The Royal Society of Edinburgh
Franco–Scottish Science Seminar Series
Computer Science

The New Alliance in Science

Thursday 13 February 2014

Report by Peter Barr

Executive Summary

The seminar on computer science was one in a series of events organised by the Royal Society of Edinburgh (RSE) and the French Embassy focusing on areas of mutual interest and collaboration in science, reflecting “the strong and productive relations” between the two countries.

The major topics covered were modelling, security, algorithms and learning, with eight individual presentations covering a series of related topics, including biological and stochastic simulation, probability and refinement in the verification of programming code, using models to predict the behaviour of hackers, statistical learning and computer science education.

The main themes were selected to highlight areas of excellence and close collaboration between researchers in the two countries, and also underlined the common themes which run throughout so many different subjects in computing, such as stochastic modelling, security and machine learning.

Collaboration between computer science researchers in Scotland and France was a recurring theme in several papers, including projects involving the University of Edinburgh and INRIA (the National Institute for Research in Computer Science and Control) in France, from the 1970s to the present day. Several speakers also embodied this “alliance of science” – researchers who have moved from France to do their work in Scotland.

Although the individual presentations may have appeared on the surface to be very different in technical terms, ranging from cell simulation and computer security to the esoteric details of programming code and machines that can learn how to think for themselves, there were also many points of common interest, such as bridging the gap between theory and practice, the status of computer science in academia, the need for education and issues of public perception – not just our attitudes to the development of intelligent learning machines, but also our strategy for “learning humans” and “learning about learning.”

The human dimension was also discussed in a number of papers, including the fact we are different from learning machines, the suggestion that we are a “nightmare” for cyber security and the need to track “ordinary” human activity in social media to analyse patterns of behaviour and prevent potential terrorist attacks.

Overview

The connections between France and Scotland are not just the historical links between two friendly nations who formed a political alliance in 1295, but also the "science alliance" which has lasted for hundreds of years and continues to thrive, especially in areas such as computing.

In his introduction, Professor Alan Alexander OBE FRSE, the General Secretary of the RSE, said that some of the linguistic connections between France and Scotland may have been lost through the years – e.g. French words used in Scots which are no longer in common usage today – but the intellectual and scientific connections are still very strong, especially in the "fast-moving and exciting" field of computing, where there is a lot of pioneering and innovative work going on.

Dr Cyrille van Effenterre, the Science and Technology Counsellor of the French Embassy, described his own department's mission as an interested observer of UK science and technology policy, a promoter of collaboration in science and technology, and a promoter of French scientific achievement. He also talked about his country's "special relationship with Scotland" and how he hopes these seminars will continue this collaboration between the two countries and showcase the excellent work being done in computing and maths – e.g. in bioinformatics, Artificial Intelligence, the new generation of self-learning, thinking machines, and cyber security. "We need to share our expertise at an international level to strengthen the resilience in the area of security," he added.

There were several examples during the seminar of historical links in science between Scotland and France. Dr Jean Krivine of the PPS Laboratory in Paris talked about the long co-operation in research into signalling pathways in France, Scotland and the US, and Professor Andrew Gordon of the University of Edinburgh discussed the importance of the work of Thomas Bayes (who studied in Edinburgh) and Pierre-Simon Laplace, who although they were from different generations, made enormous contributions to probability theory in the mid- to late-18th Century.

Several contemporary collaborative projects were also highlighted during the day, including one to improve computer security involving the University of Edinburgh, INRIA (the National Institute for Research in Computer Science and Control) in France and Microsoft Research. Professor Gordon also talked about a new "Auld Alliance" in cyber security, using probability and refinement to improve the verification of digital transactions. In addition, several speakers hold joint positions in Scotland and France, or started their careers in France and are now based in Scotland, keeping up the tradition of academic exchange between the two countries.

The intellectual links between different topics emerged clearly during discussion, particularly the contribution made to synthetic biology by computer modelling of biological systems, the importance of stochastic modelling and algorithms not only in biology but also in security and simulation (including applications such as self-driving cars, and testing plane designs before construction, as well as biophysical systems), and the increasing importance of learning machines. One speaker also suggested that biophysics or synthetic biology will eventually become a branch of machine learning.

