

# ERCIM NEWS

[www.ercim.eu](http://www.ercim.eu)

Special theme:

## What is Computation?

Alan Turing's Legacy

Also in this issue:

**Keynote**

The Impact of Alan Turing  
by Andrew Hodges

**Research and Innovation**

Ensuring Profitability of Commercial  
Long Term Digital Preservation  
by Stephan Kiefer and Michael Wilson

Cybercrime and the Security of Critical  
Infrastructures  
by Florian Skopik and Thomas Bleier

<b>KEYNOTE</b>		
<b>3 The Impact of Alan Turing</b> by Andrew Hodges	<b>14 Recent Advances in the Formal Verification of Cryptographic Systems: Turing's Legacy</b> by Benjamin Grégoire	<b>30 GoodShape: Towards Flexible Mesh Generation</b> by Bruno Levy
<b>JOINT ERCIM ACTIONS</b>		
<b>4 Cooperation with Georgia's ICT Research Centres: The Second GEO-RECAP Networking Event</b> by George Giorgobiani and Givi Kochoradze	<b>15 From Discrete to Continuous: Turing's Morphogenesis</b> by Nadia Pisanti	<b>31 Math Strengthens the Swedish Olympic Cross-country Team</b> by Kersti Hedman
<b>5 Joint ERCIM eMobility and MobiSense Workshop</b> by Torsten Braun	<b>16 Alan Turing and Systems Biology</b> by Anna Gambin, Anna Marciniak-Czochra and Damian Niwinski	<b>32 LUDUS: Serious Gaming Initiatives in South East Europe</b> by György Kovács
<b>6 MUSCLE Working Group Co-organised Mass Data Analysis Conference</b> by Petra Perner and Emanuele Salerno	<b>RESEARCH AND INNOVATION</b>	<b>33 A New Robotic Laboratory at SZTAKI</b> by György Kovács and Imre Paniti
<b>7 17th International Workshop on Formal Methods for Industrial Critical Systems</b> by Radu Mateescu	This section features news about research activities and innovative developments from European research institutes	<b>35 Engineering Asset Lifecycle Optimal Management: WelCOM Approach to E-Maintenance</b> by Christos Koulamas, Petros Pistofidis and Christos Emmanouilidis
<b>7 Annual Meeting of the ERCIM Working Group on Models and Logics for Quantitative Analysis</b> by Diego Latella	<b>18 The Parallel Heartbeat of Statistical Text Analysis</b> by Tobias Berka and Marian Vajteršic	
<b>SPECIAL THEME</b>		<b>EVENTS, BOOKS, IN BRIEF</b>
This special theme section "What is Computation — Alan Turing's Legacy" has been coordinated by the guest editors Gilles Dowek, Inria, and Samson Abramsky, University of Oxford	<b>19 Ensuring Profitability of Commercial Long Term Digital Preservation</b> by Stephan Kiefer and Michael Wilson	<b>36 CLEF 2012 and Beyond: Perspectives for the Conference and Labs of the Evaluation Forum</b> by Nicola Ferro
<b>8 What Makes Alan Turing a Great Scientist?</b> by Gilles Dowek and Samson Abramsky	<b>21 Learning to Recall</b> by Jaldert O. Rombouts, Pieter R. Roelfsema and Sander M. Bohte	<b>37 Turing Year in Spain</b> by Juan José Moreno Navarro
<b>10 Viruses in Turing's Garden</b> by Jean-Yves Marion	<b>22 The Computer and the Brain, Synergies and Robots</b> by Martin Nilsson	<b>38 HCI International 2013</b>
<b>11 When Turing Meets Milner</b> by Jos Baeten, Bas Luttik and Paul van Tilburg	<b>23 Advances in Model Driven Software Engineering</b> by Mark G.J. van den Brand and Jan Friso Groote	<b>38 Book Review: "A Multi-disciplinary Introduction to Information Security"</b>
<b>12 How to Compute with Metabolism in Bacteria?</b> by Claudio Angione, Pietro Liò and Giuseppe Nicosia	<b>25 Software Engineering for Multi-core Platforms</b> by Farhad Arbab and Sung-Shik Jongmans	<b>39 ERC President visits Turing Exhibition</b>
	<b>26 Cybercrime and the Security of Critical Infrastructures</b> by Florian Skopik and Thomas Bleier	<b>39 PROMISE Retreat 2012 report on Prospects and Opportunities for Information Access Evaluation</b>
	<b>28 VoterBallot - A New Application for ICT in Elections</b> by ZazaTabagari, Zaza Sanikidze and George Giorgobiani	<b>39 PLERCIM's Unit Awarded the Status of Leading National Research Centre in Poland</b>
		<b>39 Obituary for Horst Santo</b>

## The Impact of Alan Turing

The 2012 centenary of Alan Turing's birth has enjoyed a level of public awareness that is remarkable for any scientific figure. This is in part due to the great change in the perception of his homosexuality since the 1990s: young people can scarcely believe that British criminal law was as it was in 1952, and there has been much agitation for some sort of posthumous adjustment to his conviction. Unfortunately, just as Roger Bannister and the Comet crashes represented that particular era in Britain, so too did Turing's conviction. This is an immutable historical fact.

In parallel, also since the 1990s, the public understanding of computers has changed. Computers are not remote installations, but hubs of free personal communication, and are becoming increasingly in tune with Turing's vision. In fact it is only recently, with general-purpose chips taking over ever more functions, that the idea of the universal Turing machine has really been vindicated. When he died in 1954, Turing would hardly have been seen as a towering figure — and not just because his war work remained totally secret until the 1970s. When elected Fellow of the Royal Society (FRS) in 1951 for his 1936 work on computability, his work was barely appreciated outside a small academic field, and was not considered of practical importance. In 1953 the first British book on computers ridiculed Turing machines as "incomprehensible".

Perhaps what is most distinctive about Turing is that although in 1936 he addressed the very abstract and unfashionable material of mathematical logic, unmotivated by any prospect of economic benefit, he never spurned down-to-earth application. The commonplace picture of him as a dreaming theorist misses the mark. His codebreaking work, turning logic and probability into engineering, made a critical contribution to 1945 and indeed the post-1945 world of Anglo-American dominance. His 1945 design for the ACE computer, now the centrepiece of a special Science Museum exhibition, makes visible his eagerness to engage with technical electronics — although his prospectus for what would now be called software development was really the most powerful aspect of his plan.

This was never really followed up and one weakness of Turing's scientific career was that he did not publish more of his far-sighted ideas. When he chose, he could make a great impact with published papers, and he was not shy about explaining Artificial Intelligence in radio talks. But he was impatient with the more routine work of pressing home his arguments, always eager to move on to new explorations. Many scientists today, serial writers of research proposals, will sympathise.

The questions that most excited Turing are still alive and well. After writing a classic work on Artificial Intelligence in 1950, he turned to mathematical biology, and his models of growth are now the focus of much exciting research. In his last period, Turing was also looking afresh at quantum mechanics, and the connection of logic and physics remains a fundamental problem in modern scientific thought. There is much about Alan Turing's centenary that speaks not to 1912, but to 2012.



*Andrew Hodges, mathematician and author of the book "Alan Turing: The Enigma".*

*Andrew Hodges*

# Cooperation with Georgia's ICT Research Centres: The Second GEO-RECAP Networking Event

by George Giorgobiani and Givi Kochoradze

*The 2nd Networking Event of the GEO-RECAP project (Re-creation and building of capacities in Georgian ICT Research Institutes) was held at the Georgian Technical University (GTU) on 27-28 June, 2012. The event was held together with the EC funded Ideal-ist project (an international ICT partner search network) twinning event, which was organized by the EC ICT National Contact Point (NCP) in Georgia.*

The guests were invited from Belgium representing the ERCIM office, Sweden (KTH - Royal Institute of Technology); Ukraine (Kiev Technical University), Moldova and Azerbaijan (EC ICT NCPs). The event was attended by 30 or so scientists. At the opening presentation, the speaker emphasized that ICT is no longer a luxury but a necessity for developing countries and that developing countries are already creating new ways of communication, doing business, and delivering services. Through extending access and use of ICT, the European Commission FP7 ICT program aims to stimulate sustainable economic growth, improve service delivery and promote good governance and social accountability in developing countries as well.

The first day of the meeting was devoted to presentations concerning the implementation of the GEO-RECAP project. Dr. G. Giorgobiani, Prof. R. Grigolia and Dr. G. Kochoradze highlighted the input of GTU-MICM, GTU-IC and ICARTI in the process of the project implementation and discussed success stories and new project opportunities.

Dr. Pierre Guisset underlined the role of ERCIM as an instrument to facilitate European cooperation for excellence in ICT research and spoke about its role in the GEO-RECAP project.

The presentations and discussion on the Round Table of the Ideal-ist project concerned the activities of National Contact

Points and the ICT policy in EECA countries. Mutual suggestions and recommendations were presented.

Professors A. Grishin and H. Bergqvist gave very interesting presentations about the Centre of Excellence in Technology at KTH and highlighted the achievements of the institute in the fields of ICT and material sciences.

Dr. R. Kvataladze, head of the Georgian research and educational networking association GRENA, spoke about the development of E-Infrastructure in the schools and scientific organizations of South Caucasus countries and presented joint ICT projects (between the EU and South Caucasus countries).

The second day of the event started with the review of the Draft Strategy Paper for two Georgian institutes, which was moderated by the ERCIM representative.

Presentations by Georgian researchers followed. These focused on the idea of participation in FP7 upcoming calls. Eight proposals were suggested, encompassing various fields of ICT and its applications such as: early diagnosis of prostate cancer, metamaterials and nanoparticles, polarimetry and spectropolarimetry, ICT in elections, process-modelling standards for software development, information resource management system, control and management of large scale networks. Each presentation was followed by questions, discussions and advice.

Georgian researchers benefitted enormously from this event. It gave them the opportunity to become informed about the ICT policy in European and neighbouring countries; to get an insight into the state of the art of the advanced European technological research centres; to present their current research interests and achievements to European and local colleagues; and to establish direct contacts with European scientists. The Event facilitated the achievement of one of the main objectives of the GEO-RECAP project – to foster the integration process of Georgian research centres into the European Research Community in the frame of the FP7 and other programs.

Several informal social activities were also held as part of the event, including an evening reception and sightseeing in Tbilisi and the old town in eastern Georgia. These provided additional valuable opportunities for collaboration.



Pierre Guisset



The GEO-RECAP family

GEO-RECAP is supported by the FP7 INCO ERA WIDE program. ERCIM is a partner of the project, together with DFKI and GIRAF PM Services GmbH (both from Germany), two research institutes from Georgian Technical University – N. Muskhelishvili Institute of Computational Mathematics (GTU-MICM) and V. Chavchanidze Institute of Cybernetics (GTU-IC); and the International Center for Advancement of Research, Technologies and Innovation (ICARTI) also from Georgia.

**Link:** <http://www.georecap.eu>

**Please contact:**

George Giorgobiani, Coordinator of GEO-RECAP,  
GTU-MICM, Tbilisi, Georgia.

Tel: +995 593 129107

E-mail: [bachanabc@yahoo.com](mailto:bachanabc@yahoo.com)

Givi Kochoradze,  
EC ICT NCP, Head of ICARTI  
Tel: +995 599 292516  
E-mail: [gcp@ip.osgf.ge](mailto:gcp@ip.osgf.ge)

## Joint ERCIM eMobility and MobiSense Workshop

by Torsten Braun

*The 6th ERCIM Workshop on eMobility, together with the MobiSense workshop, was held at the Petros M. Nomikos Conference Centre at Fira, Santorini Island. The joint workshop included an invited session, a keynote talk and eight technical talks based on papers, which were selected from eleven submissions. All papers were carefully reviewed and selected in a peer review process by the joint workshop technical program committee.*

The workshop started with an invited session on delay- and disruption-tolerant networks, organized by the Space Internetworking EU project. This was a joint session with the 10th International Conference on Wired/Wireless Internet Communications (WWIC 2012). Various issues including energy-efficiency, security, cloud computing, and data streaming related to delay- and disruption-tolerant networks were discussed.

A keynote talk on Resilient Distributed Consensus was given by Prof. Nitin Vaidya (University of Illinois at Urbana-Champaign, USA). The talk discussed algorithms to achieve consensus, for example, to calculate average sensor values or synchronize clocks in decentralized wireless networks such as sensor networks with rather challenging wireless channel conditions.

Two technical sessions with the eight peer reviewed papers followed the keynote talk. The topic of the first session was routing. It included three presentations about current work in this area. Vasilios Siris (Athens University of Economics and

Business, Greece) presented OptiPath, a system to select routes for vehicles based on available real-time information on travel times. The system can also be used for offloading data from cellular to WiFi networks. Enrica Zola (Universitat Politècnica de Catalunya, Barcelona, Spain) proposed a modification to the DYMO routing protocol in mobile ad-hoc networks. In particular, she proposed to use the beacon-less routing protocol for route discovery. Finally, Geert Heijenk (University of Twente, The Netherlands) proposed a concept for vehicular networks to use tall vehicles for relaying packets for other smaller cars and presented simulation results.

The second technical session was on various aspects of mobility. Torsten Braun (University of Bern) presented concepts for topology control mechanisms in highly mobile ad-hoc networks consisting of a set of small unmanned aerial vehicles. Remco Litjens (TNO Delft, The Netherlands) gave two talks in this session. The first evaluated the novel idea of Cooperative Multi-Point (CoMP) transmission in multi-operator networks; the second examined the feasibility of using the emerging LTE standard for communication in intelligent transportation systems (ITS). One conclusion was that LTE might meet most ITS requirements, except in cases where network capacity limits have been reached. Yun Won



Keynote speaker Nitin Vaidya

Chung (Soongsil University, Seoul, Korea) talked about adaptive energy-efficient location management in multi-tier wireless networks with end systems including both wireless LAN and cellular network interfaces. Finally, Desislava Dimitrova (University of Bern, Switzerland) presented work that is being conducted in collaboration with two Universities in Bulgaria. The paper presented measurement results for video transmissions in heterogeneous wired/wireless networks such as (W)LAN and 3G cellular networks.

The evening prior to the workshop a social event was held at a beach on Santorini Island. Local organizers, under the guidance of Vassilis Tsoussidis (Demokritus University of Thrace, Xanthi, Greece), did a fantastic job of organizing this event. The PDF workshop proceedings (ISBN: 978-3-9522719-3-3) can be downloaded from <http://wiki.ercim.eu/wg/eMobility/images/d/d1/ERCIM-WS2012-Proceedings-6.pdf>.

**Link:** <http://wiki.ercim.eu/wg/eMobility/>

**Please contact:**

Torsten Braun, E-mobility Working Group chair  
University of Bern, Switzerland  
E-mail: [braun@iam.unibe.ch](mailto:braun@iam.unibe.ch)

# MUSCLE Working Group Co-organised Mass Data Analysis Conference

by Petra Perner and Emanuele Salerno

*The seventh international conference on mass data analysis of images and signals, MDA 2012, was held in Berlin from 13 to 20 July 2012. ERCIM was one of the sponsors and the ERCIM MUSCLE Working Group (Multimedia Understanding through Semantics, Computation, and Learning), participated in the organization.*

Invited talks were given by Xiaoqing Ding of Tsinghua University in Beijing (Face recognition), Patrick Wang of Northeastern University in Boston (Intelligent pattern recognition), and by Petra Perner, director of IBAI in Leipzig (Quantitative measurement of cellular events). IBAI, the Institute of Computer Vision and Applied Computer Sciences, has been a member of MUSCLE since 2004. Jacques Lévy Véhel and Michel Tesmer won the MDA 2012 Best Paper award, for “A New Method for Multifractal Spectrum Estimation with Applications to Texture Description”. Proceedings published by Ibai-Publishing, Fockendorf, Germany, ISBN 978-3-942952-15-6.

The enormous volume of data that characterizes our era is not useful in itself but does contain useful information. Large multimedia databases containing still pictures, videos, audio, and text documents are typical examples of highly redundant data sets, wherein very rich and informative patterns can be identified from both elementary pieces of information and mutual relationships between them. Pieces of information and related patterns, however, are often hidden in huge sub-



Prof. Patrick Wang (left) and Prof. Xiaoqing Ding

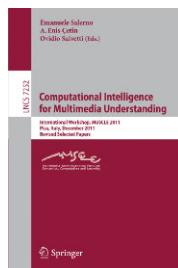
sets of data and increasingly sophisticated techniques are needed to bring them to light, consequently, data mining techniques represent one of the most important toolkits at our disposal. As the size of our data records increases data mining techniques require continual refinements and improvements.

This was the theme of MDA 2012. The subjects ranged from discovering regularities in human faces or brain activity records to finding suitable metrics in complex multivariate processes, such as the behaviour of living cells, the output of a gene expression experiment, a speech signal, or the motion of multiple objects in a natural scene.

Finding useful information patterns in large databases means reducing redundancy, but redundancy must also be exploited to increase the robustness of the analysis and the reliability of the results. The analysts are thus faced with small-size, robust and reliable processed data. They also need to be sure that no useful information is lost and that the results are easily accessible and understandable. The latter requirement entails that imaging and visualization are essential complements of data mining techniques. All the fundamental problems in data mining have found some (not always general) solution. Today's mass data, however, challenge us to find ever more innovative approaches and practical procedures. High-level approaches should take precedence over methods based on low-level features and, whilst statistics and computation are still essential, including semantics is becoming inescapable. Working on classes and categories through means such as ontologies, case-based reasoning, and other knowledge-based strategies can be the key to face the new challenges. This affects equally the development of theory, the design of new methods and the practical applications. And this is why MDA is devoted to both the research and the industry audiences.

## 2011 MUSCLE Working Group workshop proceedings published

Emanuele Salerno, A. Enis Cetin, Ovidio Salvetti (Eds.)



## Computational Intelligence for Multimedia Understanding

This recently published book constitutes the refereed proceedings of the International Workshop MUSCLE 2011 on Computational Intelligence for Multimedia Understanding, organized by the ERCIM Working Group in Pisa, Italy in December 2011.

The 18 revised full papers were carefully reviewed and selected from numerous submissions. The papers cover the following topics: multisensor systems, multimodal analysis, crossmodal data analysis and clustering, mixed-reality applications, activity and object detection and recognition, text and speech recognition, multimedia labeling, semantic annotation, metadata, multimodal indexing and searching in very large data-bases, case studies. All papers are freely available online.

Springer 2012, <http://www.springerlink.com/content/978-3-642-32435-2>

### Links:

- <http://www.mda-signals.de>
- <http://wiki.ercim.eu/wg/MUSCLE/>
- <http://www.ibai-institut.de>

### Please contact:

Petra Perner, IBAI Institute, Germany  
E-mail: [perner@ibai-institut.de](mailto:perner@ibai-institut.de)

Emanuele Salerno, ISTI-CNR, Italy  
E-mail: [emanuele.salerno@cnr.it](mailto:emanuele.salerno@cnr.it)

# Annual Meeting of the ERCIM Working Group on Models and Logics for Quantitative Analysis

by Diego Latella

*The fourth annual meeting of the WG on Models and Logics for Quantitative Analysis (MLQA) took place on 8 September 2012 at the Laboratory for Foundations of Computer Science of the University of Edinburgh. The meeting was organized in cooperation with The Scottish Informatics & Computer Science Alliance (sicsa\*).*

The event consisted of a number of invited talks, a number of contributed talks, a poster session and a business meeting. The meeting, organized by Jane Hillston, Gethin Norman and Flemming Nielson, brought together experts in the area of process algebra, fluid flow analysis, compositional verification and analysis. Its theme was Compositional Modelling and Analysis of Quantitative Systems. This year, the programme consisted of both invited and contributed talks.

Invited talks:

- Pedro D'Argenio: Security analysis in probabilistic distributed protocols via bounded reachability
- Andrea Marin: Compositional model specifications and analyses via product-forms
- Mirco Tribastone: Exact Aggregation for Fluid Process Algebra Models
- Jaco van de Pol: Symbolic Manipulation of Markov Automata

Contributed talks

- Vashti Galpin: Stochastic hybrid modelling with composition of flows
- Saray Shai: Coupled adaptive complex networks
- Chris Banks: A logic for behaviour in context

The next MLQA meeting will take place in Rome, on Sunday 24 March 2013 as satellite event of the European Joint Conferences on Theory and Practice of Software (ETAPS). The meeting will be organized by Flemming Nielson and the theme will be Cyber Physical Systems. A session of MLQA 2013 will be shared with the 11th Workshop on Quatitative Aspects of Programming Languages (QAPL 2013). The Working Group will also co-organize the QAPL-MLQA summer school which will take place in Bertinoro (IT) 17-22 June 2013. The school will be organized by Alessandra Di Pierro, Erik de Vink and Herbert Wiklicky. The theme of the school will be "Dynamical Systems".

## Links:

Further details can be found at:

[http://wiki.ercim.eu/wg/MLQA/index.php/MLQA\\_2012](http://wiki.ercim.eu/wg/MLQA/index.php/MLQA_2012)  
<http://workshops.inf.ed.ac.uk/mlqa2012/>

## Please contact:

Flemming Nielsen, Technical University of Denmark  
E-mail: nielsen@imm.dtu.dk

# 17th International Workshop on Formal Methods for Industrial Critical Systems

by Radu Mateescu

*The 17th International Workshop on Formal Methods for Industrial Critical Systems (FMICS'2012) was held on 27-28 August 2012 in Paris, as a satellite event of the FM'2012 conference. FMICS seeks to foster the dissemination of formal methods in industry by providing a forum where scientists and engineers can exchange their experiences in the development and industrial usage of these methods.*

FMICS'2012 took place at the CNAM (Conservatoire National des Arts et Métiers), a French institution of long-standing and deep scientific tradition, founded in 1794 during the French Revolution.



