Science & Technology Facilities Council

Technical Report

DL-TR-2010-008

Computing Grand Challenges

RJ Allan

November 2010

Chadwick Library
Science and Technology Facilities Council
Daresbury Laboratory
Daresbury Science and Innovation Campus
Warrington
WA4 4AD

Tel: +44(0)1925 603397
Fax: +44(0)1925 603779
email: library@dl.ac.uk

Science and Technology Facilities Council reports are available online at: http://epubs.stfc.ac.uk/

ISSN 1362-0207

Computing Grand Challenges

Rob Allan
Computational Science and Engineering Department, STFC Daresbury Laboratory, Daresbury, Warrington WA4 4AD

Contact e-Mail: robert.allan@stfc.ac.uk

October 14, 2010

Abstract

This report discusses the computing grand challenges which are of relevance to computer based research, particularly high performance modelling and simulation for engineering and the natural sciences. The report mainly draws on three UK sources: (i) The Computing Grand Challenges Initiative; (ii) Microsoft 2020 Science Group report; (iii) The Century of Information Research Strategy; plus (iv) The EU FP7 ICT Future and Emerging Technologies Programme; and (v) The DARPA Physical Intelligence programme.

This report is dedicated to the memory of Robin Milner who died in March 2010.

Contents

1 Introduction

This report discusses the computing grand challenges as identified by (mostly) computer science researchers. The report mainly draws on the following sources.

- The UK Computing Grand Challenges Initiative;
- Microsoft 2020 Science Group report;
- The Century of Information Research Strategy;
- EU FP7 ICT Future and Emerging Techologies Programme;
- The DARPA Physical Intelligence programme.

At the 2008 Computing Research Challenges Conference [9] Lesley Thompson of EPSRC noted: *The EPSRC has been looking at grand challenges. What is a grand challenge? It must be grand: a significant societal or economic problem that requires the application of research. It requires community effort, beyond the stakeholders. It should have the ability to enthuse society: this is hard in some areas of computing science, but a grand challenge offers more opportunity to do this. We quite like the idea of a bottom up approach. We would like to encourage other communities to self assemble around grand challenges in this way. We are going to try this approach in chemistry, where we've picked some champions to start things rolling. We also like the top down approach, and we will try this in nano-science and nano-engineering.*

In compiling this report we have tried to take descriptions of the computing grand challenges and extract information which is of relevance to those engaged in computer based research, particularly high performance modelling and simulation for engineering and the natural sciences. We refer to this as computational science.

2 The UK Computing Research Grand Challenges Initiative

The mission statement for the UK Computing Research Committee (UKCRC) includes: *The UKCRC aims to promote the vitality, quality and impact of computing research in the UK.*

In October 2002, under the auspices of UKCRC, Sir Tony Hoare and Robin Milner initiated discussions of "grand challenge" research projects in computing. Seven main proposals came out of those discussions, listed here (GC-1 to GC-7): `http://www.ukcrc.org.uk/grand_challenges/index.cfm`.

The Web site includes pointers to summaries of each proposal produced early in 2003 (so now seven years old), along with archives of discussion lists for each proposal, and an "overview" or "master" discussion list (GC-0). It is important to note that these were primarily scientific research projects aimed at increasing knowledge and understanding rather than practical engineering projects aiming to have immediate impact, although it was expected that advances in scientific understanding would inevitably lead to important new engineering advances (engineering in this sense means computer science).

Conferences were held in 2004 to discuss progress and make plans for promoting and extending the Grand Challenges Initiative. As a result of continuing discussions during 2004 a booklet edited by Tony Hoare and Robin Milner was published by the British Computer Society (BCS) summarising the grand challenge proposals. A short paper also summerised them [16]. At conferences in 2006 and 2008, progress was reviewed and additional challenges were added [9]. The 2010 conference was held on 16th April, at which a number of proposals were considered for new grand challenge areas including the following.

Exploiting division by zero;
Synthetic sensory motor systems for new ecological niches;
A community project to create a shared open information security knowledge base;
Privacy restrictions in social semantic web research;
Hatching a phoenix – re-claiming success from software failure;
Innovation everywhere – computing for 9 billion people;
Software engineering challenge – achieving zero carbon buildings by 2019;
Understanding the Quran;
Delivering the healthcare we deserve;
Modelling personalised value based healthcare;
Towards elastic sensor networks;
Supporting independent living for older people;
Assisted living – home care technologies;
Allowing blind people to see and deaf people to hear;
Digital signatures – "What are you" versus "Who you are";
Socially embodied technology;
Grass roots distributed healthcare;
The AI4FM approach for proof automation within formal methods.

The new proposals are still being discussed.

Each existing challenge moderator with his or her collaborators has set up a network of researchers, or is exploiting an existing network, together with a series of workshops. They aim to reach broad agreement on how best to promote high quality research in the topics of their challenge. The intended outcome in each case is to develop background documentation, feasibility studies and a roadmap of stategic directions.

The moderators of challenges GC2 and GC4 (involving different aspects of ubiquitous computing) agreed to combine them and to add a new theme on human experience – the resulting challenge GC2/4 has a new title.

A small committee for each challenge has been set up to carry it forward, together with an overall steering committee. Each committee has elected a chair person. These committees and chair persons are listed below.

GC0	Steering (Initial membership)	Dame Wendy Hall (Southampton, Chair) Alan Bundy (Edinburgh) Tom Rodden (Nottingham) Joe Sventek (Glasgow)
GC1	In Vivo – In Silico	Andrew Bangham (East Anglia, Chair) Luca Cardelli (Microsoft)
GC2/4	Ubiquitous Computing: Experience, Design and Science	Morris Sloman (Imperial, Chair) Dan Chalmers (Sussex) Jon Crowcroft (Cambridge) Marta Kwiatkowska (Oxford) Tom Rodden (Nottingham) Vladimiro Sassone (Sussex)
GC3	Memories for Life	Nigel Shadbolt (Southampton, Chair) Andrew Fitzgibbon (Oxford) Dame Wendy Hall (Southampton) Ian Horrocks (Manchester) Ehud Reiter (Aberdeen)
GC5	The Architecture of Brain and Mind	Murray Shanahan (Imperial, Chair) Mike Denham (Plymouth) Steve Furber (Manchester) Mark Lee (Aberystwyth) Aaron Sloman (Birmingham) Leslie Smith (Stirling)
GC6	Dependable Systems Evolution	Jim Woodcock (York, Chair) Cliff Jones (Newcastle) Peter O'Hearn (Queen Mary)
GC7	Journeys in Nonclassical Computation	Susan Stepney (York, Chair) Samson Abramsky (Oxford) Andrew Adamatzky (W. of England) Colin Johnson (Kent) Jon Timmis (Kent)
GC8 (New)	Learning for Life	Josie Taylor (OU, Chair) Anne Anderson (Glasgow) Tom Boyle (London Metropolitan) Mike Sharples (Nottingham)
GC9 (New)	Bringing the Past to Life for the Citizen	David Arnold (Brighton, Chair) David Duce (Oxford Brooks) Andrew Day (UEA) Alan Chalmers (Warwick) Phil Willis (Bath)

The following sections give a brief summary of each challenge of relevance to computational modelling and simulation in engineering and the natural sciences.

2.1 GC1 – iViS: *in Vivo, in Silico*

GC1 is a grand challenge for computational systems biology. iViS aims to realise fully detailed, accurate and predictive computer embodiments of plants, animals and uni-cellular organisms. The iViS challenge developed from a proposal to UKCRC by David Harel to model a complete multi-cellular organism [13], for example the nematode worm *C. elegans* and Ronan Sleep's proposal to attempt computational modelling of gastrulation.

iViS progress is in the form of a number of key demonstrators, culminating in whole life form models covering a wide range of phenomena. Intermediate demonstrators will cover a narrower range. Modelling the development of form during early growth is one example.

Initially the aims were restricted to simple and much studied life forms such as the nematode worm, the *Arabidopsis* weed, and single cell organisms such as streptomyces and baker's yeast: hence the sub-title "The Virtual Worm, Weed and Bug". Rather remarkably, human genes can replace about

35% of the worm's, so there is reason to hope that many of the complex processes involved in a worm developing both form and function from a single cell may help us understand much about the workings of higher life forms.

Potential impacts of iViS include an understanding of regeneration processes in plants and animals, with potentially dramatic implications for disease and accident victims. iViS may also lead to revolutionary ways of realising complex systems: instead of designing and programming in every detail, perhaps we can just grow them in a suitable medium. We know it's possible, because that's exactly what nature does.

iViS offers a powerful vision in which future life scientists can take virtual fly through tours of a plant, animal or colony of cells, studying what is happening at scales ranging from whole life form to what goes on inside an individual cell and stretching or shrinking the passage of time. Filters control what is seen by the observer, allowing concentration on specific aspects such as cell division, motility or chemical potential.

iViS thus offers an attractive way of browsing all we know about a life form. It may provide more – with sufficient effort, it might be possible to raise the faithfulness of the underlying model to the point where it becomes predictive as well as descriptive. If this happens, it will become possible to perform meaningful observations and experiments *in silico*. Specifically, an iViS model should faithfully exhibit the following phenomena: development from an initial fertilized cell to a full adult, cell function and interaction, motility and sensory behaviour, including possible interactions with other life forms. Virtual experiments, e.g. moving a virtual cell during development, should lead to the same outcomes as in real life.

