Issue 1

HUB

Computer Science Research at the Leading Edge

www.dcs.qmul.ac.uk/research/

Avoiding human error

Catching crime on camera

Interaction and all that jazz

Less risky business

Department of Computer Science

Queen Mary University of London

HUB

Welcome to *The Hub*
Each issue introduces some of the leading edge Computer Science research at Queen Mary. We focus on core Computer Science and also inter-disciplinary research in dynamic emerging areas such as at the boundary between Computer Science and Biology.

We specialise in both theoretical and applied Computer Science research focusing on:

- **Computer Science Theory**
- **Vision and Interaction**
- **Information Engineering**

IR

I've found it, but what does it actually say? It's not often that products gain such prominence that their name becomes the verb to do the activity: 'to hoover' is one example and 'to google' a more recent one. Google's sudden rise to prominence has meant that some people have assumed all the problems of Information Retrieval have been solved – "just Google it". Nothing could be further from the truth. Information Retrieval has matured a lot but that just means there are now many new and exciting problems to solve, not least as the way people use the Internet changes.

A crucial area of research linked to Information Retrieval concerns translation between different natural languages. After all not everyone speaks English and perhaps the information you are searching for is there on a web page in Spanish or Japanese. Traditionally text has been translated between languages by people fluent in both. That is of no use for a search engine though. They need an automated approach, which is where IR group member Christof Monz's research into machine translation comes in.

Recent approaches to machine translation have involved professional human linguists manually writing lots of translation rules for computers to follow to do the translation. The translation is then only as good as the rules used. The complexities of human languages and the fact, for example, that concepts from one may not exist at all in another, mean that devising sets of rules is far from trivial. Christof's expertise is in an area known as statistical machine translation where great advances have been made recently. In this approach the machine learns the translations rules automatically. It does this using a parallel corpus: just lots of pairs of sentences; one a sentence in the original language, the other its translation. Parallel corpora are extracted from multi-lingual news sources like the BBC web site where professional human translators have done the translations. Such corpora are publicly available and continuously evolving.

In the language Boro, 'mokhrob' means to express anger by a sidelong glance

This approach is an active research area so there is still lots to do before it works perfectly, but it shows great promise. Below is an example translation to illustrate the current state of the art and the problems remaining. It shows a sentence from an Arabic newspaper then its translation by the Queen Mary, University of London's statistical machine translator, and finally a translation by a professional human translator.

بغداد ١-١ (اف ب) ذكرت وكالة الانباء العراقية الرسمية ان نائب رئيس مجلس قيادة الثورة في العراق عزة ابراهيم استقبل اليوم الاربعاء في بغداد رئيس مجلس ادارة المركز السعودي لتطوير الصادرات عبد الرحمن الزامل.

…Its Machine Translation:

Baghdad 1-1 (AFP) - The official Iraqi news agency reported that the Chinese vice-president of the Revolutionary Command Council in Iraq, Izzat Ibrahim, met today in Baghdad, chairman of the Saudi Export Development Center, Abdel Rahman al-Zamil.

…and the Human Translation:

Baghdad 1-1 (AFP) - Iraq's official news agency reported that the Deputy Chairman of the Iraqi Revolutionary Command Council, Izzet Ibrahim, today met with Abdul Rahman al-Zamil, Managing Director of the Saudi Center for Export Development.

The statistical translation does allow a reader to get a rough understanding of the original Arabic sentence. There are several mistakes, though. Mistranslating the "Managing Director" of the export development center as its "chairman" is perhaps not too much of a problem. Mistranslating "Deputy Chairman" as the "Chinese vice-president" is very bad. That kind of mistranslation could easily lead to grave insults!

For now the human's are still the best translators but the machines are learning from them fast! One day soon you may be able to search the vast number of web pages that exist, written in languages other than your own without worrying about the language problem at all.

IR

The Department's Information Retrieval (IR) group is tackling a range of IR problems. For example, Professor Mounia Lalmas, Head of the Information Retrieval group is working with Yahoo! Research in Barcelona on identifying the most informative properties of search results to help users see what is there. Previous projects of the group have also included the evaluation of XML-based digital libraries and investigating how interactive Digital TV services can be personalised for a range of devices from TV sets to hand-held PDAs.

