

A a n v r a a g
Intentioneel Apparaat Schema
(IAS)



1988 - 1993

Centrum voor Wiskunde en Informatica



Aanvraag Intentioneel Apparatuur Schema (IAS)

1988 – 1993

Centrum voor Wiskunde en Informatica

Appendices

november 1987

- Appendix 1: Analysis and reconstruction of images**
- Appendix 2: Symbolische berekeningen in artificial intelligence**
- Appendix 3.1: The Amoeba project**
- Appendix 3.2: Een multi-computer faciliteit t.b.v. het database onderzoek 1990-1995**
- Appendix 4: Computergrafiek en gebruikersinterfaces**
- Appendix 5: Parallele numerieke algoritmen**
- Appendix 6: Automatisering bibliotheek**

Analysis and reconstruction of images

Abstract. All phases of image reconstruction and image processing (compression, enhancement and restoration, description and analysis) pose challenging problems which lie at the interface of mathematics, computer science and engineering. The CWI is embarking on a major and long-term research programme to study the mathematical aspects of these problems. Although the research will not be directed at computer science or engineering aspects, it will be necessary to have a flexible and powerful capability for experimental image analysis built around a network of workstations.

1. Characterization of the research area.

Computer image processing is conveniently and conventionally described as having phases of:

compression

enhancement and restoration(noise removal)

reconstruction (back transformation in e.g. tomography)

and last but not least

description and analysis (measurements, classification, decisions)

The research at CWI will concentrate on the last two topics: *reconstruction* and *description and analysis*. However the various phases of image processing can sometimes be profitably combined or exchanged, and in fact the whole process is circular since a good description leads to better compression techniques. Topics covered under *reconstruction* will be: reconstruction methods for moving magnetic resonance (NMR) images; inverse scattering of seismic images. Under *description and analysis* the research effort will focus on: the theory and application of cellular automata and mathematical morphology; multiresolution methods; probabilistic modelling and statistical analysis. We discuss the proposed CWI research in the various areas, which cannot always be clearly separated.

a. *Reconstruction*

Advances in computer aided tomography in recent years have been spectacular; for instance the Nobel prize for medicine went in 1979 to A.M. Cormach and G.N. Hounsfield for their work on computerized tomography (CT) (also called computer-assisted tomography: CAT). There are many techniques (emission and transmission tomography; NMR, X-ray, gamma-ray, ultrasound, electron beam, ...) all however presenting a similar mathematical problem: how to reconstruct an n -dimensional object given (some of) its $n-1$ (or less) dimensional projections. (Here n is usually 2 or 3, but can also be 4 if time plays a role too). A mathematical solution in the form of an inversion formula was given in 1917 by Radon. However this hardly ever supplies a practical solution since these reconstruction problems are 'ill-posed' in Hadamard's sense, and many are even quite horrifically ill-posed: a small perturbation in the given data leads to a totally different result with the inversion formula. Numerical methods to solve these problems fall into two classes: algebraic reconstruction techniques and transformation techniques (harmonic analysis, integral equations, special

functions).

Especially challenging as well as mathematically difficult problems arise in the reconstruction of *moving* objects. e.g. imaging a moving human heart by the NMR technique. The development of truly dynamic reconstruction algorithms is a wide open field.

Among the industrial applications more and more importance is gained by *seismic data and image processing* with the aim to reconstruct the internal structure of the earth by inverse scattering techniques. Because of the noise and the limitations of the measuring process statistical and system-theoretic considerations must be taken into account as well.

b. *Mathematical morphology & cellular automata*

An important development within the field of image processing is formed by mathematical morphology, which is both a mathematical theory and a source of practical image processing algorithms, which are part of most currently available image processing software packages. Within mathematical morphology images are considered as sets of points; logical relations between neighboring points are used to define so-called morphological transformations with a clear geometrical interpretation. In this way a method is obtained to perform global quantitative measurements on images by the successive application of several morphological transformations. In practice one is again confronted with noise and other uncertainties, so that probabilistic and statistical ingredients (random sets, random fields) are being incorporated in the theory. There is an intimate connection between mathematical morphology and *cellular automata*, which is called cellular logic image processing (CLIP). Generations of ever more advanced image processing machines are based upon this principle (Texture Analyser, CLIP1-CLIP4, DAP: distributed array processor, MPP: massively parallel processor).

A related area is that of *multiresolution image processing*, which considers simultaneous levels of resolution in an image, inspired by theories of how the human vision system functions. The question of image decomposition which arises here is a natural one to be studied in the context of mathematical morphology.

c. *Probabilistic modelling & statistical analysis*

The probabilistic approach has led to revolutionary advances even in areas where a 'statistical approach' does not seem at all applicable. For instance, in *restoration*, though a (cleaned up) LANDSAT picture of a familiar part of the Netherlands may not *globally* look at all like the result of some kind of homogeneous stochastic process, and in any case one is hardly interested in how the Netherlands would have looked in 'any possible world' than the real one, it is in fact the case that *locally* small parts of the image do look very much like (small parts of) the realization of a *Markov random field*. This fact has in the last five years been exploited by D. and S. Geman (Brown University, USA), by J. Besag (Durham, UK), P. Switzer (Stanford Univ.) and others, who have devised vastly superior image restoration algorithms based on an assumed stochastic model for the 'underlying scene', and in particular its local statistical spatial dependence, together with a probabilistic model for the noise which added to the scene produces the available picture. The image restoration problem can now be reformulated as a (Bayesian) *statistical estimation problem*. Novel techniques have been developed to actually efficiently solve the resulting optimization problem. In very recent work the approach has been applied to segmentation problems (line detection) using Markov line processes, and to inverse (tomography) reconstruction problems. As soon as the noise component of such images is more significant and has special structure - e.g. in positron emission tomography one is dealing with Poisson distributed counts, not Gaussian noise - a statistical approach can lead to far better reconstruction and/or enhancement algorithms.

However the truly *statistical* part of the project is aimed at exploring a vast and challenging area: developing a statistical approach to the *description* and *analysis* of images. By this we mean solving typical statistical problems - e.g. what is the effect of treatment with epidermal growth factor on the spatial distribution of receptor molecules in the cell membrane? what is the difference in the texture of a fractured surface of a new composite material manufactured under different conditions and how is this related to its physical properties? ... - when the data consists not of a series of numerical measurements but of a few digitized images from e.g. electronphotomicroscopy, ultrasound tomography, or whatever. Our strategy is to seek out some stimulating case-studies and to work in close collaboration with applied field scientists with real statistical questions. A statistical approach is especially called for when a small number of pictures have been painstakingly obtained and need to be thoroughly and exhaustively analysed, and when even then a *certain* conclusion will not be possible. Rather, the uncertainty in the best possible conclusion needs to be quantified (is the difference statistically significant? what is the standard error of the estimate?).

Typical problems confront the statistician with *one* very highly dimensional observation: a digital image could consist of 1 million numerical values, with very high low range statistical dependence and often large long range dependence too. Even given a model and an estimator, standard statistical techniques (requiring large numbers of independent observations) do not exist to give an assessment of its accuracy. However even before this stage, we are confronted by extreme numerical complexity of the problem of producing a best estimate; before this still, by a lack of models for describing anything but the most simple spatial processes.

The time is definitely ripe for an explosion of advances in this area. Ingredients will come from such topics as

- * *stochastic geometry* (the theory of random geometrical objects), which will provide us with models (e.g. random tessellation) and also answers to problems of *digitalization* (what happens when we intersect a fixed object with a random 'observation' grid).
- * *mathematical morphology*, a toolkit of image transformations which will produce summary statistics and diagnostic plots for evaluating models; also it is intimately connected with the theory of random sets and random fields.
- * *Markov fields, Gibbs states*, as mentioned before in connection with the work of Geman & Geman, Besag, and others; important problems are to assess the propriety of using such *scale dependent* (through pixel lattice mesh size) and *anisotropic* (through the orientation of the pixel lattice) models for essentially scale invariant, isotropic phenomena.
- * *spatial statistics*, an already well established field of statistics especially for the analysis of agricultural field trials, geophysical and ecological surveys, generalizing classical time series analysis to the analysis of measurements taken at an irregular collection of points in \mathbb{R}^2 or \mathbb{R}^3 .
- * *fractal theory*, another rich source of stochastic models for e.g. textures.
- * *stereology*, the mathematics of making inference from sampled 2-D sections to a 3-D object.

2. Innovative effect of the required apparatus.

The envisaged computer system (hard- and software) must form an experimental laboratory for image processing, allowing evaluation of new algorithms, simulation and modelling experiments, and general statistical and numerical analysis of results of image processing. The ideal environment for such activities is a network of workstations (one for each researcher on the project) allowing high resolution graphical display and a large memory storage (a typical image consists of a megabyte of data and one will often want to manipulate several images at the same time). Without this experimental

capability the research would be strictly 'academic': theoretical and at a high level of abstraction. However with the capability the research will be effectively aimed and there will be direct feedback on the usefulness of new approaches. Also it will make meaningful international cooperation with similar research groups elsewhere possible (below we mention groups in the USA, UK, Denmark, France and Norway who have similar configurations and with whom we already have close contacts; a definite advantage would be the possibility of exchanging software). In short, the apparatus is a necessary condition for the success of the project.

Apart from the innovative effect of the apparatus on the research into the mathematical aspects of image processing itself, we can expect a strongly stimulating effect on CWI activities in statistical data analysis, research into computational statistics, and mathematical modelling and analysis.

3. Relation with research programme.

The image analysis and reconstruction project at CWI has been set up within and between the departments of Applied Mathematics and Mathematical Statistics. Here there is the required background expertise in such technical ingredients as harmonic analysis, integral equations, special functions, numerical linear algebra, stochastic process theory, and statistical analysis with high-dimensional parameter spaces. Image analysis is a strongly multi-disciplinary subject and we are therefore at a unique position at CWI to also use the expertise of numerical mathematicians, system theorists and computer scientists. In particular we mention the interactive graphics group at CWI, the department of operations research and systems theory which is participating in the project on seismic reconstruction and the department of numerical mathematics which is interested in parallel computing.

Image analysis lies at the interface of mathematics, computer science, and engineering. The last few years have seen, internationally, exciting advances in the mathematics of image analysis. Sharing these in the Netherlands, and contributing to future progress, is a task at the heart of CWI's research programme.

4. Relevance for government and society.

Image processing has applications in such diverse fields as radio-astronomy, remote-sensing (satellite pictures, weather and harvest prediction), radar and sonar imaging, medical imaging (ultrasound echo scans, CAT scans, NMR scans), robotics, manufacturing, computer vision, The mathematical aspects of image processing are evident, exciting and challenging. In short, image processing is a vital part of modern high technology in the fundamental aspects of which the Netherlands must participate.

5. Position of the research.

This is a recently started research project which has just concluded its initial phases of inventarization and survey. We already have held extremely successful national colloquia and seminars, and have established serious research contacts with (among others) Shell Research Labs (seismic imaging), Philips Medical Systems (NMR imaging of moving objects), ITI-TNO, AKZO Research (electronphotomicroscopy images of surfaces of a composite material), and Organon. We also have established relationships with similar research groups internationally (at the Universities of Aarhus, Bath, Durham, North Carolina, and Oslo; at CSIRO in Australia, and with the École des Mines in Paris). Finally we have contacts with the research groups in the Netherlands (already) working in image processing, especially the Delft Image Processing Centre where the accent is on the complementary engineering and computer-science aspects.