In practice, it takes years of experience to gain familiarity with even one sort of biological data, let alone the many new forms emerging in systems biology. We urgently need integrating frameworks for organising the data. Making it all fit together into a coherent and useful picture presents a major challenge for biologists working in the life sciences.

This problem became so pressing that BBSRC established six Centres for Integrative Systems Biology in partnership with relevant universities (Imperial College London, Manchester, Newcastle, Edinburgh, Nottingham and Oxford). These Centres will need the vision, breadth of intellectual leadership and research resources to integrate traditionally separate disciplines in a programme of high quality international research in quantitative and predictive systems biology. iViS offers a challenging focus of attention for such centres.

Part of the answer to the data problem may lie in the way in which the World Wide Web is revolutionising our approach to knowledge organisation. Before the Web, the Internet was used mainly for e-mails and to transfer files of raw data. The Web revolutionised the potential of the Internet by allowing data owners to advertise the availability of their data together with information about how to display it in a browser. Today's search engines mean we can within seconds browse and access knowledge from a truly global repository. The Semantic Web movement will have an even more profound impact on this process of facilitating the conversion of data into information.

The Web is already a global source of information for scientists. Teams from laboratories that previously worked independently and communicated only via journals and the conference circuit now collaborate and share data via the Web, even swapping massive datasets to compare results. Scientists have begun to exploit the Web by establishing global virtual knowledge repositories that share data, theories and models. This is all part of the process now known as e-Research [1].

An example in the USA is Physiome `http://www.physiome.org`, which supports *the databasing of physiological, pharmacological, and pathological information on humans and other organisms and integration through computational modelling. "Models" include everything from diagrammatic schema, suggesting relationships among elements composing a system, to fully quantitative, computational models describing the behaviour of physiological systems and an organism's response to environmental change. Each mathematical model is an internally self consistent summary of available information, and thereby defines a "working hypothesis" about how a system operates. Predictions from such models are subject to test, with new results leading to new models.*

Such virtual knowledge repositories will help reduce the proliferation of models and theories for parts of the problem. A deeper challenge will be to make the models fit into a consistent larger picture. Sometimes this will be easy, for example when there is a simple input-output relationship between subsystems. More often however combining two models may show unexpected interactions inconsistent with *in vivo* data. There will be a process of development and refinement to obtain realistic behaviour.

The Semantic Web may help us to find new ways to give meaning to the data. There is now provision to enhance raw data with additional metadata. This can tell the recipient what the data means, how it is represented and organised, the way in which it was generated. Models, which often come in the form of a computer program, can be tagged with metadata. In this way it may become easier to combine data from different models and different sources.

Work has already begun to establish metadata dictionaries (ontologies) for the life sciences. There are already over 50 of them plus 20 related ontologies concerning the scientific processes used: see Open Bio-medical Ontologies `http://obo.sourceforge.net`. Whilst there is a drive to create bio-ontologies, there is not the same need to bring them together into a unified whole. The iViS challenge provides a possible driver, because the *in silico* modelling of a complete life form, will require working across all relevant ontology boundaries in the whole domain.

Even if we can build a simulation of a life form that successfully integrates all known data and models, it is not at all clear that genuinely interesting new predictions would emerge. The process may suffer from over fitting – namely, we can build a model consistent with known data, but it may merely fit the data and lack any predictive power. Avoiding such over fitting is a very strong reason to complement data driven modelling work on iViS with more abstract top down approaches.

There is also a considerable body of work underway to explore ways of organising life science data. One example is EMAP, the Edinburgh Mouse Atlas Project `http://genex.hgu.mrc.ac.uk`. *EMAP is a digital atlas of mouse embryonic development. It is based on the definitive books of mouse embryonic development, yet extends these studies by creating a series of interactive three dimensional computer models of mouse embryos at successive stages of development with defined anatomical domains linked to a stage by stage ontology of anatomical names.*

It can be expected that growing numbers of life science (and other) virtual knowledge centres will follow EMAP in adopting some form of spatio-temporal framework. The role of iViS is to expand this vision to a dynamic 3D working model, initially targeting much simpler life forms.

Even with an appropriate mix of approaches, it is not clear that iViS will ever realise its controversial predictive aims. Computer simulation may work for aeroplanes and bridges, but living systems may be too complex for them to work in the same way. There may be aspects of some biological domains which remain well outside the scope of even the most sophisticated computer modelling for many decades. There may however be some domains for which the iViS whole life form modelling approach works – developmental biology looks promising.

There are a number of research strands in the computing sciences needed to support the aspirations of iViS. This could be entitled "Computational Models and Scaleable Architectures for *in Silico* Life Sciences". Some strands will work bottom up, paying great attention to biological data. Other strands will work top down, studying minimal abstractions capable of generating the phenomena exhibited in vivo. Many will work "middle out", balancing the desire to be simple, elegant and general with the desire to be faithful to the data. Indeed this is a general feature of systems biology [19] and computational techniques such as agent based modelling.

Key to success could be the development of a new breed of computer languages for representing and manipulating biological data in a meaningful way and using it do drive a realistic, highly detailed, simulation which can be explored using advanced interfaces.

Groups of computer scientists are already exploring new languages, architectures and system design and analysis tools for the life sciences. See for example [6] for a good picture of such work. Cardelli and others are tackling the complexities of life science developing industrial quality models aimed at handling the masses of detail in a living system.

These efforts are complemented by ones looking to discover simple more abstract computational systems from which emerge various life like properties. Such abstract models are particularly useful when viewing some particular aspect of a plant, animal or cell. For example Prusinkiwicz *et al.* [11] have almost created a new art form for constructing good looking pictures of plant growth, using remarkably simple abstract models called L-systems. Of course these models may capture only a small part of the truth, but iViS may need such simplifying frameworks to structure the great mass of detail and present it to the public.

What about raw computer power and advances in haptic and other interface technologies? Solutions are already emerging from the IT industry in response to a variety of requirements. The critical problem is still to get the underlying computational models right before serious work can start on designing the framework and finding the best contemporary technologies to use in constructing a real iViS prototype.

A number of groups are already involved in developmental modelling, see the iViS Web site. The proposed, but speculative timeline, was as follows.

Within 5 years: early results, on e.g. developmental phenomena in plants and animals, and the first uni-cellular demonstrations.

Within 10 years: first prediction of a known result from an assembly of component models; models of Arabidopsis meristem growth; models of simple animal development.

Within 15 years: 100 years after the publication of D'Arcy Thompson's book *On Growth and Form* [12], first demonstration of iViS, e.g. for *Arabidopsis*.

Within 20 years: iViS models form the core of many Virtual Knowledge Centres for the life sciences.

2.2 GC2/4 – UbiComp: Ubiquitous Computing

There is a growing population of "effectively invisible" computers around us, embedded in the fabric of our homes, shops, vehicles, farms and some even in our bodies. They are invisible in that they are part of the environment and we can interact with them as we go about our normal activities. However they can range in size from large plasma displays on the walls of buildings to microchips implanted in animal or human bodies. They help us command, control, communicate, do business, travel, entertain, and are far more numerous than desktop systems.

By 2020 computers will be ubiquitous and globally connected, many in sensor networks. How many computers will there be in 2020? How will we communicate with them? What data will they collect and what else will they do? Shall we be able to manage such large scale systems, or even understand them? How will people interact with them and how will this new pervasive technology affect society? How can people without computing skills configure and control them? What tools are needed for design and analysis of such constantly adapting and evolving systems? What theories will help us to understand their collective behaviour? These are the sort of issues which make ubiquitous computing a grand challenge.

The Ubiquitous Computing Grand Challenge (UbiComp) was formed by merging two of the original grand challenges GC2 *Science for Global Ubiquitous Computing* which focused on theory and GC4 *Scalable Ubiquitous Computing Systems* which focused on engineering aspects. UbiComp is formulating a research manifesto which postulates the need for combining theory, engineering and social issues related to building ubiquitous systems. So far, most research in the UK and elsewhere has focussed on the engineering with very little attention given to the theory required to underpin the design and analysis of ubiquitous systems which are intrinsically large scale and complex. Some of the work in the EQUATOR IRC project addressed social aspects and how people interact with ubiquitous systems, see `http://www.equator.ac.uk`.

UbiComp is currently led by University of Nottingham `http://www.nottingham.ac.uk/bridging/ubicomp/`. Its activities are promoted by the EPSRC funded UK-UbiNet Network grant, see `http://www-dse.doc.ic.ac.uk/Projects/UbiNet`.