The speakers also discussed the importance of probability in the verification of code – e.g. to ensure the safety of banking transactions – and attempts to create visual maps of system and network activity which look like simulations of particles or cell interactions, and many other complex processes. In the first case, scientists are trying to evaluate the chances of success for a particular drug, and in the second case they're trying to detect "security violations" or predict the criminal actions of hackers by studying their patterns of behaviour. There were also parallels between the challenges involved in interrogating data and investigating breaches of cyber security – sometimes trying "to detect what we don't know."

Another theme mentioned on several occasions was how to turn these highly complex theories into practice – converting smart ideas into practical products or applications that people, government and businesses will use in the future.

The need for education also came up during proceedings – not just the need to train young people for the digital future or promote the study of statistics among students in general, but the need to win more recognition of computing as a science on a par with other scientific disciplines.

Presentations

1. Modelling

a. Broadcast Signalling in Cells

Professor Vincent Danos
Directeur Recherches at CNRS (Paris, PPS lab),
Professor at the University of Edinburgh and Director of Synthsys

Professor Danos described his research into neuroparticles, or the modelling of cells. This is a relatively new field which can be viewed in several different ways, including “macroeconomics” or engineering. His work is also concerned with “concepts rather than results” – for example, treating cells like crowded cities full of traffic jams. On the surface, the behaviour of the cells looks chaotic, but they do produce something meaningful and it is possible to make sense of their “noisy” environment, through simulation. There are global pressures on the cells and competition between them, but the question is how to exploit their different properties and interactions “to do something better” – i.e. change their behaviour in order to cure a disease. And first, researchers have to understand their behaviour and how they will react to the introduction of different new ingredients, incorporated into pharmaceuticals.

Cells can also be viewed as circuit diagrams, complete with input and output – leading to a series of chemical reactions such as glucose transformed into proteins, or DNA replicating itself. This is a “very convenient and predominant view” of cells, Professor Danos continued, looking at them from the perspective of an electronics engineer.

Theoreticians “forget all this technical detail” and look at what it means to be self-replicating, studying the principles involved, treating the cells as if they are “minimal replicating machines,” using a template they get from the information in their genetic material. This “unique machine reads all the instructions,” he said, “and can make more of itself,” and this view is now coming back into fashion.

The challenge in synthetic biology is the need to design new drug therapies which act “in sympathy with the host.” You can take “off-the-shelf” chemical reactions and add to the host – e.g. to provoke or prevent a toxic reaction – or create new enzymes to inhibit a process. This can lead to unexpected “interference” or reactions, and is a dynamic process involving “synthetic pathways that adapt to changing environments,” switching receptors and repressors on and off to trigger a synthesis of what is required.

Professor Danos then discussed what he described as “thought experiments,” and the attempt to understand the translation or expression of genes, in terms both of velocity and how successful they are at translation (growth rate and optimisation), in different conditions – e.g. when there are lots of resources available or there is a lot of competition for resources. He compared this to the complexity of understanding weather conditions, and went on to discuss the “macroeconomics of the system” and the “GDP” of the cells. And as in economics, synthetic biologists have to make certain assumptions about the behaviour of cells in order to run simulations.

All components of the cell react very differently, Professor Danos continued. “The faster the cell grows, the fewer transporters it grows.” This could be compared to social planning, he added, and the optimal allocation of resources. “You assume that the host has evolved so that it will allocate resources optimally,” he said. Most microbes can switch from being slow and doing things well to acting fast and being inefficient, and that is why it is important to understand the dynamic processes involved.

Ultimately, no matter which view of the cells is adopted to understand what's going on and how the cells will respond to the introduction of new proteins, the challenge is how to exploit the signals used by cells and manipulate or engineer them for therapeutic purposes – to cure diseases and make people healthier.

b. Non-local Causality Analysis in Rule-based Models

Dr Jean Krivine
CNRS researcher (CR) at PPS Laboratory

Dr Krivine began by drawing attention to the long history of co-operation between researchers in France, Scotland and the US in the field of signalling pathways in cells, and the use of computer-based models to help build new drugs – using computing science to describe and analyse complex biological systems.

