jan karel lenstra | the traveling science man

liber amicorum

Redactie
Marlin van der Heijden
Barry Koren
Rob van der Mei
Coby van Vonderen

Eindredactie
Coby van Vonderen

Ontwerp omslag
Kitty Molenaar

Vormgeving en voorbereiding drukwerk
Coby van Vonderen
Jos van der Werf

Foto's en illustraties
Lenstra familiearchief
CWI fotoarchief
Archief auteurs

Drukwerk
Zwaan printmedia, Wormerveer

Publicatiedatum
4 november 2011

Copyright © 2011 Centrum Wiskunde & Informatica
Per bijdrage: Copyright © 2011 auteur(s)

Centrum Wiskunde & Informatica
Science Park 123, 1098 XG Amsterdam

ISBN 978 90 6196 559 6
www.cwi.nl

Voorwoord

Het samenstellen van een liber amicorum staat op enigszins gespannen voet met de maatschappelijk nut huldigende CWI-missie, zeker als Jan Karel de liber-hoofdpersoon is. Jan Karel is zeer alert op het vermijden van overdaad en weelde. Tegelijkertijd stelt hij de hoogste eisen aan het geschreven en gedrukte woord. Volgens Lex Schrijver – in zijn bijdrage aan deze bundel – is één foutje in het colofon al voldoende om Jan Karel te doen besluiten een boek terzijde te leggen en niet meer te bekijken. Kortom, het maken van een vriendenboek voor Jan Karel, dat qua kosten en moeite te rechtvaardigen is en zijn doel bereikt, is een uitdaging.

Voor wat betreft de kosten denken we te zijn geslaagd, door mild redactiewerk en economisch drukwerk. Of we daarmee de vereiste hoge standaard hebben gehaald betwijfelen we, vooral voor wat betreft de typografie. Over de inhoud zijn we echter zeker. Onder de bijdragen bevinden zich pareltjes die Jan Karel vast en zeker zullen uitnodigen om deze bundel te lezen. De auteurs hebben met veel plezier bijgedragen, ieder op zijn of haar manier. Jan Karel – met zijn vele verdiensten voor het CWI – verdient een liber amicorum!

Als redactie danken wij alle auteurs hartelijk voor hun bijdragen, en wensen wij u veel lees-, kijk-, puzzel- en kookplezier!

Marlin van der Heijden, Barry Koren, Rob van der Mei, Coby van Vonderen
CWI, 12 september 2011

Inhoud

Interview

IONICA SMEETS
"Eer is niet meer waard dan een schep kerkhofzand" ***1***

Familieleden

WILLEMIEN LENSTRA, TOM WANSBEEK, JACOB WANSBEEK, FOKELETTE COELINGH, LOTJE WANSBEEK, CLAZINA WANSBEEK EN HAN THOMAS ADRIAENSSEN
Jan Karel: een rijmkroniek in tien strofen ***13***
HENDRIK LENSTRA
Het hoofd en de boekenkast ***21***
ARJEN LENSTRA
Patent nonsense and the Joined Keys Lemma ***24***

CWI-collega's

FRANK ROOS
"Kom jongens, we gaan staffen!" ***29***
DICK BROEKHUIS
Bij het afscheid van Jan Karel als directeur van het CWI ***31***
HANS HIDSKES
Jan Karel ***39***
CARINA KLERX
Beste Jan Karel ***40***
ANGELIQUE SCHILDER
De schoonheid van het woord ***41***
KARIN VAN GEMERT, MARLIN VAN DER HEIJDEN, IRMA VAN LUNENBURG EN LÉON OUWERKERK
Eton Mess ***43***
LIEKE SCHULTZE
Van de wonderwerkster ***44***
SUSANNE VAN DAM, MARTINE ANHOLT-GUNZELN EN NADA MITROVIC
Jan Karel, 8 jaar CWI ***45***
ANNETTE KIK
Jan Karel ***46***
MONIQUE LAURENT
A leisurely traveling salesman tour through several facets of cuts ***47***

LEX SCHRIJVER
Op een goede dag ***60***
HERMAN TE RIELE
Beste Jan Karel ***62***
PAUL VITÁNYI
Jan Karel, The Decider ***65***
HARRY BUHRMAN
I'm sorry, Dave. I'm afraid I can't do that. ***68***
PETER GRÜNWALD
Connecties met JKL ***69***
RONALD CRAMER
Vakant(ie) ***75***
DICK BULTERMAN
An Appreciation of the Absurd ***78***
JAN VAN EIJCK
Afscheid van JKL ***81***
LYNDA HARDMAN
How Jan Karel made CWI more attractive for female researchers ***84***
KRZYSZTOF APT
Beste Jan Karel ***86***
ROB VAN DER MEI
Over wetenschap, besturen, Chinese folklore en respect ***87***
BERT GERARDS
Köln Hauptbahnhof ***90***
BERT ZWART
Grijns ***91***
GUNNAR KLAU
Operations Research and the Life Sciences – a Healthy Combination ***92***
ROELAND MERKS
Discrete explorations of multicellular growth and morphogenesis ***95***
SANDER BOHTE
Detecting Community ***106***
DAAN CROMMELIN
OROV ***112***
JASON FRANK
FOR JKL ***113***
JOOST BATENBURG
De JKL-2080 challenge ***114***
BARRY KOREN
Recht door zee ***117***
KEES OOSTERLEE
Kinderen en Wiskunde houden je jong! ***120***
PIET HEMKER
Over het nut van opa's ***121***
KO ANTHONISSE
Transparantjes en krijtborden ***123***

JAN HEERING
Een wet van Moore voor software? ***125***
JOS BAETEN EN JAN WILLEM KLOP
Typography in Process Algebra ***127***

TU/e-collega's

HARMA KOOPS
Hallo Jan Karel ***141***
ONNO BOXMA
Idolen en idealen ***142***
WIL SCHILDERS
Beslissende momenten ***145***
JAN KORST, GIJS GELEIJNSE, VERUS PRONK EN EMILE AARTS
Chains of Fame from Catherine of Siena to Jan Karel Lenstra ***147***
KEES VAN HEE
Behind the screen it happens: from DSS to TSR ***152***

Promovendi

LEEN STOUGIE
Een uitnodiging tot samenwerken ***165***
LEON WILLENBORG
Statistiek, minimal music, en optimalisering ***168***
MARJAN VAN DEN AKKER EN HAN HOOGEVEEN
Jan Karel aan de zijlijn (en soms als speler) ***173***
TJARK VREDEVELD
De lokaal zoekende Jan Karel ***177***

Vakgenoten

ALEXANDER RINNOOY KAN
Een cursief eindpunt ***183***
GEORGE NEMHAUSER
Liber Amicorum for Jan Karel Lenstra ***185***
DAVID SHMOYS
A randomized variation on a theorem of Lenstra ***187***
ROLF MÖHRING
To Jan Karel ***189***
GERT-MARTIN GREUEL
Jan Karel Lenstra and the Mathematisches Forschungsinstitut Oberwolfach ***191***
JACQUELINE MEULMAN, ELISE DUSSELDORP EN BART JAN VAN OS
An Exact Dynamic Programming Algorithm for Regression and Classification Trees ***198***

Bestuurscollega's

PETER VAN LAARHOVEN
Het Afscheid ***211***
SYLVIA ROELOFS
Elke seconde telt ***212***
JOS ENGELEN
Het Centrum voor Wiskunde en Informatica ***215***
PETER NIJKAMP
Actions for creating an attractive research climate in Europe ***216***
RON DEKKER
Beste Jan Karel ***225***
WIM HUTTER
Prima la matematica, dopo la musica ***226***
PETER STEVENHAGEN
Een cryptokromme voor Jan Karel ***230***
ROB TIJDEMAN
Beste Jan Karel ***231***
PAUL 'T HOEN
Synopsis van een filmscript ***232***
JOHAN VOS
Beste Jan Karel ***234***
JEAN-PIERRE BOURGUIGNON
Dear Jan Karel ***236***
KEITH JEFFERY
Informatics Challenges in Future Internet ***237***
MAURICE ROBIN
Jan Karel LENSTRA: Chairman of the Digiteo Scientific Council ***251***

Rekenonderwijscollega's

ARIE KORBIJN
Jan Karel en het rekenonderwijs ***255***
KEES VAN PUTTEN EN MARIAN HICKENDORFF
De wondere wereld van het rekenonderwijs op de basisschool ***257***
LIEVEN VERSCHAFFEL EN JOKE TORBEYNS
Hoofd- versus cijferrekenen in Vlaanderen. Resultaten van een pilootstudie met de "choice/no-choice" methode ***258***

Omslagontwerpen

KITTY MOLENAAR
Cover voor liber amicorum jkl ***273***

Interview

Interview

Ionica Smeets

“Eer is niet meer waard dan een schep kerkhofzand”

Jan Karel Lenstra neemt na acht jaar afscheid als directeur van het Centrum Wiskunde & Informatica (CWI). In dit interview kijkt hij terug op zijn loopbaan en de ontwikkeling van het CWI. Over boodschappen doen, cleane problemen en onbezorgd onderzoek.

Je werd geboren in 1947, wat voor kind was je?

Ik was nummer twee in een groot gezin en ik was een braaf kereltje. Het was de tijd van de wederopbouw met rust, reinheid en regelmaat.
Ik gedijde daar goed bij. Mijn vader was wiskundeleraar en we verhuisden van Zaandam naar Drachten toen mijn vader daar directeur van de HBS werd. Later werd hij rector aan een lyceum in Groningen en verhuisden we daarheen.

Drachten (1952)

Deed je vader thuis ook aan wiskunde?

Hij gaf altijd wiskundesommetjes bij het eten, hij was heel handig met rekentrucs. Hij leerde ons bijvoorbeeld dat je 98^2 snel kunt berekenen door te gebruiken dat $a^2 - b^2 = (a+b)(a-b)$. Dus je neemt 100 maal 96 en telt daar vier bij op. Je ziet dan in één keer dat 98^2 gelijk is aan 9604.

Ik merkte dat mijn broers bij die sommen sneller waren dan ik. Ook hier op het CWI zit ik tussen mensen die Spinoza's en eredoctoraten hebben, die Veni's, Vidi's en Vici's binnenhalen. Ik denk niet dat mij dat ooit gelukt was.

Waar was jij dan goed in?

Als kind deed ik de boodschappen voor mijn moeder. Mijn broers waren slim, maar ik zorgde voor structuur. Dat lijkt een rode draad in mijn leven. Als ik dingen regelde, dan had iedereen daar plezier van. Dat deed ik niet bewust, het ging vanzelf zo.

Wat wilde je worden in die tijd?

Mijn vader vond dat ik accountant moest worden. Hij gaf me een brochure: "Accountant, een beroep dat boeit" [lacht hartelijk].
Later dacht hij dat ik wel burgemeester van een middelgrote gemeente kon worden. Hij zag in mij geen wiskundige, dat was iets voor hemzelf en mijn broers. Hij dacht dat ik beter leiding kon geven en zorgen voor een goede context. Misschien zag hij dat wel goed.

Toen ik later hoogleraar in Tilburg werd, vroeg mijn vader in welk vak ik hoogleraar werd. Hij was verontwaardigd toen ik "in wiskunde" antwoordde.

Hij zei eerst even niets en zat daarna te sputteren. Dit was tegen de orde der dingen, ik was helemaal geen wiskundige.

Hoe vond je dat?

Ik lachte erom. De rest van de wereld ziet mij als wiskundige, maar thuis zien ze dat niet zo. En ik denk dat thuis gelijk heeft.

Als student (1968)

Koos je daarom bewust voor econometrie als studie?

Ik had geen zin in wis- en natuurkunde. Ik dacht, achteraf gezien ten onrechte, dat dit voor slimmere mensen was. Ik koos voor econometrie aan de Universiteit van Amsterdam. Binnen mijn studie ging ik de kant op van operations research en daarbinnen de optimalisering. Toch wiskunde dus. Halverwege mijn doctoraalfase kwam ik binnen bij het Mathematisch Centrum [MC, de voorloper van het CWI]. Dat was een openbaring.

Wat was er anders dan op de universiteit?

Op de universiteit was het chaotisch. Er werd onderwijs gegeven en je had niet door dat de hoogleraren achter de deur van hun kamer ook nog onderzoek deden. Het MC was een oase van rust en intellect, een kweekvijver van talent en ideeën. Er liepen uitvinders rond die computers en programmeertalen bedachten.

Waar werkte je zelf aan?

Aan sequencing en scheduling, problemen waarbij taken aan machines moeten worden toegekend. Ik zocht naar relaties tussen die problemen. Ik liep eens tegen een artikel met een raar clusterings-probleem aan en ik zag gelijk dat dit een handelsreizigersprobleem was. Dat soort dingen vond ik leuk. Een bezoeker suggereerde dat ik eens moest kijken naar het in 1972 verschenen artikel *Reducibility among combinatorial problems* van Karp.
Dat was een buitengewoon baanbrekend artikel waarin P en NP werden gedefinieerd met probleemreducties. Karp liet zien dat sommige problemen polynomiaal equivalent waren.

Het artikel ging er bij mij in als Gods woord in een ouderling. Dit was precies wat ik nodig had. Er waren problemen die makkelijk waren op te lossen voor twee machines, maar heel moeilijk voor drie machines. In de besliskunde was dat volstrekt normaal. Voor het ene probleem had je een truc, maar bij het andere moest je aanklooien en niemand vroeg zich af hoe dat kwam.

Met deze theorie kon je orde aanbrengen in de chaos. In 1971 werd Alexander Rinnooy Kan op me afgestuurd om iets nuttigs te doen met wiskunde. Samen hebben we een paar jaar besteed aan het trekken van de lijn tussen makkelijke en moeilijke problemen in sequencing en scheduling. Voor problemen die niet in polynomiale tijd waren op te lossen, keken we hoe dicht bij de optimale oplossing je dan wel kon komen.
In 1979 schreven we met Ron Graham en Gene Lawler een overzichtsartikel met de structuur van het vakgebied (*Optimization and approximation in deterministic sequencing and scheduling: a survey*).

"Ik houd van cleane problemen"

Het was een stukje dienstverlening aan anderen, wij boden als eersten een helder overzicht. Dit artikel is inmiddels meer dan 1000 keer geciteerd, terwijl er geen origineel idee of eigen onderzoek in staat.

Hoe werkte je samen met anderen?

Gene Lawler werkte bijvoorbeeld totaal anders dan ik: hij was brille en chaos. Hij werkte een nacht door, rookte twee dozen sigaren weg en kwam met een krankzinnig idee. Alexander Rinnooy Kan draaide soms in een middag een artikel in elkaar. Ik was meer de man van de saaie structuren, ik was maanden bezig om de zaak uit te werken en te polijsten. Uiteindelijk stond er geen letter meer op zijn plaats, waren alle schroefjes en bouten aangedraaid en zat het stevig in elkaar. Zonder een globale visie was het dat nooit geworden, maar zonder mijn gedetailleerd micromanagement ook niet.

Heb je ooit getwijfeld om naar het bedrijfsleven te gaan?

Ik houd van cleane problemen. Een handelsreizigersprobleem is heel clean, maar in de praktijk heb je vijf reizigers in plaats van één en ze moeten stoppen om te slapen. Je krijgt allemaal smerige randvoorwaarden. In de praktijk kom je allerlei dingen tegen waar andere mensen gewoon beter in zijn.

In 1983 werd je deeltijd-hoogleraar in Tilburg. Hoe beviel het daar?

Ik gaf alleen colleges en had geen aansluiting bij de economen die daar werkten. Na vier jaar haalde Rinnooy Kan me naar Rotterdam, maar daar zat ik weer tussen de economen. Collega's werkten voor de havens en zagen niet in waarom je artikelen in internationale wetenschappelijke tijdschriften zou publiceren.

Eind jaren tachtig werd ik onrustig, ik kende het CWI inmiddels van haver tot gort. Toen ik deeltijdhoogleraar in Tilburg werd, wilde directeur Cor Baayen me niet laten gaan en maakte me afdelingschef. In zijn plaats had ik dat nooit gedaan. Ik stuur veel meer aan op doorstroming, er zijn maar weinig mensen die op het CWI hun pensioen moeten halen.

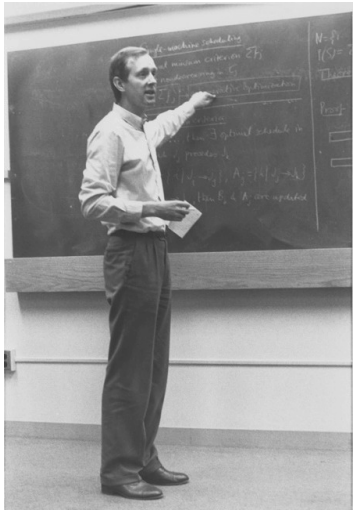
Afscheid CWI (1989)

Ik heb er zelf de eerste keer misschien wel te lang gezeten, twintig jaar. Ik wilde weg, maar wel naar een faculteit wiskunde en informatica, waar de waarheid op de eerste plaats staat en geld op de tweede. Daarom ging ik naar de Technische Universiteit Eindhoven. In Rotterdam snapten ze niet dat ik naar zo'n provinciale plaats ging, maar voor mij scheen in Eindhoven de zon.

In Eindhoven werd je uiteindelijk decaan, hoe kwam je in die functie terecht?

Het was een klus die min of meer automatisch op me afkwam. Ik was al voorzitter van een wetenschappelijke vereniging geweest en vond het normaal dat ik moest gaan besturen. Zo kon ik anderen de kans geven hun wetenschap te beoefenen. Ik heb mijn carrière nooit gepland, ik liet me door toeval leiden en rolde van het ene in het andere. Dat begon al op mijn middelbare school in Groningen toen de vereniging voor welsprekendheid een nieuwe voorzitter nodig had. Ik was helemaal niet welsprekend, ik was juist heel verlegen.
Maar er was gedoe en niemand wilde voorzitter worden.
Anderen toonden vertrouwen in mij en dus waagde ik de sprong in het diepe.
En dat bleek goed te werken, ik vond het zelfs leuk.

In 2002 vertrok je naar Georgia Tech, heb je nooit eerder overwogen om naar het buitenland te gaan?



College op Sloan School, MIT (1988)

Ik reis graag en bracht veel zomers door in Berkeley, op Cornell of MIT. Ik vond de combinatie met mijn baan op het CWI altijd ideaal. Daar kwam ook bij dat ik een dochtertje had. Ik was rond 1980 gescheiden en ze woonde gedeeltelijk bij mij. Ik had veel contact met haar en daarom wilde ik niet permanent weg.

In 2002 was die dochter inmiddels 25, en goed terechtgekomen, en ik was hertrouwd. Mijn vrouw wilde graag naar Georgia Tech. Hun besliskundeschool is de grootste en beste op zijn gebied.

Voor mij was er een interessante vacature als school chair. Tot mijn stomme verbazing kreeg ik de baan niet. Bij mijn tweede sollicitatiegesprek had ik mijn dag niet en heb ik het verknald. Ik kreeg wel een aanbieding om hoogleraar te worden, een eervolle plek met veel geld. Maar ik kon er niets regelen en dat is wat ik het liefste doe. Ik mocht er niet de boodschappen doen.

Toen ik even terug was in Nederland omdat het Wiskundig Genootschap het predicaat Koninklijk kreeg, werd ik gevraagd te solliciteren naar de baan van CWI-directeur. Dat was mijn droombaan, maar mijn vrouw had háár droombaan op Georgia Tech. Dat was lastig en mijn vrouw heeft een offer gebracht door mee terug te gaan naar Nederland. Het is gelukkig goed uitgepakt, ze is inmiddels hoogleraar in Delft. En achteraf gezien is het, ook voor onze kinderen, beter dat we zijn teruggekomen.

Geef je je eigen kinderen net als jouw vader sommetjes?

Af en toe leg ik iets eenvoudigs uit, liefst iets meetkundigs. Bijvoorbeeld dat het verschil tussen $(x+1)^2$ en x^2 gelijk is aan $2x+1$. Dat kun je eenvoudig zien door vierkanten neer te leggen. Zo kun je kinderen heel makkelijk bijbrengen wat een bewijs is.
Ik herinner me dat mijn oudste dochter mij een keer belde omdat haar wiskundeleraar had beweerd dat iets waar was en zij niet wist waarom ze dat zou moeten geloven. Ik gaf haar het bewijs in één zin. Het was even stil en daarna zei ze dat ze haar vriendinnen ging bellen om het uit te leggen. Het is toch onzin dat zo'n bewijs op school uit de wiskunde is verdwenen, dat is juist het wezen van de wiskunde.

Is het CWI veranderd in de tijd dat je weg was?

Nadat ik in 1989 naar Eindhoven vertrok, ging het CWI door een moeilijke periode. Het Informaticastimuleringsplan hield op en er liepen te veel mensen in vaste dienst rond. Het instituut is daar goed uitgekomen. In de jaren negentig was er de angst dat het instituut zou veranderen van onderzoek doen en toepassingen waar dat kan, in een instituut dat toepassingen doet - en onderzoek als dat aan de orde is. Dat is nooit gebeurd, het onderzoek is altijd op de eerste plaats blijven staan.

In de jaren negentig legde directeur Gerard van Oortmerssen meer verantwoordelijkheid bij de onderzoekers zelf. Groepen moesten voortaan hun eigen broek ophouden. Zijn model gebruiken we nog steeds: als er 2 euro van NWO komt gaat er 1 euro in de organisatie en 1 euro in het salaris van de vaste onderzoekers. De groep moet zelf voor die ene euro nog een euro genereren voor PhD-studenten en postdocs.

Wat heb jij als directeur toegevoegd aan het instituut?

In 2009 kwam de nieuwe vleugel gereed. Dat was een belangrijke mijlpaal. In het hele gebouw is een open sfeer ontstaan, op ontmoetingsplekken wordt veel gediscussieerd. Dat is mooi om te zien.

Nieuwe vleugel CWI-gebouw (2009)

Ik heb extra geld van NWO gebruikt voor interactie tussen groepen die elkaar nog niet goed genoeg kenden. Doorbraken gebeuren op grensvlakken. Als een bioloog met een probleem komt, heeft een wiskundige de neiging om die

bioloog zijn wiskundige toolbox te geven en hem dan zo snel mogelijk de deur uit te werken. Hij denkt dat hij niets leert van het toepassen van zijn methoden op andere problemen. Dat is misschien in eerste instantie zo, maar binnen een week komt die bioloog met een probleem dat de bestaande wiskundige toolbox niet aankan. Dan moet je nieuwe tools gaan ontwikkelen. Zo gaat het altijd, ook in de ICT. Er komen altijd nieuwe problemen die de ICT veranderen.

Ik ben trots op de nieuwe groepen waarop je geen wiskunde- of informatica-stempel kunt drukken, zoals cryptologie van Ronald Cramer en life sciences van Gunnar Klau. Ze zorgen met hun Januskop van wiskunde en informatica voor een hechter instituut.

Wat is er niet gelukt tijdens je directeurschap?

De organisatie van het CWI bestaat uit vier clusters met zestien groepen. Ik wilde dat versimpelen en probeerde de clusterlaag ineens te elimineren. Er waren grote voor- en tegenstanders en dat gaf spanningen in het instituut. Dus het was beter om het niet te doen.

"De grote uitdaging is om ervoor te zorgen dat de autonome onderzoekers hier nog onbezorgd fundamenteel onderzoek kunnen doen"

Vond je het vervelend om je plan af te blazen?

Ik vond het lastig in de tijd dat het niet lukte om het plan er doorheen te krijgen. De spanningen in het instituut zaten me dwars. Maar om het af te blazen vond ik een uitstekend idee, er zijn altijd dingen die niet lukken.

Ik heb gemerkt dat dit een instituut van kleine stapjes is, dus nu grijp ik steeds als het kan de kans om groepen samen te voegen. Er komt vast nog eens een moment dat de clusterlaag verdwenen is.

Hoe zie je de toekomst van het CWI?

We zoeken nu geld via wetenschappelijke projecten, maar we zijn genereus naar universiteiten en bedrijfsleven. We zijn een open-source instituut, we geven dingen weg. De luxe van onbekommerd onderzoek van de 20ste eeuw is aan het verdampen en we moeten van contacten en generositeit naar contracten en verdienen. De grote uitdaging is om ervoor te zorgen dat de autonome onderzoekers hier nog onbezorgd fundamenteel onderzoek kunnen doen.

We moeten verder meer samenwerken binnen Europa. Het CWI werkt nu samen met het Franse INRIA. Waarom zou er niet één Europees instituut kunnen zijn

met vestigingen in verschillende landen? Dat is alleen maar goed voor de interactie en kwaliteit. In de Verenigde Staten werkt een heel continent samen. Het geld van conservatieve belastingbetalers in het Midwesten wordt aan de kust door democraten besteed. Als Europa mee wil doen in de wereld, dan moeten we naar Europa-brede samenwerking en competitie. Dan kunnen wij Pools belastinggeld besteden aan goed onderzoek.

Wat is het leukste aan je baan als directeur?

Dit soort veranderingen, de dynamiek van het onderzoek. Dat je de gelegenheid krijgt om, naast alle dingen die goed gaan, veranderingen aan te brengen. Ik zit ook in allerlei commissies en organen om mee te praten over beleidsontwikkeling op het gebied van wiskunde en informatica, bijvoorbeeld de commissie rekenonderwijs van de Koninklijke Nederlandse Akademie van Wetenschappen (KNAW).

Die commissie kwam op een vreemde manier tot stand. Alexander Rinnooy Kan merkte in zijn omgeving dat mensen niet meer konden rekenen. Hij vond dat de discussie in de krant tussen traditionalisten en realisten nergens toe leidde. Toen mailde hij Robbert Dijkgraaf en mij dat wij er, namens de KNAW en het CWI, iets aan moesten doen.