It is expected that components of future distributed systems will be far more sophisticated than current mobile phones, but they will always have limited storage, processing, display capabilities and battery power compared to fixed PCs. There is thus a need to adapt information and applications so that they are compatible with such limited capabilities but also to provide information or adapt services that are relevant to the context of the user. Sensors in the environment, possibly in conjunction with personal devices, could discover a user's actual activity. The context aware ubiquitous computing environment could then react to the user's requirements, e.g. by adjusting controllable aspects of their environment to suit their needs.

There is clearly considerable interest in ubiquitous computing and research projects worldwide. Two journals have recently started covering the area of pervasive computing (IEEE Pervasive Computing and Springer Personal and Ubiquitous Computing) and there are others which cover mobility. There are also a number of conferences addressing issues of mobility, wearable computers, wireless communications, etc. There are a few isolated UK projects in universities and two relevant large collaborative projects: Mobile VCE `http://www.mobilevce.com/index2.htm` and the Equator IRC already mentioned. The DTI has recently started a new funding initiative called "Next Wave Technologies" which is meant to foster collaborative projects involving academic and industrial researchers `http://www.nextwave.org.uk`. There have also been past projects related to Virtual Society funded by ESRC. Many of the research issues relating to ubiquitous computing and the socio-economic impact are discussed in a report from an EPSRC working group [2]. There is a growing need to coordinate a large and inclusive academic community across the UK to raise awareness, coordinate research initiatives and to ensure that a healthy community of researchers can be sustained.

There are at present three largely separate academic communitiesresearching into: mobile computing in computing departments; wireless mobile communications in electronic engineering departments; and human factors and social science in various departments. These communities have comparatively little interaction and it would be useful to foster a common understanding across the groups. There is also increasing interest within computing departments on investigating the theoretical and software engineering issues relevant to ubiquitous computing.

For this reason, the UK-UbiNet network brings together academic researchers within the UK to build a new multi-disciplinary community covering the broad aspects of ubiquitous computing, including the technology, human factors and socio-economic issues. Although the emphasis is on a network of academic researchers, they welcome involvement from industry and will work closely with the DTI Next Wave Technologies initiative to involve industry and to have joint workshops. UK-Ubinet seeks to address the following key objectives and research issues.

- To establish and support an inclusive multi-disciplinary community which will collaborate on ubiquitous computing research and foster interaction with similar communities in Europe and the USA.
- To provide a forum for the exchange of ideas, best practice techniques and software related to Ubiquitous Computing.
- To hold workshops, seminars and summer schools which UK research students and RAs can easily attend as a means of training young researchers and future leaders of the community.
- To publish a manifesto specifying the research challenges which the community should address over the next 15 years.
- To provide a focus for cooperation with EU initiatives such as the Disappearing Computing Program and other EU Networks.
- To formulate and submit collaborative proposals to UK and EU funding sources.

Although there has been considerable progress in some areas such as the wireless technology to support mobile communications, there are many research issues which need to be addressed in order to realise a ubiquitous intelligent environment which is able to support people in their normal activities. UK-UbiNet has initiated discussions and collaborations to address issues such as the following.

- Use of mobile and public networks requires integration and management of disparate communication technology to support seamless and ubiquitous connectivity with quality of service guarantees.

- Issues of trust, security and privacy become even more problematic in ubiquitous computing environments which are able to track your movement and activities at all times. Portable devices will combine the functions of mobile phone, keys, credit cards, passport and medical record so mechanisms to cater for loss or theft are needed.
- The provision of context aware applications and services is key to successful ubiquitous computing. For example, the system should be able to distinguish between a possible heart problem or exertion due to running for a bus and to provide the information or services relevant to your current activity.
- New approaches are needed to provide flexible and adaptable software both for mobile devices and the intelligent environment. The scale of these ubiquitous systems necessitates autonomic (self organising and self managing) systems which can dynamically update software to cater for new services and applications.
- Human factors arise to cater for new modes of interaction and invisible ambient technology which must be useable by non-technical people. There is a need to support transient organisations and dynamic, potentially mobile work arrangements including virtual teams, *ad hoc* collaborations and virtual organisations.
- Human, social and organisational issues that arise in creating new forms of interaction based on ambient technologies and deploying those systems within everyday environments either in the home, the workplace, or more public arenas such as shops, restaurants and galleries. There is a need to support a diverse range of activities, within and across different settings and organisational environments, performed by users with different interests, skills and abilities.

These issues also raise questions concerning new approaches to requirements analysis, design, development and deployment of ubiquitous computing. Other aspects of using ubiquitous computing systems are mentioned in Section 5.1.

2.3 GC3 – M4L: Memories for Life

Memories for Life is a project funded by the EPSRC bringing together a diverse range of academics in a bid to understand how memory works and to develop technologies to enhance it, see `http://www.memoriesforlife.org`.

In today's technology rich society, human memory is supplemented by an increasing amount of personal digital information; e-mails, photographs, Internet, telephone calls, even GPS locations and television viewing logs. In research situations, published journal articles and books are complemented by Web sites, Wiki entries, Blogs and electronic notebooks. This is often referred to as "grey literature" by information specialists, but still needs to be safeguarded and managed. Bringing together psychologists, neuroscientists, sociologists and computer scientists could lead to a more effective ways to use and manage both human and computer aided memory.

The challenges that lie ahead in this field include the development of prosthetic memories, the storing and retrieval of a lifetime's worth of memories and the issues of trust and privacy that such handling such date will require. GC3 aims to produce an understanding of what is common in memory systems and use that knowledge to improve efficiency, recall and information management across human, personal, social and work domains. These challenges are of international scope, beyond what can be achieved by a single research team or grant, but offers the possibility of revolutionary advance.

M4L has identified challenges in the following areas [24]: memory representation; types of memory; memory content; content organisation; information storage, access and retrieval.

We note that there are a number of other projects tackling related issues, for instance portfolio management for lifelong learning. In e-Research the JISC-funded CREW project, Collaborative Research Events on the Web, is concerned with recording, annotating and managing access to information about research events such as workshops, conferences and meetings. This includes recording activities leading to decision making, see `http://www.crew-vre.net`.

2.4 GC5 – Architecture of Brain and Mind

GC5 is concerned with a multi-disciplinary attempt to understand and model natural intelligence at various levels of abstraction and demonstrating results of our improved understanding in a succession of increasingly sophisticated working robots. The Web site is currently hosted at University of Stirling `http://www.cs.stir.ac.uk/gc5/`.

This project brings together work in neuroscience, cognitive science, various areas of AI, linguistics, philosophy, biology and other relevant disciplines. It aims to produce a new integrated theory of how a single functioning system can combine many human capabilities, including various kinds and levels of perception, different kinds of reasoning, planning, problem solving, curiosity, many varieties of learning (including grasping new abstract concepts and developing new skills), many kinds of action of varying complexity, different uses of language, varieties of affect including motivation and emotions, social interaction and various forms of creativity.

Current robots can perform many tasks but each one is capable of only a very limited range of behaviours. Usually they do not combine perceptual and manipulative skills with the ability to communicate and co-operate, and they do not know what they are doing, why they are doing it, what difference it would make if they did things in a different way, etc., and they cannot give help or advice to another robot or a person performing such tasks. They do not have the variety of competences, the integration, or the self understanding of a 4-5 year old child (or even a much younger child that can move about manipulate things, and interact with other people, not necessarily using language) and current robots can barely learn anything a child can learn.

The summary of GC5 from 2004 listed some of the gaps in our knowledge. Many processes in the brain are not yet understood, including how we: see many kinds of things around us; understand language; learn new concepts; decide what to do; control our actions; remember things; enjoy or dislike things; become aware of our thoughts and emotions; learn about and take account of the mental states of others; appreciate music and jokes; sense the passage of time. These all involve both abstract mental processes and concrete physical processes. The project aims to explain how both kinds of processes work and to demonstrate this in robots with a large collection of human like capabilities, unlike current robots which are very limited.

Robots produced within GC5 should at least have an interesting subset of the capabilities of a child aged somewhere between 2-5, including the ability to go on learning and the ability at least some of the time to understand what they are doing and why. One way for such a robot to demonstrate that functionality would be being capable of helping a disabled person who wishes to avoid being dependent on other humans without the robot first having to be programmed explicitly with knowledge about the environment and the person's needs and preferences.

The aim is not merely to understand how such diverse functions can be integrated in a single system at a high level of abstraction which might be modelled on computers or future information processing machines, but also to explain how they can be implemented in actual biological mechanisms. The project will continue developing our understanding of brain mechanisms (e.g. chemical, neural, etc.) including showing how they are able to support the high level functionality required by a child or robot. For this purpose, natural minds can be viewed as virtual machines implemented in brains. Since human minds surpass artificial minds in many ways at present, we may discover that this is partly due to using a different kind of physical implementation from current computers. There could be other reasons – it may be that our current designs for AI systems are too simple because we have not yet identified the required functionality or kinds of architectures, forms of representation and what algorithms can provide those kinds in an integrated system.