He then discussed what are called "signalling cascades," or what's going on in the cell in terms of the reception, transmission, repression, amplification and integration of information, as inputs trigger responses in cells. One signal leads to different responses, he continued, depending on the number of ribosomes in the environment. Dr Krivine also mentioned the role of "abstract interpretation," and described the cell's behaviour as "seeking consensus," as the signal cascade makes things happen and cells interact with each other.

The cells are a small ecosystem and the functions of the individual proteins can be interpreted in various ways, some of which can be misleading. The analogy is when we study a field of potatoes to find out (and improve) the growth rate and health of the plants. To understand what's happening, we study different processes and elements, including the worms, which help to activate the growth of the potato. The worm could be described as a "potato activator," but it has many other functions which should not be ignored. Therefore, said Dr Krivine, it is wise to take an "agnostic" or reductionist view of what the worm is doing. In the same way, we should not relate the protein to one final outcome. "We should trust facts not functions," he added. And in his work he uses simulation to determine if a particular gene is activated by a particular protein, applying a set of stochastic rules to model the process and find out what happens. Dr Krivine is also concerned with "which transitions are causing the appearance of the observation," and compared this to a security violation in a computer network.

"Classic computing techniques can't be applied to biology," said Dr Krivine, "so we have to do something different," when it comes to studying "the causal dependencies between rules" and working out the rules to describe what is happening inside the cell. You can run a simulation for enzymes and proteins, with various values or functions, but sometimes this can produce "very weird results."

The model is "a metaphor" of what the proteins are doing, and Dr Krivine then used a real-world example, discussing how researchers may want to prevent a cancer cell growing by targeting a protein. To get the best results, some rules are removed from the model, and the aim is to have a "minimal explanation of some rules" in the simulation, to establish the pathway.

Q&A:

Q: Can simulations be used to study diseases and faulty cells, as well as healthy cells?

A: Looking at simple bacterial cells, there are huge problems, because the infections may do very well but the humans may not. That is why there are no value judgements – just analysis.

Q: Future trends? Achievements? What's the Holy Grail of modelling?

A: In the future, it should be possible to design drug treatments for specific individuals. We provide biologists with executable models – computer science can deliver knowledge and the 'execute' button. Eventually, biophysics will be a branch of machine learning.

2. Security

a. Probability, Refinement and Security: An Auld Alliance, A New Future

Professor Andrew Gordon
Principal Researcher, Microsoft Research, and University of Edinburgh

Security techniques are based on probability and refinement of programming code, and Professor Gordon's talk described the challenges involved in the 15-year joint project to develop F7, "an enhanced typechecker for the F# programming language, a dialect of ML which pioneers the static checking of security properties expressed with refinement types."

This joint project is an excellent example of the scientific Auld Alliance in action, involving a collaboration between Microsoft Research, the University of Edinburgh and INRIA, which receives substantial funding from the French Government.

Professor Gordon provided some historical context by describing the pioneering work of Thomas Bayes, who studied in Edinburgh and developed the theory of "conditional or inverse probability" in the early 18th Century, and the French mathematician Pierre-Simon Laplace, whose name is linked today with "Laplacian noise," or the issues faced in "differential privacy" – how to query a database without giving up too much information. He also explained conditional probability – how the apparently equal probability of tossing a coin and predicting heads or tails can be affected by knowledge of a prior event or an update of the probability. Another connection highlighted by Professor Gordon was the invention of the ML programming "meta language" by Robin Milner in Edinburgh in the 1970s – a project which also involved collaboration with INRIA.

Refinement is the "Holy Grail" of computer programming, because it aims to stamp out all possible bugs in the code. It is also becoming increasingly important because software code is so embedded in everyday life, a dependency which can lead to fatal results. "At last, we are making real progress on verification," said Professor Gordon, referring to what has been described as the "war on error" and the effort to "take the battle to languages."

Professor Gordon then reviewed the history of cyber security and crypto-protocols, and the early (failed) attempts to create a secure tunnel between the parties engaged in transactions – e.g. to transfer funds from one bank account to another, via the Internet. "The web service was intended for use by programs, not humans," he said, explaining that the problem started with trying to identify the sender of any instruction. To get around this problem, programmers added a security header and a signature block to the signature sign-in or password, but that didn't work because a hacker could still edit the message to fake it by creating bogus headers, changing the account name and the amount. Encrypting code is hard to get right, so we have to look at the actual code itself to verify the transaction, he added.

