



## Applying natural computing to systems engineering

**Systems Engineering Innovation Centre, Loughborough University  
19–20 September 2007**

*The aim is to produce a system that satisfies defined customer and technical requirements within cost and timeframe constraints.*

The Systems Engineering Innovation Centre (SEIC) based at Loughborough University was established in 2002 as a partnership between BAE Systems, the East Midlands Development Agency and Loughborough University. The major benefit for BAE Systems has been the achievement of significant, cost effective acceleration of its industrial capability for Systems Engineering to support its strategic objectives. Through working closely with Loughborough University academics, research associates and undergraduates, the company has an improved ability to obtain external fully and partially funded research contracts, to embed a high performance culture and to optimise value from sustainable business and cooperative working with our customers.

BAE Systems is a leading global defence and aerospace company delivering a full range of products and services for air, land and naval forces, as well as advanced electronics, information technology solutions and customer support services. It has been a keen supporter of NCAF for many years, and staff from its Bristol site organised the September 2006 meeting, with the theme of image processing.

A system can be defined as a combination of elements which form a unitary whole. These elements can be complex in nature and show emergent behaviour. Systems engineering is an interdisciplinary approach to problem solving applicable across many sectors. It spans the whole system lifecycle from concept through design and manufacture into use and finally disposal. It helps define customer needs, required functionality, design synthesis, and aids system verification and validation by considering the entire problem domain. The aim is to produce a system that satisfies defined customer and technical requirements within cost and timeframe constraints. Other important issues such as customer training, product maintenance/upgrade and product disposal are an integral part of the overall systems engineering process.

Due to the nature of the SEIC, this NCAF meeting will be entitled 'Utilisation of Natural Computing in Systems Engineering.' The wide spectrum of systems engineering will provide an opportunity for numerous specialists to contribute, and develop the system-wide point of view. As an example, utilising genetic algorithms for optimisation needs to look at the overall system rather than individual sub-systems and needs to consider the whole lifecycle from conceptualisation through design and manufacture to disposal.



Photo: BAE Systems

The Systems Engineering Innovation Centre.

It is intended that the following themes will be addressed:

- Uncertainty, be it in the models themselves, in the data, in the design concept or due to environmental effects, has an important consequence on overall system design where the preference may be for a robust system rather than a system optimised in a cost versus performance sense. The best system is usually a trade-off between several areas: system architecture and function, business value, training and personnel costs, etc.
- Autonomous vehicles including a demonstration of autonomous system interaction.
- Utilising Bayesian belief nets to determine Integrated Modelling Environment dependability.
- Utilising decision making in model selection and the ranking of training competencies for given scenarios.
- Utilising Bayesian belief networks in fault analysis, resolution and management.

Loughborough University is easy to reach, being just off Junction 23 of the M1 motorway. I look forward to seeing many of you at the September meeting to discuss this topic and view the interaction between Systems Engineers and Natural Computing aficionados. For those of you who are still undecided, I seem to remember that several NCAF members are real ale lovers and so we intend to hold the social night in an excellent real ale pub in the centre of Loughborough with a meal and a game or two of skittles. There may also be an opportunity to try some archery at the university, but only before the consumption of any alcohol.

**Steve Whittle**  
BAE Systems

### INSIDE

- Dealing with multimodal and high-dimensional data
- The case against electronic patient health records
- Puzzle Corner No. 36

### BACK PAGE

- On *Networks'* 50<sup>th</sup> issue
- Diary Dates

# Dealing with multimodal and high-dimensional data

Some of the most powerful approaches use Self-Organising Maps (SOM) to build low-dimensional models within densely populated data spaces.

The spring NCAF meeting was hosted by Liverpool John Moores University (LJMU) and held at the idyllic location of Burton Manor, near Chester. The weather also decided to join in the meeting, so it was possible to enjoy the warming sun on the terrace of the venue after lunch, whilst admiring the beautiful surroundings.

