

## **Complex Systems Research in FP7**

This document is the result of various consultations and meetings that were held in preparation of the 7<sup>th</sup> framework programme. It is based on contributions from the research community, in particular from the following researchers

O. Babaoglu  
C. Barret  
P. Bourgine  
Ch. Diot  
A.E. Eiben  
D. Floreano  
D. Hales  
A. Hoekstra  
J.H. Johnson  
S. Kirkpatrick  
M. Morvan  
C. Nikolau  
N. Packard  
M. Schoenauer  
P. Slot  
L. Steels  
A. Vespignani

**Motivation:** Nothing in traditional science and engineering that seek to understand and design systems by ‘divide and conquer’ - by breaking components into smaller parts - has prepared us to face challenges posed by systems that the ICT revolution pushed to the centre of technological and scientific interest: We fail to manage today’s massively distributed ICT systems composed of billions of autonomous components that exceedingly penetrate other systems including social and mission critical technological systems; we equally struggle to make sense out of massive ICT generated data on intricate webs of cause and effect in living, social and man-made systems.

In contrast, the science of ‘‘complex systems’’ (CS) embraces irreducibility of system behaviour and emphasises adaptive self-organisation of components, without central design - often inspired by self-organisation in biological and social systems. It seeks to develop theoretical foundations that allow characterising the nature of transitions in behaviour – emergence - that occur when systems massively scale-up. The vision is to explore – based on these foundations - radical new venues in monitoring, modelling and designing of massive scale systems.

### **Three focus areas for complex systems research in FP7**

**Prediction and Predictability of multi-scale systems:** The radical increase in the amount of IT-generated data on living and social systems brings about new challenges related to the sheer size of data and will transform science and engineering as we know it. Combining these data with computational infrastructures and the conceptual framework provided by a *data-driven* science of CS will allow *predicting* quantitatively the behaviour of large scale systems. The challenges include understanding how one can reconstruct multi-level systems and their dynamics by integrating simulation with data platforms that connect models to massive sets of heterogeneous and often incomplete data. Conceptual, theoretical and methodological foundations are necessary in understanding multi-scale networks, dynamic regulatory networks, and predictability limits of large agent-based models. Meeting these challenges will allow the creation of computational infrastructures able to provide reliable, detailed and quantitatively accurate predictions for large multi-scale CS in disciplines such as system biology, epidemics, modelling and design of mission-critical infrastructures, or option analysis in decision making.

**Enabling self-organising information networks that respond to human needs:** The Internet, the Web, distributed P2P systems, or, more generally, systems of services form large human-computer networks of autonomous entities. The vision is to enable such networks to autonomously organise into topologies and functionalities meeting desired objectives. The challenge is to develop tools for measuring and modelling their evolving topology and dynamics. Breakthroughs will lead to deeper understanding and the ability to predict and design new generations of autonomous information systems by exploiting models of self-organisation, adaptability and social behaviour. Applications include P2P collaborative search engines, emergent cooperative ICT-mediated communities, and ‘service economies in the large’ comprising billions of heterogeneous interconnected services.

**New design paradigms for adaptive artificial systems:** A science of CS will lead to a breakthrough in artificial system design for massive-scale systems of interacting heterogeneous, noisy, and often unreliable entities. The challenge is to engineer permanent adaptive dynamics allowing these entities to achieve *collectively* desired functional objectives, making the whole artefact viable in changing environments. Design paradigms must go beyond one-time optimization at the system design stage to open-ended ‘just-in-time’ adaptation during system deployment. The need for online structural modifications implies, at the component level, the need of adaptable components with high plasticity such as evolvable hardware chips or software modules, self-organising artificial cells, or neuro-morphic chips. Applications include societies of smart artefacts that adapt to human needs, robot societies that collectively achieve goals by means of cooperation and emergent communication and novel hardware architectures.