*From left to right: Marielle Stoelinga, Yann Régis-Gianas, and Ralf Pinger*

The workshop was chaired by Marielle Stoelinga (University of Twente, The Netherlands) and Ralf Pinger (Siemens AG, Germany). It attracted over 30 participants from nine countries, both from academia and industry. Thirty-seven papers were submitted, of which 14 were accepted for presentation at the workshop (a 38% acceptance rate). The keynote lectures were given by Dimitra Giannakopoulou (NASA AMES, USA) and Hubert Garavel (Inria Grenoble - Rhône-Alpes, France). The proceedings of FMICS'2012 were published by Springer Verlag as volume 7437 in the LNCS series.

The best paper award for FMICS'2012 was granted by EASST (European Association of Software Science and Technology) to Nicolas Ayache, Roberto Amadio, and Yann Régis-Gianas for their paper entitled "Certifying and reasoning on cost annotations in C programs".

In addition to their involvement in FMICS'2012, several members of FMICS also contributed to the scientific program of FM'2012 as authors of papers, tutorials, and lectures at the Industry Day.

**Link:** <http://fmics.inria.fr>

## Please contact:

Radu Mateescu, Inria Grenoble - Rhône-Alpes and LIG  
E-mail: Radu.Mateescu@inria.fr



Introduction to the Special Theme

## What Makes Alan Turing a Great Scientist?

by Gilles Dowek and Samson Abramsky

*This Turing centenary marks a point at which we can realize that Alan Turing has become, with the passage of time, a scientific icon whose name is known by people in many countries world-wide, and far beyond the scientific community. This may seem a paradox because the genesis of computability theory, for which Turing is probably best known, was a collective effort, to which the names of Herbrand, Gödel, Church, Post, Kleene, Rosser and Turing are often associated.*

There are of course many non-scientific reasons for Alan Turing to be an icon: his short life, martyrdom, significance as a gay political symbol, ... but Turing's fame came firstly within the scientific community, as a founding father of computer Science. Thus, for example, it was decided by the ACM, in 1966, to name the highest distinction in computer science, the Turing Award, after him. So we must search for the origin of Turing's fame in his scientific achievements.

Alan Turing started his scientific work with one of the most abstract problems in mathematics: the decision problem. And he solved it by introducing an imaginary, mathematical computing machine. This work already shows the originality of Turing's work: Church independently solved the decision problem at the same time, but Church's approach focused on the concepts of a language and of an algorithm, with their roots in logic and mathematics. Turing, by contrast, introduced a decisive third concept, of a "machine", thus going beyond logic, and laying the foundations for the nascent discipline of computer science. Moreover, his analysis of computability in



terms of his notion of machine was so compelling that it was rapidly accepted as definitive; and in the form of the Church-Turing thesis, is still with us today.

Turing continued this work on machines after the war, when he designed a real machine: the Automatic Computing Engine (ACE) for the National Physical Laboratory. While many mathematicians, at that time, saw a water-tight boundary between mathematics and technology, Turing introduced a notion of a machine within mathematics, and moved freely back and forth between computability and computational machinery. While many logicians came close to inventing computer science, only Turing did it.

During the war, many of Turing's contemporaries, under the pressure of history, turned to action. Turing understood that the outcome of the war depended as much on the development of cryptanalysis as on what happened on the battlefield. This led him to join the Government Code and Cypher School in Bletchley Park, where ciphers and codes of several Axis countries, in particular the Enigma and the Lorenz machines, were decrypted. Again, Turing seemed not to pay attention to the boundary that we all customarily see between thought and action.

When Turing became interested in biology and morphogenesis, he modelled the development of living organisms with differential equations, as reaction-diffusion systems, just as one would model an inorganic object. And, when he got interested in intelligence, which many consider to be the sole prerogative of mankind, he got rid of the border between the human and the non-human, to ask under which conditions a computing system could be said to be intelligent, giving a purely behavioural definition of intelligence.

The scientific legacy of Turing is huge: on models of computations (see the paper of Jean-Yves Marion, and that of Jos Baeten, Bas Luttik, and Paul van Tilburg in this issue), on Cryptographic systems (see the paper of Benjamin

Grégoire), on morphogenesis (see the paper of Nadia Pisanti), at the border of biology and computation, on systems biology (see the paper of Anna Gambin, Anna Marciniak-Czochra, and Damian Niwinski and that of Claudio Angione, Pietro Liò, and Giuseppe Nicosia), ... But Turing not only left us with this huge scientific legacy, he also showed that borders must be crossed, as an essential part of the scientific endeavour.

Science and technology, thought and action, organic and inorganic, human and non human: we tend to believe that these boundaries are fixed and immutable, but they are really only there to divide our knowledge into small and comfortably familiar pieces. Alan Turing did not pay much attention to these boundaries and, following where his powerful curiosity led, he moved from one area to another, with an amazing intellectual ease and technical facility. He applied methods from one domain to another, candidly asking why these methods should not work, if they had worked somewhere else. He has set an invaluable and inspiring example: go where your ideas lead you, and pay no attention to artificial borders.

**Please contact:**

Gilles Dowek, Inria, France  
E-mail: [gilles.dowek@inria.fr](mailto:gilles.dowek@inria.fr)

Samson Abramsky, University of Oxford  
E-mail: [samson.abramsky@cs.ox.ac.uk](mailto:samson.abramsky@cs.ox.ac.uk)

# Viruses in Turing's Garden

by Jean-Yves Marion

**Cohen and his supervisor Adleman defined a virus as follows: "A virus is a program that is able to infect other programs by modifying them to include a possibly evolved copy to itself". This definition seems to be well accepted by the computer security community as a foundational definition. Thus, a virus is a self-replicating program, whose offspring may be a mutation of the original program. Viruses thrive in our computers, which are based on Turing's model of computation. We discuss the fundamental reasons for this.**

One of Turing's main achievements is the construction of a Universal Turing Machine (UTM), which corresponds to a self-interpreter. That is a UTM executes a program by reading and interpreting a data. Thus, a program is a data, and this feature is one of the keys to build self-replicating programs. Programs called quines, which generate an exact copy of their source code, provide an amazing illustration. To devise a self-replicating program, we can follow Thompson's Turing award lecture or use Kleene's second recursion theorem. Indeed, from a programming view, a virus may be defined by a specification such as the following:

```
Viral_Spec (virus vs)
  Send confidential information
    For each pdf file f, append f to vs
```

A virus, satisfying the above specification, will first steal some information and then infect all pdf files in the system by appending a copy of itself at the end of each file. Thus, a solution of this specification is a program  $v$  such that  $v$  is a fixpoint of  $\text{Viral\_Spec}$ , i.e.,  $v = \text{Viral\_Spec}(v)$ . We see that a solution  $v$  behaves like a virus with respect to Cohen's definition. Anyone familiar with Smullyan's works will have no difficulty seeing that such a specification is self-referential. Indeed  $vs$  refers to the virus defined by the specification. There are many methods now known of achieving self-reference specification, but Kleene's demonstration is constructive and so gives an explicit solution, that is a virus. Moreover, his demonstration is uniform with respect to a self-reference specification, and therefore, it provides a virus compiler, which outputs a virus from a specification of it.

Viruses, that we have developed above, spread between programs without changing their program: There is no mutation. However, virus program mutation is an important feature used to avoid anti-virus detection. Again,

Kleene's second recursion theorem gives us an effective construction of viruses whose descendants are all different. Indeed, a specification (and more generally a computable function) has an infinite number of fixpoints. Although the set of fixpoints is not computable, we may construct an infinite subset of fixpoints. Each fixpoint of this subset is a solution of the same self-reference specification, and each fixpoint corresponds to a virus. Then we can build a virus, which mutates when it duplicates, so no two copies are the same. Consequently, viruses cannot be detected.

In addition to self-replication and mutation, self-modification - the ability of a program to modify itself - is also an important feature. A self-modifying program is able to alter its own code. For example, it may change its instructions on the fly, or generate codes from data that it can access, and thus the overall program may be seen as a sequence of different code layers. As a result, we can say that "a data is a program". From a theoretical point of view, a model of computation of self-reproducing systems must reflect the crucial fact that a data can be activated and executed, such as in random access stored program machines of Hartmanis.

In summary, Cohen's or Adleman's model of viruses bears a resemblance to self-reproducing machines. Universal Turing machines and the second recursion theorem of Kleene form the core of the study of self-reproducing machines. There are other directions worth exploring: We may think that from one generation to another, viruses evolve and do not compute the same thing. In practice, this happens when a virus is updated either to apply a patch or to embed new functionalities. Another direction might be inspired by works on synthetic models of living organisms, beginning with Turing's paper on mor-

phogenesis and von Neumann and Burks' construction of a self-replicating cellular automaton.

## References/Links:

M. A. Arbib: "From universal turing machines to self-reproduction. In A half-century survey on The Universal Turing Machine", pages 177–189, Oxford University Press, Inc., 1988. <http://dl.acm.org/citation.cfm?id=57249.5725>

F. Cohen: "Computer viruses: Theory and experiments", Computers & Security, 6 (1):22–35, 1987, ISSN 0167-4048. [http://dx.doi.org/10.1016/0167-4048\(87\)90122-2](http://dx.doi.org/10.1016/0167-4048(87)90122-2)

S. C. Kleene: "On notation for ordinal numbers", The Journal of Symbolic Logic, 3 (4):150–155, 1938, <http://dx.doi.org/10.2307%2F2267778>.

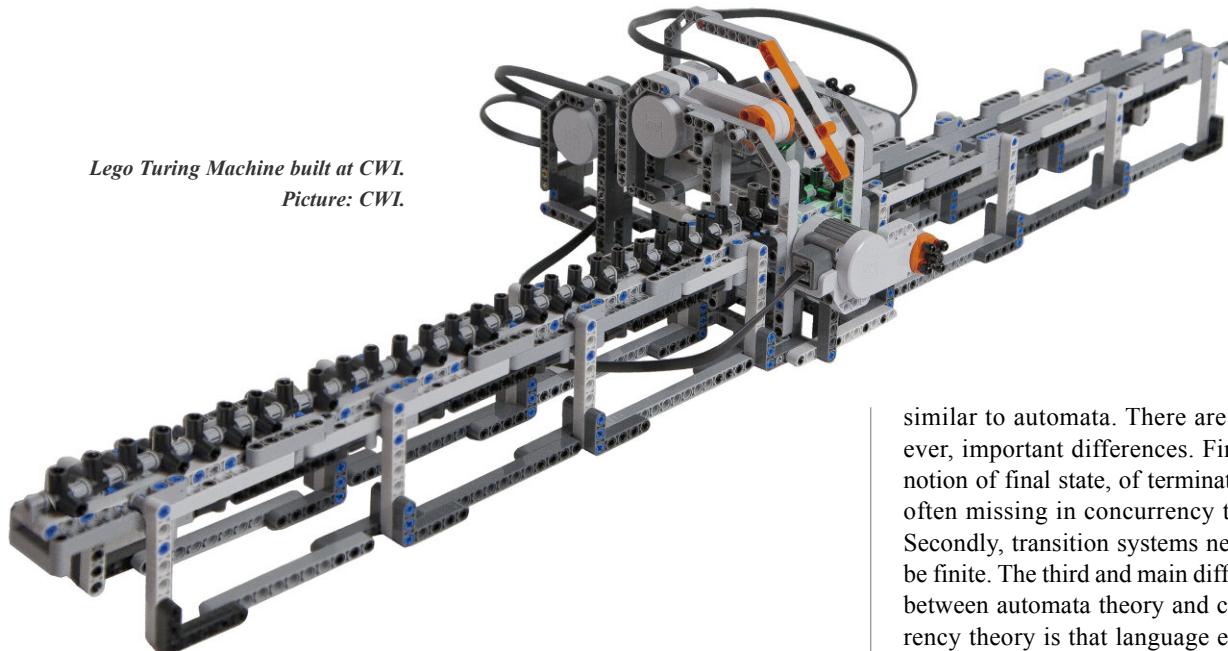
J-Y. Marion: "From Turing machines to computer viruses", Phil. Trans. R. Soc. A. July 28, 2012, <http://dx.doi.org/10.1098/rsta.2011.0332>

M. Turing: "On computable numbers, with an application to the Entscheidungsproblem", in Proc. of the London Mathematical Society, second series Turing [1937], pages 230–265

A. M. Turing: "The chemical basis of morphogenesis", philosophical transactions of the Royal Society of London Series B, Biological sciences, B 237(641):37–72, 1952. <http://turing.ecs.soton.ac.uk/browse.php/B/22>

## Please contact:

Jean-Yves Marion  
Université de Lorraine, LORIA,  
France  
E-mail: [Jean-Yves.Marion@loria.fr](mailto:Jean-Yves.Marion@loria.fr)



Lego Turing Machine built at CWI.  
Picture: CWI.

## When Turing Meets Milner

by Jos Baeten, Bas Luttik and Paul van Tilburg

**At CWI and Eindhoven University of Technology in the Netherlands, we enhanced the notion of a computation in the classical theory of computing with the notion of interaction from concurrency theory. In this way, we adapted a Turing machine as a model of computation to a Reactive Turing Machine that is an abstract model of a computer as it is used nowadays, always interacting with the user and the world.**

What is a computation? This is a central question in the theory of computing, dating back to the work of Alan Turing in 1936. The classical answer is that a computation is given by a Turing machine, with the input given on its tape at the beginning, after which a deterministic sequence of steps takes place, leaving the output on the tape at the end. A computable function is a function of which the transformation of input to output can be computed by a Turing machine.

A Turing machine can serve thus as a basic model of a computation and, until the advent of the terminal in the 1970s, could also serve as a basic model of a computer. The terminal made direct interaction with the computer possible. Nowadays, a computer interacts continuously with the user at the click of a mouse or with many other computers all over the world through the Internet.

An execution of a computer is thus not just a series of steps of a computation, but also involves interaction. It cannot be modelled as a function, and is inherently nondeterministic. In our research, we made the notion of an execution pre-

cise, and compared this to the notion of a computation. To illustrate the difference between a computation and an execution, we can say that a Turing machine cannot fly an airplane, but a computer can. An automatic pilot cannot know all weather conditions en route beforehand, but can react to changing conditions run-time.

Computability theory, which is firmly grounded in automata theory and formal language theory, is the study of languages and the sets of strings, induced by them. Language can be viewed as an equivalence class of automata (under language equivalence).

The notion of interaction has been studied extensively in concurrency theory and process theory exemplified by the work of Robin Milner, who played a central role in the development of concurrency theory. He proposed a powerful parallel composition operator that is used to compose systems in parallel, including their interaction.

The semantics of concurrency theory are largely given in terms of transition systems which, in some respects, are

similar to automata. There are, however, important differences. Firstly, a notion of final state, of termination, is often missing in concurrency theory. Secondly, transition systems need not be finite. The third and main difference between automata theory and concurrency theory is that language equivalence in concurrency theory is too coarse to capture a notion of interaction. Therefore, other notions of equivalence are studied in concurrency theory, capturing more of the branching structure of an automaton. Prominent among these is bisimulation equivalence.

We studied the notion of a computation, taking interaction into account. We defined, next to the notion of a computable function, the notion of an executable process, a behaviour that can be exhibited by a computer (interacting with its environment). An executable process is a divergence-preserving branching bisimulation equivalence class of transition system defined by a Reactive Turing Machine, an adaptation of the classical Turing Machine that can properly deal with ubiquitous interaction.

There have been other attempts to add a notion of interaction to computability theory but none have taken full advantage of the results of concurrency theory. Studies tend not to give interaction the status it deserves; in all the formalizations of interaction machines we could find, the notion of interaction itself is still very implicit.

In our research, we used process theory to consider finite-state processes and pushdown processes. We also defined executable processes, highlighted the relationship of computable functions and executable processes, laying the foundations of executability theory alongside computability theory, and defined a new grammar for executable processes,

based on the universality of the queue process.

In conclusion, we discussed the notion of an execution that enhances a computation by taking interaction into account. This was achieved by marrying computability theory, moving up from finite automata through pushdown automata to Turing machines, with concurrency theory, not using language equivalence but divergence-preserving branching bisimilarity on automata. Although every undergraduate curriculum in computer science contains a course on automata theory, an introduction to concurrency theory is rarely included. Both theories as basic models of computation are part of the foundations of computer science, and can be taught in an integrated manner, with the results of our research [2].

Finally, it is unlikely that Turing and Milner ever did meet: by the time Milner entered King's College in Cambridge as a young student, Turing had left some years earlier to take up a job in Manchester.

*With the help of Annette Kik, CWI*

#### References:

1. J.C.M. Baeten, T. Basten, and M.A. Reniers: "Process Algebra (Equational Theories of Communicating Processes)", number 50 in Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, 2009
2. J.C.M. Baeten: "Models of Computation: Automata, Formal Languages and Communicating

Processes", Technische Universiteit Eindhoven, 2011, Syllabus 2IT15

3. J.C.M. Baeten, B. Luttik, and P. van Tilburg: "Turing meets Milner", in proc. of CONCUR 2012, Springer LNCS volume 7454, pp. 1-20. 2012

#### Please contact:

Jos Baeten, CWI, The Netherlands  
E-mail Jos.Baeten@cwi.nl

## How to Compute with Metabolism in Bacteria?

by Claudio Angione, Pietro Liò and Giuseppe Nicosia

*An enzyme can be thought of as a computational element, i.e. a processing unit able to transform an input into an output signal. Thus, in a biochemical pathway, an enzyme reads the amount of reactants (substrates) and converts them into products. Here we consider the biochemical pathway in unicellular organisms (e.g. bacteria) as a living computer that can be programmed to obtain the desired output. Through an optimal executable code stored in the "memory" of bacteria, we can simultaneously maximize the concentration of two or more metabolites of interest.*

The key role of computation in the bio-inspired science was first discovered by Alan Turing in 1952. In 1995, Bray pointed out that a single protein can transform one or multiple input signals into an output signal, and can thus be viewed as a computational or information-carrying element. Following this line of thought, we provide a framework to show that bacteria could have computational capability and act as molecular machines. Accordingly, the same framework can be applied to eukaryotic cells.

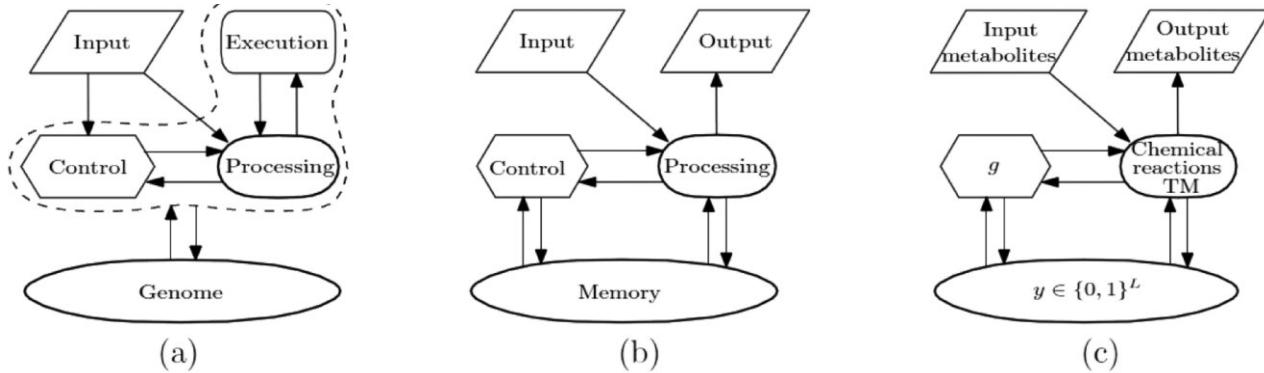
Inspired by Brent and Bruck [1], through the formalism shown in Figure 1 we describe the whole behavior of bacterial cells in terms of the von Neumann architecture. In particular, the genome sequence is thought of as an executable code specified by a set of commands in a sort of ad-hoc low-level programming language. Each combination of genes is coded as a string  $y$  of  $L$  bits, each of which represents the status of a gene set. By turning off a gene set,

we turn off the chemical reactions associated with it.

The memory unit contains the string  $y$ , which is the program to be executed. We model the processing unit of the bacterium as the collection of all its chemical reactions. In this way, we associate the chemical reaction network of bacteria with a Turing machine (TM). This relationship is based on the mapping between the cell metabolism and a Minsky's register machine RM (equivalent to a TM) [3]. In fact, an RM is a multitape TM with the tapes restricted to act like simple registers (i.e. "counters"). A register is represented by a left-handed tape that can hold only positive integers by writing stacks of marks on the tape; a blank tape represents the count '0'. Specifically, the reactions taking place in the cell can be thought of as increment/decrement instructions of the RM, where the RM registers (tapes) count the number of molecules of each metabolite.