Paulo Lisboa (LJMU) was the first speaker, giving a very interesting presentation focused on two main topics, namely, visualisation and rule generation. Cluster-based visualisation can be successfully applied to different fields, such as bioinformatics or computational marketing. However, from a practical point of view, it is more appropriate to talk about segmentation rather than about clustering. In the field of bioinformatics, cluster-based visualisation was used to visualise data from breast cancer patients, when the input space comprised 25 variables. The main advantage of this visualisation with respect to other traditional ones (PCA, Sammon mapping) stems from the fact that cluster-based visualisation offers a separation index which is invariant to any linear transformation. This is especially useful in bioinformatics, since it allows a projection onto a subspace generated by cluster centres. Cluster-based visualisation also becomes very relevant in computational marketing since it is possible to visualise thousands of customers by segment profiling and followed by aggregation of several marketing segments. With regard to rule extraction, there are two main approaches: decompositional methods and pedagogical methods. The former uses the structure of the model to extract rules, whereas the latter does not care about model structure, but only about data. The aim is usually to obtain rules as simple as possible that can be then used to visualise clusters in the data.

## Self-piercing rivets

Paul Johnson of LJMU gave a 'riveting' talk about automatic monitoring of metal surface joints in Jaguar cars. This addresses a range of possible failures in self-piercing rivets, ranging from human failure to equipment wear or breakdown. The aim was to carry out a non-destructive testing of the mechanical interlocks, thus circumventing some of the current industrial problems.

After lunch, Peter Banister of Oxford University presented a visualisation of multi-channel sensor data from aero jet engines for condition monitoring and novelty detection. The idea was to carry out a model of normality and detect any variations away from this model (novelty), and then, apply this idea to a real Rolls-Royce aero engine. A feature extraction was carried out based on three fundamental operation frequencies of the engine, being the engine characterised by monitoring combinations of these features and mapping these into a two-dimensional space for visualisation. The

potential to successfully produce prognostic warning was demonstrated with an example of a test engine which lost a component in flight.

Hujun Yin of Manchester University provided a useful tutorial on dimensionality reduction and data visualisation methods. Some of the most powerful approaches use Self-Organising Maps (SOM) to build low-dimensional models within densely populated

data spaces. This talk showed that Visualisation-Induced SOM (ViSOM) is particularly suited for direct visualisation. It preserves distances or metrics locally on a nonlinear manifold. It is metric scaling rather than relational or non-metric scaling, as in the case of SOM. Kernel SOM is related to a probabilistic mixture model and can outperform SOM when its parameters are optimised. This review cannot include all the methods Hujun described; his slides were included in the meeting handout, and hopefully, will appear on the NCAF website shortly.

Sabine McNeill of 3D Metrics gave a talk related to visualising multi-dimensional data independent of scale and application. Although the data projector did not have sufficient resolution to show the new software from Visual Data Intelligence at its best, it has been used to analyse financial data, and light particles or photons. Further details can be found at [www.3dmetrics.co.uk](http://www.3dmetrics.co.uk).

## Clustering gene expression

Asoke Nandi of University of Liverpool gave a presentation about clustering gene expression, in which two clustering algorithms were presented: Self-Organising Oscillator Network (SOON) and Radius based Clustering Algorithm (RACAL). These algorithms aim to identify the optimum number of clusters, a task often made more difficult because clusters can run into each other and genes can have more than one function. The talk highlighted the need for extensive validation and repeatability measures in assessing clustering performance.

The first day ended with Puzzle Corner. Those present learned all about the mathematics of going 'all in' during a hand of poker. The majority decided that it was probably best never to play the game for money. In the evening, there was a quiz set by Graham Hesketh. A greater number of questions seemed to be on a certain Liverpool-AC Milan game than might have been expected beforehand.

The second day started with David Lowe of Aston University giving a talk on emergent behaviour in UltraDevice (UD) systems. Such behaviour is an unwanted feature of these systems as miniaturisation brings with it the unavoidable effect of mechanical cross-talk between individual oscillators. In order to illustrate the wide issues arising in Micro-Electro-Mechanical-Systems (MEMS), David showed an ingenious demonstration of coupled behaviour in mechanical oscillators already appreciated by Huygens who in the 17<sup>th</sup> century witnessed the synchronisation of two pendulum clocks at home while ill in bed. The modern demonstration included two metronomes, two cans of cider, a tea towel and hardwood base! Depending on whether a tea-towel was used to dampen the rotation of the cans, the metronomes relaxed into beating either in synchrony or in anti-phase. The main conclusion was that UD systems have inherent complexity that cannot be designed away, thus requiring a new design philosophy that directly models these complex systems in order to be able to correctly interpret the information they give out.

