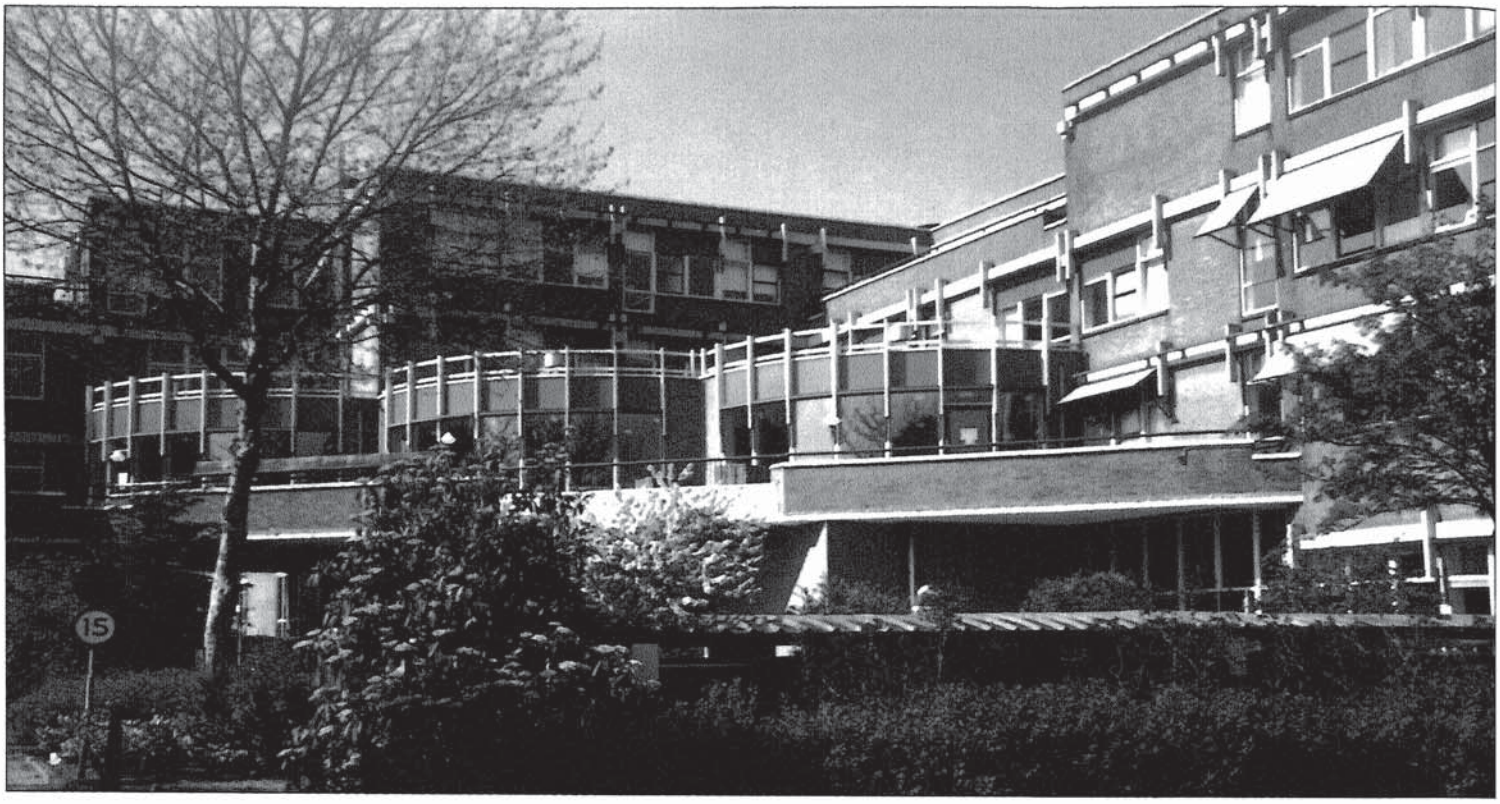
Centrum voor Wiskunde en Informatica

Annual Report

JANUARY
FEBRUARY
MARCH
APRIL
MAY
JUNE
JULY
AUGUST
SEPTEMBER
OCTOBER
NOVEMBER
DECEMBER

CWI

1998



CWI

ERCIM

Telematica Instituut

CWI is the National Research Institute for Mathematics and Computer Science. CWI is administered by the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications. SMC is sponsored by the Netherlands Organization for Scientific Research (NWO).

CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics. CWI participates in the Telematics Institute.

General Director
G. van Oortmerssen

Colophon
Issued by the Stichting Mathematisch Centrum, June 1999 ©
Production Bureau CWI
Design Tobias Baanders / Facility Department CWI
Printing drukkerij Mart.Spruijt bv, Amsterdam

Stichting Mathematisch Centrum / Centrum voor Wiskunde en Informatica
Visiting address Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
Postal address P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Telephone +31 20 592 9333
Telefax +31 20 592 4199
Website www.cwi.nl

CWI-AR 1998

Contents

This Annual Report is complementary to the Jaarverslag SMC (in Dutch), which concentrates on the Institute's management including financial and social aspects. A complete overview of CWI's research activities is also available.
This Annual Report and the other reports can be ordered at:
Mrs. D.C.M. Amende; *phone* +31 20 592 4128, *e-mail* Thea.Amende@cwi.nl

OVERVIEW

The century approaches its end and never the turn of a century was marked so literally by a scientific development in its preceding decades. The millennium problem is almost symbolic for the computerization of society. Also scientific practice has been deeply affected by the computer, not only its use in research itself, but also the way in which this research is organized. Large-scale cooperation is seen more and more often, which would be unfeasible in practice without fast electronic communication. Indeed, cooperation is a positive requirement of society, which does not like its researchers to operate from ivory towers. CWI is excellently equipped to operate in this climate. Its researchers are alert in taking up new interesting challenges, knowing themselves supported by, e.g., a very advanced computing infrastructure. For example, the rapidly increasing research needs from the financial world triggered at CWI a substantial effort in that area. During the report year five projects were in progress related to financial problems, ranging from (quasi-) Monte Carlo simulations in connection with derivatives to data mining of financial time series. The visual presentation of information is rapidly gaining popularity as mankind, well-equipped with visual information processing capabilities, is increasingly faced with information overload and computer processing power no longer poses heavy constraints on visual presentation. CWI recognized the importance of this trend already years ago and now runs several projects in visualization of complex systems, multimedial information systems and presentations, and image processing. Other examples include evolutionary computation and neural networks, with applications in, e.g., planning large-scale air traffic and classification of remote sensing images. This, and other research, and the frameworks in which it takes place, will be detailed in the survey below of CWI's activities during 1998.

Telematics Institute

The national Telematics Institute (TI) was officially inaugurated in Enschede on February 11, the very day of CWI's foundation in 1946. It was the first of a series of 'Leading Technological Institutes' initiated by the government (viz., the ministries of Education, Culture & Science, of Economic Affairs, and of Agriculture, Environmental Protection & Fishery). Here knowledge institutions like CWI closely cooperate with Trade & Industry and governmental bodies. The Telematics Institute is built around the former Telematics Research Centre. Participating knowledge institutions are the universities of Twente and Delft, the organization for Applied Scientific Research (TNO-MET), and CWI. In addition, seventeen companies are involved. So far, CWI secured participation in no less than seven projects financed by TI (see Table 1), actually not surprising in view of CWI's long standing expertise in areas as queueing theory, databases, and multimedia – all of prime importance to telecommunication. Hence, CWI expects to merge through this channel its fundamental knowledge with the industrial partners' domain knowledge, which is a necessary basis for creating innovative products.

Quality-of-service in future networks
(Borst) (CTIT, KPN, Lucent)

Domain Specific Languages
(van Deursen/Klint) (TI, ING, CAP Gemini, Lucent)

Systems Validation Centre
(Groote) (CMG, CTIT, KPN, Lucent, TI)

Digital Media Warehouse Systems
(Kersten) (UT, KPN, Syllogic)

DRUID – Multimedia Indexing and Retrieval on the basis of Image Processing and Language & Speech Technology
(Kersten) (TNO, UT)

U-WISH – Web-based Services for Information and Commerce
(Pemberton) (TNO-TM, CTIT)

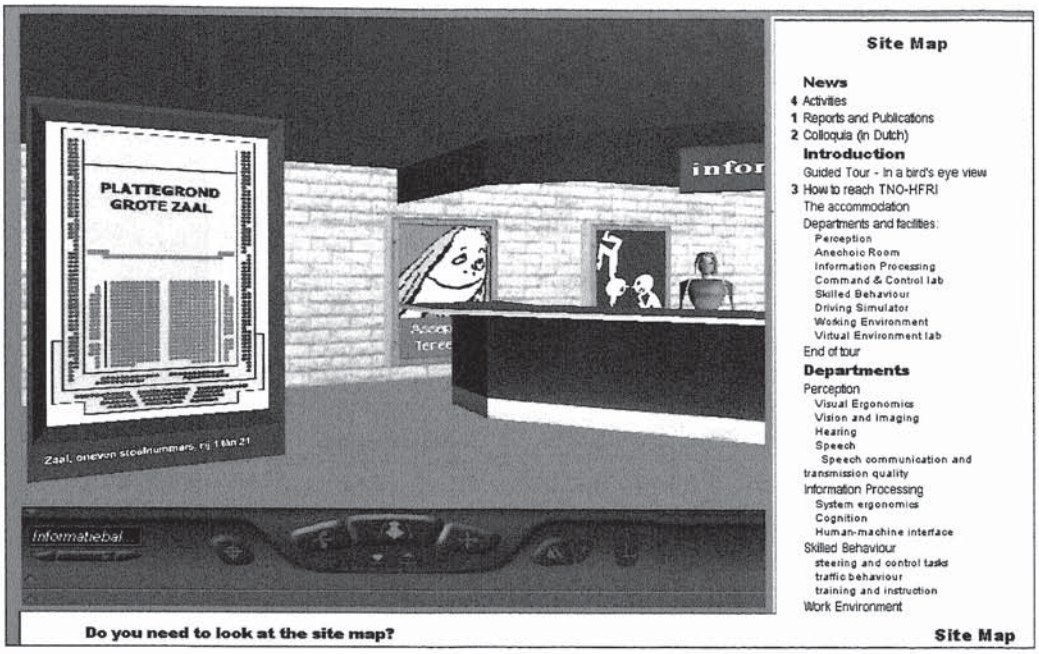
Mediated Communication
(ten Hagen) (TNO, TUD)

Table 1: Telematics Institute projects with CWI participation approved in 1998. In brackets the CWI project leader and the partners.

Multimedia presentations on the Internet

GRiNS (Graphical Interface for SMIL) is an authoring system for multimedia applications on the Internet, recently developed at CWI. The specification language SMIL (Synchronized Multimedia Integration Language) was accepted in June 1998 by the W3C consortium as a standard. (W3C comprises several hundreds of research institutions and companies involved in the development of the Web all over the globe. CWI hosts the local W3C office in The Netherlands.) SMIL, used in connection with other W3C standards such as HTML and CSS, enables a Web author to develop applications comparable to multimedia presentations on CD-ROM systems. SMIL was developed in the W3C working group on synchronized multimedia SYMM, in which CWI played a key role. Several leading streaming media companies such as Philips Research, RealNetworks, Netscape, Lucent, and Compaq (Digital) participated in this work. SMIL's theoretical foundation is the Amsterdam Hypermedia Model which was designed at CWI some years ago. The authoring system GRiNS was used to test SMIL and experimental applications were built with CWI's authoring and editing tools. CWI's effort for SMIL was funded in part by the ESPRIT-IV project CHAMELEON which ended in November 1998. GRiNS has the advantage of providing a uniform interface for composing Internet presentations in the leading standard formats for hypermedia and multimedia. Moreover, its user-friendliness enables even non-programmers to compose complex presentations without expert help. GRiNS protects return on investment, in that one presentation can be produced in various formats, thus facilitating users to keep up with the dynamic market developments of multimedia usage in computer networks. The business plan for further developing GRiNS won in the fall of 1998 one of the three main prizes in the McKinsey New Venture '98 competition. GRINS is now being marketed by the spin-off company Oratrix.

http://www.oratrix.com



The Telematics Institute sponsors a project, UWISH, in which CWI studies the usability of the interface of Web-based applications, both theoretically and in several types of applications, in particular by studying the effect of interface adaptations on various websites. The final report on usability will itself be presented as a web application.

ICES

The ICES-KIS programme (ICES = Interdepartmental Committee for Economic Structure Reinforcement, KIS = Knowledge InfraStructure) concerns a joint effort of a number of ministries, including those of Education, Culture & Science and Economic Affairs, in order to reinforce the Dutch research infrastructure and its interaction with society and Trade & Industry in particular. It is financed from the government's revenues from natural gas exploitation. CWI is involved in two large-scale projects approved in 1998. The first, Brainport, to be carried out with partners of the WTCW Science Park (WTCW = Science & Technology Centre Watergraafsmeer, CWI is located on its premises) receives 30 Mfl as support for Research & Development. CWI will focus on multimedia information engineering and the modelling of biological processes. The second project, Gigaport (innovative internet technologies and applications), consists of two parts: Gigaworks and Giganet. CWI is involved in both. Work in Gigaworks will be carried out with partners in the Telematics Institute.

ERCIM

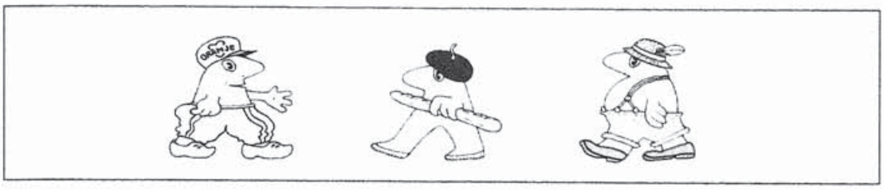
A third large-scale cooperative framework is the European Research Consortium for Informatics and Mathematics (ERCIM), in which at present national research institutions from fourteen European countries are represented (see Table 2). This consortium was founded in 1989 by CWI, jointly with GMD (Germany) and INRIA (France). In 1998 CWI's general director, Gerard van Oortmerssen, was elected President of ERCIM. Also bilateral cooperation was started with GMD, followed by preparations for the same with SZTAKI (Hungary). Cooperation with INRIA already existed since

ERCIM
years
1989 - 1999

Table 2: Present ERCIM Membership (31 December 1998)

Organization	*Country*	*Research staff (informatics/mathematics)*	*Member since*
CWI Centrum voor Wiskunde en Informatica	The Netherlands	185	1989
GMD Forschungszentrum Informationstechnik	Germany	1250	1989
INRIA Institut National de Recherche en Informatique et en Automatique	France	2100	1989
CLRC Central Laboratory of the Research Councils	United Kingdom	290	1990
CNR Cosiglio Nazionale delle Ricerche	Italy	600	1991
SINTEF Stiftelsen for Industriell og Teknisk Forskning ved Norges Tekniske Hogskole (SINTEF Telecom and Informatics)	Norway	350	1992
FORTH Foundation of Research and Technology – Hellas (Institute of Computer Science)	Greece	160	1992
SICS Swedish Institute of Computer Science	Sweden	80	1992
VTT Technical Research Centre of Finland (VTT Information Technology)	Finland	200	1993
SZTAKI Magyar Tudományos Akadémia – Számítástechnikai és Automatizálási Kutató Intézete	Hungary	350	1994
SARIT Swiss Association for Research in Information Technology	Switzerland	800	1994
CRCIM Czech Research Consortium for Informatics and Mathematics	Czech Republic	250	1996
DANIT Danish Consortium for Information Technology	Denmark	80	1996
SRCIM Slovak Research Consortium for Informatics and Mathematics	Slovakia	145	1998

Two members have left ERCIM meanwhile:
AEDIMA (Spain) 1993–1996, and INESC (Portugal) 1991–1998.



Signing the ERCIM Agreement on Scientific Cooperation in April 1989 at Schloss Birlinghoven (GMD): (from left) Alain Bensoussan (INRIA), Friedrich Winkelhage (GMD), Cor Baayen (CWI). Photo: S. Münch.

the 1980s in the framework of a general research agreement between France and The Netherlands. The joint projects with GMD started in the report year cover embedded systems, data mining, environmental modelling, computational fluid dynamics, and scientific visualization. Last, but not least, CWI will host the activities to celebrate ERCIM's 10-year jubilee in 1999.

W3C

On the global level CWI is active in the World Wide Web consortium W3C. This consortium, joined meanwhile by several hundreds of companies and institutions, plays a guiding role in the further development of the Web. For this purpose it created several working groups, in a number of which CWI plays an active role. Here we mention the work on the multimedia specification language SMIL, to which CWI made significant contributions, and which acquired in June 1998 the status of W3C Recommendation, and the working group XHTML, chaired by CWI researcher Steven Pemberton. W3C has offices in several countries; the Dutch office is located at CWI. In addition CWI hosts a 'mirror site' for the W3C server. Finally CWI has the honour of organizing in May 2000 in Amsterdam the large annual international WWW conference.