GC5 is carrying out several mutually informative tasks concurrently.

Task 1: Bottom up specification, design and construction of a succession of computational models of brain function, at various levels of abstraction, designed to support as many as possible of the higher level functions identified in other tasks.

Task 2: Codification and analysis, partly from a software engineering viewpoint, of many typical and widely shared human capabilities, for instance those shared by young children, including perceptual, motor, communicative, emotional and learning capabilities, and using them: (a) to specify a succession of increasingly ambitious design goals for a fully functioning (partially) human like system; (b) to generate questions for researchers studying humans and other animals which may generate new empirical research leading to new design goals.

Task 3: Top down development of a new theory of the kinds of architectures capable of combining all the many information processing mechanisms operating at different levels of abstraction and testing the theory by designing and implementing a succession of increasingly sophisticated working models, adding more detail at each version.

The EU has recently made available funding for research in cognition, both natural and artificial.

2.5 GC6 – Dependable Systems Evolution

In today's world, the cost of malfunctioning software is enormous. The US Department of Commerce in 2002 estimated that the cost to the US economy of avoidable software errors is between 20 and 60 billion dollars every year. Over half the cost is incurred by the users.

The Verification Grand Challenge (VGC) [23], is an ambitious, international, long term research programme that seeks to address this situation in the future by creating a substantial and useful body of code that has been verified to the highest standards of rigour and accuracy. It has a 15 year perspective, hoping in that time to achieve three major objectives as follows.

- Establish a unified framework within which different theories of program construction and verification can co-exist and be used productively;
- Build an integrated suite of tools to support all aspects of verified software construction: requirements capture, specification, validation test case generation, refinement, analysis, verification, and run time checking;
- Populate a repository of formally specified and verified codes, that can not only serve as exemplars to convince others of the utility and practicability of formal techniques, but that also are seen as being useful in practice and, ideally, that are taken up and used in anger.

Software verification is part of a branch of computer science known as Formal Methods. Formal methods are intended to explain software, both to the user and the developer. They can be used to produce precise documentation, structured and presented at an appropriate level of abstraction. By being amenable to mathematical analysis, formal methods are also intended to be used to predict software behaviour.

Given the right computer based tools, the use of formal methods could become widespread and transform software engineering. Representatives of the computer science community met in Zürich in 2005 to discuss the international grand challenge on verification `http://vstte.ethz.ch/report.html`. They committed to making verified software a reality within the next 15 to 20 years.

Tony Hoare initially suggested the Verification Challenge [14] as an effort to create a toolset that would, as automatically as possible, guarantee that programs meet given specifications. The toolset should be applicable to a wide range of software and, ideally, to itself. Hoare conceived the challenge in a broad sense, which would include testing and resolving the problems of getting the specifications right.

Although we can never bridge the gap between the formal and the informal, having a fully fledged "correctness of construction" system would allow feedback to flow from program to design, so that we could precisely track unwanted behaviours to a suitably formalised design.

2.6 GC7 – Journeys in non-classical Computation

The GC7 challenge is to produce a fully mature science of all forms of computation, that unifies the classical and non-classical paradigms. The non-classical paradigms include computation under non-classical laws of physics (quantum theory, general relativity), biological computing, embodied computing, *in materio* computing, and more. Details of the challenge and some suggested research directions can be found in [20, 21].

One of the original areas discussed in the context of GC7 is concurrency, where it was argued that classical computer science had taken a wrong turn by focussing on sequential machines and how concurrency, when approached correctly, should be the natural way to describe, model and implement systems. Not surprisingly, the case was argued for process algebra based languages, specifically, CSP [15] and Pi-Calculus [18], see also `http://en.wikipedia.org/wiki/Pi-calculus`.

Work has demonstrated that extremely efficient massively parallel simulations are readily derivable from process algebra descriptions. The work is continuing in the EPSRC supported CoSMoS project, "Complex Systems Modelling and Simulation Infrastructure", which is a case study driven approach to developing a method for modelling complex systems and deriving appropriate highly parallel process oriented simulations. Case studies include modelling of immunological sub-systems, plant ecologies and liquid crystals.

One of the major features of biological systems is their massive parallelism on multiple scales. Cells contain millions of molecules, organisms contain trillions of cells, ecosystems contain billions of organisms, all acting and interacting on different length and time scales. A parallel Turing model has no more intrinsic computational power than a sequential one, since a parallel machine can be simulated by a sequential one. However, this argument applies only under the assumptions of the Turing model, which does not include considerations of real time performance, naturalness of description, or fault tolerance and redundancy, among others.

The massive parallelism in the CoSMoS and TUNA projects is achieved by time slicing tens to hundreds of thousands of processes on a single processor and distributing over several tens of processors in a cluster. The implementation language allows this distribution to be achieved relatively transparently. True concurrency on a single processor can be achieved on FPGAs (Field Programmable Gate Arrays), programmable hardware that allows different parts of a single chip to be executing different parts of a program at the same time. FPGAs can be linked together in large numbers (for example, the BioWall), allowing massive parallelism to be achieved relatively cheaply. The CSP based language Handel-C has been used to program multiple FPGAs in a process oriented style.

Such classical parallelism is relatively straightforward technologically, it merely requires a change in mindset to move from a "sequential first" to a "concurrent first" style. More challenging is the parallelism provided by many forms of analogue computing. These can provide a more immediate massive parallelism, implemented directly by the underlying physical properties of the substrate being exploited. There is a wide range of problem specific analogue computers available, but there are also general purpose analogue computers. It is clear however, that there is still challenging work to be done to provide the level of support for programming such devices as we currently have in the classical realm. What are suitable models for programming by analogy? How expressive can these models be?

Other areas of investigation in GC7 currently include: hyper-computation; interactive computing; bio-inspired computing; embodiment and *in materio* computing; growth and self assembly.

2.7 GC8 – Learning for Life

With the emergence of mobile and ubiquitous computing, the semantic Web and the development of an e-Research infrastructure, new possibilities open up for e-learning and learning for life that take us beyond what has been conceived in this area before. Indeed, to a large extent, the challenge here is to even conceptualise how learning environments and opportunities will become manifest, how people will engage with learning events and what learning for life will be like. These new possibilities need to be understood in the context of our developing understanding of the co-evolutionary nature of learning and computing systems, so we ensure that the full potential of learning for life is realised. The future will involve much more than incrementally extending current models of teaching and learning – rather something of the order of a paradigm shift is required to halt the accretion of tools and environments that do little more than replicate face to face assumptions about teaching and learning, and which do very little to recognise the nature of learning for life.

In order to do this we need to promote a strong inter-disciplinary research agenda that brings together a broad range of previously separate domains. This has already been recognised in the USA, with the establishment of dedicated e-learning research centres and NSF support for the *A tutor for every learner* initiative. Similarly, the UK needs to develop and promote a dedicated initiative in learning for life that exploits its significant capacity to tackle the emerging relevant research issues. GC8 came together as a result of a nationwide consultation on research in this area, supported by a workshop attended by 50 contributors. It represents an area of work to which a large number of researchers are committed.

Research into e-learning is inherently multi-disciplinary, requiring partnerships between those who develop technology and a broad range of social science researchers who seek to understand the nature of learning, factors in educational attainment and the interaction and organisational effects and impact of technology. This combines perspectives, methods and theories from technical domains such as computer science, technology, artificial intelligence; design disciplines such as design and HCI; the learning sciences such as educational technology, psychology, education; and the disciplines studying communication, communities and discourse e.g. social sciences, linguistics. Establishing and maintaining these multi-disciplinary research teams is essential for successful e-learning research, particularly given the longitudinal nature of the research involved and requires us to move forward in a coordinated manner in order to build effective e-learning environments in the future.

The development of new forms of e-learning environment and the effective use of new e-learning tools and facilities requires us to consider a variety of distinct research challenges, including four main themes in particular.

Modelling and dynamic evaluation: how do we best model and represent learners within e-learning facilities and how might we assess these over time?

Informal and lifelong learning: how might we develop facilities that support learning outside formal educational settings over a learner's lifetime?

Creativity and problem solving: how might we encourage and support creativity and problem solving with and through e-learning facilities?

Inclusion and accessibility: how do we ensure that learning for life is a viable option for all and that the facilities provided reflect the diversity of learners?

Underpinning these themes are two multidisciplinary challenges in e-learning.

Technological challenges: how do we exploit the potential that emerges from new technological advances to best support learning for life?

Economic, social and cultural challenges for learning for life: how do we ensure that the benefits to emerge from e-learning are exploited to their best potential and how do we best understand and manage the social and cultural impact of e-learning research and practice?

GC8 focuses on the research questions to emerge from these themes and challenges that will require significant progress to be made in terms of the technologies, theories and methods of e-learning.