Gabor Barton of LJMU gave a clear introduction to the visualisation of gait patterns using SOM to analyse normal and pathological instances of gait walking cycles. SOM provides a neat solution to some of the problems related to Clinical Gait Analysis (CGA), such as the inherent subjectivity in some of the current analysis and the large amount of data that needs to be processed. In this way, it is possible to quantify the characterisation of complex dynamical high-dimensional data sets comprising angle, power and



A demonstration of Huygens' clocks. The metronomes started out of phase, but became in phase after about one minute.

# The case against electronic patient health records

Electronic Patient Health Records (EPHRs) are being promoted by the governments of many countries and health care providers as a solution to a host of information problems faced by the healthcare industry. An EPHR is an electronic version of the current paper based health record of an individual. Combined with automated systems to support personalised healthcare, and including predictive means to assess likely future risk to specific health problems, it constitutes a 'bioprofile' – a unique dynamic medical fingerprint characterising an individual's current and potential well-being.

If international distributed databases of individual EPHRs could be accessed for research purposes, significant advances in medical treatment could be forthcoming. The benefits of allowing this are evident. Indeed, members across NCAF are already working on different aspects of the EPHR and the more general bioprofile. These include natural computing solutions of biomedical data processing, data reduction, storage and transmission, and also methods of security enhancement and digital rights management of sensitive personal information. David Lowe's impromptu presentation at the Swansea NCAF meeting addressed this latter point – see *Networks* Edition 46, April 2006.

Whilst in principle, the EPHR is an excellent concept, and political pressures are being applied to push through their implementation, we should ask whether the technology is currently advanced and stable enough to support their implementation. Consider these recent relevant news stories:

The woman falsely labelled as an alcoholic by the National Health Service (NHS), 2 November 2006, from *The Guardian*.<sup>1</sup>

A woman who was falsely registered as an alcoholic in NHS records has responded by removing herself from the NHS so that her records could be deleted. Citing the Data Protection Act, Helen Wilkinson argued the NHS was causing her "unwarranted and substantial distress" by storing data on her. A senior health official stressed that the deletion of her records was a special case and did not set a precedent.

NHS security dilemma as smartcards shared, 30 January 2007, from *Computer Weekly*.<sup>2</sup>

other biomechanical patient measurements, in a way that seamlessly integrates with routine clinical working practices.

Bojian Liang of York University gave some light on the application of differential encoding to fast pattern matching on time-series data. Differential encoding enhances the high frequency components, and a filter can compensate for these enhancements. The use of differential encoding can reduce the data saturation, and supports similarity and distance measures. The use of Advanced Uncertain Reasoning Architecture (AURA) for fast approximate searching supports parallel searching and it is between 4 and 10 times faster than the conventional methods.

Karen Lees of Oxford University turned our attention to rice blast fungus. The fungal records contain angular data structures that correspond to certain patterns of expression changes. It was possible to use a wrapped distance in a Variational Bayesian Mixture Model to identify genes associated with aggressive behaviour.

Finally, Christos Tjortis of Manchester University gave an interesting talk on mining source code. This approach is especially relevant in large-scale systems where maintenance is time consuming and difficult due to the lack of adequate tools, the high cost, and complexity involved in carrying out these tasks. A

*South Warwickshire General Hospitals NHS Trust has allowed some staff to share smartcards used to access patient records, after concluding that log-in times for systems were too long for high-activity areas such as Accident and Emergency. Paul Cundy, spokesman for the British Medical Association's GP IT subcommittee, said the actions of the trust "drive a coach and horses through the so-called privacy in the new systems". He said, "This is precisely what we have long predicted and shows that security systems, although highly specified on paper, need to be tested against live environments before they can be said to be secure."*

Public distrust of large scale networks capable of providing details related to personal issues such as health is understandable when news stories such as those mentioned above occur. Without full public support and confidence, distributed personal information management systems such as EPHRs are doomed to failure. Investigation of complementary approaches to smart card based security systems and signing into health networks using personal identification numbers such as steganographic techniques currently applied to digital media, is necessary before thousands of people go through a situation similar to that of Helen Wilkinson. Organisations and individuals responsible for the formatting of health record standards need to consider the issues of security with the highest priority above all issues.

The development of the EPHR and the bioprofile is an area to which NCAF could contribute significantly, either technically, commercially or politically.