European and national projects

CWI's participation in new cooperative frameworks like the Telematics Institute and ICES-KIS has an added value because the delay in the EU Fifth Framework Programme causes a temporary dip in income from European projects. Even so, CWI still participated in 1998 in some twenty European projects and networks. One of the

Symbolic Techniques for Program Analysis

As a consequence of our ever increasing dependence on the proper functioning of software systems, the need for proving their reliability has to be taken serious indeed. Until recently such proofs were only feasible for systems whose size is far below what is found in practice. Recently developed techniques, however, may offer a solution. The aim is now to apply these techniques effectively, so that real-life systems can be successfully analyzed. In principle there are three approaches: proof by manual labour, fully automated proof techniques, and a mix between the two. CWI has shown that the last way enables the analysis of realistic systems. Manual proof can be carried out in the context of Process Algebra (PA). The use of PA has advantages above other verification methods such as modal or temporal logic because of its high level of abstraction and its composition properties. On the basis of PA, CWI developed μCRL (micro Common Representation Language). The idea was to create a basis for sharpening symbolic techniques, rather than adding another language to the repertoire. With μCRL one can carry out proofs manually following strict logical rules. In practice, however, several such proofs remain sloppy, because the manual method is effective only for small systems, not exceeding one page of code. Fully automated proof techniques are usually based on state automata. At CWI now systems with 10^8 states can be dealt with (in general the limit is 10^6), but realistic systems are still considerably larger (for example 10^{1000} states). Of course, it is of the utmost importance to find ways to reduce the number of states. CWI research indicates that by transforming processes described in μCRL to a normal form (Linear Process Operation) using rewriting techniques (automated induction, tree automata), exponential reduction of the number of states can be reached. By using proof checkers, which guarantees the required precision, in combination with manual control, CWI has shown that this approach can be effective for middle-sized systems. One may compare this hybrid technique with the way packages such as MATLAB and MAPLE are used in mathematical formula manipulation. This approach to putting formal proof techniques to practical use may in due course very well lead to a revolution in mathematical argumentation, as was foreseen already some thirty years ago by the eminent Dutch mathematician N.G. de Bruijn when he created his Automath system. Meanwhile several instances of faulty software have been revealed by applying formal proof checkers under manual control. Recent Dutch examples include the automated control system for the legs of car lifting installations in garages, and for the doors of the flood barrier in the Nieuwe Waterweg which protects the Rotterdam area by closing the doors in case of flood. Since many more such instances can be expected to show up in the near future, we may see before long the birth of a new profession: that of software prover.

http://www.cwi.nl/~jfg/

Software controlling the flood barrier near Rotterdam harbour was checked with formal methods.
Photo: Rijkswaterstaat, Meetkundige Dienst Afdeling Grafische Technieken, Delft.

projects concluded in the report year was CHAMELEON (multimedia document processing), to which CWI made a substantial contribution with the development of the authoring system GRiNS for multimedia applications on the Internet. Also the work on the ISO-standard PREMO (Programming Environment for Multimedia Objects) was concluded with an implementation and a description, to be published in 1999. This standard emerged from the European project MADE which ended three years earlier and in which CWI played a key role. On the national level CWI participated already from the outset in the ICES-HPCN programme for high performance computing and networking, which started in 1996 as an initiative of the ministry of Economic Affairs, with the comparatively large number of six projects, in the fields of visualization, environmental modelling, and databases (data mining, database management). This programme ends in 1999.

Apart from direct basic funding through NWO a substantial part of CWI's research funding is acquired in competition in the form of projects in special NWO programmes. It concerns a total of forty projects. Other support comes from bilateral cooperative agreements which NWO has concluded with Hungary and the Royal Netherlands Academy of Arts and Sciences with Indonesia. Finally CWI receives income from several research commissions by governmental bodies and Trade & Industry. Such commissions may range over several years, such as at present CWI's work on 'the 21th century railroad timetable', commissioned by Dutch Rail (NS).

Spin-off companies

One of the fixed dates in CWI's agenda is *CWI in Bedrijf*, the annual day at which CWI since 1992 presents its research to Trade & Industry. In the report year, the event drew some 110 participants. Its central theme was CWI's spin-off policy and practice. Creation of spin-off companies is an important instrument for institutes like CWI to turn its fundamental knowledge into market-oriented applications, providing at the same time high-tech job opportunities. After careful preparations the fall of 1998 saw the birth of two spin-off companies. Oratrix, headed by Dick Bulterman, focuses on further development and marketing of the software product GRiNS (Graphical interface for SMIL – Synchronized Multimedia Integration Language, a specification language developed in the W3C working group SYMM), which was developed at CWI. Eidetica, led by Annius Groenink, provides linguistic and mathematical expertise for the development of software supporting information and knowledge managers.

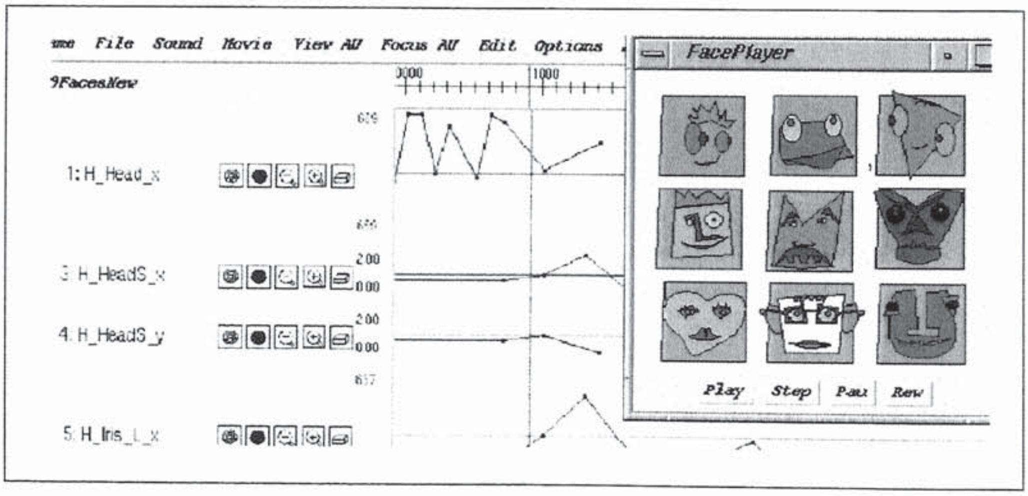
The financial sector

CWI's research activities with a financial component cover several fields. In the theme Financial Mathematics, started early 1998, Monte Carlo and quasi Monte Carlo simulations for efficient valuation and risk assessment of financial derivatives are studied, as well as statistical inference for stochastic processes and fractal analysis of financial data. This research is covered more in detail elsewhere in this Annual Report. Furthermore CWI, in cooperation with the MeesPierson bank and Cap Gemini, from 1992 on designed, formalized and implemented the domain-specific language RISLA for the description of financial products, in particular interest rate products. Meanwhile RISLA became a commercial success, which has led to a long-term cooperation with Cap Gemini. In the IMPACT project, part of the national HPCN programme which started in 1996, parallel and distributed database technology is applied to problems in the financial sector. CWI research concentrates on data mining of financial time series. Commissioned by ABN AMRO bank and software house Roccade, CWI performed

CharToon – a JAVA-based animation system

The 3-year project FASE (Facial Analysis and Synthesis of Expressions), started in spring 1997, has entered its second phase. The aim is to develop a system which can recognize facial movements in front of a camera and use the data thus produced to animate models of the human face. Basic editing tools have been implemented in the CharToon system, consisting of a collection of JAVA programs enabling the interactive construction of parametrized 2D drawings and a set of time curves to animate the drawings. The project is a joint effort of Delft University of Technology (TUD) and CWI and is sponsored by the Dutch Technology Foundation STW. FASE integrates two technologies: Facial Analysis (FA) and Facial Synthesis (FS), in order to achieve life-likeness of the estimated facial information and motion. The FA module, covered by TUD, receives as input video images of a face (the user) and possibly additional information (e.g., text and audio sources), on the basis of which the parameters describing the facial movements and expressions are estimated. The FS module, developed at CWI, contains a generic facial model and an animation model. It uses the estimated facial parameters as an input to deform and move the facial model and thus mimic the looks and movements of the user. The results are used as a feedback input for FA. In addition, a third module, Model Editing and Parameter Control Interface, not only facilitates the development and set up of the (face) models and the generation of new animation models, but also provides easy control over the whole process of changing facial parameters – a highly desirable feature, in view of the close interplay between analysis and synthesis. The integration of FA and FS is also done at CWI. The FS module allows two kinds of facial models. The 3D physically based models aim at realism, and are of interest to the television and telecommunication industries (efficient image coding). Second, 2D drawings, aiming at cartoon-like effects, stylized avatar faces or possibly also animation of 2D photographs of a real face, are of interest to the film industry (animation) and in man-machine interaction (social user interfaces: avatars). Other applications include: entertainment (games, interactive television, synthesized actors), lip synchronisation (text or voice based, possibly also adjusted for the deaf), and facial surgical planning and simulation. For the 2D models we developed a Face Editor to design the drawing to be animated, an Animation Editor to specify the time-behaviour of the drawing's animation parameters, leading to a movie script, and a Face Player to animate the drawing using this script. These three tools are integrated in the CharToon system. The implementation in JAVA makes the system suitable for Web applications.

http://www.cwi.nl/FASE

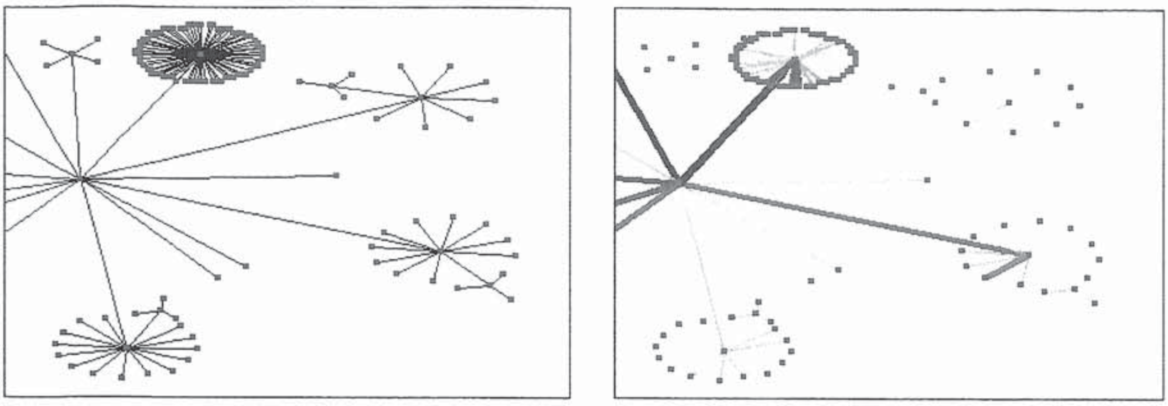


Animation Editor, showing graphical parameter specifications and their effect on faces, created by Face Editor.

research in software renovation in the Resolver project, which started in 1996. Jointly with the University of Amsterdam and ID Research, and with support from the ministry of Economic Affairs, the research included case studies of problems connected with Y2K, the Euro, and the migration from COBOL to OO-COBOL. This project was concluded mid-1998 and yielded a toolset for the analysis of real-life COBOL legacy systems. A continuation of this work has been proposed.

Visual information

Another important research area at CWI is covered by the term 'visual information'. The rapid development of this area originates in the increasing need and possibilities to render and process information visually. At CWI several projects come under this denominator. The study of complex physical phenomena, such as turbulence, requires

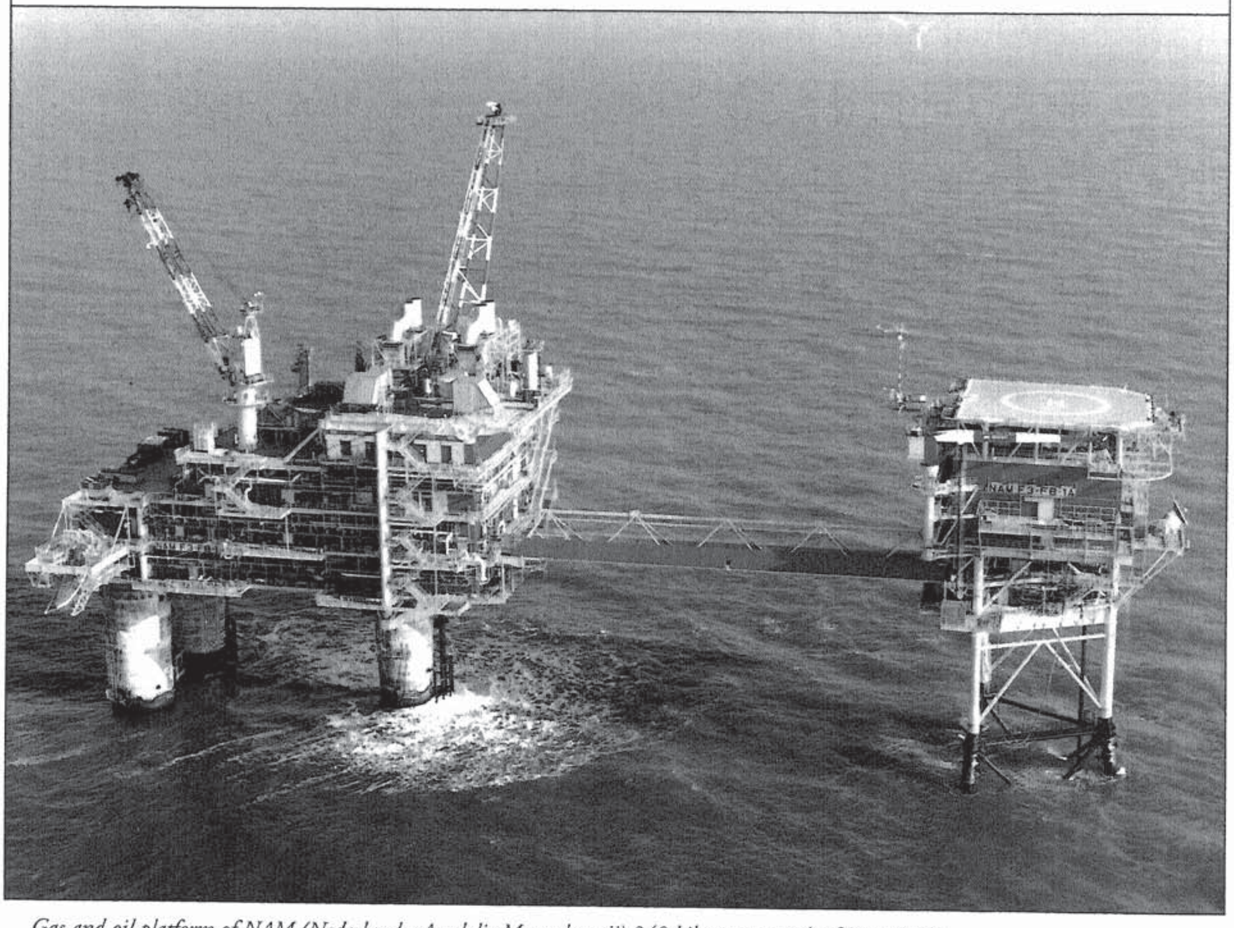


Using graph metric values, and visual tools like colour saturation, linewidth, etc, the package Latour, developed at CWI, can highlight the "back-bone" of a tree, i.e., those edges which hold larger, more complex subtrees. The image on the left is a simple zoomed-in image: the user barely knows where to move with the pan, if complex areas are searched; it is also not clear whether the node on the left which looks like a root node is indeed the root of the tree. The right-hand image shows the same portion of the graph, with the metric based visual cue. It clearly shows, for example, that the node on the left is indeed the root and that one of the edges going toward the left leads to a complex portion of the tree, whereas the other one is probably less interesting.

visualization of the underlying computational models (see elsewhere in this Annual Report). Visualization of large graphs is important in the study of, e.g., large-scale parallel computation processes and internal data structures of compilers or Virtual Reality scenes. CWI research into multimedial databases and information systems received substantial momentum during the report year. An important subject is searching large picture databases for certain image characteristics. Work on multimedia standards, which led to the ISO standard PREMO, was followed by the research into multimedial presentations on the Web already mentioned before. In the area of signal and image processing one now realizes that for a good understanding analysis over a broad range of resolutions is necessary. At CWI images are studied using three techniques in which this multi-resolution aspect is clearly present: fractals, wavelets, and morphology. The present focus is on fractal image coding applied to multimedial databases, use of the wavelet-Radon transform developed at CWI, and the construction of non-linear (morphological) wavelets. Finally CWI also pursues research into stochastic geometry. The resulting techniques are applied in, e.g., image analysis, for example the recognition of objects in noisy images.

Mathematical analysis for gelation near a well

The oil which is found in the North-Sea resides in layers (called reservoirs) about three kilometres under the bottom of the sea. To retain the oil, (brine) water is injected into the reservoir through an injection well to have the oil flow out of the reservoir through a production well. Due to inhomogeneity of the reservoir, the layers with a higher permeability are depleted soon. Since the injected water tends to flow mainly through the well-permeable (depleted) layers, the water production increases enormously. In order to shut-off these highly permeable layers in the reservoir, polymers and crosslinkers in aqueous solution are injected. The polymers, seen as large thin molecules, react with the crosslinkers and form a gel. The gel is adsorbed in the reservoir by which a decrease of the permeability is obtained. Under certain chemical circumstances, the crosslinker may precipitate and hence the gelation reaction may be delayed or even stopped. In the mathematical formulation of the problem, we consider the advective transport of the polymers, crosslinkers and gel, the gelation reaction, gel adsorption and crosslinker precipitation. We also take the adsorption kinetics into account, i.e. equilibrium is not instantaneously assumed. The resulting set of hyperbolic transport-reaction equations can be solved numerically using a higher order upwind finite volume scheme. In our research emphasis lays on the analytical study of the problem. For simple cases, we obtained an analytical travelling wave solution. For more complicated situations, when we are not able to find an analytical solution, we analyse the problem more qualitatively. The insights developed in this research can be used to validate the numerical solutions obtained from commercial simulators. The work is carried out in collaboration with Delft University of Technology (Department of Petroleum Engineering) and is supported by the project WELGEL.



Gas and oil platform of NAM (Nederlandse Aardolie Maatschappij) 240 kilometres north of Den Helder.
Copyright: Sky Pictures.

Academia

An important life-line for CWI is its extensive network of relations with the academic world. With its interdisciplinary research themes, derived from societal problems, CWI is positioned complementary to academic research groups, at the same time maintaining intensive contacts with these groups. The majority of CWI's projects involves one or more partners at a Dutch university. In the report year 22 CWI researchers had a part-time appointment at a Dutch university, including 18 full professors. Conversely, 20 academic researchers spent a part of their time at CWI, among them 7 as advisor. Eight young CWI researchers successfully defended their Ph.D. thesis at a Dutch university. CWI has concluded cooperative agreements with all seven Dutch research schools active in the research fields covered by CWI. In addition CWI participates in the programme of EURANDOM, the international research institute for statistics, probability theory and operations research located at Eindhoven University of Technology (TUE). Finally CWI concluded agreements with TUE and the Free University of Amsterdam separately concerning mutual secondment of researchers.