2.8 GC9 – Bringing the Past to Life

The past is all around us. We live our lives, whether consciously or not, against a rich backdrop formed by historic buildings, landscapes and other physical artefacts from the past. The historic environment is however more than just a matter of material remains. It is central to how we see ourselves and to our identity as individuals, communities and as a nation. It is a physical record of what our country is, how it came to be, its successes and failures. It is a collective memory, containing an infinity of stories, some ancient, some recent: stories written in stone, brick, wood, glass, steel; stories inscribed in the field patterns, hedgerows, designed landscapes and other features of the countryside.

Much of the evidence that we have of the past relates to contentious material, most often because of conflict and religion, about which there are inevitably differing perspectives, affecting perception of the events themselves (and the participants in them) as well as the modern day observers with their own current ethical, philosophical and social contexts. Cultural heritage professionals are often, rightly, loathe to settle on particular interpretations of the significance of events or artefacts, preferring to present a range of interpretations. They are also suspicious of the media trivialising and misrepresenting the past in the interest of a more entertaining and profitable re-interpretation. The emphasis is therefore often on preservation and custodianship rather than research.

Computer scientists have become widely involved in attempting to assist cultural heritage professionals in their tasks but, as with any data management, the computing professionals look for additional value to be obtained from the existence of digital records. At times these additional uses have appeared to cultural heritage professionals as crass and insensitive, failing to address the real requirements of their disciplines and spreading more confusion than understanding.

The long term vision of GC9 is that the citizen should be able to witness events of the past replayed interactively. This is more than mere recreation, it would allow the viewer to explore and discover more about the circumstances and motivations of the participants, linking the reconstruction to the modern day evidence if they choose and receiving explanations of the differing socio-political perspectives which are relevant to the events.

This is a truly multi- and inter-disciplinary challenge. There are many intermediate steps to achieving the long term vision and many technological challenges to meet on route, touching on a widely spread set of computing sub-areas and other disciplines. At the extreme the challenge re-activates the Turing test.

Computing science has already much to offer to the many interim applications with stages of discovery, recording, analysis, cataloguing, reconstruction, interpretation, story telling and communication of physical artefacts and records of the past. Currently these offerings are somewhat fragmented with a huge range of intermediate formats, many (mostly local) formats for classifying artefacts and cataloguing collections and little by way of interchange formats ensuring the persistence of the information specifically targeted at the preservation of the cultural heritage content. Instead there are common formats for general geometric information or GIS content or database and archive structures, but little specifically targeted at the preservation and re-use of the cultural heritage content. There is also a huge body of knowledge already archived in incompatible formats and often where the original data collection cannot be repeated since the original sources have been lost or destroyed, whether by acts of war, terrorism or simply the ravages of time or normal processes of archaeological discovery.

Breaking down some of the component challenges that the complete vision would need to address, the following seem key areas.

- An integrated infrastructure from data capture to deployment in cultural heritage research and scholarship. The main challenges are the definitions of data formats to allow interoperability of tools and long term applicability of the base data.
- Digitising and preserving existing collections, for example, the challenge of digitising and preserving the estimated 100M hours of audio-visual material from the 20th century is a significant challenge to production automation and deployment, quite probably involving research into viable automation and preservation techniques.
- Intelligent interactive tools for use by non-IT professionals, which are tailored, so that the cultural heritage professionals can work in their domain of expertise rather than fighting to achieve particular effects using general purpose tools.
- Modelling and visualisation systems which differentiate interpretation and evidence supported fact, so that viewers are not mislead by prettified presentations into mis-conceptions of cultural heritage knowledge. Enormous progress has been made in some areas of recording, e.g. laser scanning of sites and artefacts, but there are very significant challenges in analysing such data as the basis for interpreting the original state of the sites or objects. Visualisation also involves reconstruction of the environment under which the originals would have been viewed so that the context is established and preserved.
- Management of very large data sets, for example viewing whole cities in detail including not only models of the original architectural detail but also the population at different periods in time.
- Algorithms, data structures and systems for efficient visualisation of very large, animated, and detailed multi-media datasets are therefore required. Whilst this challenge is shared with other potential applications, one aspect that will improve the results is an understanding of the common characteristics of historic artefacts.
- Natural language technologies for interpreting historical accounts and related events in a historically meaningful context.

Some of the technological challenges have arisen from the requirement that the credibility of the objectives of GC9 is clearly demonstrated to cultural heritage professionals.

3 Towards 2020 Science

The 2020 Science Group met for a three day workshop in Venice, July 2005. This group comprised of some 30 distinguished scientists from fields including biology, physics, chemistry, biochemistry, astronomy, genetics, medicine, mathematics and computer science from 12 different countries. The group, led by Stephen Emmott and Stuart Rison of the Microsoft Research Centre Cambridge, focused largely, although not exclusively, on the natural sciences rather than the physical sciences, engineering or social sciences. In particular, their report focuses on the biological sciences broadly defined, from molecular biology to systems biology to organisms biology and ecosystems science. It debates the role and future of science over the 14 year period towards 2020, in particular the importance and impact of computing and computer science on applied science.

The workshop and report were sponsored by Microsoft. The report is organised as follows.

3.1 Laying the Ground

Computing has rapidly established itself as essential and important to many branches of science, to the point where "computational science" is a commonly used term. Indeed, the application and importance of computing is set to grow dramatically across almost all the sciences towards 2020. Computing has started to change how science is done, enabling new scientific advances through enabling new kinds of experiments. These experiments are also generating new kinds of data – of increasing complexity and volume. Achieving the goal of being able to use, exploit and share this data most effectively is a huge challenge. The group considered trends and developments already under way in computing and computer science and additional requirements needed to achieve this aim. These will lay the ground for a far more fundamental impact on science.

This section covers issues of: computational science; semantics of data; intelligent interaction and information discovery; transforming scientific communications; and computational thinking.

3.2 Building Blocks of a Scientific Revolution

Concepts, theorems and tools from computer science are now being developed into new conceptual tools and technological tools of potentially profound importance, with wide ranging applications outside the subject in which they originated, especially in sciences investigating complex systems, most notably in biology and chemistry. The group believed these tools have the potential to have a fundamentally radical impact in science and especially in the biological sciences. They suggested such tools will become integrated into the fabric of science and are the potential starting point for fundamental new developments in biology, bio-technology and medicine, as well as other branches of science towards 2020.

This section covers issues of: the fundamental role of computer science concepts in science; integrating theory, experiments and models; from complexity to coherence; new conceptual and technological tools; codification of biology; prediction machines; artificial scientists; molecular machines; new software models for new kinds of tools; and new kinds of communities.

The conclusion to this section notes: *it is clear that the computer science community and the science community need to work together far more closely to successfully build usable, robust, reliable and scalable tools for doing science. This is already happening in some areas and in some countries, but the scale of the integration required is not going to happen by accident. It will require a dedicated effort by scientists, numerous government initiatives to foster and enable such an integration, and the co-involvement of commercial companies such as Wolfram, MathWorks, IBM, Apple and Microsoft. It is our recommendation that government science agencies take the initiative and introduce schemes and initiatives that enable far greater co-operation and community building between all these elements.*

3.3 Towards Solving Global Challenges

The 21st Century is already starting to present some of the most important questions, challenges and opportunities in human history. Some have solutions in scientific advances, e.g. health, while others require political or economic solutions, e.g. poverty. Some require significant scientific advances in order to provide the evidence necessary to make fundamental political and economic decisions, e.g. our environment. Addressing any of them is non-trivial. The report outlines some of the primary global challenges for the first half of this century that can be foreseen now that science can help to address and how the advances in computing and computer science can accelerate advances in science in order to enable science to do so.

This section identifies grand challenges of: Earth's life support systems; understanding biology; global epidemics; revolutionising medicine; understanding the universe; origin of life; future energy; and building blocks of a computing revolution.

4 The Century of Information Research Strategy

A strategy document arose from the e-Science Directors' Forum in 2008 in response to a speach from then Prime Minister Gordon Brown which included the statement *this is the Century of Information* [1].

The ensuing discussion was led by e-Science Envoy, Malcolm Atkinson [3]. Through a year long consultation with UK researchers, a coherent strategy was developed and presented in the paper which notes: *In the Century of Information, there is rapid change in both the research and digital-economy worlds. More data will be created in the next five years than through human kind's entire history. The challenge for society is to be able to search for and find specific datasets, to combine the datasets, to undertake new types of computationally intensive analysis, to visualise the output, and thereby to extract significant information and knowledge. This is transforming not only how researchers from science, arts, humanities and social science do their work, but also what they will discover, with whom they will collaborate, how they will share work, how they will report their findings, and what know-how they will require. These changes are already reflected in business, with a growing awareness of the opportunities and risks associated with our ability, or lack thereof, to process increasing amounts of disparate data.*

This statement is endorsed by Bell *et al.* [4] who note that *As simulations and experiments yield ever more data, a fourth paradigm is emerging, consisting of the techniques and technologies needed to perform data intensive science. ... The demands of data intensive science represent a challenge for diverse scientific communities.*

A selection of examples of the anticipated digital infrastructure needs by 2020 were illustrated as follows.

