Networks 27 – June 2000

Welcome Return to Manchester

NCAF at Manchester University, 5 - 6 July 2000

The Summer 2000 meeting sees a welcome return to the University of Manchester. It is to be sincerely hoped that the weather we enjoyed on our last visit (April 1997) is repeated this year, rather than the traditional inclemency served up by the Mancunian skyline.

The meeting takes the established format of a themed day and a general papers day, but this time the order is reversed. Having the general day first allows us to include an introductory tutorial on data fusion as a gentle precursor to the more detailed presentations of the second day.

Of course, the main attraction of the first day is the long-awaited keynote talk by Professor Chris Bishop on probabilistic graphical models. This presentation was scheduled for September last year in Cambridge, but Chris was unable to give it due to his well-publicised court appearance (on the jury!).

The second day is a joint NCAF/IEE colloquium on Data Fusion, and is introduced by Professor Lionel Tarassenko from Oxford University who says:

"The aim of fusing data from multiple sensors or knowledge sources is to enhance the intelligence of a diagnostic system. Data fusion not only delivers improved system reliability and robustness but also provides more information than the analysis of each source of data on its own. The combination of multiple independent measurements of the same quantity leads to reduced uncertainty and improved resolution. Different sensors, on the other hand, cover different domains of the parameter space and their fusion provides more complete information leading to reduced ambiguity and better discrimination between available hypotheses. In some cases, certain phenomena may only be detectable by combining the information from different sensing modalities. The classical approach to data fusion over time has been to fuse the sensor measurements into a system model, usually within a Kalman filter framework. There is not always sufficient knowledge to derive realistic system models and so data- and information-driven approaches (for example, neural networks or Bayesian inference networks) are being developed as alternatives."

The speakers will be covering a variety of issues, both theoretical and practical, with some interesting, diverse applications. A Chinese meal with a conjurer (hold onto your chopsticks) rounds off the programme. With IEE members eligible to attend at the NCAF members' rate, this should prove to be a very popular meeting. Check out the NCAF web site (www.ncaf.co.uk) for programme details and the new, improved on-line registration process.

Graham Hesketh
Rolls-Royce plc

Neural and more at DTI 2000

A review of NCAF at DTI, London, 26 - 27 April 2000

The Spring meeting of NCAF, the first to be held in London for some time, was an interesting mix of traditional neural computing talks and applications using a broader spread of smart computing technology. Once again our host Ray Browne introduced the forum and welcomed the audience.

Ian Lang from Assistum gave a demonstration of an intelligent decision support tool that has been used to develop systems that support developers in deciding on the choice of smart technology for an application, give business advice on outsourcing, and aid gas and oil well development. The key features are fuzzy logic, visualisation (showing links between topics and questions) and linguistics (to provide a reasoning capability).

This was followed by an excellent presentation by Tom Bayes from Satra, which guided the audience through the complexities of cutting leather from hides for shoes (complete with demonstration hides). Computer-based visualisation allows skilled cutters to lay-plan the cuts for a whole hide but the automated algorithm for lay-planning requires 3 hours (for a single skin) to do better than a cutter. Tom is working with genetic algorithms to speed this up to provide a good solution in about 5-6 minutes. Early results are promising.

Giles Oatley, from the University of Sunderland, in the first of two presentations, showed how repeat victimisation is profiled using crime data from Cleveland Constabulary. He hypothesised that people who have suffered a crime are likely to suffer another, possibly within a short time. This has important consequences in policy for policing strategy, crime prevention and victim support. Neural network tools have been used to predict the likelihood of such victims.

Rajan Amin from ERA presented a session on clinical data analysis on behalf of the Smart Software Technologies Club. His paper on the prediction of ulcer healing dealt with the problem of choice of treatment based on patient parameters. He showed how a neural network model improved the performance of prediction over the already effective linear model.

Giles Oatley, in his second presentation, showed how visualisation and clustering could be used to detect business clusters and determine what are the common features of different clusters. His work showed that common conceptions, such as a high-tech M4 corridor, were not borne out by the data. A similar application, involving detecting patterns in criminal activity, was discussed by Chris Kirkham from Brunel University. A particularly difficult obstacle is the use of large amounts of free text in crime reports; this is analysed by extracting key words and encoding them as binary variables. Unfortunately, this seems to lose a lot of contextual information which provides evidence for a criminal's modus operandi.