Dynamics

The internal document MOBILE (1992) formed the starting point for the FIT operation (Flexible, Interdisciplinary, Thematic) which was carried out in 1996. From that time, in the mid-1990s, CWI's research field was increasingly set in motion, requiring more alert adaptation to interesting new developments. This has led to the termination of research into biomathematics, cryptography, computer algebra, and dynamical systems, not because these subjects were exhausted (on the contrary), but rather because CWI had made innovative contributions to them for years and preferred to continue doing the same in other, newly emerging areas. Hence, research in environmental modelling (Hans van Duijn, Jan Verwer) and in signal and image processing (Mike Keane, Henk Heijmans) was considerably reinforced, and new research was started a couple of years ago in Quantum Computing (Paul Vitányi), followed by Evolutionary Computation (Han La Poutré), and most recently by the aforecited Financial Mathematics (Hans Schumacher) project.

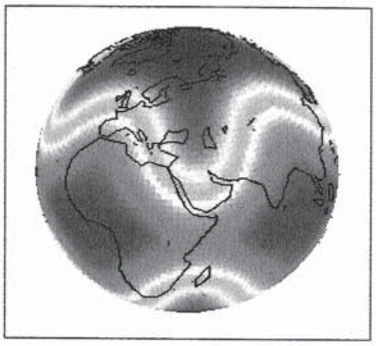
Interest from society

CWI enjoyed considerable attention and appraisal from society during the report year. Several politicians wanted to be informed about the developments at WTCW Science Park, and CWI hosted them on some occasions. Minister Ritzen of Education, Culture & Science, and his successor Hermans, minister Jorritsma of Economic Affairs, and member of the European Parliament ms Plooij were successively received. This interest clearly contradicts the generally prevailing impression that science in The Netherlands is threatened to become a suppositious child in the eyes of politicians. Furthermore CWI hosted the Dutch computer science community for an exchange of thoughts with Tony Hoare, one of the pioneers of theoretical computer science. The meeting followed Hoare's Huygens Lecture, organized by NWO.

Numerical Atmospheric Circulation

For long, people have tried to forecast the weather. In the early years, predictions were made by observation of current and historical meteorological data. Nowadays, forecasting is mostly done by numerical simulation with circulation models based on the atmospheric primitive equations. During the sixties, the approximation methods in circulation models mainly consisted of gridpoint methods, including finite difference, finite volume and finite element methods. With the introduction of the spectral transform method in global atmospheric modelling by Orszag and Eliasen, in 1970, this accent shifted. Because spectral methods proved to be very accurate and efficient, they started to dominate the field. Recently though, the discussion has been renewed, and spectral methods are no longer considered ideal. Progression in atmospheric modelling, on the meteorological as well as on the computational side, demands much higher grid resolutions than used in the past. At high resolutions, however, other methods become competitive to the spectral ones. Furthermore, in some cases, the spectral method contributes to unphysical structures in the numerical solution. Finally, the global character of spectral methods has also shown some drawbacks, for example in parallelizing spectral codes on parallel machines with distributed memory. In 1998, CWI started a co-operation with ERCIM partner GMD aimed at developing a new numerical gridpoint method for high-resolution, future generation circulation models. Initially the focus is on the 2D spherical shallow water equations. These equations are widely acknowledged as a first prototype of a circulation model. To avoid the problem of pole singularities, which arise when gridpoint methods are applied on a full longitudinal-latitudinal grid, a stereographical grid in the polar areas is used. To respect the upwind character of the equations on the two space grids, an Osher-type approximate Riemann solver with a higher order accurate state interpolation is applied. Advantages of the Osher scheme are its robustness, and its logical extension to more realistic primitive equations containing also the energy balance. In the next stage time integration issues will be considered, again for the 2D spherical shallow water equations. An efficient time integrator must be able to cope with the large CFL numbers emerging from the polar areas. For this purpose special operator and time-splitting techniques will be studied. A clear demand is that a successful technique must be extendible to the full set of primitive equations.

http://www.cwi.nl/-gollum/MAS1.1/GOA.html



Rossby-Haurwitz waves are frequently used as meteorological tests for the Shallow Water equations which can be used to describe atmospheric flow. Though not analytically known, their flow pattern is well-recognizable. This illustration shows the height of the atmosphere for a Rossby wave with wave number 4. The height runs between 8000 and 10560 metres with red indicating a high height region and blue a low one.

Factoring record starts attack on RSA code

A 186-digit number was factored at CWI early September 1998. This number, compactly written as $32633^{41} - 1$, was factored with the Special Number Field Sieve, a method particularly suited to numbers of this form. The previous factoring record with this method, a 181-digit number, was realized about one year ago, also at CWI. Although apparently not a spectacular progress, the acquired experience enables CWI researchers Stefania Cavallar, Peter Montgomery (visitor from the USA) and Herman te Riele to attack now a 512-bit RSA key – the present most prominent challenge in this field using the still more powerful General Number Field Sieve method. This will be done in cooperation with Microsoft and a few other institutions disposing of abundant computing power. The RSA cryptosystem, still considered as virtually uncrackable, is based on the difficulty to factor certain large numbers. The researchers expect to finish the job still in this millennium. The factoring was completed within one and a half month using 88 SGI/Cray computers at CWI and the Cray C90 supercomputer at the Academic Computing Centre SARA. The factored number was not just another toy for CWI's number addicts. Last year Warren D. Smith of NEC Research Institute in Princeton asked CWI – a world leader in this field – to factor this number for him. At NEC Research one is interested in fast and reliable algorithms generating random number sets which are applied, e.g., in physics (experiment simulation) and the financial world (assessing the value of investment portfolios). Testing the quality of certain random number generators requires the knowledge of the prime factors of numbers having the form of Smith's number.

http://dbs.cwi.nl/cwwwi/owa/cwwwi.print_projects?ID=12

Minister Jo Ritzen of Education, Culture and Science, visiting the Science Centre Watergraafsmeer in February, is welcomed by CWI director Gerard van Oortmerssen. Photo: Henk Thomas.

The eminent computer scientist Tony Hoare during the 1998 NWO Huygens Lecture. Photo: Arie Wapenaar.

Awards

Several CWI researchers received awards during the report year. Hans van Duijn, leader of the Modelling, Analysis & Simulation cluster, and professor in the Mathematical Analysis of Flows through Porous Media at Delft University of Technology, received in Bonn (Germany) the Max Planck Research Award for International Cooperation, in recognition of his outstanding research results over the years. The group researching multimedia authoring systems and distributed applications, led by Dick Bulterman, won one of the three main prizes in the McKinsey New Venture '98 competition (over six hundred participants) for the best business plan made by researchers aiming at a future career as an entrepreneur. The group, consisting of Dick Bulterman, Lynda Hardman, Sjoerd Mullender, and Jack Jansen, won the prize with the afore-mentioned authoring system GRiNS for multimedia applications on the Internet. Ronald Cramer received the Christiaan Huygens Award, granted by the Royal Netherlands Academy of Arts and Sciences for the most innovative Ph.D. thesis over the past three years in The Netherlands in the field of Information & Communication Technology. Cramer worked at CWI under Paul Vitányi. After having completed his Ph.D. thesis on secure, yet practical cryptography systems he went early 1997 to ETH Zürich to continue research in this direction. Jointly with Victor Shoup (IBM Research Laboratory, Zürich) he devised a 'hacker-proof' encryption system which protects, e.g., Internet transactions against so-called active attacks. The result was announced at the Crypto'98 conference held in Santa Barbara, California, and received wide publicity. Finally, Debby Lanser, now a Ph.D. student in CWI's research theme on Numerical Algorithms for Air Quality Modelling, led by Jan Verwer, received at Delft University of Technology the annual prize for the best Master's Thesis in the field of Technical Mathematics. The work concerned the modelling and computing of a complex industrial flow problem.

Several CWI researchers received awards in 1998: Hans van Duijn (1, left); Dick Bulterman, Jack Jansen and Lynda Hardman (2, from left to right, Sjoerd Mullender not present) receiving their award from Morris Tabaksblat (Chairman and CEO Unilever N.V.); Ronald Cramer (3, left); and Debby Lanser (4, 8th from left).

Organization

Finally we mention some developments in the organization. As a consequence of the reorganization of NWO, which was completed in 1998, CWI now reports directly to the NWO Board. Following this reorganization CWI will be evaluated in 1999 by an international panel. Preparations to provide its members with the necessary information, including a Progress Report 1993-1998 and a Strategy 2000-2005 document, were in full swing at the end of 1998. The NWO foundations SWON and SION, for academic research in mathematics and computer science, respectively, were dissolved and their duties transferred to advisory bodies coming under NWO's Board for the Sciences. After careful preparations a CWI Advisory Board was installed (see appendix Organization). Onno Boxma, leader of CWI's Probability, Networks & Algorithms research cluster, was appointed professor at Eindhoven University of Technology from

September 1. He remains attached to CWI as an advisor. As a cluster leader he was succeeded by Lex Schrijver.

Rapid changes in society, for example in the field of (tele)communications, give rise to several new research questions in mathematics and computer science. CWI's interaction with the external world through an extensive network of cooperative frameworks, forms one of the best guarantees that its research direction remains on course and that its research results are disseminated as broad as they deserve. The Strategy 2000-2005 document radiates CWI's ambition to play an even more important role in the new century than it is already doing in the current one. The institute: its researchers, its supporting staff, and its management, are ready for it. Or, after the words of Winston Churchill: 'Give us the tools, and we will do the job'.

Gerard van Oortmerssen

General Director

Resource Allocation
in Integrated-Services Networks

Research Project: Communication and Computer Networks
Researcher: S.C. Borst
E-mail: Sem.Borst@cwi.nl

Introduction

Service providers race to develop next-generation communication networks, integrating a wide variety of services, such as voice, video, and data, onto a common infrastructure. While offering potential synergies, however, the integration of heterogeneous service classes also raises several fundamental issues. Different services not only have drastically different traffic characteristics, but also extremely diverse quality-of-service requirements. The heterogeneity in services requires advanced allocation mechanisms to govern the efficient usage of network resources. Performance evaluation research at CWI focuses on models and techniques from queueing theory for evaluating the relevant quality-of-service measures, such as delay performance. Important issues in integrated networks which have received particular attention include the impact of long-tailed traffic characteristics and the interaction between best-effort and real-time services.

Integrated-services networks

World-wide, the use of communication services is experiencing revolutionary growth. The growth is fueled not only by the expansion of conventional telephone services, but also the advance of data communications, the spectacular development of the Internet, and the proliferation of wireless communications.

Driven by these demands, service providers rush to enhance their networks. A major trend is the integration of voice and data services. Telephone operators work to accommodate data services onto their networks so as to improve connectivity between company branches and provide better Internet access capabilities. Cable companies have started to offer phone services, Internet access, and video-on-demand. Internet standards are being developed to support real-time services, such as telephony and videoconferencing.

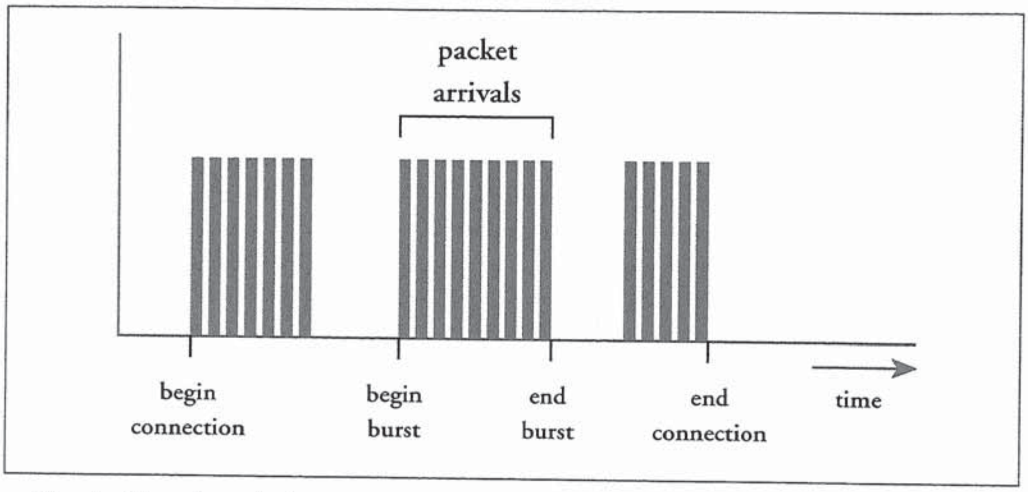
It remains to be seen how successful these efforts will be. The telephone network has been highly customized to carry voice calls, cable systems were specifically designed for broadcasting, and only great visionaries may have anticipated the Internet ever to support real-time services. The inherent limitations of legacy systems offer opportunities for start-up companies to leapfrog by capitalizing on the latest technology.

Traffic characteristics

Eventually, the current trend is expected to result in the consolidation of a wide variety of services onto a common platform. While offering potential synergies, however, the integration of heterogeneous service classes also involves several fundamental problems.

First of all, different services may have radically different traffic characteristics. To characterize traffic processes, it is convenient to adopt a three-level hierarchy. The highest level is the *connection* level, or call level, at which connections are established and terminated. (For example, a connection is set up when one dials to make a phone call.) The lowest level of granularity is the *packet* level, or cell level, at which, for the duration of the connection, packets are generated and transmitted through the network. (For example, during a phone call, a stream of packets are transmitted, each carrying a speech fragment.)

Packets are typically generated in bursty patterns. (For example, during a phone call, speech bursts usually alternate with silence periods. In Web browsing, clicking on a page triggers such a burst.) This bursty behaviour defines an intermediate level, the *burst* level. We know from experience that some persons produce larger and longer speech bursts than others. If one contrasts speech with Web browsing, it is not difficult to imagine that burst characteristics across different services exhibit even greater variation.



Three-level hierarchy: packet-burst-connection.

The three-level hierarchy described above induces an ordering of the associated time scales: packets arrive on a very fast time scale, bursts at a moderate rate, and connection requests on a relatively slow time scale. For the purpose of analysis, it is convenient to assume a complete separation of time scales. For example, a burst typically contains a large number of packets, so that at the burst level, packet arrivals occur at a very fast rate, and may approximately be modelled as a fluid flow. Also, a call usually consists of a large number of bursts, so that at the burst level, the population of calls is slowly varying, and may approximately be assumed to be fixed.

Quality-of-service

Besides different traffic characteristics, different service classes may also have extremely diverse quality-of-service requirements. One major quality-of-service measure is the amount of packet delay incurred during transmission. A second crucial performance characteristic is the fraction of packet loss caused by buffer overflow. Voice traffic is rather sensitive to delay – if packets are delayed by more than a few hundred ms, conversation is virtually impossible – but can sustain some packet loss, thanks to the error-correcting capabilities of human hearing. In contrast, data traffic can tolerate some amount of delay – file transfers for example are usually not that time critical – but is quite vulnerable to packet loss: if a few bytes are altered, that may have dramatic consequences.

Resource allocation instruments

The heterogeneity in quality-of-service requirements and traffic characteristics requires sophisticated allocation mechanisms to regulate the usage of network resources (link bandwidth, buffer space). The main resource allocation instruments include *admission control*, *routing*, *scheduling*, and *flow control*. The admission control strategy, which operates at the connection level, determines for each connection request whether it should be accepted or rejected, depending on the amount of network congestion. In case of acceptance, the routing algorithm identifies a path of network links for the packets to traverse. Routing may help balance the load across the network. The scheduling mechanism, which operates at the packet level, determines the sequencing of packets at the network nodes. Scheduling may help achieve differentiation in quality-of-service. Flow control may be used to adapt the transmission rate of the various connections, depending on the degree of network congestion.

The relative importance of the various instruments is a matter of great controversy. One school of thought advocates centralized control, with a primary task assigned to admission control, so as to provide 'hard' quality-of-service guarantees to end users. A second community promotes distributed control, with a pivotal role played by flow control, and reliance on adaptive applications and intelligence with end users, to achieve 'soft' performance degradation. In both views, a key role is reserved for routing and scheduling algorithms, although implementation details may differ.

Performance evaluation

The design of efficient resource allocation algorithms requires techniques for evaluating the relevant quality-of-service measures, e.g., delay performance, cell loss, call blocking. There are two basic approaches for performance evaluation, namely mathematical analysis and (Monte Carlo) simulation, each with its own pro's and con's.

The main attraction in simulation lies in the amount of detail that may be captured in the models, although the complexity of realistic systems may render detailed simulation prohibitively demanding. Besides, some performance measures, such as cell loss, involve simulation of extremely rare events. Thus, excessively long running times may be needed to obtain statistically significant results. This stimulated the development of special techniques to speed up the simulation of rare events.

The strength of analysis is that performance measures may be characterized as an explicit function of the system parameters. A disadvantage is that exact mathematical analysis is usually restricted to relatively simple models. However, the insights obtained may be used to construct approximations or develop algorithms for more complex scenarios.

Queueing theory

Performance evaluation research at CWI has traditionally focused on mathematical analysis, using techniques from queueing theory. Queueing theory is a field of applied mathematics concerned with the study of congestion phenomena in stochastic service systems.

The basic queueing model consists of a server, or a group of servers, where customers arrive who require some kind of service. In the context of communication networks, the server usually represents a transmission link, or a buffer or port on a switch. The customers typically correspond to transmitted packets, bursts, or offered calls, in accordance with the three-level hierarchy described above.

If customers cannot be taken into service immediately, they may join a queue and wait (queueing models), or they may instantly leave the system (loss models). The arrival times and service durations are usually described through stochastic processes, since it is typically uncertain, or intrinsically random, exactly when customers arrive, and precisely for how long they will need to be served. Queueing theory provides techniques for evaluating the performance measures of interest, e.g., the average delay of customers, or the fraction of customers that are being lost.