- community support for advanced work dealing with more complex data and combining higher precision models, with increased collaboration;

[1] G. Brown, Speech on Liberty, University of Westminster, 2007. `http://www.number10.gov.uk/Page13630`

- research dependant on sustained, easily used and well curated digital data resources that have common access policies, providing a balance between open access to encourage scrutiny and wide investigations, and constraints that protect emerging results, privacy and ethical standards;
- one million users in the UK, using between them 10 petaflop/s of computing power, which will have to be provided using a minimum carbon emission strategy;
- professional decision makers and researchers sharing a rich ecosystem of evolving tools and services that compete and cooperate to provide the analyses, modelling power, information and advice that users seek;
- each student developing skills and judgement in using these systems with universities producing graduates who have well developed strategies for exploiting them;
- all popular information manipulating tools, such as browsers, spreadsheets, image processors, bio-informatics systems, statistical packages, design tools and mathematics packages drawing on the computation and information services in ways that are virtually invisible to the user;
- new community behaviours emerging across both public and private sector groups, people collaborating and competing using the new skills, capabilities, tools, data and computing power to build the information collections, models and services that address their group's vision and needs;
- real time control of experiments; reaching femto-second observational and control cycles that can only be managed by direct coupling with computational models - ITER is a prime example;
- multi-core computers, wearable computers and disposable computers will be prevalent, pervasive and inter-connected by a rich hierarchy of wireless, electrical and optical digital communication networks.

The challenges were identified as follows.

Ease of use – a pre-requisite for expanding and accelerating the take up of computational and data intensive methods. As more researchers from more disciplines try to use diverse digital resources in combination they may be thwarted by apparently arbitrary policy and interface variations that may already be accepted by the original users. As researchers gain experience of well supported Web based services, such as Google, Wikipedia, YouTube, Flikr and FaceBook, many of their expectations for ease of use and interfaces will rise.

Sustainability is crucial to adoption and effective use, but not every service can be sustained indefinitely; hence a long term plan is needed encompassing provision, funding and withdrawal of superseded services and those that are no longer cost effective.

Capacity balanced across the infrastructure is essential for fluent use by a growing community of researchers.

Evolution of the services, facilities, tools and knowledge of the support teams must keep pace with the advances in methods so that the majority of researchers can meet their needs with modest effort proportional to the complexity of their goals.

Efficiency, cost-effectiveness and green computing have to be demonstrably addressed without compelling researchers and institutions to use the pooled or common provision.

Support for extreme e-Science must empower those gifted researchers who choose to pioneer new ways of building and composing models, of exploring and analysing data, of working collaboratively, of pushing the scale and complexity boundaries. It is essential for seeding the next rounds of innovation.

5 EU FP7 ICT Future and Emerging Technologies Programme

The FET Programme in Europe was developed to promote "high risk" research which may have high technological or societal impact. It is an incubator for new ideas and themes for long term research. There is a combination of exploratory open research and more strategic themes. See `http://cordis.europa.eu/fp7/ict/programme/fet_en.html`.

A preparatory report [10] pointed to examples of previous projects which in many cases are large scale ICT infrastructures. The list is shown in the table.

USA	**Elsewhere**
Assembling the Tree of Life	Blue Brain Project
CaBIG: Cancer Biomedical Informatics Grid	EU-US RFID Lighthouse pilot projects
CBOL: Consortium for the bar coding of Life	GBIF: Global Bio-diversity Information Facility
DARPA Autonomous Vehicle Grand Challenge	Japanese Earth Simulator
Deep Thought/ Deep Blue	CERN Large Hadron Collider
Earthscope	OMII-UK
GEON: Geoscience Network	Super Kamiokande: Kamioka Nucleon Decay Experiments
Hubble Space Telescope	Japanese 5th Generation Computing Initiative
Human Genome Project	Virtual Physiological Human Network of Excellence
Human Micro-biome Project	PRAGMA: Pacific Rim Applications and Grid Middleware Assembly
iPlant	**USA contd.**
LTER: Long Term Ecological Network	Man on the Moon Challenge (Apollo Programme)
NASA Centennial Challenges programme	National Ignition Facility (construction phase)
NNI: National Nano-technology Initiative	NCEAS: National Center for Ecological Analysis and Synthesis
NIGMS Protein Structure Initiative	Sematech: Semiconductor Manufacturing Consortium
Strategic Computing Initiative	SDI: Strategic Defense Initiative (Star Wars)
Super-conducting Super Collider	TeraGrid
ASCI: Accelerated Strategic Computing Initiative	War on Cancer
X Prize Foundation	

The FET Programme is complementary to the larger FP7 ICT programme. The latter lists seven challenges as follows.

1. Pervasive and trustworthy network and service infrastructures;
2. Cognitive systems, interaction, robotics;
3. Components, systems, engineering;
4. Digital libraries and content;
5. Sustainable and personalised healthcare;
6. Mobility, environmental sustainability and energy efficiency;
7. Independent living, inclusion and governance.

Within the FET Programme, a call was made in July 2010 for flagship proposals with a final deadline of 2/12/2010. On 1/10/2010 the satements of interest were published, see `http://cordis.europa.eu/fp7/ict/programme/fet/flagship/`. Up to 10M EUR is available with the possibility that two or three projects will be funded. The list of submitted projects is as follows.

1. Autonomic Network Computing Science (D.R. Avresky) – for a society that depends on information, this project aims to develop self healing networking technology that will avoid disruptions.
2. FuturICT (Steven Bishop) – aims to develop a decision support system to pre-empt and manage societal challenges on a planetary scale. See description below.
3. Federal Water Administration (Sergio Brandano) – defending human right to water, including autonomous and robotic installation and management of supply networks.
4. Computed Medicine (Anthony J. Brookes) – knowledge management and workflow infrastructure for personalised health care.
5. ICT Beyond Limits (Tommaso Calarco) – quantum based information technologies.
6. Robot Companions for Citizens (Paolo Dario) – developing a new generation of soft bodied sentient machines which act and interact with their physical and social environment.
7. The Social Computer (Fausto Giunchiglia) – Internet scale human problem solving.
8. Unravel the human brain (Bart Haex) – understanding integrated brain activities and neurological disorders.
9. The Web Time Machine (Wendy Hall) – understanding the past and present to build the future.
10. Ubiquitous Powering (Adrian Ionescu) – energy aware nano-electronic personal companions.
11. Graphene Science and Technology for ICT and Beyond (Jari Kinaret) – innovative device design using graphene.

12. The Human Brain Simulation Project (Henry Markram) – a facility for scanning, using supercomputers to study the brain from gene to behaviour and building brain derived technologies.

13. Sustainable Personal Living Technology (Norman Packard) – computer controlled personal fabricators

14. Matrix Re-done (Giuseppe Riva and Brenda K. Wiederhold) – VR for self consciousness, cognition and learning (virtual embodiment)

15. Non-deterministic Polynomial Computing (Lutz Schubert) – overcoming issues of scalability in solving non-polynomial (chaotic) systems. We note that Keith Jeffery (RAL) is the coordinator of this bid.

16. Uncovering the Human Cell Lineage Tree in Health and Disease (Ehud Shapiro) – a high throughput infrastructure mapping somatic mutations

17. S-Gaia/ Smart Society (John Sutcliffe-Braithwaite) – intelligent multi-media systems and services embedded in society

18. The Transition to Real World Computing (Peter Van Roy) – using machine learning and elastic computing to process "dirty" data, grey literature, audio and video.

19. Ubiquitous Complex Event Processing (Rainer von Ammon) – event driven BPM for choreography of internet services

We now illustrate the potential scope of some of these projects by describing relevant aspects of the proposed FuturICT Knowledge Accelerator.

5.1 The FuturICT Knowledge Accelerator

Social systems increasingly feature crises leading to unstable and potentially dangerous situations that are characterised by abrupt and large scale changes. Such chaotic disruptions are very hard to predict with any accuracy and even harder to control. The reasons behind the recent failures stem from the inadequacy of current theories in the social sciences and the inefficacy of tools and mechanisms for shaping new assumptions capable of establishing a more rapid and effective innovation process.

There is growing momentum in Europe for an EU ICT Flagship proposal to explore social life on Earth, see `http://www.futurict.eu`. Preliminary investigations are being carried out in the VISIONEER project (EU FP7 Support Action) `http://www.visioneer.ethz.ch`. There is a white paper [8] explaining the background and potential impact of such a project which would run for at least 10 years and ultimately be on the billion EUR scale. A number of members of the UK community have already expressed interest (the bid coordinator is Steven Bishop from UCL).

The goals of the FuturICT project are to try to make sense and efficient use of the large amounts of data about society. It proposes to do the following.