Getting results

The first day ended with a lively panel discussion on the theme 'Getting Results with Smart Technologies'. As usual, discussion became increasingly animated the closer we got to the end of the session. All the panellists emphasised just how hard it was to turn a good idea into a successful product, and that building up awareness of success stories is key, particularly through the use of good reference sites. The technology will not sell itself. In fact, its use is often not a feature of a product selling; it is the benefits to the customer that matter. Tom Harris said that Cardionetics Ltd had spent about half their funds on market analysis and had changed the aims of their product a lot since the initial final year student design project that they had started with. The day ended with a high flying expedition on the 'London Eye', giving spectacular views only bettered by British Airways' more elevated forms of transport.

On the second day, Mark Plumbley of Kings College explained the links between information theory and learning in neural networks, particularly for unsupervised networks. The problem of data processing can be formulated as a communication problem, and learning then becomes the extraction or reconstruction of information. This gives a framework for analysing many different learning algorithms, including Principal Component Analysis (PCA) and Independent Component Analysis (ICA).

The keynote talk from Professor Geoff Hinton, Director of Gatsby Computational Neuroscience Unit, continued the theme of unsupervised learning. Mixture models (for example, mixtures of Gaussians) are commonly used to model data density or for clustering. Professor Hinton proposed used products (instead of sums) of experts. This has a profound impact on the characteristics of the model: experts are AND'd together rather than OR'd. This means that the model is more precise about the location of features than individual experts, which is important in high-dimensional feature spaces (such as images). There being no such thing as a free lunch, training these models is more complex (and slow) than the EM algorithm for mixtures, but the results in digit and face recognition are impressive.

Curvaceous Software

Robin Brooks from Curvaceous Software gave an interesting demonstration of multivariate inspection of process data. In this technique all the data is plotted at once in 'parallel co-ordinates' and individual samples are linked by a line. The pattern of the lines is quickly learnt even by unskilled or less numerate operators as a kind of signature, without having to understand mathematics or data processing. Application examples showed how patterns of control data showed clear empirical constraints in real, operational inputs in order to satisfy output requirements.

Mark Nixon from Southampton University gave an outstanding multimedia presentation, revealing his TV stardom and proving that the 'Ministry of Silly Walks' can now identify you. Biometrics are big business for security and surveillance operations. As well as fingerprinting, voice recognition and retinal characteristics, humans also have distinct gaits when they walk. The measurement of the position of the legs and joints is by no means trivial, but the data fits good cyclical models. In many instances the gait is more recognisable than the face, e.g. a bank robber takes care to disguise himself, but cannot change the way he walks.

Ray Frank from the University of Hertfordshire showed an application of neural networks in clone and fraud detection in telecommunications. 'Cloned' blocks of code appear in software systems which have evolved over many years. If one section of code needs changing, then many others may also need attention. Input data based on keywords, indentation and line length were fed to a self-organising map, which successfully identified the clones. In the use on cellular telephones, fraudulent use is an expensive and increasing problem. Fortunately, cellular phones leave a trail of information behind them. Some patterns are easily detected because they almost never happen in normal use, but a neural solution has been used to identify patterns in the data which show call transactions are anomalous relative to historical use.

The final talk, given by Chris James from Aston University, was an application of component analysis to EEG for epilepsy detection. The aim is to determine the location of spikes or the onset of seizure to aid surgeons. In the longer term there is the possibility of early warnings of seizure so that pacemaker-like implants can prevent the attack. EEG is hard to analyse as the levels of noise are high; for example, muscle movement such as eye blinking or gum chewing (mainly an North American problem!) give much larger signals than those of interest. Traditional methods remove the noise by averaging signals, but this makes the detection of local events (such as spikes) inaccurate. In this work ICA has been successfully applied to the original data to isolate the correct number of components in data simulated from numerical models, and gives plausible results on real data. This gives the promise of much more accurate event detection in the future.

