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 2000 ©
Production Bureau CWI
Design Tobias Baanders / Facility Department CWI
Printing drukkerij Mart.Spruijt bv, Amsterdam

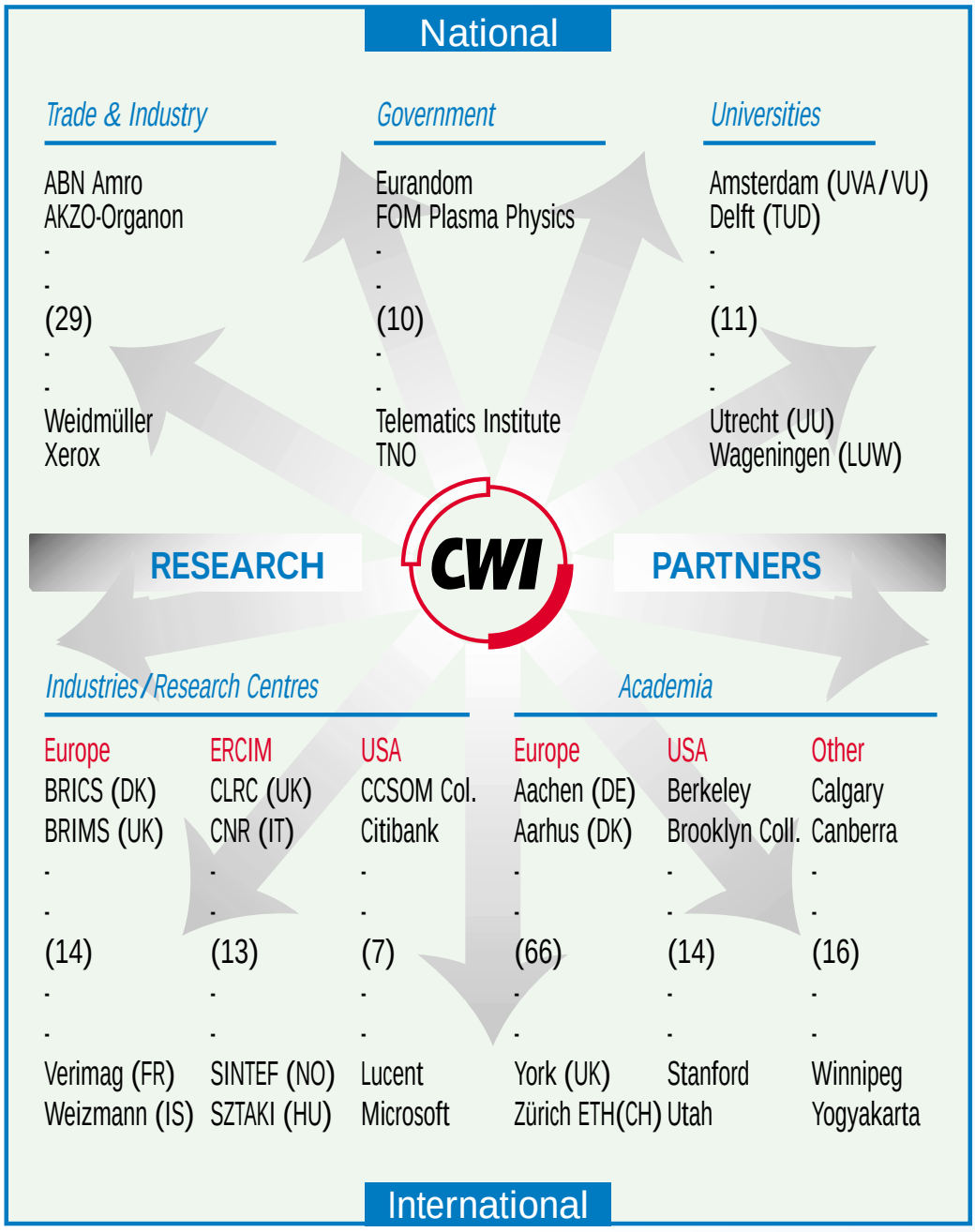
Stichting Mathematisch Centrum / Centrum voor Wiskunde en Informatica
Visiting address Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
Postal address P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Telephone +31 20 592 9333
Telefax +31 20 592 4199
Website www.cwi.nl

OVERVIEW

The turn of a year is often used for looking back and into the future, and this inclination is even stronger in a special case like the transition to a new century. CWI used this opportunity in 1999 to review its position in the scientific world and in society. Pivotal was the evaluation of the institute requested by the Dutch Science Council NWO, CWI's major funding organization. As a preparation for this evaluation CWI produced two documents: Progress Report 1993–1998, *and* Strategy 2000–2005. *The outcome of the evaluation was very positive for the institute, but did not lead to better financial prospects. In the strategy document CWI chooses for a scenario where the expected growth of externally financed projects is coupled with an increase in basic subsidy by NWO. CWI considers this as the only way to keep long-term fundamental research (basic funding) in balance with application-oriented projects (external funding). Such a balance is necessary to warrant the unique profile and quality of CWI. (An externally funded project often brings with it a financial commitment by 'matching'.) Meanwhile, however, NWO has reduced CWI's basic subsidy, starting from the year 2000, as a part of a general reduction of expenses. Hence, CWI has started to look for other ways to realize its growth scenario.The overview below of CWI's activities in 1999 yields ample evidence that these attempts are fully warranted.*

Evaluation

A committee chaired by J.C.M. Baeten (Eindhoven University of Technology) has evaluated CWI at the request of NWO early June. Key questions were the positioning of the institute, and its financing by NWO. Mission, strategy, and performance of CWI were reviewed. The other committee members were: F. Baccelli (École Normale Supérieure, Paris), C. Hankin (Imperial College, London), K.H. Hoffmann (Technical University Munich), and E.J. Neuhold (Technical University Darmstadt). The committee judged the institute on the whole as very good, and half of its research themes even as excellent. Given its limited financial means, CWI's activities have a considerable impact. The mission of CWI remains as it was and should be: to perform frontier research in strategically chosen areas of mathematics and computer science, and to transfer new knowledge in these areas to society in general, and Trade & Industry in particular. At the organizational level the institute has been very active to initiate new research themes in larger groups and to ease interdisciplinary research. Increased flexibility is apparent from the regular starting of new pilot projects (and the termination of projects which have run their course). Ties with industry were reinforced by participation in the national Telematics Institute, as well as by the successful launching of new spin-off companies. Contacts with the academic world were intensified by extending mutual secondments and new cooperation agreements with two graduate schools in geophysics and plasma physics. Internationally, CWI's leading role in the European Research Consortium for Informatics and Mathematics ERCIM was highly appreciated. The bilateral cooperation agreement with GMD (Germany), which started in


National
Trade & Industry
ABN Amro
AKZO-Organon
-
-
(29)
-
-
Weidmüller
Xerox
Government
Eurandom
FOM Plasma Physics
-
-
(10)
-
-
Telematics Institute
TNO
Universities
Amsterdam (UVA/VU)
Delft (TUD)
-
-
(11)
-
-
Utrecht (UU)
Wageningen (LUW)
RESEARCH
CWI
PARTNERS
Industries/Research Centres
Europe
BRICS (DK)
BRIMS (UK)
-
-
(14)
-
-
Verimag (FR)
Weizmann (IS)
ERCIM
CLRC (UK)
CNR (IT)
-
-
(13)
-
-
SINTEF (NO)
SZTAKI (HU)
USA
CCSOM Col.
Citibank
-
-
(7)
-
-
Lucent
Microsoft
Academia
Europe
Aachen (DE)
Aarhus (DK)
-
-
(66)
-
-
York (UK)
Zürich ETH(CH)
USA
Berkeley
Brooklyn Coll.
-
-
(14)
-
-
Stanford
Utah
Other
Calgary
Canberra
-
-
(16)
-
-
Winnipeg
Yogyakarta
International

1998, was seen as a model example worthy to be followed. Finally, the encouraging developments concerning WTCW – the complex of research and education institutions and ICT companies in the Amsterdam district Watergraafsmeer which received substantial financial support for strengthening the knowledge infrastructure – make CWI also geographically very well positioned for the future.

Strategy

CWI enters the new century with a strong starting position. The need for a national research institute in mathematics and computer science is undisputed: many drastic changes in science and society originate from new research in these fields. The organization of CWI's research is flexible, interdisciplinary, and dynamic, enabling rapid adaptation to scientific and societal developments. This becomes manifest in, for example, an active policy to create spin-off companies. The report year saw the birth of two such companies: Oratrix and Eidetica. A point of lasting concern is the recruitment of highly qualified researchers in sufficient numbers to carry out the many ambitious projects which CWI has in stock. Dutch academia produces too few talented graduates in mathematics, whereas in computer science the attraction of the industry, in particular the ICT industry, is a formidable factor to be taken into account.

The institute aims at an ever wider recognition as a centre of excellence, on the national as well as on the international level: a favourite partner for universities, research organizations and companies to perform joint research. To this end CWI reinforces its ties with universities and national graduate schools and extends its international cooperation, for example by bilateral agreements with ERCIM partners. All along its existence CWI has entered new research areas, and played in The Netherlands a pioneering role in several of them. Examples include: computers, statistics, operations research, biomathematics, and cryptography. CWI continues starting up new research themes which, after proven viability during a period as a 'pilot' theme, last 10–15 years on the average. The report year saw the start of new research in the field of mathematical modeling and visualization with application to biology and medicine.

Topics in which CWI played a pioneering role in The Netherlands.

computers

statistics

operations research

biomathematics

cryptography

ERCIM Jubilee event on November 5th.

Audience listening to Heikki Hämmäinen (Nokia).

Blowing the candles: President Gerard van Oortmerssen (left) and Vice-President Stelios Orphanoudakis (right).

Lunch in the main hall of the Beurs van Berlage.

Demonstration of the Cycab (INRIA).

10 years ERCIM

The European Research Consortium for Informatics and Mathematics ERCIM was formally established in April 1989, when the directors of GMD (Germany), INRIA (France), and CWI signed a cooperation agreement. In November 1999, CWI hosted several hundreds of guests, from ERCIM and beyond, who attended a two-day symposium in the architecturally unique setting of the Beurs van Berlage in Amsterdam. In the course of the first ten years ERCIM has grown into a consortium which includes leading research institutions from 14 European countries, representing together over 7.000 researchers. ERCIM researchers cooperate in twelve Working Groups, and regularly workshops are organized. A Fellowship Programme has offered up to now about a hundred young promising researchers from the whole world the opportunity to work at one or more ERCIM institutes. ERCIM sponsors conferences on a regular basis, for example the annual World Wide Web conferences which attract in general well over a thousand participants. CWI hosts *WWW9*, which will take place in Amsterdam in May 2000. Every year – in 1999 for the fifth time – the Cor Baayen Award is granted to a researcher of the ERCIM community, on the level of postdoc with some years of experience. The Award is named after the first ERCIM President (and former director of CWI). Research performed at ERCIM institutes is reported in the quarterly *ERCIM News*. ERCIM is regarded as an important actor on the European level. This became apparent when in the report year the European Commission asked ERCIM to establish, on behalf of the European Union, a cooperation programme with the American National Science Foundation NSF. CWI has played a comparatively prominent role in ERCIM from the outset. For example, the first and the present Presidents are from CWI. The choice of Amsterdam as the location for the jubilee festivities underlined this role once again. The first day of the symposium was targeted at ERCIM scientific and administrative staff. Some twenty ERCIM researchers and invited speakers covered research related issues and gave state-of-the-art presentations of their research. Particular attention went to the presentation by Dennis Tsichritzis (director of GMD) 'The Changing Art of Research'. On the second day, leaders from the field of Information and Communi-cation Technologies, Manufacturing, and Information Content gave their vision on the future of European R&D. The speakers were: Gerard van Oortmerssen (CWI, President of ERCIM), Heikki Hämmäinen (Nokia), Dieter Klumpp (Alcatel SEL), Roger Needham (Microsoft Europe), Alexander Rinnooy Kan (ING Group), and Jacques-Louis Lions (Institut de France). The day was concluded with a plenary discussion session led by Dick Bulterman (Oratrix), on the basis of a short video presentation by Tim Berners-Lee (MIT) about the future of the Web and the role of Europe, in particular ERCIM. During the two days there were non-stop demonstrations of current research taking place at several ERCIM institutes.

Computer infrastructure

CWI received during the period 1994–1996 means from NWO to carry through a much-needed upgrading of its internal electronic communication. A 228 km glass fibre network connecting about 200 desks enabled information to be transferred at a rate of 155 Mb/sec and higher. Research into multimedia and visualization profited in particular. Investments in computer infrastructure substantially increased again from 1998. This boost is strongly connected with the new generation Internet and the processing of multimedia information flows that require considerably larger (Gigabit) bandwidths and server capacity. The move into the Gigabit area takes place at three

Constraint satisfaction problems

Constraints have been studied in Artificial Intelligence starting from the seventies. The central notion is that of a constraint satisfaction problem (CSP). A CSP consists of a finite set of constraints, which are simply relations over some domains. A solution to a CSP is a sequence of values from the underlying domains that satisfies each constraint after projecting it on the relevant domains. The task of constraint programming consists of formulating the initial problem as a CSP and of solving it by means of general or domain specific methods. 'Solving' can mean finding a solution, all solutions or the best solution with respect to some cost function.The general methods are usually concerned with the techniques of reducing the search space and with specific search methods. The main idea is to reduce a given CSP to another one that is equivalent (i.e., has the same set of solutions) but is smaller. This process is called constraint propagation and the algorithms that achieve such a reduction are called constraint propagation algorithms. So these algorithms reduce the search space and consequently attempt to limit the combinatorial explosion. Which type of constraint propagation is used depends on the initial choice of constraints and domains and on the applications. In the literature literally dozens of constraint propagation algorithms were introduced and investigated, the most known being so-called arc consistency and path consistency algorithms. In our recent work at CWI we provided a uniform framework for these algorithms that allowed us to present and explain them in a uniform way. In this framework we proceeded in two steps. First, we introduced a generic iteration algorithm on an arbitrary partial ordering with the least element $\bot$ (see below), and proved its correctness in an abstract setting. Then we instantiated this algorithm with specific partial orderings, functions and the update definition to obtain specific constraint propagation algorithms. In this analysis various properties of the scheduled functions, to wit monotonicity, inflationarity, idempotence, commutativity and semi-commutativity, were of importance. So specific constraint propagation algorithms could be obtained simply by suitably instantiating a general scheme with specific functions to be scheduled. This was used for instance in a recent work of Monfroy for the case of non-linear constraints on reals. In another recent work, Monfroy and Réty generalized this framework and showed how such generic iteration algorithms can be parallelized. This led to a framework for parallel and distributed constraint propagation. Such a general framework has several advantages. Using it we can now more easily derive, verify, compare, parallelize, modify or combine constraint algorithms. Currently we are studying how this general framework for constraint solving can be realized using the coordination language MANIFOLD developed at CWI by F. Arbab.

F is a finite set of monotonic and inflationary functions on a partial ordering with the least element $\bot$. The following nondeterministic algorithm computes the least common fixpoint of the functions from F.