- develop novel ICT systems including applications and infrastructures which combine the best of human and computational abilities to support the understanding, integrative design, and management of complex systems;

- apply these to model techno-social and economic, transport, environmental and other global systems;
- create instruments to support the self organisation, decision making and governance in politics, business, industry, and academia, with the aim to foster societal goals, e.g. robust techno-social and sustainable economic systems;
- develop principles and tools that will facilitate the emergence of high quality processes, products and institutions in techno-social networks.

The three grand challenge areas outlined in this paper are as follows.

1. socio-economic challenges – multi-disciplinary research to understand and mitigate extreme events;
2. massive data mining and reality mining – sensor networks, i.e. ubiquitous computing, using data to forecast future events;
3. social super-computing – simulate, optimise and manage large complex systems.

Quite a lot of the discussion focusses on multi-level modelling and the use of multi-agent systems and agent based modelling, in particular as follows. Besides massive data mining capabilities, it is required to build up suitable super-computing capacities for the simulation, optimisation, and management of sustainable techno-social and economic systems. Gigantic computer power is, for example, needed for large scale computational analyses in the following areas.

- Massive data mining, e.g. real time financial data analysis;
- Network research, community detection;
- Monte Carlo simulations of probabilistic system behaviour;
- Multi-agent simulations of large systems (e.g. "whole earth simulation", which may involve up to 10 billion agents and complementary environmental simulations);
- Multi-agent simulations considering human cognitive and psychological processes, e.g. personality, memory, strategic decision making, emotions, creativity, etc.;
- Realistic computer simulations with parameter rich models, e.g. coupling simulations of climate and environmental change with simulations of large techno-social-economic environmental systems;
- "Possibilistic" multiple world view modelling, e.g. to determine the degree of reliability of model assumptions and to improve the overall prediction capability;
- Calibration of parameter rich models with massive datasets;
- Scanning of multi-dimensional parameter spaces;
- Sensitivity analyses;

- Parallel worlds scenario analyses, e.g. to test alternative policies, etc.;
- Visualisation of multi-dimensional data and models of complex systems
- Optimal real time management of complex systems, e.g. "guided self organisation", "self optimisation".

It should be underlined that most challenges addressed by the FuturICT Flagship concern a combination of the above points.

As contributions to these areas STFC staff are or have been involved in projects such as EURACE, an EU FP7 project for agent based modelling in macro-economics (Chris Greenough) and NeISS: National e-Infrastructure for Social Simulation (Rob Allan).

6 DARPA – New Theories of Intelligence

This news item is from 8/5/2009, see `http://www.darpa.mil/dso/solicitations/sn09-35.htm`.

DARPA's latest venture, called "Physical Intelligence" (PI) is to prove, mathematically and by demonstration, that the human mind is nothing more than parts and energy. In other words, all brain activities – reasoning, emotions, processing stimuli – derive from physical mechanisms at work, acting according to the principles of thermodynamics in open systems. Thermodynamics is founded on the conversion of energy into work and heat within a system, which could be anything from a test tube solution to a planet. The processes can be summed up in formalised equations and laws, which are then used to describe how systems react to changes in their surroundings.

If funded, projects will implement a variety of solutions to extract energy from their environment and spontaneously evolve and adapt in relation to the energy available in the environment. They are then expected to demonstrate the emergence of complexity and ultimately physical (artificial) intelligence.

The military wants a new equation: one that explains the human mind as a thermodynamic system. Once that's done, they're asking for abiotic, self organising electronic and chemical systems that display the PI principles. More than just computers that think, DARPA wants to re-envision how thought works – and then design computers whose thought processes are governed by the same laws as our own.

The PI program plan is organised around three inter-related tasks: (1) creating a theory (a mathematical formalism) and validating it in natural and engineered systems; (2) building the first human engineered systems that display physical intelligence in the form of abiotic, self organising electronic and chemical systems; and (3) developing analytical tools to support the design and understanding of physically intelligent systems. If successful, the programme would launch a revolution of understanding across many fields of human endeavor, demonstrate the first intelligence engineered from first principles, create new classes of electronic, computational and chemical systems and create tools to engineer intelligent systems that match the problem and environment in which they will exist.

There is a Defense Science Office Web site where project members can interact. The main part requires registration for access but there is an FAQ, see `http://www.sainc.com/PITeaming/pi_faq.asp`.

Solicitation workshops were held in mid-2009 since which there have been no further updates to the site.

For a number of reasons, this initiative was severely critisised in the press. There is however reference to one project which has been funded for three years from Feb'2010. It is led by Yong Chen at UCLA, see `http://www.mae.ucla.edu/news/news-archive/2010/`.

7 Conclusions and Recommendations

We here draw together themes from the programmes described above and note their relevance to applied computational science and engineering.

7.1 The UK Computing Research Grand Challenges Initiative

Despite their high profile participants, most of the UKCRC areas are after five years still in a community building phase, as was to some extent the original intention. They have set up networks and have organised workshops and meetings, often bringing in other disciplines. Whilst some areas will remain as a discussion forum, it is time for others to move forward to become actual projects with published aims and a roadmap or landscape leading to the delivery of testbeds and software. It is perhaps timely to become engaged in this work as and where appropriate.

GC1 *in Vivo, in Silico* is clearly relevant to the computational modelling and simulation agenda. It has technology themes of data sharing, combining models, metadata, ontologies, semantic Web, model validation, spatio-temporal frameworks, visualisation, virtual knowledge centres, scaleable architectures, new languages, interface technologies (e.g. haptics).

GC2/4 Ubiquitous Computing. It is at present not clear how computational simulation and modelling will benefit from the issues being addressed in this initiative. We note however that a strength of the NW-GRID project was in linking all the way from micro-sensor networks up to Grid based supercomputer simulations of the environment, for instance in risk management around flooding events [17]. GC2/4 has been influencial in setting the agenda for the EPSRC WINES Programme: Wired and Wireless Intelligent Networked Systems.

GC3 Memories for Life notes that there are potential applications to science because the technology would apply perfectly well to corporate or group memories as well as those of an individual. An obvious application area of that sort is the corporate memory of a scientific research group or consortium, which is generally now spread over Web sites, e-mail correspondence and individual holdings of papers, slides, notebooks, datasets, etc. This is relevant to knowledge management as part of the research life cycle which is growing in importance as challenges become more complex. Themes include large scale complex self modifying systems, data representations, inference, security, trust, coupled models, languages, information management, Web 2.0 mashup, natural language processing, artificial intelligence, cataloguing, annotating, information discovery and retrieval, RFID, ontologies, folksonomies, data reduction, managing distributed data, digital curation.

GC5 Architecture of Brain and Mind. Although many practical applications are possible, their primary goal is to increase an understanding of the nature and variety of natural and artificial information processing systems. This is likely to influence many areas of research including psychology, psychiatry, ethnology, linguistics, social science and philosophy. It could also transform ideas about education in the longer term.

GC6 Dependable Systems Evolution. Some practitioners in industry and researchers from universities believe it's now practical to use formal methods to produce software, even non-critical software. And that this will turn out to be the cheapest way to do it in the longer term. Given the right computer based tools, the use of formal methods could become widespread and transform the practice of software engineering. We note that Juan Bicarregui of STFC is a GC6 committee member.

GC7 Journeys in non-classical Computation is investigating a number of areas of computing which move away from the Turing machine or embody analogue or bio-chemical processes. Process driven parallelism such as CSP and Pi-Calculus could be re-visited, e.g. on multi-core systems where thousands of threads must be supported. The feasibility of large scale applications development using languages supporting this approach should be investigated.

GC8 Learning for Life. There are important applications of lifelong learning in the research context, for self motivated learners to acquire new skills and for teaching about applications, tools and technologies. Themes include: semantic Web, information management. We have already been involved in projects which embrace this goal, through ReDRESS: Resource Discovery for Researchers in e-Social Science (which developed a Web based learning repository) and Sakai, a leading open source teaching and learning portal which includes portfolio management.

GC9 Bringing the Past to Life may be of relevance to computational science and engineering for a number of reasons. Firstly, there will be a need to employ similar visualisation methods for large datasets arising from simulations. Secondly CSE could be applied to interpretation of historical artefacts, e.g. in analysing data from experimental studies such as those carried out on synchrotrons. Finally computer re-construction of artefacts could be carried out facilitating studies of 3D models, as already done for architectural sites. This could be extended to engineering studies of artefacts such as ships to support theories of design and usage. This could be included in the work of the Virtual Engineering Centre.

7.2 Towards 2020 Science

Recommendations from this report were to do the following: establish science and science based policy high on the political agenda; urgently re-think how we educate tomorrow's scientists; engage the public in science; re-think science funding and science policy structures; create new kinds of research institutes; re-energise computer science to tackle grand challenges; a call to action to develop new conceptual and technological tools; develop innovative public private partnerships to accelerate science based innovation; find better mechanisms to create value from intellectual property.