6. Cooperation with other research areas.

As mentioned under point 5., a number of interesting research relationships with 'applied field' researchers have already been built up. Here we can also mention contacts with the department of ophthalmic surgery at the University of Nijmegen (ultrasound imaging with very high noise levels), the Hubrecht laboratory for cell-microbiology (electronphotomicroscopy images of the cell membrane), and the KNMI (remote sensing for weather maps). This type of contact will at a later stage lead to more applied work being done in the context of STW, Esprit, or NATO projects. We note that Shell has already donated a research fellowship to the CWI for this project and that a staff-member of ITI-TNO is presently seconded to the project. There are also research contacts with researchers in stochastics in the Netherlands, in particular in stochastic geometry and random fields.

7. Criteria for the desired apparatus.

The criteria below are set up with the aim that at medium term (within five years) a research group of about six senior researchers will be active at CWI, together with various programmers, junior researchers (oio's), stagiaires, industry-visitors, foreign visiting scholars, etc. The research will concentrate on two closely related areas, but will eventually require two slightly different sub-configurations. Image input will be mainly done by applied field co-researchers with their own special input devices; at CWI only some simple apparatus for digitizing e.g. monochrome photographs will be necessary. Hardcopy output is very important however, and also network connections (national and international).

7.1. Very short term. 2 large workstations with floating point accelerator and large disc storage; video input (frame grabber); software (TIPS, HIPS, VISIX, VISILOG).

7.2. Middle term. Ethernet network of workstations, 6 smaller workstations, 50 Mbyte disc per node, software; perhaps a permanent (network) line to some applied field co-researchers.

7.3. Long term. The *reconstruction* sub-project will need a more powerful system to handle the large data sets occurring in medical or seismic reconstruction problems. The same holds if one wants to study 3-D or 4-D images. Here one will have to think of parallel processors. For output one will need a good frame buffer, e.g. a PIXAR machine.



Symbolische berekeningen in Artificial Intelligence

P. Klint
P.J.F. Lucas

Centrum voor Wiskunde en Informatica
Kruislaan 413
Amsterdam

Voorgesteld wordt gespecialiseerde apparatuur aan te schaffen ten bate van onderzoek waarin symbolische berekeningen een primaire rol spelen. Hierbij wordt vooral gedacht aan onderzoek op het gebied van expertsystemen, kennisrepresentatie, knowledge engineering en programmageneratie. Hier worden de globale apparatuurbehoeften voor dit onderzoek geschetst in de context van IAS.

1. Karakterisering van het onderzoeksgebied

Het belang van het gebruik van computers en programmatuur bij numerieke berekeningen is in de afgelopen tientallen jaren ruimschoots gebleken. De laatste jaren neemt in diverse gebieden van de informatica de belangstelling voor symbolische berekeningen snel toe. Enkele van deze gebieden zijn artificial intelligence (expertsystemen, automatisch redeneren met diverse vormen van kennisrepresentatie, knowledge engineering), programmaverificatie (dataflow-analyse, automatische controle van correctheidsbewijzen) en programmaspecificatie (prototyping, analyse van de consistentie van specificaties). De ontwikkeling van programmatuur in deze gebieden vereist een combinatie van nieuwe computerarchitecturen, krachtige software-omgevingen en een performance die aanzienlijk beter is dan die van traditionele computerapparatuur. In deze aanvraag wordt derhalve voorgesteld dergelijke apparatuur ten bate van het bovengenoemd onderzoek, waarin symbolische berekeningen een primaire rol spelen, aan te schaffen.

2. Vernieuwend effect van de gevraagde apparatuur

De gevraagde apparatuur zal ertoe bijdragen dat de voorzieningen voor het beoogde onderzoek een orde van grootte toenemen in vergelijking tot de mogelijkheden die de reguliere ZWO-financiering toestaat. Hiervan is een hoge output te verwachten die gebaseerd zal zijn op omvangrijke experimenten op het gebied van:

- o Object-georiënteerde implementatietechnieken van expertsystemen;
- o Knowledge-engineering omgevingen voor de constructie

van expertsystemen;

- o Prototyping van qua omvang realistische problemen;
- o Ontwikkeling van nieuwe technieken ten bate van programma-generatoren.

De output van deze experimenten bestaat uit wetenschappelijke publikaties en ontwikkelde programmatuur. Voor het uitvoeren van deze experimenten is de gevraagde apparatuur essentieel.

3. Relatie met het onderzoeksprogramma

Het onderzoek dat bij de projectgroepen 'Uitbreidbare Programmeeromgevingen' (AP3) en 'Expertsystemen' (AP5) plaatsvindt, maakt voor de ontwikkeling van prototype systemen voornamelijk gebruik van de talen PROLOG en LISP. Het ontwikkelen van grote software systemen in deze programmeertalen -- bij beide projectgroepen is hiervan sprake -- vereist apparatuur die enerzijds een grote verwerkingssnelheid biedt en anderszijds deze talen op adequate wijze ondersteunt.

Bij de projectgroep 'Uitbreidbare Programmeeromgevingen' wordt in het kader van Esprit (GIPE) onderzoek verricht naar de mogelijkheden om interactieve programmeeromgevingen te genereren op grond van formele specificaties. Hierbij worden allerlei nieuwe technieken voor programmageneratie ontwikkeld om, onder andere, syntax-directed editors, typecheckers en interpretators te genereren uitgaande van een formele specificatie. De ondersteunende programmatuur om dit doel te realiseren is volledig in LISP en PROLOG geschreven.

Bij de projectgroep 'Expertsystemen' wordt onderzoek gedaan naar de toepassing van methoden van kennisrepresentatie en automatisch redeneren in expertsystemen. Hierbij worden hulpmiddelen ontwikkeld die de knowledge engineer ondersteunen bij de constructie van expertsystemen, zoals bijvoorbeeld de detectie van inconsistenties in een kennisbank. Tevens wordt in het kader van het, door SPIN gesteunde, PRISMA-project onder leiding van het Philips Research Laboratorium gewerkt aan methoden voor gedistribueerde inferentie in expertsystemen. De programmatuur wordt geschreven in de talen POOL-X, een parallelle objectgeoriënteerde taal, PROLOG en LISP. Bij dit project dient de ontwikkelde programmatuur voldoende performance te kunnen halen om grote knowledge bases te kunnen interpreteren.

In de nu volgende deelparagrafen wordt het bovenstaande onderzoek nader beschreven.

3.1. Onderzoeksthema's Projectgroep Uitbreidbare Programmeeromgevingen (AP3)

Er is een algemene behoefte waar te nemen aan meer "intelligente" hulpmiddelen bij het construeren van

programma's. Dergelijke hulpmiddelen kunnen kennis bevatten over bepaalde programmeertalen, over toe te passen ontwerp-, implementatie- en testmethode, of over het probleemgebied waarvoor programmatuur ontwikkeld wordt. Een voorbeeld van deze ontwikkeling vormen de programmeeromgevingen voor specifieke programmeertalen: de editor in zo'n omgeving kent de syntax en de typeringsregels van de programmeertaal en kan direkt, d.w.z. op het moment dat een syntaxfout gemaakt wordt of een qua type incorrecte expressie wordt ingevoerd de gebruiker hiervan in kennis stellen. De voordelen van zo'n omgeving zijn groot, maar de ontwerp- en implementatiekosten van dergelijke programmeeromgevingen voor diverse programmeertalen zijn dat ook! Tegelijkertijd zet de ontwikkeling zich voort dat er steeds meer gespecialiseerde (programmeer)talen ontstaan -- zoals, bijvoorbeeld, querytalen, vierde-generatie talen, besturingstalen voor allerlei softwarepakketten, specificatietalen, en talen voor kennisrepresentatie -- die adequate, gespecialiseerde, omgevingen behoeven. Voor elk van deze talen zou opnieuw een editor, een incrementele typechecker, e.d. gemaakt moeten worden. De vraag doet zich nu voor, of het mogelijk is om de verschillen en de overeenkomsten tussen diverse taalafhankelijke programmeeromgevingen te identificeren. Als dit mogelijk is, hoeven de taalafhankelijke aspecten maar een keer ontworpen en geïmplementeerd te worden. Dit leidt tot een generator voor programmeeromgevingen die geparametriseerd is met een beschrijving van de gewenste taal.

Het ESPRIT-project GIPE (Generation of Interactive Programming Environments) stelt zich ten doel dit probleem nader te onderzoeken. In dit project werken BSO (Nederland), CWI (Nederland), INRIA (Frankrijk) en SEMA (Frankrijk) samen.

Het GIPE-project heeft in het bijzonder tot doel de mogelijkheden te onderzoeken om automatisch interactieve programmeeromgevingen te genereren op basis van formele definities van programmeertalen. Met een interactieve programmeeromgeving wordt hier bedoeld een geïntegreerde verzameling hulpmiddelen voor het creëren, manipuleren en compileren van uiteenlopende, gestructureerde objecten zoals programma's in een programmeertaal, specificaties in een specificatietaal, of zelf technische documenten. Voorwaarde is dat de vorm (syntax) en de betekenis (semantiek) precies, formeel, beschreven zijn. Enkele vragen die hierbij aan de orde komen zijn:

- o Hoe kan een programmeertaal formeel gedefinieerd worden?
- o Hoe kunnen, uitgaande van een formele definitie, componenten van een programmeeromgeving gegenereerd worden?
- o Hoe moet een dergelijke generator geïmplementeerd worden?

- o Hoe moet het user-interface van een dergelijk systeem ontworpen en geïmplementeerd worden?

Dit onderzoek heeft -- naast allerlei wetenschappelijke publikaties -- tot nu toe de volgende concrete resultaten opgeleverd:

- o VTP (Virtual Tree Processor), een database management systeem voor manipulatie van programma-objecten in boomvorm.
- o Ontwerp en implementatie van diverse specificatieformalismen, zoals ASF (Algebraisch Specificatie Formalisme) en SDF (Syntax Definitie Formalisme). Bij implementatie van deze formalismen wordt gebruik gemaakt van (a) compilatie van specificaties naar Prolog; (b) nieuwe technieken voor incrementele programmageneratie.
- o Een eerste prototype van de generator voor syntax-directed editors.

De in het kader van dit project ontwikkelde programmatuur is volledig in LISP en PROLOG geschreven en maakt gebruik van het X-WINDOWS systeem voor (gedistribueerd) window-management.

Bij de uitvoering van dit project is aanzienlijke reken capaciteit nodig voor:

- o het uitvoeren van realistische testcases in de vorm van input voor grote, naar Prolog vertaalde, specificaties.
- o het verkrijgen van acceptabele performance van het (gegenereerde) man-machine interface van gegenereerde programmeeromgevingen.
- o het genereren van omgevingen voor bestaande ("grote") programmeertalen.

3.2. Onderzoeksthema's Projectgroep Expertsystemen (AP5)

3.2.1. PRISMA-project

PRISMA (PaRallel Inference and Storage MACHine) is een door SPIN ondersteund samenwerkingsproject tussen het Philips Research Laboratorium, de Universiteit Twente (prof.dr. P.M.G. Apers), de Universiteit van Amsterdam (prof.dr. L.O. Hertzberger), de Rijksuniversiteit Utrecht (prof.dr. J.A. Bergstra) en de DAISY-goep van het CWI (dr. M.L. Kersten). Het centrale thema van dit onderzoeksproject is de toepassing van parallelisme voor de opslag en manipulatie van kennis en data. Het belangrijkste medium van dit onderzoek is een multiprocessor computer, de PRISMA-machine, waarop een prototype relationeel database-systeem en een prototype expertstelsel zullen worden geconstrueerd. De ontwikkeling

van deze machine, op basis van specificaties afkomstig van de database- en expertsysteem-applicaties, vindt ook plaats in het kader van dit project.