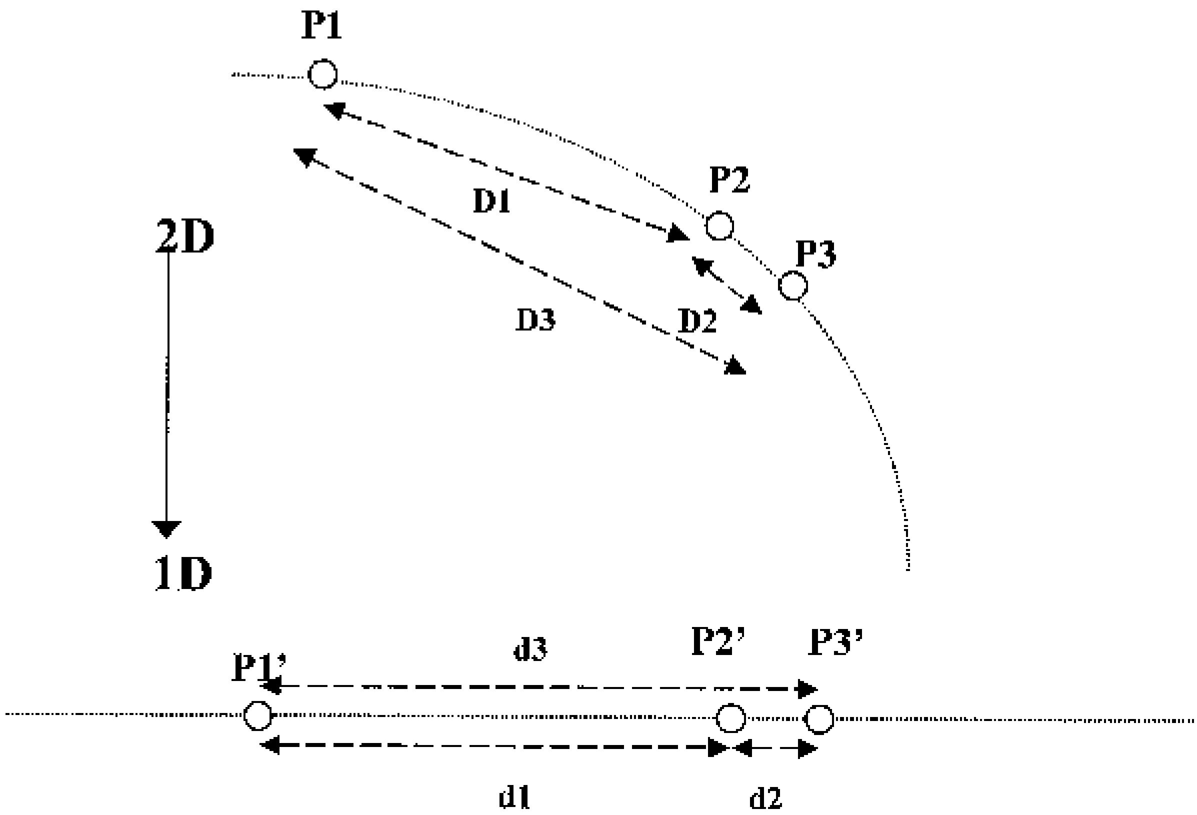
The DTI proved an excellent venue, and we thank Ray for his hospitality. Visiting Westminster always gives one a flush of pride, reminding us that our taxes are well spent, and that a visit to London really is worth triple the train/air fare! See you at the next meeting in grim but economical Manchester.

Ian Nabney
Aston University
(aided and abetted by Andrew Starr, Manchester University)

An overview of the Sammon mapping technique

The successful outcome of using neural computing techniques for any problem will depend on many factors. However, one important criterion is having a good level of understanding of the data and its origin. Often, we find that the data represents a high-dimensional view of a process. For example, in plant control and monitoring applications it is not unusual for the state of the plant, at any given time, to be represented by several hundred parameters (pressures, temperatures, fuel flow, etc). Clearly, it is important to be able to make decisions based on this data by describing its structure in engineering terms. The most trivial approach would be to construct a family of graphs each displaying the behaviour of pairs of parameters. A tedious solution to say the least. Ideally, we would like to use a visualization technique that can display the whole data in just two-dimensions, whilst still preserving its underlying structure in a meaningful form. There are a number of techniques that attempt to do just that. For this short review, we will consider the role of 'Sammon mapping'.

Sammon mapping is a non-linear technique that attempts to project points from a high-dimensional space to points in a lower-dimensional representation (usually 2-D). Unless there is a high degree of redundancy in the co-ordinates of the data, we can never hope to achieve a perfect projection without suffering a degree of distortion in the projected representation. Sammon mapping attempts to preserve as much of the original data structure as possible by positioning points in the lower- dimensional space such that the distance between any two points is a close approximation to the distance between the two corresponding points in the higher-dimensional space. This is more easily seen in the diagram below.