THEORY

No not that button!

Humans don't just blindly follow instructions in the same way as a computer. If a computer does the wrong thing, to find the problem you look to those instructions and the hardware that executes them. Humans on the other hand are unpredictable. Even the same person doing the same thing will do it differently, making different mistakes, from one day to the next.

We like to think we have some free will rather than being programmed from birth. Unfortunately that makes us very error prone!

So what happens when you apply tools and techniques developed to find bugs in the implementation of computer systems to the behaviour of people? That is exactly what Paul Curzon's Theory group team have been working on. Surprisingly perhaps, it turns out you get a useful way to predict human error. More to the point, you get a way to identify when a computer system needs to be redesigned to prevent the people using it from making mistakes.

Software and hardware bugs are a big problem. Computer applications are so complex that you cannot check they work in every situation by testing. That is why a lot of effort has been put into developing mathematical approaches to finding the problems. Formal verification is one approach. It involves developing a mathematical model of the system and of its specification and proving that the model does describe the behaviour specified. If you can't prove the model has the right behaviour, you may have found a bug. To handle the complexity of the systems individual parts are modelled separately.

These separate models are then combined, modelling how information flows between them, to reason about the correctness of the whole.

Most computer systems do not just have a computer component though. A human is needed to push the buttons! Think not only of computer applications like word processors and drawing packages, but of all those everyday devices we interact with like cash machines and mobile phones, as well as more complex safety critical systems like an aircraft cockpit. They do not work alone. The idea is that a person interacts

with them to achieve some goal (like get cash). The combined person-computer system will only achieve this desired result if the person does the right thing too. Unfortunately, some designs of interactive systems are easier to make mistakes with than others (think about how hard some video recorders are to use).

Wouldn't it be nice if we could mathematically model a person's behaviour in the same way as the rest of the system and treat it just like a computer component? That way we could find out if human error could lead to the combined system not working as planned in the same way as for computer errors.

In fact, humans are not as unpredictable as you might imagine. Results from cognitive science have shown that our behaviour follows patterns to do with our goals and knowledge as well as limitations about the way our brains are wired. A set of principles of cognition can be developed that give an outline of the way we behave in general. For example, if we are expecting to press a particular button, we may do so even if instructions on the screen say otherwise. A key idea is to model our behaviour non-deterministically. That just means that if several options are plausible the model does not tie down which will be taken.

> We use experiments based on games like SpaceInvaders and fire engine dispatch to work out the behaviour to model.

In some situations – with badly designed interactive devices, for example – this 'rational' behaviour can lead to our making errors in predictable ways. Such 'systematic' errors won't be made every time, so they are not predictable in that sense. However, we all have the potential to make them and so eventually someone will. If a design harbours such systematic errors then it probably should be improved.

By modelling the human component of an interactive system based on these general principles of cognition, we can use this 'user model' within the formal verification process just like any other component. If our verification tools suggest the user model makes mistakes when combined with the model of a particular application, we can deduce that humans are liable to make the same error systematically too. Time to rethink the design of the interface!

So next time you make some stupid mistake pushing the wrong buttons on your video recorder, have the comfort of knowing there is a mathematical model that probably would make the same mistake. One day that model could help ensure future gadgets are easier to use.

THEORY

The research of the Queen Mary Theory Group cuts across a wide range of inter-linked areas. It is a major centre for the development of Separation Logic, one of the hottest topics in current Theoretical Computer Science, addressing the 25-year old problem of reasoning about shared memory programs. This work is now feeding into Microsoft's quality control process. Other major areas of the group include information theory and security, formal verification, making fundamental advances in classical logic, proof and algorithms, reasoning about control systems, and the foundations of concurrency. The group's research on the latter, concerned with the pi-calculus and web services, is feeding into the work of the WWW consortium in developing a framework to support powerful and flexible business tools for the web.

RADAR

Software defect prediction revisited

The developers of any new complex software system will confirm that, no matter how much testing they perform, there will still be plenty of defects (or 'bugs') yet to be found. The hope is that, when the software is released, any defects found by end-users will have minimal impact.