"Dus deed ik de boodschappen weer en werd voorzitter van die commissie"

In eerste instantie dacht ik dat dit wel erg ver van een instituut voor fundamenteel onderzoek afstond. Later bedacht ik dat het juist ook een taak van het CWI is om bij dit soort dingen te kijken wat er aan de hand is en orde op zaken te stellen. Ik wist er geen snars van, maar vond dat iemand dit moest aanpakken. Dus deed ik de boodschappen weer en werd voorzitter van die commissie. Het was hartstikke moeilijk, maar ook heel erg leuk. Uit ons rapport blijkt dat er qua effect geen verschil is aangetoond tussen de realistische en traditionele didactiek. Binnen één didactiek zijn er grotere verschillen dan tussen die didactieken. Eindelijk konden we het hebben over waar het wel over gaat: dat de leraar zelf niet meer kan rekenen. Dat ligt niet aan de leraar, maar aan zijn opleiding en het gereedschap dat hij meekrijgt. Op de PABO is er te veel aandacht voor de tere kinderziel, ten koste van vakkennis. Terwijl je het allebei zou moeten doen. We zijn teveel doorgeschoten.

Op mijn basisschool in Drachten was er helemaal geen didactiek. Met pijn in mijn hart denk ik terug aan de rekencompetitie die mijn onderwijzer hield. Hij zette twee kinderen voor de klas en gaf ze een som als 87 +15. Wie het eerst het goede antwoord gaf had gewonnen en mocht blijven staan. Aan het eind had

hij alle kinderen op het bord geordend in twee kolommen, “dommen” en “knappen”. De knapste van de dommen werd als beloning de domste van de knappen, ik zou zelf trouwens liever de knapste van de dommen zijn.

Wat voor mensen bewonder je?

Iemand als Aldus Manutius uit Venetië rond 1500. Die man stelde zich ten doel, net na de uitvinding van de boekdrukkunst, om de Klassieken uit te geven. Hij was een van de eerste uitgevers.

Maar zoiets klinkt als regelwerk, dat was ook wel iets voor jou geweest toch?

Het werk wat hij deed had me wel gelegen, maar de grootsheid van zijn visie is iets anders. Hij moest een enorm apparaat opzetten met auteurs, letter-ontwerpers, apparatuur en papier. Hij heeft de cursieve letter uitgevonden, want daarmee paste meer tekst op een bladzijde, zo kon hij pocketboekjes maken. Die man heeft impact gehad.

Iets dat altijd is blijven hangen is wat mijn oudoom Jan tegen me zei in 1976. Alexander [Rinnooy Kan] en ik waren net gepromoveerd en gingen na een conferentie in Pittsburg op bezoek bij de kwekerij van mijn oudtante in Albany. Ik was genoemd naar haar broer en deze oom Jan was daar toen ook op bezoek. Hij moet tachtig geweest zijn.
‘s Avonds liepen we over de kwekerij. Hij schopte tegen wat steentjes en zei toen: “Eer, mijn jongen, is niet meer waard dan een schep kerkhofzand”. Dat vond ik zo’n ontzettend mooi citaat, het relativerende dat erin zat. Doe nou maar je werk.

Varøya, Noorwegen (1997)

Familieleden

Jan Karel: een rijmkroniek in tien strofen

Elk lid van de Lenstra-clan
Heeft een heel bijzonder gen.
Want een feestje of partij
Is onaf, zo vinden zij
Zonder vrienden-exposé.
Ik, een Lenstra, doe graag mee
Met dit nieuwste vriendenboek.
Vele uren bracht ik zoek
Met gerijmel en gedicht
Dat uitsluitend is gericht
Op Jan Karels leefverhaal.
Leest en huivert allemaal.

Knaapjes van een jaar of zeven
Staan vaak vrolijk in het leven.
Maar al in zijn jongste jaren
Kon Jan Karel scherp ontwaren
Dat men wereldse problemen
Heel erg serieus moet nemen.
In een korte broek met jasje
Overhemd en zijden dasje
Vulde J zijn kinderdagen
Met het op zijn schouders dragen
Van de mondiale lasten.
Iets dat menigeen verraste
Want Jan Karel Lenstra's maten
Speelden vrolijk in de straten
Zonder zich echt druk te maken
Over serieuze zaken.

Duitsland (1968)

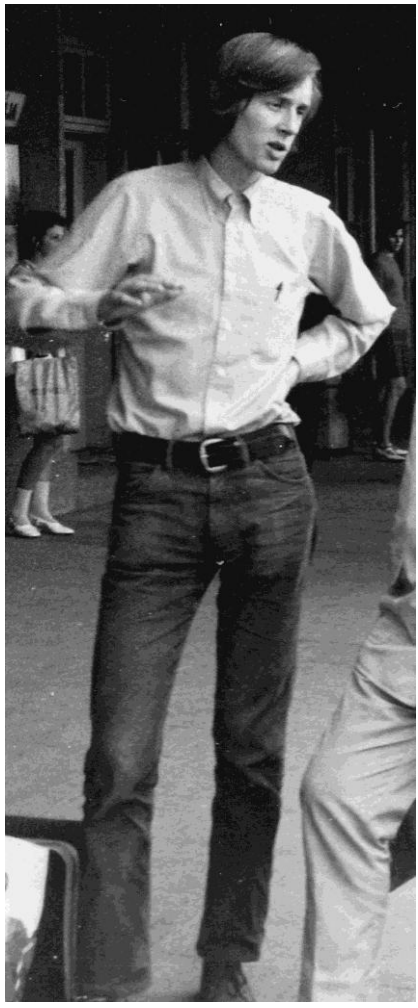
Moeder Lenstra zag al snel:
"Ah, Jan Karel kan ik wel
In gaan zetten voor de klus
Om zijn broertje en zijn zus
Kalm te houden, zoals het hoort."
Spoedig hadden zij een woord
Waarmee zij hun broederlief
Treffend en zeer expressief
Konden schetsen in de taal.
Maar JK was helemaal
Niet hiermede in zijn sas.
Want zijn geuzentitel was
Niet een uiting van gevlei:
Baaskop, ja zo heette hij.

In Jan Karels studiejaren
Kon men echter iets ontwaren
Dat men provoschap kan noemen:
J kan zich erop beroemen
Dat de Mokumse politie
Met hun Ha-twee-o munitie
JK flink hebben beschoten.
Onze held werd natgespoten
Door zijn rol als provo-vechter.
(Critici die zeggen echter
Dat Jan Karels strijdbijdrage
In die veelbewogen dagen
Slechts bestond uit dapper kijken.
Niet heldhaftig mag dit lijken
Maar dat wil ik ferm bestrijden.
Want JK heeft moeten lijden
Door de subversieve daden
Van zijn studiekameraden.

Graf Marx (1970)

Goed, hij is dan niet gestorven
Maar zijn parka was bedorven.)

Na die mooie studietijd
Stortte J zich met veel vlijt
Op promotieonderzoek.
Na voltooiing van dit boek
Ging hij zoeken naar een baan
Want de jaren braken aan
Dat JK behoefte had
Aan een ambt of baantje dat
Echt salaris genereert.
Want JK heeft veel ontbeerd
Tijdens zijn studententijd.
Maar, zo voor de eerlijkheid
Laat ik hier niet onvermeld
dat JK veel extra geld
Door het volgende ontving:
Voor zijn broer- en vriendenkring
Kookte hij met regelmaat.
Geen foie gras of wildgebraad
Maar spaghetti met tonijn
Met een glaasje hoofdpijnwijn
En een rode appel na.
Fijn en goed, etcetera
Maar wat J uitstekend wist
Als goed econometrist:
Slechts de zon krijgt men voor niks,
Meedineren was een riks.
En de gasten van 't diner
Namen braaf dit muntstuk mee.



Hongarije (1971)

Zoals u wel goed zult weten
Wil Jan Karel niets vergeten.
Hij mag dus graag lijstjes maken
Van de te onthouden zaken.
Ik kan ook een lijstje geven,
Maar dan over JK's leven.
Vele academiesteden
Mochten ooit in het verleden
Zich er pauwtrots op beroemen
Om JK te kunnen noemen
Als één van al die personen
Die in deze woonkern wonen
En daar bovendien ook werken
Om de stad zo te versterken.
Tilburg, komt zo bij mij boven,
Rotterdam en Philipshoven.
Eén van JK's vele banen
Was bij de Amerikanen.
In Atlanta, in het oosten
Kon men zich met moeite troosten
Toen JK de bond der staten
Na een tijd weer ging verlaten.
Nu zult u wellicht gaan zeggen:
"J heeft wel iets uit te leggen.
Mocht hij nergens langer blijven?
Wilde men hem graag verdrijven
Om gebrek aan kwaliteiten?"
Dit kan men hem niet verwijten.
Nee, integendeel, zijn gaven
Zijn gemakkelijk te staven
Door de vele publicaties
En de lange lijst citaties
Die hij op zijn naam mag schrijven.
En dit wetenschap bedrijven

Hilversum (1973)

Deed JK vaak met een ander.
Met die partner, Alexander,
Schreef hij in slechts één decade
Stukken voor de beste bladen.
Ja, de gammawetenschappen
Maakten voortgang, grote stappen,
Door het werk van beide heren.
Weinig man kan dat presteren!

Groningen (1986)

Wat men tevens zeggen kan:
J is een familieman.
Hm, dit klinkt wel reuze wee
U krijgt nu vast een idee
Van een eindeloze stoet
Feestjes waar men komen moet
Zodat hij niet zielig is.
Nou, dan heeft u het goed mis.
Want slechts eenmaal in het jaar
Komen Lenstra's bij elkaar
In het huis van onze held.
Iedereen staat steeds versteld
Van de grote hartlijkheid
Die Jan Karel dan altijd
Aan zijn kerstmisgasten biedt.
Want, heus, ik benijd hem niet.
Veertig gasten, potverdrie,
Is een horrorfantasie!
Maar Jan Karel vindt het goed,
Blijft gastvrij en welgemoed.
In zijn overvolle huis
Kruipt hij achter zijn fornuis
En kookt als een waar traiteur
Ieder jaar dezelfde keur
Aan decemberheerlijkheid.

Het menu bestaat altijd
Uit een zalmhap als entree
En daarna, dus als gang twee,
Volgt een weldoorbraden dier
Dat altijd met veel plezier
Door een broer gekocht wordt bij
Een Speciale Slagerij.
Na het beest volgt steeds de spoom
En tot slot is er een droom
Van een chocoladetaart.
Ja, ik voel me wel bezwaard
Dat ik heel het feestmenu
Hier beschrijf, maar nondeju,
Wat J kan op kookgebied,
Nee, dat kunnen velen niet.

'k Mag misschien de indruk geven
Dat Jan Karels hele leven
Slechts bestaat uit kokerellen
En statistische modellen.
Wel, dit moet ik ferm bestrijden
Want zo in de loop der tijden
Is ook in private sferen
Roerigheid te signaleren.
'k Zal niet erg lang uit gaan weiden
Want ik wil al dat vermijden
Dat maar enigszins mag lijken
Op gemene kletspraktijken.
Mensen die dit rijmwerk lezen
Zijn misschien nu aan het vrezen
Dat Jan Karels vroege jaren
Vol ellendigheden waren.
Maar die indruk mist een reden
Want er was in het verleden

Zeker iets met blijdschapswaarde:
Nicht Catrien kwam op de aarde!
Maar met A. is 't niet gelopen
Zoals ieder pleegt te hopen.
Maar er kwam weer nieuwe vreugd:
J begon een tweede jeugd.
Op een leeftijd dat een man
Zonder gêne zeggen kan:
"Ik ga lekker met pensioen."
Ging JK iets anders doen.
Want het lukte Don Juan om
Een charmante, jonge blom
Te verleiden tot de echt.
Heel erg knap, dat moet gezegd.
En bij trouwen, raar maar waar
Hoort altijd een kinderschaar.
Binnen heus een mum van tijd
Werd de Lenstra-clan verblijd
Met een lieve nieuwe nicht.
En dit vrolijke bericht
Kreeg een vrolijk supplement:
Nog een kleine Lenstra-vent.
(Lenstra-Aardal moet het zijn
Maar dat loopt helaas niet fijn.)

Zwitserland (1989)

Lezer, u hoeft niet te vrezen,
Dat u nog erg lang moet lezen.
Deze ellenlange mare
Handelt over al J's jaren
Die tot dusver zijn verstreken.
De auteur heeft niet gekeken
Naar de jaren die gaan komen.
't Is wel aardig om te dromen
Over JK's toekomstplannen.

Leert hij ooit zich te ontspannen?
Gaat hij ooit eens cello spelen?
Leert hij gramschap te verhelen?
Kan hij Ivans completeren?
Zal hij eindelijk eens leren
Dat men T-shirts heus mag dragen?
Deze, en nog veel meer vragen
Moeten nog beantwoord worden.
Maar 't is zeker, een verdorde,
Saaie, donkergrijs behaarde
(Dat is Wansbeeks voor bejaarde)
Zal men J nooit kunnen noemen.
Hij kan zich nog steeds beroemen
Op veel vuur en levenskrachten.
Niemand zal daarom verwachten
Dat J langzaam uit zal doven.
Wij in Groningen geloven
Dat JK een mens zal blijven
Waar veel over valt te schrijven!

Willemien Lenstra
Tom Wansbeek
Jacob Wansbeek
Fokelette Coelingh
Lotje Wansbeek
Clazina Wansbeek
Han Thomas Adriaenssen

Het hoofd en de boekenkast

Opgedragen aan Jan Karel

Hendrik Lenstra

De bovenburen

De bovenburen hebben mij een sleutel van hun appartement gegeven, zodat ik als ze op reis zijn de post op hun keukentafel kan leggen. Deze missie pleeg ik met een summiere veiligheidsinspectie te combineren, en hoewel ik niet in alle hoeken en gaten gekeken heb, meen ik nu te weten: *mijn bovenburen hebben precies zeven boeken.* Deze vormen een stapeltje op de vloer, naast een grote bloempot, en getuigen van weinig bibliofiele gezindheid van de kant van hun eigenaren.

Zeven boeken, dat geeft 5040 mogelijke volgordes! Zouden ze het merken als ik er twee omwissel? Zijn ze gerangschikt op formaat, op kleur, alfabetisch op auteursnaam, chronologisch op jaar van verschijnen? Of wisselen ze zelf elke dag twee naburige om, in bijna 14 jaar een Hamilton-cykel in de betreffende permutatiegroep doorlopende? De problematiek van het ordenen van een privé-bibliotheek is het onderwerp van het onderhavige opstel.

De ton van Diogenes

De zojuist genoemde problematiek confronteerde mij in zijn volle omvang toen bij een recente transatlantische verhuizing verschillende onderdelen van mijn boekenverzameling naar één punt convergeerden, en wel naar mijn Leidse appartement. Dat laatste was er gelukkig op uitgezocht; de lezer stelle zich een ton van Diogenes voor, met genoeg ruimte voor een paar honderd strekkende meter boekenplank.

"Heeft U ze nou allemaal *gelezen*?", verzuchtte één van de met boekendozen sjouwende employés van *De Gruijter & Co., Royal Dutch Movers.* Opeens wist ik mijn antwoord op deze veelgestelde vraag: "Welnee, wat ik gelezen heb, gooi ik weg!" Boeken heb ik om ze te *kunnen* lezen, niet om ze gelezen te *hebben.* Maar ik wil ze wel kunnen *vinden.* In Californië was het me maar al te vaak overkomen dat ik vergeefs naar een boek zocht, en dacht: dat staat in Amsterdam. Tijdens de kerstvakantie leek dat dan niet zo te zijn, zodat ik *back home* met verdubbelde ijver verder zocht. Enzovoort—het kon jaren duren voor ik een boek te pakken had. Daar moest in Leiden een eind aan komen.

De hoofdregel

Bij mijn anderhalve plank *books on books* staan *Theorie van den catalogus* en *Praktijk en theorie der titelbeschrijving*, beide uit de jaren veertig, en beide geërfd van mijn oom.

Dichter bij de theorie van de indeling van een bibliotheek komt mijn boekenverzameling niet. Bovendien, die theorie houdt zich voornamelijk bezig met bibliotheken voor meer dan een enkele gebruiker en is daarom maar beperkt van toepassing op mijn geval. Jan Karel bracht het essentiële verschil met een privé-bibliotheek treffend onder woorden toen ik hem over de ordening van mijn boeken consulteerde: *"zet ze neer zoals ze in je hoofd staan."*

Jan Karels *hoofdregel* is kennelijk pas bruikbaar voor een openbare bibliotheek als door een van staatswege geregisseerde strakke opvoeding ieders hoofd op dezelfde manier ingericht is. Dat is niet wenselijk. Ik beschouw de inrichting van mijn hoofd als een strikte privé-aangelegenheid, en heb de neiging iemand die mijn nu inderdaad volgens de hoofdregel gerangschikte boeken inspecteert, als een indiscrete gluurder te beschouwen.

Ik wil dan ook niet ingaan op de details van mijn implementatie van Jan Karels hoofdregel: op de indeling in categorieën die ik hanteer; op de grens tussen categorieën, zoals klassiek Latijn versus Neo-Latijn, of reisgidsen versus geschiedenis; op de plaatsing van werken die tot meer dan één categorie behoren, zoals een Frans leerdicht over de boekbindkunst; op de rangschikking van boeken binnen één categorie; en op de omgang met seriewerken, met meer dan één schrijver tellende bundels, en met boeken van afwijkende formaten. Bij al deze kwesties hielp de hoofdregel mij op weg, al had ik soms wel de indruk *twee* hoofden te bezitten: één hoofd dat ik op heb als ik een boek wegzet, en een ander dat ik op heb als ik er één zoek.

De uitzondering die de regel bevestigt

Volgens een bekende natuurwet heeft iedere regel een uitzondering, een wet die zichzelf op schone wijze bevestigt door, bijvoorbeeld binnen de wiskunde, zelf talrijke uitzonderingen te hebben.

Het literaire deel van mijn collectie is in overeenstemming met de hoofdregel onderverdeeld naar de taal van het origineel, zodat bijvoorbeeld *Kening Lear*, "yn Fryske oersetting fan Douwe Kalma", bij het Engels staat. Het totstandbrengen van deze onderverdeling vereiste meer werk dan ik gedacht had. Tegenover mijzelf wilde ik namelijk volhouden dat mijn hoofd logisch in elkaar zit, en de logica dicteerde, anders dan ik tot dan toe gedaan had, *Grieks* van *Latijn* te scheiden. Evenwel, bij lang niet alle vertalingen staat de taal van het origineel vermeld. Tussen de verhuisdozen mijn handboeken raadplegende, ontdekte ik dat iemand met een onvervalst Romeinse naam als *Flavius Josephus* in het Grieks geschreven had, en dat ook de kerkvaders gemeen konden uitpakken.

Binnen een taal staan de boeken alfabetisch naar auteursnaam, en dezelfde zucht

om consequent te zijn, dreef mij, voor de afdeling *Grieks*, hiertoe het Griekse alfabet te hanteren. Terugtransliteratie was hierbij het geringste probleem, zelfs bij de zojuist genoemde Flavius Josephus, maar het is moeilijk gebleken eraan te wennen dat ik voor *Chariton* achteraan moet kijken, waar eerst *Xenofanes* stond, die nu na *Nicander* zijn plaats heeft. Dat er een heus verschil is tussen mijn ideale hoofd en mijn echte hoofd, realiseerde ik me pas toen ik ook de Griekse wetenschap op dezelfde wijze ordende, met *Theon van Smyrna* direct volgende op *Hero van Byzantium*. Met verontrusting zie ik de dag tegemoet dat *Kalidasa* niet meer de enige auteur in mijn *Sanskriet*-collectie zal zijn.

Evaluatie

Er valt nog het een en ander te zeggen over de evolutie van de gehanteerde systematiek naarmate de bibliotheek zelf evolueert en, voor de computerbezitter, over de mogelijkheden en moeilijkheden van digitaal boekenbeheer. Ik wil me hier evenwel beperken tot een kort klanttevredenheidsonderzoek.

Kan ik een boek dat ik meen te bezitten inderdaad terugvinden? In verreweg de meeste gevallen is het antwoord *ja*. Als dat niet zo is, dan is de kans het grootst dat het boek bij de binder is, hetgeen ook na te gaan valt, of dat ik het ooit aan iemand cadeau heb gegeven. Er blijkt evenwel ook een kleine maar positieve kans te bestaan dat het boek op een plaats staat die pas bij nader inzien voor de hand ligt. Wanneer men samenwoont, al is het dan met een verzameling boeken, dan zijn kleine onenigheden nooit te vermijden. Dat ze in mijn geval op zo beperkte schaal voorkomen, heb ik te danken aan de hoofdregel van Jan Karel.

PATENT NONSENSE AND THE JOINED KEYS LEMMA

ARJEN K. LENSTRA

Dedicated to Jan Karel Lenstra on the occasion of his retirement as General Director of the Centrum Wiskunde & Informatica

Abstract. This true story sketches the developments that took place as a result of a previous *liber* contribution: patent nonsense, the *Joined Keys Lemma*, and the resulting rogue authority.

Key words: JKL, chaos and whimsical inertia, rogue authority.

Cooking RSA. One of the contributions to the *liber* [2, Page 75] consists of two 438-digit numbers (n and m below) followed by a poem singing the praises of A.J. Lenstra. Upon superficial inspection, the number n looks suspicious. Indeed, writing n as the concatenation $u|\ell$ of 214 leading digits u and a 224-digit tail ℓ, and applying the transformation `01` $\rightarrow$ a, `02` $\rightarrow$ b, ... from [1], it is found that u is a Dutch sentence. It says how to use n as *RSA modulus* and that its factorization can be found "in polynomial time ...". After a search in the family archives [3, Page 132] the number m can be decrypted into a limerick befitting [2]. However, not the limerick but the RSA modulus is the subject of this story.

A regular RSA modulus is the product of two large randomly chosen prime numbers. Because the resulting leading digits cannot be expected to be coherent Dutch, the RSA modulus n was evidently not constructed in the regular fashion. A moment's thought – a moment that occurred between two of many inspirational sips meant to facilitate the creative process of above poem and limerick – learns that an RSA modulus may be constructed in an entirely different way: just fix its initial part u upfront (for instance, as done for n, in such a way that u is a Dutch sentence), randomly select a large prime p, and then craft ℓ so that $u|\ell$ becomes a large prime multiple of p. The resulting moduli do not seem to be easier to factor than regular RSA moduli (unless, as in n's case, the initial part u gives the factorization away), and they are barely harder to construct. As the construction works for any u and p, the number of different RSA moduli that share their u-value is about the same as the number of primes of size similar to p.

If one accepts the presumption that an RSA modulus consists of orderly and meaningful information, chaos, or its numerical equivalent, can now easily be dealt with: any random number may play the role of u above and can thus be turned into the leading digits of a respectable RSA modulus $u|\ell$. Because transforming chaos into order has been Jan Karel's lifelong hobby, another application of this idea is described below. First, however, a brief detour to another consequence of [2].

Patent nonsense. In RSA moduli $u|\ell$ as above, u may contain information (as in [2]) or may be random, but u may also be chosen to facilitate RSA-related calculations. For instance, for $2N$-bit RSA moduli the choice $u = 2^{N-[\log N]}$ or, at the expense of slower generation, $u = 2^{N-1}$ (i.e., blowing $N-1$ zero bits into the RSA moduli), leads to any number of RSA moduli that allow a substantial speedup and smaller footprint of the RSA engine. This is potentially useful.

Intellectual property – even when invented during off-hours – is normally speaking owned by one's employer. So, the threat of endless meetings with patent attorneys was imminent, and could only be staved off if similar prior art could be shown to exist. After all, the idea is so straightforward that it is almost inconceivable that someone had not come up with it before. That suspicion turned out to be correct. At least three independent different parties had done so much earlier: to save space and gain speed a proprietary 1984 French banking standard uses $u = 2^{N-\cdots}$, Adi Shamir had used the trick at a consulting stint, and Don Coppersmith knew about it as well. None of this prior art, however, sufficed to keep the lawyers at bay, since none of it was published. Despite vigourous attempts to unearth even the least shred of evidence, no relevant publication could be found and the unavoidable had to be faced. The rest of this part of the story can be found in US patents 6,404,890 and 6,496,929, assigned to Citibank.

Hash collisions. In August of 2004 Xiaoyun Wang showed how for any *prefix*, different values u_1 and u_2 can quickly be found so that *prefix*$|u_1|$*suffix* and *prefix*$|u_2|$*suffix* have the same MD5-hashvalue for any *suffix* (cf. [7]). Bad news, because the security of the widely used cryptographic hash function MD5 relies on *not* being able to find such *collisions*. Nevertheless, cryptographers rejoiced in the breakthrough and in the job opportunities it implied. Others inertly shrugged their shoulders. The collision-causing u_1 and u_2, though carefully constructed, have no structure that would occur in practice and are hundreds of bits long. MD5-applications could safely be left untouched; at least so it was – whimsically – believed.

Few realized that the safety margin was thin. Upon certification, RSA moduli are hashed, and RSA moduli can be made to usurp hundreds of bits of unstructured information, as shown above. Thus, for any *prefix* that one would typically use when certifying an RSA modulus, different u_1 and u_2 can be constructed such that *prefix*$|u_1$ and *prefix*$|u_2$ collide under MD5, after which ℓ is chosen so that $u_1|\ell$ is an RSA modulus. This leads to colliding c_1 =*prefix*$|u_1|\ell|$*suffix* and c_2 =*prefix*$|u_2|\ell|$*suffix* for any *suffix* that would be appropriate when certifying the RSA modulus $u_1|\ell$. A miscreant who had the sensible string c_1 certified, may claim that the certificate is valid for c_2 instead. This is not a big deal because c_2 is useless, but it is undesirable nonetheless.

