

The Ring

The Newsletter of the University of Cambridge Computer Laboratory Graduate Association

Governing Council: Chairman: Dr Ian Pratt (K92)

Members: Stephen Allott (T80), David Colver (CHR80), Peter Cowley (F77), Richard Mason (Q84), Lorenzo Wood (CHR93)

Careers Committee Chairman: Peter Cowley

Members: Youssef Bouguerra (PEM98), Nathan Dimmock (JE01), Richard Hadden (Q97), Richard Mason

Ring News

There have been a few changes recently so there are some welcomes and goodbyes to make.

Following his appointment as Pro-Vice-Chancellor with responsibility for research (see *The Ring Issue VI*), and Professor Andy Hopper's subsequent appointment as Head of the Computer Laboratory (see page 14), Professor Ian Leslie has stepped down as Chairman of the Ring. Many thanks to Ian for sponsoring the start-up of the Ring and for his support and commitment.

And, a warm welcome to Dr Ian Pratt who takes his place. Ian is a Senior Lecturer at the Lab and a Fellow of King's College. He also leads the Systems Research Group.

Stephen Allott, Ring Council Member, has set up a London Ring Committee to organise more London events. Youssef Bouguerra, Tom Tanner and Matt Wiseman have joined this committee. Please contact Stephen if you would like to help.

Ring Careers Committee Report

The Ring Careers Committee has been set up to provide assistance to members in helping them achieve or enhance their career potential. We do this by providing information on the Ring website, telephone or email advice from

other members, who have volunteered their services, and shortly a mentoring scheme.

There are 10+ positions currently on the job bulletin board and we would welcome more – this is free advertising to a group of highly qualified people. We have provided careers advice to some Ring members, and would like to increase the usage of this – again this is at no cost to our members and within the Ring, we have a huge breadth and depth of experience. There is a separate article about our new mentoring scheme – please get in touch with Richard (*see article below*) if you wish to get involved.

Our committee was officially launched at the drinks party in June, where I met many interesting people. I look forward to meeting more of you at forthcoming Ring events.

Peter Cowley, Careers Committee Chairman

Could you be a mentor?

September will see the pilot launch of our mentoring programme. Ring members will be paired up with experienced mentors who will assist them in their professional growth and development. Fifty-eight Ring members have already expressed an interest in becoming a mentor, but we're still keen to find more. Do you have what it takes? A successful mentoring relationship involves regular meetings as the pair "gets to know one another" and builds mutual respect and understanding. The

mentor then uses their wider experience to guide the "mentee", helping them to overcome development hurdles and realize their career potential.

As a Ring mentor you will enjoy seeing your mentees grow in stature, and in turn hone your own mentoring skills. You can take pride in having contributed to the community of Cambridge Computer Laboratory graduates. The Ring uses the power of its network to find suitable matches, and has a rigorous selection and oversight process to ensure the success of the programme. If you think you have the experience and passion required for this volunteer position, contact **Richard Mason** (www.mason.ca), programme director, for more information.

Ring Events

RAUCOUS LONDON DRINKS PARTY AT OYSTER'S COOL OFFICES

15 June 2004

Report by Stephen Allott,
Member, Governing Council,
Cambridge Computer Lab Ring

50 people crowded into the playroom at Oyster Partners' cool offices in Clerkenwell for the 2004 Ring London Drinks party. Despite an official end time of 19:30, people were still there at 21:00 and they still didn't want to leave.

Oyster, one of the UK's leading web consultancies, have real dotcom offices. A huge playroom, with sofas, pool table, ping pong and bean bags, was the venue for the party. The playroom is attached to their staff restaurant which forms most of the ground floor of the Oyster building. The loft style high ceilings, probably 25 feet high, were useful on what was a hot day in London. The fridge was working hard to keep the beer, water and white wine cold but everything was consumed by the end of the party.

Here is the news from some of the people I chatted to. **Lorenzo Wood**, who is from Oyster and who arranged the use of their offices told me that they have tons of work at the moment.

Matt Wiseman from McKinsey's Business Technology Office is also very busy. **Frank Wilson** has returned from almost 20 years abroad in the US and Germany and is now job hunting in the UK. **Andrew Wallace** had come down from his home in Yorkshire for the party. Formerly the Marketing Director for Pace Electronics, Andrew is now doing consultancy for some leading UK consumer electronics firms. **Thomas Tanner** from Bloomberg was in good form. **Glen Slade** is excited about his new start-up StegoStik which has nearly closed a round of seed funding. StegoStik make an elegant encrypted storage device for PCs. **Rend Shakir** is also excited about her new start-up, Cambridge Matrix, who do mesh wireless networking and was chatting to **Faisal Ahmed** who is a broadband expert.

Kamiar Sehat from chip designer ARC was talking to **Andrew Bower** from Velocity Semiconductor about the chip industry. **Peter Radford** from LogicaCMG said that the amount of work for his telco software group has recently picked up and that Logica are recruiting again. **Chris Oswald** from communications software firm Equisys was there. We had held the 2003 London drinks in Chris's office. **Giles Nelson** from Apama has just won several new customers in the City for his real time correlation software and he talked for some time to **Martin Coen** who has just moved into a new sales role in a City software firm. **Ewan Mellor's** games company has a brilliant new approach to games development and he was talking to **Adam Martin** of Grex Games who have a growing order book for their multi-player platform. **John Mannix** of Governor Technology also has a growing order book from some of the world's largest computer companies which is most impressive. **Roderick Lumsden** had a long chat to **Richard Jebb** who is organising the awards for the 2005 Ring Annual Dinner.