"There are lots of security vulnerabilities," Professor Gordon continued, explaining how responses to multiple requests can get mixed up during transactions, thus requiring extra data to be added to the code to ensure message verification.

Verification of the actual protocol code arrived in 2005, with the development of TLS (transport layer security), but Professor Gordon also asked: "How secure is TLS?" He then discussed the next stage in security with the introduction of what is called "probabilistic verification" in the new F7 typechecker being developed by the new "Auld Alliance."

"F7 achieves a flexible and scalable verification of security code," he explained, and is also a highly effective "alliance" of probability and refinement which enables significant progress in verification.

b. Digital Trust & Security Innovation Roadmap: Implementation around French SYSTEMATIC Cluster

Emmanuel Miconnet
Research & Innovation Director at Thales

Mr Miconnet gave a wide-ranging talk on recent innovations in cyber security, starting by describing the SYSTEMATIC Cluster, a project based in Paris which was set up to bring new ideas in digital technology to market, bringing together an alliance of researchers, industry and government. The cluster focuses on system design and the development of new tools for a range of applications, including smart energy management, healthcare, automotive and transport, sustainable cities and digital trust, Mr Miconnet's specialist area – a total of more than 350 projects with a budget of over 600 million euros. Another key aim is to support the growth of innovative SMEs (small to medium-sized enterprises).

“Digital trust” brings together 155 partners, including 42 large companies, 24 research labs and numerous SMEs, focusing on safety and security – risk management, prevention and resilience in critical infrastructure. One example of a commercial product being developed is a new toolbox to detect and investigate credit card fraud on the Internet.

According to Mr Miconnet, there is a convergence of physical and logical security – e.g. using a password to access a network is similar to wearing an ID badge to enter a building. “There is no mismatch between what I do in the physical world and the logical world,” he said. The work of the researchers also models fraudulent behaviour and studies the interactions in networks, as well as looking forward to the Future Internet (FI), when “FI ware” will deliver new Internet services. The researchers also monitor and simulate “business ecosystems” and the need to connect applications to the physical world. Even in the field of security, business will “benefit from open innovation,” he added.

We need to change the way we protect our information systems, said Mr Miconnet, moving from the old approach based on perimetric (which involves too much data) and static (too costly) defence to a new more proactive approach. This involves identifying emerging threats and studying behaviour patterns, tracking cyber criminals and injecting “real-time intelligence” into our systems.

The cyber threats to information assets in business include denial of service (DoS) (which can bring down a network and lose lots of money), sabotage (which can damage the supply chain and cost a lot to rebuild), and espionage (which compromises confidentiality and leads to a loss of competitiveness, and propaganda (loss of reputation and confidence).

Mr Miconnet then described a cyber attack on the French infrastructure in 2011, which was penetrated by an unsecured administrator and enabled a hacker to “steal” information on a major project – “a huge data leak” which took 200 people from his company (Thales) three months to fix.

There are three approaches to cyber security: static (focusing on information systems security, design and encryption), active (cyber defence focusing on education and enhanced awareness) and proactive (cyber intelligence focusing on social media and predictive analytics). Mr Miconnet also said that security has to involve “products, solutions, processes and people” in order to work.

When it comes to threat and detection, there are known attacks (which can be defended against using firewalls and rules) and unknown attacks, which can involve anomaly detection and “searching for something you don't know about.” Whichever company develops a solution for unknown attacks will have “a serious competitive advantage,” said Mr Miconnet.

The new trends in security include a move from hackers to the need for protection from potential attacks from enemy states, and a step up from network-level attacks to mission- or service-level, and the need for intelligent cyber defence as well as faster and enhanced response to attacks. We need to build “integrated cyber threat management” into our systems, not just to detect and investigate hackers but also to study emerging threats – sometimes described as setting up a “cyber rapid reaction team.”