JKL. The method from [2] to turn any u into an RSA modulus $u|\ell$ inspired a Chinese remaindering based extension that does the same for any pair (cf. [4]): the *Joined Keys Lemma* joins different u_1, u_2 into $u_1|\ell$ and $u_2|\ell$ that are both hard to factor. With the JKL the above strings c_1 and c_2 can be made equally useful, simply by picking an ℓ that "works" not just for u_1 but for u_2 as well. More in general, the JKL can transform any amount of chaos $\{u_1, u_2, \ldots, u_s\}$ into a set of orderly, hard to factor integers $\{u_1|\ell, u_2|\ell, \ldots, u_s|\ell\}$.

From JKL to a rogue authority. Because the JKL made actual threats conceivable (though they remained far-fetched), it should have been a sufficient argument to discontinue usage of MD5. It wasn't, but it triggered further developments that ultimately convinced most that giving up MD5 may indeed be better.

MD5-practitioners wriggled themselves out of JKL-caused colliding moduli trouble using a variety of arguments. The shared prefix, for instance, implies a single, apparently untrustworthy owner who can easily be traced if irregularities occur. It led to the more general MD5-collisions in [6] which, at considerable computational expense, allowed different prefixes. This enabled c_1's miscreant owner to target anyone as owner of c_2; it also resulted in abnormally large RSA moduli, thus obviating realistic threats.

Another argument was that the certifier includes hard to predict data such as certification time and serial number in *prefix*. Combined with the computational effort involved in constructing collision-causing u_1 and u_2, this offered adequate protection against attacks; at least so it was – still – believed.

Cryptanalysts, however, are a dogged bunch. One certification authority that happened to insert not-so-hard-to-predict data in *prefix* was incentive enough to further sharpen the knives: the RSA modulus length and computational effort were sufficiently reduced, and levelheaded persistence did the rest. The resulting *rogue certification authority certificate* – and the potential to seriously undermine internet security if even a single certification authority keeps using MD5 – is described on [5]. It should be noted that JKL, though crucial in the developments leading to the rogue authority, plays no role in the latter: [2] suffices.

References

[1] M. Gardner, *Mathematical games, A new kind of cipher that would take millions of years to break*, Scientific American, August 1977, 120–124.

[2] R. Koning, H.W. Lenstra, J.K. Lenstra, T.J. Wansbeek, *Andries en zijn kameraden*, ISBN 90-9011507-2, March 1998.

[3] A.K. Lenstra, *Polynomial-time algorithms for the factorization of polynomials*, Ph.D. thesis, 1984.

[4] A.K. Lenstra, X. Wang, B.M.M. de Weger, *Colliding X.509 certificates*, March 2005, `http://eprint.iacr.org/2005/067/`.

[5] A. Sotirov, M. Stevens, J. Appelbaum, A.K. Lenstra, D. Molnar, D.A. Osvik, B.M.M. de Weger, *MD5 considered harmful today, creating a rogue CA certificate*, 25th annual Chaos Communication Congress, Berlin, December 2008, `http://www.win.tue.nl/hashclash/rogue-ca/`.

[6] M. Stevens, A.K. Lenstra, B.M.M. de Weger, *Chosen-prefix collisions for MD5 and colliding X.509 certificates for different identities*, Proceedings EuroCrypt 2007, Springer LNCS 4515 (2007), 1–22.

[7] X. Wang, H. Yu, *How to Break MD5 and Other Hash Functions*, Proceedings EuroCrypt 2005, Springer LNCS 3494 (2005), 19–35.

CWI-collega's

"Kom jongens, we gaan staffen!"

Een schets van het jachtige bestaan van de directeur van het CWI. Elke gelijkenis met bestaande personen berust niet op louter toeval. Situaties en feiten zijn echter gefingeerd, uit hun verband gerukt of niet berustend op de waarheid.
Het is een willekeurige maandagmorgen vijf over elf, vijf minuten later dan de 'staf' - het overleg van de directeur met de managers van de ondersteuning en de directiesecretaris - normaal begint. Jan Karel komt hijgend aanlopen.

Jan Karel klapt in zijn handen: "Kom jongens we gaan staffen! Waar blijven jullie nou?" Dick en Frank snellen naar de kamer van Jan Karel. Angelique en Niels zijn reeds ter plaatse.

Jan Karel begint het overleg: "Sorry dat ik wat later ben, maar ik moest eerst nog langs de lijstenmaker." Hij kijkt op zijn horloge. "Ik heb weinig tijd want ik moet uiterlijk om twaalf uur naar de TWINS-raad. Dus laten we snel ter zake komen. Laten we linksom beginnen. Angelique heb jij iets te bespreken?" Angelique wandelt door haar stapel papieren en pakt er een concept-memo uit gericht aan het GLO. In de discussie die er op volgt is er even verwarring over het aantal onderzoeks-groepen. Er is de laatste tijd ook zoveel geschoven, men is even de tel kwijt. Jan Karel: "Over tellen gesproken, kennen jullie deze? Je hebt drie soorten wiskundigen, zij die niet kunnen tellen en zij die kunnen tellen." Iedereen lacht ontspannen, Jan Karel loopt licht rood aan van plezier.
Jan Karel kijkt weer op zijn horloge: "Jongens we moeten verder. Ik heb weinig tijd." Angelique: "Ik zou nu nog even de structuur van [noemt een kwestie waar zij mee zit] willen bekijken." Jan Karel: "Je weet toch? Eerst komen de processen en daarna de structuur." Frank veert op en onderbreekt: "Maar een structuur kan toch ook eerst komen als je …" Jan Karel: "Nee lieverd, dat kan niet." Jan Karel begint vervolgens te vertellen over een conferentie waar hij eens was en waar een collega een indrukwekkend betoog hield over wat er allemaal fout kan gaan bij het bestrijden van grote rampen. Jan Karel: "Gebleken is uit onderzoek, dat planning in zulk soort gevallen niets helpt en dat de structuur uit de processen moet voortkomen. Planning blijkt een excuus voor interactie. Dat moeten jullie goed onthouden." Vervolgens loopt hij naar zijn indrukwekkende rijen boeken in de kast en pakt er één uit. "Kijk, dit is het boek van die collega. Gedrukt in het lettertype Garamond." Hij zet het boek weer terug. "Ik kwam daar overigens [noemt een naam die niemand van de aanwezigen iets zegt] tegen. Wat een kwaakdoos is die zeg. Ik zal jullie eens vertellen …" Jan Karel begint over een bizar verlopen proefschrifttraject. De term tekstuele incontinentie valt.
Als het verhaal afgelopen is kijkt Jan Karel weer op zijn horloge, het is inmiddels half twaalf: "Dick, heb jij nog iets?" Dick: "Ik heb een paar punten." Dick somt de punten op. Jan Karel: "Ja, dat zijn inderdaad belangrijke punten, maar we hebben

weinig tijd. Ik stel voor alleen [noemt een kwestie] te bespreken en de rest buiten de vergadering te doen." Jan Karel peutert wat aantekeningen op een papiertje van 2x2 cm. Dat lijkt te gaan lukken, want Jan Karels handschrift is economisch zoals een ERCIM-collega eens opmerkte. Dick doet inmiddels een relaas over iets wat niet zo lekker loopt in het instituut. Iedereen knikt en valt Dick bij. Jan Karel: "Kom jongens zo gaan die dingen nu eenmaal. Niet klagen maar dragen. Dick, verder nog iets?"
Omdat het zo verschrikkelijk warm is in de directiekamer - er schijnt iets niet in orde te zijn met de vloerverwarming in de nieuwe vleugel - zijn de aanwezigen toe aan wat drinken en snellen de kamer uit om bij de koffieautomaat iets te gaan halen. Jan Karel roept ze nog na: "Jongens, ik heb nog wat blikjes cola in de koelkast van Euro Shopper …"
Als iedereen terug is neemt Jan Karel de draad weer op: "Over blikjes fris gesproken, kennen jullie de overeenkomst tussen vrijen in een kano en Australisch bier?" Jan Karel vertelt de pointe. Gelach. Jan Karel kijkt weer op zijn horloge: "Oh, we moeten verder, het is al kwart voor twaalf. De taxi staat zo voor. Niels heb jij nog wat?" Niels schudt het hoofd. Jan Karel: "Hoe staat het met de verwarming?" Niels: "Nou, er zijn nog wel wat problemen. Op sommige plaatsen in het gebouw is het te koud en op andere plaatsen is het te warm." Jan Karel: "Tja, dat moet wel opgelost worden. Laten we daar buiten de vergadering nog even op terugkomen. We hebben nu weinig tijd meer. Frank jij nog wat?"
Frank vraagt naar een bijdrage aan een of ander document, die Jan Karel beloofd had. Jan Karel: "We zouden die tekst door een theezeefje gaan drukken hè? Ik ga binnenkort naar het eiland en dan kom ik er aan toe. Verder nog iets?" Frank schudt ontkennend het hoofd.
Jan Karel: "We hebben nog 5 minuten. Dan kan ik verslag doen van het recente overleg met NWO, de stand van zaken met ERCIM, het directeurenoverleg van het Science Park, de situatie m.b.t. ICTRegie [gaat voort met het opsommen van allerlei onderwerpen] en niet te vergeten de huisstijl natuurlijk." Frank probeert Jan Karel nog iets te vragen. Jan Karel: "Ja, mag ik nu even? Val me toch niet steeds in de rede. Oh jongens, ik moet weg, het is twaalf uur. We bespreken een aantal zaken wel bilateraal. Ik zal de secretaresse vragen om de afspraken te maken. Aan de slag!" Jan Karel pakt zijn schooltas, doet zijn jagershoedje op (want het regent buiten) en beent de kamer uit.

Amsterdam, 21 juni 2011
Frank A. Roos

Bij het afscheid van Jan Karel als directeur van het CWI
Ongetwijfeld zal dit boekje veel verhalen bevatten van Jan Karel óp het CWI (of zelfs nog van op het MC). Het leek mij daarom aardig om een inkijkje te geven in een paar dingen buiten het CWI, waar Jan Karel en ik actief waren.

Oost West Thuis Best
Jan Karel werd directeur op het CWI na een kort avontuur in Atlanta, wat lang genoeg was om de familie daar te vestigen en erheen te verhuizen. Zijn appartement in Utrecht was nog wel in bezit, lag voor het vorige werk van Karen en Jan Karel centraal genoeg, maar door de mindere verbinding met het CWI en vanwege gezinsuitbreiding was hij op zoek naar iets ruimers. Tips waren welkom.

En zo gebeurde het dat bij mij in het dorp een villa te koop stond, die voldeed aan de eisen. De afmetingen waren in orde, de foto's zagen er leuk uit en de vraagprijs was OK. Een bezichtiging werd gepland. Het was toen begin 2004. Kort daarvoor hoorde ik bij de bakker (ons dorp is te klein voor een kapper) dat het met dit huis niet om zo maar een huis ging, maar om de voormalige villa van één van de ontvoerders van een bekend biermagnaat. Deze man was begin 2003 geliquideerd. Bij aankomst verwoordde de makelaar de reden van verkoop met "de vorige bewoner is plotseling overleden". Een dergelijke opmerking past, net als "voor de handige doe-het-zelver" en "met authentieke details", naadloos in het bekende makelaarsjargon. Toegegeven, de vraag of dit overlijden 'onvrijwillig' was gebeurd werd zonder aarzelen instemmend beantwoord. Jan Karel ging vervolgens met meetlint en opschrijfboekje de woning door.

Uiteindelijk viel de keus van de familie Lenstra niet op dit huis maar na lang zoeken op een huis in Utrecht dat na een flinke verbouwing (het was voorheen een kantoorpand) betrokken werd. Gelukkig hebben we de foto's nog.

Meten is weten (2004)

Amsterdamse humor: stoeltjes met 'Amstel Bier' logo in de tuin (2004)

Op naar Kreta

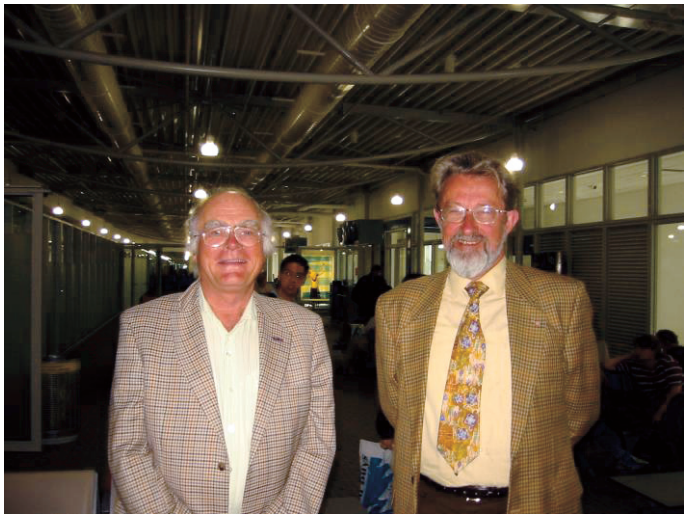
CWI is lid van ERCIM en was in 1989 één van de oprichters van dit netwerk van instituten die onderzoek doen in informatica en wiskunde. In die hoedanigheid gingen we in mei 2004 naar een vergadering van de Board of Directors, georganiseerd door het Griekse lid, FORTH op Kreta. Het was Jan Karels eerste *BoD* meeting maar niet zo maar een vergadering; ERCIM bestond 15 jaar en dat was reden voor de Grieken om flink uit te pakken. Ongeveer 100 personen waren aanwezig voor deze gelegenheid, waaronder twee van de drie oprichters van toen. Cor Baayen vloog als speciale gast met ons mee, Gerhard Seegmüller van het vroegere GMD was er ook. Alain Bensoussan van INRIA was helaas verhinderd, anders waren de *"founding fathers"* compleet.

Een "*celebration event*" met diverse sprekers van binnen ERCIM en de EU DG-Research werd gehouden, een grote taart werd door de "*founding fathers*" aangesneden en als slot een mooi diner dat met vuurwerk werd afgesloten. Kosten noch moeite werden gespaard (met de kennis van nu over Grieken en financiën vermoed je een sigaar uit ERCIMs doos, maar het organiserende instituut FORTH zit goed in de slappe was: ze beheren het .gr-domein en hebben nog een aandeel in de door hen mede opgerichte Griekse internetprovider Forthnet).

Het was naast deze festiviteiten een boeiende vergadering. Jan Karel kon zijn lol op met de Griekse manier van voorzitten en organiseren. De terugreis was vol vrolijke verrassingen. Het was vroeg opstaan: we vlogen zaterdagochtend al om 7:00 vanaf Kreta naar Athene voor de overstap naar Amsterdam, dus dat was om half 5 opstaan. Bij de overstap op het vliegveld van Athene liep Cor Baayen opeens Reind van de Riet tegen het lijf die een week vakantie had gevierd. Terwijl ze met elkaar aan het bijpraten waren liep een andere bekende van hun tweeën, Ton Levelt, ook voorbij. Het werd zo een gezellige boel en de tijd tot de overstap was op die manier snel voorbij.

Met Cor Baayen op de vliegtuigtrap (2004)

Er is er één jarig, hoera, hoera...(2004)



Reind van de Riet en Cor Baayen (2004)

Aan het diner met Keith Jeffery (STFC), Martin Prime (STFC), Peter Kunz (ERCIM), Jan Karel, Cor Baayen, Eckart Bierdümpel (Fraunhofer) en Peter van Roy (UCL) (2004)

In de lichtstad

Naast ERCIM bestaat er ook ERCOM, een comité van de European Mathematical Society, waarin diverse wiskunde-instituten zijn verenigd.

Per 2006 nam Jan Karel het voorzitterschap over. Herman te Riele vervulde uitmuntend de rol van "*scientific secretary*" en mijn taak was het runnen van de groep "*administrators*". Met zijn drieën togen we naar het lommerrijke Bures-sur-Yvette bij Parijs, naar het Institut des Hautes Études Scientifiques voor de jaarlijkse vergadering.

Voor Jan Karel was deze vergadering een uitdaging want de groep luisterde met dusdanig veel respect naar haar voorzitter dat ze zelf nogal kalm en stil bleef. Jan Karel probeerde met alle macht de groep aan de praat te krijgen. Bij een rondje langs de instituten gaf elke deelnemer zo kort wat informatie, dat het Jan Karel nog steeds niet gegeven was even écht achterover te kunnen leunen en te vertrouwen op de dynamiek van het gezelschap. Tot het woord bij de gastheer kwam: Jean-Pierre Bourguignon hield een fantastisch verhaal over IHÉS en een recente fundraising campagne. Dat bood perspectief voor een minuut of 15... Helaas voor Jan Karel was hij toch gedwongen ook hier actief te luisteren omdat Jean-Pierre dit plenair bedoelde verhaal grotendeels in een monoloog gericht tot hem en enkel met oogcontact met hem hield.

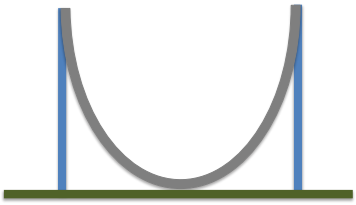
's Avonds was er in de Tour Montparnasse een diner in het restaurant op de 54^{e} verdieping, op 200 meter hoog, met een fantastisch uitzicht over de stad en haar beroemde plaatsen. Gelukkig zit je dan zelf in het gebouw dat alle andere vergezichten over Parijs verpest. Onder het genot van een goed glas wijn en een prima maaltijd kon Jan Karel de dag van zich af laten glijden. Bij latere meetings in Triëst, Coimbra en Djursholm was er meer ruimte voor ontspanning en bij de laatste meeting was er een rondleiding in de bibliotheek van het Instituut Mittag-Leffler waar Jan Karel zijn hart kon ophalen aan het bewonderen van enkele oude boeken.

Eén van de gebouwen op de campus van het IHÉS (2006)

Dôme des Invalides met le Tour Montparnasse (2006)

Op de eerste rang. Anders Björner toont een zeer bijzonder boek (2009)

Tot zover een kleine schets van de activiteiten van Jan Karel en ondergetekende buiten het instituut. Door de jaren heen zijn er ook nog diverse wiskundige puzzeltjes op mij losgelaten, zoals deze, inclusief het "behulpzame" plaatje: Er staan twee palen in het veld met een hoogte van 6 meter. Tussen de palen hangt een touw met een lengte van 12 meter. In het midden hangt het touw precies tot op de grond. Hoe ver staan de palen uit elkaar?



Puzzeltijd!

Ook heb ik diverse stellingen of problemen gepopulariseerd uitgelegd gekregen. Vandaar dat ik in ieder geval wel iets met een wiskundig luchtje moet opnemen in dit stuk: het jaar 2003, waarin Jan Karel begon als directeur van het CWI, en het jaar 2011, waarin hij stopt als directeur, zijn beide slechts deelbaar door zichzelf en het getal 1 en dus priemgetallen. Er zit geen ander priemgetal tussen, toeval?

Jan Karel, voor de toekomst:
veel succes en geluk, in goede gezondheid. En nu aan de slag!

Dick Broekhuis

Jan Karel

Het eerste wat je me vertelde was: "voor mij moet je hier niet komen werken!" Dat was tijdens het sollicitatiegesprek waarin je me liet weten dat dit je laatste jaar als directeur zou worden. Ook vertelde je aan een punt te kunnen zien of die cursief gedrukt was. Dat was even schrikken. Gelukkig bleek je je ook te kunnen vinden in de uitdrukking "perfect is the enemy of good". Wat ik fijn vond, want ik zie dat nog steeds niet, die cursieve punt. Wel ben ik inmiddels een goede dubbele spatievinder geworden. Nu ik erover nadenk is dat sollicitatiegesprek het op één na langste gesprek geweest dat ik met je gevoerd heb. Je hebt niet veel woorden nodig. Je instructies zijn 'crisp and clear' en je agenda weerspiegelt je geordende brein. Ik merk dat het prettig is die duidelijkheid. Ik heb voor vele directeuren gewerkt maar jij bent mijn eerste professor directeur. Van de spreekwoordelijke verstrooidheid is echter geen sprake. Jij hebt mij geholpen mijn eerste stappen te zetten binnen dit wetenschappelijke instituut. Daar wil ik je voor bedanken. Straks zie ik je niet meer voorbij flitsen maar waarschijnlijk zal ik je zo nu en dan bij een 'CWI event' tegenkomen. Ik wens je alle goeds.

Hans Hidskes

Beste Jan Karel,

Graag maak ik van de gelegenheid gebruik om een klein stukje te schrijven in je liber amicorum.

We hebben slechts een jaar samengewerkt maar dat was, in ieder geval voor mij, een bijzonder en enerverend jaar waarin veel is gebeurd en toch ook veel bereikt. Mijn besluit om te gaan 'emigreren' naar Brabant viel bijna samen met jouw aankondiging om eind 2011, na de afronding van de evaluatie van het CWI, het directeurschap over te dragen.

Ik heb de samenwerking met jou als heel plezierig ervaren en bewaar goede herinneringen aan mijn jaar bij het CWI. Wellicht krijg je nu meer tijd en rust om je aan je gezin en het doen van onderzoek te wijden. Ik wens jou en je gezin nog vele mooie jaren toe en zal je blijven volgen.

Hartelijke groet,

Carina Klerx

De schoonheid van het woord
oneffenheid stoort
geen tante betjes
alles moet netjes
Symmetrie in woord en vorm
dat is zijn norm

Gelukkig kunnen soms normen
vervagen en had ik een kans van
slagen door te mogen zijn wie ik ben

bedankt daarvoor Jan Karel!

Beste Jan Karel,

Nu ik dit voor jou aan het schrijven ben, realiseer ik mij maar al te goed dat ons tijdperk van samenwerken bijna op zijn einde loopt. Het is tijd voor reflectie. En als je mij een beetje kent (en dat doe je), weet je dat dit niet zozeer zal gaan over inhoud maar meer over beleving en gevoel.
Ik weet nog goed dat ik samen met Bert Gerards bij jou op de kamer zat voor mijn tweede sollicitatiegesprek. Ik ging heel onbevangen het gesprek in omdat ik nog een ander sollicitatiegesprek in het vooruitzicht had bij een bedrijf dat me heel erg aansprak (meer dan het CWI eigenlijk). We begonnen met elkaar te praten. Ik had een hele waslijst met vragen, ingegeven door mijn ervaring tijdens het eerste sollicitatiegesprek. Ik voelde aan dat P&O een moeilijke positie innam binnen het CWI en ik wilde graag weten hoe jij daarin stond. De verwachtingen waren hoog gespannen over wat de nieuwe manager P&O allemaal moest gaan doen. Het was geen gespreid bedje. Het gesprek liep bijna op zijn einde en opeens zei jij woorden van ongeveer deze strekking: "Ik ben er uit, ik wil je wel hebben".
Met deze directheid overviel je mij totaal en ik bekende toen ook eerlijk dat ik nog geen "ja" kon zeggen. Ik had, zoals gezegd, nog een andere sollicitatie lopen en wilde dat gesprek afwachten. Op de weg terug naar huis belde ik Gert-Jan en vertelde ik over het CWI en over jou en ik merkte dat ik steeds enthousiaster werd. Gert-Jan concludeerde dat ik mijn keuze al had gemaakt en dat klopte. Ik ben zelfs niet meer naar dat andere sollicitatiegesprek gegaan! Voor jou wilde ik werken en waarom? Het voelde goed, je humor, je oprechtheid, eerlijkheid en gedrevenheid. Dus ik zei "ja" en sprong het diepe in ..... En dat viel echt niet mee in het begin. Met vallen en opstaan heb ik nu het gevoel dat ik mijn plek gevonden heb.
We hebben samen veel meegemaakt op het CWI, intensieve processen doorleefd en doorstaan. Het voelde af en toe als bloed, zweet en tranen en

juist op die momenten bracht jij jouw humor en relativeringsvermogen ten tonele en voelde alles veel lichter als ik je kamer uitliep. Zoals jij altijd tegen mij en anderen zei: "Het is hier geen koekjesfabriek". Veel overleggen en sparren met elkaar. Je stond altijd open voor mijn mening of advies. Soms waren we het eens met elkaar en soms niet. Je gaf me ruimte, vertrouwen en loyale steun, dé ingrediënten die ik nodig heb om goed te kunnen functioneren. Al ben je niet het type leidinggevende dat veel complimenten geeft, toch voelde ik af en toe dat je blij met me was, vooral wanneer je tegen mensen aan het praten was en mij met een bepaalde toon 'mevrouw Personeel' noemde.

Wat zijn voor mij nog meer typische Jan Karel oftewel JKL herinneringen? Voor jou was het belangrijk dat jouw actiepunten op jouw actielijstje werden weggestreept. Zo hield je overzicht op de zaken die je moest doen. Ik denk wel Jan Karel, dat je in de laatste jaren hebt geleerd dat iets wegstrepen op jouw lijstje niet meteen betekent dat het is afgehandeld. En wat heb ik af en toe genoten van de momenten dat je helemaal in extase kon zijn over een bepaalde zinsnede, woordspeling of tabel die je zelf had bedacht of die iemand anders had gemaakt. Je probeerde me soms al tekenend wiskundige zaken uit te leggen en heel braaf luisterde ik en probeerde ik het te snappen. Ik moet nu maar eerlijk aan je bekennen dat ik het vaak niet begreep. Je liefde voor boeken niet te vergeten. Je reist er het liefste de hele wereld voor af en geld speelt daarbij geen rol. De keer dat ik je per se moest spreken voor een heikele zaak en jij in Noorwegen de kettingzaag moest stopzetten om mij te woord te staan heeft erg veel indruk op me gemaakt. En wat kun je goed speechen, petje af, hoe je iedere keer weer een vlotte tekst in elkaar zet met een vleugje humor er in. De gezellige momenten tijdens een borrel of diner waar we dan uitgebreid staan te kletsen over belangrijke en minder belangrijke zaken. De foto's die je ziet zijn gemaakt tijdens je laatste officiële optreden voor het CWI, de BBQ. En ja, ik moet nu een einde maken aan mijn tekst en afscheid van je nemen. Daar ben ik nooit zo goed in. Ik hoop dat je nog een fijn jaar binnen het CWI beleeft waarin je in meer rust aan schrijven toekomt en aan andere zaken die je belangrijk vindt voor de wetenschap. Ik ga je missen.....