Nigel Horne was chatting to several younger members about their new business ideas, as well as to Russell Haggart who is a VC with Prelude in Cambridge. Richard Hadden, also a VC with NESTA, has been very busy and is pleased with his investment in Tim Coote's company Tideway Systems. Vishal Doshi is working on printing software at TTP. Peter Cowley is just about to move to Cambridge and talked to lots of people about the Ring Careers Committee which he has recently taken over. David Colver brought along colleague Aaron Lee from Operis. Pete Chown is also moving to Cambridge and has exchanged contracts on a house. Pete had sold some of his technology earlier that day and was very upbeat. Pete Campton from CSFB is working on risk management systems. Nick Bolton of HRM Software is winning plenty of new business and is hiring new people.

That's some of the gossip I managed to pick up. Hermann Hauser, our guest of honour, said that Cambridge had come through the IT recession pretty well and now looked set fair for further growth. Hermann's companies have some 100 daughter companies and he has done as much as anyone for the Cambridge technology business cluster. He said that he thought the Ring was a great organisation and he was keen to hear from members with good business ideas. Hermann has become an Associate Life Member of the Ring and we are grateful for his support of the Ring.

GROWING FROM 2 to 758 PEOPLE

Lessons learnt from the growth of
Sophos Plc

Speaker: Dr Jan Hruska, co-founder

20 April 2004

Report by Stephen Allott
Member, Governing Council,
Cambridge Computer Lab Ring

Sophos Plc is currently top of the Ring Hall of Fame league table for business success. Co-founded by Dr Jan Hruska,

(who graduated in 1978 from Downing having done the Part II Computer Science course), Sophos is now a world leader in anti-virus software and is head-quartered at Abingdon, near Oxford. A large audience of Ring members, Lab members and outside guests gathered in Lecture Theatre 2 in the Lab to hear Jan's charming, entertaining and inspiring account of the company's success over the past 19 years.

Sophos Plc today

Founded in 1985 by Jan and his business partner and co-CEO, Peter Lammer, Sophos employs 758 people and exports to 150 countries. The company sells anti-virus software to the enterprise market via channels. In the UK, 70% of the FTSE 100 are customers. They do not sell to the consumer market. There are 8 foreign subsidiaries and development laboratories at HQ in the UK, in Boston and in Sydney to provide 24 hour around the world technical capability.

Turnover in financial year 2003 was £41m with profits of £12m, a handy operating margin of 29%! Headcount in 2003 grew by 35% year on year. Sophos has been increasing market share at the expense of competitors. Exports are 59% of turnover and growing. Sophos is a highly profitable specialist software company which has grown almost exclusively organically and which is export orientated. Sophos is the largest anti-virus vendor in Europe and the 4th largest globally. A true success story and one with accelerating growth. How did it come to be where it is today?

The early days

After being an undergraduate at Cambridge, Jan moved to Oxford to do a PhD in medical engineering. This required data capture on portable devices so Jan, and his future business partner Peter, launched their own portable computer company, Executive Computers Ltd, to provide them with the necessary hardware for their research. Jan brought along the prototype and showed it to the audience.

He had raised £100,000 of VC from an angel investor to fund the development. They were heroic days. The entire machine was designed using CAD, still then in its infancy. Jan regularly flew from Oxford to Essex to use the CAD equipment. They had to redesign their machine rapidly when their chip supplier changed specification. The ECL laptop had 256k of RAM compared with 256 Meg of RAM on his current machine, a growth of 3 orders of magnitude in just over 20 years. These early years however were unsuccessful and the project was wound up in 1983.

This led to Jan's first lesson:

"Failure precedes success".

Jan was then conscripted into the then Yugoslav army before coming back to the UK, re-uniting with Peter and, in 1985, founded Sophos using an early microcomputer, the DEC Rainbow. Although Jan then had a dog called Sophie, contrary to reports, this was not the inspiration for the company's name, which comes from the ancient Greek word for wisdom. Sophos was formed to enable multiple users of a single PC to keep their files secure from other users. It was a data security rather than an anti-virus company. They wrote their own DES, RSA and SPA (Sophos Proprietary Algorithm) and did secure erasure of files.

In 1987, one of the first crucial steps for the company came from their Norwegian distributor suggesting they develop a vaccine against some new computer viruses which had just appeared. Their technique was to look for changes in the bootstrap sector of a PC. Jan reasoned that this was an elegant and logical way of detecting the presence of a virus. Jan recounted however how the market wanted virus identification. In other words, it wanted to know *which* virus was attacking a machine, not merely that a virus was attacking a machine. Jan didn't like the virus identification model because it required continuous updates to maintain a library of known viruses on each user's machine. The Sophos bootstrap sector scanning model was more elegant. Eventually Sophos

succumbed to the market's demand and this proved to be one of the most decisive steps in the company's future financial success. Annual updates mean recurring revenues and that means, as Jan showed with a chart, cumulatively increasing revenues. A conventional software sales model with new sales and a mere 15 or 20% charged for annual maintenance requires large volumes of fresh customers each year. The Sophos subscription model, given strong levels of renewals, requires only a modest number of new customers to be won each year to show strong overall growth in turnover. This has proved to be a key factor in their subsequent growth and high profitability.

Jan's second lesson:

"Follow the market even when you think it's wrong"

Originally the company was intended to be only 10 people. People management has been a further factor in Sophos's growth. Jan said that "people are not like computers: they have minds of their own." They always want to know precisely what the corporate goal is. Often Jan would see a member of staff with a gloomy face and ask what was wrong. The reply would come that the company wasn't what it used to be and they didn't know where the company was going. To cure this problem, Jan produces notices each year, stating the company's goals for the year, and puts up copies in every lavatory in the office.

