

InfoSam2020

The Information Society of 2020
-an exercise in planning for the future

NTNU, April 19-20, 2004
Pre-conference proceedings
Arne Sølvberg (ed.)

2nd Edition, July 1, 2004

Co-arrangers:



Faculty of Information Technology,
Mathematics and Electrical Engineering



The Norwegian Board of Technology

Preface 2nd edition of INFOSAM2020 proceedings.

InfoSam2020 is an ongoing exercise at the Norwegian University of Science and Technology (NTNU) in planning research and education for the future Information Society. The first phase of the exercise resulted in the presentation of a number of position papers at a conference in Trondheim on April 19-20, 2004, see <http://www.ime.ntnu.no/infosam2020>

After the InfoSam-conference this April some position papers have been polished, and one new position paper on Network systems has been added. The papers have been collected into this proceedings, which can be downloaded as a whole from the above web site. Individual papers can be downloaded as well from the website. Future developments in the InfoSam project will be reported on the infosam2020 website.

Trondheim, July 1, 2004

Arne Sølvberg

-editor-

Table of contents

Preface 2nd Edition	
Table of contents.....	5
Introduction.....	7
Enabling technology / Basic ICT infrastructure.....	11
Electronics - Prof. Einar J. Aas et. al.....	13
Networks and networked systems - Prof Bjarne Helvik et.al.....	25
Communication Technology - Prof. Geir Øien et.al.....	33
Computer Systems - Asc. Prof. Kjetil Nørvåg et.al.....	41
The organization of human affairs / The creative human.....	49
Information Security and Safety - Prof. Torbjørn Skramstad et.al.....	51
Information Systems - Prof. Jon Atle Gulla et.al.....	61
Language Technology - Prof. Torbjørn Norgård et.al.....	69
Learning with ICT - Researcher Leif Martin Hokstad et. al.....	75
Enabling technology / Computation.....	85
Mathematics - Prof. Kristian Seip et.al.....	87
Computational Science - Asc. Prof. Anne Cathrine Elster/ Prof. Einar Rønquis et.al....	91
The material world.....	97
Electrical Energy - Prof. Arne Nysveen et.al.....	99
Gas technology - Prof. Morten Hovd et.al.....	109
The biological / medical world.....	117
Bioinformatics - Asc. Prof. Finn Drabløs et.al.....	119
Health Informatics - Asc. Prof. Øystein Nytrø et.al.....	127
Life support/medical instrumentation. - Prof. II Kjell Arne Ingebrigtsen et al.....	135
Invited talk Chief Scientist Michael Brodie.....	141
"Holistic Solutions From Technologies that Converge, Shift, Disrupt, and Disappear"	
Invited talk Professor Keith Jeffery,.....	143
"The Future is GRIDs"	
Invited talk Dr. Johan de Kleer.....	145
"Smart Matter"	

Introduction

The Information Society of 2020

“The Hitchhiker’s Guide to the Galaxy” was first published in 1979. This is 25 years ago. The major alien character Ford Prefect from one of the planets of Betelgeuse had been stranded on Earth for 15 years. In his “survival kit” he kept many futuristic gadgets. One of them was a flat gadget with a 4x8 cm screen, which could carry one million book pages of information! Such was the science fiction gadget of 25 years ago. Today we can buy a comparable PDA/PC for less than a thousand euros.

How fast information technology develops! Performance/price ratios for central ICT components have doubled every 18 months for many years. This seems to continue for the foreseeable future. That means that the performance/price ratio increases 100-fold over 10 years, and 1000-fold over 15 years. Current processing capacities of the largest computers are comparable to 1-2% of the processing capacity of the human brain. In 2012-15 we should expect that the processing capacity of a large computer is comparable to that of the human brain, and in 2020 this capacity will probably be available on desktop computers. An even faster cost decrease can be expected for information storage.

This does not mean that computers will become human-like, and even compete with humans. Computers will (probably) not fall in love in ways comparable to how humans fall in love. They will (probably) not reproduce all on their own. They will (probably) not be able to develop radical new solutions to ill-defined problems by exercising fantasy and creative thinking. But they will develop into even more useful tools for satisfying human desires than what they constitute today.

The ongoing radical changes in cost/performance bring possibility for new applications of ICT. To be in the forefront of discovering what the new possibilities are we need to encourage our ability to exercise fantasy in an informed and creative manner. We need to give the young generations relevant knowledge, to make it possible for them to see opportunities when Lady Luck crosses their path, to see possibilities for doing old things in a new manner and also to do completely new things.

This is what the INFOSAM2020 exercise is about.

The Faculty of Information Technology, Mathematics and Electrical Engineering (IME) at The Norwegian University of Science and Technology (NTNU) has an annual educational output of approximately 500 master students and 30 PhDs. The rapid pace of technological change impacts strongly on curricula and research. Technological change is also a key driver for changing the way we live, due to the deep penetration of computers in all realms of society. This impact on student recruitment, as well as on the relevance of research themes. The quality of our teaching depends on the quality and relevance of our research. In order to keep relevant we have to relate to future needs.

To this end the project InfoSam2020 has been established. The overall objective is to determine which knowledge profile our students should have in 2020, and the associated teaching and research. So we have to chart a course for the Faculty for the next 15-20 years. This may seem a long time, but it is not when we compare it to the 10 years it takes to set up high quality research in a new field.

The first step in the process is to explore the space of future possibilities, to describe possible technological changes and subsequent changes in our society. The next step comprises the charting of our course. The initial InfoSam2020 activity is a two-day seminar for exploring relevant features of the future information society, from now on until 2020.

InfoSam2020 is arranged in cooperation with The Norwegian Board of Technology, an independent body for technology assessment and foresight established by the Norwegian Government in 1999. The Board aims to assess impacts and options of technology in all areas of society; to stimulate public debate on technology; and to support the political decision-making process and shaping of technological change. Insights from InfoSam2020 will be thus be presented by the Board to relevant political authorities.

To serve as a framework two major themes have been selected. The first theme is devoted to the use of “intelligent” artifacts in various realms of human endeavor. The second theme is devoted to enabling technologies, in particular within the fields of mathematics, ICT and electrical engineering. The themes are outlined in the sequel.

Working groups have been established to prepare position papers on a limited number of subthemes within the very wide framework. The position papers are published in this report, and are also available on <http://www.ime.ntnu.no/infosam2020>

Theme 1: The Application View, Artifacts in the human environment

The material world

Keywords: Artifacts, devices, sound and sight in/out (man-machine interaction), energy, heating, lighting, energy sources and distribution, new materials, transport, travel, vehicles, roads, housing, smart houses, buildings, cities, clothing

The biological world

Keywords: Demography, life support, sustaining health/healthiness, medical diagnosis and treatment, old age, food, food safety, nutrition, cosmetics, exercise/fitness

The creative human

Keywords: Communication, entertainment, education/training edutainment/infotainment, art, games, man at work, man at leisure, mobility, fashion, lifestyle, experiences, meaningfulness, personalized services to relieve/simplify/sort/save time

The organization of human affairs

Keywords: Organizations, companies, government, social (support) systems, schools, universities, hospitals, banking, commerce, sustainable systems, service provisioning, CRM, politics, democracy, voting, organization of leisure, safety and security, law and order, sustainability, reliability, vulnerability, environmental issues

Theme 2: The Discipline View, Enabling Technologies

Information systems

Keywords: Cooperation technologies / CSCW, information management / digital libraries, human-computer interfaces, software engineering / information systems engineering, conceptual modeling, workflow management, semantic web, databases, datamining / information retrieval, artificial intelligence, system operation, competence management

Computer systems

Keywords: computer architectures, supercomputers, micro- and nanocomputers, quantum computation, computer graphics and visualization, database management systems, programming languages, program correctness management, system security protection

Communication technology

Keywords: wireless/wired communication, mobility, broadcasting technology, multimedia communications, convergence fixed/mobile / broadcast/unicast, unified infrastructures / IP technology, ad hoc networks, personal and local area networks, ambient computing, ultra high speed networks, optical switching, user interfaces, quality of service

Electronics

Keywords: Ambient intelligence, system-on-chip, embedded systems, “smart dust”, wireless systems and components, signal processing, VLSI, nanotechnology, optics/photonics, quantum communication, spintronics, mechatronics, functional materials, navigation, remote sensing, sensor/transducer technology, audio/video technology

Life support/medical technology

Keywords: Bionics, gene technology, medical equipment/systems

Mathematics

Keywords: Cryptography, coding theory, computation science / numerics, signal processing/wavelets, statistics

Electrical Energy

Keywords: Sustainable electricity production, power electronics, energy (systems) optimization, intelligent energy distribution, electricity storage systems, battery technology

Arne Sølvberg
(ed.)

Enabling technology / Basic ICT infrastructure

- Electronics**
- Networks and networked systems**
- Communication Technology**
- Computer Systems**

Electronics

Professor Einar Johan Aas, Department of Electronics and Telecommunications enar.aas@fysel.ntnu.no
Professor Harald Krogstad, Department of Mathematical Sciences Harald.Krogstad@math.ntnu.no
University Lecturer Morten Olavsbråten, Department of Electronics and Telecommunications olavsbra@tele.ntnu.no
Professor Börje Forssell, Department of Electronics and Telecommunications borje.forssell@tele.ntnu.no
Professor Peter Svensson, Department of Electronics and Telecommunications peter.svensson@tele.ntnu.no
Associate Professor Lars Lundheim, Department of Electronics and Telecommunications lundheim@tele.ntnu.no
Professor Il Jens Hjelmstad, Department of Electronics and Telecommunications jens.hjelmstad@ericsson.no
Professor Einar J. Aas, Department of Electronics and Telecommunications enar.aas@fysel.ntnu.no
Associate Professor Johannes Skaar, Department of Electronics and Telecommunications
johannes.skaar@fysel.ntnu.no Associate Professor Thomas Tybell, Department of Electronics and
Telecommunications thomas.tybell@fysel.ntnu.no Post Doc Mikael Hammer, Department of Electronics and
Telecommunications mikael.hammer@tele.ntnu.no

Scope

The scope of this paper is limited to encompass electronic, acoustic, photonic and micro/nano-electromechanical devices, circuits, components, sensors, actuators, systems and their interfacing with the rest of the world. Typically, transmitters and receivers are employed for such interfacing. The research embraces methods for modelling, simulating, prototyping and manufacturing the devices and systems mentioned. It also includes a large variety of applications, e.g. remote sensing and surveillance, localization and positioning, micromechanical operations, medical and automotive electronics.

Most importantly, we try to identify trends with considerable impact on the direction and rate our society is moving. The key words are: *convergence*, *integration*, *communication everywhere* and *at all times*, and *ubiquitous computing*. *Monitoring* and *acting* are important keys to a vision of an *ambient intelligence* that we may live in. Smart sensors, speech recognition systems, and intelligent actuators, will be helpers in this vision. We will concentrate on the “hard” side of this trend, but will include embedded system software whenever appropriate to achieve proper design tradeoffs. One spotted trend in this area is “software radio”, where traditional hardware parts may be implemented in software. Moreover, whenever necessary for future research within the scope of “Electronics”, we will envision development of new applications to our research areas, fueled by progress in computational science.

1. Introduction

Nanotechnology can be seen as the technology based on devices taking advantage of quantum effects and size effects occurring in nm-dimension objects. It is a cross-disciplinary endeavour, relying on subjects such as physics, chemistry, electrical engineering, biology and medicine. Moreover, the ability to control or manipulate matter on the atomic scale is central to nanotechnology, and a basis for future electronic applications.

Transmitters and receivers, including components and software, are vital to connect system functions in a network.. The transmitters and receivers in question are to be regarded as sub-systems, where the main system functions are connected to communications and navigation. The views concerning receivers apply mainly to the functions of receivers, but to some extent also to the purpose and use of these functions.

Optics and photonics include:

Fiber optical communications: dense wavelength-division-multiplexing, optical routing and processing.

Optical sensors: applications in medicine, structure monitoring, and offshore.

Quantum communications: Quantum cryptography, teleportation.

Quantum information processing (quantum computing).

To increase the speed of optical communication systems, more data processing will be performed at the optical layer. The technology for optical processing will be applied and further developed for the next generation of communications; quantum communications, i.e., exchanging quantum information (quantum bits). Due to low decoherence, photons may also be interesting for quantum computing if sufficiently efficient nonlinear interactions are discovered (nanotechnology). Quantum memory may be achieved using techniques for "stopping" the photons. Optical components used for classical and quantum communication will also be interesting in sensor technology, e.g. for application in medicine.

Acoustics spans a wide range of research fields and application areas. Within InfoSam2020, we concentrate on the areas *engineering and environmental acoustics*, *underwater acoustics* (remote sensing and other applications), and *music and audio technology* (sound recording, transmission, reproduction etc).

2. Authoritative foresight sources

Below are given a few references to international foresight or review papers of relevance for "Electronics".

A study on "Teknisk Framsyn" (Technological foresight) in Sweden produced a series of reports, developed and presented in 1999. It was performed as a joint effort between The Royal Swedish Academy of Engineering Sciences (IVA), the Swedish National Board for Industrial and Technical Development (NUTEK), the Swedish Foundation for Strategic Research and the Federation of Swedish Industries, and is conducted in close cooperation with the Government, companies, authorities and other interested parties. One of the reports considered "Information and Communication Systems", with key areas: *ubiquitous IT*, *complex systems*, *infotainment* (Upplevelseteknologi), *information safety and security*, *architecture and infrastructure*, and *interfacing*. An update, called **Nytt, bättre och säkrare – IT i framtidssamhällets tjänst** (*New, better and more secure – IT at the service of future society*) was published in 2003, and can also be found on [2]. The three keywords are not new to us: *convergence*, *integration*, and *transparency* (ubiquitous computing). This trend is fueled by Moore's law, and progress in wireless communication. From our point of view, this report also embraces R&D & education in micro- and nano-electronics, wireless communication, RF devices, and more within the realm of "Electronics".

2.1 The International Technology Roadmap for Semiconductors

The famous Moore's law from 1965 states that the number of transistors that can be economically integrated on a chip doubles every year. More recently, the doubling takes about 18 months. Several corollaries of this "law" (really an observation and a 10 year prediction) offer better computing performance, and less power consumption, as feature sizes decrease. The semiconductor industry, and the EDA industry (EDA – Electronic Design Automation), both work hard to continue fulfilling Moore's law. To do so, roadmaps are needed, to identify and eliminate possible stumbling blocks. The most authoritative map available is the so-called *International Technology Roadmap for Semiconductors*. This Roadmap (ITRS) is an assessment of semiconductor technology requirements. The objective of the ITRS is to ensure advancements in the performance of integrated circuits. This assessment is a cooperative

effort of the global industry manufacturers and suppliers, government organizations, consortia, and universities.

The ITRS identifies the technological challenges and needs facing the semiconductor industry over the next 15 years, including possibilities for nanotechnology devices. The Semiconductor Industry Association (SIA), the European Electronic Component Association (EECA), the Japan Electronics & Information Technology Industries Association (JEITA), the Korean Semiconductor Industry Association (KSIA), and Taiwan Semiconductor Industry Association (TSIA) sponsor ITRS. The 2003 roadmaps are found via [1].

2.2 Other micro/nano foresight initiatives

The MEDEA+ Design Automation Roadmap 2002

This roadmap outlines the future requirements for *design automation* solutions for Europe. The steering group had representatives from all major semiconductor companies in Europe including STMicroelectronics, Philips, Infineon, Alcatel, and Ericsson.

The National Nanotechnology Initiative

The American initiative for research and development based on nanotechnology, based on the document NSTC/NSET Report, July 2000, see [NNI04].

Biomedical applications for MEMS and micro fluidics

Proc. IEEE, Vol 92, No. 1, January 2004 is a special review issue on biomedical applications for **MEMS and micro fluidics**. Many of the technologies reviewed are not very common in clinical practice today, so one possible future prospect would be further development and more widespread deployment of these technologies. At any rate, the reviews point out the major applications and motivation for research in the field.

Nanotechnology at NTNU

The report *Nanotechnology at NTNU 2003*, initiated by Rector, addresses status and future possibilities for nanotechnology research at NTNU. An international panel led by Prof. B. Stokke, Dept. of Physics, wrote the report based on a site visit at NTNU, and surveys of current research. Moreover, the report reviews current trends in nanotechnology initiatives in Sweden, Denmark, Europe (EU), USA, and Japan.

2.3 Acoustics

Computational acoustics

Many fields in applied acoustics rely heavily on the computation of sound field propagation through fluids and structures. Most computational methods have a complexity, which grows very rapidly with the studied bandwidth. In environmental noise control, the so-called EU noise directive states that strategic noise mapping must be carried out for all large European cities, railways and roads no later than 2008 [EUNo03]. This is a highly complex task, which is computationally very demanding when large areas should be studied, and many noise sources need to be taken into account.

Marine acoustics

Underwater or marine acoustics is very important as a key element in the preservation, development and exploitation of marine resources, and therefore of particular importance for Norway, as reflected by the national thematic research area of marine technology. Furthermore, acoustics is the primary tool for surveying the amount and distribution of biomass, fish and plankton in the sea. Recent work has been in connection with an industry project aiming to develop acoustic sensors for gathering acoustic backscattered data of plankton from moored or drifting buoys.

Traffic noise

Noise from road traffic and from aircrafts is dominating the noise problem in Norway, and it has been estimated that approximately half a million people in Norway are strongly disturbed by noise sources (*Regjeringens miljøvernpolitikk og rikets miljøtilstand, St meld nr 8, 1999-2000*). Research is needed for developing less noisy cars and trains, and an understanding of the road-tire noise generation. Models for sound propagation from a noise source in realistic environments need to be developed. Finally, better models of how noise is perceived are needed, as discussed below under Perception of sound.

Perception of sound

A common understanding in several fields is that we have insufficient knowledge of how a noise situation is perceived by a listener. As an example, when increasingly accurate predictions of noise situations can be made, inaudible details might be computed at a high cost. Consequently, models for the annoyance as caused by various noise sources need to be developed.

A similar situation occurs in communication acoustics, whether it is speech or general audio signals. The available bandwidth for communication increases rapidly and makes new multimedia services possible [AES]. Knowledge of the perceived Quality of Service by the end users will be essential for an efficient use of available resources [ITU]. The rapid development of audio and video codecs in the last decade, such as in the MPEG standards, are examples of that.

3. Drivers

3.1 Nanotechnology

There is an increasing number of research centres worldwide devoted towards nanotechnology. Some initiatives are presented briefly below.

NNI

This initiative was launched in 2001, and is the basis for the United States government's investment in nanotechnology. The goal is to ensure the leading role of United States in the first half of the 21st century. The long-term perspective aims at novel breakthrough within areas such as medicine and nano-electronics. President Bush has proposed to spend almost 1 billion US dollars on nanotechnology R&D in 2005.

EU

Within the European Union there is an increasing funding of nanotechnology. Out of 17.5 billion € allocated to the 6th Framework program, **1.3 billion €** is allocated to funding within nanotechnology, knowledge-based materials, and new industrial processes and devices [EUNA04]. Three priority areas have been suggested, including 1) information society technologies (nano-electronics, optoelectronics, micro-nanotechnology) and 2) nanotechnologies and nanosciences, knowledge based multifunctional materials and new production processes and devices.

The Nordic countries

In our neighbouring countries there are several examples of nanotechnology initiatives: for example **MC²** at Chalmers and *The Nanometer Consortium* in Lund in Sweden. It is worth noting that at Lund there is a new master program in nanotechnology. This is a cross disciplinary education, with one basis in electrical engineering. Sweden has established a "Swedish Nano Network".

3.2 MEMS

For MEMS devices, the main applications have traditionally been automotive air bag accelerometers and cheap disposable blood pressure sensors. Upcoming market drivers include feedback mechanical sensors for minimally invasive and robotic surgery, acoustic chemical sensors for smart buildings, and switches

and filters for high frequency integrated circuits.

Minimally invasive surgery offers benefits both in economic terms and in terms of reduced patient trauma. In the coming 15 years, the share of minimally invasive surgery is predicted to increase to about 80%. The main problem hampering a more widespread use of the technology is that it constrains the 3D vision and manual dexterity of the surgeon. With MEMS devices, it would be possible to measure applied force and also provide a means of measuring the mechanical properties of different tissues.

MEMS devices are also finding applications in high frequency electronics. Currently, the main application is for switches. MEMS switches offer benefits such as lower losses, better isolation and better linearity than PIN diode switches. High quality MEMS resonators and mechanically tuned capacitors have also been demonstrated. Both switches and resonators are predicted to become important as means of increasing the integration level.

Passive micro arrays based on hybridization and optical readout of signals from fluorescent tags are currently dominant research tools in genetic research. These analysis tools call for sample preparation and can often only be used in order to assess the final result. In order to integrate a sequence of biological or chemical processes, micro fluidic devices (lab-on-a-chip) are being developed. Important drivers in this field are genetic research and drug development, and the commercial deployment of newly discovered bioassays for point-of-care and field applications. The development of this kind of bioassays could be regarded as an important step on the road to personalized and predictive medicine.

In order to put the acquired biomedical knowledge to a more widespread use, it is of interest to develop cheap disposable micro fluidic devices that are rugged enough for applications outside research laboratories. One way of achieving this could be to break down the analyses in elementary operations that could be implemented by general-purpose programmable micro fluidic hardware.

Telecommunication, sensor- and medical systems are examples of application areas for analog and mixed-signal (AMS) integrated systems. The evolution of AMS sub-systems will be determined by the interplay between cost and performance, where performance is often measured in terms of speed and complexity. Important system drivers are:

- System-on-chip (SOC)
- High speed wired and wireless communication systems
- Portable battery operated equipment
- Magnetic and optical storage
- Automotive, drive by wire, environmental and safety
- Data converters

3.3 Transmitters and receivers

The drivers of the development of transmitters and receivers for wireless communication are the mass markets. Development costs are so high that they can only be justified by very large quantities of units sold. As an example, about one million equipments containing GPS receivers were sold in Europe in the year 2000, but this volume is expected to reach six millions in a year or two from now. About $\frac{3}{4}$ of this latter number concerns GPS receiver cards mounted in mobile phones, and the rest is mainly car-mounted units. A consequence of the attraction of the mass markets and the subsequent stiff competition is the declining price per unit, which in its turn further increases the size of the market. The exponential growth of the mobile-phone market is another example of the same kind.

4. National partners

MEF

Mikroelektronikk-forum (MEF) is a national consortium of approx. 10 companies and the Department of electronics and telecommunications. It focuses on recruitment and education of microelectronics designers

NMC

A micro-technological research laboratory is established in Gaustadbekkdalen towards micro and nanotechnology based on silicon [SINT04]. This initiative is headed by A. Hanneborg, SINTEF Elektronikk og kybernetikk.

NFR

The Norwegian Research Council (NFR) has launched a program called NANOMAT [NAMA04] towards nanotechnology and novel materials. This program is an important basis for funding of future experimental research within novel electronic materials for nano-scale detectors and sensors.

NTNU

A cross-disciplinary NTNU Nanotechnology laboratory is being established at the Gløshaugen campus, with emphasis on **four** subjects: 1. Nano-electronics, -photonics, and –magnetism, 2. Bionanotechnology, 3. Nanostructured materials, and 4. nanotechnology for energy and the environment.

The IME Faculty will identify possible national partners in this foresight work on electronics. We can select partners for discussion; partners we believe will stay “alive” for a long time. Such partners are found in companies, research institutes and in governance. But we should also think about the so-called “yet unborn enterprises” that will be born out of innovation from our faculty.

SINTEF IKT

Research collaboration on research project in communication acoustics and environmental acoustics.

Statoil AS

Financing of adjunct professor's position in marine acoustics.

The Norwegian Defence Research Establishment

Collaboration on research projects in marine acoustics/computational acoustics.

5. Possible scenarios

Nanotechnology is cross disciplinary as subject, and future electronic devices will not only be based on a continued downscaling of standard silicon technology. Novel electronic phenomena encountered in nanometer (**nm**) sized materials, biological, and vivid materials will be important, and increased functionalization is expected. This will lead to novel applications. It is worth quoting NNI in this context: *“The impact of nanotechnology on health, wealth, and lives of people could be at least as significant as the combined influences of microelectronics, medical imaging, computer aided engineering, and man-made polymers developed in the century just past”*.

Photonics

To increase the speed of optical communication systems, more data processing will be performed at the optical layer. The technology for optical processing will be applied and further developed

for the next generation of communication systems; quantum communications. Due to low decoherence, photons may also be interesting for quantum computing. Optical components used for classical and quantum communication will also be interesting in sensor technology. In particular, application of photonics in medicine (diagnosis and treatment) seems to be a promising field.

6. Important areas for R&D & education within the scope.

CAS (Circuits and Systems)

*The area **Circuits and Systems** encompasses (in Norway) design of digital and analog devices and systems. Manufacturing is mostly done abroad, and we do not believe that Norway will get to produce microelectronics components. Today, that major companies are “fabless”, implying that production is done where suited, based upon computer-generated mask data*

Research within **AMS (Analog and Mixed Signal) integrated circuits** will continue to focus on techniques and architectures for further increase in speed/frequency and accuracy. At the same time, the power consumption must be reduced. Key areas will include analog IPs for SOCs, sensor interfaces, data converters, and wireless system components.

With the advent of **nanotechnology, new electronic devices and sensors** will be available. Also today’s emerging applications involving bio-inspired systems and interfacing to biological systems are expected to play a much more important role in the future, which will clearly have an impact on R&D trends within AMS circuits.

An increased need for multidisciplinary knowledge among future students is expected due to, for example, the emerging application mentioned above. Hence, the selection of basic courses may change.

Furthermore, demand for ever decreasing design cycle time, and retargeting of AMS circuit blocks into other process technologies and variants, makes it necessary to develop efficient re-use methodologies for AMS functions.

Other important areas:

- **Digitally assisted analog functions** and background calibration: Scaling of digital functions will continue for each technology node. While AMS-functions will scale to much lesser extent because noise and matching considerations dictate larger component geometries. This will lead the way to new architectures where the balance between analog and digital functions is shifted even more towards the digital domain by use of digitally assisted analog functions.
- **AMS design for test, DFT, and built in self testing, BIST**: Analog and RF tests are time consuming and account for a large part of the test cost compared to the often small chip area occupied by AMS functions. To alleviate this problem it is necessary to develop new and efficient methodologies for built in self-test and design for test for AMS functions.
- **Fault modeling** of analog failures in digital and analog devices.
- **Lower signal voltage** requires better resolution in time domain (jitter, phase noise): Power supply voltage today (2004) is already at the level of 1 Volt. This severely limits the obtainable signal to noise ratio of AMS functions on a complex SoC. Power supply voltage will continue to scale down and is expected to scale down to 0.3-0.4 Volt by year 2020. This will lead the way for new architectures, which process the signal with greater accuracy in the time domain. To achieve

better time domain accuracy it is necessary to develop better and more accurate time references and oscillators for increased sampling accuracy of low-level analog signals.

- **Noise and interference:** Substrate coupling, inductive/capacitive coupling from interconnect.

The future development in **microelectronics and MEMS process technology** will cause parasitic phenomena and high frequency effects to become more marked. In order to exploit the inherent performance of a process technology, it will be necessary for the future designer to have a wide knowledge in fields ranging from physical phenomena like electromagnetism, device physics and modelling, to high level systems aspects like signal processing, modulation, communication and information theory. No single person could be expected to be an expert in all the relevant fields, but they should, at the very least, be capable of communication with experts in other fields.

6.1 Nanotechnology

It is interesting to note that the report *Nanotechnology at NTNU 2003* specifically recommends focus on research to develop tools for **nm**-structuring of functional materials for use within spintronics and electromechanical MESA structures for future NEMS applications, nanoscale detectors and transducers. This fits well to the experimental activities at IME, e.g. Department of Electronics and Telecommunications, where there is a large activity within the fields of micro- and nanotechnology. A large part of the activity is centred on developing low-dimensional electronic materials for future use in nano-electronics and nano-electromechanical systems (NEMS) applications. There is also ongoing work towards RF-MEMS and sensor technologies. The research aims at becoming a foundation for novel application within information and communication technology.

6.2 Photonics

will continue to be an important area for some Norwegian companies. In particular, optical sensor technology is currently a strong field in Norway. There are several companies working in this field, in addition to the academic effort at NTNU. During the last years there has been an increasing activity in the field of medical applications of photonics and lasers. This trend is expected to continue. In addition it is expected that quantum phenomena in optics will be important during the next 10-20 years. To be a part of this progress, more emphasis on quantum effects should be incorporated in the educational program.

6.3 Radio systems

6.3.1 Transmitters

The transmitter in a system for communications or navigation is one of the most power-consuming parts. In recent years, the drive for lighter and smaller units with more functionality has been pronounced. This increases the challenges for the transmitter. Properties like linearity, efficiency, bandwidth and power consumption in a transmitter chain are therefore increasingly important. This trend will no doubt continue in the future.

One important trend in communications systems is the use of increasingly advanced modulation schemes. This can clearly be seen when looking at the trend from “old” analogue systems (e.g. NMT, AMPS) to the digital systems (e.g. GSM) of today, and to the future systems (UMTS and beyond). The use of advanced

modulation techniques to transmit more information per unit of bandwidth requires improved linearity and efficiency of the transmitter. The power amplifier is the main source of distortion and power consumption in a transmitter chain. In the future, there will certainly be an increased focus on power amplifiers and transmitters [TARGET].

In future wireless systems, the flexibility of transmitters will be one of the keys to integrate different systems in one unit. Today, GSM implements “tri-band” receivers and transmitters, which can switch between three different frequency bands. In the future, the ability to implement several wireless systems like GSM, GPS, Bluetooth, UMTS and future systems in one unit will become even more important. Reconfigurable transmitters, which have the ability to change their behaviour to match several different systems, will require less volume, weight and power than separate transmitters for each system. Reconfigurable transmitters can be achieved by electronically changing the behaviour of components by tuning or switching components in and out of operation.

6.3.2 Receivers

There is a general trend in receiver development, which is assumed to continue for considerable time. A generic term often used is *software radio*. By this term, we mean receivers where some functions and a lot of parameters are controlled and can be changed by means of software, without any hardware modification. In order to exemplify this trend, receivers of satellite navigation systems are chosen below. In generic terms, there is little difference between receivers in navigation and communications, which justifies the exemplification.

A rough sketch of a typical receiver design looks as follows (in this order) and will certainly remain unchanged because of the laws of physics:

Antenna with preamplifier and filter(s), down-converter (to intermediate frequency), digitiser (i.e. A/D-converter), signal processors (including demodulation and filtering functions), control and computing processors, display(s).

The trend has long been to move digitisation towards higher frequencies, and this trend will no doubt continue as a consequence of improved synthesisers and faster A/D-converters, even if there may also be a rise in the number of quantisation levels. After digitisation of the signal(s), the only performance limits are the capacity of the processors and the software used.

There is no doubt that satellite navigation receivers will be able to utilise signals from all satellites above the horizon, regardless of whether they emanate from GALILEO, GLONASS or GPS satellites. The average consumer will buy navigation equipment based on the requirements of the application in question, without any consideration of the origin of the signals or the name of the satellite system(s).

If it is desirable to handle signals at very different carrier frequencies, the bottleneck of receiving equipment as sketched above will usually be the antenna and the RF front end. In order to maximise signal-to-noise ratios, it is generally advantageous to use multiple antennas, perhaps physically stacked on top of each other, instead of very wideband antennas. A similar principle can be used for the preamplifier and subsequent RF stages, i.e. several narrowband stages in parallel instead of one wideband channel. After the digitisation of the signals, however, all signals are canalized in the same manner. From this point of view, it is obvious that digitisation should be carried out as early as possible in the chain, and this has also been a trend for some years now.

6.3.3 Iteration

A logical consequence of the ability of receivers to handle signals from different systems is integration. In this context, integration means combined and optimum use of information supplied by available systems in order to increase the performance of the integrated system beyond the level of the best participating system. With navigation as an example again, this means that position accuracies given by the integrated receiver using signals from GALILEO, GLONASS and GPS satellites at the same time are better than they would be if just one of the systems were used. Optimised use in this respect requires weighting of measurement results taking known (or at least expected) error characteristics into account.

Mobile telephones represent a striking example of integration of navigation and communications. This implies not only a physical but also a functional symbiosis. In some applications, i.e. emergency calls from indoor positions and automatic location, both functions depend on the other.

As to physical integration, the trend is towards “receiver-on-a-chip”. This means that miniaturisation can be brought forward from the (so far) usual partitioning between RF and IF-LF chips to one-chip solutions with dimensions in the order of some mm². In some of these applications, the problem of mutual interference through the chip itself, resulting in increased noise factors, has to be solved.

6.3.4 Integrity

With increasing use of electronic systems for so-called safety-of-life applications, e.g. aircraft landing and navigation of high-speed passenger craft in narrow waters, the ability of electronic systems onboard to check and report about their own performance must receive great attention. This ability is called integrity. Requirements of this kind also influence receiver design, and more attention will have to be paid to such functions in years to come, not only in navigation but also in other applications as e.g. timing and synchronisation.

6.4 Components

6.4.1 RF CMOS technology

The operating frequency of components implemented in CMOS (Complementary Metal-Oxide-Semiconductor) process technology is steadily increasing. Currently, the technology is employed mainly for ISM (Industrial Scientific Medical) band radio transceivers for short-range communications, but attempts to employ this technology for mobile telephone and satellite-navigation transceiver building blocks are also being made. Published results have also demonstrated the performance of the technology for circuits operating at several tens of GHz. The authors of the International Technology Roadmap for Semiconductors [ITRS2003] predict that the choice of implementation technology in the future will be determined more by cost and performance requirements than by operating frequency.