De vorm van parallellisme die in het project voor het bereiken van performance-winst bij symboolverwerking onderzocht wordt, is grofkorrelig, expliciet parallellisme. Hierbij wordt in sterke mate gebruik gemaakt van kennis met betrekking tot specifieke probleemgebieden om het parallellisme in de representatie van de data en de kennis te kunnen expliciteren. In het algemeen betreft het probleemgebieden waarbij veel data-intensieve berekeningen plaatsvinden.

Ten bate van de ontwikkeling van programmatuur voor de beide applicaties in het project is een parallelle, objectgeoriënteerde programmeertaal ontwikkeld, die de naam POOL-X draagt. Deze taal maakt het in principe mogelijk een grofkorrelige vorm van parallellisme af te beelden op de hardware van een specifieke parallelle computer, in dit geval de PRISMA-machine.

De projectgroep 'Expertsysteem' is verantwoordelijk voor de ontwikkeling van de parallelle expert system shell. In een parallelle programmeromgeving, zoals de PRISMA-machine, is de noodzaak gebruik te maken van kennisrepresentatievormen die een heldere declaratieve semantiek bezitten groter dan in een traditionele sequentiele omgeving. De declaratieve betekenis van kennisrepresentatievormen is daarom een van de onderzoeksthema's. Anderzijds wordt in het project ook onderzoek gedaan naar de uitdrukingskracht van formalismen in expertsystemen in relatie tot specifieke probleemgebieden, en naar de mogelijkheid een bepaald formalisme te gebruiken als uitgangspunt voor parallelle inferentie. Op grond van deze overwegingen is besloten parallellisme toe te passen voor het realiseren van functionele uitbreidingen op de kennisrepresentatie- en inferentievormen van traditionele expertsystemen, waardoor, in een aantal opzichten, krachtiger expertsystemen ontwikkeld kunnen worden. Een van deze uitbreidingen betreft het waarborgen van de consistentie van een knowledge base tijdens de consultatie van een expertsysteem. Hiertoe wordt na elke inferentiecycclus een consistentietest op de knowledge base uitgevoerd. Een andere toepassing van parallellisme die onderzocht wordt, betreft de parallelle extractie van diagnostische kennis uit case databases.

3.2.2. Validering van kennisbanken

De ontwikkeling van een expertsysteem geschiedt meestal op incrementele wijze door na elke uitbreiding van de knowledge base met produktieregels, of elke wijziging hiervan, na te gaan welke veranderingen zich voordoen in de diagnostische conclusies ten opzichte van een vaste verzameling testgegevens, die de leerpopulatie wordt genoemd. In samenwerking met het Academisch Ziekenhuis Leiden (Afdeling Gastroenterologie) en de Erasmusuniversiteit Rotterdam (Afdeling Maatschappelijke Gezondheidszorg) wordt onderzoek gedaan naar methodieken om de ontwikkeling van kennisbanken

inhoudelijk te sturen. Hierbij wordt uitgegaan van het zogenaamde HEPAR-systeem, een medisch-diagnostisch expert-systeem, dat in samenwerking met deze instellingen ontwikkeld is. Als onderdeel van dit onderzoek wordt door de bovengenoemde instellingen gewerkt aan de constructie van dergelijke leerpopulaties.

Van de expertsystemen die tot nu toe ontwikkeld zijn, heeft slechts een gering aantal (bijvoorbeeld MYCIN en CENTAUR) in beperkte mate een toetsing ondergaan. Hiervoor zijn diverse redenen aan te geven, die voor een deel terug te voeren zijn tot de problemen die men ook ondervindt bij de toetsing van elke procedure. Voor een betrouwbare toetsing van een expertstelsel dient, bij voorkeur op dubbelblinde wijze, gebruik te worden gemaakt van prospectief verkregen gegevens. Dit is de testpopulatie. Deze benadering is echter kostbaar en tijdrovend. Ook zal volledige toetsing vaak niet haalbaar zijn omdat relatief zeldzame aandoeningen niet voldoende in de toetsing betrokken kunnen worden. Een alternatieve benadering waarbij gebruik gemaakt wordt van retrospectief verzamelde gegevens kan echter nuttige informatie verschaffen over de kwaliteit van het systeem. Daarom wordt door de eerder genoemde instellingen, in samenwerking met het CWI, gewerkt aan de ontwikkeling van een bestand dat voor toetsing van het HEPAR-systeem gebruikt kan worden.

3.2.3. Plausible reasoning

In expertsystemen blijkt het bij het modelleren veel probleemgebieden noodzakelijk te zijn om in de kennisrepresentatie aan te kunnen geven hoe zeker men is ten aanzien van de frequentie van voorkomen van bepaalde relaties tussen bepaalde fenomenen. Tevens blijkt het omgaan met ontbrekende informatie bij de raadpleging van een expertstelsel in sommige probleemgebieden van groot belang te zijn. In de MYCIN-achtige expertsystemen wordt met onzekere en onvolledige informatie gemanipuleerd door de toepassing van het onzekerheidsmodel van Shortliffe en Buchanan. Dit model bezit echter een onduidelijke mathematische basis, zodat het niet zeker is welke betekenis aan de conclusies van een MYCIN-achtig expertstelsel moet worden toegekend.

In de projectgroep wordt onderzoek gedaan naar deze onzekerheidsmodellen, met het doel om bepaalde eigenschappen hiervan op te helderen door relaties te leggen met axiomatisch beter onderbouwde modellen van onzekerheid, zoals de waarschijnlijkheidsrekening.

4. Relevantie voor overheidsbeleid en maatschappij

De programmatuur die in het kader van het onderzoek bij de genoemde projectgroepen ontwikkeld wordt, stelt het bedrijfsleven in de gelegenheid kennis te nemen van moderne hulpmiddelen en technieken op het terrein van expertsystemen en software engineering. De expertise met de constructie van expert system shells en met knowledge engineering bij de projectgroep 'Expertsystemen' kan ten dienste worden gesteld van het bedrijfsleven om aldaar de introductie van expert-

systemen te begeleiden.

5. Positie van het onderzoek

Het besproken onderzoek is zowel nationaal als internationaal bekend, zoals blijkt uit de participatie in onder andere het SPIN-project PRISMA, en de ESPRIT-projecten 348 (Generation of Interactive Programming Environments) en 415 (Parallel Architectures). Over dit onderzoek wordt uitvoerig gerapporteerd middels wetenschappelijke publikaties, bijdragen aan congressen, e.d.

6. Instandhouding van het onderzoek

Door de gevraagde apparatuur wordt een kritieke massa overschreven in de apparatuurvoorzieningen. Zonder deze apparatuur zal het onderzoek op veel bescheidener schaal uitgevoerd moeten worden.

7. Samenwerking met andere onderzoeksterreinen

Om inzicht te verwerven in de toepasbaarheid van bepaalde principes in expertsystemen bestaat contact met diverse instellingen, waar onder de afdeling Rechtsinformatica van de Juridische Faculteit en de afdeling Medische Fysica van de Medische Faculteit van de Universiteit van Amsterdam, de stichting PALGA (Pathologisch Anatomisch Geautomatiseerd Archief) en de afdeling Gastro-enterologie van het Academisch Ziekenhuis Leiden. De beschikbaarheid van voldoende apparatuur zal deze instellingen, in samenwerking met het CWI, in staat stellen de mogelijkheden van expertsystemen in relatie tot het eigen vakgebied te onderzoeken.

In meer algemene zin is de aangevraagde apparatuur in principe geschikt voor het uitvoeren van onderzoek op de volgende, aangrenzende, onderzoeksgebieden:

- o Kennisbanken;
- o Logisch programmeren;
- o Mens-machine dialogen;
- o Symbolische wiskunde;
- o Programmaverificatie;
- o Natuurlijke-taal verwerking.

Waar mogelijk zal geprobeerd worden de apparatuur ook in deze gebieden in te zetten.

8. Eisen aan de gevraagde apparatuur

Gezien de zeer snelle veranderingen in de markt voor computers en werkstations volstaan wij hier met een opsomming van de eisen waaraan de gevraagde apparatuur dient te voldoen. Het ligt voor de hand om bij de uiteindelijke selectie van

deze apparatuur vooral te kijken naar zeer krachtige werkstations die -- naast grafische faciliteiten en aansluitingsmogelijkheden in een lokaal netwerk -- gebaseerd zijn op, bijvoorbeeld, een RISC architectuur of op speciaal voor de verwerking van LISP of PROLOG ontworpen architecturen. De meer specifieke eisen zijn:

- o Bedrijfsstelsel: onder andere UNIX.
- o Mogelijkheid tot aansluiting op een local area netwerk met Ethernet Interfaces (TCP/IP protocol) en een gedistribueerd filesysteem (NFS).
- o Locale diskruimte per werkstation circa 300 Mbyte.
- o Intern wergeheugen per werkstation minimaal 16 Mbyte.
- o Klokfrequentie van CPU minimaal 16 MHz.
- o Bitmap display van voldoende omvang en resolutie (implementatie van X-WINDOWS is vereist).
- o Programmeertalen: PROLOG, COMMON LISP, SMALLTALK en POOL-X.
- o Hulpmiddelen die de ontwikkeling van programmatuur in deze talen ondersteunen.

THE AMOEBA PROJECT

Sape J. Mullender

*Centrum voor Wiskunde en Informatica
Centre for Mathematics & Computer Science
Amsterdam*

and

*Cambridge University
United Kingdom*

Summary

The Amoeba Project, a joint project of the Vrije Universiteit and the Centrum voor Wiskunde en Informatica, both in Amsterdam, is now six years old. The project has already resulted in many publications in journals and conferences of international repute. It is a distributed systems research project which has resulted in the design and implementation of a general-purpose distributed operating system called Amoeba.

Amoeba is now becoming a usable distributed operating system. In this proposal we ask for the funding to turn the Amoeba Distributed Operating System from an experimental system into a production system, usable and used for the day-to-day work in a research institute.

We propose to build a general-purpose distributed operating system, which supports the development and use of distributed, fault-tolerant applications. Existing UNIX† software at CWI will run on Amoeba. Thus, the users are not required to learn a new operating system interface and the enormous body of software available under UNIX can be exploited.

The proposed hardware for the project consists of workstations, equipped with a bit-mapped display, file servers, a processor pool, and other servers giving access to laser printers, typesetters, gateways to other networks, etc. All this equipment must be connected by a communication network.

1. Distributed Operating Systems

Operating systems of the sixties were batch systems. Batch systems were inconvenient for people, but very convenient for computers: The computer – an extremely expensive machine, occupying a large air-conditioned room – had to be kept busy all the time. No time was wasted switching between processes the way a time-sharing system does; no time was wasted either waiting for input or output; that was done off line.

In the seventies, when the prices of these machines came down a little bit, and when computer science had made some progress, time-sharing systems took over. Here, a sizeable number of people could use a single computer simultaneously, the users could pretend to have

† UNIX is a trademark of AT&T Bell Laboratories

the whole computer to themselves. The only way in which the sharing was noticeable was that the response time varied with the number of users.

Prices are still falling rapidly so that now, in the eighties, giving each user a personal workstation has become cost-effective. Today, a processor of sufficient power to serve the needs of a single person costs less than one tenth of a processor powerful enough to serve ten. Time sharing, in fact, is no longer always a satisfactory way to use a computer system: the arrival of bit-mapped displays with graphical interfaces demands instant response from the system, a response that can only be provided by a dedicated, thus personal, processor.