Remarkably, as the biological system grows larger, reaching the desired multiple input/output, performance becomes a difficult task, necessitating some sort of machine optimization. To this end, we provide a novel algorithm called Genetic Design through Multi-objective Optimization (GDMO), with the aim of programming bacteria to maximize the yield of desired metabolites. The multi-objective optimization aims at exploiting the computational capabilities of bacteria in order to allow the maximum production of metabolites of practical or industrial interest. The solution of a multi-objective problem is a potentially infinite set of points called Pareto optimal solutions or Pareto front.

GDMO finds the genetic strategies that obey control signals, and optimizes multiple biological functions. Each point of the Pareto front provided by GDMO is a molecular machine that executes a particular task. The Pareto optimality enables not only a wide



**Figure 1:** Comparison inspired by [1] among biological systems (a), von Neumann architecture (b), and bacteria (c). The string  $y$  is a program stored in the RAM. The function  $g$  represents the control unit: it interprets the binary string  $y$  and turns gene sets off. The processing unit is the metabolism of bacteria, composed of all the chemical reactions taking place within it. The goal is to produce desired metabolites as output of the molecular machine.

range of Pareto optimal solutions, but also the best trade-off design (see Figure 2).

Robust genetic interventions in cells, framed as optimal programs to be run in a molecular machine, can be exploited to extend and modify the behavior of cells and cell aggregates. For instance, programs can instruct cells to make logic decisions according to environmental factors and current cell state. A program embedded in a cell could allow its metabolic network to work with a specific user-imposed aim. The complexity and performance of this type of computing could be described using the Pareto front; specifically, each axis represents a metabolite, while the shape of the area underlying the front indicates the ability of the organism to specialise. The metabolic computing (see links [a] and [b]) proposed here complements Cardelli's DNA computing (link [c]) and Bray's enzyme computing.

#### Links:

[a] <http://www.easychair.org/publications/?page=2080395222>

[b] <http://research.microsoft.com/en-us/events/2012summerschool/claudioangione.pdf>

[c] <http://www.springerlink.com/content/g16345330209/?MUD=MP>

#### References:

[1] R. Brent and J. Bruck: “2020 computing: Can computers help to explain biology?”, *Nature*, 440(7083):416–417, 2006

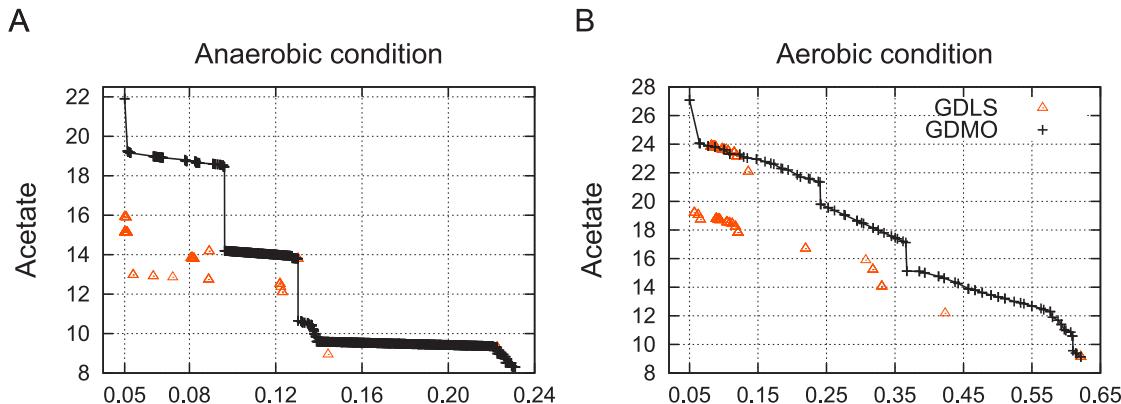
[2] D. Lun et al: “Large-scale identification of genetic design strategies using local search”, *molecular systems biology*, 2009

[3] D. Soloveichik et al.: “Computation with finite stochastic chemical reaction networks”, *natural computing*, 2008

#### Please contact:

Claudio Angione, Pietro Liò  
Computer Laboratory, University of Cambridge, UK  
E-mail: claudio.angione@cl.cam.ac.uk, pietro.liò@cl.cam.ac.uk

Giuseppe Nicosia  
Department of Maths and Computer Science, University of Catania, Italy  
E-mail: nicosia@dmi.unict.it



**Figure 2:** Maximization of biomass formation (x axis) and acetate production (y axis) in anaerobic (without oxygen, A) and aerobic (with oxygen, B) conditions with glucose uptake rate of 10 mmolh-1 gDW-1 in the *Escherichia coli* model iAF1260. The Pareto fronts obtained by GDMO are in black. We show how GDMO overcomes GDLS [2], whose results are in red.

# Recent Advances in the Formal Verification of Cryptographic Systems: Turing's Legacy

by Benjamin Grégoire

***Combining program generation and formal proof verification, we are now able to discover new cryptographic systems and prove their resistance to many classes of attacks.***

Alan Turing's legacy in computer science spans many domains including: computability theory, cryptology, and foundations of software engineering. These three domains contribute to recent results with exciting implications for the study of cryptographic systems. In work started six years ago, Gilles Barthe, César Kunz (initially at INRIA, now at IMDEA, Madrid), Benjamin Grégoire, Sylvain Heraud (INRIA), Yassine Lakhnech (University of Grenoble), and Santiago Zanella-Béguelin (initially at INRIA, now at Microsoft Research) have developed several methods to study cryptographic systems with increasing levels of automation.

This joint research effort has resulted in new tools for the design and study of cryptographic systems. The traditional approach, which relies on pen-and-paper exploration and human reviewing to design and build trust in cryptographic systems, is not only seriously flawed, as shown by the many errors routinely discovered in published proofs, but also hopelessly ineffective given the myriad of design possibilities for cryptographic

systems. This research aims at speeding up the process of both exploring new designs and verifying their robustness.

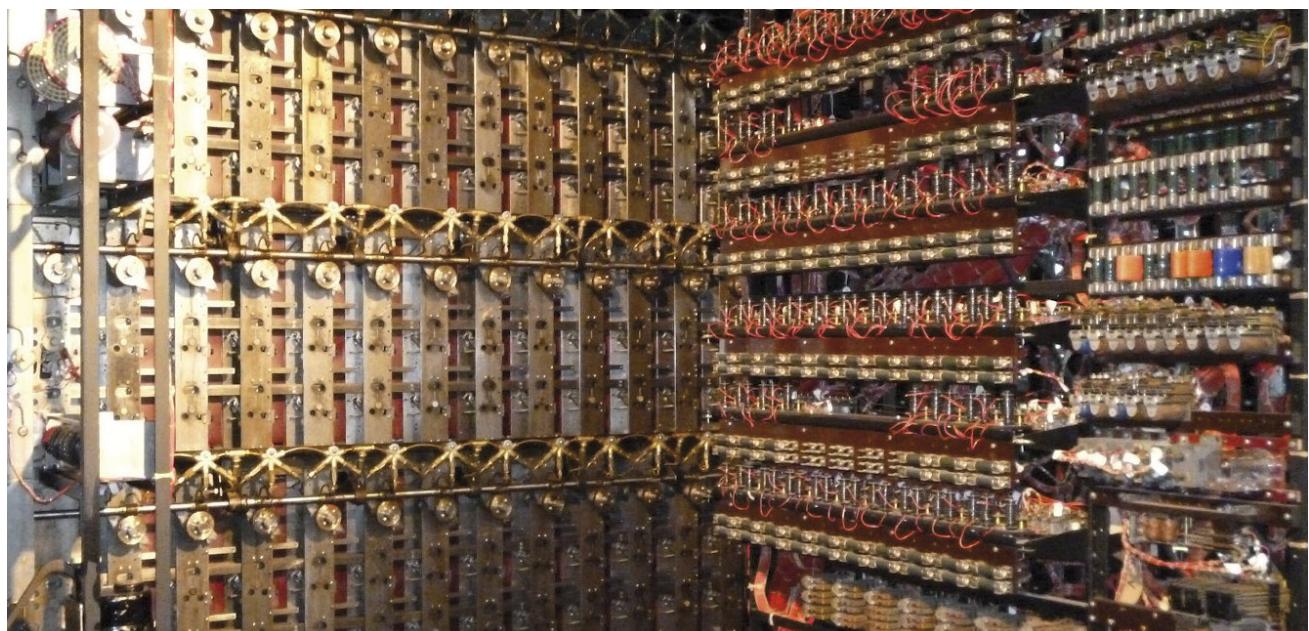
The research started by developing libraries for the Coq interactive theorem prover that capture precisely the mathematical notions that serve as foundations to prove that, for instance, an encryption algorithm preserves the confidentiality of messages. These notions build on programming language theory, because a language is needed to describe cryptographic systems and the possible attacks against them. This also requires some computability theory, because feasible attacks can only make use of "reasonable" amounts of resources (time and memory). All this must also be combined with probability theory, because cryptographic systems are probabilistic in nature and they must hold against attacks that succeed despite a small probability of success - as the old adage goes: "attacks always get better, they never get worse".

The library that they obtained, called CertiCrypt, made it possible to verify

formally the proofs of known cryptographic systems, interacting with the Coq theorem prover to develop the proofs. Significant results obtained in this manner include proofs of the security of the widely deployed OAEP encryption scheme and the Full-Domain Hash signature scheme [1,2].

The experience and firm foundations gathered from CertiCrypt, led to a new tool, EasyCrypt, which greatly automates the task of building proofs by relying on automated, rather than interactive, theorem provers. As an added benefit, this also requires less expertise and effort from users. This technology relies on two languages: one to describe cryptographic systems and attacks, and one to describe proofs of security. Proofs are by reduction and structured as sequences of games played between a challenger and an attacker. The initial game encodes the security goal, while the final game encodes a problem assumed to be computationally hard. The final statement gives an upper bound on the winning probability of the attacker in the initial game as a function

*Details of a cryptanalytical machine designed by Alan Turing to help decipher German Enigma-machine-encrypted signals during World War II.*



of the probability of solving the problem encoded in the last game. An impressive outcome of this approach is a machine-checked proof of the Cramer-Shoup encryption scheme, which previously seemed inaccessible to formal analysis. This work received the best paper award at the CRYPTO 2011 conference [3].

The latest advances have evolved from studying individual cryptographic systems to studying whole families of systems. Encryption schemes, for instance, are generally built by combining a few basic blocks in a few possible ways. All possible combinations can be represented as a formal language, which makes it possible to systematically explore each of them and analyze their

security. Most schemes turn out to be weak (they are vulnerable to known attacks), but others are secure; for many of these, an independently verifiable proof in EasyCrypt can be generated automatically. This new development, which is pending publication, gives a systematic way to discover new cryptographic systems with nice properties and machine-checked proofs.

*With the help of Yves Bertot, Inria*

**Link:**

<http://easycrypt.gforge.inria.fr/>

**References:**

- [1] G Barthe et al.: “Formal Certification of Code-Based Cryptographic Proofs”, in proceedings of POPL 2009, ACM

2009, pp.90-101, <http://dx.doi.org/10.1145/1480881.1480894>

[2] G. Barthe et al.: “Beyond Provable Security. Verifiable IND-CCA Security of OAEP”, in Proc. “Topics in Cryptology - CT-RSA 2011”, Springer LNCS vol. 6558, pp. 180-196, 2011

[3] G. Barthe et al.: “Computer-Aided Security Proofs for the Working Cryptographer”, in Proc. “Advances in Cryptology - CRYPTO 2011”, Springer LNCS vol. 6841, pp. 71-90 (Best Paper Award)

**Please contact:**

Benjamin Grégoire

Inria, France

E-mail: [benjamin.gregoire@inria.fr](mailto:benjamin.gregoire@inria.fr)

## From Discrete to Continuous: Turing's Morphogenesis

by Nadia Pisanti

**What do a leopard skin and a sunflower head have in common? Very little, should they not both exhibit patterns whose generation is the result of the interaction of components, the morphogens, according to mathematical laws described sixty years ago by Alan Turing.**

Patterns have always been observed in nature, and Turing certainly was not the first to detect mathematical models and geometrical schemes in, for example, plants' phyllotaxis. In the days of Turing, half a millennium had passed since the geometric studies of Leonardo da Vinci on plants [1], and D'Arcy Thompson Wentworth seemed to have almost exhausted, with his book On Growth and Form [2] in 1917, whatever could be mathematically told about morphogenesis, that is, the biological process of the creation of shapes in living organisms.

When, in 1952, Alan Turing wrote The Chemical Basis of Morphogenesis [3], the morphogenesis was that of D'Arcy Thompson Wentworth, whose contribution was mainly that of highlighting the importance of mechanisms and physical laws in determining certain shapes in living beings. In his book [1], D'Arcy Thompson lists an enormous number of correlations among observed shapes in living organisms, mechanical phenomena, and their corresponding mathematical models. The contribution of

Alan Turing to morphogenesis is less extensive than that of Thompson, but no less important, as it exhibits several innovative ideas. First of all, as the title of his paper (The Chemical Basis of Morphogenesis) suggests, it moves the attention from the mechanics of morphogenesis to the chemistry of its components. Such components, the morphogens, are the actual agents of a system in which they co-operate for the mechanisms of morphogenesis.

Morphogens are substances that diffuse, and thus propagate, depending on their concentration, signals that control cell differentiation. Turing gives precise mathematical models for the fluctuations of these chemical quantities and applies the corresponding laws. The model he suggests was a novelty for the theory of morphogenesis (Turing introduces it as “mathematically convenient, though biologically unusual”): a continuous system of differential equations representing non-linear dynamics. With it, Alan Turing captures his brilliant intuition that patterns observable in nature, with their repetitiveness and

variability, are the result of the interaction between chemical substances whose concentrations have iterative reciprocal effects. Such effects are, at each iteration, qualitatively similar but quantitatively slightly different, thus driving the pattern creation.

Turing describes his model as a reaction-diffusion system: starting from a conformation determined by the quantities of the chemical components in an initial state, it evolves by changing these quantities according to mathematical laws. Within these fluctuations, the system can reach various kinds of stable equilibria, and stability is due to a substantial symmetry. Breaking this symmetry leads to a (continuous, not discrete) transition from one stable state to another. Turing observes that the variety of irregularities that can break the symmetry is large and hard to investigate, but that fortunately the variety of equilibrium states that can be reached is limited to six. In his paper Turing analyses them all, outlining hypotheses on the kind of morphogenesis they can possibly give rise to. In particular, Turing's

non-trivial observation is that in a system with several components, the diffusion of a morphogen in space can itself be a cause of instability of the global system, whereas, in the local system of each single component there would otherwise have been stability.

In the conclusion of [3], Turing modestly talks about the “relatively elementary mathematics used in this paper”. This is unfair: among the remarkable aspects of his model, there is the pioneering notion of exponential drift that he uses for the onset of instability. Turing observes that “a drift away from the equilibrium occurs with almost any small displacement from the equilibrium condition”: a very similar concept to what later (after Ruelle's work in the seventies [4]) will be named sensitivity to initial conditions.

Earlier, in [5], Alan Turing had explained discrete-state machines as “machines which move by sudden jumps or clicks from one quite definite state to another”. By adding a few lines

later that “strictly speaking there are no such machines. Everything really moves continuously”, he somehow anticipated the motivations of his work on morphogenesis. The inventor of the discrete-state machines had caught, besides their potential, their limits for modelling physical processes. With the reaction-diffusion system he describes in “The Chemical Basis of Morphogenesis”, Alan Turing somehow fills this gap.

This extraordinary scientist, whose polyhedral abilities are reminiscent of those of Leonardo da Vinci, has made fundamental contributions in very distinct scientific disciplines, with discoveries based on very different concepts, such as discrete and continuous. Maybe the key of his genius lies precisely in the depth of the observations that can emerge by investigating and comparing such opposite concepts.

#### References/Links:

- [1] Leonardo da Vinci, manuscript: “Studi di geometria, ritratto di pianta

in fiore, e studi di saldatura”  
<http://brunelleschi.imss.fi.it/genscheda.asp?appl=LIR&xsl=paginamanoscritto&chiave=101407>

[2] D.W. Thompson: “On growth and Form”, Cambridge University Press, 1917

[3] A.M. Turing: “The Chemical Basis of Morphogenesis”, Philosophical Transactions of the Royal Society of London 237(641), 37-72, 1952.

[4] D. Ruelle, F. Takens: “On the nature of turbulence”, Communications in Mathematical Physics 20(3), 167-192, 1971

[5] A.M. Turing: “Computing Machinery and Intelligence”, Mind 49, 433-460, 1950

#### Please contact:

Nadia Pisanti,  
 University of Pisa, Italy and Leiden  
 University, The Netherlands  
 E-mail: pisanti@di.unipi.it

## Alan Turing and Systems Biology

Anna Gamin, Anna Marciniak-Czochra and Damian Niwinski

***Alan Turing's achievements in the theory of computation led to his recognition as the father of modern computer science. The most prestigious award in the field is named after him. He is also widely recognized in cryptography for his work on “Cryptology bombs” - code-breaking machines that were used by the Allies during the Battle of the Atlantic. Turing's fascination with the process of thinking has led to the formulation of the definition of an abstract computing machine. By proposing a test to measure the machine's ability to exhibit human-like cognition and consciousness he initiated the field of artificial intelligence. It is less well known that he spent the last few years of his life developing mathematical theories to describe biological processes.***

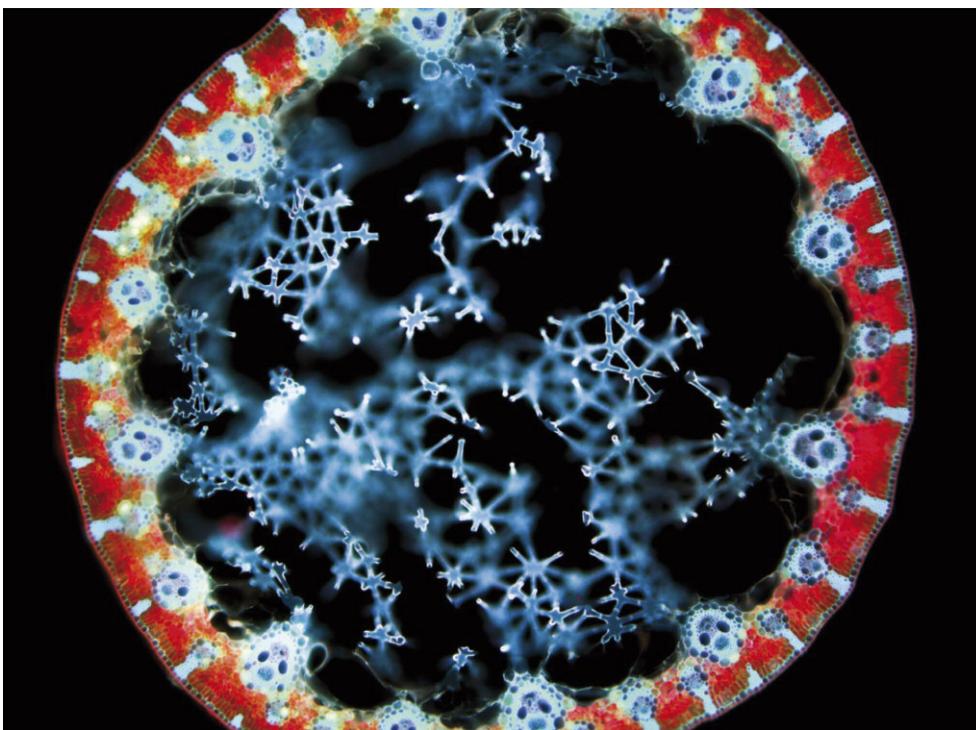
This year we celebrate the 100th anniversary of Turing's birth and the 60th anniversary of the publication of the seminal work The Chemical Basis of Morphogenesis [1], which provides the basis of the mathematical theory of pattern formation. The Turing reaction-diffusion systems had predicted phenomena such as oscillating reactions, discovered only 10 years after his premature death. The father of modern computer science not only had an enormous influence on the progress of computation technology - he also benefited from it: while working on morphogenesis and phyllotaxis, Turing used computer simulations to obtain solutions. Modern developments in biosciences provide tools to

study molecular processes in living organisms that now are carried out on an unprecedented scale, a continuation of the research initiated by Alan Turing.

If Turing, the boy fascinated by the sophisticated design of a daisy flower on the playing field, were alive today he would have a previously unavailable opportunity to observe and understand nature. Doubtless, he would use the opportunity to reconcile the beauty of mathematics with the truth of life. For this reason, Alan Turing can also be considered to be a forefather of Systems Biology - a new fast-growing field of knowledge based on collaboration between biologists, mathematicians and

computer scientists. The essence of Systems Biology, the study of systems ranging over multiple scales (such as impact of signaling pathways on cell division and further on tissue architecture), is embedded in Turing's theory of pattern formation. The multiscale study of the movement of molecules (idealized as diffusion) has been proven by Turing to contribute to the diversity of patterns in living beings, from amoebae to cnidarians to vertebrates. The implications are wide-reaching.