## Appendix 1: Prediction and Predictability of multi-scale systems

Over the last decade great progress has been made in providing the theoretical foundations and conceptual framework for the science of complex systems, including emergent phenomena, co-operative behaviour, complex networks, adaptability, evolution, and so on. Theoretical understanding has evolved from simple toy models into structured models in which the heterogeneities and details of the system under study are becoming increasingly included. The challenges for the best prediction and predictability include understanding how one can reconstruct multi-level systems and their dynamics by integrating simulation with data platforms that connect models to massive sets of heterogeneous and often unreliable data.

The new modelling approaches are increasingly based on *actual* and *detailed* data on the activity of individuals, their interactions and movement, as well as the spatial structure of the environment, transportation infrastructures, traffic networks, 'omics' data, etc. Also, a huge amount of data, collected and meticulously catalogued, has become available for scientific analysis and study by the complex systems community. Regulatory networks capturing the mutual interactions between proteins in cells, networks which trace the activities and interactions of individuals, social patterns, transportation fluxes and population movements on a local and global scale have been analyzed and found to exhibit complex features encoded in large scale heterogeneity, self-organization and other properties typical of complex systems. Along with the advances in our understanding and characterization of systems' complexities, increased CPU power has led to the possibility of creating multi-scale dynamical models by using thousands of coupled stochastic equations and allows agent-based approaches which include millions of individuals. Complex systems science is finally maturing and should now become data-driven and able to *predict* quantitatively the behaviour of very large scale multilevel natural and artificial systems.

Despite such an integrated approach being in its infancy, the level of realism and detail achievable today allows us for the first time to ambitiously imagine the creation of computational infrastructures able to provide reliable, detailed and quantitatively accurate predictions for complex systems. These will be based on large, possibly distributed, data collection and multi-scale computational models as now used in weather forecasting. In other words computational approaches are now ready to interface with the complex features of biological systems, infrastructure networks and social dynamics, entering the era in which they must become a major predictive tool. Such an approach will provide radical new ways of understanding the physical, biological, ecological, and social universe.

Complex systems science bridges the gap between the individual and the collective: from genes to organisms to ecosystems, from atoms to materials to products, from notebooks to the Internet, from citizens to society. It cuts across all the disciplines. It is part of every discipline. Complex systems science is also *computer enabled* and ICT will be part of *all* the research programmes of FP7. Indeed, the radical increase of *in vivo* data on living, social and environmental systems enabled by ICT creates completely new challenges. The enormous *size* of databases (petabytes,  $10^{15}$ ) will transform science and engineering. Real-time data from many kinds of sensors in ecosystems, the human body or in business networks, combined with new and more powerful ways of exploiting this wealth of data, will lead to major advances in our knowledge of the multi-level dynamics of living and social systems. This will greatly increase our capacity in design, planning, and management. For example, in Health the new science of complex systems will revolutionise the medical treatment of diseases, and revolutionise the delivery of treatment. Individual problems of individual people will be treated, creating the long-term Grand Challenge of 'Personalised Health'. This requires (i) huge distributed databases of every individual's

genotype, phenotype, medical and general history, (ii) new ways of searching, communicating and processing this information, and (iii) new and more efficient ways organising the delivery of food and health services to Europe's half billion inhabitants.

In this context, there is a strong drive for complex systems in all domains to generate new computational disciplines such as *computational neurosciences*, *computational biology*, *computational physiology*, *computational ecology*, *computational economics*, *computational sociology*, and so on. Necessarily these computational theories are 'local' to their domains and do not address the traversal questions of complex systems science. In particular they do not address the generality of reconstructing multilevel dynamics of systems from data collected at many levels. The CS community has to organize the convergence of these domain-based computational sciences into a more unified framework. This involves tackling a set of basic challenges that includes:

- System measurement or 'harnessing data': Cope with the data explosion in science (share, store, categorise, reduce data) by providing computational networks able to cope with Peta-bytes of heterogeneous data and by providing tools to analyse these data. Here the CS community can implement and share new kinds of *European Platforms*, similar to the CERN platform in physics, for sharing huge distributed peta-byte databases and distributed peta-flop clusters implementing new computational methods for reconstructing and simulating multi-level dynamics, connecting new kinds of multi-level models to massive sets of heterogeneous and often unreliable data.
- The creation of new *formalisms* capable of representing and modelling complex multilevel systems and networks and their dynamics. This involves understanding how descriptions at *all* different levels relate to each other, *e.g.* how does a gene network influence the organism (bottom-up), and how does the environment change the genetic expression of cells (top-down).
- Development of new formalisms for the *reconstruction* of formal system-specific models from data in order to obtain the best 'predictive' multilevel dynamics, where 'prediction' is usually probabilistic relating to uncertain distributions in the state space of the system.
- Simulation and Peta-scale computing: executing a model (*if available*) of highly dynamic systems to understand their dynamics, to predict probabilistically their dynamics, to control their viability and resilience, to "optimize" or better to ameliorate its robustness or adaptation, to design new ones.
- Use the modelling approach and formal model representation to understand and predict the consequences of a given level(s) *control strategies* on the multi level dynamics, in order to keep them in their viability domain and to be resilient after an important perturbation?

Meeting these challenges will allow the community to face a set of *domain specific grand challenges* where the integration of the complex systems set of tools with information technology, computer science and specific domain disciplines is a necessary conditions toward a predictive scientific approach. Among these grand challenges we can list *brain modelling*, *epidemic and social response forecast infrastructures* etc. Advances concerning the above challenges and grand challenges will also stimulate new ideas and tools in computing and networking and will drive the next generation of ICT systems. Data explosion and heterogeneity of data and sources of data will need novel tools to ensure use across distributed sets of data, means to endow data with semantic meaning (meta data, categorisation), means to store data in an easy accessible way (distributed storage in P2P systems), and to analyse data in a distributed way (SW oriented architectures and P2P systems for distributed computing for data analysis). Computing will need to adapt to distributed nature of data and novel computing paradigms based

on P2P (asynchronous, fault/delay-tolerant) or Grid type of environments. To achieve real understanding we need to base computational tools on *sound foundations* and avoid the trap of performing ‘mathematically illiterate’ simulations from which we cannot draw any insight. Particularly pertinent will be research in systems where interaction between parts determines system dynamics and structure.

In this context, ICT and computer science deliver new paradigms to formulate problems and provide constructive solutions, computationally tractable. This integrated approach will elaborate synergies that will arise by combining layers that are often treated as independent: (i) the data selection, integration level defining the considered phenomenology of interest (ii) the declarative layer for a precise definition of the problem (of reconstruction, adaptive control, viability, resilience ..) using the adequate representation formalisms and dealing with the data (iii) the layer of computational methods for solving the problem, (iv) the layer of the computational substrate (HPC, clusters, grid, dedicated HW, evolvable hardware, adaptive computation...).

An outcome of this research under FP7 will be the creation of *Big Instruments* to support complex systems science, involving massive multi-petabyte ( $10^{15}$ ) distributed databases and massive distributed multi-petaflop processing facilities. By the 2013 some two-thousand highly trained will be employed by these CERN-like installations. To support this, Europe also has an urgent need to increase its *human resources* in complex systems, training at least 300 complex systems scientists per year.

The potential impact of this new Complex Systems Science appears in four ways (i) a better understanding of many complex systems and their dynamics to support the pressing needs to engineer and manage complex systems, *e.g.* cancer, multinational companies, drugs, transport, and climate change; (ii) better control of the means of fabrication as dynamic complex socio-technical systems, *e.g.* new processes and materials, multi-site factory production, and supply chain dynamics, (iii) a better understanding of the complex environment in which engineered systems exist, *e.g.* regulation, ethics, markets; and (iv) a better understanding of the design, engineering and management process which is often itself a creative complex multilevel complex human system, capable of great successes but inherently liable to spectacular failures.