The decision about when to stop testing and release the software must always be balanced by the likely number (and criticality) of remaining defects. It follows that the ability to produce accurate predictions of 'residual' defects in software systems is one of the most important and challenging tasks confronting software engineers. It is especially relevant for safety critical software (such as in transport and medical systems where software that is released with too many defects can have life-threatening impact); but the business of any commercial software producer can be devastated if they get their release decision wrong.

In 1999 RADAR's Prof Norman Fenton and Dr Martin Neil wrote a paper about software defect prediction that was published in the IEEE *Transactions on Software Engineering*. The paper was a critique of the commonly used approaches to the problem and they proposed a novel approach based on causal models. Although the 1999 paper provided no validation of the idea, it became one of the most cited papers in computer science. According to Essential Science Indicators the paper is in the top one per cent most influential papers.

AgenaRisk the Bayesian network software package based on our research is available from www.agenarisk.com

In early 2006 the editors of the journal *Information and Software Technology* invited the authors of the 20 most cited papers in software engineering since 2000 to submit a new paper for a special edition of the journal. The intention was that the papers (which would still be subject to the journal's rigorous reviewing process) should reflect on the work that was done and report on new developments. Fenton and Neil were keen to respond because their work on defect prediction had recently taken a new and exciting turn.

Since the 1999 paper Fenton and Neil had been working with Philips Consumer Electronics to evolve and validate the causal models. Philips saw the work as especially important because they developed complex software that was embedded in electronic devices like TVs and DVDs. Being able to improve their decision-making about when to release the software was critical from a business perspective because faulty software could lead to the recall of entire batches of TVs or DVDs. Using models that were built and executed in the AgenaRisk toolkit Philips were eventually able to perform a comprehensive validation on 36 major projects based in Bangalore and Eindhoven. The 'headline' result was that Fenton and Neil's 1999 conjecture that causal models could outperform traditional approaches was well and truly proven. The predictive accuracy of the causal models was 93% (comparing predicted with actual numbers of defects); nothing like this can

RADAR is supported by ITRS Group plc, one of the leading providers of real-time monitoring software to the financial sector with their flagship product Geneos.

The computer controlled guided missile cruiser USS Yorktown was left drifting for hours because of a software bug that meant it tried to do an impossible calculation – dividing by zero.

NASA's $125 million Mars Climate Orbiter was lost in space because of a software bug. One programming team used Imperial weight measures. The other was using metric.

be achieved with the best of the traditional approaches, which rely on having access to extensive project data not normally available. But even more important than the predictive accuracy was that the causal models and tools enabled project managers to do genuine risk assessment and 'what-if' analysis that simply was not possible before.

In addition to the exciting results with Philips the new paper also describes how the approach can be easily tailored to handle arbitrary software development processes. Hence the approach can be used by organisations whose development processes are very different to those of Philips.

RADAR

The Risk Assessment and Decision Analysis research group (RADAR) takes a unique focus on problems of decision-making under uncertainty, and is a world-leader in the development of Bayesian net technologies. Research areas include the assessment of software risk, vehicle reliability, risk assessment in air traffic management systems, complex legal evidence, financial and operational risk in the banking sector, as well as personalisation and recommender systems. The group's current work on the problems of decision-making under uncertainty focuses on high-stake environments.

VISION

Watching…

CCTV cameras are everywhere, monitoring the streets, shops and offices. Soon they will be smart enough to know when there is a problem.

CCTV cameras are now everywhere, silently monitoring the streets, shops and offices. With every street under the watchful eye of a camera, the police and security staff can now instantly spot crimes as they happen… At least, that's the idea!

The reality is that it is impossible to watch all those videos. Only a few will ever be viewed by a human. After a crime happens the police can collect and view the footage to help track down the criminals, but even that takes a vast amount of time, so will only be done for really serious crimes.

It is ever more important that potential incidents are flagged in time to allow the good guys to intervene immediately: investigating unattended bags as soon as they have been left, for example. It is just not possible for humans to do this without help.

If only computers could 'see', really see, and make sense of what they were recording then they could help by spotting potential problems automatically and alerting security guards. It's not that easy though. In fact a massive research effort is devoted worldwide to making computers 'see'.