The workstations of the nineties will be even more powerful than those of the eighties. We envision that a typical workstation will be a multiprocessor with on the order of ten processors and a hundred megabytes of primary memory. Such workstations will have a high-resolution colour display, voice input and output, a keyboard and one or more pointing devices. Network interfaces will allow communication at rates of many megabits per second.

The workstation users of the eighties have lost something compared to time-sharing systems, however: a single, shared environment which allows people on the same project to read each other's files, to compile each other's programs, to share a common printer. Each user has a private environment in which sharing must often be done by exchanging diskettes.

The maintenance problem is an even greater one. In the time-sharing days, the operators could back up the file system every night; the system administrators could allocate the available processor cycles where they were most needed, and the systems programmers could simply install new or improved software. In the workstation environment, each user must be an operator, a system administrator and a systems programmer, all rolled into one. In a building with a hundred autonomous workstations, the operators can no longer go round making backups, the systems programmers can no longer install new software by simply putting it on the file system.

These problems were only partially recognised when workstations made their debut. Here, at CWI, with perhaps 20 workstations installed, it has already become clear that only a system's expert can really use a workstation to advantage.

Some solutions to the problems have been attempted, but none of the current solutions are as satisfactory as was the shared environment of a time-sharing system: They either consist of the addition of network-copy commands with which files can be transferred from one workstation to another, or – slightly better – of a *network file system*, which allows some real sharing of files. In all solutions, however, the user must be aware of the difference between local and remote operations. The real problem is that the traditional operating systems which still form the basis of the workstation software were never designed for an environment with many processors and many file systems. For such environments a distributed operating system is required.

In the computer system of the nineties – a distributed system – the distinction between local and remote operations disappears. Programs do not necessarily execute on the workstation where the command to run them was given. There is one file system, shared by all the users. Peripherals can be shared. Processors can be allocated dynamically where the computing resource is needed most.

A distributed system is a system with many processing elements and many storage devices, connected together by a network. Potentially, this makes a distributed system more powerful than a conventional, centralised one in two ways: It can do more work in the same amount of time, because many computations can be carried out in parallel. And it can be more reliable, because every function is replicated several times. When one processor fails, another can take over the work. Files can be stored on several disks, so a crash does not destroy any information.

These two properties, *parallelism* and *fault tolerance* make a distributed system much more pleasant to use than a traditional operating system. In the Amoeba project, a major goal is to build such a distributed system, a system that exploits parallelism and fault tolerance, but one that is still easy to use.

2. Distributed Systems Laboratory

It has often been said that computer science research in The Netherlands lags behind with that in the U.S. and Japan. This is partly true. Although excellent computer science research is being carried out in The Netherlands, there is not very much of it. Distributed Systems research, as far as I know, is only carried out in Amsterdam (Amoeba), Twente (Tumult) and Eindhoven (Doom). Yet, it is an important area of research, as is indicated by the fact that the four top universities in computer science in the U.S. (MIT, CMU, Stanford and Berkeley) each have distributed systems groups of at least 25 people.

To support the distributed systems research in The Netherlands it is essential that research here is not carried out in isolation. Frequent and intensive contact is needed between researchers here and in other countries. The casual contacts made at conferences and workshops, although extremely valuable, are not enough by themselves.

It is our intention to make CWI an attractive place for computer science researchers – and distributed systems researchers especially – to visit. To make this so, facilities for doing research must be available in the form of a state-of-the-art system. Much as high-energy physics research depends on the largest accelerators does computer science research depend on the best and most advanced computer systems. Fortunately, these systems are not as expensive as accelerators and are, in fact, quite cheap compared to the computing equipment used in physics research. Today, state-of-the-art workstations cost around 20,000* guilders, state-of-the-art file servers (serving 10 workstations) around 200,000 guilders. A share of a printer, tape drive, etc., plus the required cabling costs another 10,000 or so guilders. Thus, an investment of about 50,000 guilders is needed per computer science researcher. Computer equipment is written off over a period of approximately three years, so an investment of in the order of 17,000 guilders per researcher per year is needed, only a fraction of the salary cost.

In systems research, slightly more equipment per researcher is required, because, obviously, one cannot do the program development for the system on the system being built. Spare machines are needed for the necessary experimentation.

3. Similar Projects

Four projects of international stature are closely related to the proposed project and have similar aims. The oldest of the four is the Cambridge Distributed System [Needham82] which has been operational now for many years. The second project is the V-System [Cheriton83] whose aims and structure are closest to the Amoeba system of these four. The two others are both very large projects. Both the ITC System [Satyanarayanan84] of CMU and Project Athena are distributed systems with hundreds of nodes (Athena now has 1,500 workstations) for a student environment.

In The Netherlands, two distributed systems projects are being carried out besides the Amoeba project: the Tumult project of the Technische Hogeschool Twente [Jansen85], and the DOOM project of Philips Natuurkundig Laboratorium. Both projects are sufficiently different from the Amoeba project to merit their independent existence. Tumult was designed and implemented for a specific purpose: processing hand-written giro forms (cheques) and other images. DOOM assumes a homogeneous computing environment built up of identical processor/memory modules connected by a network. One central machine is special and deals with I/O. The environment is monolingual (POOL-T) and thus unsuitable for general-purpose computing. At this time, I am not aware of any publications about DOOM.

4. Collaboration

The primary partner of CWI in the Amoeba project is the Amoeba group of the Vakgroep Informatica at the Vrije Universiteit, led by Prof. Andrew S. Tanenbaum. The Amoeba project is carried out as a joint project between VU and CWI. Fortnightly meetings guarantee that no work is done twice and that the two parts remain coherent. This collaboration is essential. A successful distributed systems project needs a critical mass of a dozen people at least. Some research is possible with fewer people, but it is not possible to build a distributed system and collect invaluable experience for commercial deployment. Neither the VU alone, nor CWI

alone could carry out the Amoeba project successfully.

A collaboration with the Computer Laboratory of Cambridge University has started in September. Dr. Sape J. Mullender will spend six months in Cambridge during the academic year 1987/1988, and a student exchange programme is being organized as well. Research in Cambridge has recently been concentrated in two areas: designing and building very fast networks (the 160 megabit per second Cambridge Fast Ring, the Universe project connecting local networks via satellite and the Unison project, a project in multimedia networks) and distributed programming environments (the Mayflower Project). The Amoeba project moves in the area between these two: the distributed operating system. Thus, although the research at Cambridge is in different areas from that at CWI, the projects in Cambridge and the Amoeba project in Amsterdam complement each other very well indeed.

A first exchange with Cambridge has taken place in August, when one of the Amoeba systems programmers installed the Amoeba Kernel at the Computer Laboratory in Cambridge. Three Cambridge Ph.D. students visited CWI in October to discuss principles of design of a very fast file server. An experimental version of this file server will then be implemented on Amoeba.

Through a sabbatical at Digital Equipment Corporation's Systems Research Center (SRC) in Palo Alto, collaboration has come about with DEC as well. SRC has displayed an interest in sponsoring the Amoeba project by making one or two Fireflies available to the project. A Firefly is a powerful 5-processor workstation with a single 16 megabyte memory, a 210 megabyte disk subsystem, a bit-mapped display, keyboard, mouse and network interface.

DEC Europe has also become interested in our research and through the DEC European External Research Programme, DEC has consented to sponsor the Amoeba research by allowing CWI a substantial discount on DEC equipment. In return, DEC wants permission to let some of their personnel temporarily join the Amoeba project as interns. DEC imposes no restrictions on the research, nor on CWI's ownership of the Amoeba software.

Funded by the European Community's Cost-11 fund, a joint research project has started in the area of distributed systems for wide-area networks. In this project, the Mandis Project, some 15 institutes from nine European countries are designing and implementing a European distributed operating system, based on Amoeba. The implementors' group of the Mandis Project constitutes the University of Tromsø in Norway, Hatfield Polytechnic, AERE Harwell and Cambridge University (the latter unofficially) in England, and VU and CWI in The Netherlands. This collaboration has been useful both technically and politically. Technically, it has given us much practical insight into the construction issues of high-performance gateways; politically, it has served to give the Amoeba project some visibility in European standards organizations.

Through the membership of Dr. Sape J. Mullender in the Communications and Protocols Task Force of the U.S. Distributed Systems Architecture Board, a connection is laid with the U.S. standards bodies. Simultaneously, through collaboration with the Advanced Network Systems Agency in Cambridge, an important link is made with European distributed systems standardization.

The distributed systems researchers in the Netherlands are now joining forces to form a '*werkgroep gespreide systemen*' within SION's division of *Programmatuur en Architectuur*. Within this working group, research on distributed systems in the Netherlands will be coordinated so as not to duplicate research efforts and make software mutually available. The Amoeba group, in close collaboration with their partners in the Vrije Universiteit, has greeted this plan with enthusiasm and will try to participate.

Within CWI, the SPIN-sponsored Prisma Project has shown interest in Amoeba. The memory management facilities offered by Amoeba are well-suited to the main-memory database design of Prisma. The Prisma architecture is designed to be portable and a port to Amoeba is an attractive option

Een multi-computer faciliteit ten behoeve van het database onderzoek 1990-1995

M.L. Kersten

*Centre for Mathematics and Computer Science
Kruislaan 413, 1098 SJ Amsterdam*

1. Inleiding

Deze korte notitie beschijft het lopende onderzoek op het terrein van Gespreide Adaptieve Informatie Systemen bij het CWI en een kader voor het onderzoek in de periode 1990-1995. Hieruit kan worden afgeleid dat de computer infra-structuur van het CWI, bij gelijkblijvend investerings- en exploitatie-budget, onvoldoende is om dit onderzoek met voldoende kans op succes te kunnen uitvoeren. Succes moet hier worden opgevat als aantoonbaarheid van de door wetenschappelijk onderzoek verkregen inzichten, vormgegeven in een prototype implementatie, waaraan feitelijk gedrag kan worden bepaald en mogelijkheden tot commercialisering kan worden getoetst.

Het project gespreide adaptieve informatiesystemen (DAISY) is in september 1985 opgestart. Er zijn momenteel 3.8 fte onderzoekers bij betrokken en er is een vakature voor 1 fte. Het aantal onderzoekers zal naar verwachting toenemen tot 5.5 in 1992. Hierbij is geen rekening gehouden met financiering van nieuwe deelprojecten en het meewerken van studenten en stagiaires in de lopende projecten. Voor een algemene plaatsbepaling van het database onderzoek binnen het CWI verwijzen we naar het meerjarenplan 1988-1992.

Hoofdstuk 2 beschrijft kort het DAISY project vanuit een wetenschappelijke probleemstelling. In hoofdstuk 3 wordt geschetst in welke richting dit onderzoek na 1990 zal evolueren. Het laatste hoofdstuk geeft aan welke computer infrastructuur gewenst is voor dit onderzoek.

2. Probleemstelling en wetenschappelijk belang

Een geautomatiseerd informatiesysteem kan beschouwd worden als een geïntegreerde verzameling hulpmiddelen ten behoeve van de organisatie, beheer, veredeling en distributie van gegevens in de meest ruime zin van het woord. Traditioneel richt een informatiesysteem zich op de opslag en efficiënte verwerking van gestructureerde bestanden; de benaming 'databasesystemen' is hiervoor dan ook beter op zijn plaats. De kennis vastgelegd in de applicaties (gegeven de beperkingen van database- en besturingssystemen) en de organisatie van de gebruikersinterface zijn echter in een ruimere visie wezenlijke bestanddelen van het totale informatiesysteem. Zo kunnen 'decision support systems' en dialoogsystemen het beste als veredelingshulpmiddelen in een informatiesysteem worden beschouwd.