In this example, the points P1,P2, P3 along the two-dimensional curve are to be projected onto the horizontal line. As can be seen, the aim is locate the points p1',p2' and p3' such that the error between the interpoint distances (D1,D2,D3) on the 2D curve and their counterparts along the straight line (d1,d2,d3 respectively) is a minimum.# The process for achieving this projection from a high-dimensional space (H-D) to a lower-dimensional representation (L-D) is as follows:

1. Calculate all interpoint distances in H-D space. Note, there will be n(n-1)/2 of them, where n is the number of points to project.
2. Generate n random points in L-D space, Let's call each projected point yi
3. Calculate the mapping error over all the interpoint distances in L-D space. We shall call this error E.
4. If E is less than a pre-defined threshold, or the number of iterations through this loop exceeds some arbitrary count, then stop
5. Adjust the co-ordinates of the points in L-D space using a function of the form yi' = yi - f(E,yi)
6. Go to step 3

The choice of mapping error E usually employs a Euclidean distance measure such as the one shown below, although any distance measure could be used.

$$E = \frac{1}{c}\sum_{i<j}^{N}\frac{\left[d_{ij}^{*}-d_{ij}\right]^{2}}{d_{ij}^{*}} \qquad \text{Where} \quad c = \sum_{i<j}^{N} d_{ij}^{*};$$

d*ij is the distance between points i and j in H-D space;
and dij is the corresponding interpoint distance in L-D space.

Step 5 of our algorithm requires the current set of projected points to be revised based on the current mapping error. The function f that achieves this uses a steepest descent procedure in an attempt to search for a minimum error. We will not go into any detail as to the proof or form of this function in this short review, except to say that the reader should be aware that the function involves division operators incorporating interpoint distances. This highlights the first limitation of the Sammon projection; namely in order to avoid such errors, care must be taken to ensure that duplicate points are not included in the data set. If there are duplicate points, then only the first instance should be included in the projection. The projected point may then be replicated, as many times as necessary, after the algorithm has completed. However, there will be a choice in how the mapping error is calculated based on the value of N. This decision is largely based on whether the aim is to preserve the geometric or data density structure.

Since the above algorithm depends only upon interpoint distances, there is no restriction on the width of the input dataset. There are, however, issues of memory requirements which relate to the number of points in the input data. For a projection involving 100 points approximately 19Kbytes of memory will be required (assuming 4byte words). For 200 points approximately 77Kbytes is required, and for 1000 points approximately 2Mbytes is required. When Sammon published his paper in 1969, the algorithm had been developed on a GE-635 computer possessing only 128K or core memory! This imposed an upper limit of 250 vectors. Clearly, with more available memory in modern computers, this limitation no longer applies. However, care should be taken when processing large datasets; it may be necessary to incorporate a data compression technique (such as clustering) as a pre-processing step.

Steve King
Rolls-Royce plc

PUZZLE CORNER

Number 13

The Brothers Carry-Matsoff, triplet heirs to a multinational family business, were summoned to the boardroom by the company solicitor.

"I have some good news and some bad news", he said. "Our beloved President, your father, is dead", he continued.

"And the bad news?" laughed Amoral, the first born.

"The company is to be equally divided amongst all his sons", said the solicitor, "and it is up to all of you to agree on a mutually acceptable disposition."

This seemed like a tall order, as Amoral, Belligerent and Callous were fanatically selfish and would never agree to either of the others receiving more than themselves. Fortunately, Lisa was on hand to provide a solution (for a very modest fee).

However, much to the chagrin of the obnoxious threesome, a group of the President's extra-marital offspring arrived to claim their fair share of the spoils, complete with impressive DNA credentials.

"Is this the lot, or are there any more b******s out there?" said Callous. Lisa was unfazed by the developments, and promptly suggested a simple procedure whereby, once they had finally decided how many shares they needed to make, everyone would be guaranteed to be satisfied that they had received a fair share or better.

How did Lisa provide a general procedure for guaranteeing a universally acceptable partitioning of the company? The answer will be given at the next NCAF meeting (5-6 July 2000, Manchester University).

Fenella the Rottweiler

Book review

Artificial Intelligence: A New Synthesis. (Author: Nils J. Nilsson. Publisher: Morgan Kaufmann.)

Artificial Intelligence: A New Synthesis is a comprehensive introduction to the world of Artificial Intelligence (AI). The author introduces his subject with a discussion of one of artificial intelligence's most fundamental questions, asking if machines can ever truly be said to think. He then attempts to assess to what extent AI has reached its long term goals and examines the various paradigms that are being used to achieve those objectives.