Another HR issue is that, once you grow beyond 100 people, whatever can be misinterpreted will be. For example, if you organise an outing for staff on a Saturday, half will be grateful whilst 20% will be resentful at the loss of their personal time.

Jan's third lesson:

"Choose your hires very carefully"

Hiring is easier than firing especially in Italy where firing is almost impossible. Jan or Peter have interviewed every single one of Sophos's UK hires personally. Beware of idiots who will

recruit more idiots. Execute any change “violently”; i.e. as fast as possible, and before the uncertainty period becomes counter-productive. Above all, communicate, communicate and communicate.

Managing profit is about managing discretionary spending. Profit is turnover less committed costs less discretionary spending, which includes new staff, marketing and capital expenditure. Over budget expenses, under budget sales and profits will follow. Watch your cash flow carefully. Report sales and expenditures weekly. A monthly reporting cycle which is normal UK accounting practice is too long an interval for effective management. Be wary that staff who are not shareholders may lack a sufficiently strong incentive to manage costs tightly. Consider bonuses based on profits and make employees into shareholders. In Sophos’s case it took 18 years for this to happen. Professional advisers have published lists of signs of a company going downhill. Amusingly, Jan ran through this list in relation to Sophos. First on the list is a new HQ (which Sophos has) and second on the list is a VIP building opening (which they also had when the Queen presided). Fortunately Sophos does not have flagpoles outside the entrance, a fountain in reception, a car park full of new cars, company jets nor helicopter although they are an industry prize winner.

Things to watch out for as a company leader are yearbook salesmen who ring often offering to sell you entries in their yearbook. Jan answered his own phone from 1985 to 2003 and found he learnt more from speaking to an irate customer directly than he lost in dealing with salesmen calling.

Urgency is important. Do things now. Don’t put things off. Kill urban myths fast. Staff can sometimes invent rules as to how to do things which are not required, simply because someone at some stage expresses a preference for doing something in a certain way. If you win a prize or two, don’t relax.

As for the future, they plan to maintain their turnover and profit growth and continue to give headaches to their US competitors. Jan is not much in favour of acquisitions, may consider flotation but Jan definitely does not plan to retire. Would he do it again? Emphatically yes, but he would think more carefully before doing it. Starting a business when you are young is good because you have less fear of the dragons. Sophos is recruiting people now and welcome applications.

After many questions, Dr. Ross Anderson thanked Jan on behalf of the audience.

Event Calendar 2004/2005

October 13th

Talk and Drinks

Time: 18:00

Venue: London

“Cambridge Entrepreneurship and UWB”

Speaker: Jack Lang

Invitation enclosed

November 18th

Careers Talk

Time 16:15

Speakers: Cliff Penton, Head of

Research, Sophos anti-virus

Keith Turnbull, Vice President Product

Development, Citrix Systems (R&D) Ltd

January 27th 2005

Time: 16:30

Venue: Cambridge

AGM

Time: 17:00

“Opportunities and issues for British ICT/Telecom companies in China”

Speaker: Ting Zhang, Managing

Director, China Business Solutions Ltd

February 2005 TBA

March 14th 2005

Laboratory Update

Annual Dinner, Jesus College

April 2005 TBA

May 17th 2005

Time: 16:15

Venue: Cambridge

“Computer Assisted Radiology and Surgery”

Speaker: Professor Heinz Lemke, Technical University Berlin, Computer Graphics and Computer Assisted Medicine

June 2005

London Drinks Party

Ring Interview

The Ring interviewed Professor Andy Hopper, Head of the Lab for Communications Engineering, Fellow of the Royal Academy of Engineering and serial entrepreneur. Andy Hopper has an amazing record of developing and commercialising new technologies and to date has founded 11 companies.

Professor Andy Hopper Hall of Fame Companies

| | |
|------|---|
| 1978 | Co-founder, Orbis, sold to Acorn 1979 |
| 1979 | Founding Director Acorn Computer |
| 1982 | Acorn goes public on UMS |
| 1985 | Co-founder, Qudos, e-beam chip prototyping and CAD software |
| 1986 | Founding Managing Director of Olivetti Research Ltd (later AT&T Laboratories Cambridge) |
| 1990 | Acorn spins off ARM |
| 1993 | Co-founder, ATML (later known as Virata), DSL chip company |
| 1996 | Co-founder, Telemedia Systems (now IPV Ltd) |
| 1998 | Co-founder, Adaptive Broadband, fixed wireless equipment Adaptive Broadband sold to California Microwave |
| 1999 | Virata IPO on NASDAQ |
| 2000 | Co-founder, Cambridge Broadband |
| 2001 | Virata merges with Globespan to form Globespan Virata |
| 2002 | Co-founder, RealVNC Ltd, remote control software |
| 2002 | Co-founder, Level 5 Networks, network equipment |
| 2002 | Co-founder, Ubisense Ltd, location systems |

The Ring started by asking Andy what motivated him to become an entrepreneur. The answer was disarmingly simple: it just happened. However, from his early years as an undergraduate, Andy's interests have been multi-faceted. He read EFCS (electrical engineering and computer science) and business at the University of Wales in Swansea. Not only was he able to sate his appetite for Engineering, Maths and Computing but he also learnt about economics and accountancy.

So, how did Andy spot his first opportunity, the birth of Orbis?

Modestly, Andy doesn't view the ease with which he moved seamlessly between academia and industry as exceptional. His luck, he says, worked early. Back in the late 70s, he found himself at the Darwin College disco with his friend Hermann Hauser. On regaling Hermann with stories of the visitors he had shown around the Lab - they were interested in the work he was doing on the “Cambridge Ring” solution for his PhD - Hermann suggested they start a company. And so Orbis, a network equipment company, was born.

