interest NCAF members. Watch this space or the SGES website for updates.

As ever, if you have any questions on SGES/SGAI or its activities, then feel free to contact me.

**Richard Ellis**  
**Stratum Management Ltd**

## DIARY DATES 2002/2003

### 2002

**10-11 September NCAF University of Durham – Data Mining**

**Contact: Mark Cheeseman e-mail: [enquiries@ncaf.org.uk](mailto:enquiries@ncaf.org.uk) Tel: +44 (0) 1332 246989**

**22-25 September Computers in Cardiology, Cook Convention Center – Memphis, Tennessee.**

Contact: CINC Committee University of Memphis, Biomedical Engineering Dept., 330 Engineering Technology, Memphis TN 38152.  
Contact: 901-678-3733  
e-mail: [cinc2002@ccmemphis.edu](mailto:cinc2002@ccmemphis.edu)

**24-26 September 'Pattern Analysis and Data Fusion' at Aston University**

Contact: <http://www.ncrg.aston.ac.uk/courses/padf02/HomePage.html>

### 2003

**January NCAF – Independent Components Analysis or Grid Computing at the DTI in London.**

**Contact: Mark Cheeseman e-mail: [enquiries@ncaf.org.uk](mailto:enquiries@ncaf.org.uk) Tel: +44 (0) 1332 246989**

**June NCAF – Safety Critical Systems at Exeter University.**

**Contact: Mark Cheeseman e-mail: [enquiries@ncaf.org.uk](mailto:enquiries@ncaf.org.uk) Tel: +44 (0) 1332 246989**

**September NCAF – Human Computer Interaction with Microsoft at Cambridge University.**

**Contact: Mark Cheeseman e-mail: [enquiries@ncaf.org.uk](mailto:enquiries@ncaf.org.uk) Tel: +44 (0) 1332 246989**

*Edited and Produced by:*  
Chris Hawthorne  
Forum Communications  
Eastgate House  
Eastgate Street  
Winchester SO23 8DZ  
Tel: 01962 877833  
Fax: 01962 877988  
e-mail:  
[chrishawthorne@forum-pr.co.uk](mailto:chrishawthorne@forum-pr.co.uk)

## NEXT EDITION

Review of Durham Meeting  
Preview of DTI London  
Meeting

## MEMBERS' CONTRIBUTIONS WELCOME

Deadline for contributions for the next edition – 4 October 2002. Please send to: Managing Editor, Inspector Rick Adderley, e-mail: [Insp1908@aol.com](mailto:Insp1908@aol.com) or [r.adderley@west-midlands.police.uk](mailto:r.adderley@west-midlands.police.uk)