The sensitivity and power consumption of CMOS radio receivers for the frequency range up to about 2 GHz may be improved in the future by use of MEMS (Micro-Electro-Mechanical Systems) technology for high-Q micromechanical resonator filters and mechanically tuneable high-Q capacitors. The use of this kind of mechanical resonator filters may also facilitate a reduction of the chip area of tuned radio-frequency circuitry.

It is currently possible to integrate CMOS radio transceivers on the same chip as complex digital subsystems. Unfortunately, radio electronics and digital electronics have very different process requirements. Radio electronics relies on the use of integrated spiral inductors, which cannot be downsized. Digital electronics, on the other hand, can benefit directly from a downsizing of the minimum transistor feature size. As a consequence, it is possible that the current system integration will come to an end in the future, unless, of course, the very promising MEMS technology can be used as a replacement for spiral inductors.

6.4.2 Microwave component technology

At operating frequencies in the 1-100 GHz range, microwave components are usually implemented in gallium arsenide (GaAs). GaAs is chemically in the III-V group. There are several different implementations of components based on GaAs, e.g. AlGaAs and InGaP. At frequencies above 40 GHz, other III-V materials like indium phosphide (InP) show good performance. For high power densities there are several 'new' materials with very promising properties, such as GaN and SiC.

The use of new materials in RF front-end components will be important to achieve high linearity, high efficiency and low power consumption.

6.4.3 Components for high-performance radio electronics

Some of the current trends in high performance radio electronics include the development of reconfigurable or tunable systems and the development of components with smaller dimensions and higher power handling capability [Proc IEEE Feb 2004]. Reconfigurable systems can be implemented by means of MEMS switches. Ferroelectric materials with a voltage-dependent dielectric constant can be used for continuously tuneable systems with a wide frequency tuning range.

Transistors based on semiconductor materials with relatively high band gap energy can be used for the implementation of transistors with high operating temperatures. This could be employed for miniaturisation of power amplifiers. Another technology with high power handling capability is FBAR (Film Bulk Acoustic-Wave Resonator). This is a kind of devices based on stacked piezoelectric thin films that can be used for duplexers.

Ambient Intelligence is the modern label for a society where our electronic "gadgets" disappear more and more from visible sites, and blend into the ambience. In the book "The New Everyday – Views on Ambient Intelligence" [EASM03], edited by Emile Aarts and Stefano Marzano at Philips Research and Philips Design, we find visions for the next decade of pervasive digital technologies. "We may find ourselves living with almost invisible, intelligent interactive systems". The book presents different technologies that need to be developed, but also a thorough discussion of what roles we want those technologies to play. How will we interact with them? –And how can we ensure that they improve the overall quality of our lives, rather than just improving efficiency? The book describes more than 20 projects, where experts in technology, design, social sciences and business work together to get a broad and deep understanding of the issues of *convergence*, *integration*, *communication everywhere* and at *all times*, and *ubiquitous computing*. Obviously, the realm of "Electronics" will play a key role in such projects, and we want to contribute to the development of new devices, sensors, actuators, transmitters and receivers, and speech and audio technology to implement such systems.

References

[ITRS03] International Technology Roadmap for Semiconductors; website: <http://public.itrs.net/>

[TEF03] Teknisk Framsyn; report: ISBN 91-7082-658-7; website: <http://www.tekniskframsyn.nu/>,

[NNI04] The National Nanotechnology Initiative (NNI); website: <http://www.nano.gov/> . Also: NSTC/NSET Report: [National Nanotechnology Initiative: The Initiative and its Implementation Plan](#)

[EUNA04]<http://www.cordis.lu/nanotechnology/src/era.htm>

[SINT04] <http://www.sintef.no/>

[NAMA04] <http://program.forskningsradet.no/nanomat/no/index.html?3568>

References on acoustics:

[AES] "Emerging technology trends in the areas of the technical committees of the audio engineering society, J. Audio. Eng. Soc. 51, pp. 442-451 (2003).

[ITU] ITU-TE.800 (08/94), "Terms and definitions related to quality of service and network performance including dependability" (1994).

[EUNo03] <http://europa.eu.int/scadplus/leg/en/lvb/l21180.htm>

[EASM03] E. Aarts and S. Marzano, "The New Everyday – Views on Ambient Intelligence", 010 Publishers, Rotterdam, ISBN 90-6450-502-0.

References on receivers and transmitters:

Feng, M., Shen, S.-C., Caruth, D.C., Huang, J.-J.: "Device Technologies for RF Front-End Circuits in Next-Generation Wireless Communications". Proc. IEEE, February 2004, pp 354 – 375.

Akos, D.M., Ene, A., Thor, J.: "A prototyping platform for multifrequency GNSS receivers". ION GPS/GNSS 2003, Portland, Oregon, USA, Sept. 2003.

Ståhlberg, C., Normark, P.-L.: "A hybrid GPS/GALILEO Open-Service Software Receiver Prototype". Navconvention 2003, Geneva, Switzerland, June 2003.

Kadoyama, T. et al.: "A complete single-chip GPS receiver with 1.6-V 24-mW radio in 0.18- μ m CMOS". 2003 Symposium on VLSI Circuits. Kyoto, Japan, June 2003.

NoE - TARGET (Network of Excellence - Top Amplifier Research Groups in a European Team). EC 6th Framework Program, Contract IST-1-507893-NOE (www.target-net.org)

Networks and networked systems

Professor Rolv Bræk, Department of Telematics, Rolv.Braek@item.ntnu.no

Professor Bjarne E. Helvik, Department of Telematics and Q2S CoE, Bjarne.E.Helvik@item.ntnu.no

Professor Svein J. Knapskog, Department of Telematics and Q2S CoE, knapskog@item.ntnu.no

Professor Lill Kristiansen, Department of Telematics, lillk@item.ntnu.no

Professor Øivind Kure, Department of Telematics and Q2SCoE, okure@item.ntnu.no

Q2S CoE: Centre for Quantifiable Quality of Service in Communication Systems. Centre of Excellence

The interconnectedness of everything...

Ubiquitous computing will be the dominating factor in shaping how networks need to evolve. It implies that all persons and devices are always accessible and able to communicate and exchange data.

Communication will be viewed as a utility, similar to water, sewage and electricity. It will simply be taken for granted. We can already see elements in the move from vertical technologies, one for each application like voice and data, towards a common IP dial tone. In addition, virtually every system is becoming a networked system. Consequently, networking and networked system architectures are rapidly becoming a key enabler for most systems and applications (pervasive networking).

The following is a brief account of some major trends that can be observed today, and an outlook to where they may lead by 2020.

1. Networks

It is unlikely that there will be only one network. More likely there will be a multitude of networks.

Clearly there will be a multitude of access network technologies. A person or device will typically be connected to several of these concurrently. It is not given that the various networks are operating under the same administrative domain with a common operating policy. This has implications for the technologies, types of networks, and business model. There will both be a general competition between telco-operators operating in the same geographic area, and a new set of operators offering services.

Only a few years back did data communication surpass the volume of voice traffic over intercontinental trunks. There is no evidence that voice communication ever will retake its dominant place. However, the income of the large telecommunication corporations is still dominated by phone revenue. The technology for handling voice as just another data communication flow is maturing. Such an imbalance between traffic and revenue cannot continue over a longer times span. The resulting shift in revenue flow may change both consumer behavior and the industry itself.

Fundamental problems

Although, there will be huge changes in the realization of the networks, the fundamental problems will remain the same. Radio resources will be scarce resources; efficient use of them will remain a problem. Similarly, the trend of improving the capacity on the different mediums will continue. However, the efficient utilization of a link, either over radio or over glass fibre, is only a subset of the necessary technology development. The fundamental problems of resource allocation, robustness, reliability, service discovery, and security will remain. Current technology address theses problems only to a limited degree and they are not able to scale to the required dimensions. An additional factor is the changing importance;

solution suited for experimental networks will not be acceptable for networks that are a fundamental part of the society's infrastructure.

The transition from a network where devices were leaf nodes, to networks where the devices have multiple connections, also have implications for the scope of the problems. Efficient use of spectrum resources remains an important link layer issue. However, robustness, security, and the efficient usage of resources across multiple link layer technologies appear as an additional problem.

The change in context, in terms of scale and importance, will require changes in how applications interact with the network; many services that today are handled in the terminals may have to be shifted towards the network. Some of the driving forces are the lack of trust among the users and the lack of technical knowledge. Areas that will require new solutions are among others: service creation, service discovery, authentication, authorization, charging and payment, and auto-configuration.

Mobility is today viewed as a separate service. In ubiquitous networks it will simply be an integrated part of the network itself. The access network itself and the movements between them should be transparent. However, the geographic location will be a new dimension of communication. The production, usage and privacy of location data will be a new set of fundamental problems.

New usage

The traffic patterns will change, from today's server client pattern with a centre of gravity around personal communication, towards machine to machine communication. There are initial research projects aimed at developing prototypes of small, and cheap sensors. They will have a limited lifetime and functionality. As a technology itself, they represent a formidable challenge. However they are the first indication of a wide range of sensors that will be an integrated part of the future networks. The sensors not only gather data, they are themselves also part of the forwarding network and the collection and fusion of data. Sensors therefore form a new class of networks that are fundamentally different from today's network. Clearly, they will require new set of solutions.

Changes in architecture

So far we have discussed the increased scale and new classes of technologies. However, we are certain there will also be changes in the architecture.

Today's paradigm is to view the network as a simple switching layer, where most of the processing is in the end systems. However, it is not a paradigm that is able to meet the challenges. The network itself will have to be more flexible, since a tailoring of services forces individual treatment of packets in a flow.

Active networks and programmable networks are two new paradigms that are being explored. In programmable networks, the services of the network itself can be reprogrammed dynamically. Active networks targets an even more dynamic environment, where each packet can carry its own rules for how it will be treated in the network. Apart from the realization of the technology itself, there are formidable challenges in terms of robustness, guaranteeing separation and independence of flows, and security.

The current Internet architecture needs to be modified to meet the requirement of future network architecture. It was formed under radically different context, where there was a common trust and a near unity in terms of goals. The future internet will be designed in a different context, where trust and common goals are not part the foundation. The network will therefore develop differently. There will be a need for network services to compensate for these weaknesses. We will see a need for trusted third party services, and configuration services.

A likely network design goal will have to be that the network must be flexible enough to concurrently accommodate groups with conflicting goals; the network will be formed to accommodate a tussle space.

An example is the need for privacy which may conflict with the societies need to monitor communications. Both are legitimate goals. However, the network architecture will need to be neutral and flexible to accommodate them both.

Some general trends in modern networks:

1. A layered architecture allowing for ‘call and connection separation’, or in general allowing applications requesting multimedia flows from an underlying network. This should also allow for different business domains to act in each layer.
2. Multimedia over IP (including real time aspects that allow voice and conversational services).
3. Mobility (user mobility, terminal mobility, session mobility, etc.) becoming an integral feature.

2020 outlook

- Networking will be viewed as a utility
- Multitude of networks and network services that self configuring and adapt
 - programmable networks
 - active networks
 - plug and play services
- Multitude of terminals
- Sensor networks will be an important part of the networks
- Fundamental problems remains, will still have to be solved on new scale
 - resource allocation
 - robustness
 - security
 - dynamic adaptability
- Many new business models
- Architecture that must accommodate conflicting goals

2. Networked Service and Application Platforms

Major trends

Open Platforms for Convergent Services. For the establishment of end-user services in a scenario with converging technologies, a growing interest to work with open interfaces and standardized service capabilities which opens the value chain and paves the ground for third party service developers and service providers has been shown. Although platforms for information services and telecom services traditionally have been different, there are synergies to be efficiently exploited by joint efforts from developers from the previously separated ideological camps. Convergent platforms will need to support both, which means that in the near future, peer-to-peer technology will be integrated with client-server technology, and security aspects (availability, confidentiality, integrity) will become increasingly important. In the subsequent stages, an even more close integration of end-user service design and development and network service provision will be expected, with increasing emphasis on quantitative Quality-of-Service (QoS), including information security, and negotiated Service Level Agreements (SLA).

Service and application architectures. To serve the open platform service provision scenario efficiently, application and service architectures and execution frameworks need to become far more dynamic than today, allowing services and applications to be incrementally developed and dynamically deployed while maintaining high levels of QoS with emphasis on availability and security. These requirements also call for the timely development of compositional methods and tools for dynamic and quantitative verification and validation principles, methods and tools. One possible future benefit may be the successful development and introduction of self configuring systems (autonomic systems in IBM parlance), reducing the perceived user interface complexity and establishing adequate trust, and thereby lowering the end-user acceptance threshold for new services.

2020 outlook

- Pervasive networking meaning that communication and collaboration aspects become central to most applications and services.
- Peer-to-peer becoming the dominant paradigm.
- Increasingly dynamic and self configuring services and applications
- A transition from pre-designed systems to dynamic frameworks supporting dynamic composition, deployment and adaptation.
- Increasingly diverse end-devices adapted to personal taste and context.
- Personalization, context awareness.

3. Dependability issues

Dependability issues cover the means to ensure that we may trust, i.e., depend on, the systems and their services, as well as the threats to them and their attributes. Dependability encompasses availability, reliability, safety and has issues in common with security. Issues addressed range from the component level (HW, SW and humans) via the system, services and user (individual and “corporate”) levels to the societal level, where vulnerability of the society and regulations in addition to the economics sets the requirements.

The general technological and techno-econo-social trends towards a community, where we apply and rely on information and communication services for most of our doings and welfare is described elsewhere. To what degree this will become a reality and the productivity and quality of life in this society, relies heavily on our ability to deal with the dependability issues. Initiatives are taken toward this objective, see for instance [AMSD].

History has shown the dependability must be built into the system and can not with success, be added afterwards. At the same time new successful technologies tends to move from a toy phase, where dependability is a minor issue, to a phase where they penetrates society and dependability becomes of extreme importance, cf. telephony, cars and planes as historical examples. This knowledge should guide the research into new “ambient intelligent” vision.

Major trends

For the core network and “centralized” service provision there are two easily spotted trends:

- The transport networks, with their currently high number of layers with poorly coordinated survivability functionality, will undergo a reduction in number of layers, and a coordination of actions in one management layer. The current IP/MPLS layer is foreseen to be this. However, failures have large consequences, which demands conservatism and slows the pace. Integration and differentiation discussed below is expected to have significant influence on the emerging solutions.
- The distributed processing technology has not yet shown its ability to provide a network wide middleware platform meeting the dependability, as well as other, requirements of the services to be supported. The current trend is more toward building special purpose server clusters with an architecture tailored to purpose and dependability requirements.

These means are primarily intended to deal with physical failures and are “under control”. Although they are very important to address, they do not pose *the* major research challenges for the next 15 years.

We move toward truly service integrated network¹, both for traditional and the new “ambient services”. However, the dependability requirements of the various services will differ widely, e.g. leisure browsing vs. emergency services and business critical services. There are also differences in the various user (groups) willingness to pay for the same service. A high dependability comes at a high cost (~ an additional 100%) with respect to equipment volume, development effort, management and maintenance. The co-existence of:

- services with highly differing requirements and
- system components with highly differing dependability performance

in the same system will rule the architecture and design of future system.

The steadily increasing complexity and dynamism of the systems and the logical and operational failures this incur, poses salient threats to dependability. In this context:

- Operational failures, has become the major cause of system failures. It is the failure category that has got the least research attention hereto, and which must be focused in the future.
- The technical systems per se have an increasing complexity in number of elements, in heterogeneity and in dynamism following the various kinds of mobility and new self organizing networking concepts (ad hoc, peer to peer, etc.).
- After the demonopolization of traditional telecommunication services, the convergence with information services, entertainment, the emergence of “private” networks and services, etc., the network operator and service provider environment has become highly complex. Multiple private and commercial entities are involved in the provision of services.

Providing dependable services in this highly complex and dynamic network is a major challenge in the future.

The self-organizing networks (ad hoc, peer to peer, etc.) have no authority that ensures their dependability or other QoS attributes. Hence, an application/system function that requires certain dependability must interact with the environment, assess its capabilities, and use the current computing and communication environment in a way that ensures that the requirements are met.

The introduction of so-called ambient intelligence by the combination of ubiquitous computing and communication as well as “natural” man-machine interfaces, extends the challenges outlined above. If this ambient intelligence shall take care of vital or safety critical issues of our lives, a number of unresolved dependability issues arises. An interesting scenario study is carried out in Part-1 of [AMSD].

A last trend that we should be aware of, is that the manpower, or rather “brainpower”, needed for (re)configuration, operation and maintenance of systems increases with the complexity of the systems and the sheer number of components. Hence, with the development foreseen, we will run out of brainpower and experience increasingly undependable services, unless there are major changes in how we deal with these issues.

Our view is that a salient aspect, in coping with the challenges outlined above, is the ability of system components to autonomously deal with issues necessary to meet dependability requirements. For instance: autonomy in (re)configuration and adaptation to a (potential hostile) environment, and, when necessary, perform replication, check-pointing, use of disjoint communication channels, failure detection and

¹ A corollary to this; you have no independent substitutes. If the “network fails” there is no alternative way to achieve what you want.

isolation, as well as keeping the “user” aware of the risks involved. This is in line with our ongoing research and other research initiatives like IBM autonomic computing [IBM].

2020 outlook

- Dependability-aware applications and system components negotiate with and probe their current {usage * environment} space to provide the desired trustworthiness of the tasks/applications/services to be delivered. If requirements cannot be met, the users are notified.
- New approaches for building systems for providing dependable services (and surviving) in highly dynamic, unreliable and potentially hostile environments will have emerged.
- An ultra dependable core ICT network infrastructure is maintained to ensure the vital communication needs of the society.
- Access networks are deliberately built to enable independent diverse access channels from end user devices to the core network to cope with critical services.
- Safety related activities (like automatic vehicle operation) *may* be considered integrated within the “ambient intelligence” environment.

4. Security issues

New paradigms of ubiquitous integrated computing and communication services has a great potential for improving the QoS of existing and new network based user services, but it also opens up for new threat scenarios, emphasizing the need for development of new security techniques, mechanisms and protocols. In particular, it is necessary to take into account that most, if not all, new services will be developed for wireless communications. The wireless environment is characterized by limited bandwidth, dynamically changing signal transfer quality and network structure, handheld terminals with limited power and processing capabilities and restricted user interfaces. All of these factors contribute to a more complex overall security environment, where many of the traditional means and measures will become inadequate. Adequate network QoS, including security, will be decisive with respect to establishment and maintenance of trust in fair and equal conditions for business partners in a future competitive economic system, to protect intellectual property and to contribute to the overall acceptance of the ubiquitous use of ICT in organizations, businesses and among the general public.

5. Systems engineering issues

Being deeply rooted in conceptualization and abstraction this area develops at a slower pace than the systems and network technologies.

Major trends

The most important trend is that model driven development is being broadly accepted as an important step forward. OMG with the UML languages and the visions of Model Driven Architecture, MDA has been an important catalyst to bring this about. Although similar to languages, methods and tools that have long been used in the telecom domain, the momentum and impact is much larger simply because of the wider target community.

The big challenge is to provide tools and methods that make all the visions come true.

In the wake of this, formal methods in general get more attention and become more focused. Verification and validation techniques are maturing to a level where they are applicable to increasingly large and realistic systems.

Meta-modeling is rapidly gaining importance and maturity, both as part of tools and as part of systems. Meta-data is becoming an integral part of systems.

As systems become less monolithic and more dynamic, formal methods must follow. The focus is shifting from complete systems towards frameworks and components that can be dynamically deployed, composed, adapted and discovered. Dynamic composition and adaptation poses new challenges.

2020 outlook

- Model driven development a reality.
- Method support to make components with meta-models that self configure.
- Diversity of domain specific methods and tools supported by meta-modeling.

References

David D. Clark and Marjory Blumenthal "Rethinking the Design of the Internet: end to end arguments vs. the brave new world."

[ACM Transactions on Internet Technology, vol.1 no. 1, August 2001.](#)

David D. Clark, Karin Sollins, John Wroclawski, Ted Farber. Addressing reality: An architectural response to real-world demands on the evolving internet. [ACM SIGCOMM 2003 Workshop, August 2003.](#)

http://bwrc.eecs.berkeley.edu/Research/Pico_Radio/

Beyond Calculation *The Next 50 Years of Computing*, ACM, ISBN: 0-387-94932-1

[AMSD] Accompanying Measure System Dependability (AMSD); Project IST-2001-37553; *A Dependability Roadmap for the Information Society in Europe*.
http://www.am-sd.org/default/show.gx?Object.object_id=KM-----000000000000061B&_app.page=show-DOC.html August 2003.

[IBM] Autonomic computing, <http://www.ibm.com/autonomic>

Communication Technology

Professor Geir E. Øien, Department of Electronics and Telecommunications, oien@iet.ntnu.no

Professor Nils Holte, Department of Electronics and Telecommunications,

Professor Steinar Andresen, Department of Telematics,

Professor Torbjørn Svendsen, Department of Electronics and Telecommunications,

Post doc. Mikael Hammer, Department of Electronics and Telecommunications

1. Broadband for All and Access Everywhere – Two key challenges towards 2020

Predicting the research focus and market development - over a 15-year period - of a field that changes as quickly as communication technology is not a task that can ever be executed with 100 % accuracy. Looking 15 years into the past, many of today's most important and penetrating communication technologies (e.g., GSM, WWW) were not yet in public use. Taking into account the accelerating pace of improvements and cost reductions with regard to enabling technologies, as well as the ever-increasing impact that communication technology has on the human lifestyle, it is likely that the future may hold even greater surprises than did the past. However, some basic trends are emerging, and we will focus on these. We will concentrate on the technology dimension rather than the market dimension, since this is where our main competence lies, and it is also the dimension that we believe can be predicted with greatest accuracy. Many different groups and organisations have made predictions of future technological developments; for instance, the objective of the *CELTIC* initiative [CEL1] is almost the same as the scope of the present document.

The two main communication technological challenges for the time period 2004 – 2020 can be summarised in the slogans *Broadband for All* and *Access Everywhere*. We should immediately note that this is *not* the same as *Broadband Access Everywhere*, which for several reasons we find not to be a realistic goal. In the present context we define broadband as *full* broadband, i.e., with enough bandwidth to implement a *Full Service Access Network* (FSAN) for domestic uses, including transmission of several high-quality TV signals. To achieve this, rates in the region 30 – 50 Mbit/s are needed. Contradictory, perhaps, to some predictions, it is our belief that broadband access according to this definition will mainly be wired (say, in 90 – 95 % of all cases). This is particularly true in urban and suburban regions, and is due to a lack of frequencies. While broadband fixed access has the potential for integrating all services in the same connection, there will in most cases probably be a wireless interface for the last few meters (< 50 m); hence the connection may still be perceived as wireless by the end user. The cost of long distance bulk data transport will be almost negligible; access, services, and content will be the cost-driving factors. The other main challenge, *Access Everywhere*, focuses on the need to make basic telecommunications services available in an (almost) ubiquitous and seamless fashion, even in areas with little or no infrastructure and for high-mobility nomadic users. Examples of such services are currently voice phone, e-mail, SMS/MMS, location and navigation services, emergency calls, banking transactions and information search via the Internet, etc. However, we emphasise again that these are not services requiring bandwidths that we have defined above as full broadband. Even when taking into account all anticipated technological advances and price reductions, there will still exist technological limitations, for instance due to cost and available frequency spectrum. There are fundamental trade-offs between accessible capacity, distance, degree of mobility, and cost per bit, which imply that full broadband access in sparsely populated areas with little or no infrastructure will always be expensive. A variety of wireless solutions, from satellite links to fixed radio access, will be important for implementing Access Everywhere.

In the following, we look in more detail into some of the important basis technology trends within telecommunication technology, sorted by discipline under the headings *Wired communications infrastructure*, *Wireless communications infrastructure*, *Networking technologies*, and *Services and content*. Note that security aspects are mostly left out, since there exists a separate INFOSAM-2020 working group and a corresponding position paper on such issues.

2. Wired telecommunications infrastructure

Both from a technical and an economical viewpoint, the implementation of universal fixed broadband access is the most important challenge within wired communication for the next 15 years. Such access may provide TV distribution, video, teleconferencing, video telephony, high-speed Internet access, electronic mail, high-speed data transmission, telephony, and a variety of other services. It may carry broadcast, multicast, as well as unicast services. There are three major alternatives for wired broadband access: Cable (cable-TV networks), DSL (Digital Subscriber Lines), and optical fibres (Fibre To The Home, FTTH).

FTTH will eventually remove all access rate bottlenecks, but until recently, the equipment costs of optical fibre systems have been too high for single users. Due to new technological developments like PON (Passive Optical Networks), optical transmission systems tend to become competitive, and further significant cost reductions are expected when mass production is started. However, installation is expensive if digging is necessary, in particular in urban areas. Full deployment of fibres will probably have to be undertaken during the normal cycle of renewing underground infrastructure, which takes 30 years or more. Consequently, only a part of the broadband access lines are expected to be optical fibres in 2020.

DSL is more flexible than *Cable*, and according to recent market analyses [PTO1], *DSL* is expected to be the main technology for broadband access at least for the first 5 - 10 years. The success of current *DSL* systems (*ADSL*) is due to reuse of existing infrastructure, by exploiting the existing copper cables all the way from the central office to the subscriber. In order to increase the bit rate of *DSL* systems, the loop lengths have to be reduced, by moving the fibre/copper interface closer to the subscribers. A significant further improvement of *DSL* technology is expected. The introduction of *DSM* (*Dynamic Spectral Management*) [DSM1] is predicted to double the bit rates for a given loop length compared to current *DSL* technology. The principles of *DSM* are basically the same as those of the *Multiple-Input Multiple-Output* (*MIMO*) techniques that are applied to wireless communications. There will be a variety of *DSL* solutions for different classes of bit rates, and up to 100 Mbit/s two-way connections are feasible (up to 500 m). Hence, *DSL* will be able to meet all foreseen needs within the private market segment.

Due to low cost and easy handling, twisted pair cable is today the main transmission medium for in-house communication, both in homes and offices. Typically, up to 100 Mbit/s Ethernet is provided over category 5 cables (or higher) for new installations. Complete cabling is usually installed as part of the building process so that changes will occur very slowly. Hence, twisted pair cables are still expected to dominate internal cabling except for some special high-speed applications.

Optical fibre technology has been present only for the last three decades. There has been an overwhelming development of the performance, and numerous new techniques have been invented. The most important breakthrough has probably been *Wavelength Division Multiplexing* (*WDM*), a technique that quickly increased the bit rate of a single fibre by one to two decades. The ongoing evolution is expected to continue [OPT1]. Further reductions of cost, and bit rates of 10 Tbit/s in one single fibre are predicted within 2020 [OPT2]. Ten Tbit/s is equivalent to more than 100 million simultaneous telephone connections, or one million full quality TV channels, and means that the maximum capacity of optical fibres may be regarded as infinite for most practical purposes.

Today, optical fibres are dominating almost all long distance communications; both national and international transport networks as well as metropolitan area networks. Undersea optical cables are already crossing underneath all oceans, and most of the world's cities are connected by fibre. During the

next decade, the capacity of the existing connections will be increased, and the fibre-optical network will be extended into a much finer grid, moving much closer to all subscribers.

3. Wireless telecommunications infrastructure

Within the next 10 - 15 years, wireless communications will become almost ubiquitous, and the ability to communicate seamlessly, nomadically, and with increased mobility will be evermore important. Relevant perspectives on the development of so-called 4th generation (4G) - or “Beyond 3rd Generation” (B3G) - wireless technology are described in [WWRF, 4GMF]. Also, the *International Telecommunications Union* (ITU) has developed a Recommendation - ITU-R M.1645 - for B3G [ITU1], and a corresponding list of key research areas [ITU2] providing guidance for researchers wishing to contribute to B3G development. B3G will represent a convergence between wireless access, wireless mobile, wireless LAN, and packet-division-multiplexed networks [4GMF], with one integrated terminal with one global personal number being able to access freely any wireless air interface. The radio transmission modules will be fully software-definable, re-configurable, and programmable (*software radio, cognitive radio*). The network will be heterogeneous with regard to available and interacting technologies, content, and services; re-configurable and adaptive; with different sub-networks interconnected and with different capacity and bandwidth requirements, channel characteristics, transmission technologies, processing capabilities, desired and available services, and sub-network architectures at the different levels. The technologies, subsystems, and bandwidths involved will depend on communication distance and will include Personal Area Networks (PANs) over ultra-short distances (< 1 m), broadband WLAN technology within small cells (< 50 m), a backbone cellular system for communication over medium ranges (< 1 km), all the way up to satellite links for providing ubiquitous access to more basic services in remote and sparsely populated areas.

One of the most important issues governing future development of wireless technologies, services, and business models is the regulation of the available radio spectrum. In the US, there is currently a move towards regulatory reforms that may eventually free up enough radio frequency (RF) bandwidth to significantly influence the development of mobile telephony and wireless Internet services [SPEC]. As an example, 85 MHz bandwidth previously allocated for analogue UHF broadcasting is being released for use in future mobile communications services. Digital terrestrial broadcasting (DVB-T) exploits the bandwidth far more efficiently, facilitating both more TV programs and extra bandwidth for other services. Frequency reallocation, spectrum leases, and spectrum sharing will allow the use of several new frequency bands for future mobile communications. These frequencies will find different uses depending on range and capacity needs, with the highest frequencies being reserved for high-capacity short-range communications.

The currently prevailing paradigm for the implementation of B3G is thus as follows: The core network will evolve toward a TCP/IP-based core network, serving a wireless Internet radio access based on packet switching for all services, including voice [NEW]. The major trend will be a move towards higher frequencies (above 5 GHz), leading to a nano- or pico-cell structure. This will make it difficult, if not impossible, to design the network on the basis of the standard cellular concept to provide continent-wide coverage. The network will evolve towards an ad-hoc wireless network, where base stations are installed where they are needed, and connected to each other in a self-configuring way to transfer TCP/IP traffic, similarly to present Internet wired architecture [NEW].

The overall structure may thus be one of multi-layered ad hoc networks rather than one rigid network structure. Distributed high-speed WLANs will serve local hot spots, inter-connected by the overlaid backbone cellular network and by a wired infrastructure. A multitude of wireless sensors will be embedded and integrated in the network for dynamic and intelligent interaction and communication between users and devices, as well as directly between devices. B3G will deliver much higher data rates and more diverse services than 2G and 3G systems, and the design of suitable radio interfaces has to take into account that the dominant load will be high-speed burst-type traffic; a great challenge for all existing

radio interface technologies. The following technology components, foreseen to become central in B3G radio interfaces, will ultimately enable a potential increase in bandwidth efficiency by a factor of 10 to 100 compared to today's wireless systems (GSM and UMTS):

- *Multiple-Input Multiple-Output (MIMO) technology*: This will be the single most important factor for enabling dramatically increased capacity and link reliability, as well as improved multi-user transmission, detection and interference cancellation. By spatial multiplexing and smart antenna processing, MIMO systems enable the same bandwidth to be reused many times within the same cell.
- *Cross-layer optimisation*: for joint optimisation of transmission schemes/protocols at the physical layer, link layer, MAC/network layer etc. – and thus a more optimal use of system resources.
- *Link adaptation*: for optimal capacity exploitation in temporally and spatially varying radio channels, by adapting throughput and power dynamically to predicted channel conditions.
- *Multi-carrier modulation and access*: for spectrally efficient broadband modulation and user multiplexing with easy handling of inter-symbol interference.
- *'Turbo'-based iterative algorithms and 'soft' information*: for improved efficiency/reliability of coding, multi-user detection, synchronisation, equalisation/estimation, interference rejection etc.
- *Ultra-wideband (UWB) radio*: Although UWB's ultimate importance has yet to be determined, it is currently a hot topic in the international research community, due to its potential for providing high-capacity, low-power transmission over very short ranges without interference limitations.

In addition, the ability to accurately measure, model, characterise, and predict the properties of many different types of radio propagation environments will continue to increase in importance with regard to successful design and optimisation of novel and more advanced radio interfaces. Furthermore, the development and commercialisation of new equipment for RF communication will call for the development of hardware with improved performance and reduced component costs. In order to avoid specification of communication standards that are difficult or expensive to implement in hardware, it is important to maintain a useful dialogue between engineers working at different levels of the standardisation process.

The impact of the 3G (UMTS) technology, which is currently under deployment, is still uncertain. B3G technology is still at the research stage and will not be available until after 2010. However, B3G has an enormous potential for increase of capacity, and is expected to have a very large impact on future mobile communications.

4. Networking technologies

As the convergence of the 4 "C" (Computers, Communication, Consumer electronics, and Content) is happening, we witness a proliferation of terminals and access forms. This puts stress on standards (formal as well as industry initiatives) in order to sort things out. In the home area this has resulted in initiatives such as Multimedia Home Platform [ETSI], [MHP], which extends the existing, successful DVB open standards for broadcast and interactive services in all transmission networks including satellite, cable, terrestrial, and microwave systems; and the Home Audio Video Interoperability initiative [HAVI], joined by eight of the largest producers of consumer products in the field.

We currently witness a massive transition from traditional circuit switched networking over to Internet based technology, this also concerns signalling and control. An SRI report [VOIP] recently termed Voice over IP, and the announcements of new hardware products complying with the SIP (Session Initiation Protocol) standard, a disruptive technology threatening traditional telecommunications. The key "disruptive effect" of SIP is that it puts the much of intelligence in the terminal instead of relying heavily on central resources, thereby undermining much of the importance of Intelligent Networks (IN) principles. It can be implemented both on a LAN connected device and on any DSL line.