As to Andy's goals there was only ever one: to do something that would make a big impact and live up to the standards set by the Computer Laboratory. The gravitas of the Computer Laboratory was palpable and Andy's key motivator.

Indeed, the question of how to move forward with the serious intellectual perspective that lives up to the Computer Lab's tradition is one that still exercises Prof Hopper.

Unpretentiously, Andy says he was lucky in all sorts of ways. He liked doing practical things and found himself in a high growth field at the early stages of its development. Things are much more difficult now, more mainstream. The pioneering days, he believes, are over.

As the acceptance of innovation has become lower and more problematic so Andy Hopper believes he has an increased responsibility to engage with the public and promote a wider understanding of the issues. His position

on the Board of Trustees of the IEE helps achieve this function, as the IEE acts for the engineering profession and provides professional advice to Government and other agencies.

The one thing Andy Hopper has learnt, since he co-founded Orbis, is that Cambridge, as an innovation entity, is team driven rather than project or finance driven. Indeed, it is the importance of networks that bind Cambridge companies together. At their centre lie a number of key individuals all of whom have a link with the University and the Computer Lab.

What lies ahead for Andy Hopper?

Well, from October 1st 2004, he will become the Computer Laboratory's new Head of Department. The challenges facing life in academe in 2004 are large and he believes passionately about making the job as attractive as possible not only from an intellectual perspective but also from a business and collegiate standpoint.

Andy Hopper has been christened 'Chief dream catcher' by those who know him and who have worked with him. He will soon be helping those at the Lab catch their dreams too.

Hall of Fame Profile

Bango.net

In the latest in the series of articles profiling companies founded by Computer Lab graduates, 'The Ring' was delighted to talk to Ray Anderson, CEO and co-founder of Bango.net. Ray is a graduate of Caius College.

TR: Ray, what is the history of Bango.net? What made you decide to start it?

RA: In early 1999, a friend of mine showed me the exponential graphs of internet user growth and mobile phone subscriber growth and asked me the question "what would happen if these things came together?". I had a lot of

sleepless nights trying to think through the implications of billions of normal people having immediate access to any service they needed, in their pocket, at any time. A trip to Japan the next month opened my eyes to the fact that Internet connected mobile phones were starting to emerge, and that now was the time to enter the market as a small company – while there was a chance to make a difference.

I realised pretty quickly that the way users accessed information on the fixed internet - with "URLs" that they had to remember - was totally unsuited to mobile devices. First of all, it's hard to enter words, dots and slashes into a phone, and secondly, if you have the phone with you, you can avoid remembering anything by just entering where you want to go immediately – just as Mobile Phones meant people could call immediately rather than have to write down a number and use it at their desk. The idea was simple, enable Content providers anywhere in the world to register any internet content with a central database that would give them a simple number – a "Bango Number" – that could then be entered on any device to get immediate, unambiguous, pinpoint access to the content. Numbers are easy to show, speak, present, share and enter by keyboard, voice or scanning. They are also global in reach. To people registering with that big database, many other useful services could be provided – making the business potentially unlimited in scale – but based on an activity that only a small, neutral company could provide – a central database mapping numbers to content.

TR: How have bango numbers changed the take up and utility of WAP services?

RA: Bango Numbers have now been used by more than 20 million users world-wide, and from our data seem to be driving around ten percent of Mobile Internet Traffic in our top five countries.

TR: What has been the growth of pages accessed on the mobile internet over the past year and how do you see it developing ?

RA: The Mobile Operators publish figures every month, and these show an approximate doubling of page accesses over the past year. Our growth has been closer to a doubling every 4 months, so Bango users are obviously an important driver of the overall growth of the Mobile Internet. On the other hand, while 40% of UK mobile users have internet enabled phones, less than a quarter have yet used that capability, and only about 2 million users in the UK use the Mobile Internet more than once a week. There is obviously a huge potential upside out there!

TR: When did Bango.net start providing technology to handle mobile users' payment transactions for mobile content? What was the thinking behind the move?

RA: Content providers have to pay Bango a small fee to make their content easy to reach. The problem back in 2001/2002 was that most of them could not invest to make their services accessible unless they could get revenue back in return. Our growth was slower than we would like because we had to wait for Content Providers to figure out ways to get paid and integrate with them. We decided to solve that problem, so that Content Providers could make money, and then spend some of it with us to accelerate their own growth. We were fortunate in establishing a unique partnership with Vodafone at that time, which gave us privileged access to their billing systems. This meant that Vodafone users got easy access to a wide range of high quality content, and that content providers who registered with Bango could turn a profit very very quickly. That was the rocket booster that took us to a multi-million pound turnover and which made us a very hot company on the Mobile Internet.

TR: SMS is likely to become a powerful form of marketing in the browse and buy shopping experience for mobile content. How do you think abuse, in the form of unsolicited SMS spam, can be tackled?

RA: SMS already is a very successful marketing method. One of our content providers sends out 50,000 text

messages to their customers every week, at a cost of around £2,000 and generates well over £20,000 of revenue as a direct result – in the following few days. The good news for SMS is that the sender has to pay to send it – about 3pence a time, so they have to have a solid proposition to afford to send the spam. This either means targeting receptive, welcoming customers – as with proper advertising – or trying scams that play to people's greed or ignorance and these are generally closed down very rapidly by preventing the money flow back to the sender and punishing the company that contracted with a Mobile Operator to send the messages. It's very different to the old style PC internet where anonymity reigns.

TR: What mobile technologies have caught your eye over the last year?