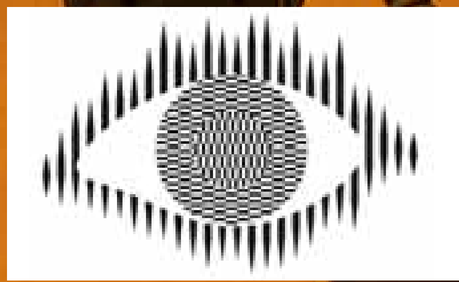
One such project is run by Professor Sean Gong, Head of the Vision Group. His team has developed new ways to process, interpret and manage the vast quantities of visual information captured by live CCTV cameras.

Computers can 'learn' what to look for. After being exposed to numerous examples of normal scenes, the computer can detect the unexpected, like detecting changes in a scene that could for example indicate a suspicious package, or theft of a bag, and the recognition of undesirable behaviour like trespassing or a pick-pocket in action.

An important problem to solve concerns what is called 'Zero Motion Detection'.

The biggest job for a computer trying to do Zero Motion Detection is to work out what is in the background and what is in the foreground. Even with simple scenes that can be harder than it sounds – many optical illusions like the Ouchi Eye (above) are the result of our own brains getting foreground and background confused.

It gets even harder in real CCTV footage where the scene changes constantly as the light changes through the day, as trees sway in the wind and so on. Just because it is background, doesn't mean it doesn't change. By programming computer systems to adapt to slow changes over time, sudden changes in the scene can be detected.



For example if a person walks into a room and places a box, temporary changes like the person moving around are ignored, but after the box has been present for a while it is flagged as a new permanent object.

To recognise undesirable behaviour, the problem is to separate normal from odd behaviour. To do this, the computer is given lots of different footage that contains the 'normal' if constantly changing state of affairs. The computer pulls out patterns in the data creating a 'model', a representation, of what is okay based on this allowable behaviour. Then when the computer detects a variation that doesn't fit the patterns in the 'model' it raises the alarm.

For example, if the computer has been trained repeatedly on sequences of a person entering a room and walking up and down, then it will thereafter ignore people entering and moving around because it is 'normal'. However if the person stops moving, that behaviour is detected as abnormal as the training footage didn't include people stopping.

Work is in progress to add the technology developed in this and related research projects into existing security software, so it may not be long before the all-seeing computers are really seeing, not just recording, and so helping protecting a street or airport near you.

VISION

The Computer Vision group is one of the largest in the UK and internationally leading in its work on the extraction of behaviour and facial information from image sequences and live video. The work has been widely applied to the detection of vehicles and people, tracking objects, counting and recognition in public space CCTV, human gesture recognition and recognition of abnormal behaviour. The group has a major focus at the moment on crime prevention via real-time surveillance and biometrics.

IMC

Augmented Human Interaction

Much of the IMC group's research is based around our state-of-the-art Augmented Human Interaction (AHI) Lab. Its facilities including, for example, full 3D motion capture, allow the group to study the minutiae of interactions within groups of people in unprecedented detail – down to millimetre accuracy.

Electronic Arts sponsor some of the work of the AHI lab

That is important because much of the communication, both in everyday interactions between people and in closely-knit teams, is non-verbal. Very small gestures and movements are used, often unconsciously, to convey a great deal of information. The lab gives IMC researchers the ability to gain deeper understanding of, for example, how dancers choreograph the fine detail of their performances, what makes a good poker player, and how artists work. The AHI Lab also allows the group to study more mundane everyday interactions such as how people combine verbal and non-verbal communication when giving directions, or what is going on in the dynamics of dinner parties as conversation groups form and dissolve. This work can inform the design of the way computers of the future do similar tasks, including interacting with people and how they help humans interact with each other.

Jazz improvisation

A typical example of the kind of research the AHI Lab enables is that of group member Nick Bryan-Kinns exploring how creative collaborations work by investigating the way jazz musicians communicate when improvising.