For more information on the promotion use and benefits of the individual's bioprofile concept, see <http://www.biopattern.org/>, and for a recent submission against the UK NHS's National Programme for Information Technology (NPfIT), see <http://homepages.cs.ncl.ac.uk/brian.randell/Concerns.pdf>.

**Rajeswari Matam**  
**Aston University**

1 [http://www.guardian.co.uk/uk\\_news/story/0,1937011,00.html](http://www.guardian.co.uk/uk_news/story/0,1937011,00.html)

2 <http://www.computerweekly.com/Articles/2007/01/30/221461/mhs-security-dilemma-as-smartcards-shared.htm>

possible solution to alleviate these problems is clustering to produce high-level organisations of source code. Alternative approaches can also be used at the file level, such as Data Mining Code Clustering (DMCC) and Mining Association Rules from Code (MARC). It turns out to be necessary to extract data from source code, a sort of knowledge discovery in code. In summary, a different kind of data mining, best described as code mining, was presented within a quantitative framework.

Since I am only a temporary visitor at LJMU, this was the first time I had attended an NCAF meeting. To be honest, I have to say that it was a fantastic experience. There were very interesting talks, really good speakers, and the format of the talks, somewhat longer than usual in a conference, allows a deeper description of the problem and interesting discussions. The conference venue was superb and overall, the meeting was a great success. Unfortunately, the evening football match did not offer the desired continuation – but before losing in the final, you have had to get there. This is my only consolation since Valencia lost two consecutive Champions League finals a few years ago!

**José D. Martín-Guerrero**  
**Visiting Research Fellow,**  
**Liverpool John Moores University**

Modesty Blaise, still smarting from the lesson Lisa handed out in Puzzle Corner 27, tried to entice Lisa with her probabilistic dice version of Rock, Paper, Scissors (RPS). Her red die had the numbers 144444, the purple die was 222555 and the silver die was 333336. Modesty and Lisa would each pick a die and then roll them against each other, the highest number winning. If Lisa had to choose her die first, Modesty would always have the edge. However, Lisa again turned the tables on Modesty by allowing her to go second so long as Lisa could then specify whether each game was decided by just one roll of the dice or the total of two rolls.

Somewhat taken aback, Modesty then produced her pièce de résistance, a three-sided RPS-type game from a pool of seven special dice. Each die had letters on the faces: die 1 was CRVCRV, die 2 was JMTJMT, die 3 was GQSQSQ, die 4 was DLZDLZ, die 5 was BPYBPY, die 6 was HKXHKX, and finally die 7 was FNWFNW. Modesty introduced her friend, Pascal, and the three of them would play this new game, where the winner would be the one showing the highest letter in the alphabet.

"Do you want to choose your die first, second or last?" Modesty asked Lisa. "That depends", Lisa replied, "on whether the games are going to be head-to-head pairings or all three of us at once." Pascal sagely nodded and concluded that Modesty was never going to outsmart Lisa the Dice-Demon.

*What was Lisa's strategy for head-to-head games and how would it change if all three dice were rolled at once? Is it possible for Lisa to be the underdog in both head-to-head matches whilst coming out on top in the three-way version?*

The answers will be given at the next NCAF meeting (19-20 September 2007, Loughborough University).

**Fenella the Rottweiler**

## COMMITTEE NOTES

### Chairman/Treasurer

Graham Hesketh  
Rolls-Royce plc

### Secretary

Jo Thomas  
Rolls-Royce plc

### Editor of Neural Computing & Applications Journal

Professor John MacIntyre  
University of Sunderland

### Managing Editor of Networks

Dr Nick Granville  
Technical Services Group  
Smith & Nephew plc

Professor Ian Nabney  
University of Aston

Dr Andy Pryke  
University of Birmingham

Dr Rajesh Ransing  
University of Swansea

Dr lead Rezek  
University of Oxford

Please contact NCAF through Graham Hesketh, Chairman – 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://www.ncaf.org.uk>

## MEMBERS' NEWS AND VIEWS

Deadline for contributions for the next edition – 1 November 2007.  
Please send to Managing Editor – Nick Granville, e-mail: [Nick.Granville@smith-nephew.com](mailto:Nick.Granville@smith-nephew.com)

*Edited and Produced by:*  
Forum Public Relations  
Westgate House, Old Ivy Lane  
West End, Southampton  
Hampshire SO30 3RX  
Tel/Fax: 023 8047 6888  
e-mail: [info@forum-pr.co.uk](mailto:info@forum-pr.co.uk)