$d := \bot$;
$G := F$;
while $G \neq \emptyset$ do
    choose $g \in G$;
    $G := G - \{g\}$;
    $G := G \cup update(G, g, d)$;
    $d := g(d)$
od

where for all G, g, d the set of functions $update(G, g, d)$ from F is such that

A. $\{ f \in F - G \mid f(d) = d \wedge f(g(d)) \neq g(d)\} \subseteq update(G, g, d)$,

B. $g(d) = d$ implies that $update(G, g, d) = \emptyset$,

C. $g(g(d)) \neq g(d)$ implies that $g \in update(G, g, d)$.

The Generic Iteration Algorithm.

levels. In the framework of the GigaPort project, carried out by the Dutch academic network organization SURFnet and the Telematics Institute, the external connectivity of the Science and Technology Centre Watergraafsmeer WTCW will be upgraded into the Gigabit range. At WTCW level this is accomplished in the framework of the WTCW project financed from ICES-KIS funds. Of course, this also requires adaptation of the local computer network, and CWI already started with replacing the present generation of workstations. At CWI level, the main innovation of the year, however, was the installation of a very powerful Silicon Graphics supercomputer (Medusa), specifically meant for multimedia data warehousing. The purchase was co-financed by the NWO foundation National Computing Facilities NCF and the ICES-KIS fund. The computer is used in research into querying large multimedia databases and in virtual reality research connected with applications in biology and medicine. Also other CWI research requiring such a powerful computer is performed on Medusa, for example software correctness proofs, evolutionary methods for E-commerce, and factoring large numbers. Moreover, through NCF Medusa is also available for users throughout The Netherlands. CWI research in parallel computation methods (number crunching) is mainly carried out on supercomputers elsewhere (Academic Computing Centre Amsterdam SARA, Delft University of Technology TUD, and in the USA), and is financed by NCF.

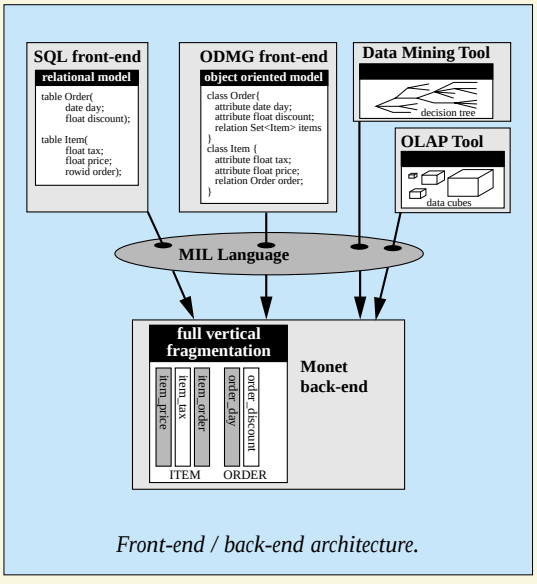
Bio-Informatics research: mathematical modeling and interactive visualization of biological processes.

Telematics Institute

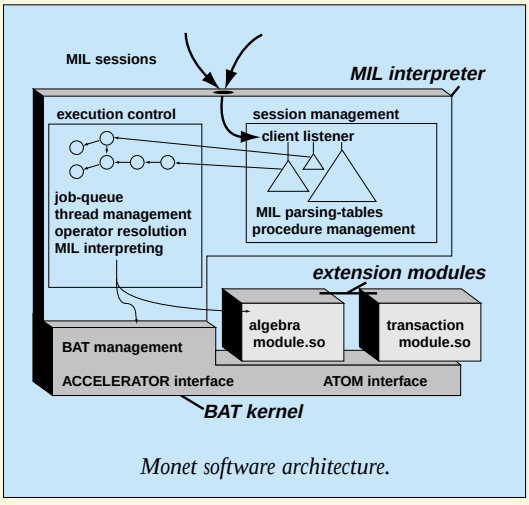
The Telematics Institute was launched two years ago in a joint effort of the ministers of Education, Culture & Science, of Economic Affairs, and of Agriculture, Environmental Protection & Fishery. It was officially opened last year. Here knowledge institutions like CWI closely cooperate with Trade & Industry and governmental bodies in the field of telematics. CWI's research profile goes very well with the starting-points of the Telematics Institute. Participating knowledge institutions are the universities of Twente and Delft, the organization for Applied Scientific Research (TNO-MET), and CWI. The Catholic University Brabant will join shortly. CWI's participation in Telematics Institute projects has increased to eight projects in 1999, the most recent ones being the Systems Validation Centre, Intelligent Agents in E-Commerce, and Domain Specific Languages.

Monet -- a main-memory database system

Most commercial relational DBMS products stem from a design line that originates from the late 1970's, and hence were carefully designed and tuned to application requirements and hardware characteristics at that time. Technically speaking, their storage infrastructure remains optimized towards the needs of OLTP (OnLine Transaction Processing), which requires high performance on large numbers of small updates. Query intensive applications like OLAP (OnLine Analytical Processing) and data mining, however, have a distinct different access pattern, as they condense large volumes of data in small, summarized results. Monet is a new extensible database kernel, specifically targeted to query-intensive applications. It uses a fully fragmented data model that consists of binary tables only, and its query processing infrastructure is optimized towards main-memory execution. Binary tables provide a canonical representation scheme for both relational, object-oriented and XML front-ends. Furthermore, a concise algebraic language MIL (Monet Interface Language) acts as a target for query language compilers and optimizers. They can be efficiently implemented with demonstrated supreme performance on both traditional (TPC-D) and modern (TPC-H,OO7,DD) database benchmarks.The focus on main-memory is warranted by the shift in price/performance of storage systems. Unlike in the 1970's multi-gigabyte random access memories are now affordable alternatives for disks as the prime storage area. The experimentation platform at CWI consists of a 32-cpu Origin 2000 machine with 64 Gbyte of RAM. Such large memories, however, call for a major overhaul of a database kernel. CPU cycles and cache management are becoming the major performance factors, rather than the number of disk accesses. The Monet database kernel has already been designed on these premises since 1993. For the coming years the research target has been set on a complete redesign of query optimization models and technology within the context of high demanding data mining and multi-media database applications. Query optimization calls for better (simple) cost metrics to predict the execution time of individual query plans in a parallel main-memory setting. Furthermore, the application domains call for advances in session-wide optimization and just-in-time optimization. These research areas remained largely unexplored in recent decades. The research activities are aligned with the needs of the CWI spin-off company Data Distilleries and the many research partners in the Telematics Institute, WTCW ICES/KIS, and Gigaport context.



Front-end / back-end architecture.



Monet software architecture.

Knowledge infrastructure

The government's investments in infrastructure, financed from the revenues of natural gas exploitation, has led to supporting the development of the knowledge infrastructure at the Watergraafsmeer campus. For this development a total budget of 70 Mfl has been allocated for four years. The government contributes with 30Mfl from its ICES-KIS programme (reinforcement knowledge infrastructure) for Research & Development. The subsequent activities are coordinated by a new company WTCW (Science and Technology Centre Watergraafsmeer). Shareholders are the University of Amsterdam, NWO, the Municipality of Amsterdam, and six companies. WTCW has defined four projects: Multimedia Information & Analysis, Biotechnology and Bio-Informatics, Biodiversity and Environment, and Virtual Lab. CWI participates in the first two projects. CWI's annual day for Trade & Industry 'CWI in the Market Place', focused in 1999 on investment in knowledge infrastructure. The day attracted with 120 participants again more attention in the target group. CWI also coordinates the annual Day for the Public of the whole Watergraafsmeer campus, as part of the National Science Awareness Week. This event attracted in 1999 more than three thousand visitors to the campus. CWI researcher Herman te Riele held a well-attended lecture about the recent breaking of the RSA-155 security code, to which he and his group had made essential contributions.

Internet / W W W

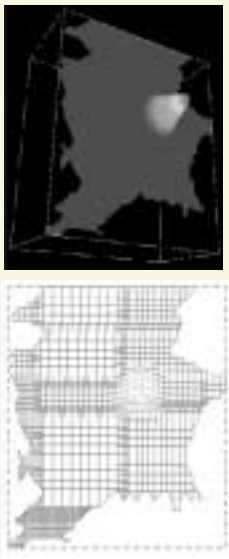
CWI was already involved in the development of the Internet and the World Wide Web from the very beginning. The institute became the first non-military Internet site in Europe (1988) and acted for several years as the gateway between the USA and Europe. Later on CWI initiated and managed the Internet domain registration for The Netherlands, which became a working model for registries abroad. This activity was transferred to a separate foundation in 1997. For his pioneering achievements in the development of the Internet in The Netherlands and in Europe, CWI staff member Piet Beertema was knighted in June (Ridder in de Orde van de Nederlandse Leeuw).

Piet Beertema and his decoration.

CWI has made important contributions to the World Wide Web ever since its inception, for example by developing the Web language Python, and by work on HTML, CSS (Cascading Style Sheets), and the multimedia standard SMIL. Furthermore CWI created spin-off companies like NLNet (the first Dutch Internet Service Provider) and General Design (the first Dutch company for designing Web sites). The institute participates in the activities of the World Wide Web Consortium W3C, in which almost

Surface water quality modeling

Shallow seas, such as continental shelf areas, play several important roles. First, they form a transitional region between land and ocean: besides input of substances from rivers, the conditions are also influenced by the entrance of sediments and (solar and wind) energy from the ocean. Secondly, shallow seas (like the North Sea) are often of economic importance. Aspects that are of human interest include land protection, shipping, fishery, recreation, exploitation of oil and gas fields, etc. A third important role concerns ecological aspects, since shallow seas are highly productive ecosystems. They continuously receive discharges from continental run-off, which contain ecologically favourable (nutrients) and unfavourable (pollutants) substances. Accidental discharges form another example of a serious pollution of the sea. A poignant disaster is the recent pollution of the coast of Brittany, France, due to the crash of an oil tanker. Obviously, for a proper management of marine ecosystems, scientists as well as politicians are greatly interested in questions like: what will be the impact of human interaction on the natural balance, especially in the long term? Since these systems are of an enormous complexity, numerical simulation will be the only feasible way to provide insight. At CWI, the research concentrated on the transport of the various species through the sea, and their interaction due to chemical reactions. These phenomena can be modeled by a set of partial differential equations, defined in three spatial dimensions. To arrive at a discrete numerical solution in space and time we first cover the sea by a spatial grid. For efficiency reasons, the sea is first divided into a number of subdomains on which grids of different resolution can be used. In subdomains where the concentrations show large variation, for example along coasts and in estuaries, a relatively fine grid is necessary, whereas a much coarser mesh can be used in 'inactive' domains, without reducing the degree of the overall accuracy. In this way, the total number of grid points, and hence the storage requirements and CPU time, can be reduced significantly. An additional advantage of the domain decomposition approach is that it allows for parallelism, resulting in a considerable speed-up on multi-processor computers. Then, discretizing the spatial derivatives on the subdomain grids results in a huge system of ordinary differential equations (ODEs), which has to be integrated in time. Although many general ODE algorithms exist, we designed a special high performance algorithm, taking into account all special features of this particular application. Such a special purpose algorithm is the only way to arrive at a feasible computation time. To give an impression: covering the North Sea with 15 domains, each having 50*50 horizontal grid points and 10 vertical layers, and taking 10 species into account, leads to close to 4 million ODEs. The algorithm developed in this project is able to simulate 1 day real time within 5 CPU seconds on the CRAY C90. This makes it possible to simulate a full year real time within half an hour computing time.



Computed concentrations of one particular species in the North Sea (above), decomposed into appropriate subdomains (below).

Oil pollution on the beach. ©The Picture Box / Nordwin L. Alberts.

four hundred companies and institutions worldwide are represented. W3C aims at an optimal use of the Web's potential, for example by the development of common protocols (standards). The W3C working group on XHTML, chaired by CWI researcher Steven Pemberton, released in the report year the specification language XHTML 1.0, which meanwhile has been accepted by W3C as a Recommendation. This means that the product has been found stable, contributes to the interoperability of the Web, and has been approved by the W3C membership, thus recommending it as an industrial standard. XHTML forms a bridge between the present mark-up language HTML and XML, the Web language of the future. CWI also runs the W3C office in The Netherlands. The office's activities in 1999 included three tutorials on XML, CSS, and XHTML, given by Steven Pemberton.

European and national projects

Apart from the ongoing cooperation in the framework of ERCIM, CWI concluded in the report year a bilateral agreement with ERCIM member SZTAKI (Hungary) for joint research in control theory and evolutionary systems. Similar agreements with GMD (Germany) and INRIA (France) had already come into operation earlier.

Signing the GMD-CWI agreement on scientific cooperation: DennisTsichritzis, director of GMD (left), Gerard van Oortmerssen, director of CWI (right).

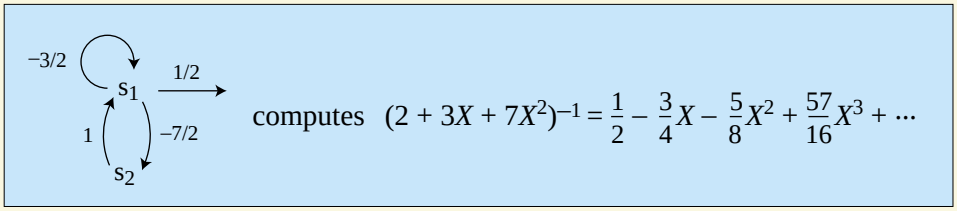
Due to the slow start of the European IST programme (successor of ESPRIT), participation of CWI was limited in 1999 to two projects. It is expected that this will increase again from 2000. CWI conducts research in several dozens of projects financed by Dutch institutions. The majority is funded by NWO, mainly through its Council for the Sciences, but also through its foundations STW (technology) and NCF (national computer facilities). Besides, CWI participates in three projects in the framework of the ICES-KIS programme (reinforcement knowledge infrastructure), eight projects of the Telematics Institute (telecommunication), and one EUREKA project (Internet). Finally, contract research is carried out commissioned by some fifteen companies and institutions.