RA: I'm very interested in applications that enable a user to easily share their location with useful services. Traditionally, the location of a phone has been provided by a Mobile Operator, but this costs the service provider 10-20p a time and involves huge "opt-in" problems. There are ways to cut the cost to zero and leave the opt-in to the user alone, and this is opening up huge new opportunities in entertainment, dating, shopping, and travel applications.

TR: What are Bango.net's plans for the future?

RA: We have two priorities this year. First, to ensure the technology scales smoothly and reliably to handle the growth that our content providers expect. Second, to relentlessly streamline the Bango user experience around the world – working in partnership with local Mobile Operators as required. Underpinning these plans is a desire to continue to build a team that can deliver on the unique opportunity that Bango offers – and that can really make the Mobile Internet both fun and profitable!

TR: With all your many years of experience as an entrepreneur, what would you highlight as the 3 most important lessons you have learned?

RA: Stay small and nimble until you have really found a workable business model. Only work with really good people who you can trust. The customer is always right, so don't waste time with prospects who are certainly wrong.

Laboratory Research

Bioinformatics

Dr Pietro Lio
Lecturer in Computational Biology

Bioinformatics involves applying computational techniques to problems in molecular biology. It focuses on probabilistic models to analyse genome sequences and other 'omic' data (for example, transcriptomics and proteomics data). The term 'genome' derives from Gene + Ome. A gene is an ordered sequence of nucleotides located in a particular position on a particular chromosome that encodes specific functional products (i.e., protein or RNA molecules). It contains upstream regulatory regions that specify how much, where (which cell, tissue, organ) and when (developmental stage, disease) a gene is expressed. The suffix 'ome' is, according to Merriam-Webster dictionary, from the Latin for "mass". The genome, therefore, represents the whole genetic information content of an organism. Genome analysis involves the description of the general rules that allow an organism to utilise and combine the different types of information present in the genomic sequence, within its physiological and environmental contexts. There is no unique way for describing the genome of even the simplest organism. The situation can be well represented by the analogy from engineering drafting where blueprints of some complicated machinery are usually replaced with an exploded view of the same, with each component shown from many different angles or in holographic view. Similarly, location and abundance of genes, regulatory sequences and 'junk' (= unknown function) DNA, are usually represented in the form of several colour-coded wheels, where high density

values of the different features are shown with darker colours. Genome sequences have been investigated on the basis of several characteristics: for example: linguistics, percentage of different classes of genes (for example genes involved in energy production, metabolism, active transport, DNA repair etc), patterns associated with physico-chemical properties of sequences, such as curved or straight rigid DNA regions, and metabolic pathways, i.e. description of ordered sequence of reactions.

Linguistics methods are usually based on the assessment of the statistical significance of over- and under-represented oligomers in complete genomes. The analysis of the bias of oligomers requires a model of how sequences are generated and the statistical distribution of the oligomer. The simplest composition analysis is the linear density of guanines and cytosines (G+C). The G+C content is an interesting statistical property of a genome since there are large differences both within and between genomes. In bacteria, genomic islands with different G+C content with respect to the neighbouring regions are generally the result of the transfer of genetic material from other bacteria. In bacterial pathogenic strains, some of the G+C islands carry virulence genes that code for toxins and virulence factors. The human genome is made up of long DNA segments, the isochores, which are homogeneous in GC content and differ in gene content. Linguistics is used to detect gene locations which is the most important step in the analysis of a new genome. This task is nowadays relatively easy for bacterial genomes but still very difficult for the human genome. In some cases a dozen human genes, although interrupted by non coding DNA, can span the same length of an entire bacterial chromosome. A starting point for gene identification is to find a similar sequence from another organism. This can be done using the BLAST and FASTA families of programs to search all the entries in sequence databases. Most sophisticated gene-finding programs incorporate inferential rules based on hidden Markov models.

There are approximately 30,000 genes in the human genome but there may be as many as 300,000 gene products. These are the result of modifications of the gene transcription in the different environments (cell types). Molecular complexity has a pyramidal shape: the hundred of thousands of existing genes correspond to few thousands or hundreds protein architectures that may contain no more than few hundred different protein motifs and very few types of protein networks. The functional characterization of genes can be obtained by using microarrays. DNA microarrays provide a simple and natural vehicle for exploring the regulation of thousands of genes and their interactions in different biological backgrounds (cells, tissues, organs). A microarray is typically a glass (or some other material) slide, on to which DNA molecules are attached at fixed locations (spots). Typical dimension of a microarray is about 1 inch or less. The spot diameter is of the order of 0.1 mm. One of the most common microarray applications allows to compare gene expression levels in two different samples, e.g., the same cell type in a healthy and diseased state. The total mRNA from the cells in two different conditions is extracted and labelled with two different fluorescent labels: for example a green dye for cells at condition 1 and a red dye for cells at condition 2. Both extracts are washed over the microarray. Labelled gene products from the extracts hybridise to their complementary sequences in the spots due to the preferential binding. The predominant colour of each spot (red, green, yellow, black) will reveal the kind of hybridisation (condition 1 for a green colour, condition 2 for red, both conditions equally represented for yellow and no hybridisation for black) and, therefore, the presence and abundance of a certain mRNA molecule. There is a growing amount of literature showing that differences in DNA sequences reflect only a small percentage of the differences between species or between persons with a genetic disease and healthy relatives. For example, chimp DNA is 98.7 percent identical with human DNA. If just the genes are compared, the similarity increases to