The paragraph describing a new kind of research institute is particularly relevant as it states: *There is, of course, no shortage of research institutes focusing on everything from molecular biology to superconductivity, and from nano-science to neuro-science. Many, however, tend to be highly discipline based (e.g. molecular biology). While such institutes are highly valuable, we argue there is a need for a new kind of institute focused on "grand challenges" rather than "grand disciplines". Such new kinds* *of institutes would be highly inter-disciplinary, combine teaching, training and research, and be focused on solving a problem or "winning" a race, rather than simply producing papers. Two good examples of this kind of institute would be the Sanger Centre in the UK and the Earth Simulator in Japan.* This clearly would also apply to the Hartree Centre.

Themes mentioned in this report include: computer science concepts and tools plus mathematical and statistical techniques, data management, evolutionary systems, resiliance, fault tolerance, adaptation and learning, semantics, multi-core, SOA and P2P architectures, symbolic computation, precision and error management, complexity and coherence, machine learning.

7.3 Century of Information Research

Atkinson *et al.* [3] noted that *Many of the obstacles to realising the full potential of the Century of Information Research will be primarily social rather than technical. An important example is the dis-incentives that researchers may sometimes perceive when contemplating participating in inter-disciplinary projects.* (Committee on Science, Engineering, and Public Policy, 2004).

The main thrust of their report is therefore on empowering the community to use and share data and information by building on the skills and ICT infrastructure which have begun to be layed down during the UK e-Science Programme.

7.4 FuturICT

Themes were already listed above and include: data mining, network research, Monte Carlo simulations, multi-agent simulations, sensitivity analysis, visualisation, real time steering.

7.5 Physical Intelligence

Concepts relevant to the objectives of the PI program can be found in numerous disciplines and areas of research including statistical physics, non-equilibrium thermodynamics, dissipative systems, group theory, collective behaviour, complexity theory, consciousness theory, non-linear dynamical systems, complex adaptive systems, systems analysis, multi-scale modelling, control systems, information theory, computation theory, topology, electronics, evolutionary computation, cellular automata, artificial life, origin of life, micro-biology, evolutionary biology, evolutionary chemistry, neuro-psychology, neuro-physiology, brain modelling, organisational behaviour, operations research and others.

7.6 Final Comments

Many of the themes from these grand challenges are the same as those in the world wide e-Science programme [2, 5, 1]. As noted by Atkinson *et al.* [3] it would be appropriate to re-purpose the outputs of that programme to support computational science and engineering. Some of the other themes noted above should also be investigated, in particular different ways to express massive parallelism and aspects of software engineering which embrace formal methods.

Finally, it is important to align the research grand challenges with real world grand challenges. Many of these are encapsulated in the priority themes of the UK Research Councils. In all disciplines, increasing amounts of data are being created, often with the use of public funds, and the expectation is that these will be curated and shared to increase their value to society. These data must be analysed in innovative ways to provide better answers to society's pressing questions, for example, those identified as cross-council themes: Energy; Living with Environmental Change; Global Threats to Security; and Ageing, Life Long Health and Wellbeing. (UK Government Department for Innovation, Universities and Skills, 2007).

8 Acknowledgements

This investigation was funded by EPSRC under its Service Level Agreement with the Computational Science and Engineering Department at STFC.

References

[1] R.J. Allan *Virtual Research Environments: from Portals to Science Gateways* (Chandos Publishing, Aug'2009) 266pp ISBN 978-1-84334-562-6 `http://www.woodheadpublishing.com/en/book.aspx?bookID=1892&ChandosTitle=1`

[2] M. Atkinson, J. Crowcroft, C. Goble, J. Gurd, T. Rodden, N. Shadbolt, M. Sloman, I. Sommerville and T. Storey *Computer Challenges to Emerge from e-Science* `http://www.semanticgrid.org/docs/Vision.pdf`

[3] M. Atkinson, D. Britton, P.V. Coveney, D. De Roure, N. Garnett, N. Geddes, R. Gurney, K. Haines, L. Hughes, D. Ingram, P. Jeffreys, L. Lyon, I. Osborne, R. Perrott, R. Procter, C. Rusbridge, A. Trefethen and P. Watson *Century-of-Information Research (CIR): A Strategy for Research and Innovation in the Century of Information* Prometheus 27 (2009) 27-45 DOI: 10.1080/08109020802657479. See also `http://wikis.nesc.ac.uk/escienvoy/Century_of_Information_Research_Strategy`

[4] G. Bell, A.J.G. Hey and A. Szalay *Beyond the Data Deluge* Science 323 (AAAS, 6/3/2009) 1297-8. DOI 10.112/science.1170411

[5] F. Berman, G. Fox and A.J.G. Hey (eds.) *Grid Computing: making the Global Infrastructure a Reality* (John Wiley and Sons Ltd., 2003) ISBN 978-0-470-85319-1

[6] L. Cardelli, Publications in Computational Systems Biology and Computer Science (Microsoft) `http://lucacardelli.name/indexPapers.html`

[7] S. Emmott *et al. Towards 2020 Science* (Microsoft, 2006) `http://research.microsoft.com/en-us/um/cambridge/projects/towards2020science/`

[8] D. Helbing *et al. The FuturIcT Knowledge Accelerator: Unleashing the Power of Information for a Sustainable Future.* (ETH Zürich, 2010) `http://www.futurict.ethz.ch/data/futurict.pdf`

[9] J. Kavanagh and W. Hall (eds.) *Grand Challenges in Computing Research 2008* (BCS and IET, 2008) 46pp `http://www.ukcrc.org.uk/grand_challenges/news/gccr08final.pdf`

[10] E. Prem (ed.) *Analysis of Flagship like Examples* (Eutema GmbH, June 2010) `http://cordis.europa.eu/fp7/ict/fet-proactive/docs/flagship-ws-june10-55-fet-study_en.pdf`

[11] P. Prusinkiewicz *et al. Algorithmic Botany at the University of Calgary* `http://www.cpsc.ucalgary.ca/Research/bmv`

[12] Sir D'Arcy W. Thompson *On Growth and Form* (CUP 1917, Dover reprint 1942) 1124pp ISBN 0-486-67135-6

[13] D. Harel *A Grand Challenge for Computing: Full Reactive Modeling of a Multi-Cellular Animal* Bulletin of the EATCS, European Association for Theoretical Computer Science, no. 81 (2003) 226-53. Reprinted in Current Trends in Theoretical Computer Science "The Challenge of the New Century", Algorithms and Complexity, Vol.1, Paun, Rozenberg and Salomaa (eds.) (World Scientific, 2004) 559-68

[14] Sir C.A.R. Hoare and J. Misra *Verified Software: Theories, Tools and Experiments* (July 2005) `http://vstte.ethz.ch/`

[15] Sir C.A.R. Hoare *Communicating Sequential Processes* (Prentice Hall International Series in Computer Science, 1985) 256pp ISBN 0-13-153271-5. Available on line at `http://www.usingcsp.com`

[16] Sir C.A.R. Hoare and A.J.R.G. Milner *Grand Challenges for Computing Research* The Computer Journal 48:1 (OUP, Jan'2005) 49-52 ISSN 0010-4620

[17] D. Hughes, P. Greenwood, G. Blair, G. Coulson, F. Pappenberger, P. Smith and K. Beven *An Intelligent and Adaptable Grid based Flood Monitoring and Warning System* Proc. UK e-Science All Hands Conference (EPSRC, 2006)

[18] A.J.R.G. Milner *Communicating and Mobile Systems: the Pi-Calculus* (CUP, 1999) ISBN 0-521-65869-1

[19] D. Noble *The Music of Life – Biology beyond the Genome* (OUP, June 2006) 176pp ISBN 978-0-199-22836-2 `http://www.musicoflife.co.uk/`

[20] S. Stepney, S.L. Braunstein, J.A. Clark, A. Tyrrell, A. Adamatzky, R.E. Smith, T. Addis, C. Johnson, J. Timmis, P. Welch, R. Milner, D. Partridge *Journeys in Non-Classical Computation I: A Grand Challenge for Computing Research* Int. J. Parallel, Emergent and Distributed Systems 20:1 (Mar'2005) 5-19

[21] S. Stepney, S.L. Braunstein, J.A. Clark, A. Tyrrell, A. Adamatzky, R.E. Smith, T. Addis, C. Johnson, J. Timmis, P. Welch, R. Milner, D. Partridge *Journeys in Non-Classical Computation II: Initial Journeys and Waypoints* Int. J. Parallel, Emergent and Distributed Systems 21:2 (Apr'2006) 97-125

[22] S. Wolfram *A New kind of Science* (Wolfram Media Inc., 2002) `http://www.wolframscience.com/nksonline/toc.html`

[23] J. Woodcock and R. Banach *The Verification Grand Challenge* Computer Society of India Comms. 31:2 (May 2007) 33-6 special issue on Formal Methods ISSN 0970-647X `http://www.cs.man.ac.uk/~banach/CSIComm-May07-FM.html`

[24] *Road Map for Memories for Life Research* (1/1/2008) http://www.memoriesforlife.org/downloads/M4L_ROAD_MAP_2008.pdf