For people to collaborate creatively, there must be highly focused points of 'mutual engagement'. These are points where participants push at the boundaries of their shared understandings of experience and expectation - they engage both with the thing being created and with each other. Whilst people can collaborate without high levels of mutual engagement, it is the key to enjoyable, efficient, creative, and high quality collaborative activities.

Collaborative music improvisation is a good place to start as it involves very high levels of mutual engagement between participants. It combines musical signals with verbal and visual cues but in contrast to say face-to-face conversation it is more about mutual engagement and aesthetic satisfaction than just information exchange. The contributions of participants also tend to happen in parallel rather than one after the other and it is a form of instantaneous production where composition is also a performance in its own right. For all these reasons, it is an exciting area to study, not least because current approaches to understanding group work don't adequately model features of group engagement.

The aim is not just to understand what is happening in such sessions, of course, but to look at how new computer technology might be developed to support those activities even, for example, when the participants are spread across continents.

Nick's approach has been first to identify the features that matter in such situations through recording collaborations, such as Jazz jamming sessions, that involve high levels of mutual engagement. The AHI Lab allows this to be done to a level of detail far greater than could be done just by normal filming. The data obtained is analysed to find out about things such as which performances were mutually engaging and why, when group improvisation occurred and when problems with communication or co-ordination occurred and why. The results are then used to design a framework for understanding mutual engagement and patterns of communication in collaboration. Based on this understanding prototype tools to support group improvisation are being developed to support richer forms of human-human and human-computer interaction.

It is from research like this that social networking could take a further leap forward with technology supported mutual engagement happening much more naturally in the future.

IMC

The Interaction, Media and Communications (IMC) group is a unique multi-disciplinary group combining computer science, philosophy, and psychology to exploit the potential of digital technologies to support novel ways for people to interact.

The group's research encompasses for example: speech and gesture; combining listening to data (data sonification) with reading text; investigating how we co-ordinate graphical and verbal dialogue and developing ways of monitoring and manipulating the interaction in chat rooms; as well as examining how future technology might support people communicating with each other when jamming in a jazz session.

Biology and Computer Science

The Hoverflies are coming to get you

Computer Science and Biology at first sight might seem to have little in common: one is concerned with the natural world, the other with man-made technology. In reality the story is much more fascinating. Computer Scientists have adapted lots of techniques from biology, such as genetic algorithms and neutral nets. They are a way computers are used to solve problems that is based on evolution or the way brains work. It runs the other way too with computer models giving biologists insights into the way nature works. Hoverflies are one example…

When hoverflies get the hots for each other they make some interesting moves. Biologists had noticed that as one hoverfly moves towards a second to try and mate, the approaching fly doesn't go in a straight line. It makes a strange curved flight. Peter McOwan, a Professor in the Vision group, who is interested in biologically inspired computer science, thought this was an interesting observation and started to look at why it might be. His team came up with a cunning idea. The hoverfly was trying to sneak up on its prospective mate unseen. The route the approaching fly takes matches the movements of the prospective mate in such a way that, to the mate, the fly in the distance looks like it's far away and 'probably' stationary.

How does it do this? Imagine you are walking across a field with a single tree in it, and a friend is trying to sneak up on you. Your friend starts at the tree and moves so that she is always in direct line of sight between your current position and the tree: as she moves towards you she is always silhouetted against the tree. Her motion towards you is mimicking the stationary tree's apparent motion as you walk past it… and that's just what the hoverfly does when approaching a mate. It's a stealth technique called 'active motion camouflage'. By building a computer model of the mating flies, Queen Mary computer scientists have been able to show that this complex behaviour can be done with only a small amount of 'brain power'.

Peter's team went on to show that humans are also fooled by active motion camouflage using a computer game where you had to dodge missiles … and some of those missiles used active motion camouflage. The missiles using the fly trick were the most difficult to spot, showing that there is such a thing as a useful computer bug!

Biology meets Computer Science

Peter's work, joint with UCL, was selected for the prestigious Royal Society Summer Exhibition in 2007. Their 'Perception Deception' exhibit focused on optical illusions and their role both in helping Scientists understand the way our brains perceive the world and to develop novel computer applications.

EPSRC Engineering and Physical Sciences Research Council

Summer Science Exhibition 07

Queen Mary University of London