The key themes today are the need to adopt a model-based approach and “big analytics” or the use of learning machines. People are also a problem – when hackers navigate their way through to vulnerable assets, it is “a nightmare” to defend the physical machines that people can change. To counter attacks, the networks of the future will have to harvest data every day to monitor system behaviour, and Mr Miconnet displayed an illustration of the visual maps (IS dynamic mapping) which highlight odd (anomalous) activities and show everyday interactions, to “detect what we don't know.” For example, he explained, you may detect unusual levels of activity on a Sunday when the network is supposed to be quiet, which alerts you to a possible attack.

Cyber intelligence is another key defence tool, analysing social media and tracking hacker communities to anticipate and prevent attacks. “You have to model what the attacker will do,” said Mr Miconnet. “The problem is: What are the limits of what the attacker can do?”

“We are fully convinced that we have huge technical resources,” he continued, “but where we fall down is how we transfer all this academic knowledge and these algorithms into commercial products. We are still far away from developing tangible applications,” although there has been lots of progress in improving security in computer code. Sometimes, he added, we simply don't know what we're looking for, while automated threats (e.g. programmed bots which infiltrate systems) can be very hard to detect. Models only give an approximate view, highlighting abnormal behaviour, so to defend more effectively, we have to merge different approaches.

Q&A:

Q: How does current work in security fit with Horizon 2020 (the biggest EU research and innovation programme ever, with nearly 80 billion euros of funding available over seven years)?

A: The technology needs to mature and have real market impact. This will require partnering with innovative SMEs.

3. Algorithms

a. Incomplete Information in Web Database Systems

Dr Amelie Gheerbrant
Lecturer, Universite Paris Diderot

Dr Gheerbrant talked about the foundations of XML, a mark-up language used in computing “that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.” The web brings new challenges to data management because web data is intrinsically different from traditional databases, because there are so many dynamic updates as well as missing variables. An XML document (designed for use on the Internet) has a different structure, or topology, which has to be considered when making a query.

“Big data interoperability leads to massive incompleteness,” said Dr Gheerbrant, and this means problems in “query evaluation and query rewriting (static analysis and verification).” When there is incomplete data, she continued, the query doesn't work any more – e.g. as the data is updated. So what should we do to prevent this? “Should we trash the semantics or change our database management systems?” she asked.

Dr Gheerbrant then used the example of a simple enquiry involving someone flying from Edinburgh to Paris then onwards to Santiago, staying in different (unknown) hotels in each city. The answer to a query may be “every possible world,” because there is incomplete data, and we should “look at the query as if it is an incomplete database.” Dr Gheerbrant then discussed homomorphism (“a transformation of one set into another that preserves in the second set the relations between elements of the first”), schema mapping and the powerful algorithms needed to cope with data complexities, and said there was a need to “reconcile theory with systems and reduce the complexity” in web database systems – as well as to “bridge the gap between theory and practice.” She also pointed out that “data cleaning” is a very costly business, so any improvement in database systems would be welcome.

The project Dr Gheerbrant is currently working on involves collaboration between her institution in Paris and the University of Edinburgh "to develop efficient algorithms for query evaluation, optimisation, approximation and static analysis." The researchers have found new classes of relational and XML theories, and are now looking at how to "rewrite queries with constraints in data values." One concrete problem is the exchange and integration of web data or "approximate" queries. The aim, she explained, is to "develop one uniform theory which accounts for various forms of incompleteness arising in all data models." The researchers have also found a way to identify classes of theories and standard query evaluation procedures which give certain answers, and are ultimately looking for applications for data interoperability.

b. Branching Stochastic Processes

Dr Kousha Etessami
Reader, Laboratory for Foundations of Computer Science, University of Edinburgh

Dr Etessami changed his theme from "branching stochastic games" to talk about some "easier-to-understand" aspects of branching stochastic processes – "a sometimes random process or collection of random variables which is often used to represent the evolution of a random value or system, over time." He also explained that his work is part of a joint project with Columbia University in the US.

Building basic models of a branching stochastic process can be relevant to studying biological processes, including the potential success rate of treatments for diseases such as cancer. The concept was first used in the 18th Century to model the disappearance of particular names in aristocratic families – i.e. the probability of daughters leading to the end of a line because their family name does not survive them.