Richten we onze aandacht op de toekomst van informatiesystemen, dan moeten we constateren dat de huidige systeembenaderingen te beperkt zijn, zowel qua functionaliteit als qua flexibiliteit. Voorbeelden hiervan zijn de onmogelijkheid relatief ongestructureerde en ongelijksoortige gegevens, zoals documenten, beeld en geluid, op eenzelfde voet te behandelen als uniforme gestructureerde informatie. Dit staat de ontwikkeling van geïntegreerde applicaties in de weg. Een ander voorbeeld is het ontbreken van de notie 'tijd': traditionele informatiesystemen geven slechts een moment-opname van de gemodelleerde, zich dynamisch wijzigende informatie. Anderzijds moeten de informatiesystemen in de toekomst in twee opzichten adaptief zijn: in de eerste plaats moeten zij zonder 'down' te gaan kunnen worden aangepast aan wijzigingen met betrekking tot zowel de onderliggende architectuur als de (hulp)programmatuur; in de tweede plaats moeten zij in hun eigen architectuur en functionaliteit het evolutieproces dat ieder informatiesysteem als onderdeel van een (dynamische) organisatie doorloopt ondersteunen.

De globale doelstelling van het DAISY project is het ontwerp van (theoretische) modellen en het ontwikkelen van software-technieken die het verwezenlijken van een adaptief en efficiënt informatiesysteem mogelijk maken. De ons voor ogen staande architectuur van informatiesystemen bestrijkt globaal de volgende niveaus:

Applications. Applicaties moeten in dit model worden beschouwd als geformaliseerde dialogen met een specifiek informatiesysteem, ontdaan van hun operationele en representatieve aspecten.

Man-Machine interface. De belangrijkste aspecten van de gebruikersinterface, te weten de dialoog en de presentatie van informatie, zijn in belangrijke mate te parametriseren. Effectieve technieken en hulpmiddelen bestaan reeds, doch zijn te vaak hetzij onlosmakelijk verbonden met het gehele systeem, hetzij met een enkele specifieke applicatie.

Operations. De bewerkingen in een toekomstig informatiesysteem zullen gedefinieerd zijn in diverse programmeertalen. Imperatieve talen (bijv. Pascal) zullen naast functionele talen (LISP, Twentel, Miranda), object-georiënteerde talen (Smalltalk, POOL), deductie-talen (PROLOG), statistische pakketten (SPSS), tekstverwerkers (ditroff, TEX) gebruikt worden. De architectuur van het informatiesysteem ondersteunt de integratie van deze verscheidenheid door verschillende 'views' (representaties) voor gegevens te bieden.

Storage of objects. Bij de opslag van objecten zal, naast de traditionele database-benadering en de eisen, gesteld door het bewerkingsniveau, rekening moeten worden gehouden met kennisregels, de integratie met commerciële kennisbanken (zoals Kluwer's Jurisprudentie), verspreide opslag en verwerking van gegevens, alsmede historische relevantie van de mutaties op de database. Integriteitseisen, ook bij falende apparatuur, en efficiëntie zijn hierbij zwaarwegende aspecten.

Het lopende onderzoek, in de vorm van drie deelprojecten, spitst zich toe op de onderste twee lagen en de formalisering van database modellen. We volstaan hier met een korte schets van de beoogde resultaten en de aanpak.

Het Godal deelproject

Het gaat hierbij om het ontwerp van een programmeertaal voor het constructie van adaptieve informatiesystemen. Hierbij kan worden gedacht aan informatiesystemen waarvoor het ondoenlijk is een logisch consistent data model te ontwerpen, omdat, bijvoorbeeld, zo'n model een te zeer tijdgebonden karakter heeft of niet uitputtend formeel te specificeren is. Anderzijds kan het gewenst zijn dat meerdere, mogelijk tegenstrijdige datamodellen binnen een systeem tot cooperatieve eenheid te bundelen.

Centraal thema in deze experimentele programmeertaal is hoe cooperatieve processen moeten worden beschreven en hoe deze effectief kunnen worden benut. In het bijzonder wordt onderzocht of het *guardian* concept, een process reageert algoritmisch op het herkennen van een (declaratief beschreven) toestand van de database, in een object-gecentreerde database benadering effectief is.

Deelproject database modellering

Een tekortkoming van de meeste, gangbare database modellen is het ontbreken van een adequate beschrijving van de voorwaarden waaronder informatie uit de database kan worden geëxtraheerd. Hierdoor is het mogelijk 'onzinnige' informatie uit de database te halen en leidt dit tot problemen bij bepaalde vormen van database wijzigingen (view updates). Dit project beoogt een wiskundig model te ontwikkelen, waarmee deze problemen formeel kunnen worden beschreven en worden geanalyseerd. Zo'n model is een noodzakelijke voorwaarde voor het toetsen van bestaande database ontwerpen en het realiseren van ontwerphulpmiddelen voor toekomstige systemen.

Het PRISMA deelproject

Het PRISMA (PaRallel Inference and Storage MAchine) deelproject beoogt kennis te verwerven van en ervaring op te doen met database machines die gebaseerd zijn een zeer grote hoeveelheid direct toegankelijk geheugen en vele processoren. Het doel is zeer concreet het vervaardigen van een hardware en software prototype waarin een reeks van innovatieve aspecten op hun gebruikwaarde kunnen worden getoetst. Hierbij kan men denken aan o.a. een database machine architectuur gebaseerd op een generatieve software benadering, exploitatie van potentieel parallelisme door middel van een kennisbank, en integratie van een vierde-generatietaal met een logische programmeertaal.

Dit onderzoek loopt van 1986-1990 en wordt uitgevoerd in het kader van het SPIN project "A Highly Parallel Machine for Data and Knowledge Base Management" waarin samengewerkt wordt met THT, Philips NatLab, UvA, CWI en RUU.

3. Het DAISY project na 1990

Het strategisch database onderzoek in de periode 1990-1995 wordt medebepaald door de problemen die zich momenteel in beide (implementatie gerichte) deelprojecten manifesteren. We zullen kort aangeven welke computer infra-structuur vereist is en welke resultaten hiermee te bereiken zijn.

3.1. Het Godal deelproject

De Godal benadering tot de construeren van informatiesystemen heeft als belangrijkste winstogmerk een verbetering van hun flexibiliteit. De taalconcepten zijn zodanig gekozen en op elkaar afgestemd dat systemen waarop voorhands geen uniforme beschrijving kan worden gegeven, niet uitgesloten worden van automatisering. Het credo is dat ook slecht gestructureerde informatie in een computer systeem moet kunnen worden bewaard en langs elektronische weg moet kunnen worden gemanipuleerd. De software systemen die zulks mogelijk moeten maken dienen uitgerust te zijn met concepten waarmee de gebruiker in de loop van de tijd de informatie alsnog kan organiseren en daardoor kan beheersen.

De keerzijde van de medaille is dat zulke systemen excessief veel reken capaciteit vergen. Immers, op voorhands kan meestal niet worden bepaald of controles en manipulaties achterwege kunnen blijven. De taal Godal speelt hierop in door typerings eisen pas te controleren tijdens programma-executie en door nadruk te leggen op de specificatie van operaties welke parallel kunnen worden uitgevoerd. Parallel in de zin dat zeer veel processen tegelijkertijd de noodzakelijke acties kunnen uitvoeren. Ter illustratie, de prototype implementatie van een abstracte Godal machine resulteerde in een 0.005 Mips systeem. Zelfs van een verbeterde implementatie mag niet verwacht worden dat deze meer dan 0.05 Mips Godal instructies kan verwerken.

De (super) mini-computers bieden onvoldoende reken capaciteit voor een gebruikersvriendelijk implementatie van Godal. Parallel computers, met name de vector processors zoals

CDC205 en GRAYs, zijn ook ongeschikt omdat deze systemen geoptimaliseerd zijn voor numerieke toepassingen. We hebben hier echter te maken met een grote hoeveelheid symbolische manipulaties. Een multi-computer systeem biedt deze vorm van rekencapaciteit wel en de taalconcepten staan een goede afbeelding van de applicaties naar zo'n machine niet in de weg.

3.2. Het PRISMA deelproject

Een primair doel van het PRISMA project is in 1990 te laten zien dat een efficiënt relationeel database management systeem kan worden gebouwd rond een multi-computer. De inbreng van het CWI richt zich op het ontwerp van een uitbreidbaar relationeel DBMS. Zo'n systeem wordt algemeen gezien als een noodzaak voor het ondersteunen van applicatiedomeinen welke vooralsnog niet tot het database domein worden gerekend. Men kan hierbij denken aan kantoor automatisering en CAD/CAM. Voor beide terreinen is de opslag en het gebruik van de gegevens vaak vastgelegd in de applicaties. Er heeft nog weinig data abstractie plaatsgevonden.

Het vervolg onderzoek zal zich dan ook meer en meer richten op het toetsen van de in het PRISMA project gehanteerde benadering in deze, uit database optiek, verwaarloosde gebieden.

Daarnaast zal de komst van WORM (Write Once Read Many) en CD-ROMs leiden tot een heroverweging van de functionaliteit en de architectuur van database management systemen. Immers, de lage kosten van gegevensopslag maakt het mogelijke de history van een gegevensbank vast te houden en deze toegankelijk te maken voor 'historisch' onderzoek. Men kan hierbij denken aan accountants die op 1 april nog steeds de toestand van 31 december j.l. kunnen inspecteren. Zij kunnen hiermee zelfs de wijzigingen traceren en zodoende eenvoudiger frauduleuze transacties op het spoor komen.

Teneinde de in het kader van het PRISMA project ontwikkelde software te kunnen hergebruiken is het gewenst een multi-computer systeem aan te schaffen.

4. Samenwerking en spinoff

De voorgestelde multi-computer biedt niet alleen mogelijkheden voor geavanceerd database onderzoek in de negentiger jaren, maar ook een medium voor gerelateerd onderzoek binnen de afdeling informatica van het CWI. Hiervoor verwijzen we naar de wetenschappelijke plannen en de jaarverslagen van het CWI. Hierbij kan worden aangetekend dat er reeds een verregaande wetenschappelijke afstemming heeft plaatsgevonden met het project gespreide operating-systemen.

Het nationaal database-onderzoek vindt zich momenteel in een opbouwfase. Het voormelde PRISMA project is het eerste grootschalige onderzoek dat op dit moment uitgevoerd wordt. Gezien de nauwe samenwerking en voortgang in het PRISMA project is te verwachten dat deze samenwerking ook na 1990 een vruchtbare vorm zal behouden.

Internationale samenwerkingsverbanden op het onderhavige terrein zijn binnen het CWI uitgekristalliseerd.

5. Een computer-infrastructuur voor database onderzoek

Als onderdeel van het PRISMA project wordt een multi-computer systeem ontworpen. Deze machine bestaat uit 64 microprocessoren, elk uitgerust met 16 Megabyte geheugen, en elk voorzien van een communicatie processor. Het betreft een prototype opstelling die toegesneden is op de software eisen binnen dit project. De totale investeringskosten worden geraamd op 2.0 M.

Uit oogpunt van beheersbaarheid en efficiënt gebruik is het niet gewenst een identieke configuratie te gebruiken voor het database machine onderzoek aan het CWI na 1990. Het is beter een meer algemeen inzetbare multi-computer te gebruiken. Gezien de snelle hardware ontwikkelingen is het voorhands onmogelijk een nauwkeurige specificatie te geven van het

beoogde systeem. We volstaan met de randvoorwaarden waaraan deze moet voldoen.