Next-generation sequencing projects (such as the 1000 Genomes Project) contribute to the understanding of the genetic diversity of the human popula-



*Brandner, D. and Withers, G. (2010) Images of fluorescence micrograph showing the cross-section of bulrush (*Juncus* sp.) leaf, autofluorescing red (chlorophyll on external side of leaf) and blue (vascular bundles). The Cell: An Image Library, [www.cellimagelibrary.org](http://www.cellimagelibrary.org), CIL numbers 10106, 10107, and 10108, ASCB*

tion. Microarray technologies have become standard in the study of the dynamics of gene expression, both in physiological processes and in disregulated cell signaling pathways. The study of proteins on the proteome scale is made possible by mass spectrometry. New research techniques provide data on the behaviour of objects such as genetic sequences, transcripts, and proteins. Systems Biology seeks to identify the interactions between these objects, which are in turn responsible for complex molecular processes. Analysis of these new data is challenging because of the combined complexity of the structure of the data as well as their size, which necessitates the use of sophisticated algorithmic techniques.

The value of rigorous mathematical modelling of molecular data should not be underestimated. Modelling can save time and money because the system's behaviour can be tested in silico, followed by a "wet" experiment involving only the most promising scenarios. To sum up, Systems Biology employs a combination of molecular biology, mathematics and computer science. Its main applications are in molecular medicine, but it also has a significant impact on agriculture and biotechnology.

We recommend the *Fundamenta Informaticae* Special Issue[2] in which various applications of mathematical theories to natural phenomena, inspired by Turing's work, are considered. The works in the volume represent the latest

trends in systems biology research, such as: stochastic and spatial models of molecular processes, self-organizing patterning in the development of higher organisms and state-based modelling.

Nearest to Turing's classic achievements is the work by Błazewicz and Kasprzak who study computational complexity issues inspired by computational biology. One of the issues they address is the similarities and differences between DNA computers and nondeterministic Turing machines.

A very modern approach to morphogenesis is presented by Setty et al. They describe the state-based modelling of morphogenesis, which results in a fully executable program for the interactions between chemical entities and morphogens. Furthermore, they discuss a variant of the original Turing test of machine intelligence, as a future means to validate computerized biological models.

Interactions of spatial and stochastic effects in a model of viral infection are studied by Bertolusso and Kimmel. Reaction-diffusion systems can be viewed as mean-value approximations of random walk based stochastic systems of interacting molecules.

A mathematical formulation of Driesch and Wolpert's positional information theory of emergence of organization in multicellular organisms is presented by Vakulenko and Radulescu.

Three-dimensional modeling of receptor signaling networks dynamics is a challenge solved in the paper by Archuleta et al. by means of a novel multi-resolution Monte Carlo technique. It allows observation of an interesting mechanism, adaptor protein hopping, that explains the increase in signaling efficiency when receptors are co-located.

The phenomenon of self-organized patterning during development of higher organisms is the focus of the work by Batmanov et al. The paper studies the dynamics of a community effect in space and its roles in two other processes of self-organized patterning by diffusible factors: Turing's reaction-diffusion system and embryonic induction by morphogens.

#### References:

- [1] A. M. Turing: "The Chemical Basis of Morphogenesis", Philosophical Transactions of the Royal Society of London 237 (641): 37–72, 1952
- [2] Watching the Daisies Grow: from Biology to Biomathematics and Bioinformatics — Alan Turing Centenary Special Issue, 2012 *Fundamenta Informaticae*

#### Please contact:

Ania Gambin  
University of Warsaw, Poland  
E-mail: [aniag@mimuw.edu.pl](mailto:aniag@mimuw.edu.pl)



*Accurate and exhaustive search is an attractive application for data analytics.*

## The Parallel Heartbeat of Statistical Text Analysis

by Tobias Berka and Marian Vajteršic

*In the future, all data centres will be parallel with racks of modular servers housing multi-core CPUs with fast interconnection fabric. We investigate how to build peak-efficiency data analytics software for parallel text analysis, scalable enough for large corpora, but responsive enough for interactive use. The goals are clear: to achieve supercomputer performance at cloud prices and to push the limits of text search.*

Automated text search and analysis require a formal representation of text, and term frequency vectors remain one of the most reliable forms in use. Most documents and queries contain only a few different words, thus very few terms have a non-zero frequency. Naturally, many search systems try to preserve this sparse, thinly populated structure and yet the nature of language suggests a different approach. Synonyms and hyponyms obviously counteract this sparse nature, but the ability to paraphrase goes far beyond relationships amongst individual words. Query expansion and query refinement by specification of desired and undesired search results are technical functions that cause the same problem: loss of sparseness.

One response is to use dimensionality reduction such as latent semantic indexing (LSI) to map the sparse indices into

vectors in a lower-dimensional space. The vectors become densely populated with non-zero entries and there is simply no sparseness left that could be lost in the processing. Now, however, we need to handle the cost of this processing. Clustering can be used to drastically limit the search space but even if we save what we can with brains, there always remains a large volume of computation that must be conquered with brawn.

We attempt to tackle the problem with computational efficiency and parallelism. Parallel computing not only addresses scalability, but also allows us to partition the data into parts that fit into fast main memory and communication costs are kept at a minimum by careful algorithm design. In order to remain committed to performance at all times, we have begun constructing components based on existing middleware, without subscribing to substantially new paradigms. This has allowed us to detect problems as they emerge and develop solutions accordingly.

We have developed a new form of dimensionality reduction called rare term vector replacement (RTVR), which serves as a basis for our text representation. On the Reuters RCV1-V2 corpus, it delivers a substantial reduction from 47,236 features to 392, while preserving, and even improving, the baseline performance. We have successfully parallelized this algorithm, allowing us to compute the underlying projection matrix for 800,000 Reuters documents in about 100 seconds using a 32-core Xeon cluster, instead of 20 minutes on a single core. The parallelization is based on a combined task and data parallel strategy. The task parallelism allows for distribution and loosely coupled processing, creating a potential

for fault tolerance and heterogeneous computing. The data parallelism can be used for data-partitioned, tightly coupled parallel computing, allowing us to scale to very large data sizes. The optimal performance was obtained with a fully parallel implementation combining both approaches.

The search on these dense vectors is based on the vector space model (VSM), using the cosine vector similarity as measure of document and query similarity. We have parallelized the query processing using data parallelism and found that it performs extremely well. The cosine similarity is based on a matrix-vector product, and the data transfer from main memory to the algorithmic unit dominates its computation. Consequently, the ability to use multiple caches, cache streams and memory channels has an immensely positive impact on the performance. The straightforward parallelization experiences super-linear speed-up between 130% and 170%, ie the performance gain is larger than the number of processors. While this is impossible in simple, theoretical models, the cost of memory access is not uniform and depends on a wide range of parameters. In this extreme case, parallel processing not only gives us a reduced response time, but also increases the throughput in queries per second. Data parallelism literally allows us to get the better of contemporary multi-core architectures.

Going beyond algorithms, we have begun to investigate how the middleware for parallel computing can be extended to deal with our requirements. We have developed a software interface for concurrent programming in parallel applications to effectively and conveniently model concurrent threads in parallel processes. This allows us to encapsulate interactivity and multi-user operation in threads. But there are many other technical challenges that must be solved to realize our vision. The fault tolerance and execution models currently provided are inadequate for our purposes. Document clustering and clustered search are two key algorithmic challenges that we need to address. Persistent data storage and management of results, including caching and server-side result set cursors, also remain interesting data management problems.

**Link:**  
<http://www.cosy.sbg.ac.at/~tberka>

#### References:

- [1] T. Berka and M. Vajteršic: "Parallel Rare Term Vector Replacement: Fast and Effective Dimensionality Reduction for Text", J. Parallel Distr. Com., in print
- [2] T. Berka and M. Vajteršic: "Parallel Retrieval of Dense Vectors in the Vector Space Model" CAI 2, 2011
- [3] T. Berka et al: "Concurrent Programming Constructs for Parallel MPI Applications", J. Supercomput., 2012

**Please contact:**  
Tobias Berka, Marian Vajteršic  
University of Salzburg, Austria  
E-mail: tberka@cosy.sbg.ac.at, marian@cosy.sbg.ac.at

## Ensuring Profitability of Commercial Long Term Digital Preservation

by Stephan Kiefer and Michael Wilson

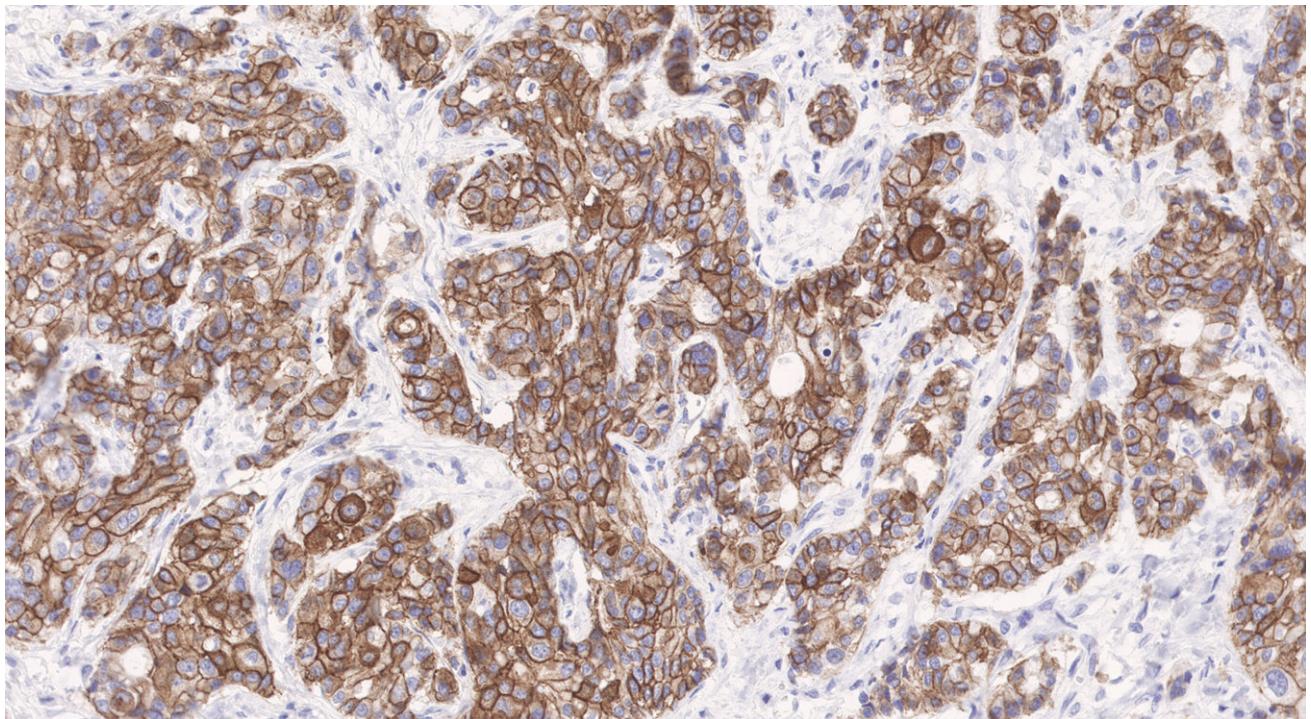
*Financial and health records are often stored in record management systems which ensure that any changes are auditably recorded to justify that the retrieved information represents that deposited, within the constraints of regulatory requirements. When the regulated retention period expires, the records are normally deleted. In contrast, data from publicly funded science is becoming preserved for the long term in repositories, to be open to discovery and access for future uses for which it was not originally collected. The experience gained from preserving scientific data over the long term is being transferred to the preservation of commercial data.*

Data preservation goes beyond backup or archiving. An important aspect of data preservation is to manage change: to the data itself resulting from corruption, in the technology which will be used on the data, and in policies applied to the data. These changes are respectively managed by: performing fixity checks on the data and replacing corrupt data, using emulators of the original technology or transforming data to formats which new technology will support, and changing the preserved data or its environment to conform to new policies.

In applying preservation technology to commercial data, it is necessary to ensure firstly that its preservation over the long term for new uses justifies the investment required [1], and secondly that the data can be discovered and accessed to meet those new uses. The EU funded ENSURE project has demonstrated a solution to the long term digital preservation of commercial data which addresses both of these issues for financial records, health records and clinical trial data. Given the current technology environment, ENSURE software supports clouds for the data storage and for the computation needed by the preservation functionality.

The approach taken to justifying profitability is to calculate the return on investment by subtracting the value of data from the cost of preservation. Cost of preservation is calculated from a hierarchical decomposition of the activities required for preservation on clouds, which can be costed individually at a low level, and then the costs can be aggregated to provide an overall cost. Private clouds have larger initial costs than commercial clouds, so the investment profiles differ accordingly.

Calculating the future value of data is a more contentious issue. Data valuation is normally based on a combination of the cost of collecting the data, the cost of not having the data, and estimates of the potential business revenue that the data could yield. The last of these can only be determined by considering potential uses of the data by a business. For scientific data there are several examples of datasets collected



*DICOM format image of a pathology slide from a patient's digital health record preserved by ENSURE.  
Image by courtesy of Philips Healthcare.*

several hundred years ago for other purposes which are now proving invaluable in modelling and predicting climate change - a purpose which was not considered when the data were collected. Similarly, in predicting future uses of commercial data such low probability high impact possibilities must be included. For example, health records can be preserved for the benefit of patients during their lifetime, and for the benefit of their descendants in the future diagnosis of inherited illness, but also for use in future epidemiological studies and medical research. Such future uses may have low probabilities, but their high potential impact contributes to the potential value of the data, and thereby to the return on investment calculations.

Predictions of future uses of data can only eventuate if the data will be discovered when it is required. Data that is indexed by its current role alone will not be easy to discover when required to meet new roles. ENSURE represents metadata for the data objects in terms of formal ontologies [2]. This supports the modelling of preservation knowledge as well as domain specific object formats and concepts effectively, in an application oriented way. The ontologies contain concepts describing general features of data objects as well as domain specific information. The metadata for data objects represent instances of the ontologies which are stored by the ENSURE software as an index. In order to ensure accessibility of structured and unstructured data by future users according to business oriented search criteria, a semantic search and query mechanism forms part of the access component. It leverages the semantic index created by the ingest component according to the ontologies in an ontology framework. The ontology framework includes an ontology registry populated by a set of initial preservation related ontologies. The ontology framework offers the flexibility required to serve future data retrieval needs. It also pro-

vides a platform to research how the evolution of ontologies over time can be exploited, both to trigger data transformations to ensure the long-term usability of the data, and to provide the basis for updating predictions of future data value to ensure the profitability of the preservation activity.

#### **Link:**

ENSURE project website: <http://ensure-fp7-plone.fe.up.pt>

#### **References:**

[1] E Conway et al: "Managing Risks in the Preservation of Research Data with Preservation Networks", International Journal of Digital Curation 7 (1) 3-15, 2012

[2] S Kiefer et al: "An Ontology-Driven Search Module for Accessing Chronic Pathology Literature", in "On the Move to Meaningful Internet Systems", Springer LNCS 7046, 382-391, 2012

#### **Please contact:**

Michael Wilson, STFC. UK

Tel: +44 1235 446619

E-mail: Michael.Wilson@stfc.ac.uk

# Learning to Recall

by Jaldert O. Rombouts, Pieter R. Roelfsema and Sander M. Bohte

*From the infinite set of routes that you could drive to work, you have probably found a way that gets you there in a reasonable time, dealing with traffic conditions and running minimal risks. Humans are very good at learning such efficient sequences based on very little feedback, but it is unclear how the brain learns to solve such tasks. At CWI, in collaboration with the Netherlands Institute for Neuroscience (NIN), we have developed a biologically realistic neural model that, like animals, can be trained to recall relevant past events and then to perform optimal action sequences, just by rewarding it for correct sequences of actions. The model explains neural activations found in the brains of animals trained on similar tasks.*

Neuroscientists have prodded the inner workings of the brain to determine how this vast network of neurons is able to generate rewarding sequences of behaviour, in particular when past information is critical in making the correct decisions. To enable computers to achieve similarly good behaviour, computer scientists have developed algorithmic solutions such as dynamic programming and, more recent, reinforcement learning (Sutton and Barto 1998).

We applied the insights from reinforcement learning to biologically plausible models of neural computation. Concepts from reinforcement learning help resolve the critical credit assignment problem of determining which neurons were useful in obtaining reward, and when they were useful.

## Learning to make rewarding eye movements

Animal experiments have shown that in some areas of the brain, neurons become active when a critical cue is shown, and stay active until the relevant decision is made. For example, in a classical experiment by Gnadt & Andersen (1988) a macaque monkey sits in front of a screen with a central cross (Figure 1). The monkey should fixate its eyes on the central cross and, while it is fixating, a cue is briefly flashed to the left or right of the cross. Then, after some delay, the fixation mark disappears. This indicates that the monkey should make an eye movement to where the cue was flashed. The monkey only receives a reward, usually a sip of fruit juice, when it executes the whole task correctly.

To solve the task, the monkey must learn to fixate on the correct targets at the correct times, and it must learn to store the location of the flashed cue in working memory, all based on simple reward feedback. The critical finding in these experiments was that, after learning, neurons were found that "remembered" the location of the flashed cue by maintaining persistently elevated activations until the animal had to make the eye movement.

## Neural network model

We designed a neural network model that is both biologically plausible and capable of learning complex sequential tasks (Rombouts, Bohte, and Roelfsema 2012). A neural network model is a set of equations that describes the computations in

a network of artificial neurons, which is an abstraction of the computations in real neurons. We incorporated three innovations in our neural model:

1. Memory neurons that integrate and maintain input activity, mimicking the persistently active neurons found in animal experiments.
2. Synaptic tags as a neural substrate for maintaining traces of an input's past activity, corresponding to eligibility traces in reinforcement learning (Sutton and Barto 1998).
3. We let the neural network predict the expected reward for different possible actions at the next time step: action values. At each time step, actions are chosen stochastically, biased towards actions with the highest predicted values.

A plausible learning rule then adjusts the network parameters to have the action values better approximate the amount of reward that is expected for the remainder of the trial. This learning rule is implemented through a combination of feed-

## Delayed Saccade Task

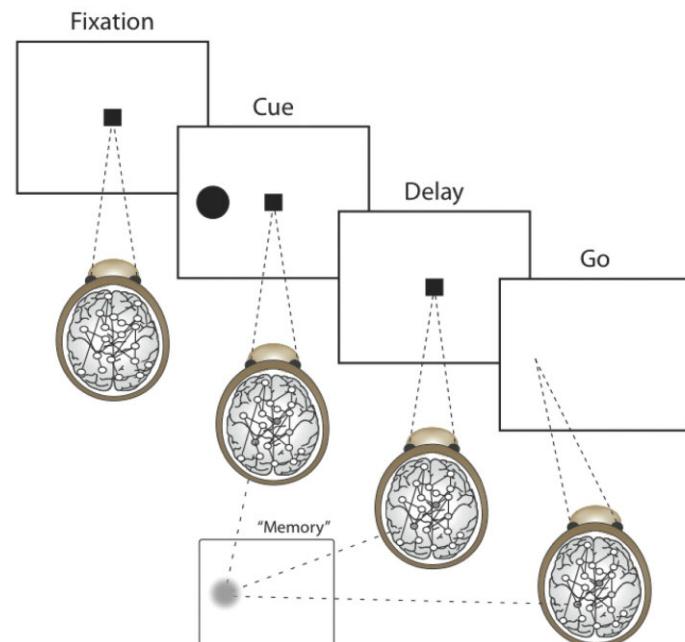


Figure 1: Delayed Saccade Task (Source: CWI)

back activity in the network, and a global reward signal analogous to the function of the neurotransmitter dopamine in the brain.

When this model was applied to complex sequential tasks like the eye-movement task described above, we find that activity in the artificial neurons closely mimics the activity found in real neurons. In the example task, integrating neurons learn to code the cue that indicates the correct action as persistent activity, effectively learning to form a working memory. Thus, the model learns a simple algorithm by trial and error: fixate on the fixation mark, store the location of the flashed cue, and then make an eye-movement towards it when the fixation mark turns off.

The neural network model solves the problem of disambiguating state information: while driving to work, some

streets look very similar; remembering the sequence of turns taken provides the information to determine your position. Mathematically, problems where instantaneous state information is aliased with other states are known as non-Markovian. Learning to extract and store information to disambiguate states is a challenge and an open problem. The neural model suggests how brains may solve some of these problems.