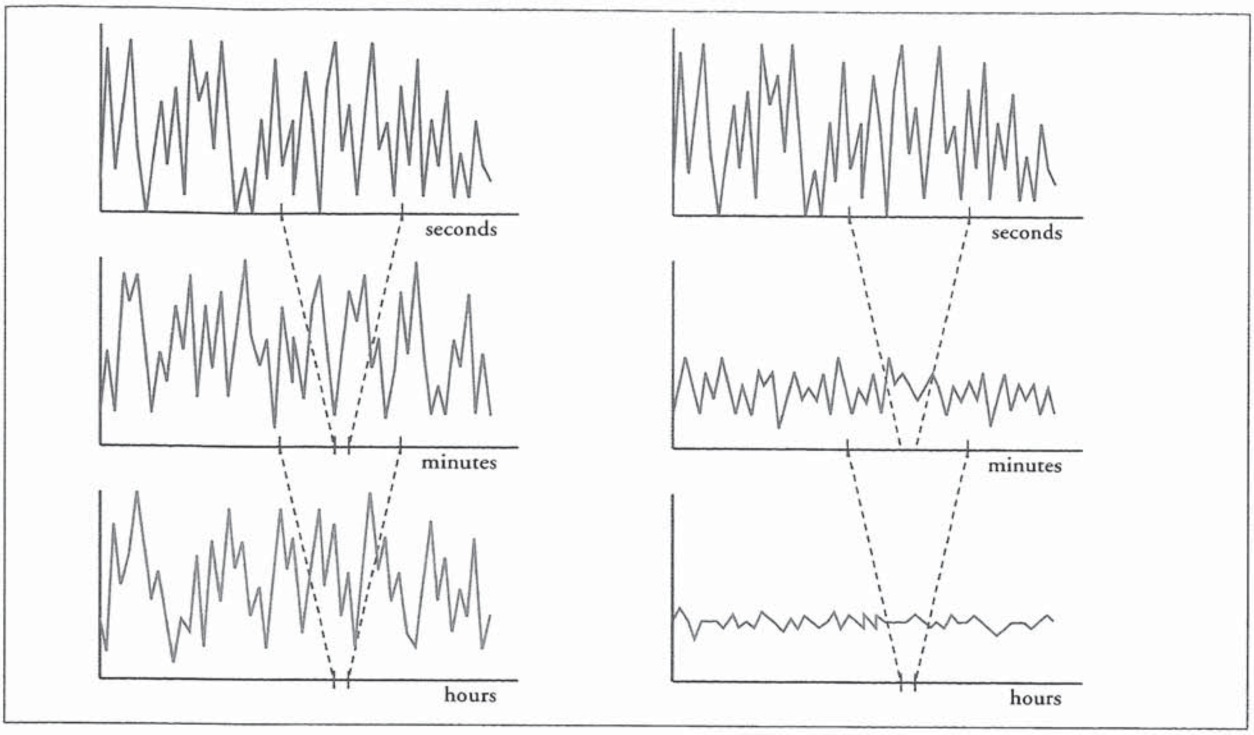
Best-effort services

Important performance issues in integrated networks which have received particular attention at CWI include the impact of long-tailed traffic characteristics and the interaction between *real-time traffic* and *best-effort traffic*. Real-time traffic is traffic which is extremely sensitive to delay, such as voice and video-conferencing. Best-effort traffic is tolerant of some amount of delay, as long as the average capacity received over somewhat longer intervals is sufficient, file transfers being a typical example. Because of the stringent delay requirements, it is common practice to give real-time traffic some form of priority over best-effort traffic in the scheduling of packets. The best-effort traffic thus receives the capacity left-over by the real-time traffic. That remaining capacity may drastically fluctuate over time because of the bursty patterns in packet flows described earlier. The available capacity is then shared among the best-effort connections in certain fixed proportions (*processor-sharing*). This motivated a detailed investigation of (discriminatory) processor-sharing models with time-varying service capacity. The results show that the average transmission time of best-effort files is asymptotically linear in the size of the files, which is commonly viewed as a desirable fairness property. In addition, comparisons were made between various admission control strategies for mixed real-time and best-effort scenarios.

Long-tailed traffic phenomena

A second major research topic concerns the occurrence of long-tailed traffic characteristics. Measurements have provided increasing evidence that packet traffic exhibits long-range dependence and self-similarity over a wide range of time scales. Long-range dependence means that correlations in the traffic volume extend over extremely long time intervals. Self-similarity is a closely related phenomenon, which implies that traffic patterns look quite similar on different time scales. These phenomena are commonly attributed to long-tailed or subexponential characteristics in the underlying traffic processes (connection times, file sizes, scene lengths in videos). This signifies a radical departure from traditional exponential traffic models, where correlations decay quite rapidly, and traffic fluctuations quickly smooth out over longer time intervals.

Queueing models with long-tailed input traffic offer a useful approach to evaluating the impact of self-similarity and long-range dependence on network performance. It was shown for example that if the input traffic is long-tailed, then so are the buffer content and the delay. In fact, the tail behaviour of the buffer content is 'one degree-worse' than that of the input traffic. In integrated networks, the transmission capacity is shared by several traffic streams, where some may have long tails, while others may not. In that case, even a single long-tailed traffic stream may cause the buffer content to be long-tailed, unless the link rate is larger than the peak rate of that long-tailed



One characteristic of long-tailed traffic phenomena is that fluctuations don't smooth out with increasing time scale (left), in contrast to short-tailed phenomena (right). This poses fundamental problems for network control and traffic engineering for long-tailed phenomena.

connection plus the mean rate of the other flows. These results have important implications for admission control.

Recently, the research has focused on the issue how scheduling algorithms may be used to neutralize the negative effects from long-tailed traffic phenomena. It was found that if traffic is not processed in order of arrival, but in processor-sharing fashion, then the tail behaviour of the delay is no longer any worse than that of the input traffic.

A related study considered a transmission link shared by several long-tailed traffic streams in accordance with the weighted fair queueing (WFQ) discipline. The WFQ scheme is a leading candidate among packet scheduling algorithms for integrated networks. The results suggest that WFQ scheduling algorithms provide an effective mechanism for extracting sharing gains, while protecting individual connections.

Interactive Visualization
of Turbulent Fluid Flow

Research Project: Interactive Visualization Environments
Researcher: R. van Liere
E-mail: Robert.van.Liere@cwi.nl
URL: http://www.cwi.nl/~robertl/

Introduction

Fluid flows at low velocities in general exhibit smooth (laminar) patterns. At high velocities, particularly when the viscosity of the fluid is important, the flow becomes irregular (turbulent). Notwithstanding substantial research efforts during this century, turbulence is still not very well understood. The phenomenon plays an important role in several industrial settings: turbulence decreases fuel efficiency, increases noise generation in engines, and presents a nuisance for air duct systems in homes and office buildings, to name a few examples. CWI,s research aims at developing novel, interactive visualization techniques, and integrating these in the processes of modelling, simulation, and analysis. Interaction stimulates exploration and is mandatory if it is not a priori known which phenomena are of interest. The developed visualization techniques are applied to a data set resulting from a direct numerical simulation of turbulent flow carried out at the University of Groningen.

Interactive visualization

Slow-moving fluids such as those often encountered in the processing industry have flows that are fairly simple. Flows become more complex for air or other fluids moving at high speeds, particularly when the physical viscosity of the medium is important. A flow with a low value of the Reynolds number – a constant that defines the ratio of the fluid's inertial resistance to acceleration to the fluid's viscosity – will be smooth and laminar, while a flow with high Reynolds number will be turbulent. To describe turbulent flow, the Navier-Stokes equations must be solved, taking into account the effects of viscosity. High performance computing allows computational fluid dynamics (CFD) researchers to compute solutions to problems with increasingly higher Reynolds numbers. To solve such problems high resolution grids are needed, resulting in very large data sets.

Interactive visualization of turbulent fluid flow is a very challenging problem due to several reasons. The foremost reason is that there is no natural visual representation for a flow field. We can represent and interpret geometric objects, colour and texture, but how should a flow field be presented? In addition, within the context of time dependent turbulent flow, flow phenomena occur at a wide range of spatial and temporal scales. For example, consider Figure 1, a famous sketch by Leonardo da Vinci showing a turbulent flow. This image shows various levels of spatial scale in the flow; larger *vortices* at the outside and small *eddies* in the middle. Flow experts require animated images like these for the analysis of flow phenomena in turbulent flow.

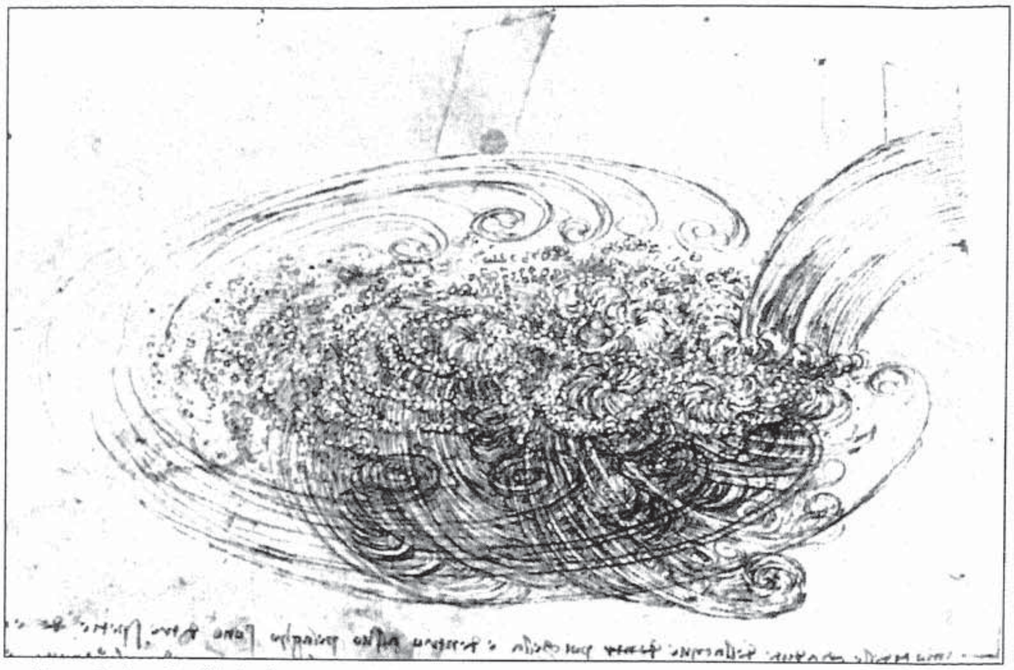


Figure 1: A study of hydrodynamic turbulence, Leonardo da Vinci (1508).

A goal of the work at CWI has been to develop an interactive environment which allows fluid dynamics experts to analyze very large flow data sets. The work has been done in the context of the national High Performance Computing programme and is centered around three fundamental beliefs. First, scientific visualization should be integrated in the process of modelling, simulation and analysis. Ultimately, we envision that visualization will be fully integrated in the process of modelling complex flow phenomena. Second, novel visualization techniques are required for the analysis of large scale data sets. Flow phenomena can now be simulated at such a level of detail that traditional visualization techniques do not suffice. We believe that hierarchical visualization techniques are required that can represent various abstraction levels in the data. Finally, the visualization techniques themselves must be highly interactive. Interaction stimulates exploration and, as such, increases the insight into the data. Moreover, interaction is mandatory in cases in which it is not a priori known which phenomena are of interest.

Our approach

What type of analysis do flow experts want to apply to the computed data? We consider two governing concepts that have driven our research. The first concerns the recognition of flow phenomena. While the data sets themselves contain only values such as velocity and pressure, the flow expert is interested in flow phenomena on a higher level, such as flow separation, re-attachment, vortex formation, etc. Techniques are sought to extract these phenomena from the data. Unfortunately, the difficulty is that these phenomena themselves are sometimes not well understood. Extracting and displaying these phenomena from the underlying data is still a largely unsolved problem. The second concept concerns levels of scale. Phenomena in turbulent flow are characterized by flow patterns of widely varying spatial and temporal scales. In state of the art simulations, pattern sizes may vary by three orders of magnitude. Visualization techniques should be able to cope with these different levels of scale.

The governing philosophy of our environment is that achieving insight into data is an interactive process. The user begins with an analysis of the raw data and iteratively progresses towards higher level concepts. Jointly with other scientific visualization researchers, we have developed two techniques that address this philosophy. The first technique is a visualization method based on *texture*. Texture is 'something composed of closely interwoven elements' (Webster). Texture based flow visualization methods map a vector field to a texture. The primary advantage over other flow visualization techniques is that texture can give a continuous view of a field opposed to visualization at only discrete positions, as with arrow plots or streamlines. The visual effect of direction in a texture is achieved by line structures in the direction of the vector field. These lines are the result of coherence between neighbouring pixels in the texture. Coherency in the texture will be higher in the direction of the vector field than in other directions. The second technique relates to *feature visualization*. Instead of directly visualizing the raw data, feature visualization techniques extract meaningful structures and depict these structures schematically. In this way high level abstract visual representations can be produced. Examples of feature visualization include flow topology analysis, vortex detection and tracking, localization of shock waves, etc.

Some Visualization Techniques

We briefly discuss some technical aspects of two visualization techniques we have researched. They serve as illustrations of techniques which can be used to explore large flows.

Interactive spot noise. Spot noise is a texture synthesis technique which can be used to present a global overview of a vector field. In spot noise, small icons – called spots – deformed according to the underlying data are used to show a flow field. If many spots are used, the individual spots can no longer be discerned and texture is perceived instead. This idea is illustrated in Figure 2. A spot noise texture is characterized by a scalar function f of position $\mathbf{x}$. It is defined as

$$f(\mathbf{x}) = \Sigma\, a_i h(\mathbf{x} - \mathbf{x}_i)$$

in which $h(\mathbf{x})$ is called the spot function. It is a function everywhere zero except for an area that is small compared to the texture size; a_i is a random scaling factor with a zero mean, $\mathbf{x}_i$ is a random position. In non-mathematical terms: spots of random intensity are drawn and blended together on random positions on a plane.

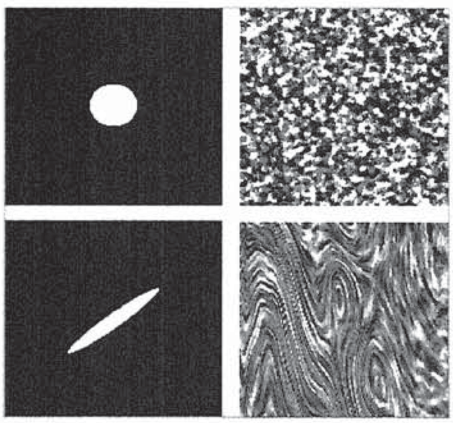


Figure 2: Spot (left) and generated texture (right). Circular spots (top) and deformed spots (bottom).

The downside of spot noise is that it is computationally very expensive. A large number of particle paths and particle positions must be calculated, spots must be transformed, textured and blended. We developed a scalable algorithm which partitions work evenly among processors, and multiple graphics pipes. Interactive speeds can be obtained for very large data sets. High texture generation speeds can be used for interactive adjustment of spot noise parameters to highlight certain aspects of the flow, or to zoom in on details of the flow. In addition, animations of time dependent flow can be generated in real time.

Interactive hierarchical flow topology. Vector field topology is a visualization technique that presents essential information by partitioning the flow field in regions using critical points which are linked by streamlines. Critical points are points with zero flow velocity. Each critical point is classified based on the behaviour of the flow field in the neighborhood of the point. For this classification the eigenvalues of the velocity gradient tensor are used. In our analysis we consider only 2D slices from 3D data sets; this is a simplification since critical points on 2D slices need not necessarily coincide with critical points in the 3D field. The velocity gradient tensor – or Jacobian – is defined as:

$$J = \nabla\vec{u} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}$$

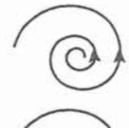
in which u and v are the velocity components in the x- and y-direction, respectively, and subscripts denote partial derivatives. Based on the two complex eigenvalues $(R_1 + iI_1, R_2 + iI_2)$, six different cases corresponding to real flows are distinguished:

Saddle point $R_1 * R_2 < 0$ and $I_1, I_2 = 0$

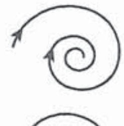
Repelling node $R_1, R_2 > 0$ and $I_1, I_2 = 0$

Attracting node $R_1, R_2 < 0$ and $I_1, I_2 = 0$

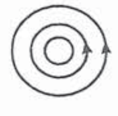
Repelling focus $R_1, R_2 > 0$ and $I_1, I_2 \neq 0$



Attracting focus $R_1, R_2 < 0$ and $I_1, I_2 \neq 0$



Center point $R_1, R_2 = 0$ and $I_1, I_2 \neq 0$



Critical points can be connected by tracing streamlines in the direction of the eigenvectors of the velocity gradient tensor. These lines will divide the flow field into distinct regions.

Phenomena in turbulent flow fields are characterized by flow patterns of widely varying spatial scales. In terms of topological information, this means that a large set of critical points results from flow patterns at small spatial scales. However, the global structure of the flow can be described by a limited subset of all critical points. The governing idea of the multi-level flow topology method is that the displayed number of critical points should be limited to characterize only those flow patterns of a certain level of scale, while the other critical points are omitted. For this purpose, a *pair distance filter* has been designed. The motivation of this filter is that an often occurring small disturbance of the flow is caused by pairs of critical points. For example, the topological structure of a two-dimensional vortex consists of a focus (repelling or attracting) or a center combined with a saddle point (see Figure 3). The size of the vortex is determined by the distance between the pair of critical points. Removing the pair from the topology does not influence the global structure of the flow.

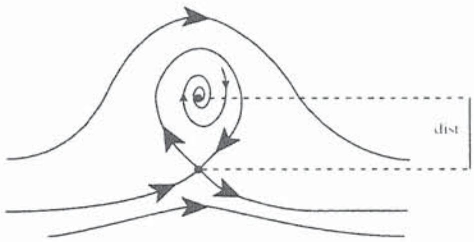


Figure 3. An example of a critical point pair: a focus and saddle point forming a vortex. The distance defines the spatial scale of the point pair.