Multicomputer

Centraal thema in het database onderzoek is het benutten van parallele verwerking. Hiervoor is een multi-computer nodig waarmee schalingsproblemen kunnen worden bestudeerd. Dit vereist een multi-computer met tenminste 16 processoren.

Primair geheugensysteem

Uit oogpunt van performance is het gewenst de componenten in de multi-computer uit te rusten met elk 32 Megabyte. Dit voorkomt o.a. dat parallele verwerking niet nadelig wordt beïnvloed door een te complexe en te trage geheugen hierarchy.

Secondair geheugensystemen

Het secundair geheugensysteem zal voornamelijk dienst doen als stabiel opslagmedium. Normaliter zal deze slechts worden benaderd aan het begin en het eind van de sessie. Dit vereist een opslagcapaciteit van 2 Gigabyte (=16*32*4) mogelijkterwijs aangevuld met WORMs.

Workstations

De ontwikkelomgeving zal moeten bestaan uit een vijftal krachtige workstations waarmee, door middel van windows, zicht kan worden verkregen op de individuele componenten van de multi-computer.

We denken aan een multi-processor met gemeenschappelijk geheugen van minstens 256 Mbyte en snooping caches ter voorkoming van bottlenecks op de geheugen bus. Een alternatieve architectuur bestaat uit vergelijkbare machine waarin echter de communicatie geschiedt via een snel lokaal netwerk. De laatste mogelijkheid is nu reeds commercieel voorhanden. Het laat zich aanzien dat de eerste architectuur in 1990 reëel haalbaar is.

Ter ondersteuning van het lopend onderzoek naar database machines is in de jaren tot 1990 behoefte aan een aantal workstations, zoals SUN 3/60s, en de beschikking over een experimentele processor-pool, zoals voorzien in het AMOEBA project.



Computergrafiek en Gebruikersinterfaces

P.J.W. ten Hagen

Samenvatting

In deze aanvraag wordt voorgesteld hoogwaardige grafische hardware en software aan te schaffen ten behoeve van onderzoek op het gebied van het genereren en manipuleren van twee- en drie-dimensionale beelden. In dit onderzoek komen onder meer aan de orde: gemengde vijfde generatie systemen, waarvan de architectuur tegelijk snelle (parallele) numerieke berekeningen en inferentie mechanismen ondersteunt; zeer geavanceerde invoer mechanismen gebaseerd op patroonherkenning en directe manipulatie; opzet en beheer van grote pictoriële systemen.

1. Karakterisering van het onderzoeksgebied

Door de snelle verbeteringen van de kwaliteit van computer grafische apparatuur, die op zijn beurt weer wordt mogelijk gemaakt door de grote reken capaciteit die in een grafisch werkstation kan worden ondergebracht, is het toepassingsgebied van deze hardware aanzienlijk uitgebreid. Het zwaartepunt is komen te liggen bij het opvoeren van de intelligentie van de computer bij de communicatie met een gebruiker. Voordien werd grafische apparatuur voornamelijk gebruikt als hulpmiddel bij het produceren van tekeningen. Intelligentie komt tot uiting doordat de computer enerzijds de bedoeling van de gebruiker aangeeft te begrijpen door momentaan de bijpassende beelden te genereren, anderzijds middels beelden ingegeven informatie vlot blijkt te kunnen verwerken. Deze mogelijkheden zullen in de nabije toekomst het karakter van computersystemen en het effect op hun omgeving grondig veranderen. De beheersing van deze middelen vereist fundamenteel onderzoek naar de mogelijkheden van formele beeldtalen, geheel nieuwe programmeer-technieken en -omgevingen alsmede nieuwe computer/werkstation architecturen. Dergelijk onderzoek vereist geavanceerde computerapparatuur waarin gecombineerd zwaar numeriek en symbolisch rekenwerk (patroonherkenning) mogelijk is. De apparatuur moet voorzien zijn van hoge resolutie grafische kleuren-beeldschermen en nieuwe invoer-apparatuur.

2. Vernieuwende effecten

De gevraagde apparatuur kan op twee manieren een vernieuwend effect hebben. Ten eerste kunnen nieuwe grafische systemen en gebruikers-interface systemen die met het oog op de verwachte hardware verbeteringen op verouderde apparatuur reeds als experiment zijn ontwikkeld, daadwerkelijk tot een prototype voor de toepassingswereld worden opgevoerd. Dit is essentieel voor de overdracht van deze hoogwaardige technologie naar de industrie (vooral de software-industrie).

Ten tweede zijn er nieuwe wetenschappelijke inzichten te verwachten van systemen met een hoogwaardige gebruikersinterface. Te denken valt aan verbetering van kennisrepresentatie-methoden door deze op te splitsen in methoden voor kennisgebruik en voor representatie van de eigenlijke kennis, nieuwe parallele rekentechnieken en hun beheersing, die via redenering rechtstreeks in verband gebracht kan worden met waargenomen extern (gebruikers)gedrag.

3. Relatie met het onderzoeksprogramma

Er zijn vijf onderzoeksprojecten die essentieel gebruik maken van grafische apparatuur.

Computergrafiek (IS 1) onderzoekt nieuwe hardware architecturen voor snelle beeldgeneratie enerzijds en structurering van beeldinformatie anderzijds voor hoogwaardige interactie. Hierbij wordt onder meer gebruik gemaakt van zelf te ontwerpen chips en andere parallelle hardware. De projectgroep heeft behoefte aan een experimenteel systeem met een open architectuur waarin deze nieuwe componenten kunnen worden ingebouwd.

De groep vervult een voortrekkersrol bij het ontwerpen van internationale standaards (ISO) voor computergrafiek. De nieuwe generatie standaards moet gericht zijn op deze hoogwaardige apparatuur.

Het project Gebruikers Interfaces (IS 2) ontwikkelt nieuwe methoden voor het behandelen van beeldinformatie, gebruikmakend van Kunstmatige Intelligentie. Deze behandeling vereist een object-georiënteerde representatie met een expertsysteem-achtige interpretatie van beeldconfiguraties. Hiermee moet in reële tijd gereageerd worden op gebruikersacties. Daarvoor is een grafisch werkstation nodig waarin in zijn geheel een nieuw type geometrische modelmaker kan worden gehuisvest. Deze modelmaker gebruikt reconstructie-technieken voor constructie van drie-dimensionale geometrie uit twee-dimensionale gegevens (bijv. handschetsen).

De projectgroep Dialoogprogrammering (IS 3) is erin geslaagd een geheel nieuwe programmeertaal te ontwikkelen voor het specificeren van mens-machine-dialogen, inclusief de bestuurlijke consequenties van een dialoogtransactie (e.g. commando). In deze taal kan de gelijktijdigheid van menselijke zintuigen worden nagebootst door het computersysteem middels parallelle invoer-mechanismen van een analoge gelijktijdigheid te voorzien. De taal leent zich eveneens voor het rechtstreeks implementeren van dialogen volgens ergonomische eisen zoals deze vanuit een specifieke toepassing worden gesteld. Het is de bedoeling voor deze taal een zgn. dialoogmachine te bouwen waarmee in een handomdraai elke gewenste gebruikersinterface kan worden gerealiseerd. Hiervoor is eveneens een hoogwaardig grafisch werkstation met nauwkeurige beheersing van het reële tijdsgedrag vereist.

De projectgroep Intelligente CAD-systemen (IS 4) wil kunnen beschikken over een twee-componenten systeem. In de ene component wordt de abstracte technische informatie behandeld op basis van kennisrepresentatie en symbolisch rekenen. In de andere component wordt de intelligente gebruikers interface voor de CAD-gebruiker ondergebracht. Dit systeem moet onder meer in staat zijn de historie van een ontwerpproces te reproduceren. Hierbij wordt gebruik gemaakt van een grafisch gebruikers-interface conform de prototypen uit IS 1, IS 2 en IS 3.

De projectgroep IS 5 zal een bijdrage leveren aan de problematiek rond zgn. product modellers, die de basis zijn voor computergeïntegreerde fabricage. Een productmodeller is een zeer groot pictorieel informatiesysteem. De enorme variatie in grafische presentatietechnieken die voor zo'n modeller nodig is eist eveneens een hoogwaardig grafisch werkstation.

4. Relevantie voor overheid en maatschappij

De afdeling IS heeft reeds met succes resultaten van vroeger onderzoek bij de Nederlandse industrie ondergebracht. Zij ziet grote mogelijkheden om deze activiteiten op bovengenoemde terreinen in versterkte mate voort te zetten. De afdeling is zeer gemotiveerd om internationaal gezien als een van de eerste de mogelijkheden van de nieuwe technologie voor het bedrijfsleven beschikbaar te maken.

5. Positie van het onderzoek

Op het gebied van computergrafiek zijn de bovengenoemde projectgroepen internationaal erkend als voortrekker, hetgeen moge blijken uit het feit dat vertegenwoordigers uit die groepen de leiding hebben gehad bij het tot stand komen van de eerste internationale standaard (ISO) op het gebied van computergrafiek.

De groep IS 1 is tevens sponsor (i.e. voorsteller) van de geaccepteerde internationale standaard voor drie-dimensionale computergrafiek.

Een internationaal opererend softwarehouse heeft zijn activiteiten voornamelijk gebaseerd op grafische systemen, voortgekomen uit onderzoek in IS 1 en IS 2. De groep heeft voor IBM-USA een grafisch werkstation ontworpen. Op het gebied van gebruikers interfaces en grafische informatie systemen wordt IS 3 gezien als toonaangevende onderzoeksgroep. Het abstracte model voor interactie dat in dit kader is ontworpen wordt gehanteerd als basis voor de volgende generatie grafische standaards.

Het project Intelligente CAD-systemen is, althans in Europa, een concentratiepunt voor vergelijkbaar onderzoek. Dit moge blijken uit de grote belangstelling voor de serie internationale workshops die in het kader van dit project worden georganiseerd.

De projectgroep IS 5 is voortgekomen uit, onder meer, een succesvolle pilot-studie die voor ESPRIT-CIM is uitgevoerd. Thans wordt samengewerkt met derden in SPIN/ISNaS- en SPIN/FLAIR-verband. Daarnaast wordt veelvuldig deelgenomen aan internationale conferenties en workshops. Een medewerker is chief-editor van een internationaal tijdschrift voor computergrafiek. Er wordt uitvoerig verslag gedaan middels wetenschappelijke rapporten.

6. Instandhouding

De instandhouding en versterking van de vooraanstaande positie die, gezien de taakstelling te fungeren als centre of excellence, op zijn plaats is, maken het absoluut nodig dat deze specialistische, geavanceerde apparatuurvoorziening gerealiseerd wordt.

7. Consequenties voor de exploitatiebegroting

In principe moeten voor deze apparatuur standaard onderhoudscontracten worden afgesloten. In bepaalde gevallen is door externe samenwerking de mogelijkheid aanwezig het onderhoud gratis te laten verzorgen.

8. Samenwerking met andere onderzoeksterreinen

De architectuurproblematiek van IS 1, met name het VLSI-aspect, heeft geleid tot samenwerking met VLSI-ontwerpgroepen van de Universiteit Twente.

De geometric reasoning aspecten van IS 2 hebben geleid tot samenwerking met knowledge engineering groepen van de University of Southern California en de Universiteit van Tokio.

De projectgroep IS 3 heeft vooral contacten met geïnteresseerde onderzoekers in Nederland, met name TNO/IBBC en Philips CAD-centre. Meerdere verzoeken van buitenlandse universiteiten het programmeersysteem te mogen gebruiken worden overwogen.