Spin-off companies

CWI created two spin-off companies in 1999: Oratrix Development and Eidetica. Oratrix Development started operations in early 1999, after having won the business plan contest New Venture 1998 for starting entrepreneurs in the (ICT) branch. The company, founded by CWI staff members Dick Bulterman, Lynda Hardman, Jack Jansen, and Sjoerd Mullender, produces streaming media authoring systems and players for use on the Internet. During the first year, the company focused on transferring the research technology developed at CWI into an initial product offering. Several prototype versions of the system were released in the course of 1999. As with nearly all companies of Oratrix's size, a considerable effort was invested in obtaining the

Coalgebra

Picture yourself sitting behind a computer: it has an internal state, given by, among other things, the contents of its memory cells and its registers. This state is not directly observable, but the screen provides certain observable information about it, by means of some specific operations. A computer is also a dynamical system in that its internal state can be modified by other operations (commands). Coalgebras are simple mathematical structures capturing this type of behaviour of systems, consisting of observations and internal dynamics. They were introduced in the early eighties to deal with infinite data structures, for example infinite sequences (streams). Suppose that only the first element (head) of a stream is observable, the remainder (tail) remaining hidden. Now we remove the head. This internal operation induces a change in the observable world as well, because now the head of the new stream, which is the tail of the old one, can be observed. Coalgebras operate in fact just the other way round as algebras, where for example two elements are associated to form one new element of a set. The notion of coalgebra is a categorical dualization of the notion of algebra. Algebra forms a well-established part of computer science, for example computer algebra and axiomatic semantics (process algebra). Also data types are often modeled as initial (term) algebras, along with a generalized definition and proof principle of induction. There are limitations, however. It turns out that algebraic descriptions of state-based systems are not so natural and successful, especially when nondeterminism is involved. (The state space then has a tree-like structure, contrary to the one sequence of states in the deterministic case.) Algebraically, one thinks of the elements of a given structure as constructed in a particular manner. Often, however, it seems more natural to view the state space of a system as a black box, where in principle nothing is known about the states or about the way they are constructed. In a coalgebra a function acts on states, and this function is typically composed of several components, meant for both the observation and the modification of states. Considering coalgebraic structures instead of algebraic ones may seem like a trivial step, but it has turned out to be a fundamental one, opening many new perspectives. In fact, many coalgebraic structures have been used for a long time in various situations in mathematics and computer science, but people usually did not identify them as such. It has only recently been recognized that coalgebras form the underlying structure of various kinds of dynamical systems, automata, transition systems, infinite data types, object-oriented systems, formal power series, and even various classes of differential equations. By now, the subject has gained worldwide recognition and interest among computer scientists, logicians, and mathematicians alike. CWI has been one of the first to put this exciting and promising field of fundamental research on its agenda. Cooperation, both internally on the use of coalgebra for the supervisory control of discrete event systems, and with the Spinoza project at the University of Amsterdam (on coalgebra as a model theory for modal logic), are representative illustrations of the highly interdisciplinary character of the study of coalgebra.



By first generalizing the notion of nondeterminism (extensively studied in various branches of computer science) coalgebraically, and then applying the resulting theory to the mathematical field of generating functions, finite nondeterministic representations of infinite power series can be constructed.

financing necessary to build a viable organization. An early round of financing was completed late 1999. Initial distribution agreements were completed with several external parties, including a joint distribution and sales agreement with Seattle-based RealNetworks, Inc, the leading supplier of streaming media technology worldwide. Oratrix is an active participant in the W3C working group defining the next version of the SMIL language. This new version is expected in mid-2000. Oratrix's product offering is available at `http://www.oratrix.com/GriNS/`
Eidetica specializes in software development for classification and indexing of text and text-specific databases. Eidetica is a Twinning company founded early 1999. Twinning is an initiative of the ministry of Economic Affairs to stimulate starting ICT companies. Eidetica's products and services are based upon a textual repository system, integrating features of a database and an information retrieval system. At the end of the year these were ready for the (business-to-business) market:
- t-find: advanced, professional and reliable Internet and Intranet search for knowledge
- t-store: large textual database and indexing services for autonomous agent systems and information routing
- t-mining: text mining and automatic classification systems.

At the end of 1999, Eidetica signed contracts with the EU for its LIMES project (starting April 2000), and with the newspaper publisher PersCombinatie for a pilot project, which meanwhile has resulted in a large follow-up project (starting April 2000). For more information: `http://www.eidetica.com`

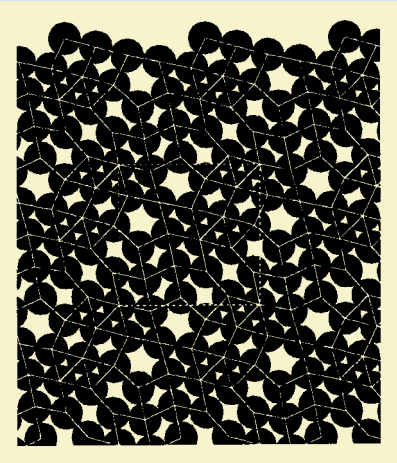
Academia

CWI's traditionally intensive relations with the academic world were further reinforced during the year. Contacts mainly consist of joint research projects and mutual secondment of researchers. A joint venture was created with the University of Amsterdam and the Free University in financial mathematics. Agreements for the exchange of researchers exist with Eindhoven University of Technology and the Free University (Amsterdam). New cooperation started with national graduate schools in geophysics and plasma physics. (CWI already has cooperation agreements with almost all graduate schools in mathematics and computer science.) The close ties of CWI with academia is also apparent from the fact that some twenty staff members combine their position at CWI with a professorship at a Dutch university. J.M. Schumacher, who led at CWI the theme on Financial Mathematics, received a full-time appointment as a professor at the Catholic University Brabant. Staff members A.M.H. Gerards and A.P.J.M. Siebes received a part-time professorship.

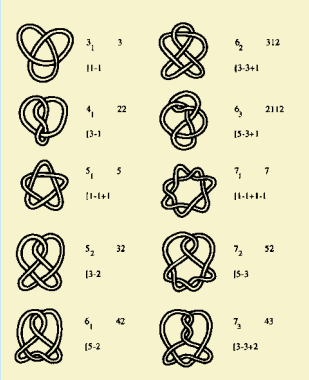
Conferences and courses

CWI was always very active in organizing all kinds of meetings like conferences, symposia, and workshops, thus underlining its central role in the Dutch research world. The aforementioned two-day symposium on the occasion of the 10-year jubilee of ERCIM drew several hundreds of professionals to Amsterdam. A fixed point is the annual vacation course for mathematics teachers, which was given without interruption from CWI's foundation in 1946 (except in 1954, when the International Mathematics Congress was held in Amsterdam). This course is still highly appreciated by the Dutch mathematics community. The 1999 subject was 'Unproven Conjectures'. In the framework of CWI's joint venture with both Amsterdam universities, courses in financial mathematics were set up. CWI organized the annual International ETAPS conference (European Joint Conferences on the Theory and Practice of Software) in

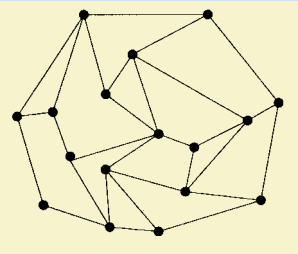
CWI's 1999 Vacation Course for mathematics teachers focused on Unproved Conjectures.



Packing problems.



Classification of knots.



Hamiltonian paths in graphs.

CWI Soirée during the 1999 ETAPS Conference, organized by CWI in Amsterdam.

Guest speakers:

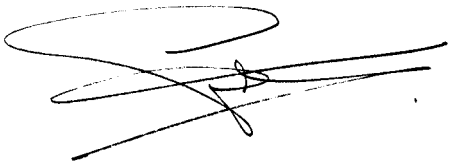
Edmund M. Clarke (Carnegie Mellon University)

Bjarne Stroustrup (AT&T Bell Labs)

Amsterdam. During this conference CWI organized a festive Soirée for the participants and the Dutch computer science community. This followed a similar event in 1996 on the occasion of the 50-year anniversary of the foundation SMC, under which CWI resorts. On that occasion Donald Knuth and Benoit Mandelbrot were the invited speakers. An audience of three hundred listened in March 1999 to Bjarne Stroustrup (AT&T Labs) and Edmund M. Clarke (Carnegie Mellon University). Stroustrup invented C++, the language most used for object-oriented programming. He spoke about 'What is an object and what isn't?', Clarke's specialty is the automatic verification of computer hardware and software. The title of his address was: 'Symbolic Model Checking without BDDs'.

In order to carry out the broad spectrum of activities, from which here, naturally, only an incomplete account is given, a dynamic, skilful and cooperative staff is required. There is no lack of that at CWI. Serious efforts are, however, required to realize the intended extension of activities. Frontier research – CWI's core business – remains of course the institute's main concern. Hence, extension of the scientific staff with excellent researchers is a necessity. Apart from the funding aspect a housing problem arises here: CWI is getting overcrowded. In order to deal with this problem, a start was made with plans for more working-space. Looking back on the past, and forward into the near future, I have the fullest confidence that CWI will succeed in realizing its ambitious plans, and will keep its research on world level.

Gerard van Oortmerssen



General Director CWI

Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights

Particle Systems and Correlation Inequalities

Research project: Probability
Project leader: J. van den Berg
E-mail: J.van.den.Berg@cwi.nl

Introduction

A more accurate, but too long, title would be 'Large systems of interacting objects, and the use of correlation inequalities'. 'Interacting particle systems' is a subfield of Probability Theory, introduced in the mid-seventies to model the behaviour of a huge collection of objects (called 'particles') each of which may be in different possible states and influence the states of other objects in its neighbourhood. Randomness is an important ingredient in these models. The motivation comes from Statistical Physics (derive macroscopic behaviour from microscopic rules), Chemistry, Biology (population dynamics/genetics, epidemics, spread of forest fires, etc.), and more recently also from Operations Research and Computer Science (several so-called probabilistic algorithms can be described in terms of interacting particle systems or, related, cellular automata). During the nineteen-eighties and -nineties this field has become one of the most active research areas in Probability Theory.

Typical questions involve the large-scale behaviour of the system. For instance, given the mechanism by which a certain type of infection is transmitted from individual to individual in a large population, can one (or a few) infected individuals, cause an epidemic?

Although the models are a simplification of reality, they lead to substantial insight in 'real-world phenomena' and give rise to the development of new mathematical tools which, in turn, can be used to handle more realistic versions. A small number of models has been developed which are considered 'prototypes' of certain phenomena. The philosophy (which we share) is that a complete understanding (in a rigorous mathematical sense) of these models higly clarifies the essence of the real phenomena.

A correlation inequality

First we introduce a correlation inequality which appears to be quite useful. Suppose we have a number of lamps and each lamp has a colour, say green, red or blue. Moreover, the colours are chosen randomly (independently). By a pattern (in fact, a better word would be 'subpattern') we mean a prescription of the colours of some of the lamps. Two patterns are called orthogonal if no lamp involved in the first pattern is also involved in the other. See Figures 1a and 1b. Now suppose we have patterns P_1, ..., P_k and Q_1, ..., Q_k and for each i, P_i is orthogonal to Q_i. It may be plausible, but is not at all easy to prove, that then the probability that there is an i for which both P_i and Q_i occur, is at most equal to the probability that at least one of the P's occurs times the probability that at least one of the Q's occurs. (Of course it is true if every P_i is orthogonal to every Q_j; in that case we even have equality.) This was

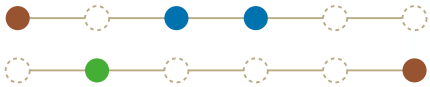


Fig.1a: Two orthogonal patterns for a collection of six lamps.

Fig.1b: A configuration in which both patterns of Fig.1a occur.

formulated as a conjecture in the eighties by Van den Berg and H. Kesten (Cornell University), who also proved a special case which appeared to be very useful in percolation theory (the study of long-range connectivity properties in a large grid with random defects). The general case was proved ten years later by D. Reimer (who received a George Polya prize for his nice proof).

The Contact Process

The Contact Process is one of the simplest spatial models of an epidemic. It is very easy to describe, yet the essential problems are hard. The progress made on this model has been very fruitful. One of the simplifications made in this model (and many others) is that space is not continuous but discrete: the objects (individuals) are vertices (points) of a 'nice' grid, for instance the cubic lattice or its 2-dimensional analog, the square lattice. Further, the interaction rules have been simplified. Each point has two possible states: healthy (but susceptible to infection) or ill (and infectious). An ill point recovers (becomes healthy) at rate 1 (we will explain below what this means).

The spread of forest fires, rumours, or epidemics can be modeled by random systems of interacting objects ('particles') – presently a very active branch of probability theory. Photo: Jonathan Blair. Courtesy: Foto Natura, Wormerveer.

Each healthy point becomes ill with a rate equal to α times the number of ill neighbours. Here α is a parameter of the model, called the infection parameter. The above formulation is for continuous time and has many technical advantages, but (for finite grids) the following, equivalent, discrete-time formulation is easier to explain (and is, in fact, essentially what is done in computer simulations). See Figure 2. Suppose we start at time 0 with a certain configuration of ill and healthy points. At each step in time we change the state of at most one point, and this is done as follows. Consider the set consisting of all points of the grid and, moreover, of all pairs of points (i, j) where i and j are neighbours. To each point we assign weight 1, and to each of the above pairs weight α. Now we simply choose randomly an element from this set,

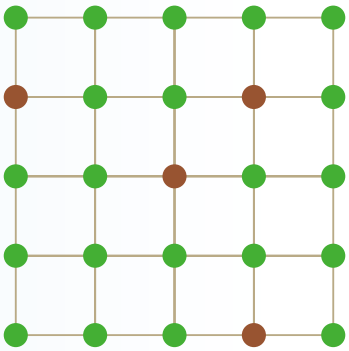


Fig.2a: Example of a configuration for the contact process on a finite 5x5 grid. denotes 'healthy' denotes 'ill'

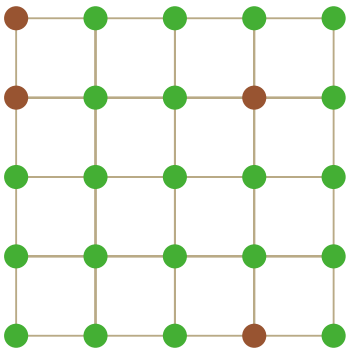


Fig.2b: Possible configuration two time steps later. Note that one of the ill individuals has recovered and one has infected a neighbour.

with probability proportional to its weight. So, if S denotes the sum of all the weights, the probability that a given point is chosen is $1/S$, and the probability that a given pair of neighbours is chosen is α/S. If the chosen item is a point, it becomes healthy (if it already was healthy, nothing changes at this step). If it is a pair of points, say (i, j), and i is ill, then j becomes ill (again, if it already was ill, nothing changes at this step). An equivalent, and visually very useful, way to describe the evolution is by a so-called space-time diagram in which on the time axis of each point i random 'recovery' symbols (on the average 1 per time unit) are put, and between the time axes of two neighbours (i, j) arrows are put randomly (on the average α per time unit). See Figure 3.

Figure 3: Contact process on a finite (size 5) one-dimensional 'grid'.

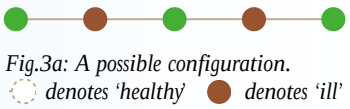


Fig.3a: A possible configuration. denotes 'healthy' denotes 'ill'

Fig.3c: The configuration at time T for the contact process starting at time 0 with the configuration in Fig.3a and evolving according to the diagram in Fig.3b.

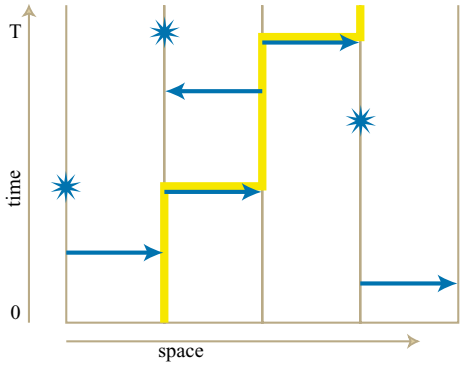


Fig.3b: Example of a space-time diagram. denotes an infection 'attempt'. denotes a recovery. The yellow trajectory is a (potential) infection path in space-time.