Liefs, Angelique

Eton Mess

Hoe komt het dat het statige, eerbiedwaardige Engelse College Eton in één adem wordt genoemd met het woord 'mess'?

Eton Mess is een klassiek Engels dessert, dat traditioneel wordt geserveerd op de jaarlijkse cricketwedstrijd van Eton College tegen Winchester College. Het verhaal wil dat een labrador bovenop de picknickmand ging zitten. Dit toetje van slagroom, aardbeien en gebroken meringue verdient dan ook geen schoonheidsprijs, maar wel één voor smaak.

Ingrediënten
1 grote meringue (ca. 75 gram)
250 ml slagroom
500 gram aardbeien
½ zakje vanillesuiker
eventueel aardbeienlikeur

Dit zijn de basisverhoudingen. Meer meringues betekent dat het zoeter wordt.
Gebruik witte en roze meringues voor het verrassingseffect.
Gebruik mocca meringues voor een heel andere, eveneens heerlijke smaak.
Eventueel aardbeien marineren in aardbeienlikeur.

Bereiding
Aardbeien wassen en in hapklare stukjes snijden.
Meringue in ongelijke stukjes breken.
Slagroom met vanillesuiker lobbig kloppen.
Aardbeien en stukjes meringue door elkaar scheppen.
Slagroom voorzichtig door mengsel van aardbeien en meringue roeren.
Eventueel wat aardbeien bewaren om bovenop te leggen.

En hoe komt het dat dit recept een plaats verdient in Jan Karels Liber?

Wij van P&O zijn heel goed in het voorkomen c.q. opruimen van 'mess', omdat wij als geen ander weten dat onze statige, eerbiedwaardige algemeen directeur niet van 'mess' houdt. Toch zijn wij ervan overtuigd dat Jan Karel déze 'mess' wel waardeert.

Karin van Gemert
Marlin van der Heijden
Irma van Lunenburg
Léon Ouwerkerk

Van de wonderwerkster

Beste Jan Karel,

In 1987 kwam ik via een uitzendbureau terecht bij het CWI, als secretaresse BS & MS. De sfeer was prima, ik voelde me snel thuis en was dan ook blij dat ik kon blijven. Jij was in die tijd hoofd van de afdeling BS en bepaalde het merendeel van mijn werkzaamheden. Je was niet altijd de makkelijkste baas, erg precies en best veeleisend, maar ik had meestal wel het gevoel net aan je eisen te kunnen voldoen. De bevestiging daarvan kwam toen ik in 1989 het CWI verliet om me te gaan wijden aan het moederschap. In je afscheidstoespraak noemde je me een wonderwerkster, een bijzonder compliment dat grote indruk op me maakte, als ik af en toe een zetje nodig heb dan denk ik daar nog wel eens aan.

Ik kwam in 1993 terug op het CWI en was na wat omzwervingen binnen het instituut terecht gekomen bij de communicatie- en publicatiedienst toen jij in 2003 terugkwam in de rol van directeur. Door omstandigheden hadden we meer contact dan ik in eerste instantie had verwacht. Nu vele afdelingsnamen, managers en herverdelingen later doe jij een stapje terug om al je tijd te kunnen gaan besteden aan onderzoek en ik hoop bij de vernieuwde I ook in rustiger en vooral stabieler vaarwater te komen.

We zullen het allemaal zien. In ieder geval heb ik altijd met heel veel plezier voor en met je gewerkt. Wel zal ik bij ieder streepje blijven denken: is dat wel de juiste lengte? Bij ieder aanhalingsteken: Jan Karel zou deze beslist anders willen hebben. En misschien krijg ik ooit nog eens oog voor cursieve punten, maar ik hoop eigenlijk van niet.

Groetjes, Lieke

Jan Karel, 8 jaar CWI, het zit erop. Het zal niet altijd gemakkelijk geweest zijn, maar het geheel overziend hopen we dat het directeurschap je veel plezier en voldoening heeft gegeven en dat je later nog eens in de lach zult schieten bij alle herinneringen. Wij zullen de gesprekjes, je humor en je leuke besprakenmissen.

We weten hoe ordelijk je bent, maar mocht je toch hulp nodig hebben bij de laatste loodjes dan staan we voor je klaar!

Tot ziens, het liefst een keer op de koffie in Utrecht!

Martine Susanne Nada

Jan Karel,

Je speeches hield je altijd kort: "Less is more". Ik zal dit tekstje ook kort houden.

We zagen elkaar voor het eerst tijdens het NMC2001 op de VU. Kort daarop vroeg je me als bestuurslid van het Wiskundig Genootschap met als portefeuille PR – geweldig leuk! We bezochten samen vormgeefster Kitty Molenaar voor het nieuwe WG-drukwerk, mooi vormgegeven. Als ik me niet vergis zette jij ook het traject in werking om het WG 'koninklijk' te maken. Dat gebeurde uiteindelijk bij het 225-jarig bestaan in Nijmegen, met Guusje ter Horst en Nobelprijswinnaar Martin 'Tiny' Veltman – een feest! Bij je afscheid van het WG nodigde je het hele bestuur thuis uit; je liet ons je mooiste muziek horen, de bijzondere ordening van je cd's zien (op uitvoerend musicus), je Rembrandt-ets en een indrukwekkende rij boeken over lettertypes. Toen je op het CWI aangesteld werd, mocht ik het persbericht schrijven en je als een van de eersten feliciteren. We verheugden ons.

Over ERCIM en ERCIM News herinner ik me een taxirit naar INRIA met jou en prof. Baayen, met boeiende discussies. Als directeur probeerde je een groot aantal 'smaken' uit wat betreft hoofden communicatie: Erik Stokhof, jijzelf (a.i.) Godelief Nieuwendijk, Jos Besteman, Monique Bekkenutte, Judith Coster, jijzelf (a.i.), Marga Sutherland en jijzelf (a.i.). Zou je soms weer onderzoeker worden omdat je anders de langstzittende communicatiemanager van het CWI bent? In elk geval: als ik teksten terugkreeg met als enige aanmerking de cursieve apostrof, dan wist ik dat het goed zat.

Jan Karel, het ga je goed!

Veel groeten ook aan
Karen en de kinderen,

Annette Kik

V.l.n.r. Seppo Valli (VTT), Annette Kik, Jan Karel Lenstra en Seppo Linnainmaa (VTT), ERCIM meeting oktober 2006, Nice

A leisurely traveling salesman tour through several facets of cuts

Monique Laurent

CWI & Tilburg University

Foreword

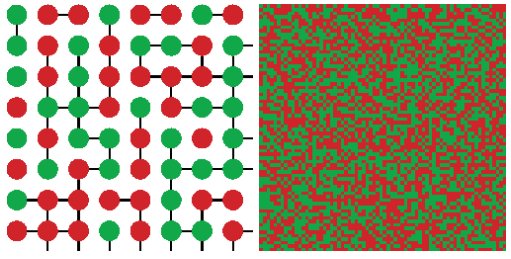
Beste Jan Karel,

Dit is een kort verhaal over sneden (cuts) in grafen, als afscheidskado en dank voor al je werk en inzet voor het CWI. Hopelijk heb jij voortaan meer tijd en kunnen we vaker van je inzicht over allerlei 'traveling salesman' vraagstukken genieten!

Veel dank en ik wens je veel plezier in je nieuwe leven na de CWI directie.

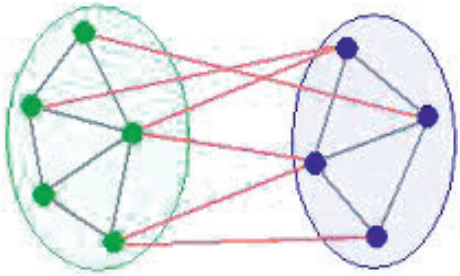
Prologue

Cuts in graphs are very basic objects. However, they often illustrate in a beautiful way several recent developments in combinatorial optimization, and they have fascinating connections to other mathematical areas like metric theory, distance geometry, combinatorial matrix theory. My co-authored book *Geometry of Cuts and Metrics* which just appeared when I joined CWI in 1997, covers some of them. [As an aside, this earned me my first experience with the generous and thoughtful work environment at CWI!]



Cuts have numerous applications, for example, to cluster analysis (cluster a set of data into groups of most similar objects), to statistical physics (minimize the energy of spin glass systems), or to VLSI design (minimize the number of 'vias' needed to route nets on several layers in chip design). My focus here will be to travel from a bird-eye perspective through some chosen mathematical questions and results where cuts play a central role. This is not a scientific paper, there is no pretention in completeness, and I will mostly omit to say precisely who has done what.

Complexity of max-cut



Given a graph $G = (V = [n], E)$, a *cut* is the set of edges that are 'cut' by some partition of the nodes into two groups. Say $(S, V \setminus S)$ is such a partition, then the corresponding cut consists of all edges $ij \in E$ with $i \in S$ and $j \notin S$ (or conversely).

If weights $w \in \mathbb{R}^E$ are assigned to the edges of G, then the *max-cut problem* asks to find a cut whose weight $\sum_{i \in S, j \notin S} w_{ij}$ is maximum. Two remarks are in order about the complexity of max-cut.

First, if all weights are non-negative and, instead of a maximum cut, one wants a *minimum weight non-empty cut*, then the problem is polynomial time solvable, using maximum flow techniques and Fulkerson's celebrated max-flow min-cut theorem.

On the other hand, max-cut is NP-hard, even if all weights are equal to 1 and the graph is cubic. Karp [1972] proposed the following reduction of max-cut from the partition problem: Given integers $a_1, \ldots, a_n \in \mathbb{N}$, define the edge weights $w_{ij} = a_i a_j \ \forall i, j$. Then the weight of a cut is $(\sum_{i \in S} a_i)(\sum_{i \notin S} a_i) \leq (\sum_{i \in V} a_i)^2/4$, with equality if and only if $\sum_{i \in S} a_i = (\sum_{i \in V} a_i)/2$, i.e., the sequence $a_1, \ldots, a_n$ can be partitioned.

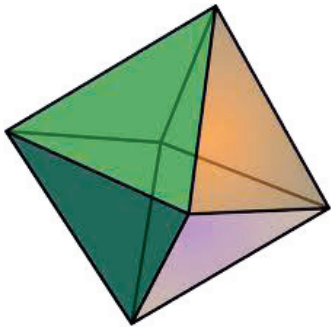
However, max-cut can be solved in polynomial time over the class of planar graphs; working with the planar dual graph, the problem can indeed be reduced to shortest path combined with maximum weight matching computations. Polynomial time solvability extends to the larger class of graphs which do not contain a K_5 minor.

Cut polyhedra

Define the *cut polytope* $\mathrm{CUT}_n^{\square}$ as the convex hull of the incidence vectors of all cuts in the complete graph K_n. Thus solving max-cut amounts to linear optimization over the cut polytope. Define analogously the *cut cone* CUT_n as the conic hull of the incidence vectors of all cuts. A beautiful symmetry property is that all the facets of the cut polytope can derived by a simple 'switching' operation from the facets of the cut cone (where 'switching' means changing signs along a cut). Hence it suffices to find the facets passing through the origin... Well, as max-cut is NP-hard, this remains probably a hopeless task! Nevertheless many classes of facets are known. In particular, facets arise within the class of hypermetric inequalities mentioned in (2) below. It is known that the subclass of triangle inequalities (3) (in fact, their projections) suffices to describe the cut cone for graphs with no K_5 minor. Gaining good information about the facets of cut polyhedra is important for branch-and-cut type algorithms for max-cut, but it is also relevant to other fields, including ℓ_1-metrics, or the study of Bell inequalities in quantum information theory, as mentioned below.

Cuts and ℓ_1-metrics

Any cut corresponds to a very simple distance on V: if two points of V are cut they are at distance 1 and otherwise they are at distance 0. Call a distance d ℓ_1-*realizable* if there exist vectors $u_1, \ldots, u_n$ in some space $\mathbb{R}^k$ satisfying $d_{ij} = \|u_i - v_j\|_1 \quad \forall i, j$. Clearly, any cut corresponds to an ℓ_1-realizable distance (then the u_i's can even be chosen within $\{0, 1\}$) and one can show that the distances that are ℓ_1-realizable are precisely the members of the cut cone CUT_n. Thus, the hardness of max-cut implies that testing ℓ_1-realizability is also a hard problem.



The octahedron: The unit ball of the ℓ_1-norm in 3D

In comparison, testing ℓ_2-realizability can be done in polynomial time - using semidefinite programming. Indeed, there exist vectors $u_1, \ldots, u_n \in \mathbb{R}^k$ $(k \geq 1)$ such that $\|u_i - u_j\|_2 = d_{ij} \quad \forall i, j$ if and only if there exists a positive semidefinite matrix X satisfying the linear conditions:

$$X_{ii} + X_{jj} - 2X_{ij} = d_{ij}^2 \ \forall i, j \in [n]. \tag{1}$$

Nevertheless the above characterization of ℓ_1-metrics as members of the cut cone implies that any valid inequality for the cut cone gives a necessary condition for ℓ_1-realizability. Among the many known classes of valid inequalities, the class of *hypermetric inequalities* is of particular interest. Given integers $b_1, \ldots, b_n \in \mathbb{Z}$ with $\sum_i b_i = 1$, the corresponding hypermetric inequality is

$$\sum_{1 \leq i < j \leq n} b_i b_j d_{ij} \leq 0. \tag{2}$$

For instance, for $b_i = b_j = 1 = -b_k$ and all other entries equal to 0, we get the *triangle inequalities*, claiming that d is a metric:

$$d_{ij} \leq d_{ik} + d_{jk} \tag{3}$$

(thus motivating the name *hyper*metric). For $b = (1, 1, 1, -1, -1, 0, \ldots, 0)$, we get the *pentagonal inequality*:

$$d_{12} + d_{13} + d_{23} + d_{45} - \sum_{i \in \{1,2,3\}, j \in \{4,5\}} d_{ij} \leq 0. \tag{4}$$

Although there are infinitely many hypermetric inequalities, the cone they define can yet be shown to be a polyhedral cone. This result is based on a deep connection between hypermetrics and geometry of numbers (Voronoi and Delaunay polytopes in lattices).

Semidefinite approximation for max-cut

Assume all edge weights are non-negative. Here is a simple random algorithm to get a good cut: Assign the nodes independently, with probability 1/2, to either side of the cut. Then an edge is cut with probability 1/2 and one gets a cut whose expected weight is half the total weight of edges, and thus at least half max-cut.

How to do better?

Goemans and Williamson [1995] introduced an ingenious construction for constructing a better approximate cut, based on semidefinite programming combined with a clever 'random hyperplane rounding' technique:

1. Find unit vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ maximizing $\sum_{ij \in E} w_{ij}(1 - v_i^T v_j)$.
2. Pick a random hyperplane, say with normal r.
3. This hyperplane induces a partition of the space, and thus a cut, depending on the sign of $r^T v_i$.

The beautiful property of this construction is that the expected weight of the obtained cut is at least 0.878 times max-cut.

The key fact in the analysis is that the probability that edge ij is cut depends only on the angle between vectors v_i, v_j, namely it is equal to $\arccos(v_i^T v_j)/\pi$. Hence the expected weight of the cut is equal to

$$\sum_{ij \in E} w_{ij} \frac{\arccos(v_i^T v_j)}{\pi} = \underbrace{\sum_{ij \in E} w_{ij} \frac{1 - v_i^T v_j}{2}}_{:=\text{SDP}} \underbrace{\frac{2}{\pi} \frac{\arccos v_i^T v_j}{1 - v_i^T v_j}}_{\geq \alpha_{\text{GW}}} \geq \text{ max-cut } \alpha_{\text{GW}},$$

where $\alpha_{\text{GW}} := \min_{\vartheta \in [0,\pi]} \frac{2}{\pi} \frac{\vartheta}{1 - \cos \vartheta} \sim 0.878$.

Indeed, the value SDP is at least max-cut. To see it, note that one can formulate max-cut as the following integer quadratic problem:

$$\text{max-cut} = \max \sum_{ij \in E} w_{ij}(1 - x_i x_j)/2 \;\; \text{s.t. } x_1, \ldots, x_n \in \{\pm 1\}. \tag{5}$$

Relaxing the condition: $x_i \in \{\pm 1\}$ by the condition: v_i lies in the unit sphere S^{n-1}, we get the following 'vector relaxation' of max-cut:

$$\text{SDP} := \max \sum_{ij \in E} w_{ij}(1 - v_i^T v_j)/2 \;\; \text{s.t. } v_1, \ldots, v_n \in S^{n-1}. \tag{6}$$

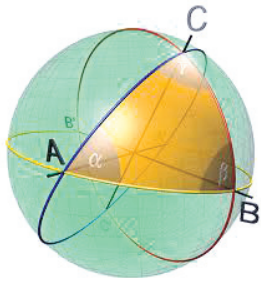
Letting $X = (v_i^T v_j)$ denote the Gram matrix of the v_i's, (6) can be reformulated as the following *semidefinite program:*

$$\text{SDP} = \max \sum_{ij \in E} w_{ij}(1 - X_{ij})/2 \;\; \text{s.t. } X_{ii} = 1 \; \forall i, \; X \succeq 0. \tag{7}$$

Thus the value SDP, together with optimal vectors v_i's for (6), can be computed in polynomial time (to arbitrary precision) using, e.g., interior point algorithms.

The Goemans-Williamson approximation $\alpha_{\text{GW}} \sim 0.878$ improves the trivial 1/2 approximation. As of today, this remains the best known polynomial-time approximation guarantee for max-cut. Whether the GW approximation can be improved is intimately related to the *Unique Games conjecture* of Khot [2002] in complexity theory. Namely, if the Unique Games conjecture holds, then the Goemans-Williamson approximation is best possible.

A trigonometric puzzle about angles of vectors



A spherical triangle:
$\alpha + \beta + \gamma \leq 2\pi$

Let $v_1, \ldots, v_n$ be given vectors on the unit sphere S^{n-1}. A natural question is what are the linear conditions satisfied by their angles ϑ_{ij} $(= \arccos v_i^T v_j)$.

A basic well known geometric fact in 3D is that the pairwise angles of three vectors v_1, v_2, v_3 satisfy:

$$\begin{aligned} \vartheta_{12} &\leq \vartheta_{13} + \vartheta_{23}, \\ \vartheta_{13} &\leq \vartheta_{12} + \vartheta_{23}, \\ \vartheta_{23} &\leq \vartheta_{12} + \vartheta_{13}, \end{aligned} \tag{8}$$

$$\vartheta_{12} + \vartheta_{13} + \vartheta_{23} \leq 2\pi. \tag{9}$$

But, what about the pairwise angles of five vectors ? Is it true that they satisfy:

$$\vartheta_{12} + \vartheta_{13} + \vartheta_{23} + \vartheta_{45} \leq \sum_{i \in \{1,2,3\}, j \in \{4,5\}} \vartheta_{ij} \tag{10}$$

as well as

$$\sum_{1 \leq i < j \leq 5} \vartheta_{ij} \leq 6\pi \; ? \tag{11}$$

Well, the answer is **yes** and the above 'random hyperplane rounding' argument permits to show this.

Indeed, say we have an inequality $\sum_{ij} w_{ij} d_{ij} \leq w_0$, valid for $\text{CUT}_n^{\square}$, and a collection of unit vectors $v_1, \ldots, v_n$. As in steps 2-3 of the Goemans-Williamson procedure described

above, pick a random hyperplane and consider the corresponding random cut, whose expected weight is equal to $\sum_{ij} w_{ij}\vartheta_{ij}/\pi$, and is at most the maximum value w_0. This shows that $\sum_{ij} w_{ij}\vartheta_{ij} \leq w_0\pi$ holds for the pairwise angles of unit vectors. As an illustration, relation (8) corresponds to the triangle inequality (3), relation (10) corresponds to the pentagonal inequality (4), and (11) corresponds to the switched inequality, claiming that a cut in K_5 has at most six edges.

Spectrahedra and the elliptope

While the feasible region of a linear program is a polyhedron, the term *spectrahedron* has been coined for the feasible region of a semidefinite program. In particular, the feasible region of the semidefinite relaxation (7) of max-cut is a spectrahedron, called the *elliptope* and denoted as $\mathcal{E}_n$.

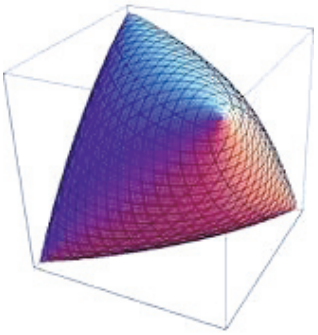


Fig. 1: The elliptope $\mathcal{E}_3$

The elliptope $\mathcal{E}_n$ consists of all $n \times n$ positive semidefinite matrices with diagonal entries one. Such matrices correspond to correlation matrices in statistics and partly for this reason their geometric properties have been much studied. For instance, it is known that a correlation matrix which is an extreme point of the elliptope $\mathcal{E}_n$ has rank r satisfying $\binom{r+1}{2} \leq n$. It is also known that the elliptope, although it is a non-polyhedral convex set, does have polyhedral faces, some of them being inherited from the cut polytope. The possible dimensions for the polyhedral faces of $\mathcal{E}_n$ have been characterized.

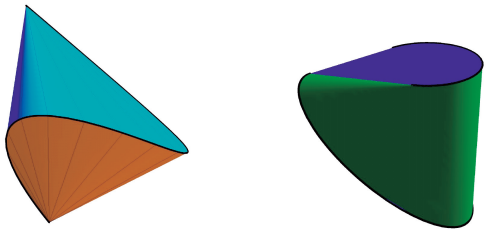
By construction, the elliptope $\mathcal{E}_n$ is a convex relaxation of the cut polytope $\mathrm{CUT}_n^{\square}$. For $n = 3$, there are four cuts in K_3 and the cut polytope $\mathrm{CUT}_3^{\square}$ is a simplex in $\mathbb{R}^3$. Fig. 1 shows the elliptope $\mathcal{E}_3$. Notice that it has four 'corners', corresponding to the four cuts. Also, $\mathcal{E}_3$ can be seen as the geometric shape obtained by 'inflating' the cut polytope while keeping its skeleton rigid (which earns $\mathcal{E}_3$ its alternative name: 'the pillow').

A basic natural question is to understand what are the characteristic properties of spectrahedra. By definition, a spectrahedron K is the solution set of a semidefinite program. That is, K can be written as the set of solutions $x \in \mathbb{R}^n$ of a so-called *linear matrix inequality (LMI)*:

$$A(x) := A_0 + x_1A_1 + \ldots + x_nA_n \succeq 0. \tag{12}$$

Clearly, any spectrahedron K must be convex, and a basic closed semi-algebraic set, which means that K can be described by finitely many polynomial inequalities (indeed, simply write that the principal minors of the matrix $A(x)$ must be non-negative). A deep algebraic result of Helton and Vinnikov [2006] characterizes spectrahedra in $\mathbb{R}^2$ (in terms of convex rigidity), but a characterization is not known in dimension $n \geq 3$.

The next figure shows another spectrahedron: the set $\mathcal{T}_4$ of all Toeplitz matrices (i.e., with identical entries on each diagonal) in the elliptope $\mathcal{E}_4$, together with its dual body. Alternatively, $\mathcal{T}_4$ is the convex hull of the trigonometric curve $(\cos(t), \cos(2t), \cos(3t))$ for $t \in \mathbb{R}$.



For the purpose of optimization one is also interested in working with projections of spectrahedra, which corresponds to lifting the problem in higher dimension. Call a set K *SDP representable* if it can be obtained as the projection of a spectrahedron. That is, there is an LMI:

$$A(x,z) = A_0 + x_1A_1 + \ldots + x_nA_n + z_1B_1 + \ldots + z_mB_m$$

such that $x \in K$ if and only if $A(x,z) \succeq 0$ for some $z \in \mathbb{R}^m$. Necessarily K must be convex and semi-algebraic, but again no full characterization is known. Characterizing SDP representable sets is a major open problem in the field of semidefinite programming and convex algebraic optimization.

As an illustration, the unit ball in $\mathbb{R}^n$ of the ℓ_2-norm is a spectrahedron, as

$$\left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 \leq 1 \right\} = \left\{ x \in \mathbb{R}^n \mid \begin{pmatrix} 1 & x^T \\ x & I_n \end{pmatrix} \succeq 0 \right\}.$$

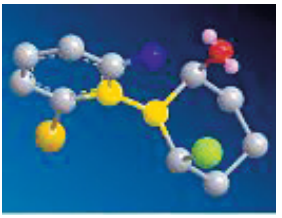
In optimization terms, optimizing over the second order cone, also known as the Lorentz cone or the ice-cream cone: $\left\{(x_0, x) \in \mathbb{R}^{n+1} \mid x_0 \geq \sqrt{\sum_{i=1}^n x_i^2}\right\}$, can be cast as an instance of semidefinite programming. On the other hand, the unit ball in $\mathbb{R}^2$ of the ℓ_4-norm:

$$K = \{(x,y) \in \mathbb{R}^2 \mid x^4 + y^4 \leq 1\},$$

is not a spectrahedron (since it is not rigidly convex), but it is SDP representable. Indeed K can be realized as the projection of a spectrahedron by adding two more variables. Namely, $(x,y) \in K$ if and only if there exists $(u,v) \in \mathbb{R}^2$ satisfying $x^2 \leq u$, $y^2 \leq v$, $u^2 + v^2 \leq 1$, which can equivalently be formulated using the LMI's:

$$\begin{pmatrix} 1 & x \\ x & u \end{pmatrix} \succeq 0, \quad \begin{pmatrix} 1 & y \\ y & v \end{pmatrix} \succeq 0, \quad \begin{pmatrix} 1 & u & v \\ u & 1 & 0 \\ v & 0 & 1 \end{pmatrix} \succeq 0.$$

Distance geometry and matrix completions



The *distance geometry problem* can be stated as follows: Given a graph $G = (V, E)$ and a (partial) set of distances $d = (d_{ij})_{ij \in E}$, find (if it exists) a realization of d in $\mathbb{R}^k$, i.e., find vectors $u_1, \ldots, u_n \in \mathbb{R}^k$ such that $d_{ij} = \|u_i - u_j\|_2$ for all $ij \in E$. This contains, for instance, the molecular conformation problem in molecular chemistry, where one wants to reconstruct the 3D structure of a molecule from the knowledge of certain interatomic distances, or sensor network localization problems.