### **GRAND CHALLENGE 3.1: Theoretical Foundations for Data-driven Complex Systems Modelling**

- Reconstruction theory:
- Mastering predictability
- Model reconstruction and Model validation
- Multilevel Modelling and Integration of multilevel data
- Theory of multiscale networks (interfacing, feedback)
- Predictability Limits and Sensitivity Analysis

### **GRAND CHALLENGE 3.2: Computational X: Harnessing Peta-scale Computing for Complex Systems**

- Domain-driven examples of X
- Modelling the brain as a Complex System (GRAND CHALLENGE)
- Computational ecology (GRAND CHALLENGE)
- Modelling Emerging Diseases
- Computational Social Sciences and Economics (GRAND CHALLENGE)
- Modelling the traffic and movement of people and goods throughout Europe
- Modelling Cultural Dynamics

From Components Biochemistry to Systems Biology  
Personalized Health

## **Appendix 2: Enabling self-organising information networks that respond to human needs**

Information is central to today's society, and the World Wide Web is its main repository. Its vast resources have altered the very way we think and organize our lives. Yet managing and navigating this immense world of information and services requires an ever more sophisticated set of tools and technical instruments. Gathering and managing large data sets is a fundamental issue in many aspects of society including individuals, industries, and governments, and will grow more problematic in the future as diverse computing entities (phone, laptops, PDA, sensors, digital cameras, etc.) will interact in distributed networks, gathering, relaying and storing vast quantities of information and providing diverse services to users. Increasingly such information is produced and consumed machines in automated processes requiring increasingly sophisticated distributed and socially intelligent computation.

The key challenge is to learn how to design Internet and Web systems that can self-organise, self-adapt and optimize their interactions and functions in a continuous and robust manner to satisfy user demand. Complex Systems Science can provide models, theories, tools, mechanisms and approaches that allow for a principled design method to be developed to address this key challenge. In the following sections we outline several problem spaces in which CS can be productively employed. These are structured around two "Grand Challenge" themes: *The Ultimate Web* and *Socially Intelligent Systems*.

### **The Ultimate Web**

Major efforts are required to keep track of the full spectrum of digital material on the Internet, discovering who participates in creating and using them, and providing a means of devising novel and improved means of information exchange and processing, entertainment, and commerce. In the coming decades, the Internet will not only connect each of us to each other and to the world's knowledge, but much of business, consumer and scholarly activity will increasingly flow through it. The Internet will continue to be a main medium for interaction across European cultures, and critical to the vitality of these cultures. By advancing the power of the internet through the application of Complexity Science we envisage the creation of the "Ultimate Web" - something that goes radically beyond current internet and web functionality.

#### *Measuring the Topology of and Dynamics on the Internet*

Internet efficiency currently suffers from an inadequate understanding of Internet topology and traffic flow. A vision motivated by complexity science, and also inspired by biological analogy, is that all systems should devote a tiny fraction of resources to creating something equivalent to a nervous system – facilities that can learn about their local environment and make that information widely accessible throughout the network. We can build on the success of DIMES in this area, as it is a first step towards an Internet that measures itself.. Studying Internet topology, traffic flow and information content together feeds naturally into a complex systems perspective, as it recognizes that each influence cannot be understood adequately in isolation. Topology influences traffic, which affects content, and vice versa.

#### *Collaborative Web Search and Semantic Overlay Networks*

Search-engine technologies provide support for organizing and querying information. But for advanced information demands, search engines all too often require excessive manual intervention – manual classification of documents into a taxonomy for a good Web portal, or browsing through long lists of results with lots of irrelevant items. Current Web and intranet search engines fail when faced with questions of the kind we tend to ask naturally based on a

context. Current technology fails because no single web site can offer a good match, or because the user can only interpret and sort through the results by using prior knowledge.

A promising approach to developing technologies that can handle natural queries is through collaborative Web search in an Internet-scale peer-to-peer (P2P) system. The idea is that every peer would have a full-fledged search engine that indexes a small portion of the Web reflecting the interests of that user. Such architecture has several advantages over a centralized server like Google since peers can become specialists in their portion of the web applying sophisticated semantic indexing schemes that is too costly for centralised systems. Achieving collaboration among peers will require strategies for routing queries to other peers and for exchanging metadata, statistics, and background knowledge to form an evolving “semantic overlay network”. Understanding the dynamics and behaviour of such a network requires analyses at different levels and scales of the overall network. To be practically viable, a P2P approach needs good incentive mechanisms to limit the influence of egoistic or malicious peers (see Socially Intelligent Systems section below).