This description makes clear that the spread of the disease takes place via certain 'paths' in space-time. This suggests connections with percolation theory, which play indeed an important role.

A basic result, obtained in the early history of the subject, is that there is a critical value α_c (which depends on the dimension) of α such that above (but not below) α_c a single ill individual in an 'ocean' of healthy individuals can cause an epidemic. Moreover, it was proved later that if α is below its critical value, the disease disappears very (exponentially) fast from the population. The special case of the correlation

inequality above plays an important role in this result (by Bezuidenhout and Grimmett).

A few years ago we have, in cooperation with G. Grimmett from Cambridge (UK) and R. Schinazi, Colorado Springs (USA), considered a modification of the above model, in which a point can no longer directly go from the 'ill' to the 'healthy' state, but only via an extra state, the 'immune (but not infectious)' state. The rate at which the state changes from 'immune' to 'healthy' is denoted by β. Our research has led to new results for such systems with small β.

In the remainder we concentrate on a quite different particle system, called Coalescing Random Walks.

Coalescing Random Walks (CRW)

We assume that at time 0 each point of a d-dimensional cubic lattice is occupied by a particle. Now each of these particles is going to make a so-called random walk: with rate 1 it 'decides' to make a jump, and when it makes a jump it goes with equal probability to one of its $2d$ neighbours. As long as particles don't meet, they don't influence each other. However, when two particles meet, they coalesce, that is, become one new particle. (An equivalent description is that if a particle jumps to a point which is already occupied, it is deleted from the system.) For a discrete-time description see Figure 4.

Figure 4: Coalescing random walk on a finite 5x5 grid.

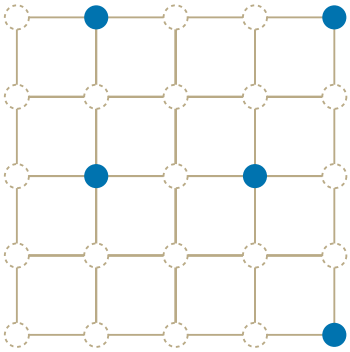


Fig.4a: Example of a configuration.
denotes 'occupied' denotes 'empty'

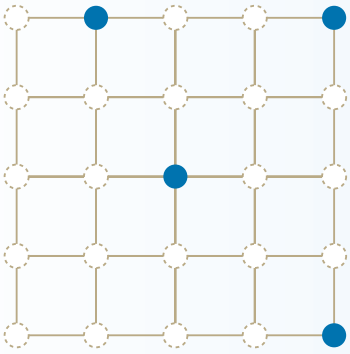


Fig.4b: Possible configuration two time steps later. Note that two particles have coalesced.

Like the contact process, it is convenient to describe this with the help of space-time diagrams. See Figure 5.

This is one in a collection of models motivated by chemical reactions. At time 0 every point is occupied: the particle density is 1. What is the particle density $p(t)$ at time t? It is not difficult to guess that it goes down and in the long run gets arbitrary close to 0. But how fast does it decrease? In the early eighties Bramson and Griffeath proved that, if d is at least 3, $p(t)$ behaves asymptotically like c/t, for some constant c (depending on d) which they could explicitly identify. They also had (different) results for d =1 and 2, but we restrict ourselves to higher dimensions. Their arguments make heavily use of a duality relation with the so-called voter model. In that model each point represents a person. At time 0 each person has a unique (different from the others) opinion about a certain case. At random times an individual randomly selects a

Figure 5: Coalescing random walk on a finite (size 5) one-dimensional 'grid'.

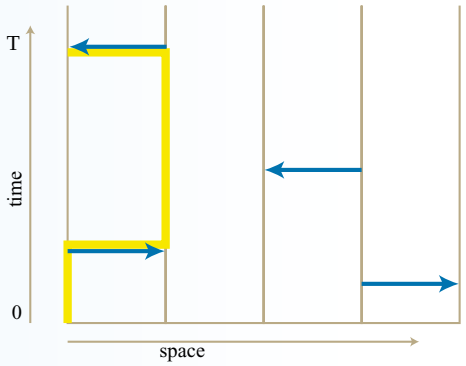


Fig.5a: Example of a space-time diagram.
denotes a jump 'attempt'.
The yellow trajectory describes the movement of the leftmost particle.

Fig.5b: The configuration at time T for the system, starting with every point occupied and evolving according to the diagram in Fig.5a.
denotes 'occupied' denotes 'empty'

neighbour and takes over his opinion. In this way opinions spread or may become distinct. The mentioned duality means that if we take a space-time diagram for the Coalescing Random Walk, reverse time and then use it for the voter model (where an arrow indicates a take-over of opinion), then the event: 'point i is occupied at time t' for the CRW corresponds with the event: 'the opinion which i had at time 0, is still present at time t somewhere in the system' for the voter model.

Modified Coalescing Random Walks

The model above is called the basic CRW model. A few years ago, we started to study, in cooperation with H. Kesten (Cornell University), the natural modification in which coalescence does not always take place. In our model the jumping particle disappears with probability $q(k)$, where k is the number of particles already present in the point he jumps to. We assume that $q(0) = 0 < q(1) \leq q(2) \leq \ldots$. Surprisingly, the key of the arguments of Bramson and Griffeath breaks down. There is no longer a suitable duality. Heuristic arguments suggested that for large t, $p(t)$ still behaves like a constant over t (but with a different constant from the one before) and new ideas were needed.

Before we go on, we revisit the basic CRW. We are interested in $p(t)$, the particle density at time t. Because of the mechanism of the process, $p(t)$ goes down because particles meet. Two particles meet when two adjacent points are both occupied and the particle located in one, jumps to the other. This fact leads easily to the differential equation $dp/dt = P$ (both O and e are occupied at time t). Here e is a given neighbour of O. This looks nice but is quite useless as long as we don't know much about the probability in the right hand side. If the events: 'O is occupied at time t' and 'e is occupied at time t' were independent, the right hand side would be equal to $p(t)^2$, and we would have a soluble differential equation, with solution $p(t) = 1/(1+t)$.
Note that this gives, in dimension ≥ 3 (but not in 2) the right order of magnitude (but the wrong constant). This very rough heuristic argument can be refined by looking at time $t - \Delta$ (for suitable Δ and arguing that if O and e are both occupied at time t, then there must have been particles at some points x and y at time $t - \Delta$ which moved to O and e respectively, *without* meeting each other. Now we can simply sum over all such possibilities, and since for most choices of x and y these points are far apart, it seems justified to do as if the events: 'x occupied at time $t - \Delta$' and 'y occupied

at time $t - \Delta$' are independent. Eventually this leads to a differential equation of the form $dp/dt = C.p(t)^2 + \mathrm{err}(t)$, where err($t$) is an error term which comes from the fact that we simply summed over all possibilities and from the independence assumption. Most of the work is to show that the error term is sufficiently small and the 'correct' asymptotic behaviour follows. The nice thing is that this works not only for the basic model but also for the more general version. The bad news was that we could control the errors only in dimension ≥ 5. Without going into details we remark that the above-mention-ed sketch suggests that it would be helpful if we could show that the probability of certain events is relatively small. In particular, that the probability that two points i and j are both occupied is at most a constant times $p(t)^2$. This was already known for the basic model, and can be derived in an alternative way by applying the correlation inequality mentioned before: consider the space-time diagram and apply the correlation inequality (where now patterns correspond with 'paths of arrows' which guarantee the occupation of x, respectively y; the orthogonality of the patterns is related to the fact that the particles moving to x and y should not meet). It is not clear how to apply this in general: usually it is not easy to describe the event that a point is occupied in terms of the space-time diagram. But it can be done, by a trick, for certain choices of the parameters q_i, by introducing colours to the particles, which randomly change when they jump. For those nice cases (which include the basic CRW) our error esti-mates work sufficiently well for $d \geq 3$ (as they should). For the other cases, many (but not all) of the estimates can be bounded by similar estimates for a suitably chosen nice case. In this way, the earlier mentioned requirement $d \geq 5$ could be replaced by $d \geq 4$ now. More work is needed to handle $d = 3$ for the general case.

Another current activity is an attempt to further generalize the correlation inequality.

Evolution, Neurons, and Agents in E-Commerce and ICT

Research project: Evolutionary Systems and Applied Algorithmics
Project leader: J.A. La Poutré
E-mail: Han.La.Poutre@cwi.nl
URL: http://www.cwi.nl/~hlp/

Introduction

Evolutionary algorithms, neural networks and applied algorithmics are important computer techniques to deal with everyday's life. Evolutionary and neural techniques are based on the biological principles of nature. They use natural principles to learn from and adapt to changing environments, and they can develop based on implicit feedback from their application settings. Various other results in applied algorithmics actually use significant knowledge about the problems at hand that need to be tackled.

Evolutionary algorithms, neural networks and applied algorithmics appear to have more and more impact for new application areas besides their current major application domains like mathematical optimization and automatic programming. An important area is that of adaptive agents. Agents can be real-world agents, like economic agents (insurance agents, merchants, companies, consumers) or social agents (people, groups, and societies), which are people or organisations in the real world. Lately, agents can also be software agents, being independent pieces of software that perform certain tasks for their owners and that can interact with other software of real-world agents. The latter field is related to agent technology.

Thus, new application areas fall on the one hand in the ICT domain, like e-commerce, design of intelligent agents, profiling, negotiation, and simulation of adaptive agent systems. On the other hand, social sciences are becoming new application fields. An example is economics, where the typically-addressed problems concern complex systems seen as agent-based systems (economic markets, grouping and collaboration of individuals, etc.). The combination of these areas with software agents yields new research in e-commerce, as an economic sub-area.

Natural computation has grown into a substantial research area in computer science. Neural networks have been studied intensively for more than two decades, while evolutionary algorithms are currently enjoying serious research efforts for more than 10 years. The types of applications that have been studied are especially optimization, classification, and regression problems.

Flower auction at Aalsmeer, The Netherlands. Auctions are examples of systems with competing bidders. Such systems can be studied with evolutionary methods for bidding strategies.
Courtesy: Capital Photos.

Research focus

The focus to problems arising from the emerging ICT society is, of course, a very recent one. The work is an extension of pioneering work that was mainly done in the USA, at institutes such as the Santa Fe Institute or the MIT Multimedia Lab. At the former institute, a basis for considering economic systems as evolutionary systems was laid, while in the latter, the role of agents in commerce was explored. Currently, attention on these fields is growing in magnitude in the economics and computer science disciplines, and in the e-commerce and AI disciplines respectively. In both fields, however, the development and usage of new, appropriate techniques from natural computation and applied algorithmics have become crucial to make the proper steps in their developments and for the actual application in human settings.

To give an idea, in order to make agents feasible to work for humans, they need to be able to learn several aspects of their interaction with humans or other agents. This is as opposed to most of the current solutions that require explicit settings of choice and strategies by humans. Since the latter is too difficult or time consuming to do on a large scale (which consumer can efficiently describe his relative preferences for a collection of ten options and accessories on a new car, in combination with his budget and service desires?), more flexible approaches using automated learning are needed. Likewise, in order to model real-world agents in a market, realistic learning and interaction techniques in simulation are needed.

The types of research that is needed in these areas concern applicable research as well as fundamental research. This is because for many of the anticipated application areas,

What is an evolutionary algorithm?

Evolutionary algorithms (EAs) are strongly inspired by the genetic evolution theory in biology, as developed by Darwin. EAs typically work as follows. First, a random 'population' (set) of possible candidate solutions for the problem at hand is generated. This population is subsequently changed and improved in a number of rounds ('generations') by means of evolutionary concepts. These concepts are 'survival of the fittest', selection, mutation, and recombination (like 'crossover'). This is repeated several times, like for 1000 generations. The final, typical result is a population with solutions for the problem that are as good as possible, and from which the (almost) best solutions can be selected.

A more precise description is as follows.

An evolutionary algorithm consists of the repeated generation of new populations, starting from a random population. A population consists of candidate solutions for the problem, where a candidate solution is e.g., a set of parameters that needs to be determined. This can be represented as a string of bits, resembling the structure of a chromosome in nature.

The generation of a new population consists of a number of phases.

- The first phase is the selection phase, where in a stochastic way, candidate solutions in the population are selected as 'parents' for the new population. This is done such that better candidate solutions have a better chance to be selected. This corresponds to the concept of 'survival of the fittest' in nature.

- Subsequently, parents are coupled in pairs. From each two parents, two new candidate solutions ('children') are generated by recombination as follows. One child consists of parts of (parameter) values of one parent and the remainder from the other parent; and the other way around for the other child.This correponds to the reproduction process in nature, where a child inherits genetic material from both parents and recombines this into new genetic material. In nature this is called cross-over. In this way a new population is formed. This part forms the recombination phase.

- Finally, the new candidate solutions are changed slightly in a random way ('mutation'), to enable diversity in the population. This phase is the mutation phase.

The above scheme is repeated until a number of generations has been created (e.g., 1000), or some termination criterium is satisfied.

only a limited amount of fundamental research has been performed as yet. Regarding the fundaments, research on the usage of evolutionary techniques is done with respect to computational feasibility and complexity of the phenomena that are studied in the application concerned. Also, and not less important, modeling and assessing the relation between computational techniques and the actual real-world mechanisms are of the utmost importance.

Interdisciplinary approach

For the actual application, cooperation with experts in the application fields is essential. In our current and future research we have strong interest in multi-agent systems, as they can occur in economics, electronic commerce, logistics and social systems. Several new projects (externally funded) fall into these areas. Examples are the large project 'Autonomous Systems of Trade Agents in E-Commerce', funded by the

Telematics Institute and with partners TNO, ING, and KPN; and the more fundamental project 'Evolutionary Exploration Systems for Electronic Markets', funded by NWO. These observations also lead to new cooperations, especially with experts in application and technology areas. Examples are economists, agent technology experts, (e-)commerce specialists, business researchers, and people in the field of multimedia and communication technology. Several of such cooperations and contacts currently exist for the research group.

In our research activities, several evolutionary systems were constructed and investigated for various types of development and simulation systems of interacting agents.

Negotiations

An evolutionary system, called EMINE, has been built for developing negotiation strategies for agents and for obtaining emergent properties of systems of negotiating agents. EMINE stands for Evolutionary model for Multi-Issue NEgotiations. In this system, two populations, one of buyer agents and one of seller agents, compete against each other in order to obtain the highest pay-off, using an alternating-offers negotiation protocol. The approach is on negotiations of deals with more than one issue: buyer agents have an other pay-off value for a specific combination of issue values than seller agents. This thus allows the development of win-win situations. The factors which have been studied are e.g., the presence of deadlines and time pressure (time discount). Regarding the evolutionary techniques, the influence of the different selection schemes was investigated.