The "simple" example used by Dr Etessami to illustrate branching stochastic processes concerned cells which produce different offspring or do not produce offspring and die. The stochastic process, he explained, is the probability that "generates the random label tree." His graphic illustration focused on a very small number of cell types and the probability of individual cell types becoming extinct, but he pointed out that with cancer, even a single tumour contains a huge variety of genetic material and cell types which undergo rapid mutation, so analysing or targeting this "mess" of many different and fast-changing cells can be extremely complex, because different cell types may require different drugs. "Stochastic processes have many applications," he explained, not just in biology but even in predicting nuclear reactions – asking if a bomb will fizzle out or explode.

Dr Etessami then talked about the complex mathematics involved in branching stochastic processes, covering topics such as the "Jacobian matrix of partial derivatives, non-linear fixed-point equations and polynomial systems of equations (PPS)," and explaining how the extinction probabilities in his example are "the least fixed point." Part of his explanation also concerned the usefulness of Newton's "worst-case behaviour for PPS."

In practical terms, taking the example of cancer again, Dr Etessami explained that the desired outcome may be the extinction with the highest possible probability, and that stochastic models simulate what happens when we introduce a drug. Using linear programming, we can calculate and model very efficiently what happens to the cells which we control, he said, but when the same model is applied to game theory, the process becomes much more complex, because we are now concerned with the unpredictable actions of opponents who control their own resources.

According to Dr Etessami, he and his colleagues have "established the first polynomial time algorithms for one of the most basic analytical problems – i.e. extinction probability – for multi-type branch processes." Cancer would be harder to study, he explained, because it involves continuous time models and would also require converting the model to thousands of variables. But the model would still be able to cope, he suggested, despite the complexity, and rise to the challenge. Would a single drug work better, or several drugs simultaneously? If the tumour has 10,000 different cell types, how to target individual cells?

The work by Dr Etessami and his colleagues is now moving forward from theory to practice. "Now that complex problems are resolved," he said, "we now want to work with practical applications," and researchers in the University of Edinburgh mathematics department are already using the model for studies of cancer.

4. Learning and Future Collaboration between France and Scotland

a. Statistical Learning and New Challenges of Large-Scale Inference

Professor Olivier Cappe
TCI, Telecom Paris Tech & CNRS

Professor Cappe focused on the basic principles, methodologies and mathematical ideas underlying statistical learning – "learning to perform tasks from examples."

He began by describing different ways of statistical learning, including supervised (using only examples), unsupervised (using only unlabelled data) and sequential or online (using data only available sequentially) approaches. He also talked about reinforcement and the issues involved in observing and interacting with data. "In large-dimensional models," he asked, "how can we deal with 10,000 (or even 100) parameters based on only one observation?"

Statistical learning needs well-defined tasks and goals, Professor Cappe continued. It is "largely disconnected from learning by humans" and has been the dominant approach in machine learning for the last 20 years. Statistical learning is not the same as Artificial Intelligence (AI), which suggests an anthropomorphic process. Humans can learn without examples (e.g. learn how to play guitar by following instructions) while machines can learn by analysing tens of thousands of examples (e.g. by watching videos on YouTube) and mimicking what they observe.

Professor Cappe then reviewed the history of statistical learning, starting in the 1980s when speech recognition, natural language processing and handwriting recognition were the flagship applications. In the 1990s, attention turned to learning theory and "convex optimisation," and in the 2000s, methods became more important, including "sparsity, compressed sensing and non-negative factorisation." Professor Cappe also mentioned the Netflix Prize, a million-dollar prize for scientists who could solve a particular problem in statistical learning, taking data to perform a particular task. Netflix did not get an application for its business from the winning solutions but it did get a lot of publicity, he added.

The next topics covered were the theoretical basics of supervised learning, including observation and probability distribution, decision rules (e.g. best possible decision or the best you can do) and the issue of risk (with the objective to minimise risk). Professor Cappe then explained that using a "regulariser in the equation could prevent the decision rule becoming too complex," adding that "numerical optimisation is a key ingredient" of any calculation. This is how machines learn, he added.