Successfully addressing this difficult issue requires combining expertise and methods from multiple scientific fields such as game theory, sociology and evolutionary biology, statistical physics, and computer science - all of which naturally come together within Complex Systems models and approaches. Ultimately, success in semantic overlay technology will allow the world's disparate cultures and software systems to talk to each other. The bible gives a relevant example of the consequences of uncontrolled growth without solving this problem.

### **Socially Intelligent Systems**

Services deployed over the Internet will become increasingly distributed and open. Such systems can not rely on centralized control, truth or authority. Yet it is desirable for them to operate in an efficient way, optimizing the global welfare rather than just individual welfare. As is well studied in economics and the wider social sciences these two welfare levels are not always in alignment. Deployable mechanisms that produce the spontaneous emergence of cooperative and evolving communities in such systems require new kind of information-systems engineering, spanning the social and biological sciences and software engineering. In a number of key problem areas the CS approach offers the potential to advance this engineering agenda.

#### *Cooperation for Efficient Resource Distribution*

Peer-to-Peer (P2P) systems are coming to dominate the Internet. As much as 35 percent of all Internet usage is now produced by just one peer-to-peer system – BitTorrent – and although most systems currently support user-level file swapping (often of an illegal nature), legitimate and successful P2P applications are now emerging. With the increasing speed of broadband domestic Internet connections, even television and media networks could become aspects of P2P technology. P2P technology also offers the potential to empower individuals and groups by redistributing computational resources dynamically on-demand from idle machines and servers to be utilized more effectively elsewhere. Many of the user machines on the "periphery" of the Internet are often idle or under-utilized. Efficient mechanisms for effectively redistributing these "latent resources" could, therefore, deliver very high returns at little cost. However, the problem of ensuring high levels of cooperation while suppressing selfish and malicious behaviour within open systems is holding back progress.

Similar difficulties of social cooperation affect the Internet at the network level. The Internet is comprised of many Autonomous Systems (ASes), each of which is a sub-network administered by a single organization. Internet routing isn't scripted by some social optimal performance, but is rather the outcome of many individual, self-interested decisions and this can lead to suboptimal global results. Many other aspects of Internet optimisation arise out of similar situations – where

the collective or global benefit needs to be optimized based on individual (often selfish) local actions.

Recent work has adapted CS models from the social sciences to produce protocols encouraging the spontaneous evolution and maintenance of cooperative P2P communities. This is an example of how a strongly interdisciplinary Complex Systems perspective - integrating the ideas from biological and social sciences with computer science - can improve the efficiency of key IT systems.

### *Security and Robustness*

In distributed open systems, like the Internet, the aim of security engineering is to design for resilience and robustness, i.e. for the ability to withstand a broad spectrum of potential attacks and to fail gracefully. Inspiration can be taken from higher biological organisms, which base their security around a fully distributed and robust immune system. In this regard, recent developments suggest that self-organised distributed systems may be the best way to tackle seemingly insurmountable security problems, such as those stemming from SPAM or viruses. One approach to stopping SPAM that is currently proposed relies on the cooperative activity of millions of computers, sharing information, to act as a kind of self-organising “collaborative filter” or immune system against such unwanted messages.

### *Science of Services*

We are experiencing a fundamental shift in the structure of economies, businesses and social communities from data-rich to information rich and to service rich. With the help of ICT, organisations and infrastructures are rapidly evolving into a dominating web of networked structures with various levels of complex and intertwined services offerings across multiple technological, societal, governmental and business systems. To ensure optimal use and dependability of these new structures it is important to ask: What is the impact of these ICT induced transformations? The challenge is to better understand the structure and the dynamics of these novel organisations and infrastructures that ICT helped create and transform and ultimately to understand how they impact our daily lives.