Meanwhile the traditional incumbents and other heavy players are working busy to define the brand new network of tomorrow, the Next Generation Network (NGN), which rely on the IP protocol for transport of all services and is a combination of the Internet, and the ISDN/PSTN of today. Different vendors, providers, and operators, trying to cooperate to create new business opportunities, seek ways and means to establish both the judicial and technological interfaces needed to enable flow of services, content, and accounting information.

VoIP can also be introduced on top of wireless data solutions. However, this has so far not been done in the cellular/mobile business, with the given bit rate and latency limitations. The deployment of 3rd generation mobile systems, UMTS, has just started and it is to a large extent based on the IP protocol. Wireless LAN (WLAN) is a nomadic (session mobility) type of broadband systems that is already in extensive use. The IEEE 802.11 family of standards is still being extended, and it defines a number of different WLAN systems, which are all based on the IP protocol. There are many different initiatives working towards 4th generation fully mobile systems (B3G), and one of them is the IEEE 802.20 Mobile Broadband Wireless Access initiative, that is based on similar techniques as the IEEE 802.11 family. The complexity and cost of building networked applications can be alleviated by the use of a highly flexible, efficient, dependable, and secure *middleware*, which is system software that resides between applications and the operating networks. Middleware as a field of component-based systems will be further developed to cater for applications that can be efficiently distributed over a myriad of computing devices and communications networks. The self-organizing property of future applications will depend on ways of effectively managing distributed/networked objects or resources. Group communications mechanisms may be developed to meet these challenges. Middleware will also be developed not only to cater for request-response communications patterns, but also for event-based modes whereby complex patterns of events can be subscribed to, and trigger actions. Such solutions will help reducing notifications latency and save bandwidth. Application areas are surveillance services and wireless sensor networks, also on an ad-hoc basis [MIDD].

5. Services and content

The expectations with regard to use of text messaging (SMS) and Internet access (WAP) in the GSM network illustrate the difficulties in predicting the successful future services in the communications network. Still, some observations and predictions can be made.

Not contesting the placement of the intelligence, future services surely will have more intelligence. ICT will be ubiquitous and pervasive. Traditional approaches to developing services offered by communications networks have largely been specific to a single network technology, such as ISDN. This has tended to create problems in enabling services to operate across multiple network technologies. New principles for managing user-services and their QoS parameters on an end-to-end basis across many provider domains and technologies have to be developed. This is quite a challenging task, when mobility and handoffs between heterogeneous systems are considered. There will also probably be many separate domains or a hierarchy of services, as sensors and intelligent agents will aid our everyday needs and tasks, and continuous information access will be available regardless of location. Access to information services will be through a variety of devices, ranging from desktop computers and interactive TV to palm-sized wireless terminals. Available bandwidth will vary from ultra-broadband to narrowband, depending on location.

There will be a host of services yielding multimedia (video/audio/speech/data...), multimodal interfaces, entertainment, games, interactivity, browsing, streaming, location-based services, position estimation/navigation, personalization, mobility, nomadic communication, ambient intelligence, virtual reality, intelligent homes, IPR issues, and billing mechanisms. This gives rise to a growing demand for specialists producing intelligent middleware solutions that ensures interoperability and assured QoS. The authorities will increasingly provide information services through the Internet. The individual citizen will be required to return information, e.g. for taxation purposes, via online forms. Interactive online

processing of e.g. applications to local (and maybe national) authorities, guided by automated decision support systems, will simplify straightforward application processing. Commercial information services will continue to grow. In addition to information services, the entertainment sector will dominate commercially, ranging from low-bandwidth applications like electronic books to more demanding pay-per-view live broadcasts and movies, music, interactive games, and other multimedia content. E-commerce with virtual showrooms and on-line secure shopping is another important area.

The various services and content must handle variations in bandwidth and terminal properties (e.g. screen, keypad size and type) seamlessly, providing transparent services for different communications platforms. This poses particular challenges to information presentation and to the design and capabilities of the user interfaces. General basic design principles can improve overall quality of design, in addition to providing access for all with the purpose of minimising the digital divide. The different requirements and capabilities of e.g. a laptop with 100 Mb/s connection, and a palm-sized device with connection speed of 64 kb/s, will necessitate context-sensitive presentation and user input.

The human-machine interaction will exploit multiple modalities such as speech and sound, touch-sensitive screens and pull-down menus, graphics and video, and of course typed and hand-written text. The preferred combination of modalities will depend on the situation, personal preferences, and terminal size and connection speed. Speech and language technology (e.g. speech recognition and synthesis, natural language understanding) will play an important role for enabling user friendly, multi-modal user interfaces [COX]. Localisation issues, e.g. presenting the information in the user's language and understanding the written or spoken input, will be important for the universal citizen.

References

- [CEL1] The CELTIC initiative – Cooperation for a European sustained Leadership in Telecommunications, "CELTIC Purple Book, Part two: Technical Scope of the CELTIC Initiative," <http://www.celtic-initiative.org/Publications/purple-book.asp>.
- [DSM1] K. B. Song, S. T. Chung, G. Ginis, and J.M. Cioffi, "Dynamic spectral management for next-generation DSL systems," *Commun. Mag.*, vol. 40, pp.101-109, Oct. 2002.
- [MIDD] Gordon Blair, Andrew T. Compell, Douglas C. Schmidt, "Middleware technologies for Future Communications Networks," *IEEE Network Magazine*, Jan/Febr. 2004.
- [OPT1] OPTIMIST consortium, "Technology Trends (in Optical Communications)," <http://www.ist-optimist.org/tech.asp>.
- [OPT2] OPTIMIST consortium, "IST Roadmap for Optical Communications, May 2002," <http://www.ist-optimist.org/tech.asp>.
- [PTO1] Point-Topic, "Broadband Analysis," <http://www.point-topic.com/home/press/dslanalysis.asp>.
- [WWRF] Wireless World Research Forum (WWRF), <http://www.wireless-world-research.org/>.
- [4GMF] 4G Mobile Forum, <http://www.delson.org/4gmobile/intro.htm>.
- [ITU1] ITU Recommendation ITU-R M.1645, <http://www.itu.int/rec/recommendation.asp?type=folders&lang=e&parent=R-REC-M.1645>.
- [ITU2] "General focus areas for research and further study for the future development of IMT-2000 and systems beyond IMT-2000," <http://www.itu.int/ITU-R/study-groups/rsg8/rwp8f/docs/focus-areas.doc>.
- [SPEC] G. Staple and K. Werbach, "The end of spectrum scarcity," *IEEE Spectrum*, March 2004.
- [NEW] Description of the pan-European *Network of Excellence on Wireless COMmunications* (NEWCOM), <http://commgroup.polito.it/newcom/>, February 2004.
- [ETSI] ETSI ES 201 812 V1.1.1.
- [MHP] <http://www.hmp.org/>.
- [HAVI] <http://www.havi.org/>.
- [VOIP] The Case for Voice over Internet Protocol, Dec. 2003, D. Klemetz and M. Gold, Digital Futures, SRI, Business Intelligence.
- [COX] R. V. Cox et al., "Speech and language processing for next-millennium communications services," *Proc. IEEE*, vol. 88, no. 8, pp.1314 – 1337, Aug. 2000.

Computer Systems

Associate Professor Kjetil Nørvåg, Department of Computer and Information Science,
noervaag@idi.ntnu.no,

Professor Kjell Bratbergsengen, Department of Computer and Information Science,
Research assistant Haakon Dybdahl, Department of Computer and Information Science
Associate Professor Anne C. Elster, Department of Computer and Information Science
Assistant professor Torbjørn Hallgren, Department of Computer and Information Science
Research assistant Morten Hartmann, Department of Computer and Information Science
University lecturer Jon Olav Hauglid, Department of Computer and Information Science
Professor Lasse Natvig, Department of Computer and Information Science

The context of the computer systems workgroup incorporates research and development in all areas necessary for a complete working computer; from the design of computers and operating systems, to storage and management of data:

- Computer design and architecture
- Storage technologies database systems, and data management
- Operating systems, middleware, and distributed systems
- Computer graphics and visualization
- Supercomputing

A common feature of all these areas is the importance of thinking of the system as a whole, and the emphasis on both experimental and theoretical work. The working group has looked at future trends in the sub-areas listed above, and described how these changes will change systems as well as affect the use of the systems.

We will now describe trends in selected areas of computer systems. It should be remembered that computer systems is a very broad research area, and due to lack of space we can only cover some of the topics in this area in this position paper. We refer to the references given at the end of the paper for more detailed analysis of the summary provided in this paper. Finally, we will also outline some issues related to education in the future.

1. Computer design

From a technology perspective, today's computer system are based on semiconductor technologies. Development of these technologies have been providing a lower cost-per-function for the last four decades. Device integration, speed and functionality has increased while the power consumption per unit of computation has decreased [20].

Even though the core fundamentals of computer design will probably remain relatively unchanged for the coming 15 years, it is reasonable to assume that design methodology will focus on continuing efficient utilization of the available technologies in order to maintain momentum in the improvements mentioned above. This task is all but trivial.

Near term (-2010) challenges include power dissipation, low-power, power management, modeling and simulation of high-frequency circuits (5-40GHz) and nanometer structures, design-implementation productivity, design reuse, design integration and test as well as yield enhancement. For a long term perspective (2010-2018) the challenges include use of non-classical CMOS devices and beyond-CMOS devices, global wiring scaling issues, noise management and error-tolerant designs as well as modeling

and simulation of new materials [20].

Most pressing are perhaps the challenges of heating due to high power density (according to Vice President and Chief Technology Officer at Intel, the power consumption of microprocessors is expected to reach 18 kilowatts by 2008 [14] if power-saving techniques and power-saving programming is not employed) and memory latency [13] as well as low-power requirements for embedded mobile applications.

Intra-chip communication and synchronization is costly in terms of latency, power and area. Distributed and parallel designs are potential solutions to the problems, but the lack of programming paradigms for and understanding of such systems remains a major issue. Tighter integration of computation and memory is one example [8] of a more distributed low-level technique. The integration of different types of analog design on the same chip as memory and processor will also continue to increase, in particular in mobile devices [6]; systems on a chip provide low cost, low power and fast-turnaround times.

Research on biologically inspired architectures and methods [32, 19] in literature suggest several novel approaches to some of the challenges, such as error tolerance, adaptiveness, automated design, distributed architectures, massive parallelism etc.

Emerging technologies [19] such as nanoscale [26] devices and quantum computing may in some cases require radically different design methods. It is often claimed that fundamental barriers will very soon end the days of classical semiconductors. However, such barriers have been overcome several times in the past;

”Every time someone does a study on when we will hit those (scientific) limits, we don’t ever hit them. You hit other, more important, financial barriers,
— Lange, IBM Microelectronics.

2. Computer architecture

The Future Needs New Computer Architectures. Contemporary leading processor architectures are extremely complex, resulting in huge development cost and a problematic power consumption. This has been caused by at least three factors:

- The well known processor-memory gap (also called the memory wall), which has given deeper and more complex memory hierarchies occupying more and more of the chip area and power consumption.
- The processors have used a broad range of smart, but often complex, solutions to hide this memory latency, and to exploit a maximum of the parallelism available in the instruction stream (instruction-level parallelism (ILP)).
- The tremendous development in semiconductor technology has made room for such complex solutions and large caches on the processor chips.

To be able to continue offering improved performance in the future, beyond what can be given by ILP, multiprocessors and thread-level parallelism (TLP) must be exploited, especially in the area of throughput computing. This will give an increased interest in known and new ways of exploiting parallelism. The advances in technology are working against the memory wall problem, and the situation begs for new solutions. The combination of increased focus on parallelism and new memory solutions makes parallel computer architecture to an important topic in the future.

New Challenges in the HW/SW Interface -Rethink Compilers. Near future uni- and multiprocessors will exploit both ILP and TLP. We need a better understanding of its combination, and the future role of the compiler in this context. EPIC (Explicitly Parallel Instruction Computing) architectures exemplified by the Itanium Processor Family hand over more work to the compiler [22] and have fostered a renewed interest in compiler research [24]. The memory wall motivates for radically new solutions integrating

computations with memory (e.g. computational RAM [5]), but also a renewed interest in older techniques such as SW prefetching etc. [1]. This may result in new languages or libraries for parallel programming [27], also implying focus on the HW/SW interface.

Portable and Efficient Parallel Programming will be Needed, but is Still Missing. ILP and TLP are old, conservative and precautionary roads to parallelism. Till now, most programming models have followed one of two extremes: Either (a) designed for the programmer making it difficult for the compiler to exploit the HW, or (b) using explicit HW knowledge in language constructs and directives (pragmas) reducing programmer productivity and wrecking portability [27, 15]. Computer architects and researchers in programming languages and compiler design should cooperate to find new and portable ways to exploit the parallelism inherent in the applications.

Mastering the Complexity by Better Evaluation Methods and Tools. A major bottleneck in computer industry is mastering of the complexity in new designs. Evaluation methods and tools will continue to play an increasing role. Extensive simulations executed on HPC platforms will offer the designer powerful tools for design space exploration. Advanced solutions such as integrated performance monitors and HW performance counters will give new and better possibilities for performance studies focusing on the HW/SW interplay [27].

New Technology Gives New Solutions, but CMOS will Survive to Year 2020. The rapid advances in technology will make well established techniques less important (e.g. virtual memory). Reconfigurable HW makes radically new architectures possible. Quantum Computing is a hot topic that has attained a lot of positive press coverage, but some researchers have, based on fundamental physical arguments, stated that “for general-purpose computing, the classical computer will always perform much better than any possible quantum solution” [4]. CMOS works so well in rugged environments that we do not believe it will be replaced by “bioware” or quantum computing within the next 15 years [27].

Other Challenges to Future Computer Architecture. A major problem in the society of today is that computer systems are undependable and insecure [27, 21]. To improve fault tolerance and security we need more dependable HW but also new HW solutions and better HW/SW interfaces that help the SW designer to design better systems.

3. Operating systems

When considering the future of operating systems, one of the first problems is defining exactly what the concept of an operating system encompasses. This is made increasingly difficult as the border between applications, operating system and hardware becomes ever more unclear. For example, the Palladium security technology planned for the next Windows family (codename Longhorn) relies on specific processor extensions, and traditional third-party tools such as firewalls and antivirus programs will soon become included with most major operating systems. A different, but equally important, concern is the nature of the hardware platform the OS is set to govern. The huge span from small embedded systems via traditional desktop computers to large supercomputers indicates that a discussion about general challenges for future operating systems is difficult. As operating system development is mostly driven by industry rather than research, the following discussion is more geared towards what we think will happen rather than what we would like to have.

To first address the future development of core technologies (process handling, storage management, I/O, etc) for traditional computers, several trends are apparent. First of all, the continued development of external devices requires the operating system to keep up. Key features in this respect is plug-and-play and

self configuration. Another trend is improved support for large server architectures (as manifested by the new 2.6 kernel for Linux). As for file systems, the planned Windows Future Storage (WinFS) extension to the current NTFS represents an interesting merge of traditional file systems and database technology.

For commercial desktop operating systems, the user interface plays a vital role. Current developments show that improving usability and supporting new input and output devices will continue to be important areas. Tied to improving usability is the extension of the operating system to include extra services. Security issues (viruses, worms) and SPAM presents serious threats to current and future usability of computers and will have to be taken into account.

Light-weight operating systems can be expected to become much more commonplace within the next decade due to the increased use of non-traditional computers such as embedded systems, PDAs, TabletPCs, etc. These systems have a somewhat different set of requirements compared with traditional computers.

Key areas include wireless support, power saving, resource management, device control, real-time operation and reliability.

For more details we refer to [12, 28, 29, 35].

4. Storage technologies, database systems, and data management

The future and development of data management will depend heavily on available storage technologies. Currently, temporary data is stored in volatile RAM, while persistent data is stored on hard disk. Additionally, in applications with large data volumes some data is stored on tertiary storage like tape, CD, or DVD, and in the case of mobile devices, Flash memory is used. The future of non-volatile storage is hard to predict. For example, the death of tape storage has been announced repeatedly. Still, tape is heavily used, and capacity is increasing (for example, 1TB cartridges demonstrated in 2002, probably in market in a couple of years from now [25]). For many years now, variants of high-capacity random-access storage have been expected to be out on the market "real soon", but the development seems to take time. Examples of such technologies are IBM Millipede [18, 7] and Thin Film [30] which both should provide terabyte storage on a single chip. However, if these technologies make it to the market, the impact will be heavy:

- Everybody can afford to store enormous amounts of data. For example, much of the text now only available on the web can be stored on a small device, available to local search and mining. Large storage capacity also means that we do not have the same need for deleting data as before; we can more often store the complete history of all current and deleted data/files.
- Query execution, which is now quite mature and where now mostly incremental performance improvements are done, will have new opportunities. The most important performance bottleneck now is disk arm movement, and algorithms and strategies have as the primary goal to reduce the impact of this bottleneck. If storage technologies with fast random access make it to the market, other algorithms and strategies can be more appropriate to utilize from the new technologies.
- Data transfer between the storage chips and the rest of the system can become an even larger bottleneck than today; more capacity on both system buses and external networks will be needed. The bandwidth of the new storage chips will also be an issue.
- Actually managing and searching the stored data will be a problem. The problem of searching in large amounts of data has traditionally been solved by indexing. If we expect storage capacity to increase at a much higher rate than storage bandwidth and processing (searching) capacity, new indexing techniques might be necessary. They will probably focus less on space efficiency than current techniques. It is possible that more queries have to be pre-computed and stored together with the data in order to benefit from the storage capacity.

The traditional "core" database field has now reached a high level of maturity. Although database systems will continue to become faster, more reliable and more user-friendly, the real challenge in the

coming years will be to make them more suitable for the next generation of applications [3]. More general means a broadening from databases to data management in a broader context. This includes, for example, mobile databases (which often are relatively small), but the real challenge is the management and access of the large amounts of distributed heterogeneous data available in one form or another on the WWW. The importance of interoperability will in this context continue to increase, partly solved by the use of XML as storage and transfer format (more and more business data are now stored in, use in 2004 estimated to 60% [34]) and the use of web services (already available) and the semantic web (if it succeeds).

It is also possible that the architecture of many database (and data management) systems in the future will be very different from current technologies. We will probably see heavy distribution, no fixed home location of data, and access based on authorization and roles. Many systems will be based on peer-to-peer-like architectures (it should be mentioned that such systems already exist in a primitive form: for example, a system like Kazaa already indexes 4600 TB of data in 600 million files, distributed between over 3 million users).

Reliability and availability might be achieved by replicated storage. However the problem of guaranteeing correct data has to be solved by different means. More and more data are stored over a long period of time. Data are eroding both physically and logically. No storage media holds data for ever. To avoid loss of data, programs for rewriting or transferring data to new media has to be established. Logical erosion is more subtle. Formats might change, however this is taken care of by the rewriting programs. Changing of terminology, interpretation, etc is more difficult. Data should be interpreted according to the time of registration. This means handling aging ontologies and historical data correctly.

5. Computer graphics and visualization

The science of computer graphics has evolved since the 1960s and a solid scientific and methodical base has now been established. Only recently, however, have sufficiently powerful computers at affordable prices been brought to the marketplace making it possible for researchers to better understand the data produced by their applications.

Convenient and intuitive interaction with virtual reality systems is an important challenge. It will increasingly be a demand for tools of interaction that are small in physical dimensions and that can respond to hand and body gestures, to eye movements, to voice and even maybe to facial expressions. Virtual environments of today are mostly restricted to visual expressions. Spatially correct sound is entering the scene. It is expected that smell and other sensations will be part of the future virtual environment. Especially two market segments will be goals for new development, namely the home market and the mobile platform market.

Many aspects of life will be influenced by the virtual reality technology. Only some examples can be mentioned. A very important application will be in education at all levels both for demonstration purposes and for training. Another aspect of the same nature employing the same methods and techniques is entertainment. The word edutainment is increasingly being accepted as a common denomination of these two applications. Virtual heritage is a similar application.

Virtual presence will make it possible to virtually be in the same room for instance for a meeting even if the participants are on different continents. Research is being conducted in universities. One may also envision that future television will be stereoscopic and immersive. Maybe the fashion catalogue will contain models that the customer may put on a model of her own body and then animate to get a very realistic impression how it will look.

Solutions of mathematical problems and complicated technical and scientific problems may be visualized in a virtual environment. Surgeons may train in virtual environments and perform remote operations using virtual reality techniques.

Among the basic research topics are methods for modeling and animation, parallelization, fast photogrammetry and model synthesis, methods and tools for interaction as well as the physical equipment for virtual environments.

6. Supercomputing

Over the past decades supercomputing systems has evolved from having a focus on large dedicated vector machines like the Cray series of the early 90's to large parallel distributed memory systems based on COTS (commercial off-the-shelf) parts using cluster technologies. The cluster-based systems are generally large groups of PC or workstation processor boards networked together over a local area network or switch. Currently, the "Top 500 Supercomputers" list [31] is dominated by such machines. This trend is expected to continue for the next several years [2].

In addition, as previously predicted [9, 10], traditional HPC (high-performance computing) will probably no longer drive processor technology. Instead, the consumer market – which focuses on the gaming and entertainment world – will have a even larger impact the processor market. This trend is already apparent with Intel's February 2004 announcement of their new Pentium processor codenamed Prescott, which includes 13 new instructions several related to gaming and video [33, 23]. Supercomputing software and high-end library and compiler research will hence have to find new techniques to utilize these features, rather than new architectural features being driven by traditional supercomputing applications.

Another trend in supercomputing is the move towards GRID computing – the idea of providing computing power the way electricity is delivered: plug in your computer and run your application without regard to where the processing power and data storage may be located [11]. This requires break-through in inter-operability (incl. login, compilers, etc), security, as well as streamlining user policies across national borders. Large companies such as Hewlett-Packard and IBM [17] are currently commercializing grid technologies.

Tighter coupling with visualization will become increasingly more important in order to make sense of the Petabytes of data streams generated in the future. See the visualization section for more details.

Finally cluster and GRID computing as well as high-end graphics are also evolving in the area of MMOGs (Massively Multi-player On-line Games) where thousands, and eventually millions, of players play games in a virtual world. E.g. Sony's EverQuest has currently over 460 000 subscribers that are supported by a server farm each serving up to 2000 players simultaneously. [16]

7. Computer systems education in 2020

Education in studies related to computer systems in general aims at giving a broad introduction to the whole area during the first part of the study, with further specialization later. This is likely to be the case also in the future, with curriculum based on what is actual technology at the moment.

References

- [1] 36th Annual International Symposium on Microarchitecture, <http://www.microarch.org/micro36/>.
- [2] G. Bell and J. Gray. What's next in high-performance computing? *Communications of the ACM*, 45(2), 2002.
- [3] M. L. Brodie. Panel: Future Directions of Database Research -The VLDB Broadening Strategy, Part 2. In *Proceedings of VLDB*, 2000.
- [4] Classical vs. quantum computers: And the winner is ... <http://www.newsfactor.com/perl/story/22109.html>.
- [5] Computational RAM. <http://www.eecg.toronto.edu/~dunc/cram/>.
- [6] D. Edenfeld. 2003 technology roadmap for semiconductors. *IEEE Computer*, 37(1), 2004.
- [7] E. Eleftheriou. A nanotechnology-based approach to data storage. In *Proceedings of VLDB*, 2003.
- [8] D. Elliott. Computational-RAM : The case for SIMD computing in memory. In *Workshop on Mixing Logic and DRAM: Chips that Compute and Remember at ISCA '97*, June 1997.
- [9] A. C. Elster. High-performance computing: Past, present and future. In *Proceedings of PARA 2002*, 2002.
- [10] A. C. Elster. The next level of simulations: Extreme computing. In *SIMS 2002*, http://ntsai.oulu.fi/Tapahtumat/SIMS_CallForPapers/1016a.pdf, 2002.
- [11] I. Foster. The grid: A new infrastructure for 21st century science". *Physics Today*, 55(2), 2002.
- [12] Future of operating systems: simplicity <http://www.computerworld.com/softwaretopics/os/story/0,10801,77335,00.html>.
- [13] P. P. Gelsinger. Microprocessors for the new millenium: Challenges, oppotunities, and new frontiers. In *Proceedings of the International Solid State Circuits Conferance, ISSCC*, pages 22–25, 2001.
- [14] P. P. Gelsinger. Power play. *Communications of the ACM*, 45(6), 2002.
- [15] M. Goudreau. Portable and ef? cient parallel computing using the BSP model. *IEEE Trans. on Computers*, 48(7), 1999.
- [16] <http://everquest.station.sony.com/about.jsp>.
- [17] IBM Grid Computing home page <http://www-1.ibm.com/grid/>.
- [18] IBM Millipede, <http://www.zurich.ibm.com/st/storage/millipede.htm>.
- [19] International Technology Roadmap for Semiconductors. Emerging Research Devices, 2003, <http://public.itrs.net>.
- [20] International Technology Roadmap for Semiconductors. Executive Summary, 2003, <http://public.itrs.net>.
- [21] Jim Gray Touring Award Lecture. <http://research.microsoft.com/~Gray/talks/>.
- [22] V. Kathail, S. Schlansker, and R. Rau. Compiling for EPIC architectures. *Proc. of The IEEE*, 89(11), 2001.
- [23] Intel's Prescott, Intel's Extreme Edition, and AMD's Athlon 64 3400+, <http://www.linuxhardware.org/features/04/02/02/1654234.shtml>.
- [24] Open Research Compiler project. <http://ipf-orc.sourceforge.net/>.
- [25] J. P. Scheible. A survey of storage options. *IEEE Computer*, 35(12), 2002.
- [26] Technology roadmap for nanoelectronics (microelectronics advanced research initiative), 1999 <http://www.cordis.lu/esprit/src/melna-rm.htm>.
- [27] The Future of Computer Architecture Research, Panel, ISCA-2003. <http://www.comparch.ncsu.edu/future/>.
- [28] The inevitable future of our operating systems -part 1 http://www.thetechlounge.com/article.php?directory=os_future.
- [29] The Microsoft future according to Bill Gates, <http://www.ameinfo.com/news/Detailed/33384.html>.
- [30] Thin Film, <http://www.thinfilm.se/>.
- [31] Top 500 supercomputers <http://www.top500.org>.
- [32] A. M. Tyrell, P. C. Haddow, and J. Tørresen, editors. *Evolvable Systems: From Biology to Hardware. Fifth Int. Conf., ICES 2003*, volume 2606 of *Lecture Notes in Computer Science*. Springer, 2003.
- [33] O. Valmot. Lynraske prosessorer. *Teknisk Ukeblad*, (7), 2004.
- [34] S. J. Vaughan-Nichols. XML raises concerns as it gains prominence. *IEEE Computer*, 36(5), 2003.
- [35] Windows "Longhorn" Activity Center, <http://www.winsupersite.com/longhorn/>.

The organization of human affairs / The creative human

- **Information Security and Safety**
- **Information Systems**
- **Language Technology**
- **Learning with ICT**

Information Security and Safety

Professor Stig Frode Mjøl̂snes, Department of Telematics, sfm@item.ntnu.no
Professor Torbjørn Skramstad, Department of Computer and Information Science,
torbjorn@idi.ntnu.no Professor Tor Stål̂hane, Department of Computer and Information Science,
stal̂hane@idi.ntnu.no Ph.d. Student Siv Hilde Houmb, Department of Computer and Information
Science, sivhoumb@idi.ntnu.no, Senior Researcher Lars Bratthall, DNV Research

System dependability and the future

“Where a calculator like the ENIAC today is equipped with 18,000 vacuum tubes and weighs 30 tons, computers in the future may have only 1,000 vacuum tubes and perhaps weigh only 1,5 tons.”

---Popular Mechanics, March 1949, p. 258

1. Introduction

This paper discusses the trends expected with respect to system dependability towards 2020. System dependability is trying to answer an important question, namely: "how can we make the computer systems and communication networks that we increasingly rely on more dependable?" Our lives have become critically dependent on the correct operation of these systems as we surf the web, fly in airplanes, drive our cars, manage our finances, and even heat our homes. The traditional concerns of the dependability community (e.g., inadvertent faults, errors, and failures) have now been enlarged by the massive connectivity provided by the Internet to include malicious exploitation of imperfect systems and networks, and intentional cyber-attacks on them. How can we build systems that are not vulnerable to such threats, systems that users can depend upon in transportation, financial, and e-commerce sectors?

It is said that it is easy to be overoptimistic with what you can obtain in two years, but you normally underestimate the amount of changes that will occur in a ten year period. Trying to draw up the trends towards year 2020 will therefore be an extremely uncertain prediction, although, some trends can be observed. From history we know that things develop “linearly” for some time, and is then influenced by disrupt changes, due to new inventions, new technology and new markets

Dependability is the extent to which a (often critical) system is trusted by its users. By a dependable system we usually mean a system that has at least some of the characteristics reliability/availability, security, safety, robustness etc. We therefore start with a short description of what we mean by these terms.

2. Safety and security - what is it?

By a *safe system* we mean a system that can perform its intended function without failures that might lead to dangerous or catastrophic situations. Safety is concerned with ensuring that the system cannot cause damage irrespective of whether it conforms to its specification. By *security* we mean the ability of a system to protect information against unauthorised access and intentional misuse. Security includes characteristics such as authentication, authorisation, confidentiality, integrity, privacy, non-repudiation and availability. Information is an asset which, like other important business assets, has value to an organization and consequently needs to be suitably protected. Information security protects information from a wide range of threats in order to ensure business continuity, minimize business damage and maximize return on investments and business opportunities.

Security is an essential prerequisite for reliability, availability and safety, and thus also dependability.

3. Some economical, political and technological trends

It is expected that the economic growth will continue in the Western World throughout the period, and that the trend towards deregulation and privacy will continue. People will have more spare time and will spend more time on travel, health care etc. There will also be a pressure for increasing productivity, and for shorter development times. Increased efficiency.

We will experience an increase in the terror threat towards computer intensive systems as these get more and more important for the modern societies.

Globalisation: more and more software and hardware will be developed in low cost countries such as India and China. The same will apply to Computer Operations Centres and Research as well as many office based services [1].

4. Some expected trends related to safety

Safety has for a long time been considered seriously and carefully within many traditional industries such as process control, oil and gas production, transport systems such as railways, and airplanes, and nuclear power plants. The systems were safe because technology developed slowly, and one could learn from near accidents and accidents. During the last 10-20 years computers and software has taken over more and more for electromechanical devices (such as relays and servos) and hydraulics. Traditionally, special computer hardware has been used in such applications. The software development has been very burdensome with focus on well-behaved software development processes, extensive testing and validation. The ongoing trend within these industries is that general purpose cheap processor chips are used more and more, and we see an increased use of COTS (i.e. using standard operating systems such as Windows XP and Unix) is forced in use mainly for reducing development times and costs. We also see that the systems are being more and more distributed, often over public communication networks such as Internet. Examples of this is "Remote Oil Production" where more and more of the offshore installations are being operated from onshore via complex/multifunction communication networks. It is also a trend towards interoperability – systems that are operated in close cooperation. We also see that

propulsion and control systems onboard large marine vessels are developed for being operated by 40-50 crew in order to be safe enough, while modern ships mostly are manned with less than 10 crewmembers.

The use of computers and software is rapidly increasing within industries such as automotive and medical device industry. These industries have a little tradition for thinking safety related to computers and software and are much less mature than for instance the aircraft industry.

Traditional industries have to learn how to utilize modern technology to increase the efficiency without reducing the safety. The systems get more complex due to the flexibility to develop functionality through software and because several systems are integrated and due to distribution of components geographically and network-wise. In [4] Charles Perrow, argues that the conventional engineering approach to ensuring safety--building in more warnings and safeguards--fails in complex systems because systems complexity makes failures inevitable. He asserts that typical precautions, by adding to complexity, may create new categories of accidents. He further argues that new safety analysis methods have to be developed in order to manage increasing complexity and increasing speed in technology evolution.

Typical questions are:

- How to assess safety of complex distributed systems?
- Can the uncertainty related to use of COTS and software of uncertain origin in such systems be solved by redundancy and diversity. How can we assess the diversity of COTS hardware and software? As more and more components are developed in low cost countries, how can we be sure that the integrated system is safe even if each component has high quality?
- How complex systems are we able to assess for safety? (Top Issue).
- What are the impacts on human computer interfaces related to safety for such systems?
- How to obtain safe communication over the Internet? Can we be sure that the response time is short enough? Are we sure the messages will reach the receiver fast enough and with integrity?