If the dimension k is fixed and part of the input, then the problem is NP-hard (already for $k = 1$, where there is an easy reduction from the partition problem). However, the exact complexity of deciding existence of an ℓ_2-realization in any dimension $k \geq 1$ is not known. Note that this amounts to deciding whether the semidefinite program:

$$X_{ii} + X_{jj} - 2X_{ij} = d_{ij}^2 \ (ij \in E), \ X \succeq 0$$

(already encountered in (1) above) has a feasible solution. However, the exact complexity of deciding feasibility of a semidefinite program is not known. Let us make a brief excursion to the complexity of semidefinite programming.

Consider the problem of testing feasibility of the LMI (12), where $A_1, \ldots, A_n$ are given integer symmetric matrices. Clearly, this problem belongs to NP *in the real number model of computation*, since one can test whether a matrix is positive semidefinite in polynomial time - using Gaussian elimination. However, the problem is not known to be in NP *in the bit number model of complexity* (although Ramana [1997] could show that it belongs to NP if and only if it belongs to co-NP - using a clever extended duality theory). It is indeed not clear how to come up with a short certificate. For instance, the following semidefinite program:

$$\begin{pmatrix} 2 & x \\ x & 1 \end{pmatrix} \succeq 0, \ \begin{pmatrix} x & 2 \\ 2 & 2x \end{pmatrix} \succeq 0$$

admits a unique *irrational* solution: $x = \sqrt{2}$. Deciding the complexity status of testing feasibility of a semidefinite program is a major open problem. On the positive side, Khachiyan and Porkolab [1997] have shown that one can test existence of a rational solution of the LMI (12) in polynomial time when *fixing the number n of variables*, thus extending the result of Lenstra [1983] about polynomial time solvability of integer LP in fixed dimension to SDP.

On the positive side, the distance geometry problem can be solved in polynomial time for the class of chordal graphs, and for the class of graphs with no K_4 minor. The latter result relies on the following observation. Clearly, if d has an ℓ_2-realization then it can be

completed to a metric $\hat{d}$ on V; moreover one can test in polynomial time whether such a metric completion exists. It turns out that existence of a metric completion is also sufficient to ensure existence of an ℓ_2-realization in the case when the graph has no K_4 minor. More strongly, if d has an ℓ_2-realization, then it has a metric completion $\hat{d}$ which satisfies *all* valid inequalities for the cut cone CUT_n. (In sophisticated terms, ℓ_2-realizability implies ℓ_1-realizability.)

Two examples of distances with no ℓ_2-realization are shown in Fig. 2. As a first example, consider the case where $G = C_4$ is a circuit of length 4 and d satisfies: $d_{12} = d_{23} = 1$, $d_{34} = 3$ and $d_{14} = 0$. Then d cannot have an ℓ_2-realization since d has no metric completion $\hat{d}$, for otherwise one would have $\hat{d}_{13} \leq d_{12} + d_{23} = 2$ and $d_{34} \leq d_{14} + \hat{d}_{13}$ implies $\hat{d}_{13} \geq 3$, yielding a contradiction.

As a second example, consider the distance d defined on the graph $G = K_4$ by $d_{12} = d_{13} = d_{14} = 1$ and $d_{23} = d_{24} = d_{34} = 2$. Then d is a metric (and moreover d belongs to CUT_4, since all the facets of CUT_4 come from triangle inequalities). However d does not have an ℓ_2-realization; for otherwise one could find vectors $u_1 = 0$, u_2, u_3, u_4 satisfying $\|u_i\| = 1$ $(i = 2, 3, 4)$ and $\|u_i - u_j\| = 2$, i.e., $u_i^T u_j = -1$ $(i \neq j \in \{3, 4\})$, clearly a contradiction.

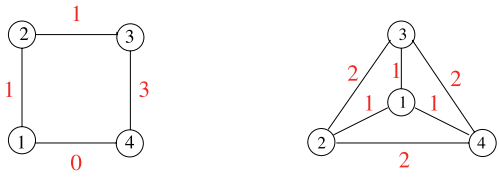


Fig. 2: Two distances with no ℓ_2-realization

The distance geometry problem is closely related to the *PSD matrix completion problem:* Decide whether a partially specified matrix can be completed to a positive semidefinite matrix. In other words, given $a \in \mathbb{R}^E$, is it possible to find a matrix $X \succeq 0$ such that $X_{ii} = 1$ $\forall i \in V$ and $X_{ij} = a_{ij}$ $\forall ij \in E$? That is, decide whether a belongs to the projection $\mathcal{E}(G)$ of the elliptope $\mathcal{E}_n$ onto the edge subspace $\mathbb{R}^E$. Again the exact complexity is not known. An obvious necessary condition for $a \in \mathcal{E}(G)$ is that the vector $\arccos(a) \in [0, \pi]^E$ can be completed to a spherical distance. More strongly, following our earlier discussion on 'angles of vectors', this spherical distance should satisfy all valid inequalities for the cut polytope. As an illustration, the following partial matrix:

$$\begin{pmatrix} 1 & \mathbf{1} & ? & \mathbf{-1} \\ \mathbf{1} & 1 & \mathbf{1} & ? \\ ? & \mathbf{1} & 1 & \mathbf{1} \\ \mathbf{-1} & ? & \mathbf{1} & 1 \end{pmatrix}$$

cannot be completed to a positive semidefinite matrix. This can be checked directly. Alternatively observe that assigning values $(0, 0, 0, \pi)$ $(= \arccos(1, 1, 1, -1))$ to the edges of C_4 gives a partial distance that cannot be completed to a spherical distance.

Grothendieck inequalities

Given a graph $G = (V = [n], E)$ and edge weights $w \in \mathbb{R}^E$, consider the quadratic integer problem:

$$\mathrm{IP}(w) := \max_{x_i \in \{\pm 1\}} \sum_{ij \in E} w_{ij} x_i x_j \tag{13}$$

and its semidefinite relaxation:

$$\mathrm{SDP}(w) := \max_{u_i \text{ unit vectors}} \sum_{ij \in E} w_{ij} u_i^T u_j. \tag{14}$$

Clearly, $\mathrm{IP}(w) \leq \mathrm{SDP}(w)$. Alon et al. [2006] showed the existence of a constant $K(G)$ such that

$$\mathrm{SDP}(w) \leq K(G)\ \mathrm{IP}(w) \quad \forall w \in \mathbb{R}^E.$$

Note that the program (13) corresponds to optimizing over the cut polytope $\mathrm{CUT}^{\square}(G)$, while (14) corresponds to optimizing over the ellitope $\mathcal{E}(G)$ (the respective projections of $\mathrm{CUT}_n^{\square}$ and $\mathcal{E}_n$ onto the edge subspace $\mathbb{R}^E$). Thus the constant $K(G)$ can be viewed as the integrality gap of the SDP relaxation or, in geometric terms, as the smallest dilation λ for which $\mathcal{E}(G) \subseteq \lambda\ \mathrm{CUT}(G)$. For instance, $K(G) = 3/2$ for $G = K_3$ (recall the picture of $\mathcal{E}(K_3)$ in Fig. 1).

The constant $K(G)$ is known as the *Grothendieck constant* of the graph G. Indeed, this concept goes back to work of Grothendieck[1] [1953] who showed that $K(G)$ is universally bounded for bipartite graphs. More precisely it is known that

$$1.6769 \leq \sup_{G \text{ bipartite}} K(G) < 1.8782.$$

These results were shown in the context of operator norms in functional analysis and can be used to design approximation algorithms, for instance, for the problem of computing the cut norm of a matrix.

More generally, Alon et al. [2006] show the existence of two absolute constants $C, C' > 0$ such that

$$C \log \omega(G) \leq K(G) \leq C' \log \vartheta(\overline{G}),$$

Alexander Grothendieck

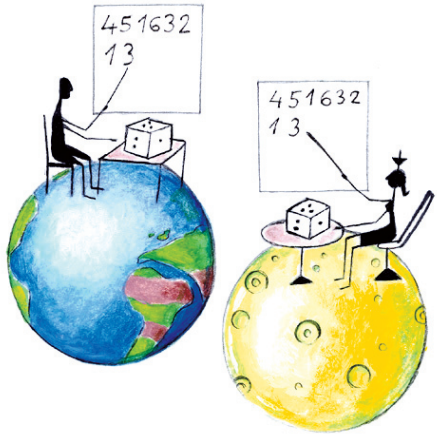
[1] Grothendieck left the academic world since the 1980's and is now said to live in reclusion in southern France. In January 2010, he wrote a "Declaration d'intention de non-publication", where he asks that none of his work should be reproduced in whole or in part, and even further that libraries containing such copies of his work should remove them.

where $\omega(G)$ is the largest size of a clique in G and $\vartheta(\overline{G})$ is the celebrated *Lovász theta number*, known to be upper bounded by $\chi(G)$, the coloring number of G. As $\chi(G) = 2$ when G is a bipartite graph, this implies again Grothendieck's result, claiming that the Grothendieck constant is universally bounded for bipartite graphs.

Interest in the Grothendieck constant comes, in particular, from its relevance to approximation algorithms in complexity theory, as well as to non-local games in quantum information theory, briefly discussed below.

Bell inequalities in quantum information

A non-local game $\mathcal{G} = (\pi, V)$ involves three parties: the *referee* and two *players*, Alice and Bob. Data include a probability $\pi : S \times T \to [0,1]$ and a predicate $V : A \times B \times S \times T \to \{0,1\}$ giving the rules of the game. Here, A, B, S, T are given finite sets, S, T are the sets of 'questions' sent by the referee to Alice and Bob and A, B are the sets of 'answers' of Alice and Bob, respectively. The players and the referee know π and V. Before the start of the game Alice and Bob may decide on a strategy to play the game, but once the game has started they are not allowed to communicate.



The game goes as follows:

1. The referee picks a pair of questions $(s,t) \in S \times T$ according to the ditribution π and sends question s to Alice and question t to Bob.

2. Alice selects her answer $a \in A$ and Bob selects his answer $b \in B$.

The players win the game if $V(a,b|s,t) = 1$ and loose otherwise. The aim of the players is to choose their own strategy so as to maximize their probability of winning the game. This probability can be computed with the following optimization problem:

$$\max_{P} \sum_{(s,t)\in S\times T} \pi(s,t) \sum_{(a,b)\in A\times B} V(a,b|s,t) P(a,b|s,t) \tag{15}$$

where P corresponds to all possible probability profiles, depending on the strategies adopted by the players.

In the *classical* deterministic case, the players simply choose a map $a : S \to A$ for Alice and $b : T \to B$ for Bob. The corresponding feasible probability profiles then form a polytope, whose bounding faces are supported by the so-called *Bell inequalities*. The optimal value of the corresponding optimization problem (15), called the *classical value* of the game, can be computed via LP.

In the non-classical setting, the players share an entangled strategy. It consists of a quantum system $\{A_s^a \mid a \in A, s \in S\}$ of bounded operators on a Hilbert space H_A for Alice, a quantum system $\{B_t^b \mid b \in B, t \in T\}$ of bounded operators on a Hilbert space H_B for Bob, and a shared entangled state Ψ which is a unit vector in $H_A \otimes H_B$. The operators A_s^a of Alice are required to be Hermitian projectors forming an orthogonal resolution of the identity, and analogously for the operators B_t^b of Bob; that is,

1. $A_s^a A_s^{a'} = \delta_{aa'} A_s^a$, $B_t^b B_t^{b'} = \delta_{bb'} \quad \forall a, a' \in A, \forall b, b' \in B, \forall s \in S, \forall t \in T$,
2. $\sum_{a \in A} A_s^a = I_A$, $\sum_{b \in B} B_t^b = I_B \quad \forall s \in S, \forall t \in T$.

Moreover, $P(a, b|s, t) = \langle \Psi, A_s^a \otimes B_b^t \ \Psi \rangle$. Hence the *quantum value* of the game can be computed via the optimization problem:

$$\max_{\Psi, A_s^a, B_t^b} \sum_{s,t} \pi(s,t) \sum_{a,b} V(a, b|s, t) \langle \Psi, A_s^a \otimes B_b^t \ \Psi \rangle, \tag{16}$$

where the maximum is taken over all possible unit vectors $\Psi \in H_A \otimes H_B$, A_s^a Hermitian operators on H_A, B_t^b Hermitian operators on H_B satisfying the above conditions 1-2.

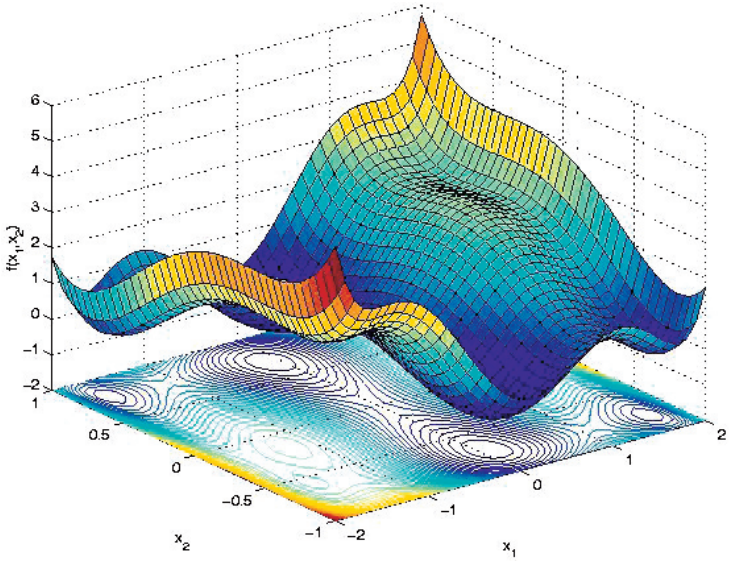
In an XOR game, Alice and Bob have two possible answers, i.e., $A = B = \{0, 1\}$, and the predicate $V(a, b|s, t)$ depends only on the XOR value $a \oplus b$ of their answers. A classical result in quantum information theory is the theorem of Tsirelson [1987], showing that the optimization problem (16) can be reformulated as a semidefinite program, of the form

$$\max_{u_s, v_t \text{ unit vectors}} \sum_{s \in S, t \in T} \pi'(s,t) u_s^T v_t,$$

after suitably defining $\pi'(s, t)$. Note that this problem is of the form (14), where the graph is bipartite with bipartition (S, T). In a nutshell, for XOR games, the classical value of the game can be cast as an instance of the integer quadratic problem (13), while the quantum value of the game can be cast as the SDP relaxation (14), both for the case of bipartite graphs. Hence the Grothendieck's constant also permits to estimate the ratio between the classical and quantum values.

Moreover, the set of feasible probability profiles corresponding to classical strategies form a polytope, closely related to the cut polytope of the complete bipartite graph. In other words, Bell inequalities can be interpreted as valid inequalities for the cut polytope. The maximum violation of a Bell inequality that can be obtained when considering quantum strategies can be computed via semidefinite programming.

For general non-local games the program (16) can be reformulated as a polynomial optimization problem with *non-commutative variables* which can then be approximated by a hierarchy of semidefinite programs. This relies on a recent extension of the techniques developed by Lasserre and Parrilo around 2000 for classical (commutative) polynomial optimization problems, based on sums of squares of polynomials and the dual theory of moments.



Op een goede dag

Toen ik in 1973 in dienst trad van het Mathematisch Centrum, presideerde Jan Karel al hoog en breed in een van de grotere kamers aan de achterzijde van de Tweede Boerhaavestraat. Ook toen al 'vermomd als heer', zoals Jan Karel je pleegt te begroeten als je zelf ook eens ergens met stropdas verschijnt. Hij deelde toen nog zijn kamer met zijn trouwe compagnon Ben Lageweg. Later, toen Jan Karel en ik deeltijdfuncties optimaal poogden te combineren, hebben we lange tijd een werkkamer gedeeld, dankzij Jan Karel steeds in goede harmonie.

Kenmerkend voor Jan Karel is zijn grote loyaliteit. Sta je bij Jan Karel eenmaal aan de goede kant van de streep, dan kom je daar heel moeilijk meer vandaan. Ook na vele beledigingen aan zijn adres blijft hij voor je klaar staan, voor en achter de schermen. Eens zat mijn knie in het gips en steeds als ik een afspraak had, reed hij met zijn wagen bij mij thuis voor en bracht hij mij waar ik wezen moest.

Een tweede kenmerk is zijn uiterste precisie en zijn scherpe oog voor detail. Dankzij een minuscuul handschrift weet hij op een A7'tje te krijgen wat een normaal mens nog niet op een A4 kwijt kan. Zijn boekenkast is geordend naar typografie, en hij is de enige die ik ken die het colofon van een boek uitspelt. Mocht daar een foutje in zijn geslopen dan belandt het boek aan de verkeerde kant van de streep en laat hij het verder ongezien. Dus nu maar hopen dat de redacteuren van deze bundel het colofon secuur hebben gecheckt, anders zullen de bijdragen hun doelwit niet bereiken.

Jan Karels nauwgezetheid gaat samen met een derde kenmerk, zijn grote efficiëntie. Als perfecte regelneef krijgt hij vele ballen en schoteltjes van de grond, en hij houdt er tegelijk nog

vele in de lucht. Doelgericht worden alle items van zijn A7'tje afgewerkt. Deze efficiëntie uit zich ook in de persoonlijke omgang, die mogelijk door nieuwkomers als wat nurks zou kunnen worden ervaren. Aan overbodige beleefdheden en small talk verspilt Jan Karel geen tijd.

Wie schetst dan ook mijn verbazing toen ik op een goede dag aan het werk was en de deur van onze gezamenlijke werkkamer geopend werd door Jan Karel, echter zonder zelf naar binnen te gaan. Even later verscheen een hoogblonde, zeer Scandinavisch ogende schone in de deuropening, en Jan Karel liet haar voorgaan!

Meteen daarop de volgende verrassing: Jan Karel hielp haar uit haar jas! Het was een lange, dikke, witharige mantel, gemaakt van een soort ijsberenvacht, en Jan Karel drapeerde hem zorgvuldig om een van onze klerenhangers.

Vervolgens, om de verrassing compleet te maken: Jan Karel stelde ons aan elkaar voor! De pronte dame bleek Karen Aardal te heten, een naam die ik kende doordat haar triple-A achternaam steeds de deelnemerslijst van elke conferentie aanvoerde.

De twee gingen na deze verwarrende rituelen dicht bijeen aan Jan Karels bureau zitten, haalden een voorraad smørrebrød te voorschijn, en begonnen op fluistertoon uitgebreid te keuvelen en te tortelen. Ik kreeg sterk de indruk dat Karen Aardal niet zo maar een collega was.

Toen ik een paar weken daarna bij Jan Karel thuis werd uitgenodigd voor een diner — hij is ook nog eens een excellente kok — en ik bij de deurbel hun beider namen op het naambordje zag staan, viel, nadat een gelijk arriverende andere gast had uitgeroepen "Nou, dat gaat snel!", alles op zijn plaats. *And the rest is history.*

Lex Schrijver

Amsterdam, 1 juli 2011

Beste Jan Karel,

Graag benut ik deze gelegenheid om bij je afscheid als directeur van het CWI enkele lovende woorden tot jou aan het papier toe te vertrouwen.

Dat is geen plichtpleging, maar oprecht gemeend. Zo mocht ik in de periode 2006-2009 onder jouw voorzitterschap secretaris zijn van ERCOM (voor de niet-ingewijden: ERCOM staat voor “European Research Centres on Mathematics”) en dat heb ik met veel plezier gedaan, in belangrijke mate omdat jij zo'n goeie voorzitter was. Die delegeren het één en ander en dat betekende dat ik tijdens een kort verblijf in Barcelona als gids mocht optreden toen we in een paar verloren uurtjes een museum gingen bezoeken. Dat ging niet helemaal goed want op een bepaald moment stonden we bij een Metro-station terwijl ik meende jou naar een Museum te hebben geleid. Als excuus kan ik aanvoeren dat ik als zuinige secretaris een pocket-versie van de plattegrond van Barcelona hanteerde en dat de logo's voor Metro en voor Museum op dat kaartje erg op elkaar leken: beide bestonden voornamelijk uit de hoofdletter M.

Af en toe zaten we naast elkaar in het vliegtuig naar of van ERCOM-bijeenkomsten, soms vergezeld van Dick Broekhuis. Leuke reizen waren dat, niet in het minst omdat ERCOM eigenlijk (in jouw woorden) een soort dinner's club was (waar vanzelfsprekend heel goed werd gecommuniceerd door de aanwezige directeuren van de ERCOM onderzoeksinstituten).

Wat me dan opviel was dat je in het vliegtuig vaak uiterst minutieus op 1 A4-tje je werk zat te plannen en te organiseren. Ik ben zo vrij geweest zo'n A4-tje uit de tijd dat je bestuurslid was van het Wiskundig Genootschap uit het KWG-Archief op te duiken en hierbij af te drukken. Om precies te zijn: dit overzicht dateert van maandag 14 mei 2001, 23:09 uur. Interessant zijn de drie opties die je bij de actiepunten hanteerde: + *done*, o *to do*, - *forget*. Zelfs wat je wilt vergeten wil je kennelijk nog onthouden! Zeer beknopt en uiterst overzichtelijk en informatief. Toen speelde bijvoorbeeld het Koninklijk worden van het WG (actiepunt JKL) en het naar Amsterdam halen van het Fifth European Congress of Mathematics (actiepunt EJNL), zie p. 2/2 bovenaan. Beide actiepunten zijn tot een goed einde gebracht: het WG is op 28 april 2003 Koninklijk geworden en het Fifth European Congress of Mathematics heeft van 14–18 juli 2008 in Amsterdam plaatsgevonden.

Jan Karel, ik wil je hier heel hartelijk bedanken voor de uiterst prettige contacten en samenwerking in de afgelopen jaren. Ik wens je nog een vruchtbaar laatste jaar toe op het CWI als onderzoeker, en ook al vast een goede en gezonde post-CWI tijd. Ik kan me niet goed voorstellen dat je dan de wiskunde en alles eromheen vaarwel gaat zeggen maar je gezin en familie zullen het vast zeer op prijs stellen je dan vaker te zien.