The vision is to use ICT as a means to model, design, and implement novel adaptive models of businesses and infrastructures, or even novel societal structures and governmental policies. This vision necessitates a system view that integrates technological and economic processes and ICT services with social processes and human factors into one ‘fabric’ created and maintained by ICT.

Such a ‘science of services’ - or more generally a ‘science of hybrid systems’ combining technological layers with models of human behaviour - will necessitate an ambitious long-term research agenda combining expertise from computer science, high performance computing, mathematics, and engineering with expertise from social and behavioural sciences, and economics. From economics one can learn how to design incentive-compatible mechanisms for harvesting knowledge, from anthropology the importance of social networks for organisations, and social psychology might help understand how to motivate certain kinds of behaviours. Novel methods for achieving semantic interoperability between different processes could be inspired by biological or social paradigms.

## **Grand Challenges (long term)**

### *GRAND CHALLENGE 1.1: The Ultimate Web*

#### Long-term Vision

- Apply and develop theory, tools and paradigms from CS to understand and engineer the next generation of internet and web applications

- Recast the internet as a co-evolving ecosystem of interacting entities under continuous change and reconfiguration
- Engineer the "ultimate web" supporting changing user needs with the properties of self-organisation, self-management and self-repair ("self-star" properties)

#### Intermediate Milestones

- Measuring the topology and dynamics of the Internet
- Keeping an Internet History
- Ultimate Google - distributed, massive and robust collaborative web indexing and search
- Ultimate Gnutella - mass content distribution for a broadband and wireless world
- Ultimate Akamai - self-organised content caching for a faster web

#### *GRAND CHALLENGE 1.2: Socially Intelligent Systems*

#### Long-term Vision

- Apply and develop the new "bottom-up" emergence-based social mechanisms and theories from CS science to understand and engineer socially intelligent ICT systems
- Engineer systems that guide self-organising ICT systems toward cooperative collaboration between potentially selfish or malicious system subcomponents
- Understand, refine and deploy mechanisms for dynamic and emergent semantics and hence allow socially interacting subunits to evolve their own communication languages
- Develop a "Science of Services" which integrates self-organising ICT systems and human organisational and economic behaviour – a new paradigm for the e-economy.

#### Intermediate Milestones

- Self-organising P2P network infrastructures and protocols
- Efficient resource sharing through emergent cooperative protocols
- Better formal models of human social organisation in relation to ICT – modelling humans in the loop
- Network security via continuous adaptive infrastructures - e-mail SPAM and virus control
- Observatories: monitoring of millions of activities across business, society, and infrastructures

### Appendix 3: Adaptive artificial systems

A science of CS will lead to new approaches in artificial system design to engineer adaptation in massive-scale systems of interacting heterogeneous, noisy, and often unreliable entities. The challenge is to develop novel tools and methodologies to instruct a huge number of components to achieve collectively desired functional objectives, making the whole artefact viable in changing environments, by permanent adaptive dynamics. At the global level, the "how" of design must switch from one-shot optimization to open-ended on-the-fly adaptation during the actual use of the system, while the "what" must exceed simple parameter identification and optimization to reach on-line structural modifications. At the component level, this in turn implies the need to build adaptable components with high plasticity such as evolvable hardware chips or software modules, self-organising artificial cells, neuro-morphic chips. Applications include systems of smart artefacts that adapt to human needs, robot societies that are able to collectively achieve goals by means of cooperation and emergent communication, new architectures for hardware systems required by the predicted end of Moore's law, and adaptable and scalable software systems.

The two main characteristic of complex systems, both natural and artificial, are the high level of interactions between a very large number of independent components on the one hand, and the multi-scale aspect of those interactions: a rescue team or a distributed recording system will be made of hundreds of small robots or sensor; each one of those entities will have sensing and acting capabilities thanks to some "brain" made of hundreds or thousands of interconnected chips - but each of those chips can in turn be built from numerous artificial cells, each cell being itself driven by thousands of nano mechanisms, ... Moreover, because of the different time scales, but also of the different technologies involved, all those levels will operate at different time scales, too. Hence the design of such systems must be in itself "scale free", allowing the designer to handle design issues both at the component- and at the system level simultaneously: a system built upon self-repairing and adaptive chips should itself be, at its own level, self-repairing and adaptive ...