Analysis of a turbulent flow

Researchers at Rijksuniversiteit Groningen have studied methods for direct numerical simulation (DNS) of turbulent flow. DNS is an accurate technique for computing turbulent flow. Flow experts use the resulting visualizations to test hypotheses about flow phenomena and – after a detailed inspection of the animation – as a means to pose new hypotheses. Of particular interest is the detailed visualization of vortex formation and the transition from laminar to turbulent flow.

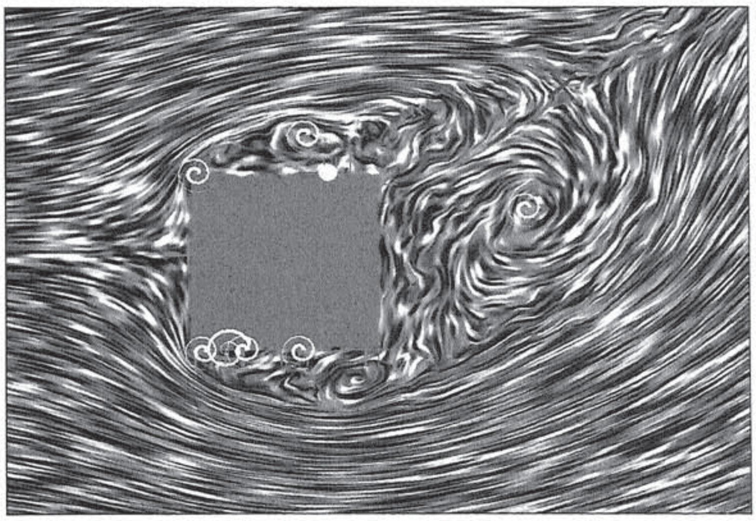


Figure 4. A view of global flow structure around a square cylinder. Flow from left to right.

In this particular problem, a DNS of a turbulent flow around a square cylinder at Re = 22,000 (at zero angle of attack) has been performed. The computation and the size of the resulting database is impressive: the computation took three weeks on a 12 CPU Cray C90. The resolution of the rectilinear grid was 316×540×64 and the total number of time steps computed was 100,000. It was impossible to store all data on disk; thus a selection of the data was made. 7500 time steps of a XY-slice were taken, resulting in about 50 Gigabytes of data.

Figure 4 shows a 2D slice of the flow. The pair distance filter was used to remove many critical points from the image. Small coloured icons are used to display the set of critical points: a yellow spiral icon denotes a focus, a blue cross denotes a saddle point, and cyan/magenta disks denote repelling/attracting nodes. Streamlines can now be drawn without excessive cluttering of the image while, simultaneously, maintaining the global structure of the flow. Note, for example, the large vortex (consisting of a saddle and attracting node) behind the square cylinder.

Figure 5 shows two zoomed-in views of the previous image. The distance threshold for the pair distance filter is adjusted to reflect structures at a smaller scale. The image on the left is a section of the upper right corner of the cylinder. The image on the right is zoomed in even further. Note that the images provide additional detail when zooming in on a section of the data. Providing this detail in the top level image would result in cluttering.

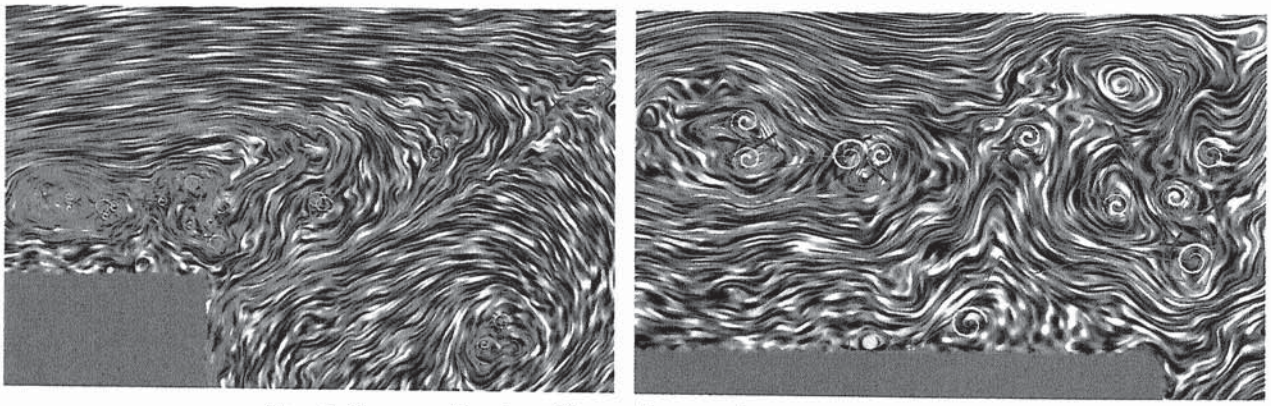


Figure 5. Two zoomed-in views of flow topology around a square cylinder.

In the near future DNS can be applied to flows with a Reynolds number in the order of 10^5. Insights from such fundamental simulations can lead to better practical designs to improve fuel efficiency, lower noise generation in engines, and reduce frictional drag in air duct systems in homes and office buildings. Interactive visualization will be vital to gain these insights.

Finance
changes face

Research project: Mathematics of Finance
Project leader: J.M. Schumacher
E-mail: Hans.Schumacher@cwi.nl
URL: http://dbs.cwi.nl/cwwwi/owa/cwwwi.print_themes?ID=15

Introduction

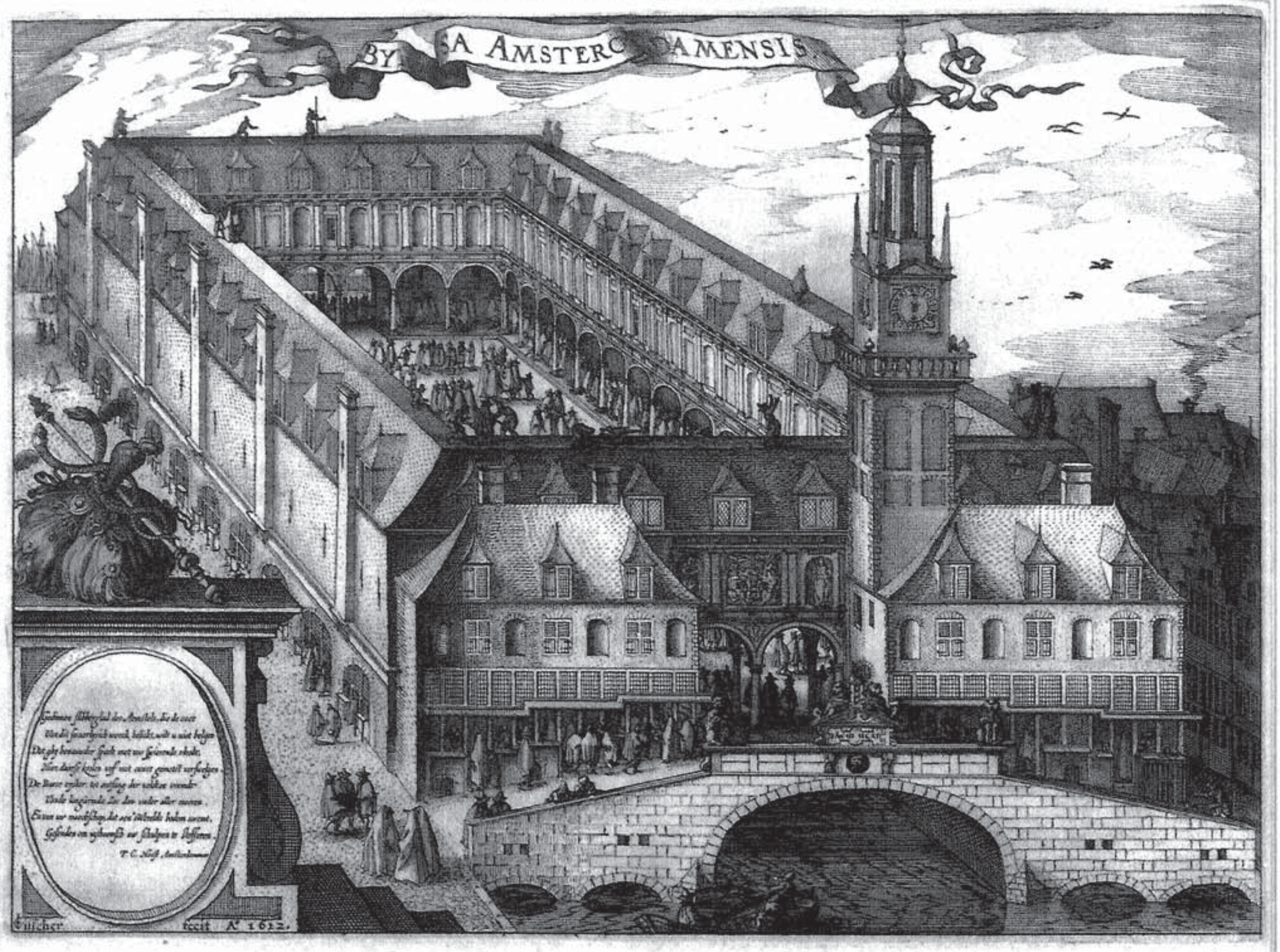
Over the past quarter of a century, a revolution has taken place in the world of finance. It is said in the US that once there was a time in which one could run a bank by the two-three-four rule: pay interest at two percent, lend at four, and arrive at the golf course by three in the afternoon. Those days are definitely gone. Instead, finance has become a quantitative discipline in which mathematical models play a leading role. The key factor that is responsible for this major turnover has been the introduction of the new financial instruments known as derivatives.

To the general public, derivatives are mainly known as culprits in a number of highly publicized affairs: the reputed British bank that could not cover the losses of just one of its traders (Barings), the Nobel prize winners that jeopardized the American economy by losing a bet on interest rates (Merton and Scholes, by their involvement in Long Term Capital Management). Eye-catching as these events may be, they are actually just a small part of a much bigger story. The total value covered by derivative contracts was estimated in 1998 by the Bank for International Settlements in Basel at the phenomenal amount of 84 trillion dollar (84.10^{12}). That is more than three times the value of the world's yearly production of goods and services. In the vast majority of cases, these derivative contracts serve to reduce risk rather than to speculate.

Derivatives

An example of a derivative is a *forward contract* in which one party agrees to deliver a certain good to the other party at a certain time in the future for a price that is already agreed upon in the contract. In situations where prices are changeable, such a contract may reduce risk for both parties. For instance, when a farmer agrees with a miller on the price for which he will deliver wheat next year, then the farmer is protected against a fall of the price of wheat, and the miller is protected against a rise of the price. Such contracts have existed since ancient times. The trading of forward contracts can be made easier by the introduction of a 'clearing house' that effectively finds counterparties for anyone who would like to make a forward deal. A forward contract that is established through a clearing house is called a *futures contract*. For instance, in the middle of the nineteenth century a futures market for wheat developed in Chicago, in those days the agricultural center of the United States.

A futures contract limits downside risk but also constrains upside opportunities. It may be attractive to have a contract that provides protection against losses but leaves possibilities for gains intact. Such a contract is an *option*: it gives the right, but not the



The Amsterdam Exchange (building by Hendrick de Keyser) in the 17th century. Copyright: Capital Photos.

obligation, to buy a given asset (i.e., anything that is traded) for a given price at a certain time in the future. Options were traded at the *Beurs* in Amsterdam already in the beginning of the seventeenth century.

An increased interest in protection against fluctuations of financial indicators such as exchange rates and interest rates came up in the early 1970's following the collapse of the Bretton Woods fixed exchange rate system and the oil crisis. Foreign exchange rate futures were introduced in Chicago in 1972 and other financial derivatives, including options, followed. In the years after, futures and option exchanges were established in all major financial centers. As users found new ways to employ the flexibility of derivatives and providers responded by the creation of new instruments such as interest rate swaps, the derivatives markets showed a spectacular growth.

The role of mathematics

If it would be unknown how the price of a derivative relates to the fluctuations of the underlying variables (which can for instance be prices of stocks, exchange rates, or interest rates), then banks would be reluctant to market new derivative products since it would be difficult to assess the risks involved with the trading of such products. A high price would be required to cover the unknown risks, and derivatives markets might not even get off the ground. The key to the growth of the derivatives markets is that there is a theory available that predicts how the price of a derivative will vary with the underlying variables.

The early developments in the theory of derivatives are due to Fischer Black, Robert C. Merton and Myron Scholes. In 1973, the *Journal of Political Economy* printed a paper by Black and Scholes on the pricing of options. The paper presents a formula that gives a relation between the price of an option and the price of the underlying asset. The expression, which is now of course known as the Black-Scholes formula, not only provides a *price* for derivatives. Just as importantly, it also contains information about how to *hedge* against the risk of writing (i.e., selling) an option. By trading the underlying asset during the lifetime of an option, the institution that has sold the option is able to reduce (theoretically even eliminate) the risk involved in the option contract. The trading policy which achieves this is defined from the Black-Scholes formula.

Black and Scholes used a mathematical argument in their paper which was based on a stochastic model for the evolution of the price of the underlying asset. The model they used is known as 'geometric Brownian motion'. Noting that risk could be eliminated in a properly constructed portfolio of options and underlying assets, and arguing that the return on such a riskless portfolio must be the standard interest rate for riskless investments, they wrote down an equation that should be satisfied by the price of the derivative as a function of the price of the underlying asset and the remaining lifetime of the option. The equation obtained in this way is a partial differential equation of a standard type for which an analytic solution is known. From that, Black and Scholes were able to write down their celebrated formula.

Later developments

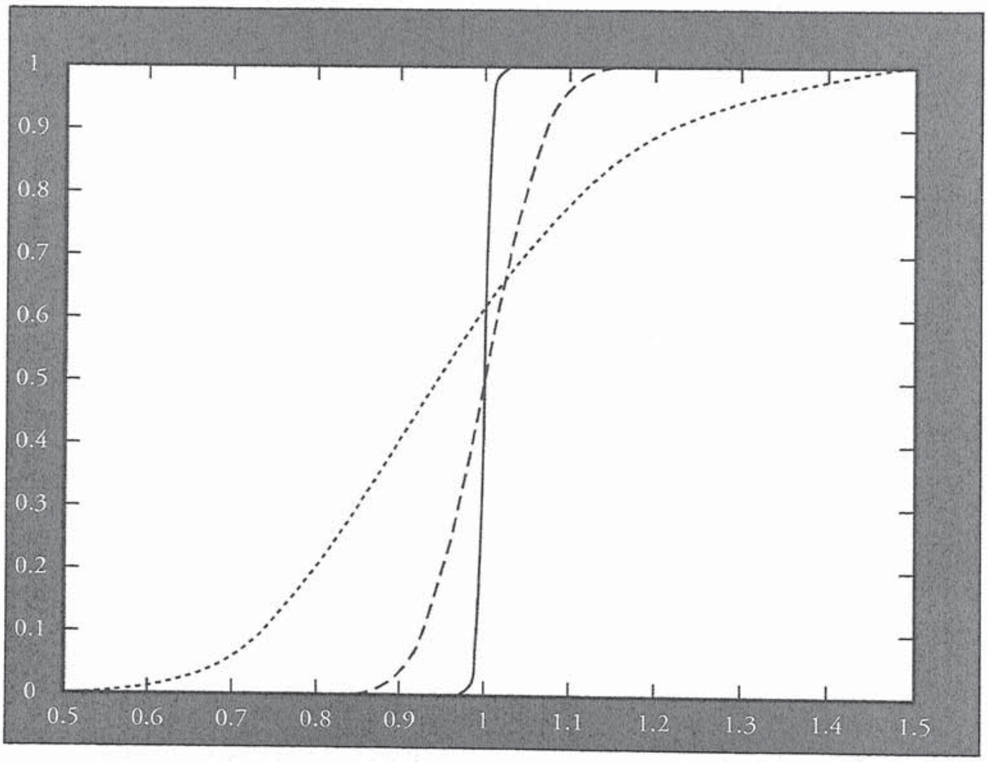
From a strictly mathematical point of view, there was nothing new in the work of Black and Scholes. After all, the relation between Brownian motion and partial differential equations was well known, and so was the analytic solution of the PDE in question (a diffusion equation). The reasoning applied by Black and Scholes, based on the construction of theoretically riskless portfolios, however opened up a connection between derivative pricing on the one hand and, on the other hand, stochastic models and partial differential equations. This connection can be used in many more situations than just in the relatively simple case studied in the original Black-Scholes paper.

The option contract for which an analytic price formula was derived by Black and Scholes is known as a European option written on a single underlying asset. The term 'European' does not have a geographical meaning here; it just refers to the fact that the option gives the holder the right to buy a specified amount of the underlying asset for a specified price *at* a specified time. If 'at' is replaced by 'at or before', then one speaks of an *American* option. Already in the case of simple American options, there is no analytic solution available and one has to resort to numerical methods.

American options occur quite frequently. So do options which depend on several underlying variables, and derivatives whose value is determined by the values of the underlying asset at many different points in time. Moreover, there are good reasons to replace the geometric Brownian motion that Black and Scholes used by other stochastic models. Further complications arise when one considers aspects like transaction

Management of market risk

Financial institutions manage large portfolios of many different types of contracts. According to regulations, these portfolios should be such that 'the book is balanced', which means that the total value of the contracts should not be too sensitive to unpredictable market forces. As long as the contracts that are traded by the institution are built up from simple standard elements it is not too difficult to tell when the book is balanced. Things change however when more complicated contracts are involved, such as derivative securities which depend in a nonlinear way on the value of some underlying asset. A typical example of a derivative instrument is the European call option, which gives the holder the right to buy a certain asset (for instance a stock, or a given amount of foreign currency) at a specified time (the *time of maturity*) at a specified price (the *strike price*). Of course the value of the call option is sensitive to the price of the underlying asset, as well as to the time that remains until the option matures. One of the main purposes of modern mathematical finance is to quantify these dependencies. The graph below shows sensitivity curves for a call option which have been computed from the Black-Scholes formula. The horizontal axis gives the current asset price divided by the strike price; the vertical axis gives the fraction of the underlying asset that has the same sensitivity to price changes as the call option. The dotted, dashed, and solid curves correspond to options that will mature in one year, one month, and one day respectively. For instance: a bank holds one unit of a certain stock and has sold four call options on the same stock (0.25 on the vertical axis). The options will mature in one year and have a strike price that is 20% higher than the current asset price (about 0.83 on the horizontal axis). Then it can be read off that the position of the bank is approximately neutral. In the plot the following parameter values have been used: interest rate 4%; volatility of the underlying asset 20%. The steepness of curves is directly related to the nervousness of traders.