### NEXT EDITION

Review of Loughborough meeting  
Preview of Aston meeting

# On Networks' 50<sup>th</sup> edition

## – a personal view from the Chairman

NCAF started nearly 17 years ago. It began as a support group for users of 'NeuralWorks', a popular neural network package, and it had its first meeting in Southampton in 1990. So it had already been going a few years before I was invited along to my first meeting in the summer of 1993 at Fitzwilliam College, Cambridge. And what an eye-opener it was! Really talented, entertaining people, fully engaged in all the activities, contributing everywhere in a common spirit of adventure in a newly developing field. Chris Bishop (then Chairman, and my boss at AEA Technology) said I would enjoy it, and he wasn't wrong.

At the next meeting in Brunel I met Ila Patel. The Committee had been in consultation with Chris Hawthorne at Forum Public Relations, a firm engaged in publicising and promoting NCAF's activities, and had decided to run a quarterly newsletter to coincide with our regular meetings, and Ila was to be the Managing Editor. *Networks* was born, and the first issue went out in November 1993.

Ila was always keen to get good copy for the newsletter and she very pro-actively encouraged all of us to contribute where possible. Reviews of books, software and conferences frequently appeared alongside the regular preview of the upcoming meeting and review of the past event. Then came the moment that forever changed my involvement with *Networks*. Ila, desperate for copy for issue 15 (May 1997) asked me to do a small piece. My contribution took half an hour – Puzzle Corner. I had no idea how popular it would become, but the pantomime solution enacted at the Cambridge meeting that Autumn sealed its fate, and it has been a regular piece in *Networks* and meetings ever since.

In May 1999 the NCAF web pages made their appearance, and the organisation started to move to a more electronic format, for which we owe a great deal of thanks to Mark Cheeseman, then Secretary. 2000 was an auspicious year. In March we changed the name from 'Neural' to 'Natural', in line with the broadening of our membership's interests, in July Peter Cowley stepped down and I was invested as Chairman, and in September we celebrated our 10<sup>th</sup> anniversary.

In March 2001 Ila eventually stepped down as editor and handed the reins to Rick Adderley, who did a sterling job until handing over to the current editor, Nick Granville, in March 2004. All three deserve our thanks. Indeed, there are so many

people to thank, but so few column-inches. Suffice to say we are indebted to both the Engineering and Physical Sciences Council (EPSRC) and DTI for their continued support over a long period, and apologies to all those worthy people I have singularly failed to thank.

On a personal note, I have thoroughly enjoyed every moment of my involvement with NCAF. It has been illuminating, entertaining and, on occasion, embarrassing (for me) and I regret only that I have missed three meetings since my induction 14 years ago. 2007 is not only *Networks*' 50<sup>th</sup> issue, it also brings up the big 5-0 for both Nick Granville and myself, and next January will be my 50<sup>th</sup> NCAF meeting! Documenting almost that whole period has been the function of *Networks*. Like a comfortable friend it reminds absentees of what they are missing and beckons us to join in the next event.

I would like to dedicate this 50<sup>th</sup> issue to Chris Hawthorne. He has acted as editor and publisher of *Networks* for most of its lifetime, and it is only through ill health that he has had to hand over responsibility to his wife, Sally, who continues to run Forum. Our thanks and best wishes go to both of them.

**Graham Hesketh**  
NCAF Chairman

## DIARY DATES 2007

**11-14 November** – ANNIE 2007, Artificial Neural Networks in Engineering. St. Louis, Missouri, USA.  
<http://web.umn.edu/~annie/annie07/information.htm>

**3-6 December** – NIPS 2007, Neural Information Processing Systems. Vancouver, Canada.  
<http://nips.cc/>

**10-12 December** – AI-2007: 27<sup>th</sup> SGAI International Conference on Artificial Intelligence. Cambridge, England.  
<http://www.bcs-sgai.org/ai2007/>

**Mid-January** – NCAF meeting (theme: tba) at Aston University. For information, email [enquiries@ncaf.org.uk](mailto:enquiries@ncaf.org.uk) or telephone +44 (0)1332 246989.

**11-13 February** – AIA 2008, Artificial Intelligence and Applications. Innsbruck, Austria.  
<http://www.iasted.org/conferences/home-595.html>