De projectgroep IS 4 heeft contacten met instellingen in binnen- en buitenland. Gezien de korte tijd dat het project bestaat zijn intensieve samenwerkingsverbanden bewust uitgesteld, m.u.v. de samenwerking met TNO/IBBC. De projectgroep IS 5 werkt samen met groepen die zich bezighouden met robotics en fabricage-technieken (o.a. Universiteit Twente en Universität Karlsruhe). Onderhandelingen over een mogelijk ESPRIT-project zijn gaande.

In elk van deze gevallen wordt in het kader van deze samenwerking van ons verwacht dat realistische demonstreerbare resultaten kunnen worden geleverd. De geschetste apparatuur is daarvoor onontbeerlijk.

9. Eisen aan de gevraagde apparatuur

De gevraagde apparatuur moet als basis een zeer krachtig werkstation hebben dat kan functioneren als een standaard UNIX™-systeem in de CWI infrastructuur, e.g.:

local area netwerk
transparant filesysteem
electronic mail faciliteit.

Daarnaast moeten de experimentele systemen als coprocessoren aan deze UNIX™ machine gekoppeld kunnen worden.

Tevens is het noodzakelijk voor sommige experimenten een aantal van deze grafische werkstations in een onderling apart net te kunnen schakelen (voornamelijk om verstoring van het algemene net uit te sluiten). De grafische processoren moeten beschikken over (special purpose) cpu-snelheden tot meer dan 200 MIPS. De enorme gegevensbestanden, eigen aan beelden, eisen een diskcapaciteit van 300 Mbyte per werkstation, met een centrale faciliteit van ca. 300 Mbyte "scratch" ruimte voor experimenten. De interne werkgeheugens voor twee installaties moeten minimaal 32 Mbyte zijn. Voor de overige zou 16 Mbyte voldoende zijn.

Voor registratie en analyse van real-time gedrag is koppeling aan een video-installatie nodig.

Na deze globale randvoorwaarden kan het apparatuurbeleid voor Computergrafiek als volgt geformuleerd worden:

9.1. Computergrafiek - basistechnologie

De fundamenteel nieuwe benadering van interactieve 3D-grafiek resulteert in een nieuwe klasse van algoritmen, gegevensrepresentaties en interactie-technieken.

Hiervoor zijn twee experimenteeromgevingen vereist. De ene richt zich op nieuwe hardware architecturen, eventueel op basis van VLSI, waarmee voor de drie bovengenoemde componenten (algoritmen, representaties en basistechnieken) een optimale implementatie kan worden gerealiseerd.

De tweede experimenteeromgeving bestaat uit minimaal één hoogwaardig state-of-the-art 3D-grafisch werkstation (bijv. silicon graphics display), waarmee de meer algemene beginselen uit bovengenoemde benadering kunnen worden gerealiseerd als een proto-type machine-onafhankelijke implementatie. Hiermee kan het bereik van de nieuwe generatie internationale grafische standaard worden vastgesteld.

Gezien de reeds gemaakte vorderingen bij dit onderzoek heeft de apparatuur voor deze activiteiten (vnl. IS 1 en IS 2) de hoogste prioriteit. Bijvoorbeeld is thans de vraagstelling actueel waar de grens ligt tussen realisering van componenten op basis van eigen VLSI-ontwerp dan wel op basis van beschikbare multiprocessor systemen.

9.2. Dialoogmachines

De experimentele omgeving voor dialoogprogrammering alsmede de toepassing van AI-technieken (IS 3 en IS 2) bestaat uit een vijftal hoogwaardige grafische werkstations waarmee naast de verdere experimentele implementaties van dialoogsystemen voor een aantal systemen (min. twee) ook een krachtige symbolische rekenfaciliteit gekoppeld kan worden (evt. erin ondergebracht). De dialoogbasismachine moet op een grote verscheidenheid van installaties gerealiseerd kunnen worden met als resultaat dat interactieve grafische programma's overdraagbaar zijn tussen dialoogmachines met behoud van hoogwaardige interactie.

Op deze installaties zal een groot aantal interactietechnieken worden uitgetoet, enerzijds om het huidige (te beperkte) basisrepertoire te vergroten, anderzijds om methoden te vinden om voor een gegeven applicatiegebied snel de meest adequate, gebruikersvriendelijke dialoog(grond)vormen te vinden.

In het kader van project IS 5 worden onder meer hulpmiddelen ontwikkeld om een gebruiker in staat te stellen parallelle rekenprocessen te besturen. Ook hiervoor kan een dialoogmachine als basis gebruikt worden. Van deze werkstations zal een aantal in 1989 reeds moeten worden aangeschaft. Voor het overige kan de aanschaf in latere jaren plaatsvinden, waarbij dan gedeeltelijk sprake is van vervangen (e.g. doorschuiven verouderde apparatuur).

9.3. Twee componenten machine

Het doel van deze experimentele installatie is een enge koppeling tussen symbolisch rekenen en numerieke behandeling van grafische informatie te bewerkstelligen. Hierdoor kan het redenerende vermogen van het systeem resultaten leveren die direct bijdragen aan de gepresenteerde informatie. Andersom kan een gebruikersreactie op aldus gepresenteerde informatie direct het inferentie mechanisme beïnvloeden. Bijvoorbeeld, evaluatiefuncties uit zoekprocessen kunnen een door de gebruiker te bepalen component bevatten. Voordat deze installatie kan worden ingericht zal zeker nog één à twee jaren vooronderzoek nodig zijn. De experimenten die daarbij worden uitgevoerd kunnen op een eenvoudiger twee componentensysteem (grafisch werkstation en symbolische rekenmachine) worden uitgevoerd. Op bouw van deze installatie is derhalve voorzien vanaf 1990.

Zowel de experimenten uit 9.1 als 9.3 maken een co-processor op basis van dataflow noodzakelijk, die volgens een standaard netwerkverbinding toegevoegd kan worden aan enkele van de grafische werkstations. Co-processoren van dit type zullen worden aangeschaft als onderdeel van enkele van de geavanceerde grafische werkstations.

Tenslotte zij vermeld dat het basis UNIX™-werkstation minimaal een 16 Mbyte geheugen met een cpu van minstens 16 MHz en snelle floating point processor moet bevatten.

Minimaal één van de grafische co-processoren moet vanwege de hybride programmering (numeriek en symbolisch) bij voorkeur van het data flow type zijn.

Parallele Numerieke Algoritmen

H.J.J. te Riele

Motivering en achtergrond voor aanschaf van parallelle computerapparatuur ten behoeve van het onderzoek van parallelle numerieke algoritmen.

1. Karakterisering van het onderzoeksgebied

Nu de single-processor computers de grens van wat, qua snelheid, fysisch nog mogelijk is, beginnen te naderen, komt het in de zeventiger jaren gestarte onderzoek van parallelle numerieke algoritmen (met name voor het oplossen van problemen op het gebied van de lineaire algebra en van partiële differentiaalvergelijkingen) in een stroomversnelling. De komst van praktisch bruikbare en goed toegankelijke parallelle computers, zoals de Alliant FX/8 (met 8 vectorprocessoren) en de Sequent Balance 8000 (met 24 processoren), betekent een extra impuls voor dit onderzoek. Een complicerende factor is de grote variëteit van parallelle architecturen die momenteel voorhanden is (MIMD vs. SIMD, shared vs local memory, programmable vs. fixed application, etc.). Dit heeft tot gevolg dat er een scherpe scheidslijn begint te ontstaan tussen, enerzijds, onderzoek en ontwikkeling van parallelle algoritmen en, anderzijds, de implementatie van parallelle algoritmen op parallelle machines. Het laatste vraagt veel meer inspanning dan het eerste, maar is niettemin absoluut noodzakelijk om de praktische waarde van parallelle algoritmen te kunnen beoordelen. Als gunstige ontwikkeling kan hierbij genoemd worden een duidelijke groei in het onderzoek en het ontwikkelen van methoden ter ondersteuning van het schrijven van machine-onafhankelijke parallelle programma's (met name in Fortran).

2. Vernieuwend effect van de gevraagde apparatuur

Parallele processoren liggen aan het front van de nieuwste computer-ontwikkelingen en verwacht wordt dat deze machines een vitale rol zullen gaan spelen bij de realistische grootschalige simulatie en modellering van fysische en chemische processen. Ook voor real-time toepassingen zijn de verwachtingen van parallelle processoren hooggespannen.

3. Relatie met onderzoeksprogramma

In het onderzoeksproject NW 4.2 (Ontwikkeling van numerieke algoritmen en software voor vector- en parallelle computers) heeft het accent tot nu toe gelegen op vectoralgoritmen en -software, mede door de beschikbaarheid van de Cyber 205 vectorcomputer (bij SARA) en de toegang tot een Cray X/MP 24 computer (te Bracknell, UK). Verwacht wordt dat binnen afzienbare tijd ook toegang verkregen wordt tot een NEC SX/2 vectorcomputer (via SARA bij het NLR). Aangezien vectoriseren en paralleliseren van algoritmen zeer nauw verwante bezigheden zijn, kan het bovengenoemde vector-onderzoek als een verstandige investering beschouwd worden voor parallel onderzoek, mits natuurlijk parallelle apparatuur beschikbaar komt.

Technieken die een rol spelen, maar waar nog zeer veel onderzoek aan dient te worden verricht, staan bekend onder de namen: divide en conquer, data reordering, cyclische reductie, recursive doubling, etc. Het expliciet maken van impliciete methoden kan ook tot goede resultaten leiden.

De algoritmen die bestudeerd zullen worden liggen in het verlengde van de omvangrijke expertise op het gebied van het oplossen van partiële differentiaalvergelijkingen, die de afdeling NW in het verleden heeft verworven. Met name wordt gedacht aan multigridmethoden voor elliptische problemen, ADI- en Hopsotch-methoden, Fast Poisson Solvers, FFT-based methoden, etc. De bovengenoemde technieken en vermoedelijk nog vele andere zullen hierbij een grote rol spelen.

In het onderzoeksproject NW 4.1 (Ontwikkeling van numerieke programmatuur in Ada) wordt

momenteel een Esprit-proposal opgesteld voor het ontwikkelen van Ada- programmatuur voor het oplossen van partiële differentiaalvergelijkingen. Hierbij wordt gebruik gemaakt van de mogelijkheid om in Ada parallelle processen (tasks) te definiëren.

4. **Relevantie voor overheidsbeleid en maatschappij**

De expertise die in dit onderzoek wordt verworven zal, zoals gebruikelijk, zijn neerslag vinden in publicaties, presentaties op congressen, en in software die in programmabibliotheken beschikbaar zal komen. Deze expertise is voor overheid en bedrijfsleven van belang in verband met de noodzaak om te komen tot grootschalige simulatie en modellering van fysische en chemische processen, zoals 3-dimensionale stroming rond vliegtuigen, rivierstromingen, getijdenwerking, grondwaterstroming, vervuilingseffecten, moleculaire structuren, veel deeltjes problemen.