5. Some expected trends related to security

[2] has some interesting reflections:

- There will be a worldwide broadband network based on fiber optics, communication satellites, cellular and microwave. Face-to-face, voice to-voice, person-to-data and data-to-data communication will be available to any place at any time from anywhere
- Ubiquitous availability of computers will facilitate automated control and make continuous performance monitoring and evaluation of physical systems routine
- Our homes will be integrated systems, smart and smarter integrated houses, and plugging into the communications network

[3] says that:

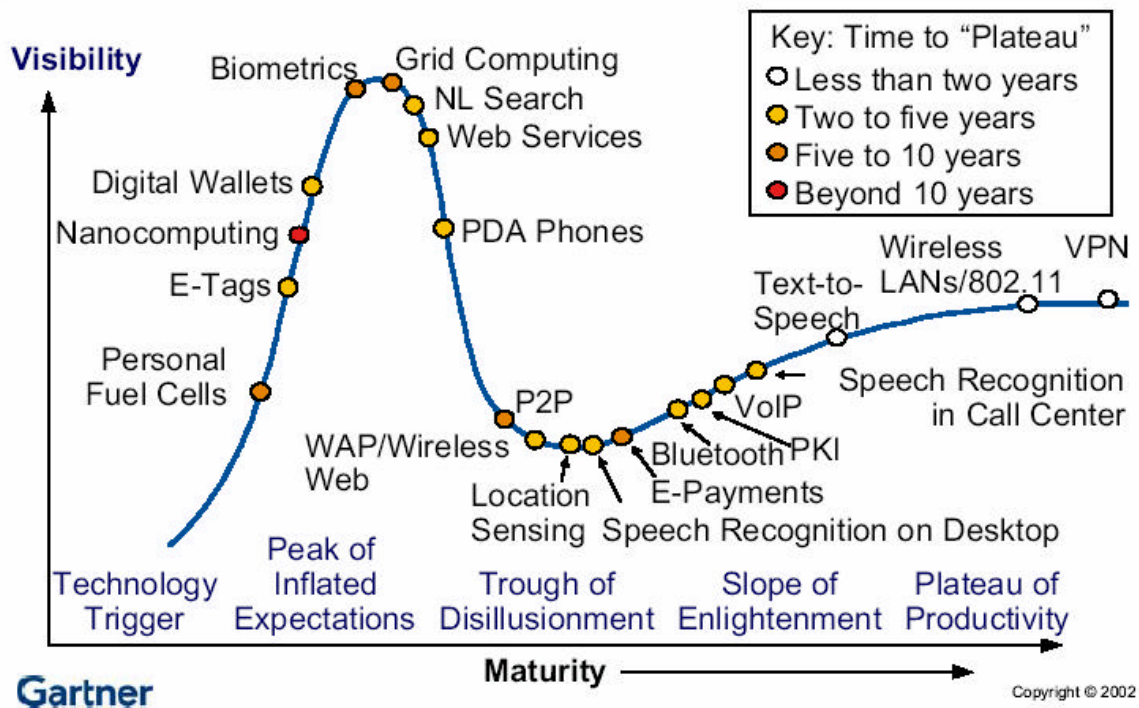
- Emerging core computing technologies (e.g. Quantum cryptography, nanotechnologies, hybrid inorganic/biology computing) will not disrupt general-purpose, semiconductor based computing through 2010. It is expected that the introduction of quantum computing will disrupt the evolution of cryptography (with probability 0.6).
- At the same time as general-purpose computers in the world are interconnected via the Internet billions of miniature intelligent devices already inhabit the world, with their number increasing faster than the human population. The next ten years will bring new capabilities: a)

many physical objects will be coded and therefore will become uniquely identifiable, b) intelligent devices will be embedded in many physical objects, and will be networked via the (mostly) wireless Internet. Gartner calls this the “Supranet”

- In the next ten years, a single, advanced integrated IP network will be handling the majority of the world’s communications needs. This converged, broadband, intelligent network will extend well beyond voice and data, local and long distance, supporting an ever-widening array of services, and blurring distinctions among networking, computing and applications. Driven by e-business requirements and facilitated by technological advances such as e-switching and next-generation satellites, the increasing externalization of networking will give rise to an environment where applications, content and data reside in the network and are dynamically handled by network service providers in real time, without user intervention.
- Content is the core of business transactions, publishing and entertainment. The diversity, volume and effect of content will grow such that during the next 10 years, we will experience unprecedented levels of interactive content, driving valuable revenue streams for publishers, corporations and media companies. Content will be accessible almost anywhere via broadband and CDNs. The effects of this will stretch from the corporation into the home, as rich media content will be stored and managed in a digital asset management system. High-value content will have to be delivered securely. In the enterprise, the ongoing digitization of more and more information, including document authorization, will ease in fully digital process management for more and more business processes.
- Mobility represents the next major business and technical discontinuity facing large enterprises. While the PC and Internet revolutionized communications systems, mobility will revolutionize information flow that will affect business users, customers and partners. By the year 2007 more than 60 % of the EU and US population aged 15 to 50 will carry or wear a wireless computing and communications device for at least six hours a day, by 2010 this is expected to be more than 75 %. By 2010, less than 5 percent of global wireless subscribers will be using true 4G technology (0.6 probability), but 15 percent will be using components of a full 3G architecture based on LAN/WAN integration and IP applications (0.7 probability).

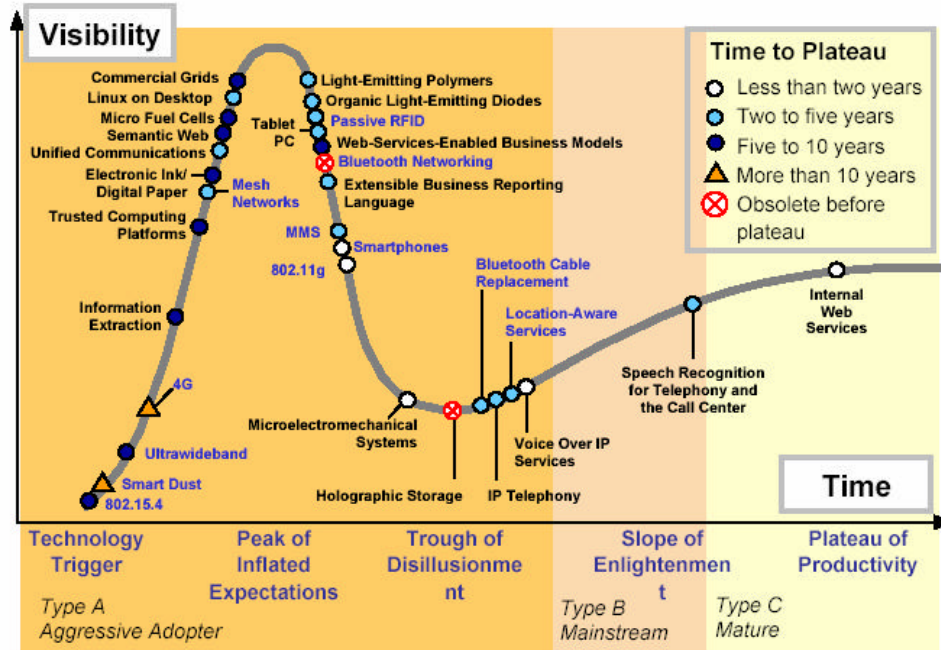
Gartner [5] uses a so-called Hype Cycle Graph to illustrate how new and promising technologies go through various phases of development. The following curve shows Gartner’s curve (from 2002) for ICT related technologies. It shows the typical progression of a technology from overenthusiasm through a period of disillusionment (because of the inevitable failures arising from inappropriate application), to an eventual understanding of the technology’s relevance and role.

Hype Cycle 2002



A similar graph from 2003 is shown below:

Gartner Technology Hype Cycle

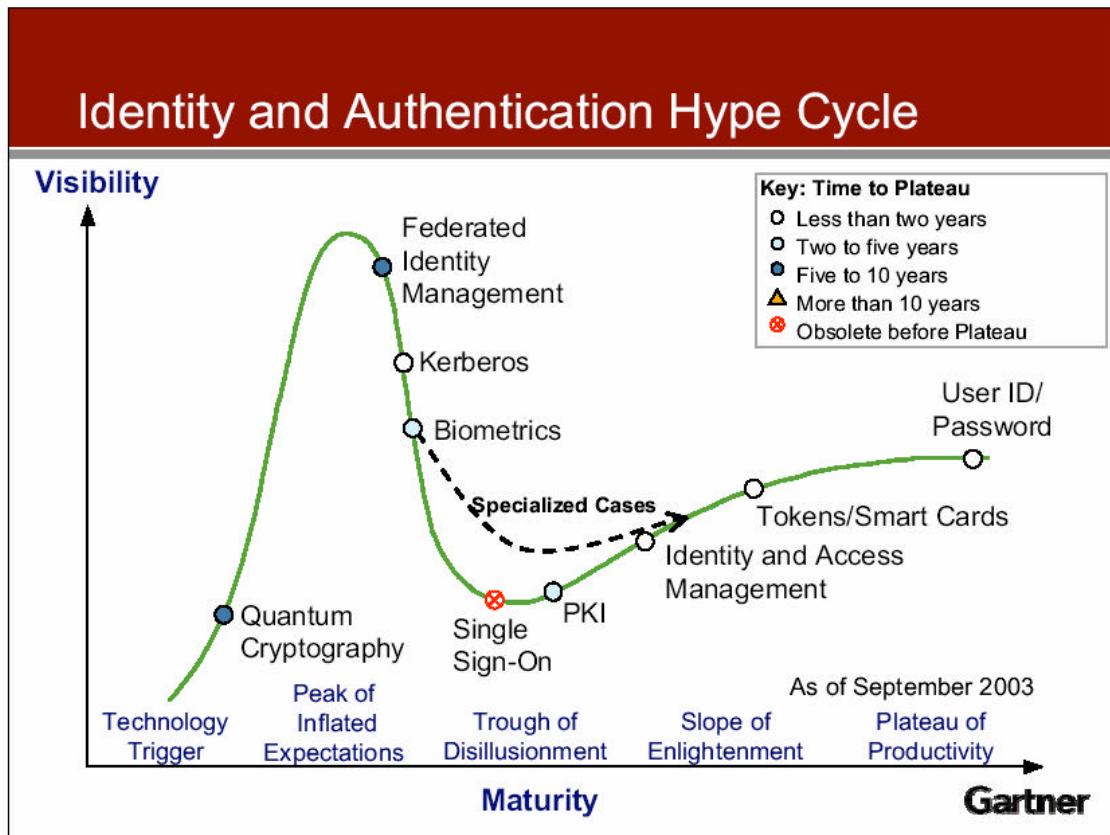


IT Wisdom

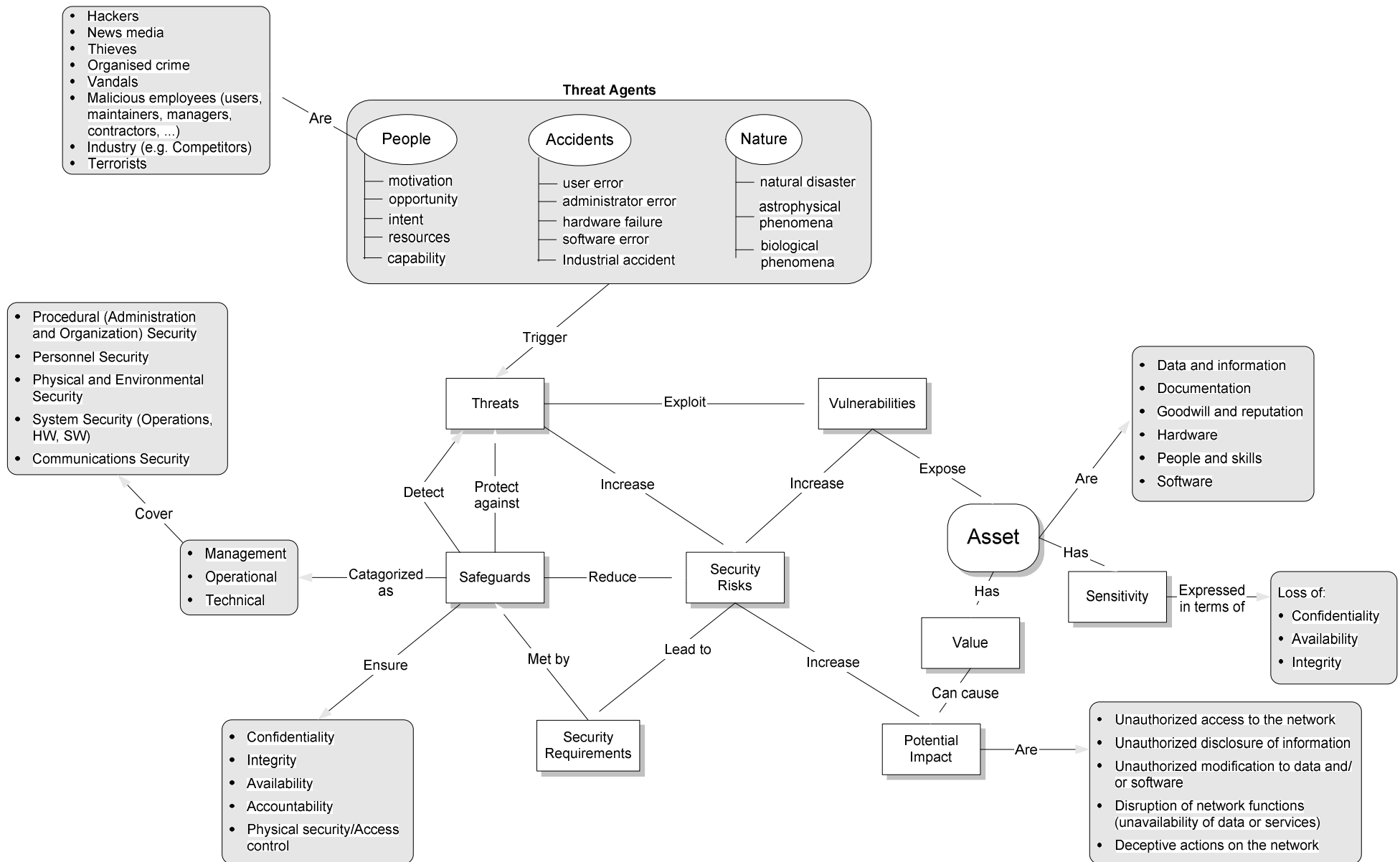
Entire contents © 2003 Gartner, Inc. All rights reserved.

Gartner

Gartner has also presented a Hype Cycle graph for Identity and Authentication related technologies. It is from 2003 and presented below:



The figure on the following page shows the relations between threats, threats agents, assets and vulnerabilities and safeguards as seen from the ISO17799 [7].



By 2020 we may see a scenario like this:

- A larger amount of the world's information assets will be digitized and access able via the Internet.
- Internet traffic will increase enormously and most human beings in the Western world will have continuous and direct access to Internet, mostly via wireless communication devices.
- Criminals and terrorists will have at least the same access to Internet as most people. They will be better educated, more sly and will have the newest technology available. Example: On a tour on a large passenger vessel in the Oslo Fjord it turned out to be possible to get control of the ships machinery from a passenger laptop with wireless communication in the vessel's conference room.
- The amount of computer crime and possibly terrorism will increase significantly. Statistics from CSI/FBI shows that incidents have nearly doubled each year in the period 1997-1999. This is expected to continue. The police need to have access to and measures to avoid and investigate crime and terrorism in the Cyberspace in the same way as in the public space today (cyberspace forensics).
- More and more of the information stored and transported via the Internet will be of significant value to individuals, organisations and nations. Use of the Internet for communication related to safety applications will increase significantly.
- More information has to be encrypted in order to protect sensitive information both while stored and while transported, but emerging computing technologies such as quantum computing may make contemporary encryption algorithms infeasible.
- Biometric access technologies will dominate authorisation and authentication

So what are the main questions we have to face:

- Research on better and new types of encryption technologies
- Develop better risk analysis methods to better understand and manage the real risks in a more complex setting
- Research on more reliable software development methods.
- How to back-up enormous amounts of data distributed across the Cyberspace.
- Which impact will the introduction of RFid have?
- Research on how safety critical systems should act when disconnected from partnering systems – transmission of “plan ahead safe states”.

References and Background literature

- [1] Dagens Næringsliv, 3-4 April 2004, p. 33m
- [2] J.F. Coates, J.B. Mahaffie, A. Hines, in “Scenarios of US and Global Society Reshaped by Science and Technology, 2025”
- [3] Gartner Group, “Emerging trends and Technologies scenario, 2003”
- [4] Charles Perrow, “Normal Accidents” (1999)
- [5] DNV Research: Technology Outlook 2003 – DNV Research’s assessment of development trends, DNV 2003.
- [6] Gartner Symposium ITXPO 2002
- [7] ISO17799: Information Security, 2003
- [8] Rebecca T. Mercuri, Security Watch: Superscaled Security, Comm. of the ACM, March 2004/Vol. 47, No. 3, p. 15.

Acknowledgements

Stig Frode Mjølunes (NTNU) has been a valuable discussion partner during the production of the first draft of this paper and giving comments to the first draft. Useful comments and improvement proposals have been supplied by the following individuals: Tor Stålhane (NTNU), Siv Hilde Houmb (NTNU), Lars Bratthall (DNV Research), and Jan Hovden (NTNU).

Information Systems

Professor Jon Atle Gulla, Department of Computer and Information Science,
jon.atle.gulla@idi.ntnu.no

Researcher Ingunn Amdal, Department of Electronics and Telecommunications,

Professor Monica Divitini, Department of Computer and Information Science,

Post.doc Helge Langseth, Department of Mathematical Sciences,

Associate professor Hallvard Trætteberg, Department of Computer and Information Science

Information systems deal with information-processing software systems that help users carry out tasks by providing relevant information and computational support. They generate, manipulate, analyze and store both structured and unstructured data, and their functionality automates, controls, monitors, or supports the tasks at hand. Central to the information systems area is the inter-dependence of software, humans and various artefacts. This complex mixture of technological and human resources is today crucial in many domains, and information systems integrate and apply more specialized technologies to meet individual or organizational needs.

Information systems include prescriptive and empirical methods and approaches for developing and analyzing both software systems in isolation and the wider concept of information systems. The area includes sophisticated user-computer interfaces, methods for mining, retrieving and managing information, support for collaboration and coordination among computers and people, adaptations to individual and organizational needs, mobile and distributed systems, knowledge discovery and management, and architectural and design considerations. Indirectly, their impact on social and legal structures, as well as related ethical issues, also falls within the study of information systems.

The information systems area is characterized by a wide range of research fields, and there are few of today's research topics that will survive the next decades. There are however trends and tendencies in this area that span several research fields and are likely to dominate both research and industry the next 20 years. The following trends are of particular importance:

- ❑ Interaction – body and environment
- ❑ Intelligent support
- ❑ Universal personalized services
- ❑ Ubiquitous computing
- ❑ Networked communities and social implications
- ❑ Emerging development processes

Each of them is discussed below.

1. Interaction – Body and Environment

With *multi-modal interaction* we refer to multiple input and output modalities as well as multiple “terminals”. The most common output modalities today are screen (text, graphics, and video) and audio. Rapid development in display technology and speech technology will result in a more diverse set of output modalities in the future. An important aspect for future output is the presentation of information to non-experts (e.g. medical information directly to the user not through physicians). Typical input devices (controlled by user, i.e. not a sensor) today are keyboard and various pointing devices (buttons, mouse, and stylus). Speech is an obvious additional modality that is becoming more important as the technology improves. The same is also the case for eye tracking, gesture tracking, biometrics, etc.

Sensor technology today includes devices for GPS, temperature, humidity, light, etc. We will see a rapid development in especially medical and military sensors. As sensor technology improves, they will be a natural input/output modality for information systems.

The terminals as we know them will probably be on the museum in 2020. All items will have sensors, and computers will be everywhere (“disappearing computers” (ISTAG 2004)). With *wearable computing* the sensors may track human condition and behavior (when you enter a zone – indoors or outdoors, biosensors) and be an integral part input and output devices.

The next 20 years we expect substantial development in sensor technology, speech technology, displays, and audio/video research. Speech technology will improve (at least languages with enough funding, small languages like Norwegian will need special care) and can be taken into account (see Language Technology position paper).

Whereas we today have mostly one interface per sensor or input device, the future will bring us interfaces that combine information from several sensors and input modalities. Essential concepts here are fusion, fission and dialogue management. As interfaces get more complicated the human centered design gets more important to let the user be in control of the interaction. Ideas from human-human-interaction will find their ways into human-computer interaction, though there are situations in which humans may want computers to behave differently from humans.

Many of these input technologies and sensors have an inherent uncertainty (speech recognition errors, the interpretation of a medical sensor regarding a specific human condition, noise in transmission, etc.). This calls for context-based interaction (intelligent dialogue), where we allow delayed decisions using probabilities and models from the different modalities as well as dialogue histories and user preferences. Fusion and fission will be more complex, and for a user friendly interaction the semantics of the input/output should be coordinated with the semantics of the information providing part of the system. The system must be able to detect and repair misunderstandings.

2. Intelligent Support

There is a class of intelligent systems that makes use of published data to analyze issues, help users solve problems or come up with recommendations. The quality of these systems grows as more data is added, provided that the system has the technology needed to deal with the data made available to it. Except from document management systems, this class also includes many learning systems, expert systems, and even tools for finding and using web services.

Tools for analyzing structured and numerical data are available today. As more data is added to these systems, the tools come up with more precise and more reliable analyses that help the users understand the issues and make decisions. For unstructured data, like text documents or images, the situation is however somewhat different. Methods for information retrieval, information extraction, text mining, and categorization and clustering have been introduced the last few years, though there are still fundamental problems in this area. The methods make use of statistical approaches like normalized word frequencies and simple morphological information about words to analyze document content. These methods address syntactic aspects of the data and does not go into the semantic logical part. Semantic approaches are still not reliable enough, and the processing limitations prevent the methods from being employed in many information management and learning systems.

The next 20 years will see huge amounts of data being published and information systems that are far better integrated than today. With this plethora of data, the potential of these intelligent systems will rise dramatically (PC Magazine 2003). At the same time, both processing capacity and semantic analysis methods will improve to the extent that the systems can start dealing with the semantics of this data rather than various statistical and morphological aspects. Methods and approaches from linguistics, statistics, and logic will be tightly integrated and used on-the-fly to help users find and interpret the information they need.

As envisioned by the Semantic Web initiative (W3C 2004b), the systems of the future will distinguish between the documents or services themselves and knowledge about these representations. Since all documents, services, rules or records are given formal logical definitions, the systems can answer questions directly rather than just pointing the user to sites that may have the information they

need. The systems can combine data from different sources by reasoning about the different semantic representations. They can also cooperate on an ad-hoc basis and make deductions about the result from analyzing the semantic description of each system. Every time a new document, service or rule is added, the system has access to another piece of knowledge that will improve the system's ability to provide the user with intelligent support.

Structured domain knowledge in the form of ontologies and models will be important to interpret the data in the right context. There will be large domain ontologies available, but they will also build ontologies for individual or more specialized use. The users may maintain their own ontologies that reflect their experiences, interests, and preferences. The same data may mean different things to different users in different contexts, and the use of these ontological representations make sure that the available and sometimes fragmented and ambiguous data is meaningful and relevant to the users. Ontologies will also allow the user to specify only what he wants to achieve and let the system be responsible for figuring out how to do it by retrieving and combining the necessary data and services.

3. Universal Personalized Services

Through continuous interaction with computerized systems, human users will bit by bit release information about their interests and preferences. They leave traces that can be collected and used to assess their past behavior. By collecting and analyzing these bits of information, automated systems (intelligent artificial agents) will help to find relevant information, carry out the appropriate transactions, and provide a computer-supported environment that is tailored to the individual user. This development brings (at least) two technologies forwards:

- Personalization
- Decision support systems

Both these technologies are already finding their commercial niches today. First-generation personalization systems built on techniques like collaborative filtering, clustering, and classification algorithms can be found on the internet. Limited decision support systems (a.k.a. expert systems) have also emerged, and are today capable of automating relatively simple decision making in limited domains. Examples include screening of medical images, classification of spam-mail, troubleshooting of electronic systems, etc.

Emerging technologies related to data warehousing and data mining make distributed learning and decision support available. In the near future these technologies will be more integrated in everyday life. Information overflow through digital television and an increasingly voluminous internet makes filtering and personalization of the available information important. Smaller and cheaper computer systems make decision support systems available everywhere. Venetian blinds with built-in intelligence and t-shirts with an opinion on which other clothing it prefers to be washed together with are but two examples.

4. Ubiquitous Computing

When Apple Research did a study on the future of computing, they tried to analyze how people in the future will carry out their tasks, which tools they will use, and how they will use them. They came up with a new concept, *ubiquitous computing*, that describes the way users, their tools and network resources will be working together:

"The Concept of computers as things that you walk up to, sit in front of and turn on will go away. In fact, our goal is to make the computer disappear. We are moving towards a model we think of as a 'personal information cloud'. That cloud has already begun to coalesce in the form

of the Internet. The Internet is the big event of the decade [...]. We'll spend the next 10 years making the Net work as it should, making it ubiquitous."

- Frank Casanova, director of Apple Computer Inc.'s Advanced Prototyping Lab

The underlying goal is to move computers away from the user's attention, where they are applied subconsciously and embedded in many small and highly specialized devices within our continuously changing environment. This vision of ubiquitous computing has been around for a while, but it is only now starting to take shape. Ubiquitous technology is a new computational paradigm that is important to embrace for assuring the widespread adoption of IS in the society. We envision that in the next decade ubiquitous applications will get out of research laboratories and into everyday life of people. This is made possible by a growing interest of the industry in these applications and in the advances at the level of connectivity. However we expect this adoption to be problematic for some years to come because of limitations in the technology, hardware and running costs, but also and foremost for the worries of the general public around these technologies. This will force the developers of these applications to take seriously into consideration the social and ethical implications that the developed technology brings along.

Ubiquitous computing has implications for research fields like

- software architectures,
- computer-supported cooperation,
- interaction design, and
- multimodality

5. Networked Communities and Social Implications

The European Union is actively working with its research funding programs to bring, within 2010, "IST [Information Society Technologies] applications and services to everyone, every home, every school and to all businesses." Even if this objective can be reached until 2010, we expect that the consequences at the societal level will take much longer to be fully visible. In the nineties the Web has allowed the emerging of web based communities. At the beginning of the millennium we are witnessing the emergence of new forms of social aggregation, called by someone "social mobs", supported by widespread usage of mobile technologies. We can expect that pervasive usage of IS better integrated into everyday life of people will lead to different social communities, probably more ad-hoc and volatile. These new forms of aggregation, bypassing traditional forms of governance and control, are promoting new and richer forms of participation and, potentially a more democratic participation of individuals to community life. Everybody is online most of the time (we are only offline when we choose to be), both in private and professional life, and we are part of numerous networked communities within which we communicate and coordinate our activities.

If on one side the widespread of IS can lead to a *more democratic* society, on the other it also brings the risk to *more control* on citizens and *lack of privacy*. We expect that the next fifteen years will be largely characterized by the struggle between these opposites and the consequent definition and continuous revision of different policies:

- Communities will support local democracy through the spread of relevant information and gathering of opinions. Networked interest groups will be formed to fight for particular issues (which both may strengthen and weaken democracies). Different groups of the population will be strengthened and weakened by the move to online democracy, e.g. education, economy and accessibility will affect how online you are.

- Our constant use of networked communication devices will leave electronic traces all over the place. A complete trace will reveal a lot about our private life, and care must be taken to protect our privacy and avoid intrusion, like spam and little brother surveillance. End-users must be able to understand and control how information about themselves spreads in the network. At the same time, both the citizens and institutions must be protected against network-based criminal activities, like identity theft, electronic trespassing and illegal spreading of protected information, through both technical and legal mechanisms (like coordination of national and international standards and laws).

Also from a legal point of view different challenges are arising. It is interesting for example to remind here the discussion about the distribution of legal responsibility connected to the exchange of copyrighted music shared over the net thanks to peer-to-peer infrastructure, or about the collection of an unprecedented quantity of individual data for different reasons, ranging from medical record to terrorism prevention. The sensitive nature of this information arises a set of issues related to security and ethical usage and sharing. We expect these issues to be only the top of the iceberg, with the full scale of the problem visible in the ten or fifteen years.

The adoption of Information Systems by new users will bring along a new set of requirements:

- “Design for all” has been a recurrent slogan for the last decade. In the next decade this has to become common practice.
- More flexible systems that are able to evolve with their users, but also be flexible enough to adapt to usages different than expected
- Clearer distinction between inscribed behavior and policies

From a process point of view this requires more interdisciplinary work and more awareness in the design team of the social, legal, and ethical implications of their systems. We also need to pay more attention to emerging patterns of usage. For example, SMS and chat started as tools for teenagers and are now getting into the workplace. We believe that the introduction of the technology in different environments will bring new patterns of usage that can possibly be adopted in other contexts, with a positive contamination between different sectors. Adoption must be based on acceptance and consensus, which will dramatically change the marketing and deployment strategy of IS.

6. Emerging Development Processes

Most large IS development projects take more time and cost more than expected, and often still do not satisfy their users. This highlights two issues that developers are struggling with: 1) Understanding their

users, their environment and what they really need, and 2) controlling the process of gaining this understanding and developing the relevant technology. Apart from reducing development time and costs, the focus the next 20 years will be on making development projects more predictable, flexible and less vulnerable.

The major trend will be to reduce the size of each project, evolve rather than revolve by support continuing change rather than major overhauls. The relationship between the customer and developer will change into closer cooperation, with more shared learning and reduced risk. The customer will not want to rely on only one developer organisation, so the process must be open, have smaller steps and be easier to change.

Systems will be based on open standards, not just for infrastructure and middleware, but also for vertical components established by specific segments. Component-based development will have matured to the point that systems development implies less programming and more configuration of standard and generic components. In some sense, every system will already exist, it will just be a matter of describing the needs and perform the necessary customization. The customers can take care of most of their own needs through visual composition of functionality and scripting.

To be able to continuously evolve and migrate systems, systems will be self-contained and self-describing (no difference between system and documentation) using executable models of function and design. Since systems will be transparent, developers know what to modify to achieve a certain effect, they can verify the consequences of the changes (ripple-effect). New methods for model integration and migration will support gradually changing running systems.

As many important "mainstream" technical problems have been solved, there will be less emphasis on making new technologies feasible and more on understanding, describing and fulfilling end-users' needs. The developers will need to switch roles between viewing technology as a system and seeing it through the eyes of its users. This kind of "stereo sight" will become even more essential in the years to come as no single perspective on technology, society and design alone is sufficient.

7. Challenges in Education and Research

The anticipated development of this field requires strong multi-disciplinary approaches combined with deep specialist knowledge. The educational paradigm should be related to how students can learn to work with people of different backgrounds while contributing with their particular competences. This is more important than teaching each student fractions of several disciplines.

The need for switching roles between viewing technology as a system and seeing it through the eyes of its users, requires a new generation of computer science professionals and researchers with a training in technology supplemented with courses from the social sciences, design, marketing and psychology. The focus of research at the same time must move from studying technology de-contextualized in the lab to also studying it "in the wild".

Strong analytical and theoretical skills are needed to push the technology forward and take full advantage of the technological basis is established. This also means that education has to be regarded as a continuous activity that has to be encouraged by employers and society at large.

A cooperative and open approach to the development and use of information systems poses new challenges to education and research. On the one hand, we need to adhere more to standards and be more efficient in applying reusable or generic software rather than developing it ourselves. On the other hand, projects will be more loosely organized, and we need to learn how to organize our work at an individual basis and get accustomed to more ad-hoc ways of cooperating with other people.

Finally, we need to encourage creativity. Whereas we in the past have improved the technology by imitating and automating manual approaches, we will soon be in a position to do things that we could never imagine before. We need to stimulate innovation and creativity both in education and research.

References:

- Horn, P. (2000). The Future of Information Technology, IBM Research.
[<http://www.cs.colorado.edu/events/lectures/horn/horn.pdf>]
- ISTAG (2004) *Information Society Technology Advisory Group in EU*. [<http://www.cordis.lu/ist/istag.htm>]
- ISWC (2003) *IEEE International Symposium on Wearable Computers*. [<http://www.cc.gatech.edu/ccg/iswc03/>]
- ELSNET (2004) *The Human Language Technologies roadmap*. [<http://elsnet.dfki.de/>].
- PC Magazine (2003). *20 Hot Technologies to Watch*, July 1 2003.
[http://www.pcmag.com/print_article/0,3048,a=43573,00.asp].
- Rheingold, H. (2002) *Smart Mobs: The Next Social Revolution*; Perseus Books.
- W3C (2004a) *Multimodal Interaction Activity*. [<http://www.w3.org/2002/mmi/>].
- W3C (2004b). *Semantic Web*. [<http://www.w3.org/2001/sw/>].
- Weiser, M. (1991) The Computer for the 21st Century, *Scientific American*:September 1991, pp 933-940.
[<http://www.ubiq.com/hypertext/weiser/SciAmDraft3.html>]

Language Technology

Torbjørn Nordgård, Department of Language and Communication Studies,
torbjorn.nordgard@hf.ntnu.no

Torbjørn Svendsen, Department of Electronics and Telecommunications,
Erik Harborg, SINTEF Information and Communication Technology
Knut Kvale, Telenor Research and Development

1. What is Language Technology?

A broad definition of language technology could be *the branch of information technology which deals with natural language information*. This information is typically conveyed in two forms – written or spoken. For this reason we might define two main branches of language technology: speech technology and textual language technology, although the latter is integrated in many speech applications. Today we find language technology in many products – spell checkers and grammar checkers in standard office suites, word prediction in cellular phones, dialogue systems between humans and machines (for instance automatic reservation via telephone), rough machine translation in internet browsers, automatic dictation (conversion of speech to text). These technologies have evolved into commercial products over the last 10 to 20 years. In the years to come we expect applications like search in speech databases (e.g. radio programs), automatic transcription of colloquial speech and TV programs, among others.