#### Link:

<http://homepages.cwi.nl/~rombouts>

#### References:

1. J.W. Gnadt, R.A. Andersen: "Memory Related motor planning activity in posterior parietal cortex of macaque", *Experimental brain research* 70(1):216–220, 1988
2. J.O. Rombouts, S.M. Bohte, P.R. Roelfsema: "Neurally Plausible reinforcement learning of working memory tasks", to Appear in *Advances in Neural Information Processing (NIPS) 25*, Lake Tahoe, USA, 2012
3. R.S. Sutton, A.G. Barto: "Introduction to Reinforcement Learning" MIT Press, 1998

#### Please contact:

Sander Bohte, Jaldert Rombouts  
CWI, The Netherlands  
E-mail: [S.M.Bohte@cwi.nl](mailto:S.M.Bohte@cwi.nl), [J.O.Rombouts@cwi.nl](mailto:J.O.Rombouts@cwi.nl)

Pieter Roelfsema  
Netherlands Institute for Neuroscience (NIN), Amsterdam  
E-mail: [p.roelfsema@nin.knaw.nl](mailto:p.roelfsema@nin.knaw.nl)

## The Computer and the Brain, Synergies and Robots

by Martin Nilsson

*Compared to a contemporary robot, the human body comprises a large number of actuators and sensors. Nevertheless, the central nervous system can efficiently extract just the right low-dimensional subsets of these for fast and precise motion control. How is this achieved? In the EU FP7-ICT project THE, scientists from neurophysiology, physics, computer science, and robotics are working together in order to try to answer this question. SICS' role is to try to understand and model some of the functioning of the mammalian central nervous system in order to apply it to adaptive control of robot limbs.*

#### The Synergy: a clever brain trick?

The human hand-arm system has on the order of 102 degrees of freedom, but studies [1] have shown that just a small number of combinations — "motor synergies" — of elemen-

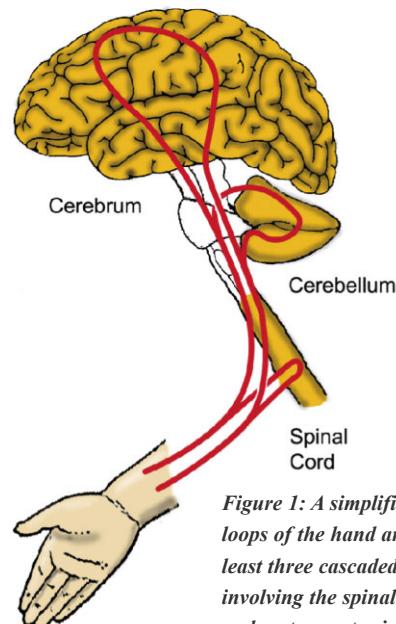


Figure 1: A simplified view of motor control loops of the hand arm system: It contains at least three cascaded feedback loops, involving the spinal cord, the cerebellum, and motor centra in the cerebral cortex.

tary movements account for most of the motion repertoire. How and why are such synergies formed? The hand-arm system also has in the order of 104 sensors. How does the brain "know" which sensor combination, "sensor synergy", best represents which motor? Knowing the answers to these questions would enable us to improve the design and control of robots. It is currently a real challenge to achieve anything approximating human agility in robots.

In the EU FP7-ICT project THE ([www.thehandembodied.eu](http://www.thehandembodied.eu)), scientists from neurophysiology, physics, computer science, and robotics are working together in order to try to find the answers. Considering the complexity of the mammalian central nervous system (CNS), this may appear to be an impossible task. However, we are beginning to see indications that much of the observed complexity in the adult is due to experience, while some of the fundamental, underlying mechanisms may be simpler than previously thought. Could it be that nature originally provides a relatively simple, general-purpose substrate, on which our interaction with the environment builds and optimizes the control circuitry? As our picture of the low-level machinery is crystallizing, some surprising properties are revealed.

#### von Neumann: ahead of his time

It is well known that when John von Neumann wrote the seminal "First Draft of a Report on the EDVAC" in 1945, he was deeply impressed by Alan Turing, but it is seldom mentioned that he was also much inspired by McCulloch and Pitts' work in neuroscience. In fact, references to the nervous system abound in the report, and the only publication referred to explicitly is their 1943 paper "A logical calculus of the ideas immanent in nervous activity". Although, for many years, there has been public debate on whether the brain can be compared to a computer, or even be understood at all, von Neumann himself considered the brain and the computer two kinds of automata. In his last work, "The Computer and the Brain", written in 1956, von Neumann compares computers with the brain, and many of von Neumann's observations are amazingly on target, more than 50 years later.

The CNS is obviously very good at controlling biological mechanical systems. If von Neumann was right, and the brain essentially is a kind of automaton, we might be able to extract useful knowledge of the brain's motor control algorithms, including the properties of synergies. This is a central theme of the THE project. SICS' role is to try to understand and model enough of the function of the mammalian CNS that it can be applied to adaptive control of robot limbs.

#### Starting with ion channels

In contrast with other attempts - which typically proceed top-down from a cognitive behavioural level - our approach is to start on the lowest, ion channel level in the structural hierarchy of the CNS. From there, we derive the function of neuronal microcircuits involved in motor control. A microcircuit is a combination of a small number of neurons operating together; it can be seen as the next level above neurons in the hierarchy. These circuits engage motor centra, the spinal cord, the cerebellum, and the pre-cerebellar structures in a complex pattern of feedback loops (Figure 1).

A serious challenge is that the macroscopic operation of microcircuits depends critically on neuronal membrane molecular dynamics, which is directly affected by thermodynamic noise. Although noise in an electronic circuit is normally undesirable, we can see that evolution has taken advantage of this in order to design robustness into circuit operation. However, the crucial function of noise in microcircuits requires non-trivial mathematical treatment, and stochastic models of neuronal activity are among the most advanced applications of the theory of stochastic processes in biology. Nevertheless, our preliminary results offer some surprises: for instance, it seems that much of the fundamental processing in the CNS is in fact — linear!

#### Project partners

SICS cooperates closely with a group of neurophysiologists led by Dr Henrik Jörntell at the Department of Experimental Medical Science at Lund University, Sweden. This group performs sophisticated electrophysiological experiments, where physical connectivity and signal transmission between neurons are recorded *in vivo*.

Other partners in the project are Centro "E. Piaggio", University of Pisa, Italy; Deutsches Zentrum für Luft und Raumfahrt (DLR), Munich; National Technical University Athens, Greece; University of Siena, Italy; Utrecht University, The Netherlands; Université Pierre et Marie Curie, Paris, France; Universität Bielefeld, Germany; and Arizona State University, USA.

#### Link:

<http://www.thehandembodied.eu>

#### Reference:

[1] M. Santello et al.: "Patterns of Hand Motion during Grasping and the Influence of Sensory Guidance", *J Neurosci* 22, 1426-1435, 2002

#### Please contact:

Martin Nilsson, SICS, Sweden  
E-mail: mn@sics.se

# Advances in Model Driven Software Engineering

by Mark G.J. van den Brand and Jan Friso Groote

*Empirical evidence shows that the use of model driven software engineering can result in an up to 10-fold quality improvement and decreased development time. Researchers from Eindhoven University of Technology tested this on a detector control system at CERN. This brings Turing's vision of software another step closer.*

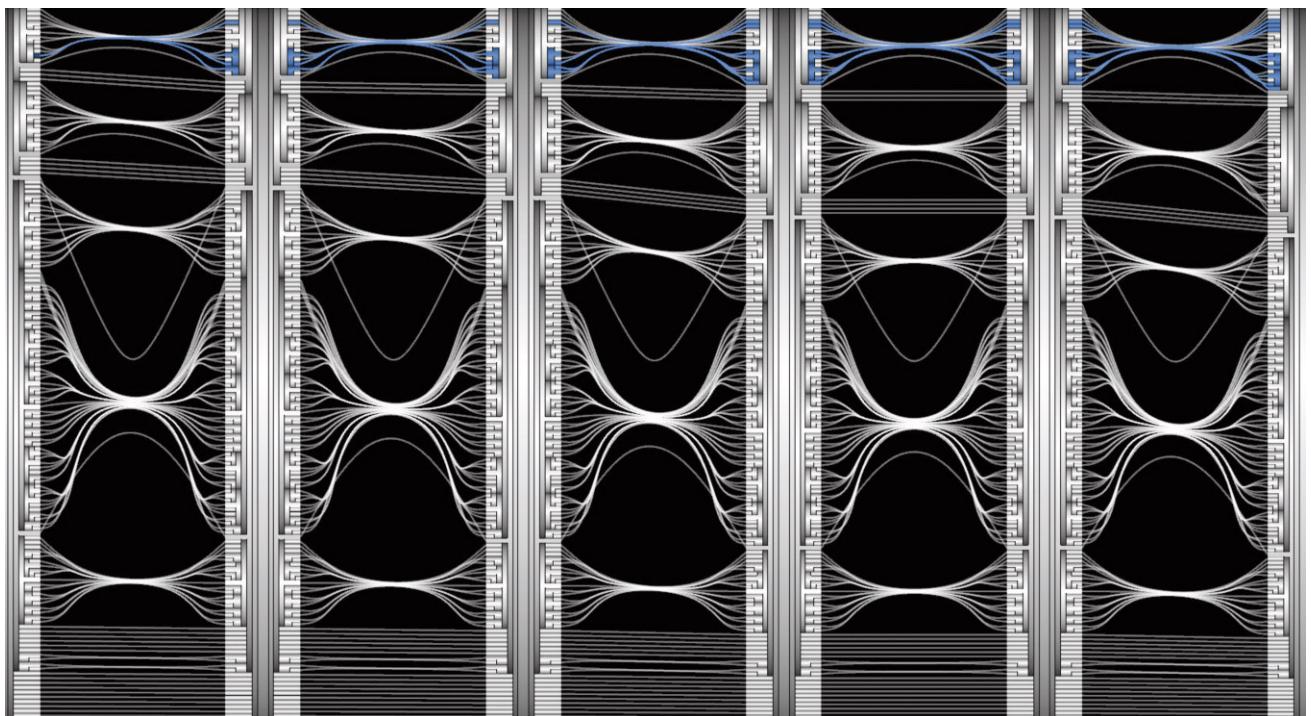
Well before the advent of modern computers, Turing anticipated the complexities of computing software. Dijkstra spent his latter life developing methods to simplify software, by mathematically deriving correct algorithms. Hoare and Milner worked on formalisms to model and understand the essence of behaviour long before concrete programs existed. Over time, the thinking on software became more abstract.

Our work is reversing this approach towards abstraction by using models as the primary step in the construction of software. First, it is shown mathematically that the models perform their intended functionality and never perform undesired and dangerous behaviour. Subsequently, these models generate software.

Careful comparison of projects that use a more classic approach compared to those that use models show an up to 10-fold reduction in bug-reports during development and an up to three-fold reduction in development times. These figures stem from the medical domain [1]. More telling are the responses by test engineers: "What have you done? Normally we find bugs in minutes. Now we find none." In reality, not all bugs are removed by model driven software engineering. But typically, the deep and intricate errors are removed and the shallow problems persist (e.g. reformulating a message for the user).

The biggest challenge of model driven software engineering is the state space explosion problem. The size of software can be described in terms of the number of states it can reach. This number is so forbiddingly large that a new name is invented to indicate the class of numbers: computer engineering numbers. Typically, the smallest such numbers are  $10^{1000}$  for a small controller through to numbers not concisely expressible with a single exponential. For comparison the largest astronomical number is 10 to the power 100.

Although models have fewer states than actual software, faster algorithms, huge computers and in particular "symbolic methods" are becoming increasingly effective in establishing the correctness of huge models. For instance, at CERN in Genève the control system of one of the detectors is modelled by approximately 20,000 interacting finite state machines from which the actual control software is generated. The model and the software suffered from a persistent liveness problem, where only part of the detector would be initialized properly. By employing symbolic methods we managed to detect and remove all such liveness problems [2].



*Figure 1: Model driven software improves both quality and development time of new software. During the process, visualizations of model transformation dependencies are used for debugging and other purposes.*

Due to the success of the applicability of model driven software engineering, a fascinating new scientific question is emerging: which kind of software modelling avoids the state space explosion problem? We provide initial answers, among which the most interesting is a preference for information polling instead of information pushing. We already see signs that these modelling guidelines are being transformed into the way actual products operate.

Model driven software engineering not only increases the quality but also the effectiveness in software development. Models are used as artifacts from which the executable code is being generated. The executable code is obtained via model transformations. The target code can be considered as a model with a lot of low-level details.

This way of working means that the overall quality of the resulting software is based on the model and the model transformations. The quality – both internal and external - of the model transformations becomes more important. The internal quality, related to understandability, maintainability and reusability, can be established via analysis of the model transformations. This can be done via metrics or visualization of dependencies between models, meta-models, and model transformations (see Figure 1). The external quality, related to correctness, is hard to determine. In order to ensure correctness of model transformations it is necessary to establish the semantics of both the input and output language and proof obligations have to be derived [3].

The research effort in the design of domain specific languages, one of the important artifacts of model driven software engineering, shifts from a syntax towards static and dynamic semantics. The research on proving model transformations correct is relatively new and unexplored but will be crucial in order to ensure the overall quality of software.

In summary, model driven software engineering is becoming increasingly effective to the point that it will soon be generally adopted.

#### **Link:**

<http://www.mcrl2.org>

#### **References:**

[1] J.F. Groote, A.A.H. Osaiweran, J.H. Wesselius: “Analyzing the effects of formal methods on the development of industrial control software”, in Proc. of the IEEE ICSM 2011, Williamsburg, VA, USA, September 25-30, pp. 467-472, 2011

[2] Yi Ling Hwong, et al.: “An Analysis of the Control Hierarchy Modelling of the CMS Detector Control System”, in Journal of Physics: Conference Series, 331(2), 2011

[3] S. Andova et al.: “Reusable and correct endogenous model transformations”, in Proc. “Theory and Practice of Model Transformations”, Z. Hu & J. de Lara (Eds.), ICMT 2012. Springer LNCS, vol. 7307, pp. 72-88, 2012

#### **Please contact:**

Jan Friso Groote, Eindhoven University of Technology  
E-mail: [J.F.Groote@tue.nl](mailto:J.F.Groote@tue.nl)  
<http://www.win.tue.nl/~jfg>

Mark van den Brand, Eindhoven University of Technology  
E-mail: [M.G.J.v.d.Brand@tue.nl](mailto:M.G.J.v.d.Brand@tue.nl)  
<http://www.win.tue.nl/~mvbrand>

# Software Engineering for Multi-core Platforms

by Farhad Arbab and Sung-Shik Jongmans

*Decades after Turing proposed his model of computation, we still lack suitable means to tackle the complexity of getting more than a few Turing Machines to interact with one another in a verifiably coherent manner. This dearth currently hampers software engineering in unleashing the full potential of multi-core platforms. The coordination language Reo, developed by the Foundations of Software Engineering group at CWI, offers a promising approach to overcome this obstacle by fulfilling the role of a domain specific language (DSL) for compositional specification of protocols.*

To utilize massively concurrent multi-core platforms, modern software engineering must develop:

- New compositional programming language constructs and paradigms that allow software developers to explicitly express the concrete interaction protocols in their applications at a high-level of abstraction, shielding them from the low-level details of the concurrency involved, and its mapping onto a multi-core architecture.
- Compositional, scalable verification, analysis, and testing techniques for reasoning about correctness and quality of service properties based on formal methods.
- Efficient, dynamic scheduling, resource allocation, and reconfiguration techniques for utilizing processor cores, memory units, and communication bandwidth, to satisfy performance and predictability requirements.

Models of concurrency embedded in conventional general-purpose programming languages offer only low-level constructs (such as threads, semaphores, locks, etc.) for building interaction protocols to coordinate the concurrent execution of computing actions. Moreover, these languages do not structurally enforce modularization of protocols as a concern separate from computation code. Consequently, dispersing protocol code among computation code comprises a common methodology for specifying structured interaction among threads.

Figure 1, a Java implementation of a typical producer-consumer scenario, exemplifies this phenomenon: lines 4-6, 9-11, 20-22, and 27-29 implement the protocol. This methodology suffers from several shortcomings. For example, intertwining computation and protocol code makes protocols nebulous, latent, and intangible. It prevents different groups from working independently on computation and protocol code, and prevents reusing code in other applications. The

negative impact of these shortcomings increases as programs grow larger and more complex.

Reo offers an alternative wherein programmers build protocols as concrete first-class constructs, by directly composing simpler (ultimately, primitive) protocols [1]. Reo [2], a visual programming language, has various formal semantics for describing the behaviour of Reo programs, called connectors, and tools for their verification and analysis [3]. These include both functional analysis (detecting deadlock, model-checking) and reasoning about non-functional properties (computing quality-of-service properties). Its declarative nature really distinguishes Reo from other models of concurrence.

```
0 import java.util.LinkedList;          16 private class Producer extends Thread {  
1 import java.util.concurrent.Semaphore; 17   public run() {  
2                                     18     while (true) {  
3 public class Main {                19       Object d = produce();  
4   private LinkedList<Object> buffer; 20       notFull.acquire();  
5   private Semaphore notEmpty;        21       buffer.offer(d);  
6   private Semaphore notFull;        22       notEmpty.release();  
7                                     23     } } } }  
8   public Main() {                  24 private class Consumer extends Thread {  
9     buffer = new LinkedList<Object>(); 25   public run() {  
10    notEmpty = new Semaphore(0);      26     while (true) {  
11    notFull = new Semaphore(1);       27       notEmpty.acquire();  
12    (new Producer()).start();       28       Object d = buffer.poll();  
13    (new Producer()).start();       29       notFull.release();  
14    (new Consumer()).start();      30       consume(d);  
15  } } } }
```

Figure 1: A Java implementation of a typical producer-consumer scenario

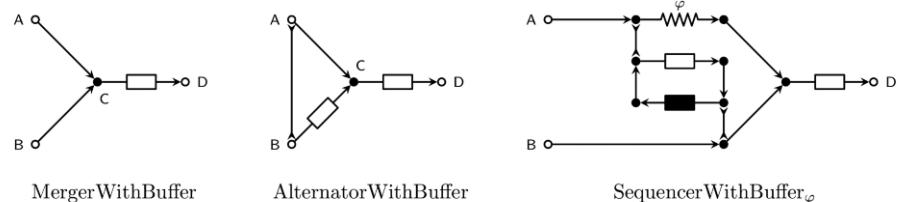


Figure 2: Example connectors

rency. Using Reo, programmers specify what, when, and why interaction takes place; not how. Indeed, Reo does not feature primitive actions for sending or receiving data elements. Rather, Reo considers interaction protocols as constraints on such actions. In stark contrast to traditional models of concurrency, Reo's constraint-based notion of interaction has the advantage that to formulate (specify, verify, etc.) protocols, one does not need to even consider any of the alternative sequences of actions that give rise to them.

Using Reo, computational threads remain completely oblivious to protocols that compose them into, and coordinate their interactions within, a concurrent application: their code contains no concurrency primitive.

The sole means of communication for a computational thread consists of I/O actions that it performs on its own input/output ports.

To construct an application, one composes a set of such threads together with a protocol by identifying the input/output ports of the computational threads with the appropriate output/input nodes of a Reo connector that

implements the protocol. A Reo compiler then generates the proper multi-threaded code for the application.

Figure 2 shows three examples of protocols expressed as Reo connectors: graphs of nodes and arcs, which we refer to as channels. Importantly, communicating parties remain oblivious to how a connector routes data: parties dispatching (fetching) data elements do not know where to (from) these elements go (come). The connector on the left, called MergerWithBuffer, specifies the same protocol as the one embedded in Figure 1.

Untangled from each other into separate computation and protocol modules, we can study, analyse, and verify the properties of each, independently of the others. We can combine the very same pieces of code that implement a set of producers and consumers with any of the connectors in Figures 2 to obtain different applications that manifest different protocols. We can reuse the same protocols, i.e., the connectors in Figure 2, with very different computational threads in entirely different applications. Reo turns interaction protocols into tangible, explicit, first-class concepts. We can connect the nodes of various instances of these Reo connectors to each other, to obtain connectors that implement more complex protocols. This protocol-level compositionality, a key feature of Reo, supports compositional, scalable construction, verification, analysis, and testing of protocols. Availability of the protocol of an application as an explicit, concrete piece of code enables efficient, dynamic scheduling and resource management.

Our on-going work on generating efficient code from Reo specifications to run on multi-core platforms has shown promising results, revealed interesting challenges, and confirms that the principles of “separation of concerns” and “modularity” apply equally well in the design, construction, verification, analysis, testing, and reuse of scalable protocols.