5. **Positie van het onderzoek**

Tijdens een door het CWI, in samenwerking met de TUD en de UvA georganiseerd landelijk Colloquium "Numerical Aspects of Vector and Parallel Processors" is veelvuldig over het betrokken onderzoek gerapporteerd. In het buitenland (USA, Engeland, Frankrijk, Japan, Noorwegen, West-Duitsland) beginnen commerciële parallelle computers succesvol te penetreren in grote bedrijven, onderzoekslaboratoria en universiteiten. Dit zijn dan over het algemeen goed toegankelijke systemen, in een hogere programmeertaal aanspreekbaar, die vrij goedkoop zijn, vergeleken met de huidige supercomputers. Voor enkele tonnen kan men bv. reeds een multiprocessor systeem aanschaffen met een peak performance van enkele tientallen megaflops. Dit heeft geleid tot een significante voorsprong van de onderzoekers die toegang hebben tot deze systemen en tot krachtige impulsen aan hun onderzoeksprogramma's. Dit vindt zijn neerslag in vele publicaties en rapportages op de vele congressen die over het betrokken onderwerp plaatsvinden. Als het CWI (tenminste nog in nationaal verband) een centrale rol wil spelen op het gebied van parallel rekenen, dan zal op zeer korte termijn parallelle apparatuur moeten worden aangeschaft.

6. **Instandhouding van het onderzoek**

Zonder parallelle apparatuur kan geen praktisch bruikbaar onderzoek op het gebied van parallelle numerieke algoritmen worden uitgevoerd. Het is van vitaal belang dat ideeën en theorieën m.b.t. parallelle algoritmen getoetst en geverifieerd kunnen worden op een bestaand parallel systeem.

7. **Samenwerking met andere onderzoeksinstituten**

Er wordt reeds samengewerkt met de TUD (Prof. Van der Vorst, adviseur) en de UvA (Prof. Dekker). In Ada-verband wordt intensief samengewerkt met NAG (Oxford, UK), Siemens en de Universiteit van Karlsruhe (BRD), NIHE (Dublin, UK), CMI (Bergen, Noorwegen), NPL en de Universiteit van Liverpool.

8. **Eisen aan de gevraagde apparatuur**

Gezien de prille staat van het onderzoek van numerieke algoritmen op parallelle machines is het lastig om concrete eisen met betrekking tot apparatuur te formuleren. Zelfs een expert als Dongarra (Scientific Director of Argonne's Advanced Computing Research Facility met drie shared memory parallelle machines (Alliant, Sequent, Encore) en twee local memory hypercubes (Intel iPSC-VX/d5, Intel iPSC/d4 with vectorprocessing nodes)) geeft geen duidelijk antwoord op de vraag wat voor numerieke berekeningen de meest geschikte parallelle machine is. Wel is duidelijk dat shared memory machines voor een grotere klasse van numerieke algoritmen geschikt zijn dan local memory machines. Daar staat tegenover, dat parallelle machines met een shared memory duurder zijn dan die met een local memory. Wat programmatuur betreft is een goede programmeertaal als Fortran 77 (evt. met Fortran 8X features) een basiseis.

Het is van groot belang om zowel grof- als fijnkorrelig parallellisme in numerieke algoritmen te kunnen onderzoeken. Grofkorrelig parallellisme zou kunnen worden onderzocht aan de hand van de processorpool waarvoor plannen klaar liggen in het kader van het onderzoek aan gespreide systemen (Amoeba). Voor onderzoek van fijnkorrelig parallellisme zou men graag over een general-purpose machine beschikken. Een dergelijke machine lijkt echter niet eerder dan in 1990 of 1991 gerealiseerd te worden. Voor het onderzoek aan parallelle numerieke algoritmen is het dringend gewenst al in

1988 over een parallelle faciliteit te beschikken wil het onderzoek geen verdere vertragingen oplopen en de achterstand op het buitenland niet groter worden.

Een in vele opzichten aantrekkelijke keuze lijkt een kleine 4-processor versie van de zojuist gelanceerde tweede generatie iPSC-Hypercube van Intel. De communicatiemogelijkheden tussen de verschillende processoren zijn zodanig verbeterd dat deze machine geschikt lijkt voor onderzoek van zowel grof- als fijnkorrelig parallellisme in numerieke algoritmen. Een dergelijk apparaat is al aanwezig op vele plaatsen in het buitenland (o.a. GMD te Bonn, CMI te Bergen, Noorwegen), en vergroot daardoor de mogelijkheden voor internationale samenwerking.

9. **Concrete wensen met betrekking tot apparatuur:**

direct noodzakelijk: een Intel iPSC-2 Hypercube met 4 vector-nodes, inclusief software (prijs: 250 Kf). Deze configuratie kan model staan voor een grote klasse van hypercube-achtige machines, en biedt daarom goede mogelijkheden voor onderzoek dat aansluit op soortgelijk onderzoek elders. De verbeterde communicatie-mogelijkheden bieden de gelegenheid om zowel grof- als fijnkorrelig parallellisme op deze machine te onderzoeken. (Op zeer korte termijn zal een programma worden aangeschaft waarmee de iPSC Hypercube kan worden gesimuleerd, bv. op de Cyber 990 van SARA of een van de CWI-VAX computers).

Op langere termijn is een goed werkend shared memory parallel systeem met tenminste 8 processoren zeer wenselijk. De voor het Amoeba-project geplande processorpool lijkt hierbij aan te sluiten. Daarnaast bieden parallelle machines met een reconfigureerbare communicatiestructuur interessante mogelijkheden voor onderzoek van fijnkorrelig parallellisme. Het 'general-purpose' karakter maakt dit soort apparatuur tot een interessante kandidaat voor investering in parallelle apparatuur op wat langere termijn (2 à 3 jaar).

Automatisering bibliotheek

Huidige stand van zaken

Het enige bibliotheekproces dat op dit moment is geautomatiseerd is het catalogiseren met behulp van het UBS-pakket, op een Cyber van de SARA. De catalogi worden beschikbaar gesteld op microfichecatalogi. Alle andere administratieve werkzaamheden worden nog manueel gedaan.

Gewenste situatie

- a. Het verbeteren van administratieve processen.
- b. Het verbeteren van de toegankelijkheid van de catalogi door deze online beschikbaar te stellen.
- c. De integratie van de administratieve processen met het catalogiseren.

Het doel is, dat diverse bibliotheekwerkzaamheden (m.n. de uitleen-, tijdschrift- en besteladministratie) geautomatiseerd worden, zodat via terminals gegevens kunnen worden ingevoerd en gewijzigd. Het invoeren van de gegevens voor de catalogi gebeurt nu reeds via terminals, zij het alleen met betrekking tot de data-entry voor het verouderde UBS-systeem. Nieuwe gegevens en wijzigingen worden namelijk via batch-verwerking in het bestand gebracht. In het nieuwe systeem zou ook de verwerking van gegevens online moeten plaatsvinden. Daarnaast is het noodzakelijk dat de catalogi online beschikbaar worden gesteld voor raadpleging, waarbij zowel bibliotheekgebruikers als -medewerkers kunnen profiteren van verbeterde zoekmogelijkheden door middel van zoek sleutels, trefwoorden enz.

Door koppeling met andere bibliotheekbestanden (bijv. het uitleenbestand) kan ook bepaalde informatie zichtbaar worden gemaakt. Dit soort informatie is van groot belang als ondersteuning voor de bedrijfsvoering c.q. beheer.

De catalogi moeten ook online raadpleegbaar zijn voor derden buiten de bibliotheek, bijvoorbeeld door het beschikbaar stellen via het Surfnet.

Recapitulatie

Het plan behelst:

- opzetten administratieve databases;
- online toegankelijk maken bibliotheek-catalogi (m.n. zoekmogelijkheden op woorden uit de titel);
- integratie/koppeling van bepaalde soorten gegevens (bijvoorbeeld uitleenbestand en catalogusbestand);
- catalogi via Surfnet toegankelijk maken voor derden.

De verschillende processen, die met het te ontwikkelen systeem uitgevoerd worden (catalogiseren, catalogus raadplegen, updaten uitleenbestand e.d.), moeten gelijktijdig uitgevoerd kunnen worden.

Samenwerking

De bibliotheken van de Katholieke Universiteit Brabant (KUB), Technische Universiteit Eindhoven (TUE) en Centrum voor Wiskunde en Informatica (CWI) hebben principieel besloten op het gebied van de informatica literatuurverzorging meer te gaan samenwerken.

De respectievelijke specialisatie gebieden kunnen als volgt globaal worden omschreven:

KUB - toegepaste informatica
TUE - hardware en robotica
CWI - theoretische informatica

Voor de concrete invulling van deze samenwerking is een aantal punten vastgesteld:

- Met betrekking tot de tijdschriftencollectie zal een "sterrenproject" worden opgezet. Dit project houdt in, dat voor een aantal met name genoemde periodieken (sterabonnementen) op het gebied van de informatica, de deelnemers de verantwoordelijkheid op zich nemen wat betreft het abonneren, bewaren, toegankelijk maken en verlenen van diensten aan derden. Bij het vaststellen van deze sterabonnementen zal afstemming plaatsvinden.
- Ten aanzien van de Excerpta Informatica database van de KUB is principieel besloten dat een nog nader te bepalen aantal onderwerpsgroepen uit de huidige CWI catalogi zal worden geselecteerd en in Excerpta Informatica zal worden ingevoerd.
Verdere mogelijkheden van samenwerking worden nog nader uitgezocht.
De TUE en KUB werken op dit gebied reeds samen. Zo werkt de Excerpta Informatica database onder het VUBIS systeem, dat door de TUE werd geleverd.

Door de drie instellingen is besloten onderling periodiek te overleggen, niet alleen over de concretisering van bovenstaande punten, maar ook over mogelijke nieuwe gebieden van samenwerking.

Voor de volledigheid wordt er hierop gewezen dat, zoals bekend, zowel de KUB als de TUE reeds enige jaren deelnemen aan de door het CWI opgezette en onderhouden Landelijke Tijdschriften Catalogus Wiskunde en Informatica (LTWI). In totaal nemen 18 bibliotheken deel aan dit project. Twee maal per jaar wordt aan de deelnemende bibliotheken een geactualiseerde versie van de catalogus opgestuurd.

Afstemming met PICA

Het CWI heeft in principe besloten deel te nemen aan "shared-cataloging" van PICA. De eerste contacten daartoe zijn met PICA gelegd.

De integratie met het kantoorautomatiseringssysteem van het CWI zal nader onderzocht worden in overleg met PICA.

De bibliotheek van CWI is overigens vanaf het begin een actief deelnemer van het NCC/IBL-systeem.

Benodigde apparatuur/software

Het systeem draait op een afzonderlijke machine, met optimale wisselwerking met de andere CWI-apparatuur en externe instellingen. De wisselwerking met externe instellingen loop via het Surf-net.

Een nieuw systeem moet naast het verlenen van interactieve toegang tot de bibliotheekcatalogi, ook de uitvoering van de huidige catalogusproducten (microfiches e.d.) ondersteunen.

De respons moet bij een piekbelasting van 40 simultane gebruikers acceptabel zijn.

Het interactief raadplegen van de catalogi door middel van zoek sleutels,

trefwoorden of combinaties daarvan, wordt door de software mogelijk gemaakt.
De machine moet een groot verwerkingsvermogen hebben.

Om de gedachten te bepalen is hieronder een DEC VAX-server 3600 opgevoerd.

* **Hardware:**

- VAX-server 3600
- microvax II-processor
- 32 Mb geheugen
- 622 Mb disk
- streamer tape
- ethernet
- Ultrix-licentie
- 12 empty slots

* **Randapparatuur:**

- X25-kaart
- terminal concentrator, 14 aansluitingen
- laserprinter
- multiplexer, 16 ttys
- console printer
- 2 bar-code lezers
- page scanner
- 12 terminals
- diversen

* **Software**

Kosten (globale raming)

- | | |
|-----------------------|----------|
| * Hardware | kfl. 300 |
| * Randapparatuur | kfl. 100 |
| * Software | kfl. 100 |
| * Conversie bestanden | pm |