The distinction between hardware and software will also become fuzzy: designing a new chip made of many cells regards hardware technology, but such chip must be programmed, not to mention the crucial routing and communications capabilities that must also be crafted, and pertain to software engineering.

Keeping all this in mind, we will now sketch some trends that we think can allow breakthrough advances in complex system design. First, at the system level, new global algorithms will be necessary to cope with the challenges raised by programming unreliable adaptive (changing) components to perform ill-specified tasks in uncertain dynamic environments. Second, new hardware components will have to be designed, too, exhibiting those properties of flexibility and adaptation that are mandatory to increase the robustness and reliability of the higher level. Third, the specifications of communications will become more and more important in decentralized systems with heterogeneous components, with clear intrications between routing between components having some

degree of adaptivity and flexibility, and emergent semantics in collaborative search and discovery within distributed information systems.

#### New algorithms

-----

At the heart of the new design paradigms for complex adaptive artificial systems there are the algorithms (or methods), being the working engines that are responsible for creating the desired global behaviour and/or structure. These methods include evolution, (machine) learning, reinforcement learning, dynamic programming, ... While most of these methods constitute well established scientific and engineering fields, the present state-of-the-art is not appropriate for the challenges posed within this programme. This implies the necessity for a paradigm shift in a number of aspects.

First, these methods are typically applied in a fashion oriented at a one-shot result, e.g., finding the optimal solution of some difficult problem, or obtaining a data model with minimal mean squared error. However, in the applications envisioned here, the emphasis lies on (open-ended) adaptation, much rather than on optimization. The shift must hence concern switching from user-specified to behavior-driven or example-based software engineering: the difficulty of complete specifications of a complex system increases with the complexity of the system itself - not to mention the additional difficulty brought by the dynamics of a system in a changing uncertain environment. On the other hand, giving examples of the expected behavior, or driving the system while in exploitation (the user in the loop) remains affordable, but requires new algorithmic paradigms.

A second aspect is the type of entities that are manipulated by these methods. As of today, most of the algorithms are to generate parametric changes, new structures emerge usually through relying on parameterized designs. The new challenges, however, require structural changes. For instance, evolving new representations, or inventing new concepts/attributes for learning seem necessary prerequisites. Third, while the methods and algorithms considered here are to empower an adaptive system they are not adaptive themselves. That is, they depend on a hand-made algorithmic setup that is not changing during execution. Considering that the application context is inherently dynamic, the method should be adjustable on-the-fly to the changes in system behaviour. In essence, some meta-mechanisms or self-adaptation mechanisms need to be developed that adapt the methods at runtime. These new requirements force a reconsideration and probably reinvention of the algorithmic cores.

Further to this, the "emergence engines" that we need to develop may be combinations of algorithms that together manipulate interactive parts with decentralized control forming a robust and flexible dynamic system. To master such emergence engines, we need to develop an understanding and an application methodology of combining various adaptive methods leading to controllable emergence.

#### New components

-----

The concurrent downscaling and increment of electronic components on a chip makes the system increasingly vulnerable to errors due to heat, electro-magnetic interference, and atomic effects. One may also wish to add or reduce the number of computing components at run-time without halting or re-designing the system from scratch. Furthermore, the distribution of computing resources across several geographic areas implies variable levels of communication and different time scales.

These features, that are typical of massive-scale electronics, require the development of novel design paradigm that can achieve robust computation with unreliable components. For example, one may consider using artificial evolution to generate large-scale electronic chips capable of achieving and maintaining a desired functionality despite occasional failure of their components or external disturbance. Such systems would also be capable of self-organization to adapt to the up-sizing or downsizing of the circuit. Evolvable chips would be particularly interesting in the domain of analog and analog/digital large-scale systems, whose functionalities are very attractive, but are very difficult to design with conventional engineering techniques because of their inherently non-linear and complex operation. Automatic synthesis of such circuits will require novel hardware and software tools that allow the exploration of topologies characterized by feedback loops and asynchronous operation.