Met hartelijke groet,

Herman te Riele

```
WISKUNDIG GENOOTSCHAP
-------------------------------------------------------------------------------
BESTUUR
  Eduard Looijenga  (db, voorzitter, inspecteur der boekerij, EMS & IMU)
  Mark Peletier     (db, secretaris, archivaris)
  Luc Florack       (db, penningmeester)
  Harm Bakker       (leraren)
  Annette Kik       (public relations) [MAP]
  Mariette Knaap    (public relations)
  Jan Karel Lenstra (publicaties)
  Jaap Molenaar     (Beta-verenigingen, Beta-federatie)
  Peter Stevenhagen (public relations, Stichting Driehoek)
  Michel Vellekoop  (studenten, ledenwerving)
-------------------------------------------------------------------------------
AGENDA                                                          + done
o REVITALISATIE     bestuur, logo, website                      o to do
                    fondsenwerving, ledenwerving                - forget
                    transformatie NAW
                    public relations MYK2 en daarna
                    relaties met andere beta-groeperingen
o CONTINUITEIT      financien: beheer & beleid
                    publicaties: structuur & proces
-------------------------------------------------------------------------------
ADMINISTRATIE & BESTUUR
+ logo, electronisch briefpapier [MAP,KMolenaar]
o website: vormgeving, redactie, onderhoud [MAP,FBeukers]
+ email-adressen: {wiskgenoot,admin,naw}@wiskgenoot.nl [MAP]
+ belegging kapitaal: fonds ABN/AMRO [LF,MHV]
+ contributieverhoging [LF]
o ledenwerfactie [MHV]
o profiel ledenbestand [NSHekster,FBeukers]
o institutionele leden [EJNL,JKL]
+ dynamiek van besluitvormingsprocessen [MAP]
o begroting in evenwicht: 2003 -Kf20, 2004 OK [db]

PUBLICATIES
+ publicatiecommissie & taak: LF,NHekster,JKL,PSh,FVerhulst,MVermeulen [JKL]
+ overzicht betaald personeel [LF]
NAW
+ CWI: vierde serie], overdracht abonnementen -> UU [PSh,FBeukers]
+ dank: WvdMeiden: NMC36; MBakker,AOng,JSchipper: brief+boekenbon [MAP,JKL]
+ CWI/UB UvA: ruilabonnementen: nieuw contract met CWI [PSh,MAP]
+ hoofdredactie: GKoole+CGZ; eindredacteur: DPik [FVerhulst,CGZ,JKL]
+ begroting [LF]; advies Kemme [JKL]
o functionaliteiten & relaties daartussen [NSHekster]
o financiele scenario's [LF,MVermeulen]
o advertentie-acquisitie [publicatiecommissie]
MWG
+ dank CThomson: NMC36 [MAP,JKL]
o papieren versie], email + web ->; aankondigen in NAW 200106
PYTHAGORAS
o begroting [LF]
o hoofd- & eindredacteur [FBeukers,JvdCraats,SKemme]
- Stichting Driehoek: bestuur, comite van aanbeveling [PSh] (eerst samenwerking)

PUBLIC RELATIONS
+ Pi [HWL,JKL,JTop,CGZ]; dank sponsors, financien, archief [LF,JKL,MAP,CGZ]
+ lezingenserie Wiskunde & kunst [EJNL]
+ NAW naar wetenschapsjournalisten [CGZ]
o collectie profielen van wiskundigen [KAardal]
o mapje ansichten [JKL,PSh,DBeckers]
o folder [JKL,PSh]
```



```
o Koninklijk WG? [JKL]
o ECM2008? ICIAM200x? [EJNL]

ANDERE BETA-GROEPERINGEN
+ niet-leden in WG-bestand: brief, NAW 5.1.1, combola+herinnering
+ ITW,NVvW,VVS: NAW-bon, brief+NAW 5.1.1, combola+herinnering
+ ITW: ITW Nieuws in NAW: nee; sectie (+NS+SWI+ICIAM): ja [JKL]
+ NVvW: reciprociteit: WG ja, NVvW nee [JM,JKL]
o VVS, Belgen, Zuidafrikanen: samenwerking, reciprociteit [JM]
o IMU: cat3 -> cat4; betaling KNAW? reductie IMU? [EJNL]
+ EMS Newsletter: bijdrage [DJBeckers,HWL]; dank: brief+drank [MAP]
-------------------------------------------------------------------------------
ERELEDEN (in leven: vdBlij,dBruijn,Coxeter,Korevaar,Zaanen)
- nieuwe (nu niet)

BOEKERIJ
- nationale coordinatie tijdschriftabonnementen

ARCHIEF, CPAW
- boekje Neugebauer
o lidmaatschap CPAW

NOCW (overleg met voorzitters NOCW & NVvW in januari)
+ AXIS-projecten [DSiersma]

NEDERLANDS MATHEMATISCH CONGRES
+ NMC36, 2000 (UM/OJVrieze)
+ NMC37, 2001 (VU+CWI/HJJtRiele)
+ NMC38, 2002 (TUE/AMCohen)
+ NMC39, 2003 (KUN/Steenbrink): stuur draaiboek

BROUWERLEZING 2002
+ mathematische fysica; dHollander,Dijkgraaf,vEnter,Takens,JKL; Aizenmann

NAJAARSSYMPOSIUM (praktijk; 1997 TUE/JKL, 1998 VU/Verduyn Lunel)
+ NS99: TUE/Schilders, 19991210
+ NS00: CWI/ITW,vBeckum, 20001201
+ NS01 en volgende: gedelegeerd aan ITW

WINTERSYMPOSIUM (leraren; HB)
+ WS00: Amersfoort, Optimalisering, 20000108
+ WS01: Amersfoort, Wiskunde & recreatie, 20010106
o WS02: Amersfoort, ..., 200201..

KALEIDOSCOOP (studenten)
+ K99: A-Es-kwadraat, UU, knopentheorie, 19991110 [JSchut]
+ K01: Abacus, UT, cryptografie, 20010305 [MHV]
-------------------------------------------------------------------------------
```

Jan Karel, The Decider

Paul Vitányi
CWI

1 Introduction

It is common (or should be) that a Liber Amicorum contains only paeans on the lucky guy. In that vein, or perhaps in contrast, here I give some anecdotes having to do with Jan Karel Lenstra. All of them also pertain to me. Therefore "I" occurs regularly.

2 In the Old Days

I met Jan Karel first in the old Mathematical Centre housed in an old school building at the second Boerhaavestraat in Amsterdam. This was a long time ago. Yet even then he gave the impression of being a "decider". This seems to be a common theme in these remembrances.

In that time Jan Karel was a member of the Operation Research group (called the equivalent of "Decision Theory" in Dutch) at the institute. I solved a self-posed problem that involved the difficulty of finding solutions to a certain problem. I recall it fondly, because it evoked the Number of the Beast—666—from graphs. This was later published in a nontechnical math journal.

Returning to the main paper, the literature I looked at told me that only decisions could be hard (or had a name exemplifying the hardness). Thus I went to consult Jan Karel. He looked at me piercingly and gave me the name—deW—of a mathematical colleague I should consult. I did so and the article was eventually published.

Quite some time later they advertised a professorial job at the University of Tilburg. It was in a subject, say X, which now escapes me. "You have to apply," Jan Karel told me. "I know nothing about X, I don't even know what X means," I said. "That doesn't matter," JK said, "They don't know that themselves, that is why they look for a professor." "I don't want to go to Tilburg," I said. But finally I let myself be convinced (that I should apply) although I secretly thought that in the unlikely event they gave me an offer, I could always refuse. But Jan Karel gave me the impression that they were dying for me and waited to see what I would do (once appointed) with X.

Jan Karel was very insistent, he even offered to drive me from Amsterdam to Tilburg, and the other way. I forget if he was associated with that university at the time. By the way, it went or goes by the acronym

"KUB" (Katholieke Universiteit Brabant). So we went together to the KUB. Once there, Jan Karel disappeared and I was interviewed by a group of about 10 people who didn't look familiar at all. "How are you going to give content to X," they asked. "First tell me what X is," I asked. "Well, we advertised for a professor in X, so a candidate can surely tell us what X is," they countered. From then on the interview went down-hill. "Well, too bad," Jan Karel said on the way back to Amsterdam "but it was worth a try."

3 Modern Times

For some time Jan Karel disappeared from my radar. He became (after being a professor there) Dean of the relevant Faculty at Eindhoven. His friend—ARK—did even better and became a power in commercial-political Netherlands. Jan Karel married to AK and moved to Atlanta (USA) to take up an appropriate position at Georgia Tech. Suddenly, the director of CWI disappeared and Jan Karel became the new director. He celebrated this with decisive action.

One of the first things he did was installing a new Crypto group. To attract a new leader he made a complicated deal with the University of Leiden where his brother—HWLjr—was ensconced. The new leader in question was my former PhD student RC. To arrange things a secret consult was called. HMB, Jan Karel and myself went to dinner in an eminent hotel. There we ate an exquisite meal and talked of the things of the day and RC. Jan Karel was a true decider; only a little time later the group with its new leader was started.

For CWI one of the important items was the matter of a new building. Discussion about this had been going on for some time. The supervising office of NWO had earmarked a modest number of millions. Plans abounded. Jan Karel cut this Gordian knot and decided on an extension of the old building. Remarkably quick a triangular new part of the old building arose. Windows from floor-to-ceiling, transparency to the corridors, air conditioning, easy seats and glass "black"boards, printing stations, kitchens, a new large triangular atrium, and so on. The old building was renovated too, to go with the new part. Everything went remarkably quick, and the result was excellent, a tribute to decisive action.

Sometime later I got a stroke in Waterloo, Canada. I was transported home and convalescenced in the appropriately named "Sun and Sea" in the coastal town of Wijk aan Zee. Interestingly, one could on one side enjoy the flames and smoke of the largest steel mill in the Netherlands together with the "Street of Steel", on the other side the dunes, beach, and sea. Due to the strong winds it is a Mekka for paragliding and ultralight aviation. Anyway, there I spent my time in a wheelchair. My roommate received many people,

among whom Prince Constantijn. He was high up in the advertising business. Some people came to visit me. Jan Karel was one of them. Rather than being a decider, it was the empathic side that took over. He took time off from a busy schedule, travelled to a difficult-to-reach place, while he should be packing and preparing for his summer stay in the country side of Norway. To me it was an enjoyable visit and we could talk about my present condition, the future, things that came to mind, and old times.

I'm sorry, Dave. I'm afraid I can't do that.

HAL de computer uit Kubricks "2001: A Space Odyssey" is mogelijk een verwijzing naar de computerfabrikant IBM. Tel een letter op bij de H, de A en de L en voilà je krijgt IBM. Kubricks meesterwerk verhaalt over de evolutie. Onder het geroffel van trommels uit "Also sprach Zarathustra" zien we in verkorte vorm de aap verworden tot mens die de ruimte in reist. HAL ontpopt zich tot een wel heel erg intelligente computer die de race aangaat met Dave op zoek naar de volgende stap in onze evolutie.

Tot zover 2001. Fastforward naar 2011. Jan Karel Lenstra directeur van het Centrum Wiskunde & Informatica, voorheen *voor* wiskunde *en* informatica, zet de volgende stap in zijn carrière. JKL is mogelijk een verwijzing naar de Koninklijke Luchtvaart Maatschappij der Nederlanden. Goed gekozen want het specialisme van de directeur is het vinden van efficiënte oplossingen voor het handelsreizigers-probleem, bekend uit de Operations Research (OR).

Beste Jan Karel, ik vind het jammer dat jouw periode als directeur en beschermheer van het CWI erop zit. Ik heb met veel plezier met je samengewerkt en heb onder andere mogen meehelpen met het initiëren van de cryptogroep en in een later stadium de lifesciences groep. Die twee onderwerpen liggen me na aan het hart en ik ben enorm blij dat het CWI nu ook deze disciplines herbergt. Ik wil je bedanken voor jouw inzet en werk.

Bij Kubrick dreigt de mens te worden ingehaald door de computer, bij jouw vakgebied zou het wel eens zo kunnen zijn dat de natuur de mens inhaalt. Onlangs las ik dat hommels goed zijn in het oplossen van het handelsreizigersprobleem. Dat past dan weer uitstekend, dankzij jou, binnen de portefeuille van het CWI. Ik hoop je de komende periode nog veel op het CWI te zien. Niet als directeur, maar als wetenschapper. Per slot van rekening, OR plus een is PS. Laat die hommels maar eens weten wie de baas is.

Harry Buhrman

Connecties met JKL

Peter Grünwald

1. **Inleiding: een ander Liber Amicorum:**

Het was eind juni 1989. Ik was eerstejaars student informatica aan de VU Amsterdam en ik bereidde me voor op mijn tentamen kansrekening. In diezelfde maand verscheen een boekje aan de TU Eindhoven met de titel "Op de Snede tussen Theorie en Praktijk". Op de achterflap lezen wij:

In deze bundel opstellen over de toepassing van wiskunde in de praktijk, probeert een uitverkoren groepje (hoog)geleerde heren zichzelf en elkaar te overtreffen in originaliteit en diepte van visie...

Het betreft hier een Liber Amicorum voor Prof. dr. J.F. Benders, hoogleraar aan de TU Eindhoven. Benders ging juli 1989 met pensioen, en tot zijn opvolger was benoemd: Jan Karel Lenstra, afkomstig van het CWI in Amsterdam! Het was daarom niet verbazingwekkend dat JKL ook als co-auteur in het Liber voorkwam. Wellicht iets verbazingwekkender is dat er ook een bijdrage in staat van mijn vader, Herbert Grünwald (destijds werkzaam bij Philips, inmiddels 80 jaar oud en nog steeds zeer actief als milieu-activist in het lieflijke Oostenrijkse dorpje Pitten). Ziehier een eerste connectie met JKL. Met enig hangen en wurgen (Bending Benders, zie Figuur 1 in Sectie 5) is het mij gelukt alle in het oog vallende connecties tussen JKL en mij aan het Benders Liber op te hangen:

2. **Organisatie van het CWI-Onderzoek**

Naast het feit dat behalve JKL en mijn vader bijna iedere auteur een grote bril en snor droeg, valt vooral de naam op van Jan Karels bijdrage aan het Benders Liber (co-auteurs K. Anthonisse en M. Savelsbergh):

De Bochtige Weg van Idee naar Implementatie

Dit lijkt wel een profetische verwijzing, twintig jaar vooruit in de tijd, naar Spinoza: JKL's plan om de organisatie van het onderzoek op het CWI eens heel flink op de schop te nemen. Ik bedoel dit geenszins sarcastisch: naar mijn mening is de verdeling van zowel onderzoekers als onderzoeksthema's over de onderzoeksgroepen onder JKL's leiding een heel stuk verbeterd: er zijn minder groepen, er wordt aan minder onderwerpen tegelijkertijd gewerkt, en er wordt beduidend meer samengewerkt. In feite zijn delen

van het Spinozaplan uiteindelijk wel degelijk gerealiseerd, en is het CWI een coherenter geheel geworden. Dit is mij pas echt duidelijk geworden toen ik, in de herfst van 2010, meeschreef aan de periodieke zelfevaluatie van het CWI. Door het meeschrijven en discussies met medeschrijvers Sander Bohte, Jason Frank, Frank Roos en vooral JKL zelf, kreeg ik een overzicht van wie op het CWI nu eigenlijk wat doet, en in welke groep. Het was overduidelijk dat er sinds het aantreden van Jan Karel een duidelijke trend is naar minder groepen en meer bundeling van onderzoek, en dat er, gezien de inderdaad immer bochtige weg van idee naar implementatie eigenlijk heel veel is gebeurd.

Desalniettemin is een van de hoofdmoten van het Spinozaplan niet uitgevoerd: de onderzoeksgroepen aan het CWI zijn zelf weer gegroepeerd als *clusters,* en Jan Karel wilde die clusters heel graag afschaffen. Dat is niet gebeurd. Maar, over clusters gesproken:

3. **Clustering**

Ook aan het begin van zijn carrière was JKL heel erg bezig met clusters, zij het van een ander soort. Eén van zijn allereerste publicaties, uit 1974, was:

Clustering a Data Array and the Traveling-Salesman Problem
J.K. Lenstra, Operations Research 22(2), pp. 413-414.

Het onderwerp van dit artikel is inderdaad het clusteren van data, samengevat in een matrix, en dit is iets waar onze onderzoeksgroep zeer actief mee bezig was toen JKL begon aan het CWI directeurschap. Mijn promotor Paul Vitányi en zijn oio Rudi Cilibrasi boekten grote successen met *Complearn.* Deze software implementeerde de zgn. *normalized compression distance.* Het idee is als volgt: neem twee bestanden *A* en *B* (bijvoorbeeld twee DNA-sequenties, of twee teksten, of twee muziekbestanden in MIDI-format). We comprimeren eerst beide bestanden apart met een standard data compressor (bijvoorbeeld *bzip* of *gzip*). Dit geeft twee nieuwe bestanden *A.zip* en *B.zip*. Vervolgens maken we een nieuw bestand door *A* en *B* aan elkaar te plakken, en comprimeren we het resulterende bestand. Dit geeft een nieuw bestand, *AB.zip*. Als *A* en *B* op elkaar lijken, dan zal de lengte van *AB.zip* in het algemeen veel kleiner zijn dan de som der lengtes van *A.zip* en *B.zip*. Als *A* en *B* volstrekt ongerelateerd zijn, dan zal de lengte van *AB.zip* ongeveer gelijk zijn aan de som der lengtes van *A.zip* en *B.zip*. Bijvoorbeeld, als *A* het volledige DNA van JKL is en *B* het volledige DNA van ondergetekende, dan is het verschil tussen *lengte(AB.zip)* en *lengte(A.zip)+lengte(B.zip)* veel groter dan als *A* het DNA van JKL is en *B* het

DNA van, zeg, een eekhoorn. Op basis van dit idee kan een afstandsmaat tussen bestanden gedefinieerd worden - hoe meer extra compressie bereikt kan worden door de bestanden aan elkaar te plakken, hoe dichter ze bij elkaar liggen in *compression space*. Die afstanden kunnen weer gebruikt worden om allerlei soorten data te clusteren - we bepalen eerst de afstanden tussen de bestanden, stoppen die in een matrix en gebruiken vervolgens (bijvoorbeeld) JKL's algoritme voor het clusteren. Dit onderzoek was 'hot' rond de tijd van JKL's aantreden als directeur van het CWI, en een van mijn eerste interacties met JKL was op de Open Dag van 2005, toen ik het algoritme uitlegde aan JKL en zijn zoontje. JKL begreep precies hoe het werkte en waarom wij het zo leuk vonden, en dat vond ik heel prettig.

4. **CAR** *of: het plezier van een directeur die begrijpt waar je mee bezig bent.*

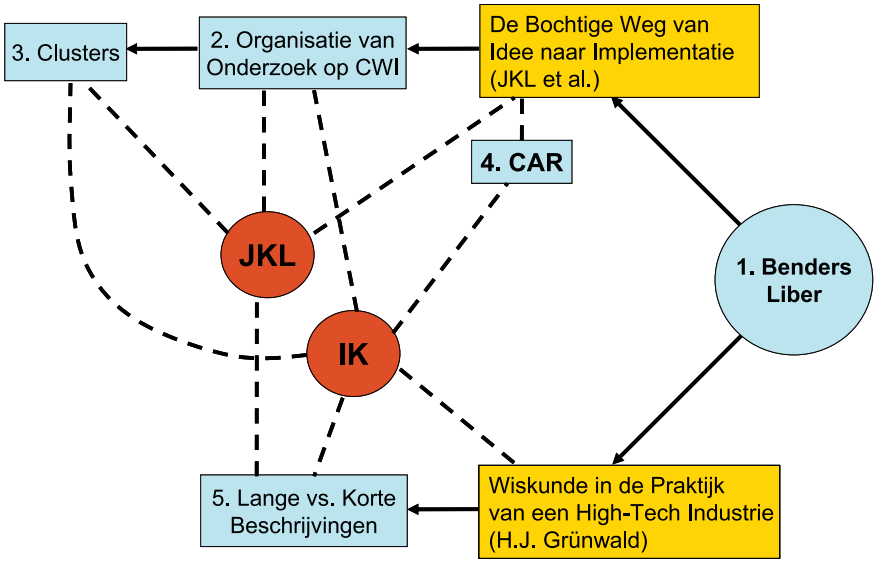
Ook later bleek vaak dat JKL affiniteit had met mijn onderzoek. In 2008 gaf ik een mini-voordracht in Leiden ter viering van de samenwerkings-overeenkomst die het CWI met Leiden had gesloten (en die leidde tot mijn snelle aanstelling als deeltijdhoogleraar aldaar). Ik probeerde in tien minuten eerst het statistische begrip *CAR* uit te leggen, en besprak vervolgens de algoritmische karakterisatie die Richard Gill en ik eraan hadden gegeven. Uit JKL's opmerkingen achteraf kon ik opmaken dat hij het verhaal van A tot Z had begrepen. Overigens betekent *CAR "Coarsening at Random",* en is er ook hier weer een verbinding met het Benders Liber: het onderwerp van JKL's 'Bochtige Weg' artikel in dat Liber is...CAR!!! Het woord CAR komt acht keer voor op de eerste pagina. Hier betekent CAR echter *"Computer Aided Routing"....*

Het komt denk ik niet zo vaak voor dat een wetenschapper die bestuurder is geworden zijn interesse in wetenschappelijke details blijft behouden. Jan Karel is dit gelukt. Tegelijkertijd is hij heel goed in het ontwaren van de grote lijnen in iemand's onderzoek. Tijdens het schrijven van de interne evaluatie viel me op dat hij een adequaat high-level idee heeft van het onderzoek van vrijwel iedere senior onderzoeker op het CWI, waarvan akte. In mijn geval bleek dit eens te meer toen ik, samen met Harry van Zanten, de Van Dantzigprijs uit zijn handen ontving: met zijn beschrijving van ons onderzoek en het belang ervan sloeg JKL de spijker geheel op de kop. Mijn ervaring is dat dit meestal niet het geval is bij prijsuitreikingen…

5. **Kort en Helder (be-) Schrijven!**

JKL is een man van korte en heldere teksten. Maar een plaatje zegt meer dan duizend woorden, en JKL is dan ook de meester van de *visuals:* tijdens

het schrijven van de zelfevaluatie was hij vaak tot in de kleine uurtjes bezig met het bedenken en vereenvoudigen van figuren en diagrammen (ik denk eigenlijk dat hij veel meer tijd aan de zelfevaluatie heeft besteed dan Sander, Jason en ik...). De zelfevaluatie van het CWI staat vol met dit soort visuals waarin een grote hoeveelheid informatie op verrassend duidelijke wijze wordt samengevat. Een voorbeeld van een visual is te vinden in Figuur 1. Wie deze visual naast de vele visuals van de zelfevaluatie legt ziet meteen dat JKL er veel beter in is dan ik.



Figuur 1: een "visual" die de hier besproken connecties tussen JKL en mij weergeeft, en tevens de superioriteit van JKL op dit gebied illustreert.

Ikzelf ben de auteur van het boek *The Minimum Description Length Principle,* een methode binnen de statistiek en machine learning gebaseerd op het idee dat het zoeken van patronen in data gezien kan worden als het pogen de data samen te vatten, dat wil zeggen op zo kort mogelijke wijze te beschrijven: *hoe korter het kan, hoe beter het is* - weer een connectie met JKL. Helaas ben ik niet consistent: mijn eigen boek is, ondanks de titel, 570 pagina's dik - als ik er een jaar langer aan had gewerkt had ik het waarschijnlijk tot 300 pagina's kunnen indikken. Maar daar had ik na vijf jaar werk geen zin meer in. Wellicht zit deze eigenschap in mijn genen: in het stuk van mijn vader in het Benders Liber staan af en toe wat lange, omslachtige zinnen. Mijn vader verontschuldigt zich hiervoor in een voetnoot, waarin hij schrijft:

Gezien "Die Knappheit der Zeit und die Vordringlichkeit des Befristeten" (Goethe) was het niet te vermijden dat de formuleringen bij het eerste lezen van dit stuk waarschijnlijk een cryptische indruk maken. Ik houd mij aan een uitspraak van mijn wiskundeprof toen hij bezig was een boek te schrijven en men hem alsmaar kwelde duidelijker te formuleren: "ik moest ook altijd moeite doen om wiskundeboekjes van anderen te ontcijferen, dus waarom zou ik het zelf duidelijk doen?")

JKL kan dit beter dan de Familie Grünwald. Toen ik mijn boek aan JKL presenteerde en hij opmerkte dat het wel erg lang was, zeker gezien de titel, verontschuldigde ik me op mijn beurt met het citeren van Pascal, die naar ik meende ooit had geschreven "Ik stuur u een lange brief, omdat ik niet voldoende tijd had om een korte te schrijven". JKL kende een soortgelijk citaat (van Benjamin Franklin, meen ik). Dit leidde tot een langer gesprek waarin wij ons afvroegen wie nu precies wat hierover had gezegd. Ik besloot het uit te gaan zoeken. Met behulp van Teddy Seidenfeld maakte ik een lijst met citaten van (inderdaad) Pascal, maar ook Thoreau, Cicero, Gauss, Nietzsche, Felelon (mij onbekend), Voltaire, Woodrow Wilson en Mark Twain. De mooiste vier repliceer ik even:

Je n'ai fait cette lettre-ci plus longue parce que je n'ai pas eu le loisir de la faire plus courte
Blaise Pascal, Lettres provinciales (Lettre No. XIV), Pierre le Petit, Paris 1656

If I had more time, I would have written a shorter letter.
Marcus Tullius Cicero

If you want me to give you a two-hour presentation, I am ready today. If you want only a five-minute speech, it will take me two weeks to prepare.
Mark Twain

You know that I write slowly. This is chiefly because I am never satisfied until I have said as much as possible in a few words, and writing briefly takes far more time than writing at length.
Carl-Friedrich Gauss.

Toen ik de gehele lijst aan JKL mailde, kreeg ik onmiddellijk het volgende korte maar krachtige antwoord:

Peter -- Dank! Twain is prachtig. Maar Cicero is het kortst dus hij wint.
--- Jan Karel

Naar aanleiding van deze citaten is het trouwens de hoogste tijd om te stoppen met dit stukje!

6. **Conclusie**

Laat ik het verder inderdaad maar kort houden en afsluiten zoals JKL het zelf zou doen:

Jan Karel -- Dank!
--- Peter

Vakant(ie)

Ronald Cramer

juli 2011

Wetenschappers en vakantie: het is soms tobben. Een paar dagen weken in warme baden kan wel eens voldoende inspiratie opwekken om bijvoorbeeld dat weerbarstige probleem via een niet eerder begane route te kunnen gaan benaderen.

Of misschien is er nog meer verheffing; de afsluiting van (je) ideeën over een zeker onderwerp knelde al een tijdje, en nu rijst er eindelijk een idee dat, na adjungeren daarvan, nieuwe mogelijkheden en vragen voorbij die eerdere horizon zal voortbrengen. Met navelstaren of reizen-per-leunstoel heeft deze wetenschappelijke soort van transcendentie natuurlijk niets van doen, want er is eerlijk en opwindend werk aan de winkel!

Onder dergelijke omstandigheden kan de vakantie ineens een nieuwe bestemming krijgen; het is allemaal wel eens voorgekomen. De onvermijdelijke vruchteloze pogingen moeten nu dan maar even buiten beschouwing gelaten worden; in zo'n geval is de vakantie voor *alle* betrokkenen een koude douche. Hoe dan ook, lang leve het “tobben”!

Het was echter niet de bedoeling dat een vakantie zoals bijvoorbeeld die in de zomer van 2001 spoedig nog eens dunnetjes over gedaan zou worden, hoe succesvol die uiteindelijk ook was. Althans, het laatste deel van die opvatting werd uitsluitend gedeeld door mij en mijn coauteur, die elders in parallel aan ons onderzoeksprobleem bezig was, alleen onderbroken door onze langdurige telefoongesprekken.

Zo hield ik mij in oktober 2003 op bezoek bij mijn vader in Zuid-Frankrijk dan ook verre van zijn badkuip, en probeerde ik door loom dobberen in het zwembad een betekenis van het woord “vakantie” (als gegeven in Van Dales Nieuw Handwoordenboek der Nederlandse Taal, 1956) recht te doen: rusttijd die jaarlijks wordt toegekend.

Nou, daar kwam helemaal niets van terecht. Als een soort deus ex machina hief Jan Karel Lenstra deze staat van inertie pardoes op. In het bericht (per email) dat ik van hem ontving maakte hij zich bekend als de nieuwe CWI directeur (nadat ik hem aan het eind van de jaren negentig in zijn

hoedanigheid van TU/e decaan reeds had leren kennen), en polste hij mijn interesse voor een positie op het CWI. Spoedig daarna begonnen wij samen na te denken over een eventueel door mij op te richten en te leiden onderzoeksgroep cryptologie op het CWI.