2. Textual Language Technology

The total amount of available text for all purposes increases extremely rapidly. Text production is to an increasingly larger extent required for documentation purposes – in businesses as well as in governmental domains, for instance within the European Union. Various aids for text production have come to the market in the last 15 years, in particular text processing tools with “enabling” facilities like spelling and grammar checkers. Automatic dictation systems (speech converted to text) are expected to be improved in the years to come, both due to improved hardware and language resources, at least for the language societies which give priority to use of native languages in technological environments. Thus the possibilities for producing text becomes better and better.

This state of affairs leads to the well-known problem of sorting out relevant information for some particular task from the huge amount of irrelevant information, or get rid of relevant information of bad quality. This problem is of course closely tied to the content of documents in the searching domain. Content analysis thus becomes increasingly more important in the areas of information management and so-called data mining. Content extraction based on more or less simple word count is the current state of the art – words are counted and often compared to normal word distributions, and unexpected distributions will in general give reasonably good clues to document content, albeit that the precision can be much better in theory if reliable information about documents can be used to classify them, cf initiatives within work on so-called semantic web. The problem for the word count approach is basically computational power when the amount of text is huge, for instance the entire world wide web. Given Moores law[10] we can expect that conventional indexing can be performed on ordinary PCs some years ahead from now, but more extensive content analysis algorithms can also be put to use, for example document indexing based on hierarchical semantic or logical structure. Resources like WordNet[11] and EuroWordnet[12] are important in this context. Entire document collections can be parsed and assigned sophisticated semantic content labels, and we can imagine search engines being able to perform logical deductions over their own updated formal models of the real world, as it manifests itself in written sources.

There are, however, multilingual aspects involved here which are difficult to handle today. Relevant documents might exist in different languages. Queries should thus be translated to other

languages if desired by the users, and relevant information in other languages should ideally be translated into the user's preferred language, perhaps as translated summaries if desired. These ideas are by no means novel – they have been quite clearly formulated in research projects, most notably within the EU. But results are poor, so far. High quality products of this type do not exist, and the best commercially available translation system is good old Systran[13] initiated in the early sixties. Unfortunately, the languages available for automatic translation do not include e.g. Norwegian. But the emergence of statistical methods in machine translation paves the way for robust translation between Norwegian and other languages, provided that manually aligned bilingual text corpora are made available for research and industry.

15 years ahead we foresee that industry and consumer markets will expect that textual applications like the ones hinted at above will be available with Norwegian language as a naturally integrated part. Research within language technology in Norway should give priority to development of formal and in particular statistical methods in textual analysis, both with respect to information management and automatic translation. Another extremely important aspect is creation of language resources for research and application development, which most naturally appears to be a task for the government, see conclusions in the report Norwegian Language Bank[14].

3. Spoken language technology

Twenty years ago, automatic speech recognition was basically treated as a signal processing problem whilst text-to-speech synthesis was very much rule-oriented, based on expert linguistic and acoustic-phonetic knowledge. During the past two decades, paradigm shifts in the technology and a move to much more complex task domains have lead spoken language technology to become a multi-disciplinary research area where insights from fields like linguistics, statistics, acoustics, phonetics, computer science and signal processing are required in order to adequately address problems involved with enabling computers to understand and produce natural spoken language. Continued progress in spoken language technology will depend on exploitation of all available, relevant knowledge sources, making close cross-disciplinary collaboration a necessity.

The past two decades have witnessed a convergence of the methodologies employed for solving the various problems in speech processing to an extent that today, data-driven statistical modeling techniques constitute the backbone of both systems for recognizing and understanding speech, as well as for production of speech from texts or concepts. This development is expected to continue, and will facilitate better interaction and utilization of knowledge between the different application areas.

The development of the core technologies of automatic speech recognition (ASR) and text-to-speech synthesis (TTS) will to a great extent govern the evolution of speech enabled systems such as conversational user interfaces, spoken dialogue information systems and multi-modal user interfaces which integrate core speech technologies with e.g. natural language understanding, decision support, data mining etc. Although substantial improvement in the performance of the core technologies are to be expected, it is doubtful whether it will be possible to design systems that are close to human performance outside narrow application domains. HAL 9000, the intelligent computer in the movie *2001 A Space Odyssey*, will be at least a few decades delayed.

In ASR major obstacles that need to be overcome mainly concern various aspects of robustness. These aspects include on the acoustical level sensitivity to ambient and electrical noise and variations on voice characteristics, on the lexical level variation in speaking rate and style and pronunciation variations (including non-native accents), and on the language model level task domain dependency, out-of-vocabulary words and variations in phrasing. As has been witnessed over the past years, the hardware development has facilitated the application of ever more powerful models and algorithms, and it is expected that this will contribute to an evolutionary development of the performance of speech recognizers. The inclusion of higher level knowledge sources, e.g. semantics will most likely be necessary to improve the accuracy of both dictation and advanced spoken dialogue systems, but this will not obviate the need for better modeling at the lower levels. The perhaps most challenging problem facing speech recognition is the handling of natural, spontaneous speech where

current systems typically exhibit word error rates in excess of 30% (transcription of conversational speech).

Text-to-speech (TTS) synthesis has evolved from being rule-based to being more and more data-driven. This was initially the case for the sound generation, but statistical methods are gradually being employed also in the linguistic layers of the synthesis engine. The quality of synthetic speech is evaluated for intelligibility and naturalness, and current state-of-the-art TTS is mainly lacking in naturalness. Improved methods for sound generation and intonation modeling will over the next 15 years lead to speech synthesis engines producing highly intelligible and natural sounding synthetic speech. Among the challenges facing TTS development in the near future, we will see development of tools for efficient production of new synthetic voices and dialects; investigations into the characteristics and modeling of emotional speech; methods for improving the flexibility of TTS engines e.g. with respect to speaking rate and style.

The task of speaker recognition is to identify *who* is speaking. The sub-domain of speaker verification is the task of verifying a speaker's claimed or assumed identity, usually for access control. The main issue governing the prospects of success for speaker verification systems will be related to reliability, in particular against impostors. The issue can currently be handled fairly reliably, but can be questioned if future TTS systems provide means of fast and inexpensive methods for generating high quality synthetic speech from a limited amount of speech data from any given individual. Telephony based and networked speech technology applications will be influenced by the projected adoption of the internet protocol (IP) for telecommunications [1]. Packet oriented speech transmission by voice over IP (VoIP) will introduce artifacts due to speech coding distortion, packet loss and round-trip delay that need to be addressed in future systems. Perceived speech quality is important to user rating of interactive voice response (IVR) systems. IVR is currently the largest application area of speech technology, and annual revenues are expected to grow to 3 billion USD by 2008 [6].

Spoken dialogue systems[2] are inherently dependent on the performance of the core modules of speech recognition and synthesis. However, correct recognition of every word is not necessary for the human/machine interaction as long as key words and phrases are correctly interpreted. Well designed dialogue strategies including error handling are essential for well functioning dialogue systems. Automated response generation, i.e. the generation of spoken responses from a response *concept* will improve the usability and flexibility of spoken dialogue systems.

Multi-linguality[5] will be an important issue. English may be the *lingua franca* (sic!) of modern society, but the national languages will prevail. However, foreign names and phrases will need to be understood by a speech recognizer and well pronounced by a speech synthesizer. Language identification is essential to switching a foreign language speaker to the correct language module in a spoken language interface. Another important aspect is the possibility of utilizing existing foreign language resources (speech and text data, statistical models) for improving the performance of speech technology systems in another language.

Multimodal[4] user interfaces integrate different I/O-modalities like speech, sound, graphics, text and user gestures such as pen/stylus on touch sensitive screens, eye movements and body movements. Such user interfaces are particularly interesting for small, mobile terminals where the terminal size precludes the use of standard I/O (keyboard, mouse...). Speech recognition and synthesis will be an important part of these user interfaces. Important issues include how to effectively integrate the various modalities such that the different information channels can be synchronized and disambiguated for maximal efficiency in use.

4. A successful Language Technology in Norway

Successful language technology infrastructure in Norway in 2020 will provide product types like the ones mentioned above, and an industry which is able to utilize national and international potential in language technology. Certain groups of disabled people will hopefully find help in language technology products like automatic dictation (blind people, dyslectics, and persons with serious physical handicap) and naturally sounding, personalized speech synthesis (people being unable to speak properly).

But there is at least one major obstacle. Given that statistical methods will continue to dominate speech technology and at the same time will become more and more important in textual language technology, it is extremely important that Norwegian language resources are made available for research and commercial use in the years to come. If not methods and solutions developed for other languages, in particular English, will be predominant and the pressure to use English instead of Norwegian will be stronger. Hopefully Norwegian authorities will make sure that relevant resources are being created.

The Norwegian Research Council has intensified research on language technology, most notably via the KUNSTI program[7] which funds two major projects in the areas of machine translation[8] and dialog modeling with spontaneous speech[9], in addition to a number of smaller projects. Unfortunately, these projects, and most other projects financed by KUNSTI, suffer from the lack of language resources. There are fundamental research issues that need to be resolved in order to improve the general performance of language technology. Some of these issues will be language specific. Continued support of language technology research, maintaining a critical mass of activity on a high international level is a vital prerequisite for the success of Norwegian language technology. We assume that the KUNSTI program will be continued after the end of the current project period in 2006, and hopefully with a better Norwegian language infrastructure.

References

- [1] R.V. Cox et al: "Speech and Language Processing for Next-Millennium Communications Services", Proc. IEEE, vol 88, no. 8, pp. 1314-1337, August 2000
- [2] A.L. Gorin et al.: "Automated Natural Spoken Dialog", IEEE Computer, pp. 51-56, April 2002
- [3] R. Rogoff: "Voice Activated GUI: The Next User Interface", Proc. IEEE International Professional Communication Conference, 2001 (IPCC 2001), pp. 117-121, Oct. 2001
- [4] S. Oviatt: "User-Centered Modeling and Evaluation of Multimodal Interfaces", Proc. IEEE, Vol. 91, No. 9, Sept. 2003
- [5] A. Waibel et al.: "Multilinguality in Speech and Spoken Language Systems", Proc. IEEE, vol 88, no. 8, pp. 1297-1313, August 2000
- [6] Datamonitor: "Voice Automation – Past, Present and Future", White Paper for Intervoice, July 2003
- [7] KUNSTI web-site, <http://www.program.forskningsradet.no/kunsti/>, accessed March 2004.
- [8] LOGON info at <http://www.emmtee.net/>, accessed March 2004.
- [9] BRAGE web-site, <http://www.tele.ntnu.no/projects/brage/>, accessed March 2004.
- [10] Moores law, <http://www.intel.com/research/silicon/mooreslaw.htm>, accessed March 2004.
- [11] The Global WordNet Association, <http://www.globalwordnet.org/>, accessed March 2004.
- [12] EuroWordNet web-site, <http://www.illc.uva.nl/EuroWordNet/>, accessed March 2004.
- [13] SYSTRAN, <http://www.systransoft.com/>, accessed March 2004.
- [14] [Consolidating and Increasing the Availability of Norwegian Human Language Technology Resources](#), report October 2002, <http://www.sprakrad.no/sbank2.htm>, accessed March 2004.

Learning with ICT

Researcher Leif M. Hokstad, NTNU LIKT, Program for Learning with ICT.

Leif.Hokstad@svt.ntnu.no

Bente Brunsgard, Telenor Bente.Brunsgard@telenor.com

Professor Torlaug L. Hoel, PLU Torlaug.Hoel@plu.ntnu.no

Introduction

This note is a contribution to the InfoSam project, and connects the proposal to considerations expressed in the mandate of LIKT, i.e. to specifically address research issues that relates to how information and communication technologies may be understood and utilized as integral and enhancing constituents of learning environments.

The project InfoSam reflects the fact that information technologies become increasingly interwoven into the social and cultural practices related to work, communication, entertainment and education. The coming generations of students, born from around 1990 and onwards, will from their early childhood have had access to game consoles, the internet, GSM and different digital devices, and have appropriated these into their lifeworld.

Consequently, the educational system will be influenced by the affordances of technology, both from within the educational institution itself, but also from the ways in which the information technologies are appropriated by learners in their activities outside the institution.

To the main axes of InfoSam, Enabling Technologies and Artefacts in the Human Environment, the research questions of LIKT borders onto several of the items on the agenda.

Through the sciences that are consulted to enhance our understanding of what learning and teaching is, a more complex understanding of learning and teaching emerges. Increasingly, learning is seen as a complex *interplay* between individual and social processes, relations and activities. More specifically related to a technological environment, this includes play and game metaphors, networks, simulations and immersion in virtual environments connected to open-ended exploratory learning assignments.

Also, the interplay with the environment, be that natural or technological, is also seen as more important. Salient theoretical points of entry are connected to the areas of learning theory commonly labelled activity theory and ecological perspectives on learning, and the tradition of *Didaktik*. Metaphors for learning and knowledge development revolve around concepts of interaction, participation and communities of practice in physical and virtual networks. Ecology is such a metaphor, applied to describe the relationship between man and his surroundings.

To this is added an understanding that the information technologies constitute a part of the very fabric of how we communicate, how we are educated and entertained. These technologies are increasingly portable and connected to WiFi-broadband internet connections. Furthermore, technologies converge, such as the convergence of computer and communications technologies and consumer electronics. Consequently, functionalities migrate towards unified or hybrid devices. More so than previously these appliances and their functions are embedded in and intertwined into cultural practices. As accounted for elsewhere in the InfoSam context, information and communication technologies are in the process of becoming ubiquitous or ambient.

In sum, the description above points in the direction of applying the term *knowledge ecology* as an analytical point of entry. How will future learners develop skills, competences and knowledge in the framework of a knowledge ecology?

To the educational system, the challenge is to understand and utilize these into an overall project of reading the world. As integral parts of learning environments, these technologies afford ideas connected to change, interaction, negotiation, participation and influence. In short, this represents a strong departure from the traditional classroom.

Previous learning methodologies have often been based on all-embracing ideas that the learner should adapt to. Technology enhanced methods will afford more individualized learning methods, enabling more learner-centred approaches that approach methods as choices to be made after assessing the context.

From the position of LIKT, a 10 – 15 year forward-looking research perspective will include the components and themes described in the below. The basic concern of an educational system is to address the questions of the *how*, the *what*, and the *why*. These questions are at the heart of *Didaktik*. The development will be characterized by two trends; one is the shift from *improvement* to *transformation*.

The first generation of ICT-enhanced learning tools and their accompanying methods have so far mainly represented an improvement aspect. In general existing methods have been incorporated into a new technological framework. The challenge of the next generation will be to transform the *what*, the *how* and the *why* in an integrated conception that reflects an interplay and coexistence with the affordances of technology.

The other trend is from *representation* to *participation*. Whereas up to the present, ICT in learning has been focused mainly on the affordance of representation, future approaches will focus more upon participation, where learning materials, learning processes and activities are seen in a context of the *what*, the *how* and the *why*.

Both trends incorporate an increased emphasis upon the procedural aspects of the learning activities.

Research themes

Below a few research themes are suggested. The net-based learning environments have applications and tools that give both the opportunity to roam or to wander, but also to gather and congregate. The themes have in common that they address these dual forces of the IC-technologies. The net lends itself to learning styles that are distributed in time and space, and in that respect become highly individual, but contains also means to connect the participants and foster collaborative work forms.

Theme 1

Perspectives upon technology

A major concern to the educational society is to develop a perspective upon technology that is integrated into the overall framework of education. This is a perspective that goes beyond the instrumentalistic notion of technology as merely a tool. Hitherto, the educational traditions have been somewhat reluctant to incorporate the presence of technology into educational design. This includes a long chain of interrelated themes, including the organization of learning, the re-conceptualization of a *Didaktik* that incorporates the net based learning environments, and the evaluation and certification mechanisms. The introduction of learning management systems (aka LMS) is but a first step.

From the perspective of learning with technologies, more research will have to be conducted in the area where the interplay between learners and technology is the focus. One point of entry will be to study the complex and many-faceted interplay as to how affordances of technologies are appropriated into activities, processes and relations that enhance or even transform learning.

Theme 2

The learning resources: Hyperlinked, multi-medial and immersive virtual environments

In the very close future a substantial portion of the learning resources, if not all, will be available on the web. To the producers of learning material, including the educational institutions, this will enable a high degree of flexibility, of both keeping curriculum texts up to date, and to produce material on demand.

However, more important to the learner and teacher is the affordance of producing highly individualized learning materials. This aspect of web-based learning has several implications, some of which are already in their emergence, especially through ideas connected to portfolio assessment.

One also needs to perceive these changes in an *ontological* perspective; the ‘objects’ of the world are changed, “the world to be known”, in the way they are composed or assembled, and accessed through digitization. Among the effects of the net based learning resources in this particular respect is the blurring between the “authorized” and “unauthorized” material, thus adding a strong component of evaluation and assessment *before* the material is included.

These characteristics of the learning environment also represent a tension between in the need for exploratory and linear learning processes. Learning activities in these learning environments create a need for addressing in a didactic way *continuity*, *coherence* and *closure*. All these are aspects of a learning process that in ICT environments needs to be addressed more specifically.

One characteristic of learning as it is described within the prevailing present paradigm, the socio cultural, learning occurs in an intricate interplay between *internalization* and *externalization*. Or in other words, an interplay between cognitive and individual processes on the one hand, and social or collective processes on the other. The internet has tools and applications that accommodate both aspects, but the exact relationship and the inner dynamics of these processes have not been given a thorough research attention.

A fundamental aspect of the internet resources to follow with regards to development is the very nature of the web itself. Metaphorically, the web has been described as a library, but one where all the books are thrown on the floor. Worldwide, research initiatives are taking place with regards to the very infrastructure of the web, concerned with redefining the inner fabrics of the web itself through initiatives such as *the Semantic Web*. Given that these initiatives will produce a “new generation” web, consequences will be substantial, and the development needs to be carefully monitored from an educational perspective.

The overall challenge will be to develop literacies that enable learners to thrive and develop in web-rich learning environments.

Theme 3:

Enhanced interfaces and ‘intelligent’ tools

This theme reflects most explicitly the relationship between man and technology, one that might best be characterized by “a two way exodus”. Future interfaces will include a wide variety of tools, agents and avatars, and immersive simulated environments. A main characteristic will be the ability to adapt to both the technological environment, and to the human. These applications or tools will thus be in a middle or third position between man and technology, and a common denominator will be the increased ability to enter into dialogues with these tools. As mediating artefacts, these may be seen as interlocutors, offering an interlanguage between the learner and technology.

Basically, one may distinguish between three different types of artefacts, although they tend to appear in clusters. Learning behaviour involves the movement between all three levels of artefacts.

Primary artefacts are the tools themselves, the “axes and clubs” and applications such as word processors. *Secondary* artefacts are the resulting practises that occur out of the interchange between the tool and the established practice, i.e. “e-mail- language”. *Tertiary* artefacts are established when the tools and practices change the social and cultural realities, and reach into the very perspective of the how we perceive the actual world. In this respect tertiary artefacts may initiate change in current practice, usually representing a deep saturation of techno cultural practices.

These three types also represent a continuum, where the third stage represents a deeper saturation of the way the tools and the associated practices are appropriated into our daily lives. This type of artefacts will occur when a critical mass of people have lived in technological settings over a long time, and technology “disappears” into the fabric of everyday life. We are beginning to see the effects upon a generation of learners who are born into this setting, thus representing a challenge to the educational system.

Theme 4:

Educational challenges

The following topics are seen as some of the consequences of the description suggested in the above.

A: From classrooms to virtual campuses

The future learning context is on the one hand characterized by a migration out of the traditional classroom context, one in which co-localization of learners, teachers and learning materials are less typical. On the other future learners will not be restricted to one physical campus, but will rather migrate between courses held at campuses in different countries, and even on different continents. Learner behaviour or learner styles will thus need to be addressed. On an on-campus level, the need for research and development into infrastructure development is required.

B: Environments for collaboration

The last few years have seen a growing interest and institutional efforts to incorporate learning management systems (LMS) as a component of the educational offer. The most used at present are *It's learning* and *ClassFronter*, both sharing a fairly similar set of functionalities. However, the present generation of LMS tend to focus upon elements of control and information representation, say, in the form of lecture material dissemination, channelling student papers to teachers and tutors, to name some examples. Future realizations of LMS will need to address more thoroughly the participation and production aspect of the learning processes, in which learners are connected in collaborative efforts, emphasizing how the individual may be “enhanced” through being connected to other people and external resources through dialogue. Added functionality will include audio, streaming video in a more technological mature realization than what we have seen up to now.

An interesting line of development is the relationship between an LMS and a various handheld portable devices. Handheld multifunction (i.e. GSM, MP3, PAD and storage in any combination) devices hold many promises seen in combination and as supplements to an LMS, thus extending the communicative and collaborative opportunities of the learner. The common denominators to the attraction of users are size and mobility. This very mobility will mean that activities connected to learning may on the one hand be more incorporated into other activities, and on the other be more cut-up and even evasive. As of yet the didactic implications of these emerging learning styles here are not fully studied and accounted for.

The incorporation of an LMS in the learning environment addresses both the two main trends pointed out in the introduction; the transition from improvement to transformation, and the transition from representation to participation.

In sum, the interchange between the learner and the computer mediated learning environments afford the learner with a gamut of possible approaches to learning, spanning from the highly individualized and solitary work profile, to partaking in communities of practices that jointly develop and exercise their domain of knowledge.

C: Multiliteracies

The themes in the above lead to the challenge of defining and fostering *multi-literacies*- the skills, competences and attitudes attributed to reading the world in a digital context. Multiliteracies build on traditional literacy (reading, writing etc), but are augmented by coping with several semiotic systems that are blended into a unified but ever changing representation on f.ex. the web, in dialogic exchange with digital agents, or immersed in virtual environments. The notion of literacy as an analytical point of entry will define and clarify the skills needed for the learner, and will give way to describe new ways in which to utilize the ICT resources. Among skills in an augmented literacy are those connected to browsing for information, such as different reading strategies, navigation skills, verification and selection strategies, reading and incorporating text, sound, stills and video in a unified understanding of a given topic. Dealing with dynamic, ever changing information systems will be one of these. These skills are both of an operative and practical nature.

At the heart of a re-conceptualized understanding of the emerging multi-literacies is the dual notion of communities of practice and genre, which seen together extends the traditional notion of dialogue.

In sum, the process of being skilled in multi-literacies, is to address questions of continuity, coherence and closure in the learning processes conducted in technology-rich learning environments.

D: Assessments, the curriculum and certification

Our traditional understanding of assessment, curriculum and certification are under pressure. Since the idea of a stable or fixed curriculum is more or less a thing of the past, the web-based resources will afford flexible curriculum and curriculum development at a higher speed than previously. Assessment and grading systems, as well as vocational certification will be influenced. The relationship between the actual ICT-rich practices on the one hand and the questions of assessment on the other will need to be addressed from a research point of view.

E: Epistemological considerations and challenges

One consequence of the themes briefly outlined in the above, is a need to re-conceptualize themes connected to epistemology, and establish the *digital epistemologies* as a field of research.

At all times, there has been a relationship between what we consider to be knowledge, and the means and methods by which this knowledge has been gathered, assessed and passed on through an educational system. In addition, the educational system has had an elaborate grid of control and gatekeeper functions, and certification mechanisms. The educational system has also been built upon the premise that information has on the one hand been limited in terms of access (distribution through teachers and paper). These premises are now challenged. The impact of digitizing relates both to the learning objects themselves (hyper structured and multimodal), their availability, the problems connected to authorisation; in short, the very process of ‘coming to know’. From the perspective of Didaktik, this means a re-conceptualization of the core questions in learning and teaching: the *what*, the *how* and the *why*.

The term the information society is by many critics now replaced by terms such as the integration society, the hyper complex society and the Network Society. From a learning perspective it becomes paramount to develop the ability to read the world, through dealing with a dynamic and ever-changing learning environment. This means a change of focus from a *standard* towards a *procedural epistemology*:

The distinctions between a standard and procedural epistemology may be sketched as follows (Lund 2004):

Linearity: knowledge carried and expressed in language, propositions and theories	Multimodality: with ICT, meaning emerges in spatial and textual mode.
Stability: knowing as carrying out acts that pertain some “truth” that is taken to exist	Flux: knowing as an ability to perform in communities of practice.
Individual cognition as a manifestation of expertise.	Collective assemblage involving minds, artefacts and contexts.
Value of knowledge determined by use	Value of knowledge determined by exchange , knowledge becomes commodified
Belief in value or “truth” of information precedes the use of it	Information is acted upon, belief in “truth” or value may follow .

How this is carried out in communities or practice, how to identify and how to assess in the educational framework becomes a major research challenge. In short, the very constitution of 'knowers' in a techno-cultural framework.

Concluding remarks

The field of learning with ICT is a field that needs to consult a large variety of research areas, both from the humanities and from the natural sciences. The main focus described above is a focus upon the interaction between man and technology. We think that it will be a main characteristic of future generations of learners that they have from early childhood on their main points of references technologies used to interact and communicate, and are accustomed to shaping technology into constituents of their own lifeworld. Their ability to develop, participate with, and share mediating artefacts thus goes far beyond the primary level, of only using the "clubs and axes".

Furthermore, the field of research has a strong relationship to the field of practice, and research here need to be conducted in a close relationship with the users themselves, both the learners and the teachers. Research from the point of view of *Didaktik* has a strong emphasis of innovation and change of educational practice.

Proposed priorities

On the "tool/technological" side:

- Research regarding the development and use of simulations and immersive virtual environments. These applications foster activities, processes and relations that are situated and contextualized, and address the learners' development of qualifications, competences, creativity and culture. These applications will cover a broad span of applications, centered on learning metaphors of participation, immersion, role play, problem solving, hypotheses building and testing.
- Research regarding the use and development of learning management systems, and applications that address collaboration in virtual environments.
- Research regarding the development and use of dialogic agents that operate as interfaces between the learner and the learning resources.

On the "cultural" side:

- Multiliteracies: what skills and competences are required in our understanding of the digital literacies, how can they be nourished and developed?
- Digital epistemology – what constitutes digital epistemology and how does it affect the traditions of the various subjects with regards to for example tradition and innovation, continuity and renewal?
- The relationship between literacies and the digital epistemology.

Coda

What won't happen:

Schools won't disappear

Teachers won't disappear.

Students will still be around.....

Learning won't happen without hard work.

References

Visions 2020 – Transforming Education and Training through Advanced Technology.

<http://www.technology.gov/reports/TechPolicy/2020Visions.pdf>

Colin Lankshear: The Challenge of Digital Epistemologies.

<http://www.geocities.com/c.lankshear/challenge.html>

Feenberg, Andrew: Perspectives upon Technology.

<http://www-rohan.sdsu.edu/faculty/feenberg/>

Lund, Andreas (2004):

The Teacher as Interface. Teachers of EFL in ICT-rich Environments: Beliefs, Practices, Appropriation.

Doctoral dissertation UiO. <http://www.ils.uio.no/forskning/doktoravhandlinger/docs/AndreasLund-avhandling.pdf>

Løvlie, Lars: Danning i støpeskjeen. <http://www.apollon.uio.no/apollon02-99/3.html>

Murray, Janet: Hamlet on the Holdodeck.

<http://www.lcc.gatech.edu/~murray/hoh/tablecontents.html>

Norris, Mason, Robson, Lefrere and Collier

A Revolution in Knowledge Sharing

<http://www.educause.edu/ir/library/pdf/erm0350.pdf> (EDUCASE Review Sept/Oct 2003)

The Semantic Web: <http://www.w3.org/2001/sw/>

The Norwegian Program for Digital Competence (Program for digital kompetanse):

<http://odin.dep.no/ufd/norsk/satsingsomraade/ikt/045011-990066/index-dok000-b-n-a.html>

Tapscott, Don: Growing Up Digital

<http://www.growingupdigital.com/>

Enabling technology / Computation

- Mathematics**
- Computational Science**

Mathematics

Professor Kristian Seip, Department of Mathematical Sciences, Kristian.Seip@math.ntnu.no
Ph.D.Student Ingmund Bjørkan, Department of Mathematical Sciences, Ingmund.Bjorkan@iet.ntnu.no
Ph.D.Student Kristian Gjosteen, Department of Mathematical Sciences, Kristian.Gjosteen@math.ntnu.no
Associate Professor Idar Hansen, Department of Mathematical Sciences, idar.hansen@math.ntnu.no
Professor Helge Holden, Department of Mathematical Sciences, holden@math.ntnu.no
Ph.D.Student Eirik Mo, Department of Mathematical Sciences, mo@math.ntnu.no
Professor Brynjulf Owren, Department of Mathematical Sciences, Brynjulf.Owren@math.ntnu.no
Professor Håvard Rue, Department of Mathematical Sciences, Havard.Rue@math.ntnu.no

1. Mathematics—a broad perspective

Mathematics is a remarkable invention of human thought, a purely logical construction independent of experience and yet the basis for our understanding of the laws of nature. Much of modern mathematics, beginning with the invention of calculus in the seveneenth century, was developed in a close interplay with discoveries in physics and other natural sciences. It is now generally recognized that the need for mathematics is increasing, as mathematics provides the appropriate tool for modeling and understanding complex phenomena in nature, technology, and society.

While the interaction with science has persisted to this day, a distinct feature of mathematics in the twentieth century is its division into “pure” mathematics as opposed to “applied” mathematics. The extensive development of abstract mathematics in the beginning of the twentieth century, leading to new fields such as topology and functional analysis, was fundamental for the blossoming of mathematics since that time. This development could be viewed as an important intellectual liberation, but the downside was that parts of mathematics lost touch with its origins. During recent years, however, we have witnessed a notable reversal; nowadays, we see ever more examples of advanced mathematics, previously considered “pure” and useless, penetrating applied areas. The lesson is that the notions of “pure” and “applied” mathematics at best are *time dependent*, and serious mathematics should never be viewed as “useless”.

Mathematics is a vast area, and its literature grows exponentially. There are currently about 600 research journals covered by the authoritative Mathematical Reviews, and some 75000 reviews produced every year. The total volume of the current literature is estimated to some 50 million pages. Mathematics is today such a formidable construction that no one is able to overlook all of it, and most mathematicians can only hope to master a tiny piece of it. Making bold predictions about what parts of mathematics will come to use is therefore virtually impossible.

A safe prediction, however, is that mathematics will be applied in new areas, and more advanced mathematics will be used in areas where it is currently not in use. This development will continue at a rapid speed. No area is “safe” from the influence of mathematics and no area of mathematics is “safe” from being potentially applicable.

An interesting example is algebraic geometry, a mathematical discipline that until quite recently was considered to be of purely academic interest. The prestigious Institute for Mathematics and Applications (IMA) at the University of Minnesota has devoted the academic year 2006/2007 to applications of algebraic geometry. The background for this initiative is that effective use of algebraic geometry is now found in applications such as optimization, control problems, bioinformatics, data communication, and computational geometry. This example illustrates the time dependency of the notion of “applied mathematics” and the potential utility of advanced mathematics of any kind.

Teaching and research in mathematics should keep the long-time perspective that serious mathematics is important and potentially useful. We should maintain and develop our excellence in the field².

2. Mathematics and computer technology

Modern computer technology has increased the need for mathematics, as well as the range of scientific problems for which mathematics is relevant. Computers have in many ways revolutionized the research in science and technology. Mathematical modeling and numerical and statistical computations are becoming essential components in an increasing number of fields, in the analysis of computer-generated data, in predictions, simulations, and verifications of theoretical models. Research in fields such as biology, chemistry, geology, and physiology depends more and more on computer simulations, and as a result, these fields are getting “mathematized”. Also, the mathematics needed in these fields becomes more advanced and sophisticated.

The 21st century will show tremendous progress in our understanding of biological systems and the human body, and ICT along with mathematics and statistics will play a significant role in this development.

Many computations that were previously unimaginable, have now become routine exercises. This is not only due to better and faster computers, but also to more efficient algorithms. This point is illustrated by the following statement from the NSF report “Mathematics and Science” [4]: “For more than 40 years, gains in problem solving power from mathematical algorithms have been comparable to the growth of raw computing speed, and this pattern is likely to continue. In many situations, especially for multiscale and chaotic problems, fast hardware alone will never be sufficient; methods and theories must be developed that can extract the best possible numerical solutions whatever computers are available.” *Thus the use and development of “supercomputers” must be accompanied by appropriate mathematical competence.*

Another important consequence of the development of computer technology is that numerical simulations are relevant and pertinent alternatives in situations in which controlled physical experiments are either impossible or too expensive. An illustrative example is that of astrophysicists studying solar activities, a field in which the amounts of data are increasing exponentially. *The use of modeling and computations, as a supplement or alternative to physical experiments, will become increasingly important, both in scientific and industrial research.*

The development of computer technology faces mathematics with new challenges that lead to new fields. An interesting example of high current interest is that of *multiscale modeling*. Strictly speaking, multiscale phenomena have been studied for a long time in mathematics, but the current development is driven by applied sciences and computational capabilities. The need for multiscale modeling can be seen in physics, where processes at different scales are governed by physical laws of different character, say, by quantum mechanics at one scale and classical mechanics at another. To quote [1]: “Our computational capability has reached the stage when serious multiscale problems can be contemplated, and there is an urgent need for science and technology - nanoscience - being a good example – for multiscale techniques”.

Mathematicians and statisticians at NTNU should contribute to the “mathematization” of scientific and industrial research and be relevant and active participants in this development. Mathematicians at NTNU should be proactive in identifying those parts of the mathematical sciences that are important for modern applications, and this should be reflected in teaching and research collaboration.

² See [6].