#### **Link:**

<http://reo.project.cwi.nl>; <http://www.cwi.nl/~farhad>

#### **References:**

- [1] F. Arbab: “Puff, The Magic Protocol”, in “Formal Modeling: Actors, Open Systems, Biological Systems”, LNCS 7000, pp. 169–206, 2011, [http://dx.doi.org/10.1007/978-3-642-24933-4\\_9](http://dx.doi.org/10.1007/978-3-642-24933-4_9)
- [2] F. Arbab: “Reo: a channel-based coordination model for component composition”, MSCS 14(3), pp. 329–366, 2004, <http://dx.doi.org/10.1017/S0960129504004153>
- [3] Sung-Shik T.Q. Jongmans, F Arbab: “Overview of Thir-ty Semantic Formalisms for Reo”, SACS 22(1), pp. 201–251, 2012, <http://dx.doi.org/10.7561/SACS.2012.1.201>

#### **Please contact:**

Farhad Arbab, Sung-Shik Jongmans  
CWI, The Netherlands  
Tel: +31 20 592 4056, +31 20 592 4241  
E-mail: [farhad@cwi.nl](mailto:farhad@cwi.nl), [jongmans@cwi.nl](mailto:jongmans@cwi.nl)

# Cybercrime and the Security of Critical Infrastructures

by Florian Skopik and Thomas Bleier

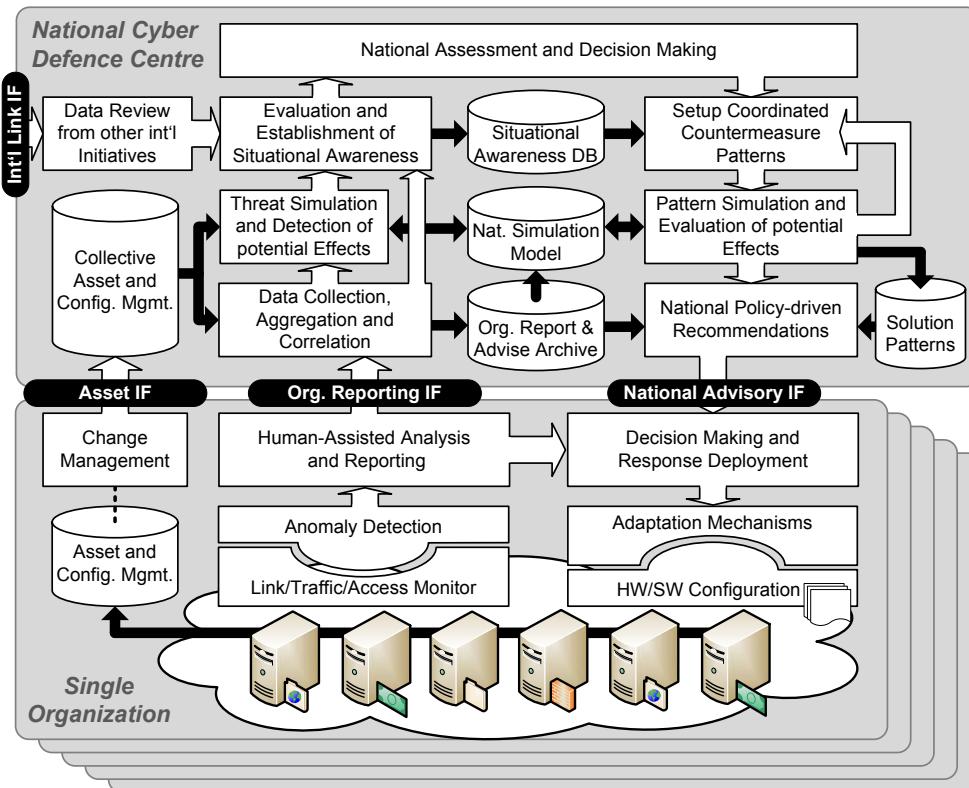
*In recent years, the Internet has rapidly expanded to a massive economic sphere of activity – not only for the new economy, where Internet-based businesses have grown from startups to multinational and billion-dollar enterprises, faster than any businesses before, but also for the “dark side” of entrepreneurship. Exploiting weaknesses in information and communications technology (ICT) systems has become a profitable business model. In order to better cope with these threats, we argue that tight cooperation between all parties in the digital society is necessary. The project CAIS deals with the implementation of a cyber attack information system on a national level, whose ultimate goal is to strengthen the resilience of today’s interdependent networked services and increase their overall availability and trustworthiness.*

In the early days of ICT, attacking other computers was mostly motivated by a desire for self-expression or competition between hackers but nowadays it has become a big business [1]. There is no clear picture of the volume of these markets, but the damage is huge. A Europol report [2] from 2011, for example, indicates losses of around € 750 billion annually. Today, spam emails are used to advertise goods and distribute phishing links or malware, viruses spread infections and carry dangerous payloads, and drive-by downloads are used to infect victims when they are accessing unsuspecting websites. Furthermore, rootkits hide the existence of other malware on a system, enabling them to act undetected for as long as possible, and botnets are used to control a large number of victim systems for malicious purposes. Even critical infrastructures, such as energy networks, transportation and financial services are becoming increasingly connected to the Internet in order to enable cost-efficient remote monitoring and maintenance. Furthermore, the pervasive use of novel computing paradigms, including mobile computing and cloud computing, makes society even more dependent on the proper functioning of ICT systems.

#### **Methodology for Protecting Networks in the 21<sup>st</sup> Century**

Traditional protection mechanisms, such as firewalls and anti-virus software are no longer able to guarantee an adequate level of security: attacks are too complex and specialized. Thus, these days we observe a major paradigm shift from prevention and remediation-focused approaches to response and containment strategies. This shift also requires organizations to move from common static policy-based security approaches towards intelligent mechanisms, incorporating identification of anomalies, analysis and reasoning of attacks, and in-time response strategies [3]. The basic properties of such approaches are:

- *Risk-based*: Prioritizing security for the most important assets. It is not economically viable for an organization or a nation to provide maximum security for all assets.
- *Contextual*: collecting huge amounts of intelligence data and use analytics to identify relevant data sources for



**Figure 1: Overall architecture of the Cyber Attack Information System**

anomaly detection. A context space is created by aggregating and correlating a wide variety of events, even if an attacker partially deleted his traces.

- *Agile*: enable fast responses to minimize the exploitable attack window and to keep (financial) losses to a minimum.

In order to cope with future advanced threats, we argue that tight cooperation between all parties in the digital society is necessary. In some domains, such as the banking sector, strategic alliances and information sharing within the community are already commonplace (e.g. to deal with phishing attacks). Furthermore, there exist relationships between organizations, such as national Computer Emergency Response Teams (CERTs), to support collaborative incident response activities. These, however, tend to be informally arranged between individual groups or are largely focused on securing infrastructures in the same operational domain. Whilst these activities have proven useful, a more comprehensive and formal approach to ensuring the security of national critical infrastructures, which spans numerous operational domains, will become necessary with the increasing use of ICT in interdependent critical infrastructure provisioning.

#### Contributions and Applied Solutions in CAIS

The project CAIS deals with the implementation of a Cyber Attack Information System on a national level (cf. Figure 1), whose ultimate goal is to strengthen the resilience of today's interdependent networked services, and increase their overall availability and trustworthiness. In particular, the following challenges are addressed and corresponding methodologies applied:

- Study of future cyber risks and emerging threats, particularly having the changing political and economic landscape in mind. Here, the well proven Delphi method – a systematic, interactive forecasting technique – is applied in consultation with an extended group of subject matter experts.

- Evaluation of novel anomaly detection techniques by composing available network tools for log file management with new pattern mining approaches inspired by models from the domain of bio informatics. Here, high performance and scalability is of paramount importance.
- Creation of highly modular infrastructure models on multiple layers, spanning hardware-centric physical aspects, over data flow and service deployment perspectives, to abstract inter-organizational dependencies.
- Innovative tools for attack simulations, using aforementioned infrastructure models and applying game-theoretic approaches as well as agent-based simulations in order to forecast the effects of attacks on interconnected infrastructures, as well as the impact of countermeasures on various levels and from multiple viewpoints.
- Investigate the deployment and instantiation of a CAIS (see Figure 1) that connects single organizations, links and coordinates isolated anomaly detection efforts, and facilitates information sharing and mutual aid between organizations.

#### CAIS Project Consortium

In order to attain these ambitious goals and finally ensure the wide applicability of developed tools, major stakeholders of Austria's security domain are involved. First, research institutions, such as the Austrian Institute of Technology and the University of Applied Sciences St. Poelten contribute their scientific expertise regarding anomaly detection techniques and infrastructure modelling and simulation. Furthermore, the OIIP Austrian Institute for International Affairs studies cyber threats and risks to national critical infrastructures caused by cyber crime. The major telecommunication service providers T-Mobile Austria and T-Systems Austria, as well as the national Austrian Computer Emergency Response Team (CERT) ensure a sound implementation on a technical layer and practical applicability and validation. The Austrian Federal Chancellery (BKA), Federal Ministry of

Interior (BMI) and the Federal Ministry of Defense (BMLVS) bring in requirements from a national security perspective. Moreover, CAIS consortium members are actively involved in international initiatives, such as the Multi National Experiment 7 (MNE7) which enables beneficial collaborations across Austria's borders. This two-year project runs from 2011 to 2013 and is financially supported by the Austrian security-research program KIRAS and by the Austrian Ministry for Transport, Innovation and Technology.

#### Links:

<http://www.kiras.at/gefoerderte-projekte/detail/projekt/cais-cyber-attack-information-system/>

<http://www.ait.ac.at/research-services/research-services-safety-security/ict-security/cais-cyber-attack-information-system/?L=1>

#### References:

- [1] J. Radianti, E. Rich, J. Gonzalez: "Vulnerability Black Markets: Empirical Evidence and Scenario Simulation", in Proc. of the 42nd Hawaii International Conference on System Sciences, pp. 1-10, 2009
- [2] Europol: Threat Assessment – Internet Facilitated Organised Crime iOCTA, 2011
- [3] EMC Press Release: RSA Chief Rallies Industry to Improve Trust in the Digital World, After Year Filled with Cyber Attacks, RSA Conference 2011, San Francisco, CA, Feb. 28, 2012

#### Please contact:

Thomas Bleier, Florian Skopik  
AIT Austrian Institute of Technology  
E-mail: thomas.bleier@ait.ac.at, florian.skopik@ait.ac.at

#### Editorial Information

*ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 8,500 copies.*

*ERCIM News is published by ERCIM EEIG  
BP 93, F-06902 Sophia Antipolis Cedex, France  
Tel: +33 4 9238 5010, E-mail: contact@ercim.eu  
Director: Jérôme Chailloux  
ISSN 0926-4981*

#### Editorial Board:

*Central editor: Peter Kunz, ERCIM office  
(peter.kunz@ercim.eu)*

#### Local Editors:

- Austria: Erwin Schoitsch, (erwin.schoitsch@ait.ac.at)
- Belgium: Benoît Michel (benoit.michel@uclouvain.be)
- Cyprus: George Papadopoulos (george@cs.ucy.ac.cy)

- Czech Republic: Michal Haindl (haindl@utia.cas.cz)
- France: Thierry Priol (thierry.priol@inria.fr)
- Germany: Michael Krapp (michael.krapp@scai.fraunhofer.de)
- Greece: Eleni Orphanoudakis (eleni@ics.forth.gr), Artemios Voyatzis (bogart@isi.gr)
- Hungary: Erzsébet Csuha-Jarjá (csuhaj@inf.elte.hu)
- Italy: Carol Peters (carol.peters@isti.cnr.it)
- Luxembourg: Patrik Hitzelberger (hitzelbe@lippmann.lu)
- Norway: Truls Gjestland (truls.gjestland@ime.ntnu.no)
- Poland: Hung Son Nguyen (son@mimuw.edu.pl)
- Portugal: Joaquim Jorge (jorgej@ist.utl.pt)
- Spain: Silvia Abrahão (sabrahao@dsic.upv.es)
- Sweden: Kersti Hedman (kersti@sics.se)
- Switzerland: Harry Rudin (hrudin@smile.ch)
- The Netherlands: Annette Kik (Annette.Kik@cwi.nl)
- United Kingdom: Martin Prime (Martin.Prime@stfc.ac.uk)
- W3C: Marie-Claire Forgue (mcf@w3.org)

# VoterBallot - A New Application for ICT in Elections

by Zaza Tabagari, Zaza Sanikidze and George Giorgobiani

*Many countries with emerging democracies aspire to a system similar to European democracy. International Organizations aim to help these countries to introduce a fairer election environment but their efforts are not always successful. Rigged elections mean that many electoral candidates are unfairly disadvantaged. In this situation, tension can escalate and, in the worst case scenario, this can lead to military confrontation.*

Researchers from the N. Muskhelishvili Institute of Computational Mathematics of the Georgian Technical University suggest a new way of applying ICT to improve the election process in a biased environment and avoid threats to democracy (see eg [1]). They propose VoterBallot, a management information communication system for electoral candidates. It includes four main components: physical infrastructure, software, structured information and training.

#### Physical infrastructure and software

At each polling station there is a trained representative of the candidate, equipped with a mobile device (or satellite phone) with extended media functions, connected through the Internet (free of any control) to the server (see eg [2]). The server is located at the electoral candidate's office (mirrors are in various places). The system's software contains: databases, forms, and various analytical programming modules, installed on the server. The main database is a three-dimen-

#### Contributions

*Contributions must be submitted to the local editor of your country*

#### Copyright Notice

*All authors, as identified in each article, retain copyright of their work*

#### Advertising

*For current advertising rates and conditions, see <http://ercim-news.ercim.eu> or contact peter.kunz@ercim.eu*

#### ERCIM News online edition

*The online edition is published at <http://ercim-news.ercim.eu>*

#### Subscription

*Subscribe to ERCIM News by sending email to [en-subscriptions@ercim.eu](mailto:en-subscriptions@ercim.eu) or by filling out the form at the ERCIM News website: <http://ercim-news.ercim.eu>*

#### Next issue

*January 2013, Special theme: Smart Energy Systems*

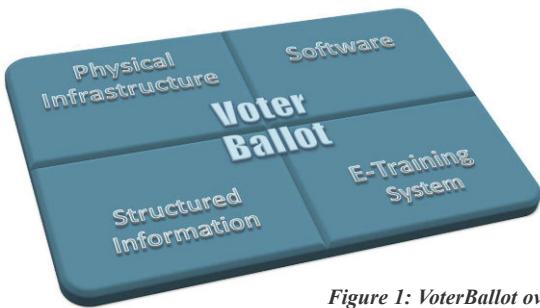


Figure 1: VoterBallot overview

sional matrix with the elements of linear arrays. The lengths of arrays are not defined in advance. The database is connected to the candidate's website. There are as many as tens of thousands of polling stations in some countries, indicating the breadth of VoterBallot's physical infrastructure.

#### Structured information

This is the concept that underpins VoterBallot. It should be designed in collaboration with the candidates and their legal teams and includes three main parts:

- 1) the collection of all scheduled events - as defined by the election law
- 2) the collection of possible irregularities - which may occur at polling stations. These depend on: the degree of influence of the governmental party on the election committees, media etc; the experience of past elections; the political education and cultural level of the population of the given country.
- 3) directives and complaints - which represent the reaction of VoterBallot to the possible irregularities as defined by point 2) above and should be designed by a professional legal team.

The total number of scheduled events and of possible irregularities ranges between 200-300 and may differ in different countries.

Formally, the element of the structured information is a linear array of data related to a given event. The structure of each event is uniform, though the actual data, collected at different polling stations, may differ. The elements of each array are the numbers or the text, photo, audio or video files. The data are time-dependent: some of the events may be repeated or adjusted for several times.

#### Implementation of VoterBallot: the training of representatives

The process of implementing VoterBallot should start several months prior to the election. During this time it is necessary to train a large number of the staff (there must be a greater number of trained staff than there are polling stations). We suggest the following e-training system:

- A Web portal for the representatives containing: the manuals, exams, etc.; an online registration facility for students (who are selected by the candidate); Webinars
- examination centers – these should be situated at each regional office

- exams – these should include theory and practical tests in near-to-real-life situations; successful students are appointed as representatives.

#### How it works

Although Election Day is the most important, the system is already loaded during the pre-election and post-election periods.

Representatives collect the information according to the forms from all polling stations and send it via the mobile devices to the server. Information is automatically analyzed and a subset is made public via the website. The matrix reacts to any irregularities and sends directives or prepared complaints with testimonies back to the mobile device. Complaints are printed and delivered to the committees or are brought to the courts later.

The benefits of VoterBallot are:

- It has the ability to react to every event, especially to the irregularities, immediately;
- the large volume of data collected makes it possible to perform various analyses (eg statistical, sociological) in real time on Election Day, and also before and after the process (see, e.g. [3]).

This work is supported by the Grant Project GEO-RECAP, #266155, FP7-INCO-2010-6. This important issue requires additional research and investigation and we hope that the idea will arouse the interest of European researchers.

#### Links:

<http://www.georecap.eu>, <http://www.compmath.ge>,  
<http://www.gtu.ge>

#### References:

- [1] S. Overton: "Stealing Democracy: The New Politics of Voter Suppression", W. W. Norton & Company, 2007
- [2] S. L. Kota, K. Pahlavan, P. A. Leppänen: "Broadband Satellite Communications for Internet Access", Springer, ISBN 978-1-4020-7659-6, 2004
- [3] J. Deckert, M. Myagkov and P. C. Ordeshook: "Benford's Law and the Detection of Election Fraud", Political Analysis, 19, 2011, p. 245–268

#### Please contact:

ZazaTabagari, Zaza Sanikidze, George Giorgobiani  
N. Muskhelishvili Institute of Computational Mathematics,  
Georgian Technical University, Georgia.  
Tel: +995 593 129 107  
E-mail: [bachanabc@yahoo.com](mailto:bachanabc@yahoo.com)

# GoodShape: Towards Flexible Mesh Generation

by Bruno Levy

*The project GoodShape, funded by the European Research Council ("Starting Grant Project"), aims at advancing the state of the art in 3D meshing.*

Mesh generation plays a central role in numerical simulation, used to predict the behaviour of engineering and physical systems under various conditions. Numerical simulation is a key to the competitiveness of many companies. It has the following benefits:

- Reducing the development cost of complex systems: in the aircraft industry, numerical simulation is an efficient alternative to wind tunnel experiments. In the automobile industry, numerical crash test simulation is clearly much cheaper than its real-world counterpart;
- Optimizing production and reducing energy consumption: in the oil and gas industry, the production of oil is influenced not only by the placement of the wells but also by a large set of parameters associated with each well. Numerical flow simulation allows evaluation of the influence of each parameter and thus optimization of the well's production of oil. More generally, it also allows managers to signifi-

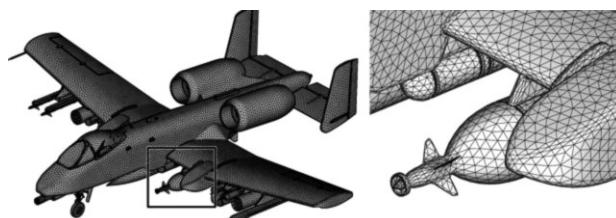


Figure 1: Mesh repair and adaptive re-meshing (data: Distene)

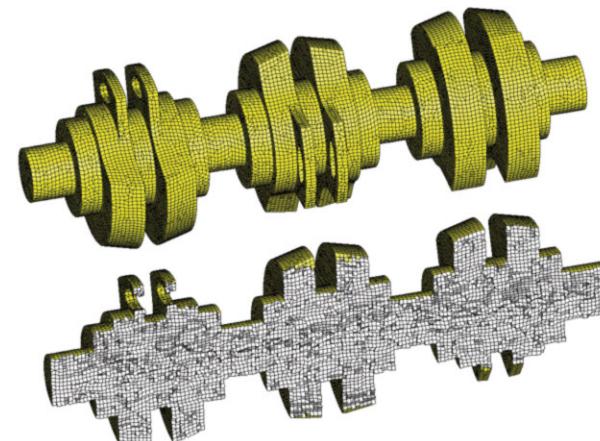


Figure 2: Hex-dominant mesh with  $L_p$  Centroidal Voronoi Tesselations (data: CM2 Computing Objects)



Figure 3: Re-meshing geological faults (data: Gocad consortium and ParadigmGeo)

nificantly optimize and reduce the consumption of energy in a wide range of systems, from small devices to large industrial infrastructures;

- Fault prevention and maintenance: by simulating the long-term evolution of a system it is possible to better understand its aging process and forecast the apparition of faults. Thus it is an efficient way of designing maintenance plans for complex systems.

For all the application domains mentioned above, the quality of the computerized representation - a 3D mesh - plays a central role in the accuracy and efficiency of numerical simulations. By "quality", engineers mean that the elements that compose the 3D mesh need to both satisfy some geometric constraints and adaptively capture the small features of the data (see Figure 1). This is a key aspect for the cited application domains. For instance, the so-called "hex dominant meshes" (Figure 2), crucial for some simulations, are notoriously difficult to generate and can require weeks (or even months) of user interaction [1].

Project GOODSHAPE takes a new approach to 3D mesh generation, based on the theory of numerical optimization. The optimal mesh generation algorithm developed in the frame of the GOODSHAPE project globally and automatically optimizes the mesh elements with respect to geometric constraints. The mathematical foundations of this algorithm, i.e. the minimization of a smooth energy function, result in practice in a faster algorithm, and - more importantly - in a higher flexibility. For instance, it will allow automatic generation of the aforementioned "hex-dominant" meshes [2],[3].

The GOODSHAPE project currently explores the industrial potential of the algorithm and conducts early technology testing in the domain of oil and gas (see Figure 3) with the GOCAD consortium, and in the domain of Computer Aided Engineering. In the long term, the goal is to push the limits of mesh generation, in terms of quality (global optimization), flexibility (resistance to error in the data, geometric constraints, hex-dominant meshing) and efficiency (parallel, multicore implementation).

## Links:

Project GOODSHAPE: <http://alice.loria.fr/goodshape>  
GOCAD consortium: <http://www.gocad.org>

## References:

- [1] Matt Staten's presentation on hex meshing (Defense & Space Sandia Labs): <http://www.scribd.com/doc/52824132/Why-Is-Hex-Meshing-So-Hard>
- [2] B. Levy and Y. Liu: "Lp Centroidal Voronoi Tessellations", ACM Transactions on Graphics, special issue SIGGRAPH conf. Proc., 2010, <http://alice.loria.fr/index.php/publications.html?redirect=0&Paper=LPCVT@2010>
- [3] Procédé de génération de maillages hex-dominants, French Patent Application, FR 10/02920 (filed 07/09/10), 2010

## Please contact:

Bruno Levy, Inria, France  
E-mail: [bruno.levy@inria.fr](mailto:bruno.levy@inria.fr)



There are two main techniques in cross-country skiing: classic –within which there are four distinct sub-techniques – and skate – which has five sub-techniques. The tool classifies a race or training session into a sequence of used sub-techniques and for each sub technique calculates a number of key performance indices. The algorithm is based on sampling three-dimensional accelerometer data, advanced filtering and pre-processing before a machine learning algorithm is used to classify the movement into correct sub-techniques.