Subsequently, we extended our negotiation models with social phenomena, in order to obtain more realistic outcomes in these evolutionary systems. An important social factor appears to be the notion of fairness: although mathematical game theory predicts extreme outcomes in cases where take-it-or-leave-it offers are possible, such extreme outcomes hardly occur in practice due to fairness norms that play a substantial role in daily life. In the research, we have introduced a fairness model based on results in experimental economics, and we show that this leads to much more realistic and stable outcomes of negotiations in our computational approach.

To validate the evolutionary approach with existing results of classical game theory, the fundamentals of evolutionary approaches on one issue have been investigated. The approach was done by Evolution Strategies, a form of evolutionary algorithms that allow real encodings and adaptive stochastic mutations. Several classical game theoretic results have been investigated and obtained by this approach, thus laying a proper and solid base for extension and elaboration of the research problems with evolutionary approaches. As we already mentioned above, mathematical game theory does not always predict realistic outcomes, which has mostly to do with the restrictedness of the modeling and the available mathematical techniques. This is one of the strong points of the evolutionary approach: it allows for more advanced modeling and black-box feedback, while still yielding good results. One example is that the total pay-off function does not need to be known to the system, but black-box evaluation of a deal is sufficient.

Agent interaction and cooperation

Research was performed on stereotyping and the formation of social groups through tag use in populations playing the iterated prisoners' dilemma. The role of tag was considered: how can the tagging (labeling) of agents help to form social groups, and what is the effect of evolutionary operators on this. Experiments were conducted using tag-mediated selection of opponents, mates, and strategies. The experiments showed that, under most circumstances, populations using tags reach a higher level of cooperation and do display more stable behaviour than populations that do not use tags. Furthermore, these effects were especially stabilized by the inclusion of recombination operators (sexual reproduction operators) in the evolutionary system. The latter operator, which is one of the standard evolutionary operators, has a common interpretation as the exchange of information, ideas, and behaviour amongst agents. This is shown to have a stabilizing effect in this situation.
Further, fundamental research was done on the influence of evolutionary mechanisms on systems of agents that interact via the prisoners' dilemma game. Amongst others, the effects of fixed-opponent evolution, ecological evolution, and co-evolution were investigated. Co-evolution together with ecological evolution yielded more robust cooperation strategies than the other approaches; i.e., strategies that arise from this approach perform better against (new) various types of agents that they are faced with afterwards. Also, the ecological evolution appears to be more realistic in its dynamics, and it was shown that it is significantly related to the applied selection scheme in the evolutionary algorithm. Thus, the proper choice of the parameters and operators in an evolutionary algorithm appears to be of much importance for these types of problems.

Agents and E-Commerce

The application of natural computation in e-commerce settings was at CWI in its first phase in 1999. Several new activities and projects started, and will be continued in the next years. Below we give an overview of some of the most important projects.

A start was made with research on electronic-market behaviour in the NWO project 'Evolutionary Exploration Systems for Electronic Markets'. A model with heterogeneous, boundedly-rational agents was implemented for an economic market of the Cournot duopoly type and first results were obtained. Various types of adaptive, 'learning agents were studied in competing with each other via the Cournot duopoly' market situation. It yielded insight in that both sophisticated agents and very simple agents (imitating the opponent's last move) appear to be strong players in such a market. The advantage of this approach is that actual behaviour types can be investigated in such systems, as opposed to the classical approach of studying (such) markets from the (mathematical) equilibrium point of view.

Also, a start was made with two projects in cooperation with industrial partners. In the field of e-commerce, the Trade Agents project ('Autonomous Systems of Trade Agents in E-Commerce') was initiated, a project funded by the Telematics Institute, which is carried out with project partners TNO, ING, and KPN. Several of the research activities described above were (also) part of this project, together with activities related to developing several business and pilot cases for the research and development activities in this project. The current cases address information brokerage agents for stock-related news, and the development of a shopping mall with flexible customer support and negotiation possiblities.

Solution of the Prisoners' Dilemma by a genetic algorithm

The Prisoners' Dilemma, is an elementary two-person symmetric game which provides a simple explanation for escalations like arms races and price wars. The name was given to this game by the American mathematician A.W. Tucker in 1950. Two suspects in crime, A and B, are held *incommunicado*. Each is given one of two choices: to maintain the pair's innocence or to sign a confession accusing the partner of committing the crime. Or, in other words: cooperate or betray. What is good for the prisoners as a pair – steadfast denial by both – is frustrated by their pursuit of individual rewards. In the game numerical pay-offs are assigned to all four possible outcomes, for example as in the Table (first score A, second score B).

Player	**B** Cooperates	**B** Betrays
A Cooperates	3 - 3	0 - 5
A Betrays	5 - 0	1 - 1

The graph below shows the solution of the game by a genetic algorithm, exhibiting five different phases. *Phase I* (initial phase, generation 0-20) starts with a population of arbitrary strategies. Here the best strategy is systematic obstruction (yielding 3 points on the average against 3/2 for cooperation). Consequently the number of traitors rapidly increases, leading to a decrease of the score: the more traitors, the more difficult to gain points. In *phase II* (evolution of cooperation, generation 20-50) the 'eye-for-an-eye' strategies emerge. Players start with cooperation and keep repeating the opponent's last move. Here a traitor never achieves lasting success. If, however, two 'eye-for-an-eye' players meet, they get a far better result (3 points) than two traitors (1 point). Hence, their number rapidly increases, as does the average score. In *phase III* (long cooperative state, generation 50-220), cooperation lasts very long, mainly because eye-for-an-eye strategies are very difficult to exploit by traitors. And yet, the traitors succeed in the long run, because all cooperative strategies perform equally well as eye-for-an-eye strategies. This includes so-called naive collaborators who always cooperate, even with traitors. Naive strategies slowly increase in number in this phase, because there are hardly any traitors. In *phase IV* (collapse of cooperation, generation 220-240) traitors emerge by a mutation in the genetic code. These traitors can efficiently exploit the naive players. The cooperative population 'collapses' under a large-scale invasion of traitors. If all exploitable players are extinct, only traitors and eye-for-an-eye strategies remain. Finally, in *phase V* (from generation 240) cooperation is restored.

Research Highlights

Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights

Factoring Large Numbers:
Fun or Applied Science?

Research project: Computational Number Theory and Data Security
Project leader: H.J.J.te Riele
E-mail: Herman.te.Riele@cwi.nl
URL: http://dbs.cwi.nl:8080/cwwwi/owa/cwwwi_print.projects?ID=12

Introduction

The ancient Greeks already realized that every integer can be written uniquely as a product of 'indivisibles larger than 1', that is, prime numbers. Finding this representation for a given number turns out to be easy for small numbers, but hard and tedious for large numbers. In fact, this implies two problems: the identification of prime numbers, and the decomposition of non-primes or composite numbers into primes.

Identification of a given, say, 200-digit number as a prime number is not too difficult, but finding its prime factors may be impossible if we are unlucky, even with the best known algorithms and computer technology. Researchers at CWI have studied, implemented and tested factoring algorithms during the last fifteen years, and used as many as possible, otherwise idle, computer cycles in order to factor larger and larger numbers. The invention in the late nineteen seventies by the MIT trio Rivest, Shamir, Adleman *of the so-called RSA cryptosystem, the security of which relies on the difficulty of factoring, has greatly enlarged the need to know the state-of-the-art of factoring at any time. In addition, it showed that factoring the largest possible numbers is not only a fascinating scientific challenge but also an indispensable activity for validating and confirming the security of RSA-based cryptosystems.*

Factoring and Cryptography

Suppose we want to factor a given composite number N. Compositeness implies that there is a prime $p \leq \sqrt{N}$ which divides N. There is a simple computational test, known as Fermat's test, by which we can decide whether a given number is composite.

The simplest algorithm to factor N is trial division by the primes 2, 3, 5,..., until success. This requires a number of trials at least equal to the number of primes $\leq p$, which is $\mathcal{O}(p/\log p)$ by the famous Prime Number Theorem. In the worst case this is close to $\sqrt{N}$. For this trial algorithm we need a list of the prime numbers up to p, or a sequence of numbers, like $6k \pm 1$, $k = 1, 2,...$, of which the primes (except 2 and 3) are a subset. The latter sequence is much easier to generate than the sequence of primes, and does not need any storage space at all.

Fermat (of Fermat's Last Theorem!) devised a factoring method, called the *difference of square method*, which is the forerunner of the fastest modern factoring methods. Fermat's method looks for two integers a and b such that $a^2 - b^2 = N$, i.e., $(a - b)(a + b) = N$ so that we have a on-trivial factorization of N unless $a = b + 1$.

This need not be a decomposition of N into *primes*, but the factors are at least smaller than N, so we have reduced the problem of factoring N to a problem of factoring a number which is smaller than N. An extension of Fermat's idea is to try to find two integers a and b such that $a^2 - b^2$ is a *multiple* of N, say $a^2 - b^2 = kN$, for which Gauss introduced the notation:

$$a^2 \equiv b^2 \bmod N. \qquad (1)$$

Because of the presence of the factor k, $a - b$ need not be a divisor of N, but instead we compute the *greatest common divisor* of $a - b$ and N, and hope to find a factor of N in this way.

For example, let N =1003 and assume that we have found (don't bother about *how*) that $55^2 \equiv 4^2 \bmod 1003$, i.e., that $55^2 = 3025$ gives a remainder 16 upon division by 1003. Then, using the relation: $\gcd(a,b) = \gcd(b, a \bmod b)$ we compute the greatest common divisor of 1003 and $55 - 4 = 51$ as follows:

$$\gcd(1003, 51) = \gcd(51, 34) = \gcd(34, 17) = \gcd(17, 0) = 17,$$

so that 17 must be a divisor of 1003. The quotient 1003/17 = 59 is a prime number, which completes the factorization of N = 1003.

The *Quadratic Sieve method* (QS)generates *many* relations of the form

$$a_i^2 \equiv b_i \bmod N, \qquad (2)$$

where b_i is not a square, but a *smooth* number, i.e., a number with only *small* prime factors. If sufficiently many of such relations have been collected, they are combined, using linear algebra methods, into a relation of the form (1). QS uses many quadratic polynomials to generate the relations (2). With the *Number Field Sieve* method (NFS) it is possible to use *higher* degree (≥3) polynomials to generate relations of the form (2), at a bigger speed than QS.

The time needed to factor a large number with help of QS or NFS depends on the size N. For example, the computational effort to factor a number of 165 digits with NFS is about 3.5 times the effort needed to factor a number of 155 decimal digits. The computational effort to factor a number of 155 decimal digits (the current world record) was about 8400 MIPS years (one MIPS year is the equivalent of a computation during one full year at a sustained speed of one Million Instructions Per Second). For QS, this growth factor is 5.5.

There is another important class of methods of which the time is not determined by the size of the number N which we want to factor, but by the size of the *smallest prime factor* p of N. The simplest such method is trial-division with the primes 2, 3, 5, 7,.... Another, faster, method is Pollard's so-called $p - 1$ method. This tries to find a multiple M of $p - 1$ where p is a (unknown) prime factor of N. Now an arbitrary number a is raised to the power M modulo N. If M happens to be a multiple of $p - 1$, then a theorem of Euler guarantees that $a^M \equiv 1 \bmod p$, and we can compute $\gcd(a^M - 1, N)$ to find p. In fact, Pollard's $p - 1$ method tries to compute a unit in the group Z_p^* consisting of the residue classes modulo p (excluding 0). The order of Z_p^*, i.e., its number of elements, equals $p - 1$ (since its elements are 1, 2,..., $p - 1$) and this is crucial for finding a unit in this group. The time to find a factor p of N with Pollard's $p - 1$ method is proportional to $\sqrt{p}$ whereas for the trial-division method this is proportional to p.

While Pollard's $p - 1$ method works in just one group, the Elliptic Curve Method (ECM) of H.W. Lenstra, Jr. works in *various* groups, which are determined by integer points on elliptic curves. Again, the method tries to find a unit in each of these groups which all have order in the interval

$$[p + 1 - p^{1/2}, p + 1 + p^{1/2}].$$

The chance of success with ECM therefore is much bigger than with Pollard's $p - 1$ method.

In the practical factorization of large numbers of, say 40–150 decimal digits, one first applies a trial method to quickly remove possible small prime factors. Pollard's $p - 1$ and some other methods (like the so-called Pollard–ρ method, and a method of Shanks) are used to trace larger factors. Next, Lenstra's ECM method is applied to detect prime factors in the range of 20–45 decimal digits. If after some time, to be determined experimentally, ECM is still unsuccessful, the big guns QS and NFS are fired at N.

In the nineteen seventies, Rivest, Shamir and Adleman have developed a public-key cryptosystem called RSA whose security depends on the supposed intractability of factoring the large numbers which are used as public-keys in this cryptosystem. Although it has never been *proved* that factoring is hard for RSA-keys, i.e., requires an amount of computational work which grows exponentially with N, practical experience has never given any evidence of the contrary. RSA is widely used today and the optimal size of an RSA public-key depends on the security needs of the user and on how long the information needs to be protected.

The American company RSA Securities Inc., founded by the inventors of the RSA method, issued a so-called RSA *Challenge List* of numbers and offers a reward to anyone factoring a number from this list. In this way, RSA hopes to stimulate researchers to push forward the state-of-the-art of factoring, thus creating a continuous process of validation and confirmation of the security of the RSA cryptosystem. RSA is widely used today, e.g., to protect E-commerce on the Internet, in smart cards, in SSL (Secure Socket Layer) handshake protocols and in PGP (Pretty Good Privacy) software for e-mail protection.

Factoring at CWI

CWI's factoring research has started with the advent, in the nineteen eighties, of vector computers. Much time and effort has been spent on the efficient implementation of QS on large vector computers like the CDC Cyber 205, the NEC SX-2, and the Cray Y-MP and Cray C90 vector computers. Two 'factors' have favoured this approach: firstly, the bulk of the computational work in QS (and in NFS) consists of adding fixed quantities to numbers in a large array at positions which lie in an arithmetic progression, so this work is extremely suitable for vectorization; secondly, CWI has always had excellent facilities for access to large vector computers through the NCF (National Computing Facilities Foundation).

Finding relations of the form (2) can be split up easily into independent tasks, so this part of factoring is well-suitable for parallelization. The advent of workstations on each researcher's desk and the availability of an abundance of low-priority CPU-time on these computers shifted the emphasis of the computational work in this project from vector computers to workstations and, later, to fast PCs.

Factoring RSA-512 drew world-wide attention in the media. During the press conference at CWI on 26 August, 1999, from left to right: Eric Verheul (PricewaterhouseCoopers, Utrecht, The Netherlands), Andrew Odlyzko (AT&T Labs-Research, USA), Herman te Riele (CWI), Gerard van Oortmerssen (CWI), Arjen Lenstra (Citibank, New York, USA), Paul Leyland (Microsoft Research, Cambridge, UK).

However, for one computational step in factoring large numbers, computers like the Cray C90 are still indispensable, namely for the step in which dependencies have to be found in a huge sparse bit matrix. This step requires an extreme amount of shared main memory which until now is only available on machines like the Cray C90.