Destijds werkte ik al zo'n zeven jaar in het buitenland (Denemarken en Zwitserland), en stond ik bijna op het punt om een leerstoel in Duitsland, waartoe ik was verkozen, te gaan accepteren ("C4" toen nog, en dus inclusief alle daarbij behorende structurele middelen om een groep op te zetten, zeer aantrekkelijk). Jan Karels interessante opening schudde de kaarten echter wel degelijk weer door elkaar, want mogelijke opvolging geven aan de teloorgegane maar legendarische cryptologie groep van David Chaum op het CWI uit de jaren tachtig en negentig, bood een bijzonder perspectief.

Er volgden enige maanden van spannende onderhandelingen voordat Jan Karel en ik tot overeenstemming kwamen.
Maar op 1 juni 2004 was mijn nieuwe cryptologie groep ("PNA5") op het CWI werkelijkheid (als ook mijn leerstoel aan het Mathematisch Instituut van de Universiteit Leiden, eveneens een belangrijk onderdeel bij mijn beslissing terug te keren naar Nederland). De rest is geschiedenis. Toen Jan Karel en ik onlangs aanzaten voor het laatste halfjaarlijkse beoordelingsgesprek van mijn groep onder zijn directeurschap en we elkaar ferm en warm de hand schudden in die wetenschap, konden we terugkijken op zeven bewogen en zeer succesvolle jaren. "Never a dull moment with PNA5" zo stelde Jan Karel terecht, en dat was geheel wederzijds!

Jan Karel, nu je per 1 oktober 2011 na precies 8 jaar terugtreedt als directeur van het CWI, wil ik je hierbij veel dank zeggen voor het vertrouwen dat je in mij hebt uitgesproken met de oprichting van de cryptologie groep in 2004 en voor onze effectieve samenwerking in de jaren daarna. De wetenschap van het CWI en onze (inter)nationale positie zijn bij jou, CWI-er in hart en nieren, in uitstekende handen geweest. Er is een gezonde basis voor een succesvolle voortzetting, die we bovendien vanuit een prachtig gerenoveerd hoofdgebouw met een nieuw aangebouwde en wederom zeer kundig doorgerekende vleugel verder vorm kunnen blijven geven.

Tegen getalenteerde jonge postdoc kandidaten die aantrekkelijke aanbiedingen hebben van in zonniger oorden gelegen gerenommeerde instellingen, plegen wij in PNA5 wel eens gekscherend te zeggen: wat wordt het, vakantie of wetenschap? En zowaar, een aanzienlijk aantal van hen

hebben we voor langere tijd aan ons weten te binden, alvorens ze nagenoeg allemaal te zien vertrekken naar vaste posities in het buitenland zoals hoogleraarschappen.
Zo'n zelfde type vraag kan natuurlijk ook heel goed gesteld worden aan van hun formele posities afzwaaiende wetenschappers. Maar die vraag is bij Jan Karel feitelijk pas aan de orde bij zijn emeritaat komend jaar, hoewel ook dan vast niet relevant: het bloed blijft toch kruipen waar het niet gaan kan. Als je nooit de diepte ingaat dan kom je nergens, zoals Jan Karel zegt, en daar sluit ik me van harte bij aan.

Jan Karel, het ga je goed!

An Appreciation of the Absurd

I believe that I first heard of Jan Karel Lenstra's appointment as director of CWI from Jan Verwer, who was pleased that a 'real' mathematician would be heading the Institute. I probably had met Jan Karel briefly in the late 1980's, when I first joined CWI as head of the Computer Systems and Telematics department, and he was undoubtedly at some of the initial MT-meetings that I had attended when that body still had a dozen or more members, but I did not have a specific physical or emotional image of the man who was about to come back from Georgia Tech in Atlanta to lead the Institute to which I had myself only recently returned.

In the early years of his tenure, our main contact came from sharing the north-west corridor of the third floor of what we would now call the old CWI-building. His arrival brought with it the usual commotion of painters and carpet layers. New was his introduction of a small refrigerator to the official artifacts of office. This seemingly simple deed signaled to me an iron will – a will that would later translate into the repositioning (and de-positioning) of topics, tasks and people across the Institute. The 'fridge' was, in and of itself, no big deal. But for the careful observer, it was as if a red flag had been planted in a virtual stroke of sand that ran across the threshold of his office. My feeling was: this is not going to be our typical CWI-managing director.

The sharing of our corridor gave my then-young children an intriguing new place to explore during their weekend afternoons at the Institute. In addition to the hot chocolate machine, the blackboard (and later whiteboard) in the third floor common room and the terrace outside of my office, they sensed a new gravitational force attracting them to JKL's office at the (for them) far end of the hall.

All of these were relatively passive encounters. My first real interaction of substance with Jan Karel came during the production of an early draft of the self-study document for the 2005 NWO evaluation of CWI. I had read the introduction that flowed from the Director's pen and thought: this was all too dry and factual to excite any reviewer. I decided to rewrite parts of the document, and submitted these as unsolicited comments from a concerned reader. To my surprise, he was open and ready to accept criticisms of his own work (which is much harder than sharing the criticisms of work done by others). What followed was a set of animated discussions that

ranged from fonts and punctuation rules to striking formulations that could capture the dynamic essence of CWI in a manner that would retain a reviewer's interest for the full ten minutes that he or she would dedicate to the introduction chapter of the document. When I presented the final text to the MT, the discussion never really got beyond the level of commas and semicolons. In the one-on-one conversations with Jan Karel, however, we were able to focus on a more abstract view of the content. This man had a delightful sense of the absurd.

This appreciation of off-beat formulations – and the love of the debate over details at the edge of nearly any subject – also defined the tone of the semi-annual theme/ group review meetings that I held with our director. In the early days, these meetings would consist of a relatively structured discussion of the short- and long-term goals of my young group, often ending with a discussion on the need for me to not only think globally but to act locally on establishing research partnerships. In later years, in the company of not only my cluster leader but also the Institute's budgetary and personnel gurus, our discussions became animated mini-evaluations of the people, projects and surpluses that were known as SEN5. In these discussions, Jan Karel always maintained meticulous facts lists, consisting of scribbles written in pen using a letter size more space-efficient than most characters in first generation microfilm documents. He was also keenly aware of the meta-process that guided our group's development. Rather than being boring, pro-forma sessions, I always left these meetings with the feeling that I had just stepped off the tennis court.

The piccolo style of writing was, in public, complimented by a neatly staccato delivery style that transformed mundane and often forced New Year's presentations into engaging mini-performances that drew crowds rather than repelling them. To me, the most engaging aspect of these talks was not the content, but the animated manner in which the content was structured and delivered. Set to music, they would have given Gilbert & Sullivan a run for their money. Even more impressive, however, was the critical self-analysis that followed each presentation: how were they received, did they 'work', what could be learned for the next show?

The official history of the Jan Karel Lenstra period at CWI will chronicle the exploits of a strong-willed director whose interests and insights did not always flow in the same direction as the generic stream of consciousness of everyone at the Institute. While there is

room for discussion on how all decisions were received, on what worked during the last decade (and what didn't), and on the lessons that have been learned by us all – and which hopefully will influence our new director – this official history will describe an Institute that is internationally regarded, intellectually alive and financially robust.

While I can imagine the frustrations of some (and the delights of others), I personally am left with an appreciation for the humor and directness of this boss. While I wasn't always happy with the results of our discussions, I was always invigorated and amused by them. His will be a tough act to follow.

Dick Bulterman

Afscheid van JKL

Als lichtvoetige afscheidsgroet aan Jan Karel gaat hierbij een poetische versie van een bewijs van een van de meest fundamentele inzichten uit de theoretische informatica: er zijn beslissingsproblemen die met geen computer zijn op te lossen. Het beroemdste onbeslisbare probleem is zonder twijfel het stop-probleem. Dat dit probleem onbeslisbaar is werd ruim voor de moderne computer werd uitgevonden aangetoond door Alan Turing.[1]

Turing's bewijs laat zien dat er geen algemene methode kan bestaan om van willekeurige programma's te zeggen of ze zullen stoppen, bij gegeven invoer. Om een voorbeeld te geven: het volgende Ruby programma stopt altijd:

```
print "Hoe heet je? "
input = gets.chop
if input == 'Jan Karel Lenstra'
    puts "Fantastisch"
else
    puts "Jammer"
end
```

Maar het volgende programma raakt bij foute invoer in een oneindige lus:

```
print "Hoe heet je? "
input = gets.chop
if input == 'Jan Karel Lenstra'
    puts "Fantastisch"
else
    puts "Jammer" while true
end
```

In 2000 publiceerde de Britse taalkundige Geoffrey Pullum een bewijs van de onbeslisbaarheid van het stopprobleem in dichtvorm. Mijn bijdrage is daarop een Nederlandse variant.

[1] A.M. Turing, 'On Computable Numbers, with an Application to the Entscheidungsproblem', *Proceedings of the London Mathematical Society*, 1936

De lusvinder genept — een logisch recept

een eenvoudig bewijs
van de onbeslisbaarheid van het stopprobleem

Geen methode voorspelt of mijn codes gaan lussen,
dat toon ik hier aan, en jij krijgt daar niets tussen.
Hoe je ook sputtert, m'n bewering gaat kloppen:
jij kunt niet voorspellen of mijn code zal stoppen.

Stel dat P, een programma, dat tóch kan voorspellen,
dus voor I, plus zijn invoer, weet te vertellen
'I stopt niet' als I in een lus raakt en doordraait en spint,
terwijl P schrijft 'I stopt' als hij géén lussen vindt.

Je geeft P mijn code, met data erbij.
En P leest en rekent, en checkt allebei.
Wat er ook in gaat, P stopt na een tijd:
P geeft altijd betrouwbaar voor alles bescheid.

Nu, het zit zo: zo'n P kan niet bestaan.
Want stel, jij schrijft hem en komt er mee aan.
Dan kan ik hem gebruiken voor een logische val
zo vilein dat je brein erop breekt met een knal.

De truc die ik uithaal vertel ik je nu.
Ik maak een recept, en ik noem het ding Q.
Q neemt een programma en gaat na of dat stopt
met P kan dat immers, dus je snapt dat dit klopt.

Voor een programma dat je Q aanreikt, zeg A,
vraagt Q P om raad, en gaat daarmee na
of wat ik A heb genoemd in een lus raken zal
als A werkt op zichzelf, een bijzonder geval!

En dit is niet alles. Als P zegt 'een lus'
dan zal Q het beamen, en Q stopt daarna dus.
Maar als P zegt 'dit stopt' gaat Q eindeloos voort
met dat steeds maar herhalen, in een lus die ontspoort!

En dit programma, dat ik Q heb gedoopt,
gaan we gebruiken, en jouw P wordt gesloopt.
Als Q opstart en vraagt om zijn invoer? Zeg: Q.
Q leest dan zichzelf, en . . . wat krijgen we nu?

Als Q in een lus raakt hebben we brokken.
P zegt dan 'een lus', want P zal niet jokken.
Maar dan moet Q dit beamen en stoppen.
Nu, je begrijpt, de methode zal floppen.

Maar als Q juist stopt, na verloop van een tijd
dan zal P dit zien, en vort met de geit,
Q moet dan een lus in, want dat was het recept.
Je ziet, ook dit kan niet, dus ik heb je genept.

Wat P doet is funest, 'k hoop maar dat je het ziet.
Want P zegt hoe het staat en Q doet dat teniet.
Als P eerlijk wil spreken dan liegt hij je wat voor.
En zo P wil liegen: 't klopt toch door en door.

Het recept dat je hier zag is bekend aan veel koks
uit de logische keuken, en het heet 'paradox'.
Toen je aannam dat die P kon bestaan
kreeg ik je te pakken, en toen ging je eraan.

Hoe kun je ontsnappen aan dit logisch debâcle?
Geloof niet in die P, want dat was het obstakel.
Van de tegenspraak die ik liet zien kun je leren:
dat zo'n P kan bestaan mag niemand beweren.

De moraal: niemand kan jou een methode vertellen
die het op 'tilt' gaan van je PC kan voorspellen.
Word niet boos en blijf vrolijk, is wat ik je brom,
en zoek zelf naar de fouten, want computers zijn dom.

Dit poëtisch bewijs van de onbeslisbaarheid van het stop-probleem is zeer vrij naar *Scooping the Loop Snooper* van Geoffrey Pullum, *Mathematics Magazine*, oktober 2000. Die versie was logisch niet helemaal accuraat; een gecorrigeerde versie is op internet te vinden: `http://www.lel.ed.ac.uk/~gpullum/loopsnoop.html`.

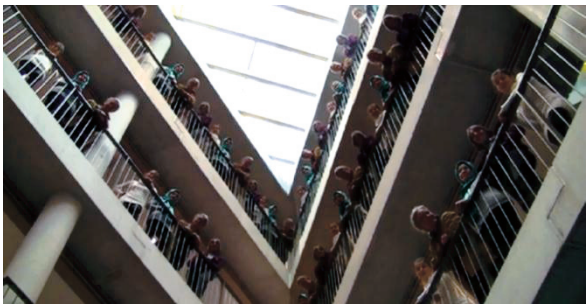
Jan van Eijck, augustus 2011

How Jan Karel made CWI more attractive for female researchers

This year, CWI has reached its 65th birthday, just a few years older than Jan Karel himself. During this time CWI has seen many female researchers and programmers come and go in its staff. In CWI's fields of mathematics and computer science, only a small percentage of the workforce is female. This can be viewed as at most unfortunate, given the fascination the subjects hold, but not necessarily something that should be changed. The true problem, as shown by the statistics, is that the percentage of female researchers becomes smaller at each step in the career ladder. Jan Karel has been active in attempting to redress this imbalance throughout his period as director. Among the positive actions that he has carried out during his directorship are the signing of the “Talent to the Top Charter” in 2010, the creation of a short video to highlight the role of women at CWI (http://www.cwi.nl/general/room-for-women) and the appointment of a female researcher to the scientific management team.

As part of maintaining the high quality of CWI as an internationally recognised institute, Jan Karel has deliberately attracted excellent talent to the institute. The selection of this excellent talent, however, often occurs at the “natural age” of starting a family, with the subsequent consequences for the selection of male or female candidates based on their CVs. The small numbers make it difficult to draw conclusions from the choices made at CWI, but they do show that there are no new permanent female members of the research staff since before Jan Karel joined as director. The dilemma is how to select an individual who excels in his/her research area, while at the same time ensuring a balanced mix of genders over a longer period of time. Selecting a female with a shorter track-record is, from the institutes’s perspective, more risky, but if only males are selected then the current imbalance will remain.

The problem is of course larger than CWI itself. In recent years, both NWO and the European Research Council have introduced a way of comparing CVs of male and female research personnel. Specifically, that the post PhD experience of a CV for female researchers is reduced by 18 months per child. This gives, to some extent, a way of compensating for the tendency for mothers to be less productive in their early postdoc years. While this construction has been introduced, those who review and compare CVs of researchers have to learn how to weigh up these differences, and the women themselves are often reluctant to be treated differently. This learning process of how to “equalise” the gender issues in CVs is individual, cultural and institutional.



This is just one example of how the assessment of female researchers is changing. Not only is the awareness of the role of women within the institute changing, so too is the role of other cultural diversity factors within CWI. Just as the sustainability of a species is dependent on the genetic diversity of its population, so is the creativity of an institute dependent on the intellectual diversity of its personnel. This applies not only to issues of gender, but also to cultural diversity. While CWI is home to researchers with highly varied backgrounds, the vast majority of the senior permanent staff is still Dutch and male.

These issues are not unique to CWI and are being addressed at different levels throughout the academic community. For example, guidelines for selecting professors in the Netherlands have been developed to give practical tips for ensuring the openness of the selection procedure, which can only be beneficial for candidates of both genders. It is now up to our new director to continue the efforts introduced by Jan Karel, so that upcoming female researchers will be ensured proportional representation throughout their career.

Lynda Hardman

Beste Jan Karel,

[ik ga door in het Engels].

Life has been good for me during your 'reign' at CWI and I certainly can attribute it to you. It was you who appointed me a CWI fellow. It provided me with new, luxurious, opportunities to pursue my research interests at CWI.

On a couple of occasions we have had discussions in which our views diverged, sometimes even strongly. But it was always within the bounds of a correct exchange of opinions. I have always been impressed that you kept the tone of these discussions correct and gave me ample time to express my views. In fact, I have never felt after these discussions that our personal relations have been affected in any way.

I am most grateful for your continuous and unfailing support for introducing (algorithmic) game theory in CWI. You have been extremely helpful and positive in this matter. It provided me and other researchers with a motivation to continue this line of research, and helped us to create at CWI a group focusing on this subject.

As an aside I should add that your support for two other new research directions at CWI, cryptography and computational biology, led to new exciting developments at CWI that are very inspiring and that made CWI an even more attractive and excellent institute to work at.

I hope you will use the last year of your CWI employment in the way which suits you best.

Best regards and many thanks for your help and support,

Krzysztof

Over wetenschap, besturen, Chinese folklore en respect

Beste Jan Karel,

Mijn eerste, en jouw tweede, leven als CWI-er lopen nagenoeg synchroon, en door de jaren heen hebben we steeds meer met elkaar te maken gekregen. Jij als directeur, ik als onderzoeker en later als groeps- en clusterleider. En ik kijk er met veel plezier op terug! Als ik terugdenk aan de laatste zeven jaar dan schieten er een aantal dingen door me heen, die ik in dit vriendenboek graag met je deel.

Directeur zijn van een instituut als het CWI is zwaar, dat realiseren mensen zich vaak niet. Veel *eigen-wijze* mensen met hun eigen belangen en overtuigingen, en die het allemaal beter weten. De leukste dagen van een directeur zijn de eerste dag en de laatste dag.

Maar daar tussenin gebeuren er een hoop leuke dingen. Ik krijg nog altijd tranen in mijn ogen van het lachen als ik terug denk aan dat verschrikkelijke optreden van die Chinese folklore-band bij de opening van de ontspanningsruimte.
En ik ben niet de enige. Je had net de officiële opening verricht en stond als directeur noodgedwongen vooraan bij het optreden van onze Chinese herriemakers. En het optreden duurde veel te lang. Maar het kon mij niet lang genoeg duren. Heerlijk leedvermaak. Je bleef geforceerd glimlachen, maar ik zag dat je ondertussen stiekem op je horloge keek, ongetwijfeld verlangend naar thuis op de bank met Karen en een goed glas rode wijn. Hahaha! Ook heerlijk: De verhalen over de douche in Risor! En: Jan Karel als scheidsrechter bij de meisjeshockeywedstrijden van Liesje. En dan die ouders van de tegenstanders die denken dat jij partijdig bent. Ik zie het al helemaal voor me. Ach, wat kan het Lenstra-leven toch mooi zijn!

Ben je een voorbeeld voor mij als bestuurder? Ja en nee, maar vooral ja. *Ja* om meerdere redenen. Jij hebt als geen ander het vermogen complexe situaties te doorzien en de vinger precies op de zere plek te leggen. Jij kunt in één zin een hele discussie terugbrengen tot de essentie, die vaak verrassend eenvoudig is. Dat is echt een kwaliteit, die indruk op mij heeft gemaakt. Jij hebt ook als geen ander het lef om

mensen direct aan te spreken en de strijd aan te gaan als dingen fout lopen. Jij bent verbaal zeer begaafd, kunt je gedachten heel scherp (en niet onbelangrijk: op de juiste toon) onder woorden brengen en je mening heel goed beargumenteren. In meerdere situaties heb ik met bewondering naar je geluisterd en heb ik gezien hoe goed je bent. Ik denk aan de GLO-discussie over de functiewaardering, de oefensessie met de SAC, de site-visit, en de moeilijke gesprekken rond een aantal personele conflicten. *Nee* in de zin dat jouw stijl van leidinggeven in wezen top-down is, en dat je anderen van jouw mening probeert te overtuigen. Ik zoek meer de consensus, ben meer van het poldermodel. Voor beide stijlen is heel veel te zeggen.

Ben je een voorbeeld voor mij als wetenschapper? *Nee*. Jij en ik zijn totaal anders. Jouw helden zijn mijn helden niet. Jij koestert nog altijd het ideaalbeeld van een wetenschapper als een genie dat met pen en papier mooie algoritmen ontwikkelt, en ziet valorisatie toch uiteindelijk als een noodzakelijk kwaad. Al kun je dat als directeur CWI natuurlijk niet hardop zeggen. Ik vind puzzelen aan modellen ook leuk, maar ga graag wat vaker naar buiten om te praten met mensen van bedrijven en overheden over wat we voor elkaar zouden kunnen betekenen. Valorisatie dus. Ieder zijn ding, maar beide zijn belangrijk. Het mooie aan het runnen van een succesvol onderzoeksinstituut is dat je, zeker vandaag de dag, beide soorten wetenschappers echt nodig hebt. Wat dat betreft vullen wij elkaar uitstekend aan.

Wat ik nog even kwijt wil. Ik weet inmiddels uit eigen ervaring dat leidinggevenden zelden of nooit complimenten krijgen over hun functioneren. Laat ik het daarom hier nog maar eens zeggen: ik vind dat jij het als directeur CWI uitstekend gedaan hebt. Ik heb van dichtbij meegemaakt hoe zwaar de functie van directeur is, en dat je er als directeur *after all* alleen voor staat en moet beslissen of we rechts- of linksom gaan. Dat anderen daarna altijd kritiek hebben op jouw beslissingen is een gegeven. Dat zijn de beste stuurlui en die hebben altijd makkelijk praten. Daarom diep respect.

Beste Jan Karel, ik dank je heel hartelijk voor alles wat je voor het CWI, en voor de Wiskunde in Nederland, betekend hebt, en dat is veel. Er breekt een nieuw CWI-tijdperk aan, zo gaat dat, en dat is ook goed. En ik weet zeker dat je de komende jaren met jouw kennis en ervaring nog heel veel voor het CWI zult betekenen. Ik wens jou, samen met Karen en de kinderen, alle geluk toe, nu en in de toekomst. Het ga jullie goed.

Rob van der Mei

Köln Hauptbahnhof

"Kijk", zei Jan Karel, en tilde bij wijze van begroeting een cassetterecorder op. "Dat is een draagbare cassetterecorder", dacht ik.

Ik slenterde over het perron op Köln Hauptbahnhof toen ik hem zag staan.
Het was op zondag 13 januari 1991, zo rond de middag schat ik. We waren beiden op weg naar Oberwolfach; ik liep op hem af om hem te begroeten.

"Kijk", zei Jan Karel, bij wijze van begroeting, "Ik ga Jack interviewen. Hij zou een stuk schrijven voor in het boekje over de geschiedenis van de mathematische programmering dat we voor het symposium samenstellen. Maar tot nu toe heeft hij dat nog niet gedaan, dat snap je, en het gaat ook niet gebeuren. Dus ga ik hem interviewen en dat drukken we dan af. Dat wil zeggen, George gaat hem interviewen, want die heeft meer gezag en met Jack weet je maar nooit. Nou, en dan bedien ik dit ding wel. Ik ben benieuwd of het wat wordt."

Hij kijkt mij parmantig en licht triomfantelijk aan, zo van: is dit nu geen mooi plan?.

"Dag Jan Karel", groet ik terug, "dat is een mooi plan, ik ben benieuwd wat het wordt." De trein liep binnen; we zochten ieder onze gereserveerde plaats op en gingen verder, richting Oberwolfach.

Bert Gerards

Nawoord: Het werd wat, vanzelf, met Jan Karel wordt het altijd wat, zie maar: Jack Edmonds (1991). A Glimpse of Heaven. Jan Karel Lenstra et al. (eds.) *History of Mathematical Programming, A Collection of Personal Reminiscences*, North-Holland, Amsterdam, 32-54.

Grijns

Als promovendus in de stochastische besliskunde aan de TU in Eindhoven deelde ik de kamer met enkele promovendi van Jan Karel. De verhalen die ik over de capaciteiten van Jan Karel hoorde (zo beweerde men dat hij binnen een oogopslag kon zien of een punt in Latex al dan niet in math environment stond) wekten de nodige ontzag en ongeloof op.

De puntigheid van Jan Karel bleek echter niet overdreven. Zo leerde ik tijdens een promotie van een collega de kracht van Jan Karel om dingen beknopt, maar treffend te duiden. Hoe effectief het gebruik van de term 'scheurkalenderwijsheden' kan zijn, zal me altijd bijblijven.

Zelf leefde ik met vertrouwen naar mijn promotie toe, maar toen ik begreep dat Jan Karel, die het inmiddels tot decaan van de faculteit geschopt bleek te hebben, de vergadering zou leiden, werd ik toch wat zenuwachtiger.

De zenuwen verdwenen al snel toen Jan Karel me vooraf verzekerde geen vraag te zullen stellen zolang ik maar de tijd zou nemen voor de vragen van de overige commissieleden. Echter, tijdens de promotie overviel de paniek me toen ik me realiseerde dat sommige vragen van de overige commissieleden nogal makkelijk waren. Wat te doen? Ik besloot bij enkele vragen net te doen alsof ze moeilijk waren en te acteren om zodoende wat tijd te winnen.

Ik bracht het tot het einde, maar was natuurlijk benieuwd hoe het gegaan was. De stoet van hoogleraren stond op en vertrok. Iedereen keek strak voor zich uit. Niemand gaf een krimp. Ik was ietwat teleurgesteld over het feit dat ik nog een minuut of twintig zou moeten wachten op de uitslag. Groot was mijn verbazing dat de laatste hoogleraar in de stoet, degene van wie ik dat het minst verwachtte, wel iets losliet. Jan Karel keek naar me met een dikke vette grijns. Hij bleek later als enige mijn acteerpoging te hebben doorzien.