After providing the reader with a brief history of AI, the text covers ground that will be well known to those readers with a previous knowledge of artificial intelligence. Stimulus-response agents, neural networks, machine evolution, genetic programming, state machines, robot vision, planning agents, search techniques, action policies, logical languages, knowledge-based systems, knowledge representation, reasoning with uncertain information, planning methods and communication and integration are all explored in the book's wide ranging material. While the subject matter has been frequently explored in other books on artificial intelligence, what makes Nilsson's work different from other introductory texts on AI is the originality of his approach to the material.

The author describes this approach as 'evolutionary artificial intelligence' and it is this which informs the organization and presentation of his material. The author makes the evolution of intelligent agents the central theme of the book, charting their development from simple reactive agents to artifacts of great complexity and power. Nilsson skillfully uses this evolutionary paradigm to provide the context for an exploration of artificial intelligence's most significant developments and ideas, using it to explore the concepts that form the basis for the development of intelligent machines.

Virtually anyone involved in the teaching of artificial intelligence will find Artificial Intelligence: A New Synthesis an extremely valuable resource. The text is extremely well illustrated with instructive diagrams and examples. It boasts over five hundred references and a comprehensive index. Each chapter offers suggestions for additional reading and topics of discussion and concludes with a set of exercises for which a solutions manual is available in PDF format via a secure FTP site.

The student of artificial intelligence will also find much of value here. AI's most important and lasting themes are explored in considerable detail. It is a work that reflects the author's deep knowledge of his subject. The extent of Nilsson's expertise will come as no surprise to those familiar with his previous work in this area. It is no less than one would expect from the author of Principles of Artificial Intelligence and the co-author of Logical Foundations of Artificial Intelligence. However, it is that depth of knowledge that casts some doubt on the text's suitability as an introductory work.

It must be observed that as an introduction to the field of AI, Artificial Intelligence: A New Synthesis is not for the faint-hearted! The rapidity with which the author moves from the elementary to the advanced might well prove daunting to the newcomer to the field of artificial intelligence. The author's style of writing is compact and concise, making few concessions to the reader.

Artificial Intelligence: A New Synthesis, is a dense work which asks more from its readers than many introductory texts. Nevertheless, those readers who are motivated to study AI by a genuine interest in the subject will have their knowledge of artificial intelligence greatly enriched by Nilsson's work. While Artificial Intelligence: A New Synthesis is not the easiest of introductions to the subject, it is a substantial and rewarding work with much to offer the serious student.

John MacIntyre
University of Sunderland

Members News and Views

What new techniques do you want to know more about? Please tell us if there are any new techniques that you know about or would like to know about so that these can also be included in Networks.

DIARY DATES 2000

July 5-6 NCAF, University of Manchester. Contact: Safia Hussain, Tel: (0)161 275 4315, e-mail: ncafsec@man.ac.uk

July 5-7 DM2000 2nd International Conference on Data Mining, Cambridge. Tel: 02380 293223 fax: 292853, http://www.wessex.ac.uk/conferences/2000/data2000, e-mail: wit@wessex.ac.uk

July 17-19 EANN2000 Engineering Applications of Neural Networks, Kingston on Thames, Tel: 0208 547 2000, fax: 547 7497, http://www.kingston.ac.uk/eann, e-mail: eann2000@kingston.ac.uk

July 24-27 IJCNN International Joint Conference on Neural Networks, Como, Italy, http://www.ims.unico.it/2000ijcnn.html

August 20-23 KDD00 Knowledge Discovery and Data Mining. Boston, USA http://www.acm.org/sigkdd/kdd2000

August 20-25 14th European Conference on AI. Berlin. http://www.ecai2000.hu-berlin.de/

August 29-September 1 ISCI 2000 International Symposium on Computational Intelligence Kosice, Slovakia. http://neuron-ai.tuke.sk/cig/isci or http://cns.bu.edu/~kopco/isci

September 3-8 International Conference on Pattern Recognition. Barcelona http://www.cvc.uab.es/icpr2000/

September 11-14 British Machine Vision Conference. Bristol http://www.cs.bris.ac.uk/Events/BMVC2000

September 26-27 NCAF, University of Southampton

Contact: Safia Hussain +44 (0)161 275 4315, e-mail: ncafsec@man.ac.uk

November 5-8 Artificial Neural Networks in Engineering. St Louis, Missouri, USA http://www.umr.edu/~annie/annie2000.htm

December 11-13 IEEE Workshop on Neural Networks for Signal Processing. Sydney http://www.eivind.imm.dtu.dk/nnsp2000

NEXT EDITION

Review of the University of Manchester meeting
Preview of the Southampton meeting (10th anniversary)
Deadline for contributions for the next edition – 21 July 2000