In the past five years, CWI's main activity in factoring was devoted to the improvement of the Number Field Sieve factoring method and its efficient implementation. Two PhD theses have been written on QS (Boender, 1997) respectively NFS (Elkenbracht-Huizing, 1997) and a third, on NFS, is being prepared (Cavallar).

The CWI Computational Number Theory group has established, or contributed to the establishment of, various new factoring world records. Most factored numbers were contributions to the so-called Cunningham Table and to an extension of this table. Several factorizations contributed to the proof of the non-existence of odd perfect numbers below 10^{300}.

For a survey of this and other work on computational number theory at CWI, see the Special Issue on Computational Number Theory of *CWI Quarterly*, Vol. 7, Number 4, December 1994.

Recent factoring records

In 1999, three factoring world records were established by an international group called 'The Cabal', coordinated by CWI 's 'Computational number theory and data security' group.

The first record was established with the Special Number Field Sieve (SNFS): a 211-digit so-called Cunningham number which, due to its special form, is easier to factor than 'general' numbers. The other two records were records for the General Number Field Sieve (GNFS): a 140-digit and a 155-digit RSA-key.

The factoring world record for SNFS was established on April 8 by the computation of the factors of the 211-digit Cunningham number $N = (10^{211} – 1)/9$ into two primes of 93 and 118 digits, respectively. This achievement also established a record for the largest penultimate prime factor (of 93 digits) ever found.

The two factoring world records for GNFS were established on February 2 (RSA140), and on August 22 (RSA155), respectively. RSA140 turned out to be the product of two primes of 70 digits each, and RSA155 the product of two primes of 78 digits each. For details on these records, see under `ftp://ftp.cwi.nl/pub/herman/NFSrecords/` the files `SNFS-211, RSA-140, and RSA-155.`

Both RSA140 and RSA155 are representative for the public keys which are used in the RSA public-key cryptosystem. Written in binary notation, RSA155 is a 512-bit number. A well-attended press conference was organized at CWI on August 26 where the factorization of RSA155 was announced. The implications for E-commerce and cryptography were discussed with four international experts (A.K. Lenstra, P. Leyland, A.M. Odlyzko, and E. Verheul). This event raised much publicity in the national and international press, and on radio and television.

Some technical information about the three factoring world records is summarized here:

number	10, 211–	RSA140	RSA155
size of the two factors (in decimal digits)	93, 118	70, 70	78, 78
sieving time (in CPU years)	10.9	8.9	35.7
calender time for sieving (in days)	64	30	110
# workstations and PCs used for sieving	125	125	300
matrix size	4.8M	4.7M	6.7M
row weight	49	32	62
Cray CPU hours for matrix step	121	100	224

Factoring expert Richard Brent from the Oxford University Computing Laboratory has given an experimental formula which ‘predicts’ the year Y in which we may expect a number of D decimal digits to be factored:

$$Y = 13 .24\, D^{1/3} + 1928.6.$$

In the derivation of this formula, Brent took into account the known factoring world record data, the known heuristic complexity formula for NFS and Moore’s law (computing power doubles every 18 months). According to this formula, a 768-bit number ($D = 231$) will be factored by the year 2010, and a 1024-bit number ($D = 309$) by the year 2018.

Research Highlights

Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights
Research Highlights

Quantum Computing

Research project: Quantum Computing
Project leaders: P.M.B. Vitányi, H.M. Buhrman
E-mail: Paul.Vitanyi@cwi.nl, Harry.Buhrman@cwi.nl
URL: http://dbs.cwi.nl:8080/cwwwi/owa/cwwwi.print_projects?ID=34

Moore's law

Computers increasingly pervade our society. This increasing influence is enabled by their ever increasing power, which has roughly doubled every 18 months for the last half-century (Moore's law). The increase in power, in turn, is primarily due to the continuing miniaturization of the elements of which computers are made, resulting in more and more elementary gates with higher and higher clock pulse per unit of silicon, accompanied by less and less energy dissipation per elementary computing event. Roughly, a linear increase in clock speed is accompanied by square increase in elements per silicon unit – so if all elements compute all of the time, then the dissipated energy per time unit rises cubicly (linear times square) in absence of energy decrease per elementary event. The continuing dramatic decrease in dissipated energy per elementary event is what has made Moore's law possible.

The future of computing: classical or quantum?

However, there is a foreseeable end to this: there is a minimum quantum of energy dissipation associated with elementary events. This puts a fundamental limit on how far we can go with miniaturization, or does it? It turns out that only irreversible elementary events (like erasing information) by the laws of thermodynamics necessarily dissipate energy; there is no physics law that requires reversible events (like negation) to dissipate energy. But so far the development of computation machinery is mostly based on the principles of classical physics and irreversible components. At the basic level, however, matter is governed by quantum mechanics, which is reversible. Further miniaturization will very soon reach scales where quantum mechanical effects take over and classical laws cease to apply accurately. The mismatch of computing organization and reality will express itself in friction: computers will generate gigantic (megawatts) of energy unless their mode of operation becomes quantum mechanical (and thus reversible). That is, harnessing quantum mechanical effects is essential for further miniaturization and hence acceleration of classical computing methods.

There is an added bonus: once we get involved in quantum effects, it appears we can go further than just miniaturizing classical computers to the quantum scale. Quantum mechanics may actually spawn a *qualitatively new* kind of computing: a kind which profits from quantum effects to boost computation to such an extent that things are achieved that would forever be out of reach of classical computers, even if these could be miniaturized to the same level.

The area of quantum computing has a great economical and societal potential. The field was born in the early 1980s through work of Richard Feynman, Paul Benioff, and David Deutsch, but only gained momentum after Shor's 1994 quantum factoring algorithm [5]. In 1995, CWI was one of the first in Europe and the first in The Netherlands to be involved in quantum computing research, and has since contributed significant discoveries to the field. Here we give a brief survey of quantum computing, its main successes so far, and some of CWI's contributions.

Superposition and interference

Suppose we have a computer with n bits of memory. From the point of view of common sense and classical physics, this computer can be in one and only one of 2^n states at the same time, namely one of the 2^n possible memory contents. A probabilistic computer (which makes random choices) or a quantum computer, however, can be in *all* of these 2^n states *simultaneously*. In the case of a probabilistic computer in a closed box (making random choices unknown to us), the state of the computer can at every point in time be described as a *superposition* or linear combination of all classical states, every state thereof having some probability of being observed if we look inside the box. It is known and commonly employed in practice that randomization can speed up some computations considerably, at the cost of obtaining the outcome with high probability rather than with certainty. This caveat is no problem if the probability is high enough – as it is made to be in practice. The next improvement can be obtained by computing quantum mechanically rather than probabilistically. The difference is contained in the fact that quantum probability has properties different from classical probability. In some sense it allows both negative and positive probabilities and has different addition properties – a feature called 'interference'.
As an example, consider the case $n = 1$. A classical computer with just one bit of memory can be in just one of two states, which we denote by $|0\rangle$ and $|1\rangle$. A quantum computer can be in a superposition of those states:

$$\alpha_0|0\rangle+\alpha_1|1\rangle.$$

Such a 2-state quantum system is called a *qubit*. The complex number α_i is called the *amplitude* of state $|i\rangle$. A quantum state can be changed in two ways. If we measure it, we will see classical state $|i\rangle$ with probability $|\alpha_i|^2$, and the quantum superposition will disappear. Since the squared amplitudes induce a probability distribution, we must have $\sum_i |\alpha_i|^2=1$.
Apart from measuring, we can also change the state by applying a *unitary transformation* to its vector of amplitudes. This is a special kind of reversible linear operation.
A simple example of a unitary transformation is the following:

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

If we apply U to the state $|0\rangle$ (which is the vector $(1\ 0)^T$) we get the state $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$. If we measure this we will see $|0\rangle$ or $|1\rangle$, each with probability 1/2, just like a random coin flip. We also get a random coin flip if we apply U to $|1\rangle$ and measure. However, if we apply U to $\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)$ and measure, then instead of getting another coin flip we get $|0\rangle$ with certainty. This is due to the fact that the positive and negative amplitudes

for $|1\rangle$ cancel out. Such *interference* is similar to light-waves canceling each other out.

If A is a classical algorithm for computing some function f, possibly even irreversible like $f(x) = x \pmod 2$, then we can turn it into a unitary transformation which maps classical state $|x,0\rangle$ to $|x,f(x)\rangle$. Note that we can apply A to a superposition of all 2^n inputs:

$$A\left(\frac{1}{\sqrt{2^n}}\sum_x |x,0\rangle\right) = \frac{1}{\sqrt{2}}\sum_x |x, f(x)\rangle.$$

In some sense this state contains the results of computing f for *all* possible inputs x, but we have only applied A once to obtain it. This effect shows one of the advantages of quantum over classical computing and is called *quantum parallelism*.

A quantum algorithm corresponds to a unitary transformation U that is built up from elementary unitary transformations, which each only act on one or two qubits. The algorithm applies U to an initial classical state containing the input and then makes a final measurement to extract the output from the final quantum state. The algorithm is 'efficient' if the number of elementary operations is 'small', which usually means at most polynomial in the length of the input. Quantum computers can do everything a classical computer can do – and more.

What quantum computers can do

The biggest success so far – and the event which ignited the current explosive growth of the field of quantum computing – was Peter Shor's 1994 discovery of an efficient quantum algorithm for finding the prime factors (factoring) of large integers [5]. By making clever use of superpositions, interference, quantum parallelism, and some classical number theory, Shor's algorithm finds a factor of a number N in time roughly the square of the length of the input (which is log N bits). In contrast, every known classical algorithm requires exponential time to factor. Since factoring is one of the most elementary aspects of number theory, the oldest mathematical discipline, and centuries of efforts by the greatest mathematicians have not yielded better methods, it is widely believed that such better methods either do not exist or are prohibitively difficult to find. In fact, this belief underlies most of current public-key cryptography, notably the RSA system, ubiquitously used on the Internet and in the financial world. Such crypto-systems can be broken if one can factor large numbers fast. Accordingly, the advent of quantum computing compromises all such systems: if a quantum computer can be built, then most of current cryptography becomes totally insecure, and, for example, electronic money can be forged.
What quantum computing takes away with one hand (classical public-key crypto), it gives back in another form with the other (quantum secret-key crypto) albeit only in part. Already in 1984, Bennett and Brassard found a scheme which allowed two distant parties to obtain a shared secret key via quantum mechanical communication [2]. Their scheme was always believed to be fully secure against any type of spy or eavesdropper, and recently this has indeed been formally proven. On the other hand, some other parts of electronic transactions, like unforgeable signatures, appear to be beyond the power of quantum methods.

A third application is Grover's 1996 algorithm for searching databases [4]. Consider finding some specific record in a large unordered database of N items. Classically, there is no smarter method than just to go through all records sequentially, which requires

expected $N/2$ time steps for a record in general position. Grover's algorithm, however, uses quantum superpositions to examine all records 'at the same time', and finds the desired record in roughly $\sqrt{N}$ steps. Examining a 10^{12} records with unit microsecond probes, this is the difference between about two months of computing and one second of computing! His algorithm also allows to solve the widespread and notoriously hard NP-complete problems (such as the traveling salesman problem) quadratically faster than known classical methods – reducing say exponential time with exponent N to exponential time with exponent $N/2$.

A fourth application was initially conceived and primarily developed in collaboration with Richard Cleve of Calgary University. It deals with the setting where two separated parties, Alice and Bob, want to compute some function $f(x,y)$ depending on x (only known to Alice) and y (only known to Bob). A simple scheme would be for Alice to send her x to Bob and then let Bob do all the work by himself, but this may take a lot of bits of communication and often there are much more clever schemes requiring less communication. The field of *communication complexity* examines the optimal number of bits that have to be communicated in order to compute the function at hand. What happens if we generalize this setting to the quantum world and allow Alice and Bob the use of quantum computers and qubit-communication? It turns out that some tasks can be solved with significantly less communication if we allow such quantization [3] – and this despite the fact that quantum bits cannot contain more information than classical bits! We have obtained similar advantages by sticking to classical communication, but allowing Alice and Bob the use of pre-established 'entangled' qubits. Both approaches beat the limits provable for just classical communication.

The above developments suggested the vision that *all* computation can be enormously speeded up by quantum computers. But not so! CWI's researchers obtained strong and general *limitations* of quantum computers as well. Recall that Grover's algorithm is quadratically faster than classical search algorithms. It was already known that such a quadratic speed-up is the best quantum computers can achieve for searching a database, so exponential speed-ups cannot be obtained for this problem. CWI researchers (in collaboration with others) recently showed that the same holds for *all* problems in the database-setting of Grover's algorithm: for all such problems, quantum computers can be at most polynomially faster than classical computers [1]. In the last year, CWI's researchers partially extended this lower bound approach to quantum communication complexity as well, but they also exhibited some new (polynomial) speed-ups, both in Grover's setting and in communication complexity.

Limiting results like the above, of course, do not preclude exponential speed-ups in different settings, like Shor's, or a clever future setting as yet unknown. Exploring this potential of quantum computation remains an exciting and important task for computer scientists and physicists alike.

How quantum computers do it

The above results are very promising, but so far mostly theory. How about actually building quantum computers which can run the fast algorithms like Shor's, Grover's, or CWI's? To date only very small quantum algorithms (and slightly bigger quantum crypto devices) have been implemented, but the physical realization of quantum computers is still in its infancy. The main problem is that quantum superpositions are extremely vulnerable and any interaction with its environment will quickly cause errors, which degrade the performance of the computer. Quantum versions of error-

correcting codes have been developed recently which to a large extent solve this problem in theory, but not yet in the brittle practice of the physical lab (let alone the brittle practice of our desktops). This is related to the development of Quantum Information Theory – the quantum extension of classical information theory. CWI has contributed to this research, and to related notions of the information in individual quantum states: Quantum Kolmogorov Complexity. Actually building large quantum computers presents formidable problems to experimental physicists reminiscent of the initial barriers to classical computing: unreliable components, physically large components, memory, organization, communication, programming. In fact, the theory of quantum mechanics is currently extended, partially by CWI research, in particular with respect to the algebraic analysis of 'quantum entanglement' – a vital notion in many quantum algorithms, apparently not yet thoroughly investigated in quantum theory. Similarly, Markov chain theory is extended into the quantum world, [6].

References

1. R. Beals, H. Burhman, R. Cleve, M. Mosca, R. de Wolf (1998).
Quantum lower bound by polynomials. *Proceedings of 39th IEEE FOCS*, 352–361.

2. C.H. Bennett, G. Brassard, A.K. Ekert (1992).
Quantum cryptography. *Scientific American*, 267(4), 50–57.

3. H. Burhman, R. Cleve, A. Wigderson (1998).
Quantum vs. classical communication and computation.
Proceedings of 30th ACM STOC, 63–68.