The training tool will be used by the Swedish team for the Nordic World Ski Championships in

2013. “This is an exciting new tool for us in skiing,” says Rikard Grip, coach of the Swedish women’s cross country ski team. “It provides major opportunities to develop and optimize training. A fundamental question for skiers today is movement economy, and this provides exciting opportunities to speed up even more.”

SICS sees the project as an outstanding chance to convert several new innovative technologies developed at SICS into a fun application that is in demand. The service uses the latest sensor technology and new findings in interaction design, modelling and pattern recognition. Working with the most talented individuals in the sport—skiers, trainers and researchers—is one of the most stimulating components of the project. There are no margins here; every hundredth of a second counts.

Hans-Christer Holmberg sees major potential in combining knowledge of physiology and biomechanics with the latest sensor technology to take training and results to the next level. Everybody is extremely satisfied with how the collaboration has gained momentum.

“As head of development at SOC, naturally I see this collaboration as a way to achieve major successes at the 2014 Olympics in Sochi,” says Hans-Christer Holmberg. “But I can also see the results being adapted for several sports, including running, rowing and kayaking. By using the results in smartphones, exercise will be more enjoyable for more people, which will make Sweden healthier and more active.”

We’re only at the beginning of an exciting development in what modern IT can do for the sport. The project is a key component in the SICS initiative on the Internet of Things, the vision of an internet that connects not only people, but also objects in a context that benefits people’s lives.

#### Please contact:

Christer Norström, SICS, Sweden

E-mail: [cn@sics.se](mailto:cn@sics.se)

Hans-Christer Holmberg, Swedish Olympic Committee

## Math Strengthens the Swedish Olympic Cross-country Team

by Kersti Hedman

**SICS and the Swedish Winter Sport Research Centre at Mid Sweden University have initiated a collaborative project to demonstrate how cutting edge technology and advanced mathematics can help to develop new training methods to achieve success at the 2014 Olympics in Sochi. The tool measures how skiers move, and helps optimize technique and training.**

The new partnership consists of a “dream team” comprising the world’s most experienced sports researchers in biomechanics and physiology and the best skiers in the country, combined with the foremost experts in advanced mathematical modelling in Sweden. Swedish Winter Sport Research Centre at Mid Sweden University, led by Professor Hans-Christer Holmberg, also head of development at the Swedish Olympic Committee (SOC), known primarily for the development of cross-country skiing, accounts for the domain knowledge, while SICS and partners are addressing the technical solutions.

In brief, the service is a cell phone application that continuously registers and provides information about how the skiers move and their movement economy. The sensors that provide the information are in an ordinary Android phone worn on the body together with a traditional heart rate monitor. The captured data are transmitted via the Internet to processing in “the cloud” using advanced algorithms. Data are then sent immediately back to the skier or trainer in the form of useful, understandable information that can optimize training.

# LUDUS: Serious Gaming Initiatives in South East Europe

by György Kovács

**The LUDUS project aims to create a European network for the transfer of knowledge and dissemination of best practices in the innovative field of serious gaming.**

The South East Europe Transnational Cooperation Programme of the EU (SEE) aims to develop transnational partnerships on matters of strategic importance, in order to improve the territorial, economic and social integration process and to contribute to cohesion, stability and competitiveness of the region. To this end, the programme seeks to realize high quality, result-oriented projects with a strategic focus, relevant for the programme area.

Serious games (SG) are well summarized in [1]. The term “serious game” was actually used long before the introduction of computer and electronic devices into entertainment by Clark Abt, who introduced the concept in his 1970 book *Serious Games* [2]. He gave a useful general definition, which is still applicable in the computer age:

“A game is an activity among two or more independent decision-makers seeking to achieve their objectives in some limiting context. A more conventional definition would say that a game is a context with rules among adversaries trying to win objectives. We are concerned with serious games in the sense that these games have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement.”

Gaming has been used in educational circles since at least the 1900s. Use of paper-based educational games became popular in the 1960s and 1970s, but waned under the Back to Basics teaching movement. (The aim was to focus on students’ poor reading, writing and arithmetic). With the proliferation of computers in the 1980s, the use of educational games in the classroom became popular, with titles that included Oregon Trail, Math Blaster, and Number Munchers.

These days, SGs are generally considered to be applications, developed with game technology and design principles, whose primary purpose is to train, to simulate situations, or to educate while entertaining the user. Serious games tend to be very effective and pervasive as they partially hide learning, the user has an active role in them, and they are almost always amusing.

The LUDUS consortium has eight partners from six countries: PROMEA and BIC (Greece), POLIMI and UNIMIB (Italy), ARIES (Romania), BSC (Slovenia), BIA (Bulgaria), and INNOSTART (Hungary). (In the last ten months of the project INNOSTART resigned and was substituted by SZTAKI in Hungary). LUDUS is a real South-East European consortium containing different types of organizations, including: a research institute, an innovation centre, an industrial association, a university, a didactical centre and a development agency.

The project had several ambitious goals and targets ranging from theoretical analytical studies to practical applications of SGs, including networking in a physical and human sense, and applying knowledge management. Some basic goals include:

- understanding and analysing the state of the art in serious gaming by bringing together SG stakeholders with IT companies and developers.
- to promote SG for different partners by means of organization of European training courses, competitions, etc., by supporting knowledge transfer, by setting up a web-based collaboration and knowledge management platform, hosting database with experts, documents, outcomes.
- support new research and development in serious gaming.
- achieve (in the longer term) a more intensive use of technology for learning/training purposes and innovation in Europe.

The target groups of LUDUS are local SMEs and enterprises, IT companies, scientific and engineering associations, chambers of commerce, regional development agencies, press, media, teachers, trainers, educational experts, decision makers, training companies, students, gamers and even the general public.

Interaction between developers and users, as well as feedback from the latter, is necessary in order to make improvements and enable the field of serious gaming to achieve its



Infoday and regional networking workshop in Athens (19 June 2012)

full potential. Users and others are encouraged to submit their comments on functionality, themes and application of SG facilities, through the project's communication tools. LUDUS wishes to offer these communication tools as a space for the exchange of opinions and the birth of new ideas, attracting and involving all stakeholders.

Some main events during the project's life-cycle to date include:

- a kick off meeting in Ioannina, Greece and a project meeting in Milan, Italy
- three info days in Italy, Romania and Bulgaria
- an open Brainstorming meeting in Slovenia
- Knowledge Sharing Regional Training Course in Italy
- The Game Barometer – A survey on the subject of SGs
- two European best learning game competitions - Romania and Slovenia
- two international conferences, Iioannina and Milan
- Serious Gaming Open Learning Lab, Milan
- Infoday and regional networking workshop in Athens, Greece(see Figure 1)

These meetings and open events attracted several hundred participants in each country, representing all target groups. Competitions were organized in different categories, for different size enterprises with several competing teams.

LUDUS is funded (up to 85%) by the South East Europe Programme under Priority Axis 1: Facilitation of innovation and entrepreneurship.

#### Links:

- <http://www.ludus-project.eu/>
- <http://www.serious-gaming.info/>
- <http://www.abtassociates.com/page.cfm?PageID=452>
- [http://en.wikipedia.org/wiki/Serious\\_game](http://en.wikipedia.org/wiki/Serious_game)

#### References:

- [1] M. Floryan: "A Literature Review of the Field of Serious Games", Computer Science Dept., University of Massachusetts, Amherst, 2009, pp. 1-16, under the supervision of Prof. Beverly Woolf
- [2] C. Abt: "Serious Games", Viking Press, 1970

#### Please contact:

György Kovács, SZTAKI, Hungary  
E-mail: [kovacs.gyorgy@sztaki.mta.hu](mailto:kovacs.gyorgy@sztaki.mta.hu)

# A New Robotic Laboratory at SZTAKI

by György Kovács and Imre Paniti

*In the summer of 2011 a new robot laboratory began operation in SZTAKI, in the Computer Integrated Manufacturing Research Laboratory, with the aim of supporting international networking and raising the public and professional profile of research and development in this field.*

SZTAKI was one of the first institutes in Hungary to initiate robotic and intelligent manufacturing research around 40 years ago. The institute achieved outstanding results in fields related to robotics and manufacturing automation, including: pattern recognition, sensor technology, controller developments and virtual and extended manufacturing. The establishment of the new robotic laboratory will allow us to join high-level EU R&D efforts in joint projects, and to extend our scientific- and application activities.

The new laboratory has two robots (one small and one large) and a small 2.5D milling machine. The small robot has a circular work envelope of 445 mm maximum horizontal reach and a maximum payload of 1.2 kg, while the milling machine has a working space of 400x250x50 mm (see Figure 1). The industrial robot (see Figure 2) has the following specifications: Axes: 6, Payload: 130 kg, H-Reach: 2488 mm, Repeatability: ±0.5 mm, Controller: FANUC R-J3. This is a rather big and powerful tool that showcases several main features that are requested from "real industrial robots".

To get the laboratory up and running, we had to design and implement several tools, including:

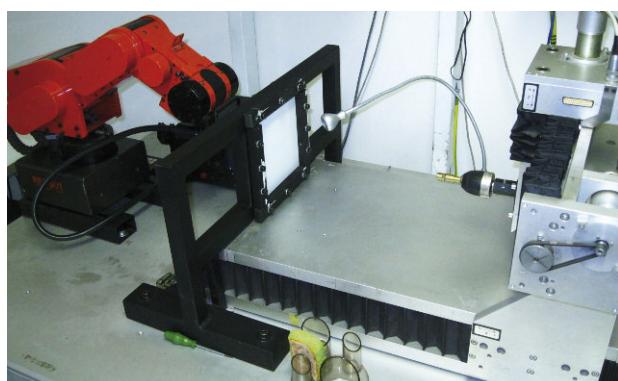


Figure 1: Set-up with small robot and milling machine



Figure 2: New set-up with the FANUC S-430iF robot

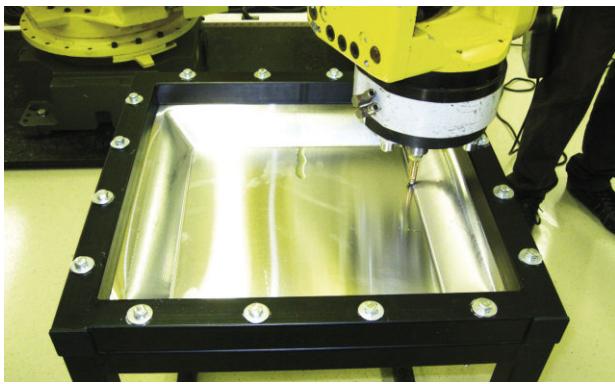


Figure 3: ISF experiment with the FANUC S-430iF robot

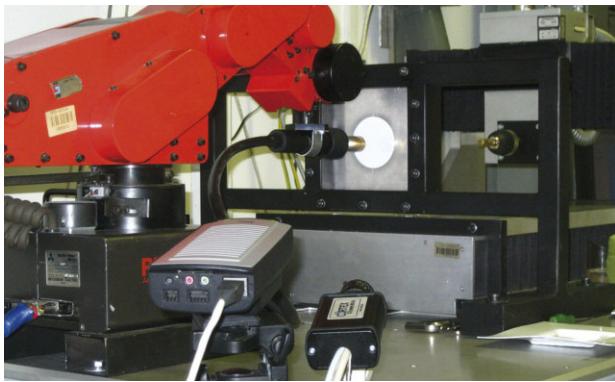


Figure 4: FANUC S-430iF robot, lights-camera box and Rubik's Cube

- mechanical stands to install the robot and the controller, enhancement of the controller with I/O modules, installation of industrial camera, mirror, etc. to prevent ceiling collision.
- appropriate frames and tools for the robot to run incremental sheet forming (ISF) experiments [1] (see Figure 3).
- an environment (lights-camera box and software to recognize the starting situation of the magic cube, and an appropriate gripper). The software to control turning sequences and robot movements was developed and installed to solve Rubik's Magic Cube puzzle (see Figure 4).

Recently, three experiments have been running on the laboratory's three tools: Two deal with ISF, while the third is a kind of a "Hungaricum" to solve the magic (Rubik's) cube.

Incremental sheet forming is a relatively new technology in sheet forming and there are various research questions still to address, for instance questions relating to thickness, stress, force and torque measurements on the sheet and on the tool. We are performing several measurements to determine the effects of heating on a sheet, and to understand the relationships among the aforementioned variables. We started to deal with sheet forming within the framework of a successfully completed EU FP VI project ("SCULPTOR") in collaboration with German and Spanish partners and we have since become world wide accepted experts in this area of research.

In our first ISF experiment, we are experimenting with heating a polymer sheet from one side (using the small robot) and forming it with an appropriate tool from the other side (the milling machine). The processes in this setup can be started and monitored in VirCA (Virtual Collaboration Arena), a loosely coupled modular, 3D Internet based interactive virtual environment for collaborative manipulation of robots

and other hardware or software equipment [2], [3]. The VirCA system has been developed within HUNOROB, a Hungarian–Norwegian research based innovation project. Optical measurements for sheet thinning and thermal behaviour investigation are started.

In our second experiment we are investigating ISF of metal sheets with the FANUC robot using a spherical edge tool and the specific frame to hold the sheet.

Our third experiment aims to control the robot and the optical box to solve the magic cube (Rubik's) puzzle by learning it and then by turning it to its requested positions/colours with the FANUC robot. The light box assists in the learning of the initial colours of all six sides of the cube, and then the same hole on the side of the box assists in turning it. This experiment allows us to investigate different mathematical algorithms to solve the cube using different robot movements, such as sudden deceleration or acceleration, travelling slowly or fast, and making different, complicated turns.

#### Some future plans:

- We plan to realize and test a recently submitted EU patent application (Device for Two Sided Incremental Sheet Forming) and a new adjustable sheet holder frame with the capacity to work with different sheet sizes.
- We plan an optical robot control to recognize and find and then grip and move known objects with the FANUC robot.
- Serious industrial experiments are planned for assembly, disassembly, part sorting and perhaps for point and line welding.

In conclusion, the laboratory will be able to take part in international cooperation as a real and a virtual laboratory, continuing our experiments and collaborating with the leading Spanish and German laboratories.

#### Links:

##### Incremental Sheet Forming:

[http://en.wikipedia.org/wiki/Incremental\\_sheet\\_forming](http://en.wikipedia.org/wiki/Incremental_sheet_forming)

##### Rubik's Cube:

[http://en.wikipedia.org/wiki/Rubik%27s\\_Cube](http://en.wikipedia.org/wiki/Rubik%27s_Cube)

##### CIM Lab. at SZTAKI:

<http://www.sztaki.hu/department/CIMLab/>

Virtual Collaboration Arena: <http://virca.hu>

#### References:

- [1] I. Paniti: "CAD API based tool path control for novel incremental sheet forming", Pollack Periodica, Volume 5, Number 2, pp. 81-90, 2010
- [2] P. Galambos: "VirCA as Virtual Intelligent Space for RT-Middleware", in Proc. of the 2nd International Conference on Cognitive Infocommunications, CogInfoCom 2011, 3-7 July 2011, pp. 140-145
- [3] J. Nacsá, I. Paniti, S. Kopácsi: "Incremental Sheet Forming in Cyberspace - a Process Oriented Cognitive Robotics Application", in Proc. of the 2nd International Conference on Cognitive Infocommunications, CogInfoCom 2011, 3-7 July 2011, pp. 1-5.

#### Please contact:

György Kovács, SZTAKI, Hungary

Tel: +36 1 279 6140, E-mail: [kovacs.gyorgy@sztaki.mta.hu](mailto:kovacs.gyorgy@sztaki.mta.hu)

# Engineering Asset Lifecycle Optimal Management: WelCOM Approach to E-Maintenance

by Christos Koulamas, Petros Pistofidis and Christos Emmanouilidis

*In the modern era of global competition, reduced financial operation margins, and rapid market changes, enterprises cannot afford to underutilize their resources. The efficiency of a production line is largely influenced and ultimately defined by the reliability of its technical equipment. Optimal and cost-effective management of engineering assets is essential. This cannot be achieved with the traditional “fail-and-fix” policies (corrective maintenance). Rather, a more proactive preventive maintenance is necessary.*

Condition-based maintenance (CBM) is a maintenance management strategy that plans maintenance actions on the basis of the actual machinery or asset condition, as opposed to pre-determined (e.g. scheduled) maintenance. A CBM strategy presupposes the implementation of condition monitoring, that is the process of monitoring machinery condition parameters in order to infer the machinery condition and offer maintenance actions recommendations accordingly. Maintenance-related data and services must be ubiquitously and transparently available at anytime, anywhere, to any authorized person. This is key towards realizing e-maintenance.

The WelCOM platform [1], depicted in Figure 1, is moving in this direction. It offers multi-layered intelligence for the optimal operation and maintenance (O&M) of the equipment during its productive period (middle-of-life). At the device level, an innovative and versatile optical sensor is designed and ported in both wireless and wired form which can be used for diverse monitoring applications.

The optical sensors are integrated in networked embedded systems forming a wireless sensor network that can sustain operation in industrial settings. The wireless sensor network nodes implement a distributed sensor-embedded intelligence based on condition state models and a novelty detection engine. The overall system integrates real-time condition monitoring and novelty detection with accurate fault diagnosis; support of prognosis and prediction of failures; and effective notification for immediate attention and maintenance.

The maintenance data can be provided by heterogeneous sources; hence data integration is a major obstacle for wide adoption. To this end, maintenance information modelling will be utilized and based on open specifications. Domain ontology has been developed to express the semantics for the maintenance problem space based on the MIMOSA OSA-CBM specifications [2]. Also, data interoperability and exchange formats and interfaces have been defined so as to allow seamless information flow between the wireless sensor network layer and the knowledge management layer.

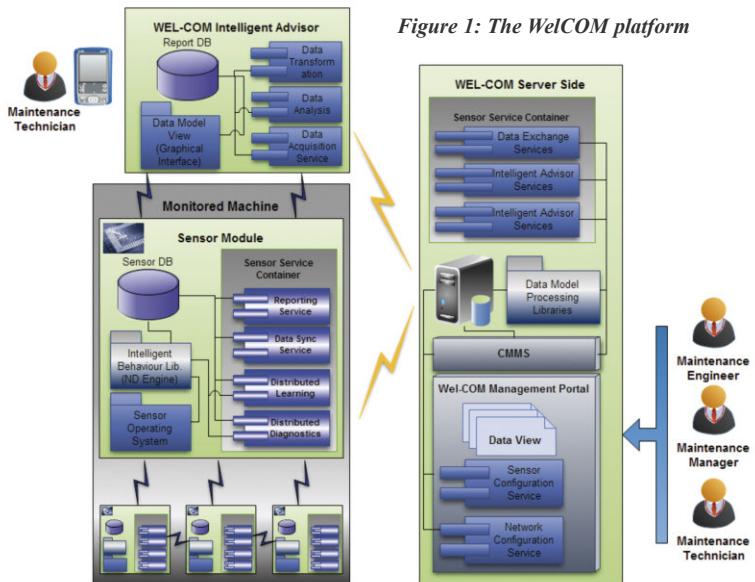


Figure 1: The WelCOM platform

The WelCOM configuration and management portal will offer all authorized personnel access to real-time, on-the-spot sensor feeds of modelled machine state data, while traditional CBM approaches allow only delayed, off-line processed information. The portal will couple services and sensor software components as to enable “zooming” from the administration level down to managing and calibrating any single sensor device of the monitoring network.

The knowledge produced within the platform is not only available in the control room. A significant component is the portable intelligent maintenance advisor. This module is available as a software for portable devices (laptops, tablets, smartphones, etc.) that scans through the knowledge portal and retrieves proper “knowledge objects” to aid the maintenance practice of a technician in the field.

Finally, a training module will be integrated into the WelCOM platform aiming to achieve a defined level of competencies for maintenance technicians by harnessing the derived knowledge. The e-training is a practical and cost-effective alternative to expensive on-the-job training.

This research is conducted by a Greek public-private partnership between the Research and Innovation Center Athena (including its Industrial Systems Institute) and the NHRF public research institutions and Kleemann Hellas Lifts, Prisma Electronics, GDT, and Atlantis Engineering as industrial partners. The project is supported by the GSRT (grant no. 09SYN-71-856).

**Link:** <http://welcom-project.ceti.gr/>

## References:

- [1] P. Pistofidis et al: “A Layered E-Maintenance Architecture Powered by Smart Wireless Monitoring Components”, presented at the 2012 IEEE Conference on Industrial Technologies, ICIT 2012
- [2] MIMOSA - An Operations and Maintenance Information Open System Alliance web site <http://www.mimosa.org/>

## Please contact:

Christos Emmanouilidis, Athena RIC, Greece  
E-mail: christosem@ieee.org

# CLEF 2012 and Beyond: Perspectives for the Conference and Labs of the Evaluation Forum

by Nicola Ferro

Since 2000, CLEF has played a successful role in stimulating research and promoting evaluation in a wide range of key areas in the information access and retrieval domain. In 2010, a radical innovation and renewal process led to the establishment of the CLEF Initiative, whose mission is to promote research, innovation, and development of information access systems with emphasis on multilinguality and multimodality. The CLEF Initiative is structured in two main parts:

1. a series of Evaluation Labs, i.e., laboratories to conduct evaluation of information access systems and workshops to discuss and pilot innovative evaluation activities;
2. a peer-reviewed Conference on a broad range of issues, including
  - the activities of the Evaluation Labs;
  - experiments using multilingual and multimodal data; in particular, but not only, data resulting from CLEF activities;
  - research in evaluation methodologies and challenges.