3. Predictions concerning boundary conditions for teaching and research at NTNU in 2020

- *No revolution will take place concerning ways of learning.* In recent years we have seen various novel ways of presenting new material: Using various electronically based methods we can display animations, illustrations, etc that make presentations more elegant and comprehensible than ever before. Furthermore, there has been considerable increase in the understanding of the human learning process. However, when it comes to the "last step", the individual understanding of new concepts, ideas, and techniques, it still requires hard and focused work by each of us. Our claim is that this will persist. In particular, there will be no effortless way of learning.
- *Printed text will still be important.* TV, radio, Internet, etc. represent competitors to the printed word, which, however, has proved to be surprisingly resistant to the fierce competition. New handheld gadgets to represent text will appear, but traditional books and newspapers will continue to be important.
- *There will be full convergence of all means of communication.* We have seen an unprecedented development of our means of communication, e.g., mobile phones and internet. There will be full convergence of Internet, TV, radio, phones, etc
- *Proper and rapid treatment of information will become even more vital.* We see that as a result of the tremendous development of the Internet we have an incredible amount of information available at our fingertips. The understanding of this information, our ability to find, check, understand, and store this information will become increasingly important.
- *Norwegian students will still mainly go to their local universities.* We now see that most students in Norway go to the nearest university. NTNU has been an exception, and it is absolutely vital to keep it that way. The claim is that Norwegian students will mostly study in Norway, and the sincere hope is that NTNU will recruit from all of Norway. The competition between Norwegian institutions for the best students will escalate. The vast majority of students at NTNU will be Norwegian.
- *All teaching of science and engineering at universities will take place in English.* More and more students will come from abroad, and will demand lectures in English.
- *At least one of the above claims will turn out to be blatantly wrong.* Looking back at bold predictions about the future show that one should approach the present task of forecasting with caution.
- *We know the potential number of students in 2020.* One of the few sure things we can say about the future is the number of people in each age group. Our students in 2020 are already born! We have to check carefully the demographic numbers.

Below are some additional references not referred to in the above text, but of relevance to our discussion of the future role of mathematics in general [2], [3], and at NTNU and IME in particular [7].

References

- [1] Weinan E and B. Engquist, *Multiscale modeling and computation*, Notices Amer. Math. Soc. 50 (2003), 1062-1070.
- [2] G. Høst, *Hvorfor satser USA på matematikken?* Kronikk, Aftenposten, October 30, 2003
<http://www.aftenposten.no/meninger/kronikker/>
- [3] *Mathematics: Giving Industry the Edge*, The Smith Institute for Industrial Mathematics and Systems Engineering, UK, April 2002. <http://www.smithinst.ac.uk>
- [4] *Mathematics and Science*, National Science Foundation, USA, April 1999. <http://www.nsf.gov>
- [5] *Preliminary Program; IMA Special Thematic Year on Applications of Algebraic Geometry*,
<http://www.ima.umn.edu/AlgGeom/>
- [6] *Research in Mathematics in Norwegian Universities and Colleges*, Research Council of Norway, August 2002, <http://www.forskningsradet.no>
- [7] *Basic Research in the Interface between Mathematics and Information and Communication Technologies*, Research Council of Norway, August 2002, <http://www.forskningsradet.no>

Computational Science

Professor Helge Andersson, Department of Process Engineering, Helge.I.Andersson@mtf.ntnu.no

Associate Professor Finn Drabløs, Department of Clinical and Molecular Medicine,

Finn.Drablos@medisin.ntnu.no

Professor Helge Drange, Director G. C. Rieber Climate Institute,

Nansen Environmental and Remote Sensing Center (NERSC), Helge.Drange@nersc.no

Associate Professor Anne Elster, Department of Computer and Information Science, Anne.Elster@idi.ntnu.no

Professor Henning Omre⁵, Department of Mathematical Sciences, Henning.Omre@math.ntnu.no

Professor Bjørnar Pettersen, Department of Marine Hydrodynamics, bjornar@marin.ntnu.no

Professor Einar Rønquist, Department of Mathematical Sciences, Einar.Ronquist@math.ntnu.no

Abstract.

Many areas in science and engineering will benefit from future progress in mathematical modelling and analysis, and from the development of new computational methods. Together with advances in computational infrastructure, and tools for exploiting this infrastructure, it will be possible to simulate currently intractable systems. While offering many exciting possibilities for academic research and industrial development, the interdisciplinary nature of this field also poses many challenging issues.

1. Introduction

Computational science comprises a wide variety of research areas where the development and the use of simulation tools play an important role. Such research areas are found in both the basic sciences and in applied research. Examples are reservoir simulation, seismic analysis, molecular dynamics, materials research, computational mechanics, bioinformatics, medical diagnostics, and climate modelling. A common feature of all these activities is a strong interdisciplinary character: physical modelling, mathematical analysis, numerical analysis and algorithm development, computationally intensive problems, software development, and visualization. The working group has looked at the future opportunities and challenges in this field, in particular, based on expected progress in mathematical, statistical and numerical methods for solving scientific and technological problems, as well as the availability of significantly more powerful computational resources and tools for exploiting these.

Since the application areas are driving much of the development of computational science, we start by briefly discussing the future of a few selected areas. We then discuss some issues related to advances in mathematical and computational methods, and in computational infrastructure. Finally, we reflect on the impact computational science will have on future education.

2. Selected application areas

2.1 Computational mechanics

As an introduction, consider first the classical example of direct numerical simulation of *turbulent flows* involving a single fluid phase with homogeneous material properties [1]. The governing equations are known from a continuum description, and provide an excellent mathematical model at a macroscopic level. However, current numerical methods and available computational resources only allow these equations to be solved for relatively simple cases (simple geometries and at fairly low Reynolds numbers). For most industrial problems, including marine applications, a direct numerical solution of the governing equations is out of reach. Even with a continued progress in algorithm

development and computational power, it is expected that such a direct computational approach will not be possible even in 2020 for many problems of engineering interest. The main reason for this is the large range of temporal and spatial scales necessary to resolve. Note that an increase in resolution by a modest factor of 2 in each spatial direction and in the time direction will give a 16-fold increase in the number of degrees-of-freedom, implying at least an order of magnitude (perhaps two orders of magnitude) increase in the necessary computational resources. And much more resolution is needed...

Despite the limitations mentioned above, simulation of turbulent phenomena are being performed today with the help of turbulence models; such approximate models make it possible to reduce the number of unknowns to a level which is compatible with current computational resources. The trend is therefore to complement (or replace) expensive laboratory experiments with simulations, also for difficult problems like combustion.

Computational mechanics has had a tremendous impact upon science and engineering in the past [2]. Current knowledge allows us to simulate complex systems comprising structures and fluids, and also the coupling of such systems (e.g., simulation of vortex-induced vibrations). The need for computational mechanics in science and engineering will continue in an expanded set of applications as indicated below.

2.2 Medical technology

Medical technology following from advances in computational mechanics offers promising possibilities. Current use of medical technology is often of a *diagnostic* nature (x-rays, ultrasound, NMR etc). New application of computational mechanics promises new tools for *predictive surgery* [3]. Computational methods can be used to simulate blood flow and vessel dynamics, test hypotheses of disease formation, evaluate devices that have not yet been built and treatments that have not yet been implemented.

2.3 Nanotechnology

Nanotechnology is an important new area in materials research. The synthesis of nanoscale materials promises lower energy usage, lower environmental pollution, structural perfection, small size, high stiffness, high strength and excellent electronic properties. Nanoscale materials may be used in material reinforcement, as sensors, and as medical diagnostic and delivery systems [4].

Nanotechnology is expected to have a profound impact on the basic research being performed in medicine, electronics, and materials science in the coming years. Engineering software which incorporates multiscale modelling and computation will be needed; see a discussion of this approach later in this document.

2.4 Inverse problems

Inverse problems refer to problems where the objective is to infer system characteristics from measurements. Examples include seismic analysis, medical tomography (ultrasound, NMR, etc), meteorology, and satellite remote sensing. Such problems can also be regarded as optimization problems, and are computationally very demanding. It is expected that improved mathematical and computational methods will make such problems much more tractable. One promising approach is based on a probabilistic framework denoted as Bayesian inversion.

2.5 Reservoir simulation

Reservoir simulation will continue to be an important tool for determining the best way to manage our oil and gas reserves. In the coming years, it may be feasible to use high-resolution geomodels, for example, based on a multiscale approach [5]. This will allow details at the microscopic level to be better accounted for on the macroscopic level. Together with improved measurement techniques, this will allow a much better system description. Such detailed models can also be used to develop and

validate low-dimensional (reduced) models, which in turn can be used in the context of real time reservoir management.

2.6 Climate modelling

Climate modelling will become an increasingly important tool in order to predict and control the effect of human activities on Earth. For global climate simulation, many submodels are of interest: atmosphere, ocean, land surface, cryosphere and biosphere. The spatial resolution of these submodels will increase significantly in the future (down to about 10 km), and new submodels will be added, e.g., atmospheric chemistry [6]. The incorporation of local effects will also be better accounted for, perhaps using multiscale modelling. Climate models will provide individual nations and the international community with increasingly important predictive tools that can possibly also be used for educational purposes.

2.7 Systems biology

Systems biology is used as a general term for the systematic analysis of complex relationships in living organisms, in particular, at a macromolecular level. The current trend indicates that system biology will become very important in the years ahead. Computational demands will most likely come in the area of simulating cellular and (macro)molecular systems, and in particular in combined, multilevel simulations, where different levels of complexity (and resolution with respect to amount of detail in model and theory) are combined in a single, integrated system. In the future, the ability to better predict and understand complex biological systems is expected to lead to advances in drug design, disease diagnosis, biologically inspired computers, and environmental health [7].

2.8 Future outlook

The current trend indicates that much more complex systems will be simulated, involving the modelling and simulation of multiscale and multiphysics phenomena. Computational science will increasingly complement theory and experiments. However, we predict that the need to obtain *good* measurements will become more important, and not less; mediocre experiments will be replaced by computerized analyses and thus become superfluous.

3. Mathematical and computational methods

In addition to developing improved models in any specific application area, it will be equally important to continuously analyze and understand the underlying mathematical and statistical properties of such models. This is important in order to have a firm and rigorous framework to build computational methods from. The numerical methods need to be both efficient and accurate. The availability of scalable algorithms will become even more important in the quest to solve ever larger problems [8]. Algorithms that may otherwise be attractive will be subject to a selection process where only algorithms that prove to be both accurate and scalable will survive as practical tools for the simulation of very complex systems. We remark that progress in developing efficient computational methods is equally important as progress in processing speed at the hardware level [8]. With the feasibility of simulating more and more complex systems, it will become crucially important to also develop methods that can be used to ensure fidelity in the computational results (error estimation). As an example, let us consider one promising computational framework in more detail.

3.1 Multiscale methods

Consider materials constructed at the nanoscale. Such materials cannot be described with standard continuum models. Some of the critical phenomena need to be described at the atomistic and molecular level. However, nanoscale materials may interact with other components that are larger and have longer response times. There is therefore a need to simulate systems over an enormous range of time and length scales, ranging over many orders of magnitude. This will necessitate a coupling of atomistic, molecular and continuum models. The mathematical and computational challenges will be to link these models together and to make the simulation of fully coupled systems feasible. To this end, the emerging field of multiscale analysis and multiscale methods promises great research opportunities. Some key issues are how to separate the various scales as much as possible and at the same time ensure the correct coupling of one model level with the next one [4,5]. This approach may also prove useful for the direct simulation of turbulent phenomena.

4. Computational infrastructure

4.1 Software development

Software, together with hardware, represents the engine of computational science. Despite tremendous progress over the past decades, software development for scientific and engineering problems still represents a huge investment of human resources; reliability and functionality have a very high cost. The introduction of parallel processing has certainly not made matters easier. In the past, scientific software has typically been developed by researchers and engineers with background from specific application areas. This is starting to change now, and the trend is to put more emphasis on creating reusable software, and to use more modern programming techniques. Nonetheless, software development is still an obstacle for many research groups. The availability (or non-availability) of software will influence decisions regarding research tools for doctoral students. The reality today is that the turnover time for a doctoral student is about 3-4 years, the life time of research software may be 1-15 years, the hardware (and enabling tools) changes every 4 years or so, and the active life time of a researcher is about 30 years. There is obviously a strong incentive to make software development easier, more robust and less human resource intensive. Perhaps some of today's efforts in this direction will be successful [9]. It is a fact that advanced mathematical ideas are more likely to be used by scientists and engineers if they are embodied in software [10]. However, such software must be created, and this requires time and interdisciplinary skills. Technical feasibility is not enough.

4.2 Hardware development

Twenty years ago, single process systems were dominant for scientific computation. In recent years, large-scale parallel processing has been the norm for challenging problems. The current trend suggests that distributed computing and storage of large data sets will also involve many computational sites. Grid computing has certainly the potential to change the way scientists and engineers work, however, it also introduces additional challenges in terms of software development, education etc. For many research groups, it may be beneficial to wait for new enabling technology to reach a certain level of maturity before investing heavily into it.

We refer to the position paper on "Computer Systems" for more details regarding the future development of computational infrastructure.

4.3 Visualization

Methods for extracting and present useful information from large data sets will become very important. We refer to the position paper on "Computer Systems" for more details here.

5. Education

Computational science and engineering (CSE) offers a computational approach to solve problems in science and engineering which complements a pure theoretical and experimental approach. The emergence of CSE is due to the rapid advance of computational methods and computational infrastructure. The field naturally combines skills from many different fields. This interdisciplinary nature will also impact future education. We already see the emergence of specific CSE-programs within universities, at least at a graduate level, perhaps not so much at the undergraduate level. The introduction of “Computational X” classes is fairly common. How to best incorporate CSE in education is currently being discussed, and some experiences have already been documented [11]. What seems clear is that the strict borders between different departments are starting to become blurred. As an example, we mention a new PhD program in Computational Systems Biology at MIT which starts this year [7]. However, it is also clear that many current university programs are not sensitive enough to this development; it is not always easy for today’s students to mix classes from different departments in order to gain a sufficiently broad set of interdisciplinary skills.

- One thing is certain: *computational science is here to stay.*
- Finally: *It is utmost important for the future credibility of computational science to educate students with a critical mindset towards the use of modelling and computation.*

References

1. P. Moin and J. Kim , “Tackling Turbulence with Supercomputers,” *Scientific American*, **276**(1), 1997, pp. 62 – 68.
2. J.T. Oden, T. Belytschko, I. Babuska and T.J.R. Hughes , “Research Directions in Computational Mechanics,” *Computer Methods in Applied Mechanics and Engineering*, **192**, 2003, pp. 913 – 922.
3. C.A. Taylor and M.T. Draney, “Experimental and computational methods in cardiovascular fluid mechanics,” *Annual Review of Fluid Mechanics*, **35**, 2004, pp. 197-231.
4. W.K. Liu, E.G. Karpov, S. Zhang, and H.S. Park, “An introduction to computational nanomechanics and materials,” *Computer Methods in Applied Mechanics and Engineering*, 2004 (in press).
5. W.E and B. Engquist, “Multiscale modelling and computation,” *Notices Amer. Math. Soc.*, **50**(9), 2003, pp 1062-1070.
6. Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., der Linden, P. J. V., Dai, X., Maskell, K. and Johnson, C. A., eds. 2001: "Climate Change 2001: The Scientific Basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, New York. http://www.grida.no/climate/ipcc_tar/wg1/index.htm
7. Computational and Systems Biology Initiative (CSBi): <http://csbi.mit.edu/> Conference on systems biology: <http://web.mit.edu/newsoffice/nr/2003/csbi2.html>
8. ScaLeS - Science Case for Large-scale Simulation, 2003. <http://www.pnl.gov/scales/>
9. Technology Review, MIT’s Magazine for Innovation, **106**(9), November 2003. Special focus on “Extreme software.”
10. M. Wright and A. Chorin, “Mathematics and Science,” Division of Mathematical Sciences, National Science Foundation, April 5, 1999. <http://www.nsf.gov/pubsys/ods/getpub.cfm?mps0001>
11. O. Yasar and R.H. Landau, “Elements of computational science and engineering education,” *SIAM Review*, **45**(4), 2003, pp. 787-805.

The material world

- Electrical Energy**
- Gas Technology**

Electrical Energy

Professor Arne Nysveen, Department of Electrical Power Engineering

arne.nysveen@elkraft.ntnu.no

Professor Olav Fosso, Department of Electrical Power Engineering

Professor Robert Nilssen, Department of Electrical Power Engineering

Professor Tore Undeland, Department of Electrical Power Engineering

Assoc. professor Eilif Hugo Hansen, Department of Electrical Power Engineering

PhD-student Michael Belsvik, Department of Electrical Power Engineering

About this paper

This paper presents the results from the working group on “Electrical Energy” within the Infosam-2020 project. The overall aim of this project is to chart the future to the Faculty of Information Technology, Mathematics and Electrical Engineering (IME) the next 15-20 years.

Electrical Energy is a vast technical area, and this paper does not cover all areas in the same depth. The topic Electrical Energy plays an important role in international and national debates on the future energy mix, such as inclusion of renewables instead of using gas fired power plants or use of hot water vs. electricity for heating of buildings. This is very important issues, but in the context of Infosam 2020 we have focused more on power technologies and power delivery systems itself.

1. Introduction

The working group has assessed how changes due to new technologies, new demands, new regulations and increased focus on sustainability impacts generation and use of electrical energy.

One of the authoritative sources for this paper is the “Electricity Technology Roadmap” made by the Electric Power Research Institute (EPRI). In addition organizations as Conseil International des Grandes Reseaux Electriques (CIGRE) and Institute of Electrical and Electronic Engineers (IEEE) have organized activities on future electricity technology.

Distributed (or embedded) generation and storage, smart devices and buildings and advanced regimes for power system operation are aspects that the working group has considered. Advances in computational engineering and material science (e.g. nanotechnology, magnetism and polymers) open new possibilities for design, analysis and operation. New requirements and possibilities given by information technology on electricity generation and consumption are discussed.

2. The value of electrical energy

Energy in form of electricity represents the most controllable, precise, convertible and flexible energy source for our modern society. Electrical energy is used to power “everything” from small devices requiring only microwatts to large factories and plants requiring 100 megawatts and more. Power grids cover whole countries and continents and make electrical energy available as a utility service we all take for granted in the western world. Electrical engineering has its basis in electrodynamics, material science and mathematics. Electrical energy provides the “muscle” and energy in automation.

Electrical energy possesses some unique features making it highly valuable in modern society:

- Conversion into work with (almost) 100% efficiency. Exergy content is 100%.
- Processing, storage and transmission of information.
- Conversion in to light.
- Electrometallurgy – e.g. production of aluminium.

- Necessary energy carrier in extracting useful energy from renewable sources as hydro, wind and sun.

3 The power system in 2020

3.1 Introduction

The political and public debate on electric power is often related to how to provide new power and/or reducing the power consumption. Less attention is paid on how and which values does the electric power create. The emerging information society (or digital society) put new requirements on the power delivery system. By failing to response to this, the underlying productivity growth with reduced power density (ratio of electricity use to gross domestic product) and increased standard of living can be reduced. In the US, EPRI has formed the “Consortium for Electric Infrastructure to Support a Digital Society” – CEIDS to provide science and technology to power a digital economy and to integrate users and markets.

In order to satisfy future demands, provide electricity in an environmentally acceptable way and to operate in commercial power markets, several predictions need to be considered:

- More diversification of the power delivery system.
 - Increased use of distributed (embedded) power generation and storage. Wind power, small hydro, micro-turbines, bio and even photovoltaics.
 - Mix of primary energy source for the Scandinavian power market. Hydro, coal, gas, nuclear and new renewables.
 - Traditional vertical utility companies are replaced by independent power producers, energy service providers, transmission and distribution network owners and system operators.
- The need for premium power. Premium power is a term used for ultra reliable, high quality electric energy service. Requires embedded production and storage facilities.
- Information system making end customers responsive to real time production costs. System making automatic changes in end consumers power demand. Today prices are only present in the wholesale market.
- Increased efficiency by more intensive use of (power) electronics to control load. On a global basis, about 50% of all electricity is used by non controlled motors.
- Advanced tools for planning and operation of the system.

3.2 Power production

The future demands for electricity depends on several factors such as technological development, economic growth, social development and political decisions. However, in order to facilitate an economy based on increased productivity with more intensive use of information technology and automation, electricity is needed. In North America, the UN forecast an increase of 1.3% per year the next 50 years. In Norway, mean annual increase in power demand from 1991 to 2000 was 1.5% per year. Gross consumption including grid losses was about 122 TWh in 2000. Annual production 118 TWh. Towards 2010, NVE expects an increase in power demand of about 14 TWh from the year 2000.

The new production can be provided from several sources: Wind, hydro, gas, bio and the sun. Future production mix depends on political decisions, such as tax and regulations, cost and technology advances. Current policy is to provide 3 TWh from wind power in 2010 and to introduce more power from small hydro power plants. Gas fired power plants are under consideration (3 specific projects are approved by the authorities) but no decision on investment has been made. Uncertainties related to gas price, investment cost for pipelines and CO₂ emissions (tax, re-injection) are some of the reasons.

While power generation in the beginning was dominated by manually performed control tasks, this has changed a lot over the years. Today hydropower plants runs on remote control. This remote control covers: generation level on each unit, automatic startup and shutdown of units, switching operation and measurement of key values. Remote handling of the power plant also required that condition monitoring of equipment became centralized.

One of the big challenges is to make the control and state monitoring systems cheap enough to deal with distributed resources. In the past, few but large units made the information volume easy to handle. With many small units maybe in magnitude 1000:1 compared with the traditional plant size, one faces a tremendous growth in data volume. At the same time the distributed system must work with information highways of lesser reliability than for large power plants, the system must not commonly fail due to shutdown of data channels.

On top of this, value measurements needed for accountancy and condition monitoring adds to the data volume. Routines for handling data series with gaps in an acceptable way are a necessity.

Advances in power electronics will play a major role in a successful inclusion of distributed generation. Large wind-mill, micro turbines, photovoltaics, full cells and energy storage units all need power electronics for grid connection. This gives rise to new challenges, but does also open up new possibilities. Reduced cost and increased efficiency of power electronics is necessary.

3.3 System planning and operation

Two major challenges will be dominating the system planning and operation:

- The availability of peak capacity
- The expansion of the transmission system

The most successful restructuring of the power supply system took place in systems with excess of production capacity at the time of restructuring. Excess of a commodity is a requirement for competition to work. However, many of these countries are now experiencing smaller margins between the peak load and the peak capacity. As it is today the generation must take place simultaneously with the consumption. The physical generation facilities must therefore be present. The peak load has a very short duration and the units covering this load are often expensive to operate. The market prices are not high enough to cover the investment in new production facilities. New generation capacity will in a competitive environment not show up before the prices are high enough and this will be too late.

Very little expansion takes place in the transmission system. It is difficult to get permission to build new lines and generally the focus on cost reduction is very strong. Only smaller expansions are done in order to enhance the use of the existing system that will imply smaller margins.

Operating closer to the limits will be a major challenge to the system planning and operation. Smaller margins call for better information and the time to react will be shorter. Competition between the different actors in the power market will reduce the information available that could support the system planning and operation.

The transmission systems have become more meshed and interconnections that mainly were built for reliability purposes are now used for heavy power transfer. The result is that significantly larger systems must be analysed to evaluate the impact within own system. The long-distance power exchange will increase the need for coordination between the system operators. In the large disturbance in USA 2003, a major problem showed up to be lack of coordination.

The power production sources will in the future be a mix with different characteristics. Some of the renewable resources may have lower availability than traditional sources, but will still be an important contribution to the total production capability. An example will be large wind parks where integration within the existing system is challenging.

The flexibility of the demand side will in the future be very important for power balance. Reducing peak load is better than investing in units for providing peak capacity.

The future system planning and operation will have to deal with many different sources and with different characteristics to provide balance between supply and demand as well as smaller margins in the transmission system operation. Advanced sensors and monitoring techniques will make the system “smart”, enabling better real time estimation of the state of the system and to continuously update contingency plans. Online analysis and control will become a critical element for a safe future operation of the system. Real time access of data and the capability to analyse and interpret the results in due time in order to counteract in case of disturbances are mandatory.

3.4 Power Transmission

After restructuring of, transmission is still defined as a monopoly and separated from the sectors with competition. The situation has introduced several challenges for the transmission system operation. The need for transmission capacity has been steadily increasing after the liberalization started. Power exchange over long distances is quite common and the power flow in the system may show paths that the system was not built for. These challenges arise in a situation with decreasing investments in new transmission facilities and with a significant reduction in expenditures on maintenance.

A major motivation for the restructuring has been cost reduction, and the transmission system operators are forced to reduce the operation costs and also to justify the maximum transfer capacity on important transmission corridors in the system. It is also very difficult and time-consuming to get permission to build new transmission lines. Operating closer to physical limits is the main topic.

There are different phenomena limiting the transfer capability. The thermal limits of the individual circuits define the maximum flow on the line. However, the stability of the system is often more limiting for the maximum transfer. Stability is associated with the dynamic behaviour of the system. The system must also be operated with certain margins in case of unexpected disturbances.

Power electronic equipment (e.g. FACTS) is now introduced in the system with the purpose to better utilize the existing transmission infrastructure and to improve flexibility of the system operation. Some of these components address the power flow distribution while others address the dynamics. Increasing capacity and lower price will cause these to be widely used. A major challenge will then be coordination between all the dynamic elements in the system where the *collecting* and *interpretation* of real-time information will be crucial.

The data network becomes very important for a secure operation of the power transmission system. Both the protection system and the decision support tools need reliable data. However, disturbances in the power system will often cause problems in the data network. Other infrastructures are also dependent of the power transmission. Degradation of the system’s present performance can therefore be extremely expensive. Widespread disturbances as observed in 2003 are very expensive for the society and will easily in short time exceed the amount saved by restructuring and competition.

Improved cables and superconducting devices are likely to be widely used in the future power transmission systems. These will permit higher transfer capacity.

Distributed generation will also be an important contribution for the future power transmission system. Local generation close to the consumer will reduce the transfer requirements while penetration of large

wind farms will introduce other challenges. A breakthrough on storage technology will become very important for the utilization of distributed and renewable generation.

3.5 Power distribution system

Medium voltage systems

The distribution system operates on the medium voltage level (6 kV – 36 kV) where 12kV and 24 kV are dominant in Norway. In power distribution, power flow direction is unidirectional, from the transmission network to end-consumers connected to the low voltage level. Some industry users connect directly to the medium voltage level. In future to accommodate the needs for increased reliability, quality, service and customization, distributed resources are foreseen. This will alter the power flow in the distribution system. New means of protection, power management/planning and new hardware are needed.

In Norway, with no gas distribution system, distributed resources near load centres are most likely introduced for premium power purposes, i.e. to provide a secure supply in case of disturbances in the main grid. New concepts such as islanding of self-sustained micro grids will be introduced for customers with high power quality requirements.

Today's medium voltage system is operated as a radial network with very simple power flow patterns. In order to increase reliability, outage time and load ability more meshed topologies will be present. This calls for new devices to handle fault situations. New devices based on power electronics combined with advanced protection system can respond to abnormal situations in less than one millisecond. Today's mechanical systems typically need 30-50 milliseconds to act. Large industry stakeholders as ABB and Siemens have already demonstrated some concepts on this route. Outage times can be reduced from typically one hour to some milliseconds. Power electronics are also foreseen to be introduced for inclusion of distributed generation, power flow control and power conditioning. Generally, power electronics will move from being more or less absent in medium voltage, to play a dominant role in some specific networks.

4. Power apparatus and devices – technology advances

4.1 Power electronics

At NTNU the possibilities opened with power electronics were identified early, just when the technology was introduced. A Licentiate disputation on induction motor drives with invertergrade thyristors took place in 1966, only 7 years after the (standard type) thyristor was invented by General Electric in the USA.

The power electronics technology spread fast in Norway to a large number of production industries. With help from NFR, Norway is still a leading country in power electronics, especially in telecom represented by the two companies Power-One and Eltek. In railway applications, locomotives got inverters for speed control of the induction traction motor (Flytoget, Signatur and Agenda all have this technology and all make use of the Undeland-snubber for voltage and current stress relief during switching).

Power electronics can help the transmission capacity and the quality of power. This means that the increasing demand of electrical power can be supplied without building new lines. Also electric power by cables to Denmark, Germany, Netherlands, UK and offshore platforms is only possible with power electronics using the HVDC technology.

4.2 New power semiconductors

In early 1970's the different types of power transistors as the BJT, the MOSFET and the IGBT started to replace the thyristor. The cost went down, and new markets opened up like inverters that have revolutionised heat pumps in the recent years.

4.3 Digital control

Until early 1990's analogue technologies were used in the control due to the high speed demands. Now DSPs and FPGAs are used all over, even for Pulse Width Modulation and the fast inner current loop. Today we can not imagine control of machines in applications from wind energy to wheelchairs without utilising the digital hardware and software.

4.4 Innovative ideas for new applications and new markets

The new power components and user-friendly digital control make prototype development quite easy. Thus establishing new spin off companies where power electronics is a part of the product is not so costly and time consuming as before.

4.5 Electric machinery

Introduction

Electrical machinery is used both as generators of electrical power and motors where electrical energy is converted into mechanical energy. In OECD countries, motors consumes about 60% of all electrical power. Norwegian industry has traditionally been a manufacturer of large generators for hydro power generation. Lately the focus of Norwegian industry has been on advanced motor drives in traditional industry, shipping and offshore applications.

Improved performance of induction machines

The induction machine is the industry standard, used in most fixed speed applications. Due to relatively low overall cost and no need for power electronics control in fix speed application, this machine will still be one of the volume products in the future. The research on induction machines focus on energy efficiency, and is not performed in Norway.

Increased use of advanced variable speed motor drives

The flexibility and performance of power electronics has revolutionized electrical motor drives. Variable speed and high quality torque control is now available at acceptable cost for industry applications. A large number of process applications utilize these new motor drives to make new energy efficient industry plants. Similarly are a lot of large modern electric drives used in ship propulsion and compressor and pump offshore applications.

Research on modern motor drives including power electronics and advanced control systems is to be continued into 2020. Reduction in cost and simplification in installation/programming causes a penetration of motor drives in the constant speed market of today.

New Permanent Magnet materials are used in new electrical machine concepts.

The new low cost materials with exceptional performance have resulted in an international research on new machine concepts. The application of transversal flux and axial flux machines have demonstrated that compact machines with very high efficiency can be manufactured at a compatible cost. The possibility for cost effective fabrication of tailor made machines due to new materials and fabrication methods, opens up the arena for new stakeholders than high volume machine manufactures. Access to a large capital base and large production facilities becomes less important in these new markets.

Our research in the coming years will focus on the new machines concepts. In close connection with advanced users, special purpose machines such as wind generators, ship propulsion motors and actuators will be designed.

4.6 Breakers and switchgear

In the next 15-20 years, most new developments are foreseen for medium voltage and low voltage systems, due to introduction of power electronics and new insulation systems. For high voltage, the voltage and power rating are still too high for power electronics to be competitive. However, due to a general wish of less visibility, more use of compact GIS-systems can be foreseen, especially in city centres.

For medium voltage systems, several new developments are foreseen:

- Power electronics for connection of distributed resources
- Breakers using power electronic devices for fast and precise operation.
- Ultra compact switchgear that can be buried in the streets (no visibility)
- New, low cost solid insulation systems (thermoplastics)
- Automation and communication system for remote metering, power flow control and operation of distributed resources.

4.7 Cables

The dominant insulation material of today is cross-linked polyethylene (XLPE). In order to increase the power capacity and size, higher field stresses and operating temperature are needed. Maximum field strength in normal service is 20 kV/mm (peak value). Theoretical breakdown strength is about 1000 kV/mm. Improvement in manufacturing and new material additives are needed to come closer to this theoretical value. XLPE cables for DC systems are available today. In 2020, cables for DC will be a mature product. Use of nanotechnology to make specialised materials of will be introduced.

Prototypes for outdoor high voltage insulators have already been presented.

For city centres with a very high power load density, superconducting cables offer the possibility of supplying a high amount of power in existing cable ducts. Nexans AS in Norway has already received an order for a superconducting cable to be delivered to New York City. Development of new low cost insulation for low temperatures and cryostat will speed up the use of superconducting cables.

5. Special applications with high national importance

This section describes some applications/sectors which are not a part of traditional power supply industry but plays an important role in our national power industry.

5.1 Offshore and ship power systems

Norway delivers about 10% of all electrical power equipment and systems to commercial ships. The flexibility in design of cruise liners, ferries, supply ships and fishing trawlers is a driver to use electric propulsion. New innovate electric machine design and system integration will advance the propulsion and manoeuvring of these types of ships. For cruise liners and ferries, comfort and manoeuvrability is essential.

The offshore oil industry sets very challenging requirements to the power industry. Today's platforms are large industrial sites with advanced power systems. In order to recover oil from mature reservoirs, smaller and complex reservoirs and reservoirs at deeper water, electric power is needed at the sea bed.

- Subsea processing plants with power requirement above 100 MW
- Voltage levels of 132 kV for power supply from land or remote platforms
- Subsea frequency converters with power rating of 50 MW at 1000 meters water depth.
- Reliable power delivery systems with transformers, switchgear and cables.

- Electric actuators, motors and power electronic converters to operate complex subsea processing units.

5.2 Power supply for telecommunication

11 % of the electric energy in industrialised countries are used to supply the infrastructure for computers and communication. It is no doubt that the number of short outages able to disturb computers increase. Even worse, long lasting blackouts are experienced more now than before. In the power supply industry, this is discussed a lot. We are invited to plan a large research project in DC distribution in houses and office buildings. DC will make standby power with a much higher reliability than the present day ac supply. In areas like California and Ireland the high tech industry declares they will move if the power supply is not reliable enough.