4. L.K. Grover (1996).
A fast quantum mechanical algorithm for database search.
Proceedings of 28th ACM STOC, 212–219.

5. P. W. Shor (1997).
Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5), 1484–1509.
First version in FOCS'94.

6. B.M. Terhal (1999). *Quantum Algorithms and Quantum Entanglement*,
PhD Thesis, University of Amsterdam.

ORGANIZATION

Research

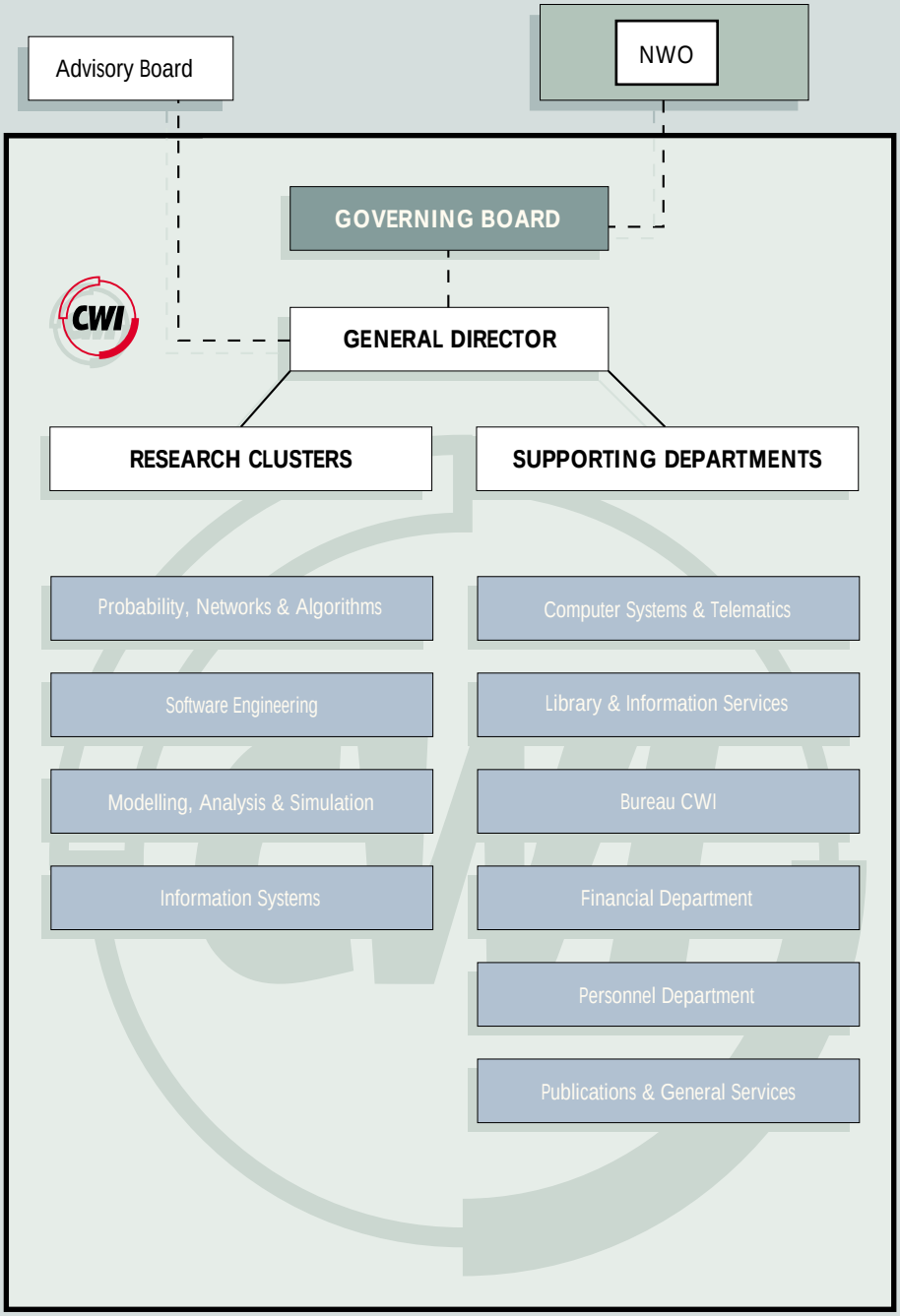
Cluster - Theme	Cluster leader Theme leader
Probability, Networks and Algorithms	**A. Schrijver**
- Networks and Logic – Optimization & Programming	A.H.M. Gerards
- Traffic and Communication – Performance & Control	J.H. van Schuppen
- Stochastics	J. van den Berg
- Signals and Images	H.J.A.M. Heijmans
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 Systems and Applied Algorithmics	J.A. La Poutré
Modelling, Analysis and Simulation	**C.J. van Duijn**
- Applied Analysis and Scientific Computing for PDEs	J.G. Verwer
- Computational Fluid Dynamics	P.W. Hemker
- Mathematics of Finance	M.S. Keane
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)
D.G.C. Broekhuis (controller)

Governing Board
L.A.A.M. Coolen (director KPN Research), chairman
P.M.G. Apers (University of Twente)
F.A. van der Duyn Schouten (Catholic University Brabant)
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)

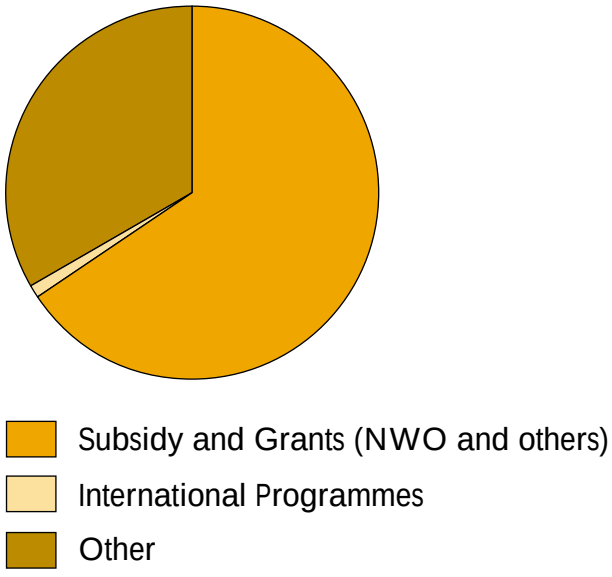

Advisory Board
NWO
GOVERNING BOARD
CWI
GENERAL DIRECTOR
RESEARCH CLUSTERS
SUPPORTING DEPARTMENTS
Probability, Networks & Algorithms
Software Engineering
Modelling, Analysis & Simulation
Information Systems
Computer Systems & Telematics
Library & Information Services
Bureau CWI
Financial Department
Personnel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

FINANCES, PERSONNEL

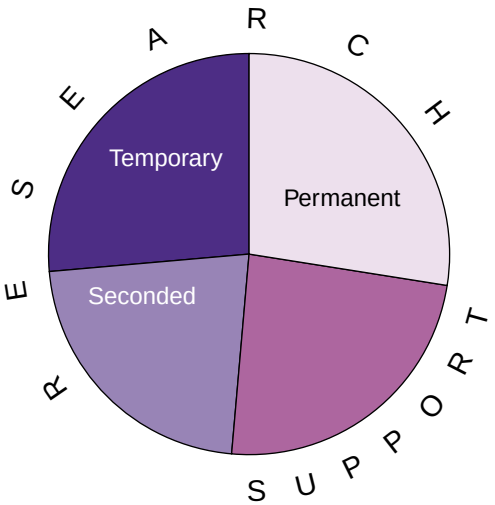
Finances 1999

In 1999, SMC spent Dfl. 32,25 million. The expenses were covered by a subsidy from NWO (Dfl. 19,75 million), other subsidies and grants (Dfl. 1,02 million), and from the international programmes (Dfl. 0,37 million). Finally, an amount of Dfl. 10,49 million (of which 2,76 million from the Telematics Institute) was obtained as revenues out of third-party-services and other sources.

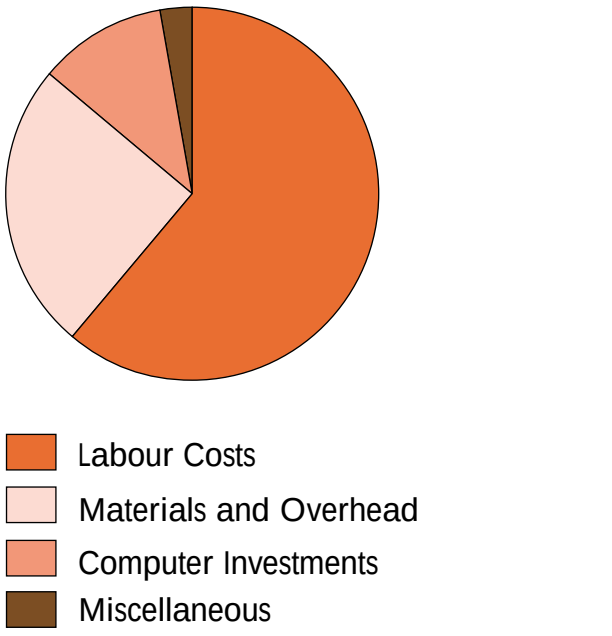
Income CWI

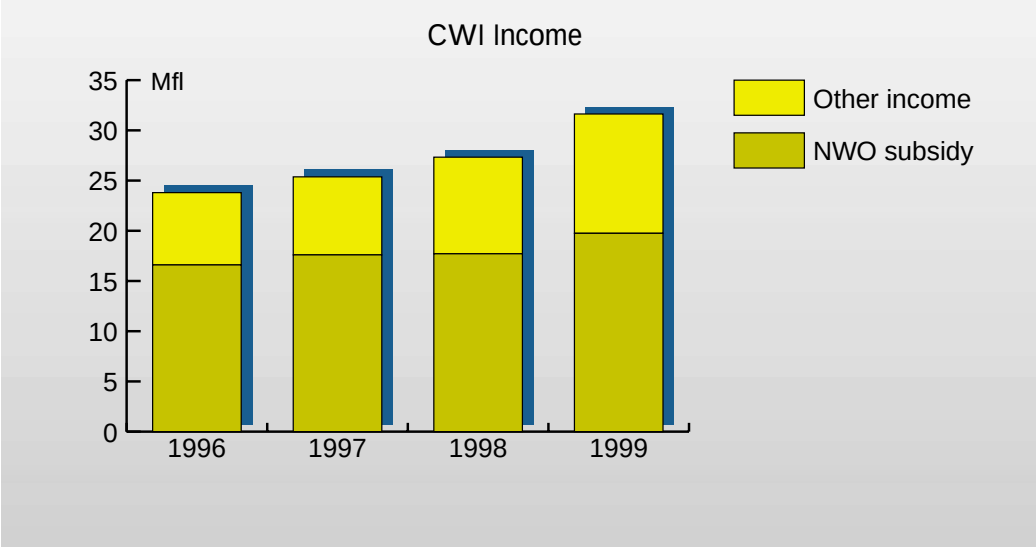


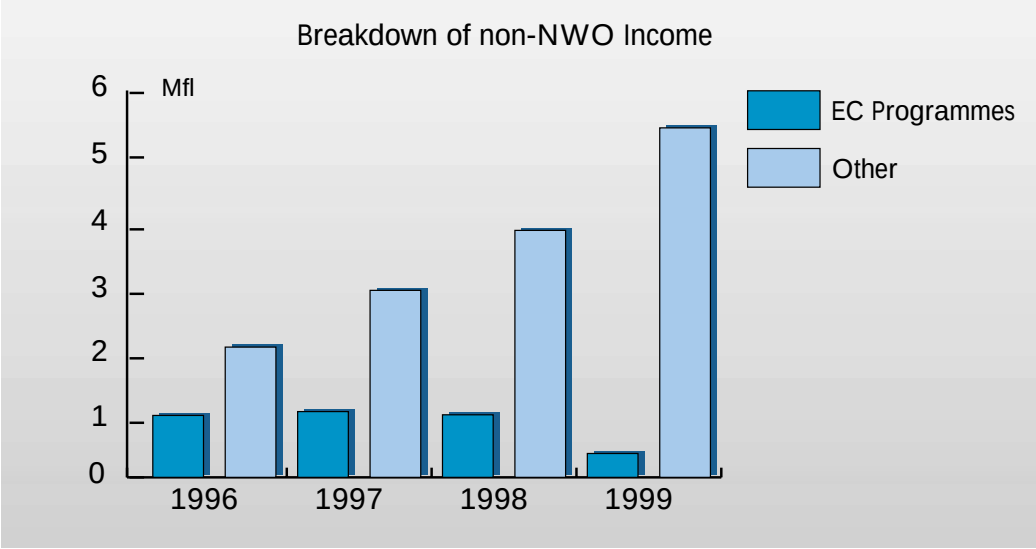
CWI Personnel:
164 fte + 47 fte seconded

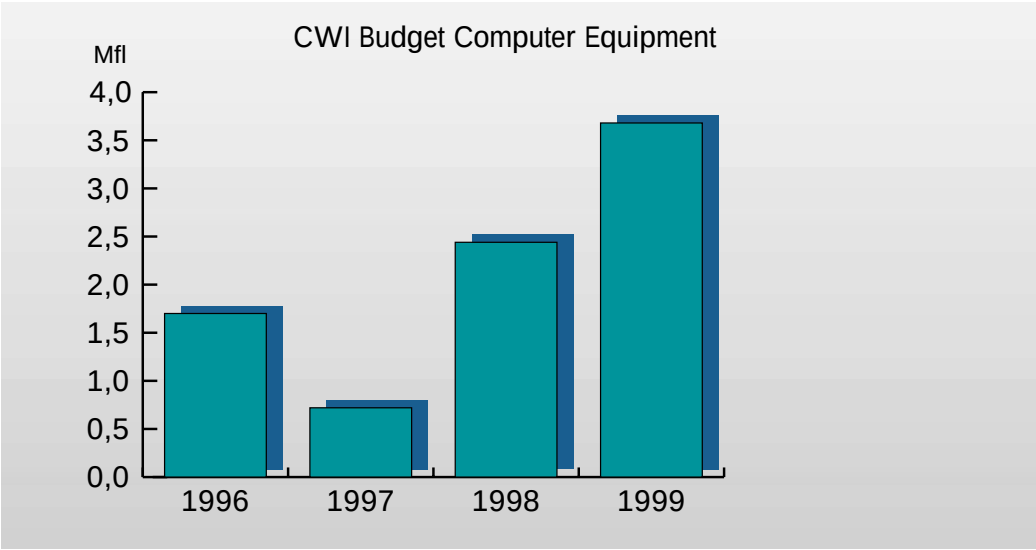


Expenses CWI




CWI Income
Mfl
35
30
25
20
15
10
5
0
1996
1997
1998
1999
Other income
NWO subsidy


Breakdown of non-NWO Income
Mfl
6
5
4
3
2
1
0
1996
1997
1998
1999
EC Programmes
Other


CWI Budget Computer Equipment
Mfl
4,0
3,5
3,0
2,5
2,0
1,5
1,0
0,5
0,0
1996
1997
1998
1999

CWI PhD THESES

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

S.A. Brands
Rethinking Public Key Infrastructures and Digital Certificates – Building in Privacy
Adi Shamir (Weizmann Institute of Science, Israel)

R.H.M.C. Kelleners
Constraints in Object-Oriented Graphics
D.K. Hammer, E.H.L. Aarts (both Eindhoven University of Technology)

C.H.M. van Kemenade
Recombinative Evolutionary Search
J.N. Kok (University of Leiden)

M.A. Reniers
Message Sequence Chart: Syntax and Semantics
J.C.M. Baeten, L.M.G. Feijs (both Eindhoven University of Technology)

J.M.T. Romijn
Analysing Industrial Protocols with Formal Methods
H. Brinksma (University of Twente), F.W. Vaandrager (University of Nijmegen)

B.M. Terhal
Quantum Algorithms and Quantum Entanglement
P.M.B. Vitányi

J.P. Warners
Nonlinear Approaches to Satisfiability Problems
J.F. Groote, J. van Leeuwen (University of Utrecht)

D.P. White
Percolation through Fractals, Backbends and Dynamic Lily Ponds
R.W.J. Meester (University of Utrecht), M.S. Keane

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, methods, and algorithms for studying congestion phenomena in communication networks, and of concepts and theory for network control by discrete event and stochastic methods. Research in performance focuses on service integration in communication networks.