Due to these changes and the broader scope of the CLEF Initiative, the acronym CLEF, traditionally expanded to Cross-Language Evaluation Forum, now translates to Conference and Labs of the Evaluation Forum. This renewal process and the organization of the annual CLEF events are partially supported by the EU FP7 PROMISE project (contract n. 258191): Participative Research labOratory for Multimedia and Multilingual Information Systems Evaluation and by the ELIAS RNP network.

## CLEF 2012: Information Access Evaluation meets Multilinguality, Multimodality, and Visual Analytics

The annual meeting of the CLEF Initiative was hosted this year by the Sapienza University of Rome, Italy, 17-20 September as a 4 day event in which



*CLEF 2012 participants.*

conference presentations, laboratory meetings, workshops, and community sessions were smoothly interleaved to provide a continuous stream of discussions on the different facets of experimental evaluation.

14 papers and 3 posters were accepted for the Conference and published by Springer in their Lectures Notes for Computer Science (LNCS) series. Two keynote speakers highlighted important developments in the field of evaluation. Peter Clark, Vulcan Inc., USA, focused on how to move from information retrieval to knowledgeable machines. Tobias Schreck, University of Konstanz, Germany, presented research challenges for visual search and analysis in textual and non-textual documents and their evaluation.

The community sessions at CLEF 2012 were organized around the presentation of the activities of other evaluation initiatives and an “Evaluation Clinic” where participants had the possibility of meeting evaluation experts and discussing with them their evaluation problems.

Seven benchmarking activities ran as evaluation labs in CLEF 2012. The results were presented and discussed in Rome in dedicated sessions:

- CHiC (Cultural Heritage in CLEF, new): investigating systematic and large-scale evaluation of cultural heritage digital libraries and information access systems;

- CLEF-IP: studying IR techniques in the patent domain;
- ImageCLEF: proposing experimental evaluation of image classification and retrieval, with a focus on the combination of textual and visual evidence;
- INEX (new in CLEF): evaluating XML retrieval;
- PAN: uncovering plagiarism, authorship and social software misuse;
- QA4MRE: evaluating machine reading systems through question answering and reading comprehension tests;
- RepLab (new): evaluating reputation management technologies.

There was also an exploratory workshop: CLEFeHealth 2012 (new) on cross-language evaluation of methods, applications, and resources for eHealth document analysis.

## CLEF 2013 and Beyond

During CLEF 2012 considerable steps were taken to further reshape and improve the overall organization of CLEF. In order to allow for more time for planning and organizing the activities the bidding process for both CLEF 2013 and CLEF 2014 was completed and the host venues of the annual meeting for the next two years were decided.

CLEF 2013 will be hosted by the Technical University of Valencia, Spain, 23-26 September 2013 while CLEF 2014 will be hosted by the University of Sheffield, United Kingdom, 15-19 September 2014.

The Call for Participation in the CLEF2013 Evaluation Labs and the Call for Papers for the CLEF 2013 Conference will both be released in November 2012. Lab registration will open in December 2012 and the expected deadline for the submission of papers to the conference is late April 2013. All details can be found on the CLEF website.

Finally, bids for hosting CLEF 2015 are now open and will close on 5th April 2013. Proposals can be sent to the CLEF Steering Committee Chair at chair@clef-initiative.eu.

#### Links

CLEF: <http://www.clef-campaign.org/>

CLEF 2015 Template for Bids:

[http://www.clef-initiative.eu/documents/71612/87713/CLEF-Initiative-Template\\_for\\_bids.docx](http://www.clef-initiative.eu/documents/71612/87713/CLEF-Initiative-Template_for_bids.docx)

PROMISE: <http://www.promise-noe.eu/>

ELIAS: <http://www.elias-network.eu/>

#### Please contact:

Nicola Ferro

University of Padua, Italy

E-mail: ferro@dei.unipd.it

## Turing Year in Spain

by Juan José Moreno Navarro

**A few years ago, Time magazine published a list of the 100 greatest minds of the 20th century, which included Alan Mathison Turing, alongside the Wright brothers, Albert Einstein, the DNA breakers Crick and Watson, and the discoverer of penicillin, Alexander Fleming.**

As clearly shown by the special theme section, we have Turing to thank for many concepts that underpin technologies that are now part of our daily lives: the first computers, the program stored, artificial intelligence, software verification and modelling. Without Turing and his brilliant ideas we would not be able to shop online, watch a video on a tablet, remotely manage our finances, play our favorite music on an mp3 device, send emails, get an x-ray or travel on high-speed trains. Like many great ideas,

such as the wheel or the arch, which in hindsight seem obvious, with his one invention, the general-purpose computer, Turing changed the world.

What is remarkable about Turing is the breadth and scope of his contributions to society. While some of the individuals on Time's list have made their mark in history with a single great contribution, Turing made numerous contributions. He was responsible for: the foundations of artificial intelligence, proposing the Turing Test, the basis of the verification and validation of software, the first uses of algorithmic modelling of natural phenomena (in this case, biological or morphogenesis patterns), connectionist networks, and more. Few scientists have been such visionaries in their field as Turing was in computer science.

A wonderful way to honour Turing is to raise public awareness of his work. Whilst the achievements of scientists like Kepler, Galileo, Newton, Darwin and Einstein are widely appreciated by the general public, fewer people are aware of the work of Turing, although his ideas arguably have had an even greater impact on our daily lives.

Unfortunately there is much work to do: even those with the best intentions, in trying to popularize Turing's work, have tended to focus on his work as an accomplished code breaker and hero of the Second World War or his contribution to Artificial Intelligence. But his contributions to the creation of the modern computer deserve a place of honour in the Olympus of the greatest scientists of all time.

The commemoration of Turing Year / Year of Informatics is coordinated by the Spanish Scientific Computer Science Society (SCIE – [www.scie.es](http://www.scie.es)) in collaboration with the Conference of Directors and Deans of Computer Science of Spain (CODDII – [coddii.org](http://coddii.org)). It began in Madrid with a formal opening ceremony on 27 July, and will end in June 2013. After that, SCIE will organize the Spanish Conference of Informatics - CEDI, a biennial event that brings together more than 2,000 Spanish researchers in computer science. An Organizing Committee is coordinating the event, with the assistance of an Advisory Committee chaired by Prince Felipe de Borbón.

As well as being a tribute to Turing, the celebration in Spain has an additional goal in showing colleagues, science policy makers and the general public that computer science (CS) is a very active area of research with excellent indicators: Spain occupies the 7th place in the world ranking, produces around 7% of Spanish scientific publications and represents around 3.6% of the CS production of the world, being, in summary, the most dynamic research subject area in Spain.

In celebration of the Turing year we have planned several activities including academic, scientific, dissemination and industry collaboration. Check <http://turing.coddii.org/turing> for a detailed list of activities.

Two of the academic and scientific activities of note include: the 7th annual celebration of the National Awards in Informatics, and a summer course in the prestigious Universidad Internacional Menéndez Pelayo entitled "1st Meeting in Commemoration of Alan Turing". This very successful and exciting meeting took place in August 2012.

In the dissemination area we have been actively engaging with media. We have had media coverage in high profile-newspapers, broadcasting and TV and there will be more to come. In particular, a blog is running in El País (the most followed newspaper in Spain). This is a Spanish-language blog that can be viewed at <http://blogs.elpais.com/turing/>. We are also planning activities in museums and schools and an ambitious exhibition on computer art developed in Spain in the sixties. The exhibition, which is usually based in Madrid, will be visiting several other cities in Spain.

Our goal is to act as exemplary and enthusiastic ambassadors of computer science and to provide Spanish society with a deeper understanding of a singular figure and of the discipline of computer science.

#### Please contact:

Juan José Moreno Navarro

Head of organizing Committee, Turing year in Spain

Technical University of Madrid, Spain

E-mail: [jjmoreno@fi.upm.es](mailto:jjmoreno@fi.upm.es)



# HCI International 2013

Las Vegas, USA, 21-26 July 2013

The 15th International Conference on Human-Computer Interaction, HCI International 2013, will be held jointly with the affiliated Conferences:

- Human-Computer Interaction thematic area
- Human Interface and the Management of Information thematic area
- 10th International Conference on Engineering Psychology and Cognitive Ergonomics
- 7th International Conference on Universal Access in Human-Computer Interaction
- 5th International Conference on Virtual, Augmented and Mixed Reality
- 5th International Conference on Cross-Cultural Design
- 5th International Conference on Online Communities and Social Computing
- 7th International Conference on Augmented Cognition
- 4th International Conference on Digital Human Modeling and applications in Health, Safety, Ergonomics and Risk Management (formerly International Conference on Digital Human Modeling)
- 2nd International Conference on Design, User Experience and Usability
- 1st International Conference on Distributed, Ambient and Pervasive Interactions - NEW
- 1st International Conference on Human Aspects of Information Security, Privacy and Trust - NEW

Extended deadline for paper and tutorial proposals:  
9 November 2012.

HCI International 2013 is expected to attract over 2,000 participants from all over the world. The program will feature, among others, pre-conference half-day and full-day tutorials, parallel sessions, poster presentations, an opening session with a keynote address, and an exhibition.

The Conference Proceedings will be published by Springer in a multi-volume set. Papers will appear in volumes of the LNCS and LNAI series. Extended Poster abstracts will be published in the CCIS series.

The best paper of each of the Affiliated Conferences / Thematic Areas will receive an award. Among these best papers, one will be selected to receive the golden award as the Best HCI International 2013 Conference paper. Finally, the Best Poster extended abstract will also receive an award.

## Links:

HCI International 2013: <http://www.hci2013.org/>  
 HCI International Conference Series: <http://www.hci-international.org/>

## Please contact:

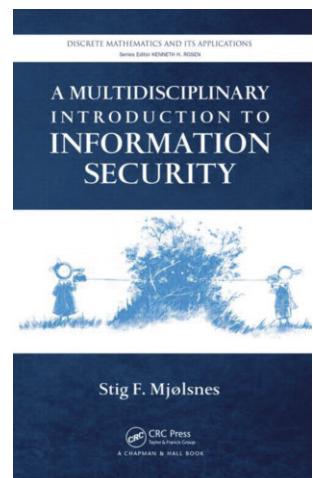
Constantine Stephanidis, ICS-FORTH  
 General Chair, HCI International Conference  
 E-mail: [cs@ics.forth.gr](mailto:cs@ics.forth.gr)

## Book review

Stig F. Mjølsnes

### A Multidisciplinary Introduction to Information Security

Computer and information security is a growing area of scientific and cultural interest. Several books have been published on the topic, including the 2011 publication, "A Multidisciplinary Introduction to Information Security" by Prof. Stig F. Mjølsnes. The book is derived from the significant experience of the author and his contributors who run a successful master-level course on the subject at the Norwegian University of Science and Technology. The main stated intention is to provide an interdisciplinary view of the broad and exciting area that is information security. The book keeps this promise.



The book's scope ranges from hardware security and cryptography to public key infrastructures and cryptographic protocol analysis. Interestingly, from these initial building blocks, the narration leads to more complex scenarios, dealing with quantum cryptography, addressing mobile device and software security, and reaching system level treatment as IT security certification, ICT forensic and risk management. The topics selected represent a truly interdisciplinary and holistic overview of the area.

The book is entertaining as well as informative. Whilst an experienced reader may find the introductory level too basic on certain topics, the wide range of topics ensures that every reader can learn something from this book. The contributors successfully manage to present intriguing aspects of the covered topics, always getting to the point promptly. This book is a page-turner, exceptional in its ability to capture and hold the reader's attention.

While the variety of topics and size of the book necessarily limit the depth of the analysis, the contributors successfully identify the key aspects of each topic and present them in an interesting way. They present them at an appropriate level of detail whilst retaining the necessary scientific rigor.

I highly recommend "A Multidisciplinary Introduction to Information Security" both for beginners and as a quick reference to the main areas of information security. The topics could also represent a nice syllabus for a master-level course. It is mandatory reading for anyone wanting to become quickly acquainted with the scope of the information security field.

2011 by Chapman and Hall/CRC Press - 348 Pages,  
 ISBN 9781420085907

Fabio Martinelli, IIT-CNR

## PLERCIM's Unit Awarded the Status of Leading National Research Centre in Poland

The Warsaw Centre of Mathematics and Computer Science comprises the Faculty of Mathematics, Informatics and Mechanics of the University of Warsaw (a member institution of PLERCIM) together with the Institute of Mathematics, Polish Academy of Sciences. This Centre has won the status of Leading National Research Centre in Mathematical Sciences (Krajowy Naukowy Ośrodek Wiodący, KNOW). In total there have been only three winning nominations (Mathematics, Physics, and Chemistry) in Science and three in Health Sciences, as the Polish Minister of Science and Higher Education, Barbara Kudrycka, announced on 12 July in the presence of the Polish Prime Minister Donald Tusk. The award comes with a substantial grant which will provide financing of the Centre for the next five years with an option of an extension for another five years. The money will be used to enhance the research potential of both participating institutions and to develop the Centre into a recognizable brand. The plans include vast extension of the graduate school aimed at the best students in this region of Europe. In addition, the grant will provide financing for a number of post-doctoral and visiting positions (both senior and junior levels).

Information regarding open positions and other activities of the Centre will be available at <http://www.wcmcs.edu.pl>

## PROMISE Retreat 2012 report on Prospects and Opportunities for Information Access Evaluation

The PROMISE network of excellence organized a two-day brainstorming workshop to discuss and envisage future directions and perspectives for the evaluation of information access and retrieval systems in multiple languages and multiple media.

Twenty-five researchers from ten European countries attended the event, covering many different research areas including information retrieval, information extraction, natural language processing, human-computer interaction, semantic technologies, information visualization and visual analytics and system architectures.

The ultimate goal of the PROMISE retreat is to stimulate and involve the research community along the identified research lines, gathering feedback and improving them, and to provide funding agencies with effective and scientifically sound ideas for coordinating and supporting information access research.

The report is available online at  
<http://www.promise-noe.eu/promise-retreat-report-2012/>



ERC president Helga Nowotny (left) and Lynda Hardman, CWI

## ERC President visits Turing Exhibition

On 20 September Prof. Helga Nowotny, President of the European Research Council (ERC), visited the Turing exhibition at CWI, which was designed for the Turing Year 2012. She was accompanied by Prof. Lynda Hardman from CWI and Prof. Marijk van der Wende, dean of the Amsterdam University College (AUC). The next day Prof. Nowotny gave the keynote speech during the opening of the new AUC building, discussing the relationship between teaching and research. The ERC President, who is a member of the International Advisory Board of AUC, is responsible for many European research funds. Picture: Helga Nowotny (left), and Lynda Hardman (right) in front of the LEGO Turing machine.

### Obituary

**Horst Santo** one the pioneers of Interactive Digital Broadcasting has sadly died in July 2012 at the age of 70.

I had the great honour and privilege of working with Horst for over 13 years. Horst Santo was truly a remarkable individual; a man of highest integrity who concealed his extensive and detailed knowledge of many subjects with an impeccable modesty.

After his graduation from the Technical University of Karlsruhe in 1975 Horst was employed at GMD (later Fraunhofer-Gesellschaft) in St. Augustin near Bonn a research centre which he served for 30 years till his retirement.

At GMD/FhG Horst Santo held various positions such as deputy head of Planning and Decision Support Systems of IPES, later FIT. This was later followed by his appointment as the head of successful R&D Interactive Television (iTV) group of Institute for Media Communication (IMK). During this time Horst Santo was also elected as a member of the Media Council of German state of North Rhine-Westphalia.

Horst's pioneering works in the areas of Internet streaming systems and interactive broadcasting technologies were numerously recognised and valued on a global scale. These included the implementation of the first European Internet "Radio & TV" systems and the development of the first ever interactive digital infotainment platform for ZDF (German second national TV channel).

Horst Santo was a man of great vision and tenacity, who always supported, motivated and encouraged any young researcher who would come in his vicinity. I like many of my colleagues who had the pleasure of knowing Horst owe him, a great deal.

Horst is greatly missed.

*Sepideh Chakaveh,  
Head of ERCIM Media Technology & Edutainment Working Group*

ERCIM – the European Research Consortium for Informatics and Mathematics is an organisation dedicated to the advancement of European research and development, in information technology and applied mathematics. Its member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry.



ERCIM is the European Host of the World Wide Web Consortium.



Austrian Association for Research in IT  
c/o Österreichische Computer Gesellschaft  
Wollzeile 1-3, A-1010 Wien, Austria  
<http://www.aarit.at/>



Consiglio Nazionale delle Ricerche, ISTI-CNR  
Area della Ricerca CNR di Pisa,  
Via G. Moruzzi 1, 56124 Pisa, Italy  
<http://www.isti.cnr.it/>



Czech Research Consortium  
for Informatics and Mathematics  
FI MU, Botanická 68a, CZ-602 00 Brno, Czech Republic  
<http://www.utia.cas.cz/CRCIM/home.html>



Centrum Wiskunde & Informatica

Centrum Wiskunde & Informatica  
Science Park 123,  
NL-1098 XG Amsterdam, The Netherlands  
<http://www.cwi.nl/>



Fonds National de la  
Recherche Luxembourg

Fonds National de la Recherche  
6, rue Antoine de Saint-Exupéry, B.P. 1777  
L-1017 Luxembourg-Kirchberg  
<http://www.fnr.lu/>



Fonds National de la  
Recherche Belgique



FWO  
Egmontstraat 5  
B-1000 Brussels, Belgium  
<http://www.fwo.be/>

FNRS  
rue d'Egmont 5  
B-1000 Brussels, Belgium  
<http://www.fnrs.be/>



Foundation for Research and Technology – Hellas  
Institute of Computer Science  
P.O. Box 1385, GR-71110 Heraklion, Crete, Greece  
<http://www.ics.forth.gr/>



**Fraunhofer**  
IUK-TECHNOLOGIE

Fraunhofer ICT Group  
Anna-Louisa-Karsch-Str. 2  
10178 Berlin, Germany  
<http://www.iuk.fraunhofer.de/>



Institut National de Recherche en Informatique  
et en Automatique  
B.P. 105, F-78153 Le Chesnay, France  
<http://www.inria.fr/>



Norwegian University of Science and Technology  
Faculty of Information Technology, Mathematics and  
Electrical Engineering, N 7491 Trondheim, Norway  
<http://www.ntnu.no/>



I.S.I. - Industrial Systems Institute  
Patras Science Park building  
Platani, PATRAS, Greece, 265 04  
<http://www.isi.gr>



Portuguese ERCIM Grouping  
c/o INESC Porto, Campus da FEUP,  
Rua Dr. Roberto Frias, n° 378,  
4200-465 Porto, Portugal



Polish Research Consortium for Informatics and Mathematics  
Wydział Matematyki, Informatyki i Mechaniki,  
Uniwersytetu Warszawskiego, ul. Banacha 2, 02-097 Warszawa, Poland  
<http://www.plercim.pl>



Science and Technology Facilities Council,  
Rutherford Appleton Laboratory  
Harwell Science and Innovation Campus  
Chilton, Didcot, Oxfordshire OX11 0QX, United Kingdom  
<http://www.scitech.ac.uk/>



Spanish Research Consortium  
for Informatics and Mathematics,  
D3301, Facultad de Informática, Universidad Politécnica de Madrid,  
Campus de Montegancedo s/n,  
28660 Boadilla del Monte, Madrid, Spain,  
<http://www.sparcim.es/>



Swedish Institute of Computer Science  
Box 1263,  
SE-164 29 Kista, Sweden  
<http://www.sics.se/>



Swiss Association for Research in Information Technology  
c/o Professor Abraham Bernstein, Ph.D., Department of  
Informatics, University of Zurich, Binzmühlestrasse 14,  
CH-8050 Zürich  
<http://www.sarit.ch>



Magyar Tudományos Akadémia  
Számítástechnikai és Automatizálási Kutató Intézet  
P.O. Box 63, H-1518 Budapest, Hungary  
<http://www.sztaki.hu/>



University of Cyprus  
P.O. Box 20537  
1678 Nicosia, Cyprus  
<http://www.cs.ucy.ac.cy/>



Technical Research Centre of Finland  
PO Box 1000  
FIN-02044 VTT, Finland  
<http://www.vtt.fi/>

## Order Form

If you wish to subscribe to ERCIM News  
**free of charge**

or if you know of a colleague who would like to  
receive regular copies of  
ERCIM News, please fill in this form and we  
will add you/them to the mailing list.

Send, fax or email this form to:

**ERCIM NEWS**

**2004 route des Lucioles**

**BP 93**

**F-06902 Sophia Antipolis Cedex**

**Fax: +33 4 9238 5011**

**E-mail: contact@ercim.eu**

I wish to subscribe to the

**printed edition**

**online edition (email required)**

Name:

Organisation/Company:

Address:

Postal Code:

City:

Country

E-mail:

Data from this form will be held on a computer database.  
By giving your email address, you allow ERCIM to send you email