Equivalent positions for a call option.

costs and limited trading frequency which were abstracted away by Black and Scholes. These questions give rise to new developments in applied mathematics, which are taken together under the title Mathematics of Finance.

Financial mathematics revolves to a large extent around the computation of functionals of stochastic processes. Due to the relation between stochastic processes and partial differential equations, it is also possible to use a formulation in terms of PDEs; another reformulation is in terms of high-dimensional integration problems.

Bundles of mortgages

To illustrate the computational problem in a concrete situation, consider a portfolio of mortgage contracts. Such portfolios can be looked at as investment opportunities, especially when constructed to meet the tastes of investors; for instance, Dutch pension funds are currently looking into the possibility of investing in mortgage portfolios to complement their usual investments in government bonds, real estate, and stocks. The cash flows resulting from a portfolio of mortgages depend on the future development of interest rates, and so to put a price on such a portfolio one has to model the behaviour of interest rates, as well as some other things like the extent to which home owners will use early redemption (which also depends on interest rates).

Future cash flows can be added up to one number by a discounting procedure, but then one still has the problem that different future interest rate paths will give rise to different outcomes. Somehow one has to average over these possibilities to arrive at a reasonable price. Fortunately, the theory of pricing of derivatives that has been developed since the seminal work of Black and Scholes does provide such an averaging rule. The rule assigns probabilities to interest rate paths in such a way that averaging according to these probabilities leads to the price that is consistent with the assumption that a riskless portfolio must earn the standard interest rate for riskless investments.

After this modelling stage, we basically have a stochastic process (roughly speaking, a collection of paths each carrying a certain probability), together with a functional, that is, a rule that assigns to each path a certain number. The theory tells us at this point that the price can be computed as the expected value of the functional, that is, the average value with weights according to the probabilities of the paths.

Integrating in 360 dimensions

To see what this means computationally, let us assume for instance that we are talking about a portfolio of mortgages over periods up to 30 years, and that we take monthly intervals to capture the movement of interest rates. Let us also assume that there is only one interest rate that plays a role; actually this is a severe simplification since in reality one has short rates, long rates and so on, and these are only loosely correlated. In this setting, an 'interest rate path' is just a sequence of 360 numbers representing a possible monthly evolution of interest rates over a period of thirty years; or, one might say, a path is a point in a 360-dimensional space. To each such point there is a number associated which represents the discounted cash flow from the portfolio under study if the interest rate follows the path represented by the data point. The computational task at hand is the determination of the average of the numbers corresponding to all data points when these data points are weighted with their respective probabilities; in more precise terms, we have to compute an integral of a function of 360 variables.

Numerical integration in low dimensions is usually carried out by averaging function values on points taken from a regular grid. In n dimensions, a grid with k points along each dimension produces k^n points. A quick calculation makes clear that the grid approach is totally infeasible in high-dimensional integration problems; even with only two points along each dimension, a 360-dimensional problem would require 2^{360} (more than 10^{108}) function evaluations, enough to keep even the fastest computer busy until the end of times. However, we can also look at the problem from a probabilistic point of view and note that we are actually trying to compute the expected value of some stochastic variable. Surely we don't need a sample size of 10^{108} to get a reasonable estimate of the population mean.

This idea is the basis of the so-called *Monte Carlo* method of doing high-dimensional integration. A random number generator is used to produce a collection of (hopefully) representative points, and the sample mean of these points is taken as an approximation of the true expected value. It can be shown that, under weak conditions, the value obtained in this way does indeed converge to the true value as the number of sample points is increased, be it only slowly; a 100-fold increase of the number of sample points brings only one extra decimal place of accuracy.

Quasi Monte Carlo

Rather than leaving the representativeness of sample points to chance, one may try to create 'evenly distributed' points in high-dimensional spaces by an algorithm which is specifically designed for this goal. This is the basic idea of what is called the *Quasi Monte Carlo* method; a misnomer actually, since randomness plays no role in this method. Not surprisingly, QMC builds on the theory of uniform distribution of numbers. (The idea is to carry out a deterministic sampling of data based on a sequence of numbers for which one can prove that, asymptotically, they are uniformly distributed on a given interval. In Monte-Carlo methods one works with random sampling of data for which uniform distribution cannot be proved in general.) This branch of number theory has a tradition that goes back well beyond the Black-Scholes days; in the Netherlands for instance contributions have been made by J.G. van der Corput (1890–1975) and J.F. Koksma (1904–1964).


J.G. van der Corput (left) and J.F. Koksma (right), both directors of CWI in the early years of its existence, made important contributions to the theory of uniform distribution of numbers, which now finds application in financial mathematics.

Among the best known modern contributors to the development of the Quasi Monte Carlo method are H. Niederreiter (Austrian Academy of Sciences, Vienna) and S. Tezuka (IBM, Tokyo). The first applications of QMC in finance were carried out in 1994 by S.H. Paskov in Ph.D. work at Columbia University in New York, supervised by J.F. Traub.

Jiri Hoogland and Dimitri Neumann, two postdocs working in the CWI research theme Mathematics of Finance, have developed a new and fast QMC generator specifically for applications in finance. In financial problems, there is often already a sizeable computational load from a single function evaluation, corresponding for instance to the computation of the discounted sum of future incomes generated by a given scenario. For this reason, asymptotic results (i.e., for a large number of evaluation points) on QMC generators are not always relevant. A better understanding of the relation between design parameters of QMC generators and their quality will probably help in improving QMC results for limited numbers of generated points. Other topics of current research in the CWI Finance Group include the combination of QMC methods with other techniques (such as 'true' Monte Carlo, or PDE methods), and, in another direction, estimation methods for stochastic processes.

ORGANIZATION

Research

Cluster - Theme	Cluster leader Theme leader
Probability, Networks and Algorithms	**A. Schrijver**
- Networks and Logic-Optimization & Programming	A.H.M. Gerards
- Traffic and Communication-Performance & Control	J.H. van Schuppen
- Stochastics	M.S. Keane
- Signals and Images	H.J.A.M. Heijmans
Software Engineering	**J.W. de Bakker**
- Interactive Software Development and Renovation	P. Klint
- Specification and Analysis of Embedded Systems	J.F. Groote
- Coordination Languages	J.J.M.M. Rutten
- Evolutionary Computation and Applied Algorithmics	J.A. La Poutré
Modelling, Analysis and Simulation	**C.J. van Duijn**
- Environmental Modelling and Porous Media Research	J.G. Verwer
- Industrial Processes	P.W. Hemker
- Mathematics of Finance	J.M. Schumacher
Information Systems	**M.L. Kersten**
- Data Mining and Knowledge Discovery	A.P.J.M. Siebes
- Multimedia and Human-Computer Interaction	H.L. Hardman
- Interactive Information Engineering	P.J.W. ten Hagen
- Quantum Computing and Advanced Systems Research	P.M.B. Vitányi

Management

Management Team

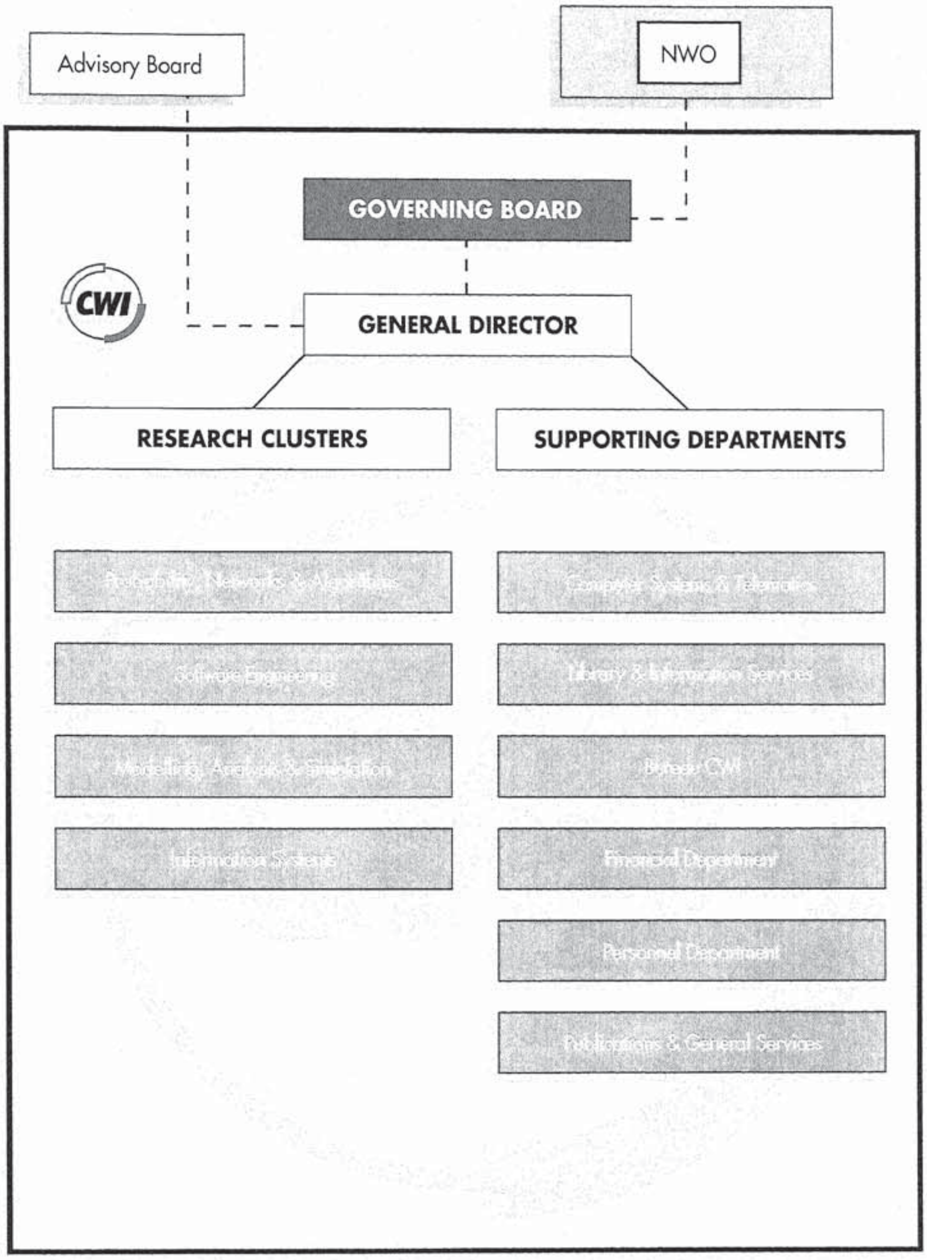
G. van Oortmerssen (director)
J.W. de Bakker, C.J. van Duijn, M.L. Kersten, A. Schrijver (cluster leaders)
P. van der Veen (controller)

Governing Board

L.A.A.M. Coolen (director KPN Research), chairman
P.M.G. Apers (University of Twente)
J.H.A. de Smit (University of Twente)
K.M. van Hee (Eindhoven University of Technology, director Bakkenist Management Consultants)
H.A. van der Vorst (University of Utrecht)

Advisory Board

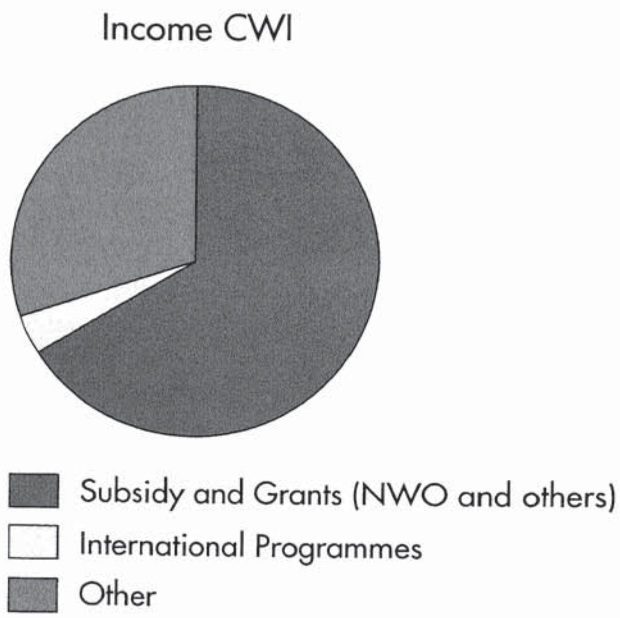
J. van Leeuwen (University of Utrecht), chairman
L.A. Peletier (University of Leiden)
J. Ridder (TNO-NITG)
G. Rodenhuis (Delft Hydraulics)
M.F.H. Schuurmans (Philips Research Laboratory, managing director)
M. Westermann
G. Wiederhold (Stanford University, USA)
B. Larrouturou (INRIA, France)
J. Gunawardena (Hewlett Packard Laboratories, Bristol, UK)

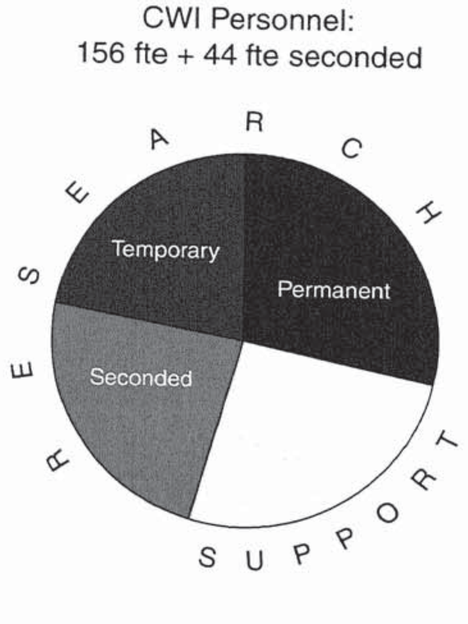

Advisory Board
NWO
GOVERNING BOARD
CWI
GENERAL DIRECTOR
RESEARCH CLUSTERS
SUPPORTING DEPARTMENTS
Probability, Networks & Algorithms
Software Engineering
Modelling, Analysis & Simulation
Information Systems
Computer Systems & Telematics
Library & Information Services
Bureau CWI
Financial Department
Personnel Department
Publications & General Services

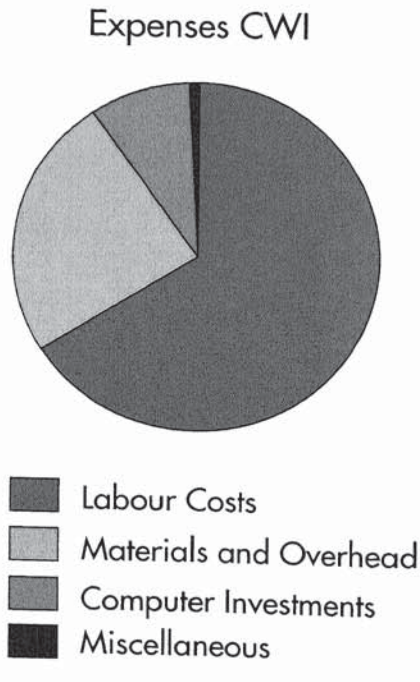
FINANCES, PERSONNEL

Finances 1998

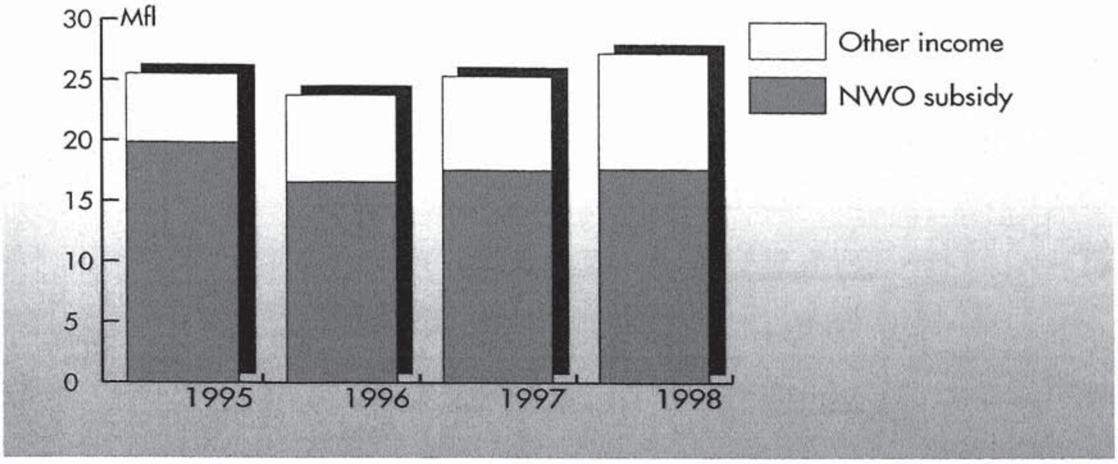
In 1998, SMC spent Dfl. 27,31 million. The expenses were covered by a subsidy from NWO (Dfl. 17,71 million), other subsidies and grants (Dfl. 0,28 million), and from the international programmes (mainly EC programmes, e.g., ESPRIT and HCM) (Dfl. 0,97 million). Finally, an amount of Dfl. 8,37 million (of which 1,26 million from the Telematics Institute) was obtained as revenues out of third-party-services and other sources.