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. The focus is on control of piecewise-linear hybrid systems, system theory of hybrid systems, and decentralized control of discrete-event 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. The focus is on computer algebra algorithms for identification and control problems for linear systems, the approximation problem for stationary Gaussian systems, and stochastic realization.

Stochastics

Theme leader: J. van den Berg

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, on mathematical models of biological and physical processes with self-organized critical behaviour, and on random spatial processes and elliptically invariant probabilistic measures, with applications to robotic control and disk scheduling.

Statistics

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

Stochastic analysis

Fundamental and applied research, in particular statistics for random processes with applications to mathematical finance, statistical methods for dynamical stochastic models, and statistical inference for stochastic processes and fractal analysis of financial data.

Ergodic theory and dynamical systems

Fundamental and applied research, in particular rigid sets, classification of Bernoulli schemes, fractal analysis, and superexponential convergence.

Signals and Images

Theme leader: H.J.A.M. Heijmans

Content-based coding, indexing, and retrieval

Image coding, feature extraction, and content-based image indexing and retrieval, in particular spatial grouping, multiresolution approaches, integration and interpretation, and applications.

Wavelets and morphology

Fundamental research on wavelets and mathematical morphology and its applications, with special attention to the construction of nonlinear multiresolution decompositions, in particular the design of morphological pyramids and wavelets.

Stochastic geometry

Parameter estimation for random sets, spatial statistics, image segmentation with appli-

cations to agriculture, and optimization problems for spatial processes.

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.

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, analysis and testing

The study of specification, analysis and testing techniques for computer controlled systems, by developing and implementing algorithms for the analysis and verification of processes with the μCRL toolset. Techniques and algorithms are as- sessed via case studies (communication protocols, embedded systems, hybrid systems, etc.), and the fundamental theory of processes is developed further.

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 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. Continued study of coalgebraic logic, which generalizes modal logic, the definitional method of behavioural differential equations, as well as coinduction and the control of discrete event systems.

Evolutionary Systems and Applied Algorithmics

Theme leader: J.A. La Poutré

Evolutionary systems

Study of evolutionary systems in management, economics and e-commerce, including: economic and commercial strategies, complex adaptive systems, adaptive agents, e-commerce, negotiation and trade, bounded rationality, interaction games, automatic programming, information filtering, optimization, dynamization, autonomous systems of trade agents in e-commerce, and evolutionary exploration systems for electronic markets.

Neural networks

Classification of data by several types of neural networks, concerning, e.g., benchmark classification problems, scaling, remote sensing, filtering, event prediction and decision support.

Discrete algorithms

Design of efficient algorithms for on-line optimization problems underlying various management and design problems in computer systems and net-works, 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

Applied Analysis and Scientific Computing for PDEs

Theme leader: J.G. Verwer

Atmospheric flow and transport problems

Numerical solution of time-dependent PDE problems from atmospheric circulation and air pollution modelling, addressing issues like linear and nonlinear advection, tailored stiff chemistry solvers, operator and time splitting, approximate matrix factorization, local refinement and sparse grid methods. More in particular: development of numerical algorithms and parallelization of the global atmospheric dispersion model TM3, numerical atmospheric circulation modelling (spatial discretization and time integration of the shallow-water equations in spherical coordinates), and advection-diffusion-reaction problems from air quality modelling.

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, and asymptotics and special functions.

Applications from the life sciences

Mathematical modelling and numerical simulation for life sciences, in particular biology and medicine. Current projects concern: mathematical modelling of biochemical processes in living cells (in particular reaction-diffusion equations describing molecular crowding), numerical methods for mixed parabolic-gradient systems occurring in models for the development of neuronal connections in the nervous system, numerical study of partial integro-differential equations modelling the growth of phytoplankton, dynamic modelling of the exchange of solutes and particles between biofilms and (drinking) water, and studying aggregated cell movement as a Hele-Shaw type problem.

Pattern formation and low temperature plasmas

Study of nonlinear dynamics, in particular the patterns of nonequilibrium plasma formed by applying a strong electric field to ionizable matter like gases or semiconductors. Patterns created by the interplay of field and ionization are studied with analytical PDE methods. Present research focuses on basic analytical questions in pattern forma-

tion and nonlinear dynamics, and on transient and stationary or periodically driven ionization patterns.

Computational Fluid Dynamics
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 numerical methods for the computation of free-surface flows, overset grid techniques for convection dominated problems, sparse-grid algorithms for 3D flow problems, and parallel solution of very large eigenvalue problems.

Computational number theory and data security

Application of new mathematical and computational techniques for the solution of problems in number theory with impact on cryptography. 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. In addition, classical conjectures like the Riemann hypothesis and the Goldbach conjecture are studied.

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.

Mathematics of Finance
Pilot leader: M.S. Keane

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.

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 focused 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. Current research concerns Bayesian Networks, mining time-series databases, mining multi-relational databases, and mining for groups with distinct behaviour.

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. At present research focuses on Hypermedia Authoring on Demand, Structured Document Languages for Hypermedia, and Hypermedia in Virtual Reality.

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 visualization

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.

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 communication complexity, quantum complexity classes, 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. Moreover, basic mathematical requirements for performance guaranties of evolutionary programs.

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.

INS-Cluster

D.J.N. van Eijck, M. Hazewinkel

Research on applied logic concerns dynamic logic, construction of electronic textbooks for logic, and interactive information engineering.
Research on digital libraries in mathematics focuses on the construction of thesauri for Knowledge Engineering, Combinatorics, and Linear Algebra.

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

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

WG CASL: Development of a Common Algebraic Specification Language
1999–2000
LORIA (Nancy), BRICS (Aarhus), ENS (Cachan), U. Bremen
M.G.J. van den Brand

IST -- Information Society Technologies

TRIAL Solution
2000–2003
U. Koblenz-Landau, Heidelberger Akad. Wissenschaften, Trinity College Dublin, U. Nice-Sophia Antipolis, FIZ Karlsruhe, Ges. f. Wiss.-Techn. Information, Open University (UK), TU Chemnitz, U. Köln, Springer-Verlag, Harri Deutsch, Shang IT
M. Hazewinkel

QAIP – Quantum Algorithms and Information Processing
2000–2002
U. Latvia, U. Oxford, U. Bristol, U. Aarhus, Lab. de Recherche en Informatique (LRI), Hebrew U. Jerusalem, Weizmann Inst. Technion, Israel Inst. Technology, U. Waterloo, McGill U.
H.M. Buhrman

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
1998–2002
U. Leuven, London School of Economics and Political Sciences, U. Pierre et Marie Curie (Paris), Rheinische U. Bonn, CNR, U. Lisbon, Société Coopérative ALMA, DASH Associates Ltd, 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
1997–2000
Moscow Institute for Numerical Economy, Moscow State U., U. Twente, U. Salerno
M. Hazewinkel

ERETIMA: English-Russian enriched Thesaursus in Mathematics
1997–2000
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

Numerical analysis of local and global bifurcations in ordinary differential equations
1999–2001
U. Gent, Russian Acad. Sciences, U. Nizhny Novgorod
M. Hazewinkel

Bilingual English-Russian thesaurus in mathematics
1999–2002
Russian Acad. Sciences, Steklov Inst. of Mathematics, Yaroslav State U., U. Utrecht
M. Hazewinkel

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

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

RTN

AMORE – Algorithmic Methods for Optimizing the Railways in Europe
2000–2004
U. Konstanz, ETH Zürich, IT-DTU Lyngby, CTI Patras, DIS-DIE Rome, L'Aquila Italy
A.M.H. Gerards

DYNSTOCH – Statistical Methods for Dynamical Stochastic Models
2000–2004
Universities of Copenhagen, Amsterdam, Berlin, Cartagena, Freiburg, Helsinki, London, Padua, Paris
K.O. Dzhaparidze, P.J.C. Spreij

Co-operation with GMD

Numerical atmospheric circulation modeling
1997–2001
J.G. Verwer

Sparse grids and overlapping grids in LiSS
1998–2001
B. Koren

Application of techniques from propositional logic for the verification of processes
1998–2002
U. Delft
J.F. Groote

Distributed Collaborative Virtual Environments
1999–2002
R. van Liere

Mining for groups with distinct behaviour
2000–2004
A.P.J.M. Siebes, M.L. Kersten

National Programmes

NWO Council for the Sciences

Dutch-Hungarian cooperation
1995–2000
OTKA
J. van den Berg

Parallel declarative programming
1996–1999
K.R. Apt

Fractal image coding
1996–2000
U. Delft

H.J.A.M. Heijmans
Convolutions on the motion group of the plane
1996–2000
Cornell . U, 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

Reinforced random walk in the plane
1996–2000
U. Amsterdam
M.S. Keane

Discontinuous dynamical systems – modeling of hybrid systems
1996–2000
Universities of Groningen, Twente, Eindhoven and Brabant
M.S. Keane

LT – Performance analysis of communication networks (long-tailed traffic phenomena)
1996–2002
U. Eindhoven, Columbia U., Lucent Technologies
S.C. Borst

A modular toolset for μCRL
1997–2000
U. Utrecht, U. Eindhoven, U. Amsterdam, U. Nijmegen, U. Groningen, U. Twente, Philips
W.J. Fokkink

Protocols, reference models and interaction schemes for multimedia environments
1997–2000
U. Utrecht
W.J. Fokkink

Dynamic algorithms for on-line optimization
1997–2000
Philips Research
J.A. La Poutré

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

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

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

CIP – Constraint and Integer Programming techniques
1997–2002
Partners in ERCIM WG on Constraints
K.R. Apt

PERS – Parameter Estimation for Random Sets
1997–2004
Eurandom, U. Utrecht, U. Berkeley
H.J.A.M. Heijmans

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

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

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
1998–2000
U. Leiden
H.J.J. te Riele

Sparse grid methods for time-dependent PDE problems
1998–2001
UU/IMAU, RIVM, KNMI, TNO, U. Iowa
J.G. Verwer, B. Koren

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

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

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

LRS – Large Random Systems and combinatorial probability
1998–2000
Cornell U., Hungarian Acad. Sciences, U. Utrecht, U. Delft, U. Leuven, U. Cambridge, Chalmers U. Gothenburg, U. Rome
J. van den Berg

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

TM3 – Parallelization and development of numerical algorithms for the global atmospheric dispersion model
1998–2002
U. Utrecht/IMAU, KNMI, CKO
J.G. Verwer

SPP – Overset grids and Singularly Perturbed Problems
1998–2002
IMM Ekaterinenburg, U. Nijmegen, U. Amsterdam, U. Dresden
B. Koren

GenTrans – Generation of Program Transformation Systems
1999–2000
U. Bergen, U. Utrecht
J. Heering

Spectral aspects of struggle for light
1999–2000
U. Amsterdam
J.G. Verwer, B.P. Sommeijer

Component based framework for constraint solving
1999–2001
K.R. Apt

Distributed imperative constraint programming
2000–2002
Russian AI Research Inst. Moscow, Inst. of Informatics Systems Novosibirsk
K.R. Apt

TokeN2000: Interaction between humans and information systems
1999–2003
Universities of Eindhoven, Maastricht, Delft, Leiden, Nijmegen, and Rijksmuseum Amsterdam

SICA – System Identification with Computer Algebra
1999–2003
Free U. Amsterdam, U. Eindhoven, INRIA Sophia Antipolis, UCAM
J.H. van Schuppen

WA – Wavelets and their Applications
1999–2003
U. Groningen, U. Eindhoven, U. Twente
H.J.A.M. Heijmans

Evolutionary exploration systems for electronic markets
1999–2003
U. Amsterdam
J.A. La Poutré

MRA – Multi-Resolution Approaches
1999–2004
U. Delft
H.J.A.M. Heijmans

Numerical singular perturbation problems (network)
2000–2003
U. Nijmegen, MGU Moscow, Russian Acad. Sciences, POMI St Petersburg
P.W. Hemker

Semi-automatic hypermedia presentation generation
2000–2004
H.L. Hardman

Special NWO projects

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

OPSS – Optimization Problems for Spatial Processes
1999–2000
U. Glasgow
M.N.M.. van Lieshout

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

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

SPINOZA – Logic in action
1997–2002
ILLC
D.J.N. van Eijck

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

XRAY – Preprocessing of seismic data: wavelet X-ray transform
1996–2000
Shell, U. Eindhoven, U. Delft, MARIN, U. Groningen
H.J.A.M. Heijmans

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

Development of a state-of-the-art Navier-Stokes solver for water flows
1999–2002
MARIN
B. Koren

Formal design, tooling and prototype implementation of a real-time distributed shared dataspace
2000–2003
Signaal
J.F. Groote

Improving the quality of embedded systems by formal design and systematic testing
2000–2003
Weidmüller
J.F. Groote

SENTER

RTIPA – Real Time Internet Platform Architectures
1999–2001
Philips, Oratrix, U. Eindhoven, EOLRING Int., France Telecom, GIP RENATER, Hitachi, INRIA, Italtel SpA, LIP6, Politecnico di Milano, Siemens AG, Telebit, Thomson-CSF
H.L. Hardman

NCF

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

Parallelization and optimization of software for very large eigenvalue problems
1998–1999
U. Utrecht
B. Koren

Parallel, distributed-memory implementation of existing sparse-grid software for 3D fluid-flow computations
1998–1999
B. Koren

Parallel adaptive mesh refinement for computational magneto-fluid dynamics
1999–2000
B. Koren

ICES HPCN Programme

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

ICES-KIS Programme

Distributed virtual reality for cell biology
2000–2001
R. van Liere

Molecular crowding – mathematical modeling of biochemical processes in living cells
1999–2003
U. Amsterdam
M.A. Peletier, J.G. Blom

MIA – Multimedia Information and Analysis
1999–2003
U. Amsterdam
M.L. Kersten

Telematica Instituut

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

Autonomous systems of trade agents in E-commerce
1999–2003
TI, TNO, ING, KPN, IBM, Bolesian
J.A. La Poutré

Summer – Secure Multimedia Retrieval
2000–2002
KPN, Min. Public Works, V2-lab Foundation, U. Twente, Object Design
Nederland
M.L. Kersten