99.2 percent. Gene expression levels in blood and liver in human and chimp are quite similar but very different in the brain tissues, where the expression of genes in the human brain has accelerated about three-fold. This suggests that the level of expression of two genes may shown large differences even if the gene sequences are highly conserved through evolution from fly to human. There are two types of microarray data: time series and steady state. In a steady state gene expression experiment, several closely related cell cultures are grown under very similar conditions, with expression data being collected only at one specific time point, usually during steady growth. A typical application is the study of the expression of thousand of genes in different human tissues and cancer cells. In a time series gene expression experiment, samples are taken at different time intervals from a single cell culture during growth. For each of these samples, microarray data can be produced revealing the changes in expression levels over time. Time series gene expression data are becoming central to all the investigations of gene expression patterns in human normal and pathological conditions. Many physiological parameters, that can be measured at cellular or at organism level, show a temporal pattern characterized by rhythmic changes along a time period, e.g. a cell division time, a day (for example the body temperature), a month, during aging or during a periodical drug treatment. It is also known that several diseases, such as cancer but also renal colic attacks and cardiovascular diseases and many others, show a rhythmic variability as well. In the near future, microarray technology will represent the appropriate tool to investigate general and personal vulnerable periods. The automation of data collection will make soon available very long time series. With the proliferation of replicated experiments by the same group, or similar microarray studies by independent groups, it is important to combine time series across studies. I am interested in developing wavelet and Bayesian methodologies for multiple time series analysis, with a particular

interest for the analysis of cell-cycle microarray data. Time series microarrays are affected by several sources of noise or errors, among which, the stochasticity of the gene expression process, the artefacts produced by the cell synchronization methods and the errors in measuring the expression levels. When a time series is affected by noise, the most obvious way to improve the signal-to-noise ratio is to replicate the experiments. Statistical analysis of replicates can be performed by first averaging gene expression measurements and then applying an usual clustering procedure. Since the standard deviation of such average expression levels will decrease with the square root of the number of replicates, it is likely that weaker associations will be detected than in the case when a single replicate is performed. However, such an approach is not using all information present in such data. The information on the variability between experimental replicates present in the data is discarded and only the information about the mean expression level is used. In the situation when experimental variability varies from gene to gene, such information can be crucial for the proper assessment of confidence in the final clustering results. It is known that the expression levels of some genes is more variable than others, and that log-transformed expression measurements of low expressed genes vary more than expression measurements of highly expressed genes. My research is focused on making the best use of all the replicates obtained in time consuming and expensive experiments, and combine them using the false discovery rate (FDR) family of techniques. The FDR is a concept for performing statistical inference when the number of hypothesis tests is large. The FDR is the expected proportion of true null hypotheses rejected out of the total number of null hypotheses rejected. Although FDR procedures are implemented in several widely used software for microarray analysis, their use for multiple time series is not statistically sound. Indeed, if we control the FDR for each microarray experiment and then we combine or intersect the results, we loose completely the control

on the overall FDR. Therefore, it is better to combine the experiments together, possibly assigning them different weights, and finally perform a FDR procedure.

The current major challenge in Bioinformatics is the accurate modelling of the genome-wide biological networks. Cells have a remarkable capacity to respond to different environmental conditions and signals. These responses can range from the induction of latent biochemical activities, to the launching of specific patterns of growth, differentiation, or death. Over the last several years, it has become clear that these responses are controlled by biochemical circuits. These circuits are composed of sequentially interacting proteins and genes whose sole function is to process and transmit regulatory information. They have many branch points, allowing for cross-talk between pathways, the detection multiple inputs, and the coordinated and interdependent formulation of appropriate responses. Thus it is increasingly appropriate to compare biological regulatory networks to computational or electrical circuits. Even single proteins have multiple domains acting as logic gates.

Due to the rapid development of large-scale gene screening techniques, particularly microarray and protein interaction data, a huge amount of new data concerning gene regulation is becoming available. My interest is to understand the principles that govern function and evolution of biological networks.

What determines the correct connectivity of regulatory circuits, given that the components are diffusable proteins? How do two communicating proteins specifically recognize one another and discriminate against the myriad of other proteins encountered in the cellular environment? How do interactions result in a transmission of the signal? How are interactions modulated/regulated by multiple signals and inputs?

One of my main goals is to investigate the possible universality

of biological network topologies and the presence of generic, emergent properties in biological networks. This means that natural selection does not operate on each gene in order to tune all its interactions and thus achieve a stable and flexible state. It only needs to tune macroscopic, average parameters such as connectivity. The development of models of evolution of gene networks may explain the high tolerance of proteins to amino acid changes, the resilience of metabolic flux to changes in enzyme activity and the robustness of developmental pathways to mutations in their constituent genes. A suggestive hypothesis is that the observed network robustness is a property that itself has evolved because it "protects" an organism against otherwise deleterious mutations.

Computational biology studies of biological networks will mark the beginning of a period of exceptional knowledge discovery in biology; we will be able to chart the wiring diagrams of every biochemical and signalling pathways, explain how genes regulate the 'integrated circuit of the cell' and how its alteration can lead to cancer and other pathologies and ultimately understand an entire organism as a vast genetic and protein network. I am developing computational models to analyze and predict network behavior in living cells and to explore the range of possible networks that combine short-term robustness with long-term evolvability. In the long-term, I hope that such studies will lead not only to a deeper understanding of how biological signalling pathways work, but also to the development of ways to block and modulate pathways for therapeutic purposes, as well as to design and engineer protein interactions and protein-based circuits for biotechnological applications. Closely related to the field of biological networks is System Biology that represents the vertical integration from molecular biology to physiology and anatomy. I am interested in understanding the building up of biological complexity, from gene networks to cell, tissue, organs, systems

and organisms. As a case study I am modeling biological networks in liver, linking liver morphology, regeneration dynamics after surgery and cancer with molecular data. Liver architecture consists of sheets of connective tissue that divide the liver into thousands of small units called lobules. A lobule is roughly hexagonal in shape, with portal triads at the vertices and a central vein in the middle. If we cut one third of liver it will grow again to roughly the original size and shape. This may suggest that size control is determined by the signalling environment and function optimization. The exagonal patterns of lobules inspired me to use cellular automata to model the spatial organization of liver. The automaton is based on local rules, for instance that the cells are more likely to divide depending on the oxygen availability and on the state of their nearest neighbouring cells. I am interested in understanding the relationship between the organization of liver cells and blood vessels using molecular data in normal and pathological conditions. The model will also give insights into the dynamics of stem cells population in liver. In conclusion, I am delighted to report that Bioinformatics is a thriving field of research. It has two important and linked benefits: an improved ability for capturing hidden components from biological data and a better link between biological systems and the computational models used to describe them.