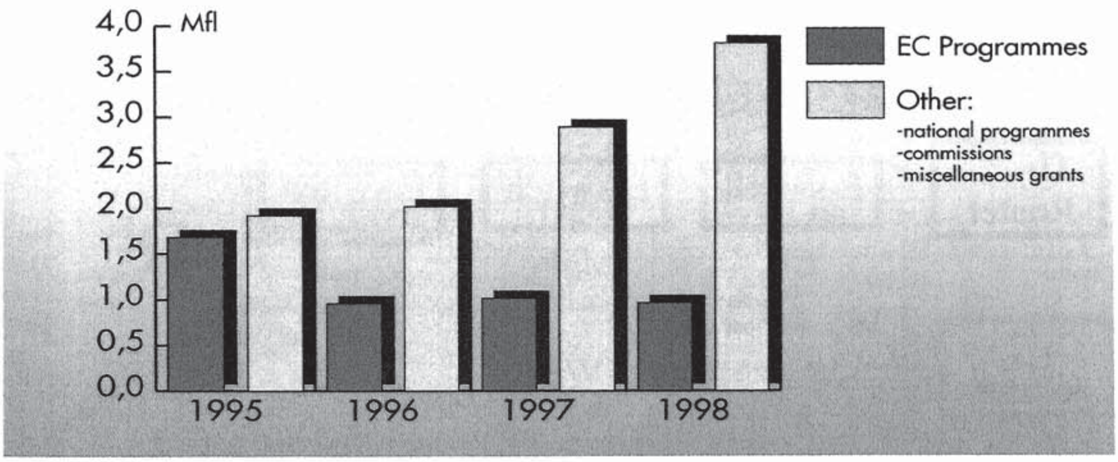




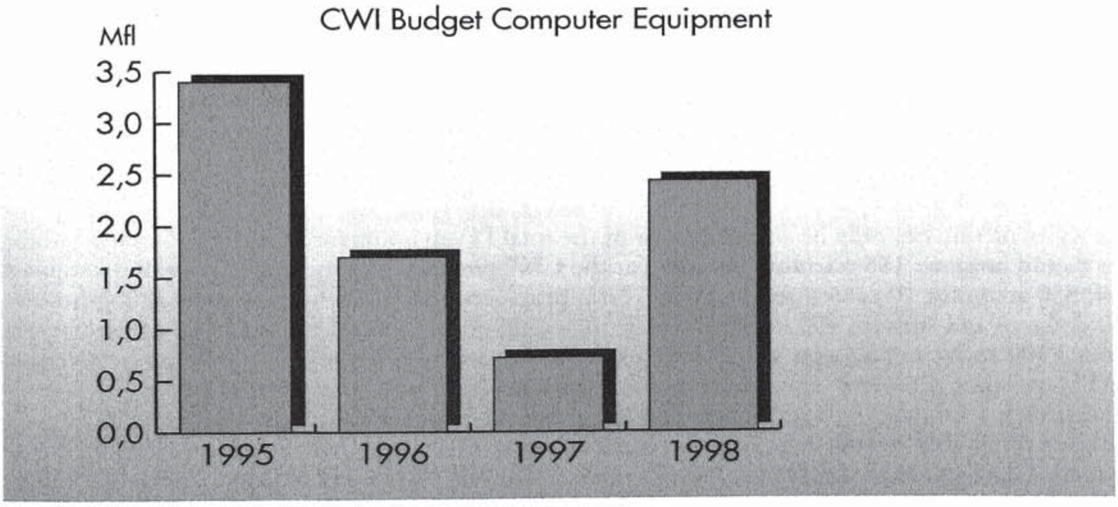
CWI Income



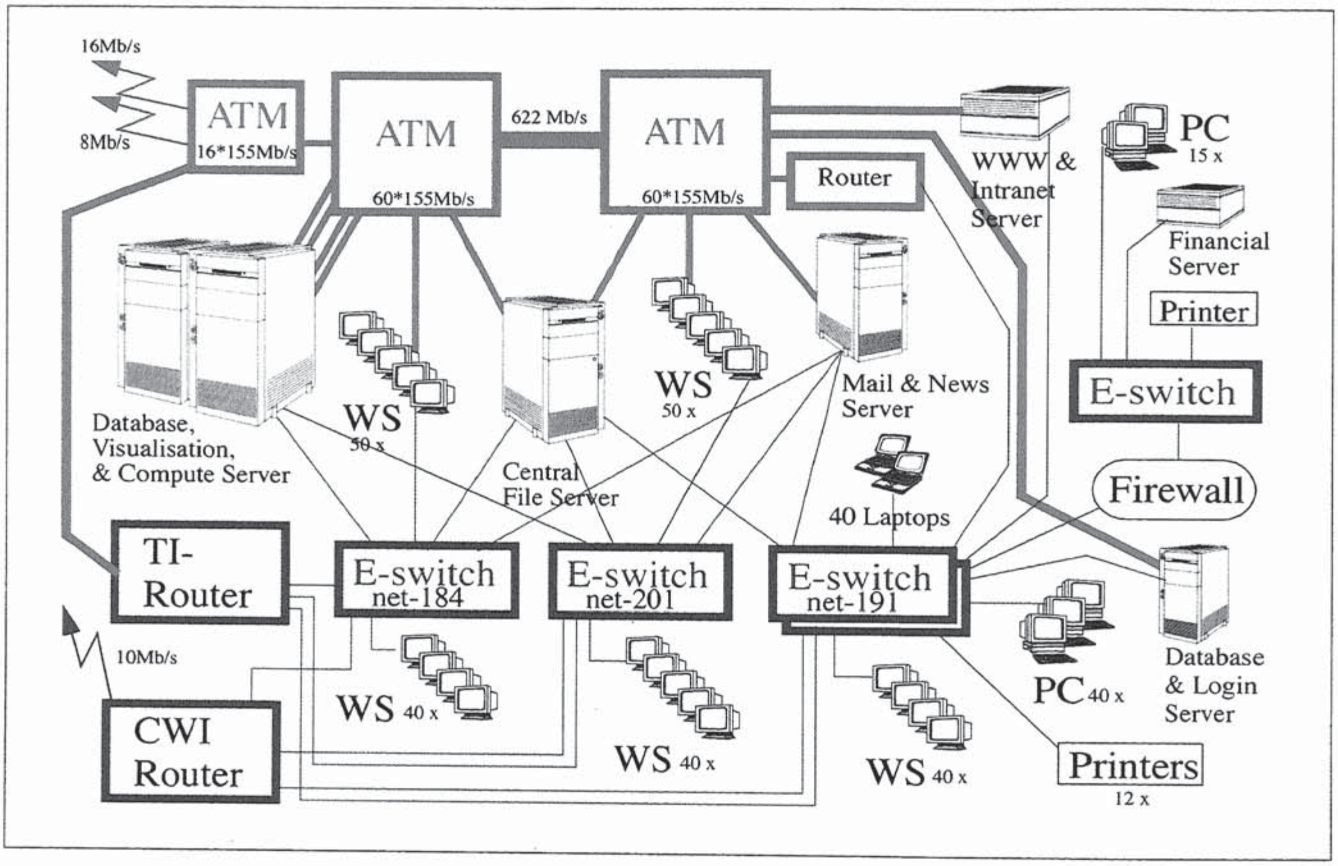
Breakdown of non-NWO Income



CWI Budget Computer Equipment



Computing Equipment Resources



The colour red designates the ATM infrastructure, blue designates the mixed 10/100 Mbit/sec Ethernet infrastructure. The latter is partly switched UTP (most of which is 10Mbit/sec only) and partly shared coax using concentrators (not depicted). The blue arrow designates a 10Mbit/sec Ethernet connection to our two Internet Service Providers (SURFnet and UUnet), via the WCW-lan campus infrastructure. The two red arrows designate a 8Mbit/sec ATM Virtual Private Network connection with the Telematics Institute and a 16Mbit/sec ATM VPN for the MESH video conferencing project, respectively.

The above picture can only be an abstraction of the total IT-environment in use. In rounded numbers, one should imagine: 185 scientists (localised at the CWI premises, not including guests); 45 supporting staff; 350 accounts; 10 central servers (DNS, NIS, print-server, connect-host, security & log-host, etc.); 250 desktops and laptops; 400 network nodes (including printers, hubs, terminal-server, switches, etc.); about 1300 software packages available to each (Unix) desktop; 3 operating systems in 9 versions; 3 ATM-switches; 3 routers (one owned by the Telematics Institute); 6 Ethernet switches (24-port 10Mbit/sec), 1 Ethernet switch (24-port 100Mbit/sec); 24 Ethernet concentrators; over 250 km coax and glass-fibre cable; supporting (industry) standards like: TCP/IP, SNMP, SMTP, SLIP & PPP, PostScript, LaT$_{\Sigma}$X, PDF, AppleTalk, Novell, NFS, SMB, NIS, DNS, HTML, etc., etc.

CWI Ph.D. THESES

Author
Title
Thesis advisor(s) (for external advisors the university's name is added)

E.J. Spee
Numerical Methods in Global Transport-Chemistry Models
P.J. van der Houwen

R.J. Schotting
Mathematical Aspects of Salt Transports in Porous Media
C.J. van Duijn

H.L. Hardman
Modelling and Authoring Hypermedia Documents
A.W.M. Smeulders (University of Amsterdam)

W.J.H. Stortelder
Parameter Estimation in Nonlinear Dynamical Systems
P.W. Hemker

J.F.P. van den Akker
DEGAS: an Active, Temporal Database of Autonomous Objects
M.L. Kersten

H.J. Elbers
Connecting Informal and Formal Mathematics
A.M. Cohen (Eindhoven University of Technology), J.W. Klop

J.D. Mulder
Computational Steering with Parametrized Geometric Objects
F.C.A. Groen (University of Amsterdam)

P.D. Grünwald
The Minimum Description Length Principle and Reasoning under Uncertainty
P.M.B. Vitányi

CWI RESEARCH PROGRAMMES

Probability, Networks & Algorithms
Cluster leader: A. Schrijver

Networks and Logic – Optimization and Programming
Theme leader: A.H.M. Gerards

Networks and optimization

Design, analysis and implementation of optimization and approximation algorithms for combinatorial problems with the help of methods from graph theory, topology, discrete mathematics, geometry, and integer and linear programming, with special attention to network problems (flows, routing and VLSI-design) and scheduling and timetabling.

Constraint and integer programming

Study of the foundations and applications of constraint programming, in particular the design and implementation of an adequate programming environment for constraint programming, and the use of constraint programming for various optimization problems drawing on integer programming techniques.

Traffic and Communication – Performance and Control
Theme leader: J.H. van Schuppen

Communication and computer networks

Development of queueing theoretic models, theory, and algorithms for studying congestion phenomena in communication networks, and of concepts and theory for network control by discrete event and stochastic methods.

Traffic networks

Performance analysis and control of traffic in urban, motorway, railway, air traffic, and other networks, in particular performance aspects of congestion, reliability and availability, and the development of control theory for discrete event and hybrid systems.

Control and system theory

Motivated by control and signal processing, the aim is to formulate classes of systems, focusing on Gaussian systems, finite stochastic systems, positive linear systems, and hybrid systems, and to develop theory for realization and system identification.

Stochastics
Theme leader: M.S. Keane

Probability

Research in probability theory and its applications, with emphasis on combinatorial probability and large random systems, in particular reinforced random walk in the plane and on finite graphs, and continuum percolation.

Statistics

Fundamental and applied research in statistics, in particular statistical inference, statistical estimation of Poisson intensity functions, and statistical methods for compound sums, with applications in finance.

Stochastic analysis

Fundamental and applied research, in particular statistical inference for stochastic processes and fractal analysis of financial data.

Ergodic theory and dynamical systems

Fundamental and applied research, in particular rigid sets, Hausdorff dimension of iterated function systems, classification of Bernoulli schemes, fractal analysis, and superexponential convergence.

Signals and Images
Theme leader: H.J.A.M. Heijmans

Coding, indexing and retrieval

Image coding, feature extraction, and image indexing and retrieval, in particular fractal image coding, statistical analysis of images by way of quadtrees, and viewer centered virtual environments.

Wavelets

Analysis of seismic data, in particular time-varying spectral estimation, and preprocessing with the wavelet X-ray transform.

Morphological image processing

Fundamental research on mathematical morphology and its applications to image and signal processing, with special attention to multiresolution approaches, in particular the construction of morphological pyramids and wavelets.

Stochastic geometry

Parameter estimation for random sets, spatial statistics, and image segmentation with applications to agriculture.

Software Engineering

Cluster leader: J.W. de Bakker

Interactive Software Development and Renovation

Theme leader: P. Klint

Software renovation

Development of new technology for the renovation and maintenance of legacy systems.

Optimization of scientific software

Design and implementation of advanced optimization tools for scientific as well as other software written in existing languages like C and C++, in particular the use of strategy scripts as an aid in generating program transformers, and algebraic techniques for language processing.

Interactive visualization environments

The study of advanced user interaction techniques (virtual reality), with applications in cell biology, and of distributed virtual environments for multi-actor scenarios, augmented with tactile feedback and audio/video.

ASF+SDF

Redesign, reimplementation, and improvement of the ASF+SDF Meta-Environment, in particular the development of a flexible and extensible generic environment to be used in language prototyping and software renovation. Specific aims are: compilation of ASF+SDF to C, unparsing, parser generation, and global architecture.

Domain-specific languages

The aim is to develop methods for selecting suitable DSL domains, and for capturing domain knowledge into a DSL and its compiler, to develop metatools for the rapid prototyping of DSLs, and to study the practical use of DSLs in various settings.

Specification and Analysis of Embedded Systems

Theme leader: J.F. Groote

Process specification and analysis

The study of specification and analysis techniques for process behaviour. Other aims are the study and implementation of (new) algorithms for the analysis and verification of processes with the μCRL toolset, increased insight in the expressivity and potential of timed and untimed μCRL through more case studies (communication protocols, embedded systems, hybrid systems, etc.), and the study of fundamental theory for μCRL language extensions, process semantics, and testing.

Proof searching and proof checking

The study of proof search in simple logical systems, viz., propositional logic, in order to increase the efficiency of symbolic verification techniques applied to verify requirements on processes. Furthermore, the development of proof checking methods to establish the correctness of programmed systems 'beyond any reasonable doubt'.

Coordination Languages
Theme leader: J.J.M.M. Rutten

Formal methods for coordination languages

Development, on the basis of transparant semantic models, of formal methods for coordination languages, and of tools for debugging and visualization for coordination languages, using the recently constructed operational model for Manifold.

Experimental testbed for control-oriented coordination

Design and implementation of an experimental testbed for practical control-oriented coordination programming on heterogeneous platforms and its programming support environment. This testbed is currently accessible through the coordination language Manifold, its utilities, and its visual programming interface, Visifold.

Coordination applications

Study and development of practically useful coordination patterns and protocols in various real-life applications, leading to program modules built on top of the above experimental coordination testbed system.

Exploratory research: coalgebraic models of computation

Further development of coalgebra as a unifying mathematical framework for (transition and dynamical) systems, with emphasis on the coordination of coalgebras, as well as probabilistic transition systems.

Evolutionary Computation and Applied Algorithmics
Theme leader: J.A. La Poutré

Evolutionary algorithms

Design of evolutionary algorithms for management-related and economic problems like optimization, dynamization, automatic programming, and information filtering.

Neural networks

Classification of data, in particular unsupervised image classification for remote sensing data, focusing on blending and mixing of substances over the earth surface, and classification, scaling, event prediction and decision support by several types of neural networks.

Discrete algorithms

Design of efficient algorithms for on-line optimization problems underlying various management and design problems in computer systems and networks, as well as the use of quality of service in managing and optimizing on-line scheduling for multimedia processes.

Modelling, Analysis and Simulation

Cluster leader: C.J. van Duijn

Environmental Modelling and Porous Media Research

Theme leader: J.G. Verwer

Numerical algorithms for air quality modelling

Numerical modelling of the long range transport and chemical exchange of atmospheric air pollutants, involving advection schemes, stiff chemistry ODE solvers, operator splitting, grid generation/adaptation and time stepping methods for PDEs, including implementations of algorithms on supercomputers and parallel computers.

Numerical algorithms for surface water quality modelling

Design of parallel numerical methods for the simulation of water pollution (calamitous releases), the marine eco-system, dispersion of river water, sediment transport, etc.

Partial differential equations in porous media research

The modelling of transport processes in the subsurface, with emphasis on the analytical study of the governing partial differential equations. Basic and applied research includes nonlinear PDEs and free boundary problems, folds in layered geological structures, polymer gel injection, optimal body forms in gas particle streams, pattern formation, and asymptotics and special functions.

Exploratory research: discretization of initial value problems

Fundamental research though possibly application driven into theoretical numerical questions and properties connected with the discretization of initial value problems for differential equations, currently focusing on the numerical Method Of Lines for temporal discretization, splitting methods for the shallow water equations on the sphere, and sparse grid methods for time-dependent PDE problems.

Industrial Processes

Theme leader: P.W. Hemker

Computational fluid dynamics

Computation of flows in gases, liquids, or combinations of these (multi-phase flows) for industrial applications. Current research includes advanced discretization methods for systems of nonlinear conservation laws, multigrid and sparse-grid solution methods, numerical methods for linear algebra problems, local grid adaptation, and parallel and distributed computing. The present emphasis lies on the development of sparse-grid algorithms for 3D flow problems, parallel solution of very large eigenvalue problems, numerical methods for the computation of free-surface flows, and overset grid techniques for convection dominated problems.

Computational number theory and data security

Application of new mathematical and computational techniques for the solution of problems in number theory. Triggered by the emergence of public-key cryptography, the project studies algorithms for factorization and primality testing, for computing discrete logarithms, and for the solution of large, sparse systems of linear equations over finite fields. Currently the research utilizes the Number Field Sieve method.