6. National partners

SINTEF Energy Research (former EFI) has since its foundation in 1953 been a major partner for research on electrical energy. The shift from “electricity research” to “energy research” reflects the broadening also in our research in this field. Within the fields of power systems and high voltage engineering the co-operation is comprehensive and integrated.

SINTEF MARINTEK is a recognized world class institution within its field. Co-operation on electric systems on ships, e.g. the NFR founded project “All-Electric Ship”.

NFR (Norwegian Research Foundation) supports long and short term research within our technical field. NFR has newly launched the RENERGI program.

Statkraft and Statnett. Statkraft is Europe’s second largest hydropower producer. Statkraft has also ambitious plans for wind power. Statnett is owner of the transmission grid and power system operator.

Nexans AS (former STK) is one of the worlds leading manufacturers of high voltage cables. The research on insulation materials has been a continuous activity for almost 40 years. Materials for superconductors and for higher temperatures are new topics.

PowerOne (former Powec) and **Eltek** are major manufacturers of power supplies for telecommunication on a global basis. PowerOne is a direct result of R&D at our department in 1960s and 1970s on power electronics.

ABB, Siemens and Aker. These are manufacturers and suppliers of electrical equipment and systems in Norway. For some products these companies have production facilities and engineering departments in Norway with world marked responsibility.

Oil companies (Statoil, Hydro). Development of new technology for power supply to sub sea production facilities.

References

- [1] Electric Power Research Institute (EPRI), “Electricity Technology Roadmap. 1999 Summary and Synthesis”, Doc. no., C1-112677-V1, July, 1999.
- [2] Electric Power Research Institute (EPRI), “Electricity Technology Roadmap: 2003 Summary and Synthesis”, Doc. no., 1009321, November, 2003.
- [3] A. Grübler *et. al.*, “Dynamics of energy technologies and global change”, *Energy Policy*, pp. 247-280, vol. 27, 1999.
- [4] K. E. Yeager, “Electricity for 21st century: digital electricity for a digital economy”, *Technology in Society*, article in press, 2004.
- [5] C. W. Gellings and R. J. Lordan, “The Power Delivery System of the Future”, *The Electricity Journal*, pp. 70-80, no. 1, vol. 17, 2004.
- [6] J. S. Graves and John D. Clapp, “The Future of Electric Transmission”, *The Electricity Journal*, pp. 11-21, no. 12, vol. 14, 2001.
- [7] J. D. Kueck and B. J. Kirby, “The Distribution System of the Future”, *The Electricity Journal*, pp. 78-86, no. 5, vol. 16, 2003.
- [8] V. S. Budhrija, “The Future Electricity Business”, *The Electricity Journal*, pp. 54-90, no. 9, vol. 12, 1999.

Gas Technology

Future need for Information and Communication Technology in the areas of Production, Transportation and Utilisation of Natural Gas

Morten Hovd, Institutt for Teknisk Kybernetikk, NTNU, morten.hovd@itk.ntnu.no

Bjarne A. Foss, Institutt for Teknisk Kybernetikk, NTNU, bjarne.foss@itk.ntnu.no

Jan Tommy Gravdahl, Institutt for Teknisk Kybernetikk, NTNU, tommy.gravdahl@itk.ntnu.no

Abstract.

Application of advanced ICT-based tools has potential for significantly increasing the economic value derived from Norway's abundant natural gas resources. Production, processing, transportation and utilization of natural gas involves a number of different engineering disciplines, and the needs for ICT as well as the challenges that must be met vary from discipline to discipline.

Scope

This memo addresses, as indicated above, the needs for Information and Communication Technology (ICT) in the areas of production, processing, transportation and utilisation of natural gas. Note that the needs for ICT in petroleum prospecting or drilling are not addressed. On the other hand, the operational challenges associated with chemical conversion of natural gas are in many aspects similar to what is encountered in other branches of the chemical processing industries. Many of the generic research needs related to gas conversion can therefore be expected to apply also to other processing industries

1. Background

Norway has abundant resources of natural gas. This gas is valuable both as a source of energy and as feedstock for the production of a number of different chemicals. The export value of natural gas is expected to surpass the export value of oil within a decade.

At present, almost all the natural gas produced in Norway is exported via transmission pipelines to continental Europe and the UK. Thus, much of the value creation from this natural resource occurs abroad. The Gas Technology Center NTNU – SINTEF has been established to coordinate research activities in Trondheim in the areas of production, processing, transportation and utilisation of natural gas.

The research challenges range from obtaining better understanding of basic phenomena, through improved operation and control, to improving the basis for investment decisions. Environmental aspects are also important. Most notably, there is the issue of greenhouse gas emissions when using fossil fuels for energy production. Minimising carbon dioxide emissions from energy production, as well as handling and disposal of captured CO₂, are therefore central research topics.

Information and Communication Technology (ICT) is essential both in performing the research as such, and in the industrial application of the research results. In some application areas the ICT tools are applied off-line, whereas other applications involve two-way interaction with operating equipment. Thus, the ICT tools are required to work on-line (in 'real time'). Both on-line and off-line uses of ICT tools typically involve extensive use of optimisation and/or simulations, and visualisation tools are necessary for presenting and interpreting results.

In on-line applications the reliability of both computations and communications is essential. The communications often have to work over significant distances – for operation of a gas transmission network extending over hundreds of kilometres this should be obvious, but also within a single production facility distance and reliability of communication is important. In addition to the safety and reliability issues associated with the on-line use of ICT tools in advanced operations, ICT tools is also used to monitor and ensure the safety of operating equipment.

In the following, we will describe briefly a number of application areas for gas technology, with operational objectives and engineering problem areas. This will subsequently be used to motivate the need for education and research in various ICT domains.

2. Application Areas

2.1 Process and product design

Traditionally, the process design problem has started from a requirement to produce specified amounts of well-defined products, within legislative and environmental constraints, with the objective of maximising the profit earned from the production.

This is still the dominant process design scenario, and presents a number of challenging research topics. Such challenges include

1. *Economically optimal design*, while taking environmental and safety-related aspects into account.
2. *Design for flexibility*. The design should allow changes in operating conditions and possibly changes in product specifications, and still enable profitable production.
3. *Design for operability*. The design must be operable in a practical sense, i.e., it should not put unreasonable demands on operators and control systems.

Valuable tools exist for addressing each of the areas above, although completely general methodologies are not available. Integrating all aspects of the design problem into a unified methodology is also hard. Typically a number of initial designs showing economic potential have to be evaluated for safety, flexibility and operability.

Some of the tools used in process design are relatively simple to conceptualise, and is at present used systematically by engineers with a Masters degree. An example of such a tool could be the “pinch curve” used in heat integration. Other aspects of the design problem result in highly complex optimisation formulations, expressed as Mixed Integer Non-Linear Programs (MINLP’s). Few Master’s level students have the background to both formulate such problems (requiring knowledge of physics, chemistry and process technology) and to solve the resulting optimisation problems (requiring knowledge of advanced global optimisation tools). A high level of competence in quite disparate areas is required to use such programs and reliably interpret and analyse the results of computer programs used to formulate and solve complex design problems.

Process design is usually formulated using static models. Although this agrees well with the operating philosophy of most process plants (requiring stationary operations, infrequent changes in operating conditions, and a control design focusing on disturbance rejection), formulating the process design problems as static optimisation problems naturally results in a vast reduction in the degrees of freedom available in the design. It is noted above that even static design problems often result in highly complex optimisation problems. Introducing dynamic operation into the design problem will typically increase the complexity significantly.

In the future, the focus is expected to shift from ‘process design’ to a paradigm more centred on the success criteria for the product in specific applications. Thus, the required physical and/or chemical properties of the product need to be specified, without necessarily specifying what type of product should be used to achieve the specified properties. To illustrate, instead of specifying that the plant should “produce X tonnes/day of polymer Y with chain length distribution Z”, the specifications may in the future detail

- o the properties of the product in the final application (strength, abrasion resistance, electrical conductivity, high/low temperature tolerance, etc.),
- o the properties of the product in subsequent processing (e.g., flow properties during moulding), and
- o the properties of the product related to recycling and/or disposal.

This clearly adds additional complexity to the design problem, and requires connections to be made between phenomena on a microscopic (possibly down to molecular) level and product properties on a macroscopic level. Effective two-directional communication between different model scales is a central issue. These techniques must reach sufficient maturity (in terms of general applicability and reliability) in order to have realistic hope of making significant progress.

2.2 Process operations

Process operations focuses on the operation of existing plants, and plant design is considered as fixed. The time scales addressed in operations range from fractions of a second (safety, regulatory control) to weeks or even months (in production planning and supply chain management). It is natural that the system descriptions and problem formulations used will differ significantly depending on the type of operational problem and time scale in question.

On the longer time scales, static plant models (e.g., describing production capacities) and discrete decision variables (what product to produce, when to produce, possibly also where it should be produced) dominate. Although alarms and automatic shutdowns may be classified as discrete decision variables, continuous decision variables and dynamic plant models dominate the operational problems in focus at the shorter time scales.

Although significant research problems exist at all time scales, the two ends of the ‘time scale spectrum’ may be claimed to be better understood than the intermediate range of time scales (from tens of minutes to days). Many problems on this time scale require both continuous and discrete decision variables. System stability cannot be ignored on these intermediate time scales, and dynamic plant models are therefore required. Thus, many problems on this timescale are inherently complex. However, significant time is also available to solve these problems, and this is therefore a promising area for advanced ICT. This intermediate timescale also contains the interface between the ‘business operations’ (focusing on the longer timescales) and the ‘plant operations’ (focusing on the shorter timescales). These are two different ‘cultures’, and very few people have deep knowledge of both cultures (or problem areas). Research aiming to bridge the gap between the two cultures should therefore have significant potential.

2.3 Operator support

The term ‘operator support’ is here used for techniques and applications for supporting decisions made by human operators. This includes techniques for

- o navigating historical data to find situations similar to the present operating scenario or operating problem,
- o performing ‘what if’-analysis by simulating the future effects of choices available to the operator,
- o providing advance warning of developing operating problems,
- o making equipment documentation and operational procedures available on demand, and
- o analysing operational data to identify poorly functioning equipment or control systems, needs for maintenance, etc.

2.4 Remote operation

Oil fields which are remotely operated already exist. Remote operation, with processing equipment located on the seabed, is clearly an attractive alternative for many deep-sea oil and gas fields. Clearly, such remote operation places high demands on the safety and reliability of the equipment and control systems.

Remote operation is also of interest to education and research as such. NTNU has received funds from the EU to make several of its advanced laboratories in the oil and gas area available to researchers from other parts of Europe. Clearly, remote operation could save a lot of travel and subsistence expenses for researchers elsewhere in Europe who perform experiments at NTNU.

3. Enabling technologies

A number of information and communication technology areas are discussed briefly below, to describe needs and potentials for research in these areas. There is necessarily some overlap with the discussion of application areas above.

3.1 Instrumentation

There are many challenges for research and development of new and improved instrumentation in various parts of the natural gas value chain. Instrumentation is here understood to include both measurement and actuation devices.

Downhole (inside a production or injection well) instrumentation must be reliable in very hostile environments, due to high temperature and pressure, erosion and corrosion. Maintenance is at best difficult and costly, often impossible. There are difficulties both in supplying energy for instrument operation, and in communication with surface facilities. Of interest are both measuring instruments (providing information on temperatures, flowrates, pressures, etc.), and actuators for active well control located downhole. Actuators of interest include active gas lift valves, as well as valves for controlling production from various zones within a well.

Instrumentation for sub-sea production facilities is undergoing rapid development. Focus is also here on reliability and cost of maintenance, harsh operating conditions, limited energy supply, etc, although these challenges in general are not as extreme as for downhole installation.

There are significant research challenges also in instrumentation of many of the chemical processes using natural gas as feedstock. For instance, some polymerisation processes operate at extremely high pressures, and even measuring basic variables like temperature and pressure is difficult. More complex properties related to product quality (e.g., viscosity, chain length distribution) often require sampling and laboratory analysis even for processes operating at more benign conditions.

Many research and development needs in instrumentation are common to all branches of the chemical processing industries. These include issues such as improving self-diagnostic capabilities of the instrumentation, ease of installation, configuration and maintenance, efficient communication, etc.

3.2 Communication

It is noted above that in many natural gas applications the ability to communicate with the instrumentation is a major challenge. This applies in particular to instrumentation inside injection or production wells, and is difficult also for subsea production facilities. Communicating with instrumentation is also a severe problem in subsea oil and gas transportation pipelines. Such pipelines can be several hundred kilometres in length, with measurements (and actuators) only at the inlet, outlet and booster/recompression platforms. The information available is thus very sparse, and this is an obstacle to efficient monitoring and optimisation of operations. This problem is more severe for gas transmission pipelines, which have slower dynamic response than oil pipelines.

On production platforms and in chemical conversion plants many difficulties arise from the sheer scale of the system. The number of signals to transmit to and from various instruments may well run

into tens of thousands. Thus, installing and maintaining communication requires substantial effort. Wireless communication may hold promise in simplifying communication, and will certainly reduce the need for communication cables. Installing and maintaining wireless communication in production facilities will certainly present challenges of its own. However, more serious are the issues of information security and potential for terrorism. The typical manipulated variable in a process control loop receives a new value from the controller once every few seconds. A patient spy or saboteur listening to the communication will therefore have access to massive amounts of data that can be used for cracking data encryption codes. Conceivably, after cracking encryption codes a saboteur may attempt to take over control of parts of a plant, cause destruction to the plant and endanger both operating personnel and people outside the plant.

3.3 Modelling

Practically all advanced tools for process design, optimisation, control and monitoring require some sort of plant model. Modelling is therefore a cornerstone in all applications of process systems engineering. Different applications require very different models. Much can be said for keeping the models used as simple as possible, but several factors point towards increased use of more complicated models. These factors include requirements for greater performance and flexibility (implying that the models must be valid over larger ranges and include dynamics), applications involving effects at different scales of magnitude (from molecule to reactor), etc.

Whereas effective tools exist for deriving linear models from experiments, there is a great need for tools to aid in rigorous models (based on physics and chemistry), fitting parameters to complex models, experiment design and model validation for non-linear models, model reduction for complex non-linear models to increase computational speed, etc.

Many real life systems are of a distributed nature, and are most naturally modelled using partial differential equations (PDEs). Modelling tools and technology is therefore needed both for systems described by ordinary differential equations (ODEs), and for systems described by PDEs.

3.4 Simulation

Mathematical simulation is heavily used in advanced plant design and operations (including decisions on when and where to drill new wells, etc.). Methods for numerical integration may be said to be well developed for models of small to moderate size, although model reduction is required in some applications to obtain required computational speed. However, reservoir simulations and simulation of large processing plants both involve systems of large scale, for which efficient and accurate simulation requires effective exploitation of model structure.

Large-scale system simulation is likely to remain an active research area, and requires understanding of both mathematics and the type of system in question.

3.5 Visualisation

Large-scale system simulations, as encountered in both reservoir simulations and process plant simulations, produce massive amounts of data. The ability to interpret and draw correct conclusions from such data rests critically on the ability to visualise simulation results and navigate in 'simulated space and time'. VR techniques have already found extensive application in reservoir simulations as an aid in visualisation. VR techniques and HMI design is also important in process simulators.

Navigating, filtering, and interpreting data

Large amounts of data are found not only in large-scale simulations, but also in operational data for large-scale systems. Although there is a great potential in using such data effectively for model fitting and validation, using historical data to monitor and diagnose present operation, etc., there are also many difficulties in doing so, including:

- The sheer mass of data available.
- Some data may be faulty or so heavily filtered that it is unusable.
- There are many dependencies between the data that are only partially known.
- Important information is not logged, e.g., manual interventions.

- o Wear, maintenance and equipment replacement means that the system is continually changing. This area will therefore remain a source of relevant research topics for years to come.

3.6 Optimisation

Optimisation is heavily used both in investment decisions, supply chain management, process design, planning and control. Optimisation is therefore a key enabling technology in a wide range of application areas. The types of optimisation problems applied range from LP problems (with a linear optimisation criterion, linear constraints and continuous decision variables), to MINLP problems (with non-linear optimisation criterion, non-linear constraints, and both discrete- and continuous-valued decision variables). The ability to find fast and reliable solutions to limits the applicability of many optimisation problems, in particular in real-time applications. The effects of uncertainties in the optimisation problems are also poorly understood in many cases. In real-time applications the effects of uncertainties are particularly difficult to handle, both due to the effects of feedback and because there is rarely time to perform any sort of ‘what if’ analysis or Monte Carlo simulation.

3.7 Control

Control is an area that makes active use of several of the other areas listed above, in particular modelling, simulation and optimisation. What distinguishes control from merely being a combination of other technology areas is the ability to fundamentally change system behaviour by the use of feedback. Using feedback control, an open loop unstable system can be made stable.

Understanding and recognising this fact has been central in many recent advances in areas of relevance to gas technology, including feedback stabilization of

- o slugging flow in pipelines,
- o unstable gas-lifted wells, and
- o compressor surge.

The above examples illustrate the importance of continued applied control research.

On the theoretical side, the problem of designing a controller for a linear plant with specified inputs and outputs is well understood. On the other hand, it was realised some thirty to forty years ago that to use a single, centralised controller for a large-scale system is in general a bad idea. Nevertheless, the problem of designing control structures is still only partially understood.

Most real-life control problems are non-linear and include constraints in controlled and manipulated variables. This is a very large research area with a number of relevant research topics.

The most widely applied control technology for problems involving constraints is what is called Model Predictive Control (MPC), which solves an optimisation problem on-line for each sample interval. Until recently, the ability to solve such optimisation problems in the time available has been a severe constraint to the range of problems for which MPC can be applied. Recent advances have dramatically reduced the need for on-line computations, at the expense of requiring heavy computations at the controller design stage. Ongoing research aims at extending these techniques to more complex non-linear systems, addressing robustness to model uncertainty, hybrid problems (involving both continuous and discrete-valued manipulated variables), etc.

3.8 MEMS technology

The emergence of Micro-Electro-Mechanical Systems technology presents a number of opportunities and challenges in production, processing and transportation of natural gas.

The aviation industry is experimenting with ‘active surfaces’ based on MEMS technology, which minimises aerodynamical drag on aircraft. One might envisage similar technology being applied to impeller blades in compressors or to the internal surfaces of transmission pipelines.

Furthermore, there is an abundance of problems where distributed sensing would be very helpful in process operations, in particular if the sensing elements can be made so small that they do not significantly influence process operation. Examples might be measurement of flow distribution in distillation columns or reactors, measuring temperature distribution in reactors, etc.

In such distributed sensing applications, it will be necessary to provide for both the necessary energy for operation, as well as providing communication with the sensing elements, without hard-

wired connections to individual sensors. The typical times between maintenance or replacement should be at least 2 – 3 years.

4. Conclusions

The efficient value creation from Norway's abundant natural gas resources is of great importance to our nation. ICT provides tools for capturing the potential value of the natural gas, by enabling decisions to be made more effectively and correctly both in off-line and on-line applications, and by improving our ability to act on such decisions.

The natural gas area holds significant potential for application of advanced ICT. Such advanced applications will require understanding both of ICT as such, and the particular application area in question. In general, needs for ICT in both education and research are recognised by the various engineering departments involved in the natural gas area at NTNU. Some of these engineering departments already have significant co-operation with the 'generic' ICT departments. However, there is clearly scope for extended co-operation. It is our conviction that extended co-operation will benefit all parties involved, and will strengthen NTNU's ability to provide education and research of relevance to Norwegian society.

The biological / medical world

- Bioinformatics**
- Health Informatics**
- Life Support / Medical Instrumentation**

Bioinformatics

Associate Professor Finn Drabløs, Department of Clinical and Molecular Medicine,
Finn.Drablos@medisin.ntnu.no

Post doc. Mikael Hammer, Department of Electronics and Telecommunications,
mikael.hammer@tele.ntnu.no

Professor Astrid Lægreid, Department of Cancer Research and Molecular Medicine,
astrid.lagreid@medisin.ntnu.no

University lecturer Jon Olav Hauglid, Department of Computer and Information Science, Email
jon.olav.hauglid@idi.ntnu.no

Professor Bjørn K. Alsberg, Department of Chemistry, bjorn.alsberg@chem.ntnu.no.

Abstract.

Future developments in bioinformatics and related research areas of relevance to informatics are discussed. The main trend is large-scale integration of data from a variety of sources, with focus on simulation and analysis of complete complex systems (“systems biology”) rather than individual parts. This requires a close integration of complementary techniques, including informatics. It is also important to have specialists in e.g. informatics, mathematics, statistics, physics, chemistry and biology with good interdisciplinary communication skills, working in multidisciplinary environments. Research and development associated with health surveys and biobanks is presented as multidisciplinary area of particular relevance to NTNU.

1. Introduction

Bioinformatics is a multidisciplinary research area at the interface between informatics and biology, with additional input from e.g. statistics and mathematics. Traditionally it has focused on molecular biology, in particular sequence analysis. This includes *in silico* gene finding, sequence analysis, sequence alignment, structure prediction etc. In parallel to bioinformatics several related and partially overlapping research areas have evolved. Computational biology focuses more on computations and simulations associated with biological macromolecules and biological systems, including simulation of cellular processes. Mathematical biology analyses e.g. population dynamics in complex ecological systems, whereas physical biology studies physical processes and phenomena found in biological systems. Nanobiotechnology combines nanomaterials and nanotechniques with biotechnology and biological molecules, using the best from both worlds, whereas nanomedicine uses nanotechnology for medical purposes. Medical informatics handles and retrieves medical data, e.g. electronic medical records or medical publications. Medical statistics and statistical genetics uses statistical methods to analyse medical data looking e.g. for correlation between genetic variation and susceptibility to specific diseases.

However, there is an increasing need to look beyond these traditional (and artificial) boundaries, and gradually these partly separate fields are interacting, merging and exchanging ideas, methods and approaches [1]. In particular there is a strong trend towards looking at large, complex biological system as a whole, rather than as a collection of specialised sub-topics. This approach is often called “systems biology” [2-4], and represents integration of informatics, biology, physics, mathematics and statistics. It also integrates several technological disciplines, as a prerequisite for getting the necessary experimental data.

We believe that this tendency will continue, with increasingly close integration of relevant disciplines in order to analyse and understand even more complex biological systems. The disciplines will probably approach a symbiotic relationship, where there is a beneficial mutual dependency between complementary disciplines. Even though large biological data sets are collected at all levels of bioinformatical research, much non-genetic information is still missing, and this must be collected

and included in the bioinformatical modelling in the future in order to arrive at new biological insights. However, most needs and requirements, e.g. with respect to informatics, will be defined by the biological and medical problems we are trying to solve.

2. Major trends

Within the main trend towards closer integration of sub-disciplines, several sub-trends can be identified.

2.1 Automation / robotisation

A high degree of automation, including an extensive use of robots for sample preparation and handling, has been a prerequisite for genomic research, in particular in areas like genome sequencing and microarray analysis. This process will continue. In the future most routine tasks in the laboratory will be handled by robots. It is also very likely that robots and computers will be involved in designing and controlling experiments as well, in the sense that they run and optimise experiments that are designed (by the computer) to answer specific questions. Proof-of-principle design of robot-controlled experiments has been published [5].

2.2 Nanoscale measurement techniques

Living tissue is a very heterogenous system. In order to really understand such systems it is necessary to do measurements on a very small scale, e.g. on individual cells, or even sub-cellular compartments. The only way to achieve this in relatively intact cells is by nanoscale measurement devices. This will give very detailed information, and will be an important research area. Advanced simulations are needed to develop these measurement techniques. Such methods can give large amounts of measurements, in principle like for example an online, continuous microarray experiment. Therefore advanced methods are also needed for data processing and analysis.

A related area is sample analysis using microfluidic devices ("lab-on-a-chip"), where extremely small sample volumes may be needed. This means that very large amounts of experimental data can be generated from small samples, both in research and in e.g. medical diagnosis.

2.3 Artificial intelligence

Research is not only experiments and data analysis, it is also a creative process leading to novel questions and inventive experiments to answer these questions. It is difficult to predict to what extent artificial intelligence, computers and robots will take over the creative process. It is certainly true that computers are better at handling large data sets and complex relationships, and can find novel patterns in such data. Intelligent computers will therefore be increasingly important for stimulating the creative process in the human researcher. In particular artificial intelligence with true creativity would be a very useful tool, as it would be able to analyse more complex data sets than the human brain. However, improved methods for data analysis, presentation and visualisation can make very complex data accessible also to human interpretation. Data mining by sorting through data to identify patterns and establish relationships is already important, and it will be essential to develop improved and novel methods.

2.4 Data integration

Advanced data analysis as described above has to be based on data from many different sources. For systems biology efficient integration of very different types of data is therefore a prerequisite. Such data may be sequence data for genes and proteins, measurement data from various types of instruments, multi-category answers from health survey questionnaires, description of symptoms and various tests from medical records, observations from publications, image data from different types of imaging experiments etc. These data need to be stored and accessed efficiently, and tools are needed to compare, classify and run statistical tests. This will be a challenging task, where standards and ontologies are needed. However, it is also important to be able to include old data e.g. from publications. These data are at present time not standardised or annotated in accordance with established ontologies. Therefore advanced data processing will be needed in combination with

construction (and continuous re-construction) of ontologies. Data modelling, in order to get a good understanding of complex data structures, can be useful in this context.

2.5 Multi-level simulations and computations

Biological systems are complex at several levels. In principle we want to simulate systems by starting at a quantum-mechanical level and go all the way through to full ecosystems. This requires simulation at different levels of theory. Full simulation of e.g. enzymatic processes requires a quantum-mechanical treatment, whereas e.g. transport processes and mechanical properties of cellular structures need various levels of approximation and abstraction. Ideally we want to treat the different elements of such complex systems at the level of theory that actually is sufficient to reproduce experimental data. However, even this can be extremely challenging.

There are also other types of computations that are relevant in this context. One example is phylogenetic estimates, where one tries to model previous evolution based on current-day data. Another example is medical statistics. New experiment types, larger population studies and research on more complex, multifactorial diseases means new challenges for medical statistics. In general, statistical methods will be essential for most future research in systems biology.

2.6 Visualisation

Visualisation is already important in bioinformatics, in particular in the area of structural biology where advanced visualisation of macromolecular structures is essential. This is reflected in the price of many commercial tools for protein structure modelling and visualisation, where the price tag may be more than USD 15.000 even for a single-computer installation. A detailed understanding of individual 3D structural properties may be essential in e.g. a drug design project, therefore virtual reality tools may be used in such projects. However, also for other types of data such visualisation will become increasingly important. Examples are visualisation of cellular 3D structures based on stacked slices from imaging experiments, flexible visualisation techniques for analysing very high-dimensional data, or techniques for highlighting correlations in complex, multi-property data sets. Here on-line, real time manipulation of the visualisation, even for very large and complex data structures, will be essential.

3. Grand challenges

Since bioinformatics by definition is intertwined with biology and medicine, it is important to look at some major challenges in biology and medicine that most likely will be important focus areas for the coming decade(s) [6]. They represent biological questions, but with important methodological challenges where informatics, mathematics, statistics, physics, chemistry etc becomes essential.

- *Comprehensively identify the structural and functional components encoded in the human genome. Elucidate the organisation of genetic networks and protein pathways and establish how they contribute to cellular and organismal phenotypes.* – This requires efficient methods for identifying DNA/protein interactions, mapping of regulatory networks, identification of alternative gene products, mapping of protein products, characterisation of DNA folding and 3D structure etc.
- *Develop a detailed understanding of the heritable variation in the human genome. Understand evolutionary variation across species and the mechanisms underlying it.* – This requires efficient methods for characterisation of SNPs (single-nucleotide polymorphisms) and possibly very high-speed genome sequencing [7].
- *Develop robust strategies for identifying the genetic contributions to disease and drug response. Develop genome-based approaches to prediction of disease susceptibility and drug response, early detection of illness, and molecular taxonomy of disease states.* – This requires improved approaches towards health surveys and biobanks, and improved methods in medical statistics and statistical genetics for handling such experiments.
- *Use new understanding of genes and pathways to develop powerful new therapeutic approaches to disease.* – This requires advanced modelling of regulatory and metabolic pathways,

experimental and theoretical structure prediction, improved *in silico* drug design techniques and better strategies for drug synthesis [8]. This will be an important aspect of personalised medicine, where treatment is optimised towards the genetic and metabolic profile of each individual [9].

Based on these and similar problems, some immediate challenges for informatics and computational biology have been formulated [6].

- New approaches to problem solving, in particular from data on physical properties (e.g. protein folding, protein-protein interactions, genetic variation).
- Reusable software modules.
- Robust methods for elucidating the effects of non-genetic factors on health and disease.
- New ontologies for description of complex data types.
- Improved database technologies and knowledge management systems, in particular for integration of many different data types across disciplines.

4. Training

What kind of training do students need in order to meet these challenges? Bioinformatics is a multi-disciplinary field, and two different strategies can be envisioned. In the “traditional” approach we train students to become experts in key areas like mathematics, informatics, statistics etc. In the “futuristic” approach we train students to become multi-disciplinary experts, so that e.g. bioinformatics students have a relatively strong background in both informatics as well as biology. This is the ideal situation. However, the “futuristic” approach is probably unrealistic. With the level of knowledge needed to become an expert, it will for most students be unrealistic to become true experts in more than one area, and to become competitive we need highly qualified students in key areas. This means that we probably have to use the “traditional” approach as a basis.

On the other hand, it is essential that these students learn multi-disciplinary communication, e.g. an informatics student must learn how to discuss a biological problem with a biologist and see how it can be solved using informatics tools. So the “traditional” approach has to evolve into a more “forward-looking” approach with more focus on multi-disciplinary collaboration. This means that some basic training in relevant areas is needed, e.g. molecular biology for informatics students, or informatics and data analysis for biotechnology students. The basis for this must be established using relevant cross-disciplinary courses, like “Molecular biology for technologists” (an existing course at NTNU). However, the most efficient training process will most likely take place in truly multi-disciplinary research groups, collaborating on actual research projects. Giving students the opportunity to do projects and thesis work in such research groups will actively stimulate cross-disciplinary communication. This means that such research groups have to be available, and they must have the necessary resources to take care of these students. However, the benefit will be large. Such research groups will have a very good potential for really novel research, and the students get excellent training. This will be a win-win situation for all parts, where multi-disciplinary collaboration becomes a symbiotic and creative process of mutual benefit.

5. Implementing the future – Health surveys and biobanks

If we want to turn these visions into reality it may be useful to identify multidisciplinary research areas that are able to integrate a large number of activities, and where research data from these areas will be of national and international importance. One such area that is of particular relevance to NTNU is health surveys and biobanks [10].

There are two major types of biobanks. Population-based biobanks are normally associated with corresponding health surveys. A specific population (e.g. the full population of a selected region) is invited to participate. They are normally examined (e.g. height, weight, blood pressure, bone mass etc), they fill in questionnaires about e.g. life style, and a blood sample may be collected and stored for future use. Based on these data as well as links to “end-point” registers (cancer, stroke etc) and analysis performed on the biological material, novel knowledge can be generated about important

diseases like asthma, cancer and diabetes. In particular for complex, multi-factorial diseases such large population-based surveys most likely is a prerequisite for new knowledge.

Patient-based biobanks have a slightly different focus. These are based on biological material collected from patients, normally as part of the diagnostic process at the hospital. This material is interesting in its own right, as it can be used for better understanding of specific diseases, in particular at a genetic and molecular level. However, it becomes even more interesting when it can be linked to a population-based survey, where it can be used to confirm and elaborate data generated from the population-based material. This may include e.g. looking for genetic differences between cancer subtypes that show subtle differences in the population-based material. Such differences may have significant implications for successful diagnosis and treatment.

NTNU and St. Olav, the University Hospital in Trondheim, is responsible for two very important biobanks, both supported by the national FUGE (functional genomics) project. The HUNT health survey and biobank is an almost complete screening of the population in the county of Nord-Trøndelag. It has been carried out two times, in HUNT 1 (1984-86) only health survey data were collected, whereas in HUNT 2 (1995-97) also blood samples were collected. A new survey, HUNT 3, is planned for 2006-08. HUNT is relatively unique on an international scale because a stable population is followed over a relatively long time. It is also important that it can be linked to the second large biobank in the region, the patient-based biobank at the University Hospital. This makes it possible to do very thorough investigations based on this material.

However, full utilisation of this potential requires a co-ordinated effort from many different disciplines.