Hall of Fame News

(The full list of companies can be found on www.camring.ucam.org)

Curious Software Company Ltd has won The Queen's Award for Enterprise: Innovation 2004 for its map animation software.

Sophos Plc has won The Queen's Award for Enterprise: International Trade 2004 for its anti-virus and anti-spam software.

On July 12th, **Artimi**, founded by Jack Lang (*EM Dip71*), announced a \$14mio Series A funding round. Index Ventures

co-led the round with Accel Partners and Amadeus Capital Partners Ltd and investment from Oak Investment Partners.

On May 4th, Tideway Systems, co-founded by Tim Coote (*CAI BA80*) and Duncan Grisby (*Q BA95 PhD99*), announced that Funds advised by Apex Parnters are providing \$5.8mio in series A funding.

Who's Who

(The following came too late to make the August 2004 Who's Who)

David Cleevly (*TH PhD83*) is leaving Analysys following its acquisition by Datatec International. David founded Analysys in 1985 and developed it into one of the world's leading telecommunications strategy consulting firms with a global client base. David is active in a variety of other areas including serving as Chairman of Abcam and being a non-executive Director of Cambridge Broadband. His links with Analysys will continue as he will serve as a consultant to the new Analysys Mason Group.

Feng Hou (*G Dip03*) has joined Nokia. He works in the Statistics group, Network division and is responsible for improving 3G network performance by analyzing traffic load among the networks.

Raymond Whorley (*DAR Dip03*) is working as a Process Engineer for MW Kellogg Ltd.

Books

"Hitchhiker's Guide to SQL Server 2000 Reporting Services"

Peter Blackburn and William R. Vaughn
0-321-26828-8, \$49.99, 608 pages,
paperback with DVD, available October, 2004.

This book helps developers and database manager to get up to speed quickly on Microsoft SQL Server 2000 Reporting Services, showing them how to build, manage, and deploy reports online. The

authors start where the Microsoft documentation leaves off, walking the reader through every aspect of report installation, management, security, creation, and programming, using in-depth tutorials and innovative demonstrations. The book features extensive coverage of the Visual Studio .NET 2003 add-in used to create and deploy reports. It also includes a DVD containing narrated Camtasia (tm) demonstrations that illustrate many complex aspects of Reporting Services as well as a wealth of code and report samples.

Peter Blackburn [*CTH BA89*] has worked in the computer industry since 1981. He holds an MA in computer science from Cambridge University and is a Microsoft MVP. Through his consulting company, Boost Data Limited, Peter has built both small and large-scale databases and reporting systems based on SQL Server. He has contributed countless hours to the development and honing of Microsoft Reporting Services and remains directly involved with the development team. Peter lives in the UK and travels all over the world to consult, speak, and provide support for his fellow developers, fellow MVPs, and Microsoft itself."

Special offer for Ring members.

Simply go to:

www.awprofessional.com/titles/0321268288.

Upon checkout, enter the coupon code "TheRing" and you will receive a special 30% discount, plus free shipping within the US. Shipping must be added for those outside the US.

"The Mathematics of Natural Catastrophes"

Dr Gordon Woo [*CHR Dip80*]

Published by Imperial College Press (1999)

This is a unique book about natural catastrophes, focusing on the mathematical aspects of these phenomena. Although academic in style and didactic in purpose, this book is practical in the treatment of the diverse issues covered, which range from hazard warning and forecasting, to risk

criteria, and economic loss estimation. This book discusses the basic algorithms on which disaster risk management software is based.

Letters

From Dr Chong Woo Kim (JN Dip99)

I completed the Diploma course in 1999, and worked for an IT company in Cambridge as a Java Application Developer for 3 years before joining Samsung in early 2003. I am currently working in TCAD (Technical Computer Aided Design) group in Samsung Semiconductor Memory Division. Prior to this, I have studied in the UK from the Secondary school education to the completion of Ph.D. research in gravitation at Imperial College, London. So my time in the UK spanned well over 20 years before deciding to return to Korea to work for Samsung. Samsung is the largest conglomerate in South Korea, and its chip making division is the second largest producer of semiconductor chips after Intel. Modelling a MOS transistor to determine its characteristics using computer simulators has been my main area of work. MOS transistor is the basic unit that makes up a chip and the performance of the chip largely depends on it. Computer simulation can save valuable development time and cost.

Since my return, I have been trying to adapt back into my own culture but the words "The Far East" do not feel as far as they used to. I would like to see more people from the Cambridge Computer Laboratory in our division and will open up collaborations with the laboratory in the future.

I do miss driving around rural English country lanes and seeing all those peaceful sheep on rolling hills of the English countryside and walking around the cobbled streets of Cambridge and popping into Belinda's cafe for a nice cup of tea. One more thing, trying to get a bag of salt and vinegar flavoured Walkers crisps here is beyond impossible.