Parallel software for implicit differential equations

Development of parallel software for the numerical solution of initial-value problems for implicit differential equations, which contain the classes of ordinary differential equations and differential-algebraic equations. The software can be used in electrical circuit simulation, constraint mechanical systems (robotica), chemical reaction kinetics, financial mathematical models, etc. The current focus is on waveform relaxation, and on further development of the CWI test set and the code PSIDE.

Mathematics of Finance
Theme leader: J.M. Schumacher

Mathematics of finance

The focus is on computational methods for derivative pricing and risk management, specifically Monte Carlo and Quasi Monte Carlo methods, and on the development of statistical procedures for stochastic processes.

Discontinuous dynamical systems

Complementarity-based modeling of hybrid dynamical systems that exhibit regime-switching behaviour. Study of well-posedness and dynamical properties. Applications for instance in electrical networks, piecewise linear systems, optimization, and mathematical finance.

Information Systems
Cluster leader: M.L. Kersten

Data Mining and Knowledge Discovery
Theme leader: A.P.J.M. Siebes

Data mining

For the data selection phase, research is focussed on structure in data, for example time-series, geographical data, multi-valued attributes, and non-universal relations, and on (non-)random samples. For data mining proper, the emphasis is on model-representation and search. An important aspect is the reformulation and generalization of well-known data mining algorithms in the KESO formalism. New research concerns multi-relational mining, especially mining multimedia data.

Database architecture

Research on main memory and parallel DBMS architectures, emphasizing facilitation of data mining and efficiency (the predominant database vehicle is MONET on the Medusa); further development of performance monitoring and prediction software; multi-query optimization and support of special data types (e.g., time-series databases and picture databases), in connection with efficient data mining.

Multimedia databases

The objective is to achieve efficient storage and retrieval of multimedia data, such as pictures, video and audio, in particular by using feature detectors to simplify and speed-up multimedia data query.

Multimedia and Human-Computer Interaction
Theme leader: H.L. Hardman

Distributed adaptive hypermedia

Investigation of models and implementation environments for the development of complex multi-/hypermedia documents. Particular attention is paid to supporting the author in creating different presentations or different output formats.

Distributed multimedia applications

[Suspended in connection with the set-up of the spin-off company Oratrix Development b.v.]

Interactive structured documents

For authoring structured documents a good design of the supporting architectures is essential. The research addresses generic aspects of the construction of interactive books, in which diverse software components have to cooperate efficiently, requiring (among other things) the development of theory for the propagation of incremental changes across changes in data representation, as well as developing proof-of-concept prototypes.

Interactive Information Engineering
Theme leader: P.J.W. ten Hagen

Information engineering framework

Objectives are the visualization of information structures on the basis of DAG (directed acyclic graph) presentations and the generation of navigation aids and presentation frameworks on the basis of the DAG layout, as well as the investigation of multimodal interaction tools for these complex presentations, and the use of intuitive, simple drawing techniques to generate illustrations in context.

Applied logic and inference engines

Application of logic, viewed as a science of information, to key issues in the processing and engineering of information, and the dissemination of this view of logic. Production of interactive textbooks where logic is presented as a core part of an emerging science of information processing and information flow analysis.

Facial animation

Production of a prototype system capable of capturing facial emotional expressions as enacted by a speaking performer and of reproducing user-controlled transformations of those expressions as part of information presentations.

Quantum Computing and Advanced Systems Research
Theme leader: P.M.B. Vitányi

Quantum computing

Investigation of quantum information and communication technology, quantum computer architectures, quantum algorithms, quantum information retrieval, quantum simulation of quantum mechanical physical systems at the elementary level (computational quantum matter) and quantum information theory.

MDL Learning and evolutionary computing

Design, implementation, and comparative analysis of a series of practical applications of machine learning techniques. Applications include automatic grammar generation from large text corpora and comparative evaluation of predictive accuracy of MDL and new forms of stochastic complexity, and GP learning of neural network governed robot locomotion and general techniques improving speed and storage requirements of GP implementations.

Advanced algorithms and systems

Design and analysis of algorithms for distributed and parallel systems. Limitations and possibilities of future systems are identified by exploiting fundamental mathematical techniques of (Kolmogorov) complexity theory. A major item is descriptional complexity leading to the `incompressibility method' and `learning by compression'. Also mobile and nomadic computing and communication are considered.

INTERNATIONAL AND NATIONAL PROGRAMMES

This appendix summarizes the major national and international projects in which CWI participates.
The following data are given for each project:
- title,
- period,
- cooperation with other institutes,
- CWI project leader(s).

European Programmes

ESPRIT

MERCURY (20089): Performance Management of Commercial Parallel Database Systems
1996-1998
ICL, IFATEC, ING, Herriot-Watt U.
M.L. Kersten

KESO (20596): Knowledge Extraction for Statistical Offices
1996-1998
National Statistical Offices of Finland, Greece (via FORTH), and The Netherlands, Infratest Burke, Data Distilleries, GMD, U. Helsinki
A.P.J.M. Siebes

CHAMELEON (20597): An Authoring Environment for Adaptive Multimedia Documents
1995-1998
CLRC, Epsilon SA, Cartermill International, Comunicacion Interactiva, Egnatia Epirus Foundation, Cycnos Systèmes Ouverts
D.C.A. Bulterman

DELOS (21057): ERCIM Digital Library
1996-1999
Elsevier, U. Michigan, all ERCIM Institutes
F.A. Roos

CONFER II (21836): Concurrency and Functions: Evaluation and Reduction
1996-2000
INRIA, ENS, CNET, ICL, KTH, Universities of Bologna, Cambridge, Edinburgh, Pisa, Sussex and Warwick
J.W. Klop

COTIC (23677): Concurrent Constraint Programming for time-critical applications
1997-2000
Universities of Utrecht, Pisa, Lisbon and Kent, SICS, CR&T
K.R. Apt
COORDINA (24512): From Coordination Models to Applications
1997-2000

INRIA, Xerox, U. Leiden, 8 European Universities, Signaal
J.J.M.M. Rutten

NeuroCOLT II (27150): Neural and computational learning
1998-2000
11 universities across Europe
P.M.B. Vitányi

VHS (26270): Verification and control of hybrid systems
1998-2001
U. Gent, INP Grenoble, U. Nijmegen
J.H. van Schuppen

DEDUGIS (28115): Deductive Constraint Databases for Intelligent Geographical Information Systems
1998-2001
CNR/CNUCE, U. Pisa, GMD - First Berlin, U. Würzburg, Sistemi Territorialli Pisa, DEBIS Berlin, INTECS Pisa, SISTEMA Grosseto
K.R. Apt

MAST Marine Science and Technology

MMARIE: Application of High Performance Computing Techniques for the Modelling of Marine Eco Systems
1995-1998
U. Leuven, Delft Hydraulics, U. Southampton, IFREMER, CRS Cagliari, U. Hamburg, U. Liège, U. Delft, RIKZ, CETIIS, U. Bradford, Hydraulic Research Wallingford, Proudman Oceanographic Laboratory Bridston, UP de Catalunya.
P.J. van der Houwen

ACTS

SEMPER (AC026): Secure Electronic Marketplace for Europe
1995-1998
Cryptomathic (DK), DigiCash (NL), Eurocom Expertise (GR), Europay International (B), FOGRA Forschungsgesellschaft Druck (D), GMD (D), IBM European Networking Center (D), Intracom (GR), KPN Research (NL), Otto-Versand (D), r3 security engineering (CH), SEPT (F), U. Freiburg, U. Hildesheim (D)
D.C.A. Bulterman

TELEMATICS

DACCORD (TR1017): Development and Application of Coordinated Control of Corridors
1996-1999
Hague Consulting, U. Delft, U. Lancaster, TNO, RWS, U. Naples, CSST, Autostrade Italia, INRTS, Ile de France, Ville de Paris, U. Crete, TCU
J.H. van Schuppen

EULER (LB5609): European Libraries and Electronic Resources in Mathematical Sciences
1998-2001
FIZ Karlsruhe, EMS, Documentaire Nationale pour les mathematiques, U. Lund, U. Göttingen, U. Degli Mari Group, U. Joseph Fourier
F.A. Roos/P.J.W. ten Hagen

TMR

DONET: Discrete Optimization: Theory and Applications
1998-2000
London School of Economics and Political Sciences, U. Pierre et Marie Curie (Paris), Rheinische U. Bonn, CNR, U. Lisbon, ALMA, DASH, Ecole Polytechnique Fédérale de Lausanne
A. Schrijver, A.H.M. Gerards

ERNSI: Systems Identification
1998-2002
KTH Stockholm, TU Wien, CNR-LADSEB, U. Leuven, INRIA, U. Rennes, U. Cambridge, U. Linköping, U. Eindhoven, U. Delft
J.H. van Schuppen

INTAS

Symmetry and cohomology approach to equations of mechanics and mathematical physics
1997-1999
Moscow Institute for Numerical Economy, Moscow State U., U. Twente, U. Salerno
M. Hazewinkel

ERETIMA: English-Russian enriched Thesaursus in Mathematics
1997-1999
U. München, Yaroslav State U., Russian Academy of Sciences, Steklov Institute of Mathematics
M. Hazewinkel

Multi-scale image analysis and applications
1997-1999
U. Leuven, U. Surrey, Institute of Mathematics (Gomel), Institute of Automation and Electrometry (Novosibirsk), Computing Center (Moscow), Institute of Engineering Cybernetics
H.J.A.M. Heijmans

Ergodic theory and dynamical systems
1998-2000
Universities of Tours, Barcelona, Marseille and Amsterdam, Ukrainian National Ac. of Science, State Pedagogical University Nizhny Novgorod
M.S. Keane

Mathematical methods for stochastic discrete event systems
1997-2001
U. Moscow, U. Novisibirsk, U. Cambridge, U. Braunschweig, INRIA
M.S. Keane

JOULE

WELGEL: Polymer Gel Injection
1998-2000
U. Delft, U. Wageningen, NAM, TNO, U. Leiden
C.J. van Duijn

INCO

Dr. Tesy: Methods and Tools for Distributed Real Time Embedded Systems Design and Analysis
1998-2001
Moscow State U., GMD Berlin, RedLab Ltd, State Research Institute of Aircraft Systems (Gosnias)
J.F. Groote

DEVIEW: Designing and Developing the Viewer Centred Paradigm in Virtual Environments
1998-2001
U. Capetown, U. College London
P.J.W. ten Hagen

Developing Software Engineering Environments for Distributed Information Systems
1998-2001
Universities of East Anglia, Manchester, and Cyprus, Space Application Services
F. Arbab

National Programmes

NWO-GBE (formerly SION/SWON)

From ideas to reality – implementing cryptography
1994-1998
L.G.L.T. Meertens

MDL Neurocomputing
1994-1998
P.M.B. Vitányi

Equational term graph rewriting
1994-1998
J.W. Klop

Generic tools for program analysis and optimization
1994-1998
P. Klint

Checking verification of concurrent systems with type theory tools
1994-1998
U. Utrecht
J.F. Groote

Constraints in object-oriented interactive graphics
1994-1998
U. Eindhoven
P.J.W. ten Hagen

Parallel declarative programming: transforming logic programs to lazy functional programs
1996-1998
K.R. Apt

Theorem proving and dynamic logic
1996-1998
ICCL, U. Stuttgart
D.J.N. van Eijck

PROMACS: Probabilistic methods for the analysis of continuous systems
1998-1999
U. Eindhoven, Free U. Amsterdam
J.J.M.M. Rutten

Classifying proof techniques for propositional logic
1994-1999
U. Delft
J.F. Groote

Large random systems and combinatorial probability
1998-2000
M.S. Keane

Monte Carlo and quasi-Monte Carlo simulation for efficient valuation and risk assessment of financial derivates
1998-2000
Universities of Groningen, Delft and Twente
J.M. Schumacher

Spectral parameters and embeddability of graphs
1998-2000
A. Schrijver

Strengthening semidefinite programming in coding and combinatorial optimization
1998-2000
A.M.H. Gerards

Parallel algorithms for solving large sparse linear equations over finite fields
U. Leiden
1998-2000
H.J.J. te Riele

A modular toolset for μCRL
1997-2000
U. Utrecht, U. Eindhoven, U. Amsterdam, U. Nijmegen, U. Groningen, U. Twente, Philips
J.F. Groote

Protocols, reference models and interaction schemes for multimedia environments
1997-2000
J.F. Groote

Dynamic algorithms for on-line optimization
1997-2000

Philips Research
J.A. La Poutré

Fractal image coding
1996-2000
U. Delft
H.J.A.M. Heijmans

Modelling of hybrid systems
1996-2000
U. Groningen, U. Twente
J.M. Schumacher

Statistical properties of movements in the plane
1996-2000
Math. Inst. Hungarian Ac. Sc., Universities of Utrecht, Delft, Leuven, Cambridge, Chalmers Gothenburg, and Rome
M.S. Keane

Integer polyhedra and binary spaces
1996-2000
A.M.H. Gerards

Discontinuous dynamical systems
1996-2000
Universities of Groningen, Twente, and Brabant
J.M. Schumacher

Quality of service for multimedia systems
1997-2001
Philips Research
J.A. La Poutré

AMIS: Advanced Multimedia Indexing and Searching
1997-2001
Data Distilleries, ICL, IFATEC, ING-Group, Tandem, Herriot-Watt U., U. Twente, U. Eindhoven, U. Amsterdam
M.L. Kersten

Quantum Computing
1997-2001
U. Amsterdam
P.M.B. Vitányi

Nonexpansive mapping and their asymptotics
1997-2001
U. Utrecht, U. Delft, U. Leuven, U. Cambridge, Free U. Amsterdam
M.S. Keane

Sparse grid methods for time-dependent PDE problems

1998-2001
UU/IMAU, RIVM, KNMI, TNO, U. Iowa
J.G. Verwer, B. Koren

Parallel computational magneto-fluid dynamics
1997-2001
U. Utrecht, FOM Inst. Plasma Physics
H.J.J. te Riele

CAM: Computational number theory and data security
1997-2001
U. Groningen, U. Leiden, U. San Rafael, Australian Nat. U. Canberra, U. Macquarie at Sydney, U. Bordeaux, U. Saarbrücken, Citibank
H.J.J. te Riele

Rigid sets
1998-2002
Free U. Amsterdam
M.S. Keane

Statistics for random processes with applications to mathematical finance
Free U. Amsterdam
1998-2002
K.O. Dzhaparidze

Connected morphological operators for image analysis
1998-2002
H.J.A.M. Heijmans

COCO: Computational Intelligence for Constraint Logic Programming
1997-2002
Partners in ERCIM WG on Constraints
K.R. Apt

COLA: Formal methods and refinement for coordination languages
1997-2002
U. Leiden, Signaal
J.J.M.M. Rutten, F. Arbab

Statistics for random processes with applications to mathematical finance
1998-2002
Free U. Amsterdam
K.O. Dzhaparidze

STW (Foundation for the Technical Sciences)

Wavelets: analysis of seismic signals
1996-1999
Universities of Delft, Eindhoven, and Groningen, Shell, KNMI, MARIN
N.M. Temme

FASE: Facial animation
1997-2000
U. Delft, Philips, NOB, Institute for the Deaf, KPN Research
P.J.W. ten Hagen

MOBILECOM: mobile communication networks
1998-2001
U. Amsterdam, Free U. Amsterdam, U. Eindhoven, U. Delft, KPN, Libertel
R.J. Boucherie

Waveform relaxation
1998-2001
Philips Research Lab., U. Amsterdam
P.J. van der Houwen

Multiresolution image analysis and synthesis
1998-2002
Johns Hopkins U., TNO, AKZO-Organon, Signaal
H.J.A.M. Heijmans

Special NWO projects

Plasma simulation
1993-1998
U. Utrecht, FOM Inst. Plasma Physics
H.J.J. te Riele

CIMS: Parameter estimation for random sets
1997-1999
U. Western Australia (Perth)
M.C. van Lieshout

NWO-MPR: Parallel solution of very large eigenvalue problems
1998-2001
U. Utrecht, FOM
H.J.J. te Riele

NWO-CAM: Number Field Sieve and related subjects
1997-2001
U. Oxford, Australian National U., Citibank New York, San Rafael, U. Groningen, U. Leiden, Macquarie U. Sydney, U. Bordeaux, U. Georgia, U. Giessen, IRI Toulouse
H.J.J. te Riele

Applied logic dissemination (Spinoza)
1997-2002
ILLC
D.J.N. van Eijck

NCF
Virtual reality
1996-1999
ECN, U. Delft, U. Amsterdam
R. van Liere

ICES HPCN Programme
High Performance Visualization
1996-1998
ACE, CAP Volmac, Arcobel, TNO
R. van Liere

NICE: Computational Fluid Dynamics
1996-1999
Delft Hydraulics, CUNY Brooklyn
F. Arbab

HPCN for Environmental Applications
1996-1999
U. Delft, Delft Hydraulics, TNO
J.G. Verwer

IMPACT: HPCN for Financial Services
1996-2000
ING, U. Amsterdam, U. Twente, Getronics, U. Delft, CAP Volmac, Data Distilleries, BIT by BIT
A.P.J.M. Siebes, M.L. Kersten

Telematics Institute

DMW: Digital Media Warehouse Systems
1998-2002
CTIT, TICO, KPN, Syllogic
M.L. Kersten

SVC: Systems Validation Centre
1998-2002
CTIT, KPN, CMG, Lucent, TI
J.F. Groote

DRUID: Multimedia indexing and retrieval on the basis of image processing and language and speech technology
1999-2003
TNO, U. Twente, OCE
M.L. Kersten

U-Wish: Web-based service for information and commerce
1999-2001
TNO-TM, CTIT
S. Pemberton

DSL: Domain Specific Languages
1999-2002
TI, ING Bank, Cap Gemini, Lucent
P. Klint

QFN: Quality-of-service in Future Networks
1999-2002
CTIT, U. Eindhoven, KPN, U. Delft, Libertel
S.C. Borst

MC: Mediated Communication in collaborative work
1999-2002
U. Delft, TNO-TM, TNO-TPD, KPN, Inst. for the Deaf, NOB, Philips
P.J.W. ten Hagen