With unsupervised learning, he continued, "loss is not really well-defined," and the role is to construct higher-level features from the data. Unsupervised learning is particularly relevant for image-related applications, he said, because the data are so complex and diverse. Professor Cappe also touched upon "news story segmentation" and said that big data and statistical learning are "two different subjects entirely."

"Is statistical learning data science or science from data?" he asked. The traditional methods used for detection of risks and reliability testing can be replaced by data-driven methods, he suggested, and one good example is plane simulation to see how the design will perform even before it is built. "This corresponds to a major evolution in several fields of science and engineering (e.g. life sciences, particle physics, geoscience and cognitive learning)," Professor Cappe added.

Statistical learning is performed very quickly and incrementally, using sampled data, and one issue is how to take the methods used for smaller problems and scale up to much larger problems, and how to assess the statistical significance of results.

b. Some Reflections

Professor Andrew McGettrick FRSE
Emeritus Professor, University of Strathclyde

Professor McGettrick focused on his work as the Chair of the Educational Board of the Association of Computer Machinery (ACM), and the "educational challenges of the day." There was a decline in interest in computer science in recent years, he suggested, but now there is a welcome resurgence of interest. The ACM is an international organisation, with about half of its members from the US and 15 per cent from Europe.

As well as issuing advice on the computer science curriculum (e.g. big data, data analytics, security, concurrency and online learning), the ACM has set aside funds for the study of cyber security and is organising a conference called L@S or "large at scale," to be held in early March in Atlanta, Georgia. The conference will focus on online learning and the impact of massive open online courses, (MOOCs), and "promote the scientific exchange of multi-disciplinary research at the intersection of learning sciences and computer science." It will also promote the study of algorithms, modelling and simulation, learning by machines and learning by humans, as well as "learning about learning."

Cyber security is another hot issue. The US Government says it needs about 40,000 security professionals, but last year there were only 67 PhD students specialising in security, and only one of them went on to teach the subject.

The status of computer scientists and software engineers is another concern. Professor McGettrick said that software engineers should be regarded on the same level as other scientists and engineers.

Computer science education is evolving all the time. How should we teach computational thinking and simulation? How deal with complexity? And how measure learning? There is a need to teach statistics to more students, including those doing computing. At the same time, we should not forget the "wonder" of it all – motivation and the "wow factor" in education. Professor McGettrick quoted the example of a teacher in the US who told her committee, "Forget about education – make programming fun." She now has 200,000 online subscribers.

Professor McGettrick also talked about his own university, and quoted the words of its founder John Anderson, who said in 1796 that it should be open to everyone, "regardless of gender, status and income." Nowadays, the university describes itself as "the place of learning," and he said that this needs to be re-interpreted, in view of changes in education.

What about the future of computer education? It has been said that "the role of computers is not to make users more intelligent but to support their ability to analyse data," and this view is still relevant today.

The themes discussed at the seminar were very important and research and education need to work together, not just to explore more sophisticated areas of computer science, but also to make the subject more accessible.

Q&A:

Q: What about deep learning?

A: "More people in cognitive science are now getting interested in machine learning," said Professor Cappe, who then compared testing the safety of planes to testing new drugs.

Professor Gordon commented that statistical models can be represented in code and that one day this will be taught in schools, because the jobs of the future will demand statistical and coding skills. It was also suggested that the teaching of statistics to scientists is also a neglected area, especially as we move into an era of self-driving cars and learning machines.

Conclusion

Professor McGettrick reminded the audience that the origins of the seminar were the Auld Alliance of 1295, when Scotland and France declared that they would help each other during times of trouble. Today, he said, we need to help each other in a lot of different areas, including computer science, particularly the subjects covered during the day:

- modelling biological systems and dealing with and understanding complex behaviour and inaccurate reflections of reality – including the “security violations” we observe in the models
- rapid progress in refinement of computer code
- modelling cyber security – active, passive and forward-looking or proactive – and machine intelligence
- algorithms
- the foundations of XML
- stochastic modelling
- statistical learning
- machine learning
-

Finally, Professor McGettrick observed that machine learning was a theme that ran throughout the day. In Stanford University, he added, all 1,800 first-year students did a programming course and this may be a lesson for us all.

Opinions expressed here do not necessarily represent the views of the RSE, nor of its Fellows