I was not fortunate enough to use the new facilities at the William Gates Building but I have been there once to attend the seminar organized by the Computer Lab Ring. I hope the Cambridge Computer Laboratory keeps on playing its central role in the fundamental research in Computer Science.

Computer Laboratory News

Head of Department

Following the appointment of Professor Ian Leslie to the Pro-Vice-Chancellorship (Research), Professor Andy Hopper has been appointed as Head of the Computer Laboratory from October 1st 2004. (see Ring Interview)

ACL Lifetime Achievement Award

Professor Karen Spärck Jones has received the Association of Computational Linguistics' Lifetime Achievement Award. Prof Spärck Jones was given the award at the 42nd annual meeting of the ACL in Barcelona on July 23rd 2004. She is only the third recipient.

Student Prize Winners

Computer Science Tripos Part II, 2004

The Microsoft Research Prize for the Best Student is awarded to Daniel Lau (Queens)

A Microsoft Software Prize is awarded to Viktor Vafeiadis (Selwyn)

The AT&T Prize for the Best Dissertation is awarded to Viktor Vafeiadis (Selwyn)

Data Connection Prizes for Outstanding Dissertations are awarded to:

David Allsopp (King's)
Ben Benfold (Jesus)
Toby Collins (Queens)
Andrew Connick (St John's)
Grzegorz Milos (Trinity)
Christian Steinruecken (King's)
Arthur Taylor (Churchill)

Job Postings

Sophos

Web Development Support Analyst

Ref: WUK203

Up to £23-27k p.a. + bonus + benefits

We are looking for a Web Development Support Analyst to join our expanding Web department as part of our new Front Line team. Duties will include providing rapid response support for our existing web applications, and programming changes and new feature requests as they arise.

This is an ideal role for a candidate with a background in supporting Apache and MySQL on Unix or Linux systems who is keen to gain structured object-oriented mod_perl development skills.

The position offers an excellent opportunity to improve your skills with help and guidance from an established team of experienced developers.

To find out more, please go to the Job Bulletin Board on the Ring website:
www.camring.ucam.org

Sophos

Virus Analyst

Ref: WUK208

£18-26K p.a. + bonus + benefits

The Virus Research role consists of three main parts.

The principal role comprises taking apart new viruses, reverse engineering the code to determine precisely what the virus does and writing a detailed analysis of it, covering its payload and method of spreading. In addition, virus researchers write identities that detect viruses. These are incorporated into the product and are used by all our customers. Whilst working as part of the laboratory team, much of this work is carried out individually, calling for self-motivation and the ability and commitment to complete whole projects.

The role also entails carrying out research into different aspects of viruses, for example spreading techniques, payload capabilities or new virus platforms and applications. This research may lead to new capabilities being added, to detect potential threats and even to the writing of white papers on the research material. This part of the role puts you at the forefront of research into viruses, requiring the confidence to face the challenge of working everything out for yourself.

Software Engineers

Ref: WUK087

£25-50k p.a. + bonus + benefits

The Software Engineering department has a number of opportunities for people with varying levels of commercial C/C++ programming experience.

Due to the continuing growth of the company, openings exist for programmers and senior programmers.

To find out more, please go to the Job Bulletin Board on the Ring website:
www.camring.ucam.org

Sophos
Software Engineering Team Leader

Ref: WUK210

From £40k p.a. + bonus + benefits
Sophos is looking to recruit a Team Leader in our Software Engineering department to supervise the development of cross-platform software in C++.

Your main duties will include designing, implementing and delivering applications, components and enhancements for software product releases. You'll contribute to the identification and formulation of the functional, architectural and test specifications and lead groups of software engineers in delivering high quality software solutions on time and to specification.

Unix Systems Administrator

Ref: WUK197

Up to £35k p.a. (depending on experience) + bonus + benefits

As part of the Unix Systems team within the IT department, you will be responsible for administering, developing, maintaining and providing internal support for Sophos's Unix-based systems. These include web servers, DNS servers, and mail gateways in the UK and overseas.

This position will primarily involve:

Set up and day-to-day administration of Unix systems; providing internal support concerning Unix systems; assisting the web team with development of web based services using Perl, PHP, MySQL, etc; support of internet services in Sophos offices overseas; building, installing, and configuring applications on Unix systems; development of scripts (shell, Perl, etc); maintenance of Unix system security; take an active role in planning future developments of Unix systems; helping to build robust, well documented, systems that will carry Sophos forward; development of system administration procedures.

Sophos
Perl Programmer

Ref: WUK169

£25-30k p.a. + bonus + benefits

As a part of the Web Applications Team, your primary role will be to develop applications for Sophos's multilingual websites in Object Oriented Perl.

We expect successful candidates for this role to be comfortable developing systems in Object Oriented Perl from specifications with minimal supervision. This is an excellent opportunity to improve your skills in an industry-leading company.

To find out more, please go to the Job Bulletin Board on the Ring website:
www.camring.ucam.org

Jagex Ltd
Web Developer

£19-23k p.a. + bonus + benefits

We are looking for an enthusiastic web developer/programmer, who has experience of programming using databases. Candidates must have extensive knowledge of core Java (not necessarily Java extensions), or C++ and good HTML skills.

The job offers an exciting opportunity to join a fast growing company working with three others in a busy web development team.

To find out more, please go to the Job Bulletin Board on the Ring website:
www.camring.ucam.org

The Ring

Issue VII September 2004

William Gates Building
Cambridge
CB3 0FD

The Ring is the newsletter of Cambridge Computer Lab Ring, the University of Cambridge Computer Laboratory Graduate Association. It is published 3 times a year.

January issue copy deadline: December 1st 2004