Another possibility is to take inspiration from the functioning of neurons. Neuromorphic engineering has been studied for some time, but very little work has been done so far in the development of large-scale systems composed of individual neuro-morphic chips, akin to a silicon cortex, capable of collectively tackling more complex problems than the sum of all individual chips. Such neuromorphic chips could be endowed with the ability to adapt at run-time and possibly rewire themselves to cope with varying demands and environmental change. Neuro-morphic chips may also be interfaced to living neuronal assemblies to form a symbiotic, large-scale system capable of synergetic interactions and collective problem solving, especially in control domains characterized by large-scale sensing and non-linear dynamics.

One may even consider drastically different computing materials, such as biological or artificial cells, capable of collectively and robustly transforming information according to some functional constraint while self-sustaining themselves through metabolic principles. Such programmable cells could incorporate mechanisms of self-reproduction, evolution, and self-repair and be readily available in large numbers. The synthesis of programmable cells will require the development or adaptation of new methods and tools, such as microfluidics and in vitro evolution. Programmable cells could be also synthesized to be used in isolation for achieving specific functions. For example, one may imagine nanoscale robots of programmable cells capable of navigating through human bodies by means of a signaling network of proteins that connect chemical sensors to nanoscale actuators.

Today we have some design principles to conceive robots at the micro- and human-scale level that are expected to operate in isolation. These robots are making their way in our homes and are used in dangerous situations. However, there are very few, or none, design

principles for conception and operation of large-scale robotic systems composed of several homogeneous and/or heterogeneous units that are expected to collectively carry out tasks that a single unit could not possibly achieve. Although today's technology allows us to build hundreds or thousands of relatively inexpensive robots, it is not clear what are the minimal and necessary sensory, communication, and physical pre-requisites for such robots to cooperate and optimally divide labor in partially unpredictable and changing environments. We call such machines "social robots" for their ability to live and robustly operate in societies that autonomously self-organize to achieve complex missions. For example, a swarm of

social robots could represent a powerful alternative to a single and much more complex robot. Also, social robots would be able to dynamically change their morphology and/or re-allocate tasks by means of emergent communication in order to maintain the functionality and viability of the society when the environmental constraints change.

#### New circulation of information

-----

The information explosion poses tremendous challenges regarding the intelligent organization of data and the effective search for relevant information in business and industry (e.g., market analyses, logistic chains), society (e.g., health care), and all sciences that are more and more data-driven (e.g., gene expression data analyses and other areas of bio-informatics). The problems arise in intranets of large organizations, in federations of digital libraries, and in the largest and most amorphous of all data collections, the World Wide Web, including numerous databases that reside behind its pages. Complex Systems research holds the promise to revolutionise the retrieval of useful data through emergent and collaborative approaches, which by their very nature will be better suited to face the heterogeneity and dynamics of the Web than current approaches around top-down ontologies. Progress in providing more flexible, dynamic ontologies has the potential to revolutionise our use of the Web and other information repositories.

Using emergent semantics collaboration can be taken one step further: encourage the establishing of mutual agreements on the interpretation of data that would permit widely dispersed sources of information, currently stored in the web under a multitude of incompatible formats, to be accessed and shared more easily. This is a very different from current approaches, which rely on imposed standards and are facing heterogeneity as a major problem.

Among the benefits of such techniques will also be that they greatly reduce the economic costs and productivity losses associated with difficult problems of software interoperability. Semantic interoperability remains a key challenge in information system technology, and today constitutes a major fraction of IT costs. Recent progress in emergent management of peer-to-peer networks and information agent systems provides a framework to address this problem. Forming an agreement is a distributed reasoning and negotiation process in a network of agents that are able to relate their local data representations with each other. With such an approach, local knowledge can be shared

through a network, the distributed agreement process allows agents to resolve ambiguities and disagreements in a scalable fashion, and by automating global agreement processes substantial human effort can be saved.