- *Robotics and low-temperature storage technologies.* Biological material in biobanks is often stored in liquid nitrogen for increased stability. The storage system should include automatic computer-controlled sample retrieval for analysis. Robotics is also important for sample handling e.g. between different instruments for sample preparation and analysis.
- *Database solutions, data integrity and security.* All data, including health survey data and data from analysis of biological material, have to be stored in a database, and FUGE has already allocated some funding to a pre-project on databases for biobanks. Such data have to be verified with respect to integrity (that data and relationships are correct), and they have to be stored according to requirements for such data (access control and encryption). Data integrity has to include data from e.g. laboratory information management systems (LIMS), like batch numbers for chemicals used during analysis. This is essential for tracking of data affected by experimental errors.
- *Analytical methods.* The novelty of the research will at least partly depend upon which analytical methods are available. Methods that can give new information on e.g. the regulatory state of cells in a biological sample will lead to unique results. Nanotechnology and advanced electronics will probably be important for the development of new analytical methods.
- *Electronic patient journals (EPJ).* Easy access to patient data will make it much easier to link patient status (e.g. diagnosis) to health survey data. The patient status can probably be given a more realistic description compared to a typical end-point diagnosis, in particular with respect to how it develops over time. NTNU has a national centre for research on EPJ systems.
- *Real-time wireless data collection.* Future health surveys may follow a subset of the population even closer than today, using a wireless system to monitor health status continuously for extended time periods. This will give more realistic data on health status. It may also be relevant to use similar systems to monitor environmental effects, like exposure to pollution or noise.
- *Statistics.* The statistical analysis is a key step in any research project based on health surveys and biobank material. In particular for the analysis of complex multifactorial diseases improved approaches are needed.
- *Bioinformatics and systems biology.* In order to actually explain how e.g. a given genetic difference may lead to increased stroke risk, and hopefully how that risk can be reduced, bioinformatics and systems biology will be essential. Statistical analysis of the health survey

and biobank material may identify the crucial differences that are associated e.g. with an increased risk, but bioinformatics and systems biology is needed in order to understand what this actually means on a molecular and cellular level.

- *Intellectual property rights, technology transfer and commercialisation.* Such projects will most likely lead to novel ideas for better diagnostic methods or treatment, including diagnostic kits, drugs etc. The technology transfer office (TTO) at NTNU will be an important partner in commercialisation based on results from biobank research.
- *Ethics.* Health-related research can easily lead to ethical questions that have to be handled properly as part of the project. This is well established in medical research, in particular through regional ethical committees. However, this is an important research area in itself, where NTNU is active.

It is clear that biobanks and health surveys has a considerable potential as an integrated multidisciplinary research project that can serve as a platform for active collaboration between different disciplines. There may be other areas with similar potential. However, the biobank and health survey area has the advantage of being a well-established ongoing activity with considerable potential for expansion and novel initiatives. NTNU is currently doing foresight-analysis on biobanks.

6. Acknowledgements

We want to thank the programme committee for Programme for Bioinformatics at NTNU for valuable input. In addition to F.D., A.L. and B.K.A. the committee members are Kjell Bratbergsengen, Arne Halaas, Atle Bones, Catharina Davies and Mette Langaas from NTNU, Per Magnus from the Norwegian Institute of Public Health and Erik Must from Fondsfinans ASA. In particular we want to thank Mette Langaas for her contributions to this paper. We also want to thank the NTNU biobank foresight group for important input.

References

- [1] V. Maojo, C.A. Kulikowski, Bioinformatics and medical informatics: collaborations on the road to genomic medicine?, *J Am Med Inform Assoc* 10 (2003) 515-522.
- [2] L. Hood, Systems biology: integrating technology, biology, and computation, *Mech Ageing Dev* 124 (2003) 9-16.
- [3] L. Hood, D. Galas, The digital code of DNA, *Nature* 421 (2003) 444-448.
- [4] E. Werner, *In silico* multicellular systems biology and minimal genomes, *Drug Discovery Today* 8 (2003) 1121-1127.
- [5] R.D. King, K.E. Whelan, F.M. Jones, P.G. Reiser, C.H. Bryant, S.H. Muggleton, D.B. Kell, S.G. Oliver, Functional genomic hypothesis generation and experimentation by a robot scientist, *Nature* 427 (2004) 247-252.
- [6] F.S. Collins, E.D. Green, A.E. Guttmacher, M.S. Guyer, A vision for the future of genomics research, *Nature* 422 (2003) 835-847.
- [7] J. Kling, Ultrafast DNA sequencing, *Nat Biotechnol* 21 (2003) 1425-1427.
- [8] E. Davidov, J. Holland, E. Marple, S. Naylor, Advancing drug discovery through systems biology, *Drug Discov Today* 8 (2003) 175-183.
- [9] R. Molidor, A. Sturn, M. Maurer, Z. Trajanoski, New trends in bioinformatics: from genome sequence to personalized medicine, *Exp Gerontol* 38 (2003) 1031-1036.
- [10] J. Kaiser, Biobanks. Population databases boom, from Iceland to the U.S, *Science* 298 (2002) 1158-1161.

Healthcare Informatics

Associate professor Øystein Nytrø, Dept. of computer and Information Science, nytroe @ idi.ntnu.no
Associate Professor Arild Faxvaag, Dept. of Neuroscience, arild.faxvaag @ medisn.ntnu.no

Abstract.

This paper discusses the future role of information technology in healthcare. It describes present and future challenges of health care, gives a status with regard to penetration and usefulness of healthcare information systems and presents elements in a strategy for future research and education in the field of healthcare informatics.

1. Background

To describe healthcare informatics in an engineering context, it is convenient to start with the basic phenomena: Diseases induce subjective discomfort and shorten life spans. Medicine as a discipline is at the intersection between three areas:

The natural sciences: Medicine studies diseases as natural phenomena in organisms.

Engineering: Medicine relies on development and use of diagnostic and therapeutic tools.

Humanities: Medicine must interpret and consider subjective, individual experiences of diseases, treatment and exposure to processes and technology. Medical problems are often closely connected to social and psychological phenomena.

Healthcare is a service provided by health care personnel to patients. Healthcare services can roughly be divided into;

Diagnosing: exploring and classifying problems which might be associated with a disease

Intervening: using medical technology and other remedies in a disease or problem process to reach a more favourable situation.

Nursing: alleviate the needs and discomfort of patients suffering from disease or effects of interventions.

Healthcare is sometimes but not always provided by health care personnel employed in larger health care organisations. Hospital services often involve very specialized, advanced, new or expensive medical technology or coordinated efforts from multidisciplinary teams. Primary care is characterized by a very large scope of diseases and problems, often with emphasis on life-long treatment and chronic diseases. An important part of healthcare is research and education.

In this context, healthcare informatics can be defined as [12]:

“A scientific discipline that concerns itself with the cognitive, information processing and communication tasks of health care practice, education and research, including the information science and technology to support these tasks”

1.1 Challenges and trends in healthcare

Predicting the future of information technology in healthcare is a challenging task. Many independent trends will influence the future:

- reliance on well developed methods for evaluating outcomes of interventions and technology and consequently aggressive marketing of “successful” new and costly medical technology.

- little interest in evaluating cost-effectiveness and clinical usefulness of interventions in real life and corresponding discrepancy between cost and usefulness.
- increased global discrepancy in healthcare service quality, spending and availability.
- a chronic discrepancy between the amount of resources that could be spent on health care and the resources actually used (the consequences of which are most catastrophic in poor countries).
- a rapidly growing and increasingly mobile population that may spread diseases (e.g. HIV/AIDS, Ebola and SARS) .
- lack of will or ability to use well established medical knowledge and cheap, highly cost-effective preventive or curative regimens (for example the use of vaccines, prevention of smoking and the use of condoms to prevent transmission of HIV)
- a large increase in the number of elderly people in some developed countries (as a result of the successful prevention of accidents and of wars, progress in medicine and health care, and the successful implementation of technologies and methods for avoiding unwanted pregnancies)
- development and global dissipation of cheap information and communication technology and resulting efficient global distribution of knowledge and information within medical communities and between providers and users of health care services.
- consumerism, patients seeing themselves as demanding and critical consumers of health care services.
- the rapid growth of biomedical knowledge and the emergence of new medical technologies on the basis of this new insight.

1.2 Present situation of IT in healthcare

Information systems have been introduced into healthcare in fairly distinct phases, motivated primarily by government, financial, legal and management needs (e.g. systems for accounting, patient management, order entry, process planning etc.). Other systems have been introduced as part of tools and technology (e.g. systems for logistics, laboratory, X-ray and other imaging). Coding systems were developed mainly for accounting and research purposes. Clinical department systems are usually developed for the purposes of quality development and research. "Electronic" medical and patient record systems have developed out of plain text editors, with little functionality and structure. In many respects, healthcare has not seen the same computer revolution as most other knowledge-intensive areas of the society.

Healthcare is worthy of special attention from an IT research and education perspective for many reasons:

Healthcare is a major part of our economy, and even small improvements are important [5]:

Healthcare averages 9% of GDP in the OECD countries, and is expected to rise with 3-4% the next 40 years. Private healthcare will increase spending considerably (The US used 13,9 % of GDP for healthcare, half of which was through public programmes only covering about 10% of the population.).

The delivery of health care is an enormously information-intensive activity.

Information handling is a major part of all clinical work. A study at St. Olav hospital found that approximately 25% of total time was used on record, notes and meetings (morning, X-ray). This is more than the time spent with patients [9].

Relevant knowledge is difficult to find, adapt and apply [8]:

One important aim of healthcare is to make clinical decisions on the basis of the best possible knowledge (Evidence-based medicine (EBM)). Systems for efficient and timely delivery of

appropriate knowledge remains to be developed. A wealth of abstracted knowledge, electronic publications, internet portals and local procedures are available "point-and-click". Availability is not an issue. The challenge is quick retrieval and adaptation of recommendations and guidelines to a specific healthcare problem. One way of achieving this is to use the same conceptual organization, structure and terminology for both user interface, computerized record content and knowledge repositories. A part of this solution is to organize information around care and patient processes and pathways.

Current clinical information systems fail to support interdisciplinary work and asynchronous communication between members of a team

Lack of support for asynchronous communication increases the possibility of making mistakes during handing over of responsibilities from one team to another.

Current clinical information systems fail to support the delivery of safe health care.

Medical errors in USA has been estimated to be responsible for 98.000 deaths each year, and the cost of preventable errors estimated to lie between 17 and 29 GUSD per year [8]. Although the literature suggests that Computerized physician order entry (CPOE) have the potential to improve patient outcomes through decrease of adverse drug events, actual improvements in medical outcomes have not been documented. Installation of such systems could actually increase the number of adverse drug events and result in higher overall medical costs, particularly in the first few years of their adoption [11]

Current clinical information systems does not support the development of high quality and efficient health care [6]:

Even with so-called electronic record systems, most Norwegian hospitals are intensely dependent upon paper records, informal communication, meetings, dictation etc. - Because the record systems are inconvenient for entering, finding and presenting information. The most prominent projects to go "electronic" have put huge effort into digital photographic representation of record documents.

Current clinical information systems contains multiple representation of the same information and do not support continuity of care [7]:

Information is entered repeatedly for different purposes and in different systems, leading to inconsistency and waste of time. A recent study shows that as much as 35% of a general practitioners time is spent on reporting information already in the EHR. Information is kept local to service providers and organization, in some cases also local to a ward or department, and is only shared through more or less formal and delayed communication and messaging. The patient, contrary to common belief, is not a reliable source of information about own health history and medication.

Information and knowledge has no value:

Healthcare economy relies on counting procedures, persons, encounters, diagnoses, hours and days. Efficient, correct, available and reliable information is not measured and valued. General practitioners in Norway are almost completely reliant on electronic patient record systems, and this may well be because the user and producer of information is the same person. Thus the information is directly relevant for their own efficiency and service quality. Timely and relevant information should be regarded as an important and valuable asset.

1.3 Relevant trends

Many events and trends may influence the needs for informatics in healthcare in the future. Some of the more relevant factors within healthcare, society and government, and healthcare informatics are outlined underneath:

Healthcare [1]

- Expanding market and increased investment in technology-intensive, expensive healthcare in the rich parts of the world.
- Increased risk of pandemic diseases.

- Lack of reliable and competent healthcare personnel, medicine, technology and infrastructure in developing countries.
- Growing awareness that systematic management, distribution, dissipation and use of knowledge and information is paramount to efficient high-quality care.
- A renewed perception of disease as an individual process, and methods integrating subjective and objective parameters.
- More information and knowledge may lead to more conditions being treated
- A broader and more specialized spectre of conditions will be treatable, and thus increase overall spending.
- Increased focus on self-treatment, lifestyle and disease prevention in the rich world.

Government and society

- Privatization and individualization of responsibility for healthcare.
- Research and technology development focused on areas of high-profit and not on areas of global importance.
- Healthcare as coming provider of "life content" in the same arena as entertainment, sports and gaming.
- A change in perspective from healthcare as production systems to supporter of the patient and disease processes.
- Much more service-aware and knowledgeable public.
- The age distribution, and growing percentage of elderly, will increase the need for both long-term and labour-intensive healthcare.

Healthcare informatics

- Focus on patient involvement and control.
- Focus on the needs of the healthcare professional.
- Global, shared, replicated, information spaces and storage with easily managed distributed access control.
- New methods for analyzing and combining information in healthcare records and biodata.
- Methods for making information, computerized support and knowledge relevant and usable for the real users, namely patients and healthcare personnel.
- Progress in representing, analysing and using relevant knowledge.
- "Knowledge-based" record systems and renewed interest in integrated decision support.
- Ubiquitous computing for both healthcare personnel and patient.

1.4 Challenges

The challenge for future healthcare informatics is to enable the healthcare services and professionals to ask the relevant questions. So far, the answer has not been clinical information systems, but

infrastructure, computers, web portals and information systems as part of equipment. The nucleus of a safe, effective, patient-centred, timely, efficient and equitable healthcare service is the patient and care professional. Software engineers must in turn be able to answer future questions from the patient and the care professional about the functionality and effect of their systems. The development of new and exciting information technology should be driven by such questions, and not by ability and profitability alone. Some relevant questions (and examples) that may be asked with respect to current and future systems are [4]:

1. For what purpose and whom was the system designed? (*Are electronic procedure-books efficient during patient encounters? Do they increase compliance? Do they increase the patients experienced quality of care?*)
2. Does the system work as designed? (*Are the alerts and reminders generated for a specific patient correct and "of use" to the clinician?*)
3. What is the impact of various system enhancements or modifications? (*Does the new patient summary display screen help clinicians quickly understand the patient's past medical conditions and treatment and allow more meaningful discussion of the current reason for the visit?*)
4. Is the system used as anticipated?
After implementation of a physician order entry (POE) system, the percentage of orders entered by physicians could be examined.
5. Does the system produce the desired results?
Does the new POE system actually reduce occurrence of adverse drug events?
6. Does the system work better than the procedures it replaced? How and why?
Does the new clinical laboratory alerting system reduce the time patients spend in a critically abnormal physiologic state compared with the previous telephone notification system?
7. Is the system cost-effective?
Does the increased time clinicians spend entering data during the patient visit lead to efficiency or improvement in quality within the overall health care system and thus justify its continued use of this practice?
8. How well have individuals been trained to use the system?
What percentage of clinicians can successfully perform a series of tasks required to manage a simulated patient encounter?
9. What is the anticipated long-term impact on how departments and organizations interact?
When routine drug/drug-interaction checking is moved out of the pharmacy and onto the clinician's desktop machine, what is the impact on the pharmacy department's morale and productivity? Will the general practitioner and hospital avoid all inconsistent medication?
10. What are the long-term effects on the delivery of medical care?
Does provision of regular health maintenance reminders to clinicians at the point of care improve the long-term health outcomes of chronically ill patients?
11. Will the system have an impact on management of healthcare?
Does incorporation of real-time reminders about the current drug formulary decrease variability in prescribing behavior and thus reduce costs of managing the organization?
12. To what extent does impact depend on the practice setting?
Can the same data entry and review screens be used by both primary care and specialty care clinicians?
13. Does the increasing complexity of modern medicine and the CIS required to implement it help or hinder clinicians and their patients? And how does use of the CIS affect the patient-provider

relationship?

1.5 IME, the future and healthcare informatics

With an IT-perspective, the need of healthcare seems endless. Superficially, even the most obvious problems such as common conceptual models, information sharing, access control, a modular service-wide software architecture and good user interfaces have not been properly addressed. Looking closer, healthcare unfolds as the ultimate challenge for all areas of research and education within information technology. Areas like software engineering, artificial intelligence, mobile computing, user interfaces, information retrieval, data and information storage, system and software architecture and computer-supported cooperative work will have to cope with intertwined problems of a new order of magnitude in both complexity and size. Solving problems for healthcare is justified both by its importance to society, and also the catalytic effect it will have on all areas of our discipline.

R. Haux & al. in "Health care in the information society. A prognosis for the year 2013" [3] writes::

The three main goals:

- *patient-centered recording and use of medical data for cooperative care,*
- *process-integrated decision support through current medical knowledge,*
- *comprehensive use of patient data for research and health care reporting*

... The research environment can expect various topics to remain of utter importance:

- *Structuring and testing comprehensive electronic patient records, which support the casuistic use of patient data for direct patient care, and which allow patient group analyses and use of data for research and reporting.*
- *Conception and testing of suitable information system architectures that support cooperative, patient-centered and cross-institutional care.*
- *Conception and testing of system architectures for 'knowledge centers' offering specialized medical knowledge world-wide via the Internet.*
- *Conception and testing of methods for medical data analysis ('medical data mining') based on modern information system architectures and electronic patient records and aimed at clinical and epidemiological research, as well as health reporting.*
- *Integration and testing of comprehensive, practical, useful and mobile ICT tools for patient care.*

Successful research in these areas, and the successful implementation of the frameworks mentioned, will set patient care and medical research on new, future-oriented foundation and will help strengthen financial developments. The term 'information society' would then, most probably, apply to health care.

In summary, healthcare can be envisioned as a major driving force and justification for research and education in most disciplines of informatics. So far, healthcare informatics has had comparably little impact. Healthcare is a very knowledge- and information-intensive arena, and the sheer size, complexity and importance of the information that drives healthcare has prevented development of information systems that can actually improve efficiency and quality for healthcare personnel and patients. The people that produces and uses the information is stuck with partial and disintegrating information systems and communication solutions.

Healthcare professionals, and patients, are realizing that a major part of healthcare is producing, using and managing information. The shared information and knowledge must be organized around the patient and the disease, and must be made available through user-specific

system interfaces, with functionality, relevance and perspective suiting the role, situation and need of the different users.

A major factor for enabling an efficient, effective, safe, patient-centered, cost effective and high-quality healthcare through informatics is by establishing and maintaining an arena between healthcare professionals and technology developers: This requires close cooperation around multidisciplinary research projects and not least mutual and shared knowledge about problems and solutions, objectives and limitations.

References

1. [The world health report - shaping the future](#)
2. [United Nations Information and Communication Technologies Task Force: Caring Communities for the 21st Century: "Age of Connectivity"](#)
3. [R. Haux & al.: Health care in the information society. A prognosis for the year 2013, Int. J.of Medical Informatics, Vol 66, Issues 1-3, 2002, pp 3-21](#)
4. [D F Sittig & al.: A Clinical Information System Research Landscape, The Permanente Journal, Vol 6, No 2, 2002.](#)
5. [Health at a glance: OECD indicators 2003](#), summary of the OECD Health Data 2003: a comparative analysis of 30 countries
6. [H. Lærum & al.: Doctors' use of electronic medical records systems in hospitals: A cross sectional survey, BMJ 2001;323:1344-1348 \(8 December \)](#)
7. [S. Rognstad & al.: Vet fastlegen hvilke medisiner hjemmesykepleien gir pasientene?, Tidsskr Nor Lægeforen 2004; 124: 810-2](#)
8. [Institute of Medicine, To err is human, USA, The National Academies Press, 2000](#)
9. [K. Røhme og L. E. Kjekshus: Når tiden telles - sykehuslegers tidsbruk og arbeidsoppgaver, Tidsskr Nor Lægeforen 2001; 121: 1458-61](#)
10. [Informatics-Review: The Future of Clinical Information Systems in Healthcare: Challenges for Researchers in Clinical Informatics](#)
11. [Berger et al. Computerized Physician Order Entry: Helpful or Harmful? J Am Med Inform Assoc. 2004;11:100-103](#)
12. [CEN/TC251 European Standards – Definitions and Descriptions](#)

Life support/Medical instrumentation

Professor II Kjell Arne Ingebrigtsen, GE Vingmed/NTNU,
kjell.arne.ingebrigtsen@iet.ntnu.no
Reinold Ellingsen, Optomed AS,
Post.doc Mikael Hammer, NTNU
University lecturer Magnus Lie Hetland, NTNU
Professor II Bjørn Olstad, Fast S&T ASA/NTNU;
Øyvind Stavdahl, Sintef/NTNU

1. The socioeconomic priorities

Medical instrumentation includes equipment for diagnostic, therapeutic, and monitoring purposes. Diagnostic equipment stands for about half of the commercial value. Therapeutic devices include implants such as pace makers, prostheses, and artificial organ support devices and replacements. Most of the technical development will be carried out through private companies. The 10 most R&D intensive medical companies are all pharmaceutical. They spend more than \$30 bill in R&D annually from a sale of \$220bill. The R&D spending in medical instrumentation is less than \$4bill out of a sale of almost \$40bill. ICT represents an important enabling technology in medical instrumentation. The priorities of the companies will be based upon return on investments. In general terms this is determined by the ability and willingness to pay for treatment and cure, and the frequency of appearance of the disease. A typical example from past decades is heart diseases in the western world. The growing cost of health care is subject to political concern all across the world. This impacts the priority of investments into new technology. Whereas substantial attention previously was devoted to “prestigious diseases” prevalent in economic wealthy societies, emphasis is turning towards efficiency in treatment, or – more health for the buck- also in these parts of the world. This trend is enforced by economic priorities such as reimbursements.

The concepts of “point of care” medicine and more recently “personalized” medicine have been raised to develop more efficient and better patient care.

“Point of care” medicine aims at making the right diagnosis and start the right treatment immediately in the first contact with the patient. By this substantial savings are expected from escalating fewer patients through the medical hierarchy. In primary care today most patients are being referred to specialists for thorough examination and treatment. For the general practitioner to be able to do more, better tools will have to be provided to assist in making the right decisions.

“Personalized” medicine aims at customizing an optimum treatment based upon detailed and specific genetic information about the patient. Important studies are being carried out among pharmacologists to better be able to tailor the medication to the individual needs. Work is carried out from the instrumentation side to narrow the gap between gene technology and anatomy. One company has introduced the terminology “molecular imaging” to promote its developments in magnetic resonance (MR) imaging.

2. The medical needs

The dominating health problems are related to life style. These are heart diseases, various cancers, and diabetes. They are all influenced by nutrition, stress in daily life, and reduced physical activity. Whereas these diseases previously were dominant in the western communities, they have more recently accelerated in frequency in the highly populated regions in Asia, and they are now more frequent here than in the industrialized countries.. With economic growth comes change in life style, and the health problems prevalent of the western world 30 years ago are now those of China and India. Since these countries are becoming economically capable of meeting the challenges, it is

obvious that this will impact the priority of investment into new and better technology for the next decades.

With increasing age of the population comes an increasing demand for prostheses and maintenance of physical capability and mobility.

Growth in car traffic has raised the frequency of accidents with human injuries and permanent disability to a level of significant economic impact.

Globalization of the economy causes more travel and transportation, and spread of contagious diseases increase. It seems also that new aggressive viruses appear with increasing frequency.

Heart failure is the major cause of sudden death in the western world (about 5 mill every year).

Whereas many heart problems can be treated and cured today (almost 15000 each year in Norway at an average cost of perhaps 50000 NOK per treatment), the major remaining issue is early detection of coronary infarction, an issue addressed by multiple companies and researchers globally.

Although admirable progress has been made over the last decade in treating cancer, it remains increasing in frequency and demand for treatment. Characteristic is the growth in female breast cancer, and male prostate cancer. Each one of these now stand for around 3000 new cases every year in Norway. Various new surgical techniques are being developed together with improved radiation and biochemical therapies.

The rapid increase in diabetes is perhaps the most directly life style related curse of the wealthy society. Important risk factors for diabetes type 2 are nutrition and physical inactivity. The economic impact is significant, and large resources will be spent on information about nutrition, and to develop effective life support means for diabetic patients.

Higher average life length, increasing average body weight, and reduced physical activity all contribute to growing demands for physical mobility support.

3. Enabling technology

Information and communication technology (ICT) may contribute in several ways to meet the needs described above.

3.1 Heart diseases

More than half of the heart procedures performed today are balloon angioplasty (PTCA) of the coronary arteries. Inflammation in the coronary artery wall is an important factor in heart failure and stroke. There is no adequate method available to precisely identify and localize the inflammatory lesion. The current professional state of opinion is that the next level of technology should incorporate intra coronary imaging in combination with sensors for identification of the inflammation.

For the general practice physician to contribute more to the diagnosis and care of heart patients, tools are required which makes the decision more reliable. Today ECG is available with some decision support available on the equipment. The next level will be ultrasound imaging. However, this requires even more expertise in making the right decision. Tools which help reducing false negative decisions, and which help making the right decision for referral, will require advanced image interpretation technology, and also access to “specialist support”. This will address the issue of “point of care” technology for heart patients.

Improved individual care and treatment for these patients will come as a result of the general physician having access to data bases with more accumulated detailed knowledge of the patient, and of similar patient problems.

The active use of “learning data bases” will be an important contributions in the development and practicing of “personalized medicine”.

Minimally invasive techniques will accelerate restitution and reduce secondary effects from heart surgery. Improved surgical procedures on the beating heart using refined robot procedures will become feasible. This is facilitated by improved real time image processing for motion estimation in combination with tactile feedback devices in the surgeon/robot tool interface.

Critical enabling ICT:

- Inflammation sensors
- Intra coronary imaging
- ICT support in primary care
- Refined interface systems for surgical robots
- Improved robot surgical tools

3.2 Cancer

The most successful way of treating cancer today is by surgery. The limitations of surgical procedures are the ability to remove the complete tumor. In order to be certain the general rule is that too much is better than too little. In many cases the tumor is so entangled and complex that total removal is impossible or unlikely. Radiation and/or chemotherapy are then the only alternatives.

Improved specificity and more detailed outline of the tumor will help to better plan a successful surgery. The development of “molecular imaging” by MR is aimed at this. Robotic surgery may help in a more precise removal of malignant tissue.

Open surgery may be dramatic for the patient and lead to long and painful recovery. The more recent laparoscopic techniques will help to reduce this problem to the benefit of the patient and the society. Better sensors such as tactile feedback devices will help to refine and expand the use of minimally invasive techniques further.

The main technical challenge in cancer treatment is the ability to diagnose it at an early state. This improves the outcome of the treatment, it reduces the severity of the surgery, and it reduces the recovery time. Today there is no affordable technique available, which can be applied widely in “point of care” approaches.

Inexpensive, specific, and highly sensitive sensors for early detection of cancer are major challenges for the future.

Critical enabling ICT:

- Better sensors for early detection
- Improved imaging
- Sensors for improved minimally invasive surgery

3.3 Diabetes

Treatment of diabetes is reasonably well controlled. Implanted devices may help diabetic patients to keep better control and thus improve their life. Artificial pancreas tissue for insulin production is another area of research. The main problem for the society is the growth in frequency of the disease, which can best be met by control of the nutrition in combination with physical activity.

A significant achievement will be an affordable and reliable non-invasive sensor for easy measurements of blood sugar.

Critical enabling ICT:

- Sensors for blood sugar measurements
- Implanted insulin support devices

3.4 Reduced physical ability.

There will be a growing demand for better technical solutions to help physically disable patients. This will require better prosthetic designs based upon improved biomechanical models. The results of this will be less pain, more normal mobility, and more complete and rapid restitution after surgery.

Recent advances in neural signal interface technology give promise of a revolution in the interaction between the user and motor assisted multi-function prostheses.

Critical enabling ICT:

- Sensors for motion and tension control
- Biomechanical modeling of the musculo-skeletal system
- Prostheses modeling
- Implanted or hybrid systems for neural communication

3.5 Molecular Biology

In April 1953 Watson and Crick published the “Molecular structure of nucleic acids”. The paper was published in Nature, and has been considered to be the most significant breakthrough in “finding the secret of life” through its presentation of the double helix structure of DNA. 50 years later, in February 2001, the sequencing of the human genome was published as two parallel works in Nature and Science. This opened the possibility to advance into the post-genomic era in which genetic information could be examined in multiple health care situations throughout the lives of individuals. Currently, newborn babies may be screened for treatable genetic diseases. The extent to which this will be applied is subject to substantial ethical debate.

For the purpose of illustrating the possibilities and the coming difficult decisions, a few scenarios are as follows: In the near future, children at high risk for coronary heart disease may be identified and treated to prevent the pathological changes in their vessel walls during adulthood. Potential parents will have the option to be told their carrier status for many recessive diseases before they decide to reproduce. For middle-aged and older populations, the ability will be there to determine the risk profiles for numerous late-onset diseases, preferably before the appearance of symptoms, which at least could be partly prevented through dietary or pharmaceutical interventions. Soon, genetic testing will comprise a wide spectrum of different analyses with a host of consequences for individuals and their families.

In its wide consequence, the perfectly controlled molecular biology’s impact on health care will reduce the needs for sensors and actuators, as action will be taken on the genetic end for correcting the disturbances leading to disease. However, it is likely that the practical approach will require extensive chemical and physical monitoring of body processes along with the possibilities of genetically based treatment and repair, thereby increasing the need for sensing systems both external and internal to the body.

Critical enabling ICT:

- Biochemical and biophysical sensors

4. The NTNU perspective

Norway represents about 1/10 per cent of the population that will have access to medical technology in the future. Our investment into R&D for the development of new technology in this field is also less than one per cent of the total. It is important therefore to make the right priorities and make certain that it becomes relevant on an international scale.

NTNU has decided to prioritize medical technology. This has been supported also by national political authorities through the establishment of national centers of competence at the medical school and at the university hospital. In a recent evaluation of the medical research in Norway, groups at NTNU in medical technology were given very good rating, notably in areas where there was tight cooperation between medical and technical researchers. The work in these groups should be further supported. There is a growing activity in sports medicine at NTNU. It seems attractive to couple this activity to the research in orthopedics and prostheses, and to the research in robotics and bio-cybernetics. Thereby one might create a research environment of mutual incitement to all these activities. Most of the novel sensor technology is based upon microelectronic and optical systems for signal conditioning and processing. Research in this field therefore relies upon access to such facilities.

The plan to establish a research laboratory in nanotechnology at NTNU may become of significant interest for the research in medical technology. This facility offers a unique possibility to carry out research in novel biochemical and biophysical sensor technologies in combination with microelectronic and optical signal processing.

5. Sources:

1. IEEE Proc Jan 2004,
2. mst/news February 2004, International newsletter on micro-nano integration, Special issue on Biomicrotechnology <www.mstnews.de>
3. Time Magazine March 2004,
4. EU 6th framework programme - Life Science in focus
5. Life Sciences and biotechnology- A strategy for Europe. Prepared by the European Commission. ISBN 92-894-3388-4
6. IEEE Spectrum Nov 2000, Special Issue on Gene Sequencing's industrial revolution
7. BRAINSTORM, Newsletter of Stanford University's Office of Technology Licensing(OTL), Vol 12.NO2, on The Bio-X program: A Merger of Knowledge.
8. Interview with medical professionals
9. Science, Vol 291, 16 February 2001

Abstract of invited talk

Chief Scientist Michael Brodie, Verizon, USA

Increasingly rapid and dramatic changes in technology and business make predicting one year close to impossible and to 2020 beyond impossible. However, it is critical to attempt to plan for 2020 to overcome the persistent failures of business and technology to address real problems, such as providing holistic, human centered solutions. Due to Cartesian basis for science, technology, and much of western thought, our disjoint technologies offer point solutions, a cell phone for communicating, a computer for computing, the Web for buying books, and a Television for entertainment. While these point solutions are most often sold as solving problems while enriching and simplifying our lives, the opposite is often true. The power of ubiquitous voice and e-mail communications is overwhelmed by the volume of communications, the ever-changing and complex devices on which to receive and send them, and the ever-changing services and service providers who offer them, even if 50% were not junk. This talk attempts to provoke thoughts on how to plan for 2020 to achieve holistic solutions in the convergence of computing, communications, entertainment, eCommerce, and more. We consider traditional factors such as constantly increasing automation and the speed of change and innovation, but will focus on factors that make prediction almost impossible, including how good technology should disappear, technology convergence producing unanticipated opportunities, disruptive technologies that displace technologies and disrupt existing products and businesses, and unpredictable paradigm shifts, such as the Internet, that change everything. The talk concludes with a vision of holistic, human-centered solutions that enhance the human experience enabled by converged, Next-Generation technologies.

See <http://www.ime.ntnu.no/infosam2020/> for slide show.

Abstract of invited talk

**Professor Keith Jeffery,
Director, Business & Information Technology Dept.
CCLRC Rutherford Appleton Lab. UK**

GRID technology, emerging in the late nineties, has evolved from a metacomputing architecture towards a pervasive computation and information utility. However, the architectural developments echo strongly the computational origins and information systems engineering aspects have received scant attention. The development within the GRID community of the W3C-inspired OGSA indicates a willingness to move in a direction more suited to the wider end user requirements. In particular the OGSA/DAI initiative provides a web-services level interface to databases. In contrast to this stream of development, early architectural ideas for a more general GRIDs environment articulated in UK in 1999 have recently been more widely accepted, modified, evolved and enhanced by a group of experts working under the auspices of the new EC DGINFSO F2 (GRIDs) Unit. The resulting report on 'Next Generation GRIDs' was published in June 2003 and is released by the EC as an adjunct to the FP6 Call for Proposals Documentation. The report proposes the need for a wealth of research in all aspects of information systems engineering, within which the topics of advanced distributed parallel multimedia heterogeneous database systems with greater representativity and expressivity have some prominence. Topics such as metadata, security, trust, persistence, performance, scalability are all included. This represents a huge opportunity for the database community, particularly in Europe.

See <http://www.ime.ntnu.no/infosam2020/> for slide show.

Abstract of invited talk

**Dr. John de Kleer,
PARC, USA**

See <http://www.ime.ntnu.no/infosam2020/> for slide show.