De stilzwijgende en tegelijkertijd alleszeggende grijns zal ik nooit vergeten en is een goede illustratie van Jan Karel: op het eerste gezicht afstandelijk, beknopt en bijna intimiderend; bij nadere kennismaking vooral eerlijk, open en gedreven. Ik zal je missen!

Bert Zwart

Operations Research and the Life Sciences – A Healthy Combination

Gunnar W. Klau GUNNAR.KLAU@CWI.NL
CWI, Life Sciences Group, Amsterdam, the Netherlands

Led by Jan Karel Lenstra, a couple of years ago, initiatives have been started to set up a Life Sciences Group at CWI. As a direct beneficiary of these initiatives I am very thankful to Jan Karel Lenstra and to those who helped him with this task.

Here, I will illustrate some links between Operations Research (O.R.), Jan Karel's home turf, and the life sciences. Indeed, now in 2011, a large number of people in CWI's Life Sciences Group are busy with developing O.R. techniques to solve problems from biology. In the following I will describe a general link between O.R. and the life sciences and explain why the same techniques can be used for problems arising in economy and in biology. I will focus on one such example, namely the protein design problem and the problem of assigning frequencies to radio links in telecommunication, which turn out to be close mathematical cousins. Finally, there are some personal remarks.

O.R. and the Life Sciences

A 2004 *Boston Globe* interview with Mark Eisner from Cornell University contains a nice definition of the field of Operations Research (O.R.): *He [Eisner] defines O.R. as "the effective use of scarce resources under dynamic and uncertain conditions." That may sound arcane, but it's pretty much the problem of living—and certainly the central problem of economic life. O.R. isn't economics, however, though most economists have some O.R. training. It's applied mathematics.*

There are three things I like in this definition:

- It clearly states that O.R. is applied mathematics.
- The definition is less dry than more "official" ones like "interdisciplinary field that develops advanced analytical methods to come up with optimal or near-optimal solutions to complex decision-making problems". The issue of invisibility that O.R. shares with other mathematical disciplines is also commented on by Eisner in the same interview by saying that O.R. is "probably the most important field nobody's ever heard of. Indeed, it's not one that's likely to come up at dinner parties."
- It provides a beautiful link to the life sciences by saying that it is "pretty much the problem of living". Going one step further one might say that problems cells face are not so different from the problems that companies face. This analogy then allows for using similar mathematical techniques to solve these problems.

There is already a rich history of applying O.R. to problems in healthcare and applied medicine. Examples include radiotherapy treatment design, robotic surgery, location of healthcare facilities, design of clinical trials, medical ressource allocation, and the optimal scheduling of vaccines.

Recently, however, O.R. is likewise emerging as a crucial component of basic research in modern biology and biomedicine. Microarray technology, next-generation sequencing and advances in proteomics make it now possible to investigate and compare entire genomes or proteomes under different physiological conditions. This leads to new mathematical challenges and has already spurred the development of new algorithms and statistical techniques.

For instance, the shotgun approach to genome assembly leads to a large-scale Traveling Salesman-like graph optimization problem the solution of which has been key to the draft release of the human genome. In phylogenetics, researchers now have moved from the study of trees to networks that allow to model evolutionary events like recombination and lead to new problems that should be addressed with O.R. techniques. More O.R. challenges arise from the study of genetic variations and from personalized medicine in general. In sequence analysis, Bayesian modeling and inference and optimization-based machine learning techniques such as support vector machines as well as data-mining techniques like classification and clustering are popular approaches. Many problems in structural biology and drug design such as protein structure prediction, docking, side-chain placement in protein design and

the comparison of 3D structures benefit from a variety of O.R. techniques.

Last, but not least, systems biology is a recent approach where O.R. has started to play a prominent role. Biologists have focused on studying interactions between components in a biological system. These interactions are frequently described as biological networks that exhibit highly dynamic and interactive behaviour. Examples are metabolic, signal transduction, genetic, and protein-protein interaction networks. Numerous computational challenges arise in the field. While some of them can already be addressed with traditional O.R. techniques like network flows or linear programming, many new problems ask for entirely novel techniques.

Protein Design and Frequency Assignment

This section highlights one example where O.R. links two seemingly unrelated problems—protein design in synthetic biology and radio link frequency assignment in telecommunications. It is a particularly suitable example for this booklet, because it also links work that has been done in CWI's Life Sciences Group with work that has been done by Jan Karel Lenstra.

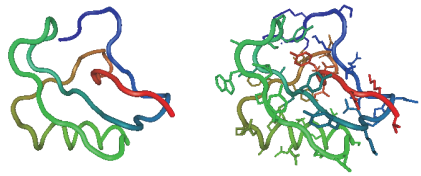
Computational protein design aims at constructing novel or improved functions on the structure of a given protein backbone. Since proteins are key players in virtually all biological processes, the ability to design proteins is of great practical interest, e.g., to the pharmaceutical and biotechnological industries. Experimental methods are time- and money-consuming. Computational approaches are an attractive alternative.

Given the modeled backbone of a protein, the amino acid side-chains have to be placed on this backbone in the energetically most favorable conformation. In protein design instances the candidate side-chains at a certain residue position correspond to more than one amino acid. The choice of the side-chain then determines the amino acid for the design. Two assumptions are commonly made: (i) side-chains adopt only statistically dominant low-energy side-chain conformations, the so-called rotamers, and (ii) the energy of a protein is the sum of intrinsic side-chain energies and pairwise interaction energies. These assumptions lead to the following NP-hard discrete optimization problem: For each residue position choose a rotamer such that the total energy of the protein is minimum. See Figure 1 for an illustration.

A formal graph-theoretic reformulation of the side-chain placement problem is as follows: Given a k-partite graph $G = (V, E)$, $V = V_1 \cup \ldots \cup V_k$, with node costs c_v, $v \in V$, and edge costs c_{uv}, $uv \in E$, determine an assignment $a : \{1, \ldots, k\} \to V$ with $a(i) \in V_i$, $1 \leq i \leq k$, such that the cost

$$\sum_{i=1}^{k} c_{a(i)} + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} c_{a(i)a(j)}$$

of the induced graph is minimum.



(a) Modelled protein backbone without side chains (b) Specific side chain placement

Figure 1. Computational protein design. The task is to determine the placement of side-chains that results in the lowest overall energy. The placement of side-chains determines the amino acids to be used for designing the protein. The right figure shows one feasible solution

Here, each node set V_i corresponds to the candidate rotamers for the set of possible residues at position i. Node costs model self energies of rotamers and edge costs model interaction energies between pairs of rotamers. A solution is given by selecting for each residue position i, $1 \leq i \leq k$, exactly one rotamer $a(i)$. Clearly, the choice of the rotamer determines also the amino acid at this position.

In (Canzar et al., 2011) we propose a novel exact method for the side-chain placement problem that works well even for large instance sizes as they appear in protein design. Our main contribution is a dedicated branch-and-bound algorithm that combines tight upper and lower bounds resulting from a novel Lagrangian relaxation approach. This makes it possible to optimally solve large protein design instances routinely. The project has been supported by a CWI internship to make the stay of PhD student Nora Toussaint possible.

Interestingly, the side-chain placement problem is mathematically very related to the frequency assignment problem in telecommunications as we will see in the following. The presentation of the frequency assignment problem is mainly taken from the report on the CALMA project (Aardal et al., 2002), in which Jan Karel Lenstra and his colleagues investigated many of its problem variants.

The problem appears when setting up and configuring a wireless communication network. Frequencies have to be assigned to radio links in order to enable communication. However, due to the dramatical increase in wireless communication, frequencies have become a scarce resource. Thus, same or similar frequencies have to be reused for different radio links. This may lead to interference problems, for example, when the involved links are close to each other. There are many variants of frequency assignment problems. In the following, we will focus on the minimum interference problem.

In the minimum interference problem we are given a finite set L of radio links, where each link i has to receive a frequency from a finite domain D_i. For pairs of interfering links the assigned frequencies must differ by more than a given distance d_{ij}. This constraint may be soft or hard, that is, its violation may either be penalized in the objective function or not allowed at all. For simplicity of the presentation we omit preassigned frequencies and special treatment of so-called parallel links. The interested reader is referred to (Aardal et al., 2002).

Formally, this simplified minimum interference problem is as follows: Find an assignment of frequencies $f_i \in D_i$ for each $i \in L$ with $|f_i - f_j| > d_{ij}$ for each pair of links $\{i, j\}$ with a hard interference constraint such that

$$\sum_{C} c'_{ij} \delta(|f_i - f_j| \leq d_{ij})$$

is minimum. Here, $\delta(\gamma) = 1$ if condition γ is true and $\delta(\gamma) = 0$ otherwise, and C denotes pairs of $\{i, j\}$ with soft interference constraints, where c'_{ij} is the cost for violating such a constraint.

Lemma 1. *The simplified minimum interference problem is a special case of the side-chain placement problem.*

Proof. Let L, D_i, for all $1 \leq i \leq |L|$, $C \subseteq \binom{L}{2}$, d_{ij} for all $\{i, j\} \in \binom{L}{2}$, c'_{ij} for all $\{i, j\} \in C$, be an instance of the simplified minimum interference problem. We can easily transform it into an instance of the side-chain placement problem as follows:

We set $k := |L|$ and $V_i = D_i$ for all $i = 1, \ldots, k$. We set all node weights c_v to zero for all $v \in V$. We now set the edge weights c as follows:

$$c_{ij} = \begin{cases} 0 & |f_i - f_j| > d_{ij} \\ c'_{ij} & |f_i - f_j| \leq d_{ij} \text{ and } \{i, j\} \in C \text{ (soft)} \\ \infty & |f_i - f_j| \leq d_{ij} \text{ and } \{i, j\} \notin C \text{ (hard)} \end{cases}$$

It is easy to see that the simplified minimum interference problem has a complete assignment if and only if the side-chain placement problem has a solution of total cost $< \infty$. In this case, solution and solution value of the two problems coincide. □

As a consequence of Lemma 1, we could use the algorithm developed for protein design instances to solve instances of the simplified minimum interference problem. It is interesting future work to see how our algorithm performs on instances arising from telecommunications and how to integrate the constraints that have been omitted in the simplified problem statement.

Personal Remarks

There are many more success stories of O.R. in the life sciences and many more will certainly follow in the future. I am happy and proud that we, the Life Sciences Group, have been given the opportunity to contribute to this exciting area at CWI, and I am personally very grateful for Jan Karel's support.

I also want to thank Jan Karel Lenstra on this occasion for the elusive conditions and for the great atmosphere we have at this institute. Among many other examples I could write about I refer to the one shown in Figure 2.

Figure 2. Jan Karel Lenstra, my son Theo, and myself at the CWI midsummer barbecue 2009. Theo has just won the first prize in his life: free entry to Artis for one year.

References

Aardal, Karen I, Hurkens, C, Lenstra, Jan Karel, and Tiourine, S. Algorithms for radio link frequency assignment: The CALMA project. *Operations Research*, pp. 968–980, 2002.

Canzar, Stefan, Toussaint, Nora C, and Klau, Gunnar W. An exact algorithm for side-chain placement in protein design. *Optimization Letters*, March 2011.

Discrete explorations of multicellular growth and morphogenesis

Roeland M.H. Merks*†‡

September 1, 2011

1 Introduction

It was a new, and somewhat scary adventure for me to start working in a mathematical environment: after my undergraduate studies in biology, I had worked on problems in theoretical biology in computer science, biology, and physics institutes, but never had I been so close to "real" mathematicians. I joined CWI via the Netherlands Consortium for Systems Biology in September 2008, and became part of MAC-4, the Life Sciences group. Would the mathematicians not frown upon our simulation methods? Our methods help develop new hypotheses for biologists, show how complex biological phenomena can be driven by simple underlying rules, and also produce beautiful pictures, but they do not produce any real proofs or theorems.

My worries were quickly taken away after having received a warm welcome at CWI, and after the introduction to Jan Karel Lenstra. The original Escher print in Jan Karel's office made me feel more at ease among mathematicians. The mathematics of life and biological morphogenesis is prominent in the work of M.C. Escher: apart from the metamorphosis from fish to bird in Jan Karel's office, Escher made many lizard tessellations, and of course the *Pedalternorotandomovens centroculatus articulosus* - the cute rolling creature similar to the extant pill millipede (Figure 1). Later on I learned that Jan Karel had been one of the main proponents of building a Life Sciences group at CWI, and also helped to establish the Netherlands Institute for Systems Biology that ensures that the models, simulation, and methods are developed in collaboration with experimental biologists, so they are applied to "real life" applications.

1.1 Biological morphogenesis

What kind of "real life" applications does my group work on? Here I would like to focus on morphogenesis, the mechanisms driving biological growth and form. A classic

*Centrum Wiskunde & Informatica, MAC-4 (Life Sciences), Science Park 123, 1098 XG Amsterdam
†Netherlands Consortium for Systems Biology / Netherlands Institute for Systems Biology, Science Park 123, 1098 XG Amsterdam
‡merks@cwi.nl

A

B

Figure 1: A "real life" and "mathematical" creature side by side. **A.** Pill millipede, possibly from order *Glomerida*. **B.** *Pedalternorotandomovens centroculatus articulosus*, or *wentelteefje* by M.C. Escher.

mathematical study of morphogenesis is Alan Turing's seminal paper "The chemical basis of morphogenesis"[16]. Turing showed how periodic chemical patterns, *e.g.*, the signals eliciting the growth of fingers or the stripes of a zebra, could arise from a homogeneous initial condition of two chemical substances, the *morphogens* A and B,

$$\begin{aligned} \frac{\partial A}{\partial t} &= a(A-h)+b(B-k)+D_A\frac{\partial^2 A}{\partial x^2}, \\ \frac{\partial B}{\partial t} &= c(A-h)+d(B-k)+D_B\frac{\partial^2 B}{\partial x^2}, \end{aligned} \tag{1}$$

with a, b, c, and d production or degradation rates for morphogen A and B. h and k are the homogeneous steady states for A and B, and D_A and D_B are the diffusion coefficients for morphogen A and B. Turing showed that uniform "stationary waves with finite wave-length", that is stripes or spots, are a steady state of this system if one of the morphogens diffuses much faster than the other.

Turing beautifully showed how patterns can arise out of "nothing" in a strip of tissue consisting of identical cells exchanging chemicals. My group builds upon Turing's pioneering work, and attempts to explain morphogenesis as a problem of self-organization of moving, growing and dividing cells. The cells in a multicellular organism follow simple rules, encoded by the information in their DNA. These rules guide cell behavior: how does a cell move, how strongly does it adhere to surrounding cells and to extracellular materials, what materials does it secrete itself, etc. Perhaps more importantly, the cellular rules determine how a cell will respond to signals from its neighbors, and to signals the cells have deposited in the extracellular space. In a typical response, they would "switch on" a different set of cellular rules: they could, for example, change their shape, start or stop moving, or tighten the adhesion to adjacent cells. Such shifts in cell behavior occur, *e.g.*, during "cell differentiation".

Using cell-based computer simulations, a form of agent-based simulations in which the agents are biophysically detailed models of cells [9, 11], we predict the emergent

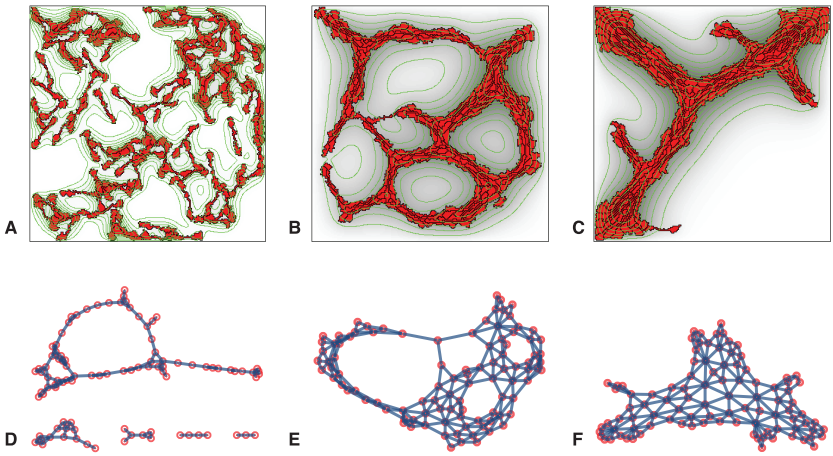


Figure 2: Snapshots from a Cellular Potts simulation of *de novo* formation of a vascular network from dispersed endothelial cells (panels **A-C**) and corresponding graph representations (panels **D-F**). Snapshots taken at 50 (**A**, **D**), 2000 (**B**, **E**), and 20000 (**C**, **F**) Monte Carlo steps.

collective behavior following from the stochastic behaviors of many, discrete cells, that follow a discrete set of cellular rules. In this way, we have studied how cells form blood vessels [7, 12], and how other cells with different rules form plants [8, 10].

1.2 Biological tissues are graphs

Did I just write "discrete"? Discrete cells, discrete rules? Yes, cells are discrete objects. Yet, many of the simulation methods in morphogenesis rely on continuum methods, including PDEs. This is a very powerful approach if we are interested in phenomena at large scales in tissues consisting of more-or-less identical cells, so we can safely "smoothen" away the individual cells. It allows us to make use of a wealth of tools in dynamic systems theory: find bifurcations, do linear stability analyses... But at smaller scales, and in more heterogeneous tissues, we cannot ignore biology's innate discreteness. For this reason, cell-based simulation methods model cells as individual entities. A challenge in cell-based models, however, is that the analytical tools currently lack the level of sophistication of tools for continuum models. To what extent could such methodology be developed using the tools available to discrete mathematicians? Cells are discrete objects, so could discrete mathematics methods be of use in analyzing multicellular morphogenesis?

In this contribution, I will explore the use of graph theory in analyzing the cellular

configurations coming from our simulations of blood vessel growth. As a first step, we transform the simulated biological tissue into a graph that represents the adjacency relations of cells. That is simple: cells become nodes. If two cells are adjacent to each other, an edge connects them. Here I only focus on topology; that is, I throw away the locations of cells, and only keep the adjacency relations. Figure 2 shows an example of such a transformation. In the remainder, I will first briefly outline the cell-based algorithms we use to simulate blood vessel growth, and then outline some preliminary explorations in the use of graph theoretic measures for characterizing blood vessel sprouting. I will end by discussing some future developments in the discrete mathematics of biological morphogenesis.

2 Methods

2.1 Hybrid cellular Potts model

We modeled cell motility at a mesoscopic level using the Cellular Potts model (CPM) [4, 5]. The CPM is a lattice-based Monte-Carlo approach that describes biological cells as spatially-extended patches of identical lattice indices $\sigma(\vec{x})$ on a square or triangular lattice, where each index uniquely identifies, or labels a single biological cell. Connections (links) between adjacent lattice sites $(\vec{x_1}, \vec{x_2})$ of unlike index, $\sigma(\vec{x_1}) \neq \sigma(\vec{x_2})$, represent bonds between apposing cell membranes, where the bond energy is $J(\vec{x_1}, \vec{x_2})$, assuming that the types and numbers of adhesive cell-surface proteins determine J. A penalty increasing with the cells deviation from a designated target volume $A_{\text{target}}(\sigma)$ imposes a volume constraint on the simulated cells. We define the pattern's effective energy:

$$H = \sum_{(\vec{x_1},\vec{x_2})} J(\sigma(\vec{x_1},\vec{x_2}))(1-\delta(\sigma(\vec{x_1}),\sigma(\vec{x_2})))+\lambda_A \sum_{\sigma}(a(\sigma)-A_{\mathrm{T}}(\sigma))^2+\lambda_L \sum_{\sigma}(l(\sigma)-L_{\mathrm{T}}(\sigma))^2, \tag{2}$$

where $(\vec{x_1}, \vec{x_2})$ is a pair of adjancent lattice sites, $a(\sigma)$ is the current area of cell σ, and $A_{\mathrm{T}}(\sigma)$ is its target area. The parameter λ_A represents the cell's resistance to compression, and δ is the Kronecker delta. The last term represents a cellular length constraint, with $l(\sigma)$ representing the current length of cell σ, L_{T} a target length, and λ_L a parameter setting the cell's rigidity along its long axis [7].

To mimic the cytoskeletally-driven cell extensions and retractions that drive cell motility, we randomly choose a source lattice site $\vec{x_1}$, and attempt to copy its index $\sigma(\vec{x_1})$ into a randomly-chosen adjacent lattice site $\vec{x_2}$. During a Monte Carlo Step (MCS) we carry out N copy attempts, with N the number of sites in the lattice. We calculate how much the effective energy change, ΔH, that would occur if we performed the copy, and accept the attempt with a Boltzmann-like probability,

$$P(\Delta H) = \{e^{\frac{-\Delta H}{M(\sigma)}}, \Delta H \geq 0; 1, \Delta H < 0\}, \tag{3}$$

where $M(\sigma)$ defines an intrinsic random cell motility M of cell σ, that drives energetically unfavorable movements.

We assume that endothelial cells secrete a morphogen $c(\vec{x},t)$ that diffuses and degrades in the extracellular matrix,

$$\frac{\partial c}{\partial t} = \alpha(1-\delta(\sigma(\vec{x}),0) - \epsilon\delta(\sigma(\vec{x}),0)c + D\frac{\partial^2 c}{\partial x^2}, \tag{4}$$

with $\delta(\sigma(\vec{x}),0) = 0$ inside cells and $\delta(\sigma(\vec{x}),0) = 1$ otherwise. Here α is the rate at which cells secrete the morphogen c, ϵ denotes the morphogen's degradation rate, and D is the morphogen diffusion. We further assume that cells chemotact towards the morphogen: they preferentially move towards higher concentrations of c. To do so, we include an additional chemical effective-energy change, $\Delta H_{\text{chemotaxis}}$, at each copy attempt [15] from site $\vec{x_1}$ into adjacent site $\vec{x_2}$,

$$\Delta H_{\text{chemotaxis}} = -\mu(c(\vec{x_2} - \vec{x_1})). \tag{5}$$

In this way the algorithm favors cellular extensions into the direction of higher concentrations of the morphogen. We solve the partial differential equations for chemoattractant diffusion and degradation (Eq. 4) numerically using a finite-difference scheme on a lattice matching the CPM lattice, with $\Delta x = 2\ \mu m$. We use 15 diffusion steps per Monte Carlo step, with $\Delta t = 2\ s$, thus equating a Monte Carlo step to 30 s real time.

2.2 Cellular networks as graphs

To study a cellular pattern as a graph, we first construct a graph. We represent patches of identical indices $\sigma = i$ (*i.e.*, a biological "cell") by node i. Nodes i and j share an edge $e(i,j)$ if,

$$\exists_{(\vec{x_1},\vec{x_2})} : \sigma(\vec{x_1}) = i \wedge \sigma(\vec{x_2}) = j \wedge \sigma(\vec{x_1}) \neq \sigma(\vec{x_2}). \tag{6}$$

So cell i and cell j share an edge $e(i,j)$ if there exists an adjacent pair of lattice sites $(\vec{x_1},\vec{x_2})$ such that $\sigma(\vec{x_1}) = i$ and $\sigma(\vec{x_2}) = j$. We exclude self-connected nodes.

The resulting graphs can be studied using standard tools from graph theory. To do so, we made use of the graph tools implemented in *Mathematica 8* and the algorithms in the *Mathematica Graph Utilities Package*. In this pilot study, we looked only at the above defined adjacency relations of cells. In more extensive studies the nodes can be given attributes, including cell type, surface area, or cellular locations. No additional notation is required to represent these node attributes, because the node indices uniquely identify the cells.

3 Results

3.1 Simulation of *de novo* formation of a blood vessel network

Figures 2A-C show an example of blood vessel network formation in our model, with an appropriate set of parameters (Table 1 lists the parameters used). After the initial formation of the chemical gradients, the cells move towards one another and form a

$\lambda_A = 50$	$A_{\mathrm{T}} = 400\ \mu m^2$
$\lambda_L = 5$	$L_{\mathrm{T}} = 120\ \mu m$
$J_{\mathrm{cell,medium}} = 20$	$J_{\mathrm{cell,cell}} = 40$
$M(\sigma) = 50$	$D = 1.0 \times 10^{-13}\ m^2 s^{-1}$
$\alpha = 1.8 \times 10^{-4}\ s^{-1}$	$\epsilon = 1.8 \times 10^{-4}\ s^{-1}$
$\mu = 1000$	
Grid size: 200×200 lattice sites	Number of cells: 100

Table 1: Parameters used for the *de novo* vasculogenesis simulation studied in this paper

network structure closely resembling real-life blood vessel networks in animals and in cell cultures [6]. Also note that the initial fine-grained network slowly coarsens over time, a phenomenon also seen in cell cultures [6]. A movie of this simulation is available at `http://www.youtube.com/watch?v=VbeRFcl6RD4`. For more details on the mechanisms driving such blood vessel sprouting, the relation to actual blood vessel growth, and extensive parameter studies of this model, see our previous and ongoing work [12, 6, 7]. Figures 2D-F show the graph representations corresponding with the blood vessel networks.

3.2 Variability of graph measures marks morphological events

What graph measures could we use to analyze the morphological transformation from a set of disconnected cells, to a fine-grained network, and into a coarse-grained network? To detect the first occurence of a vascular network, the first obvious measure to look at is the number of connected components in the graph: the number of connected components drops rapidly during the first steps of the simulation, as the cells move towards one another (Figure 3A). After around 300 Monte Carlo steps all cells have connect to one another.

After the first connections between cells have been made, the blood vessel network slowly coarsens over time; a good measure of graph complexity should capture this coarsening process. Also, we would like to have a good definition of the number of branches in the simulated networks. We first identified the *eccentricity* of the vertices in the graph, where the eccentricity of a node is the longest of all shortest paths to all other nodes in the graph. We expect nodes in the branches to have high eccentricity. Another useful measure is the *graph periphery*, defined as the set of vertices with maximum eccentricity. We noticed that many branches typically contained a set of connected peripheral nodes; so simply counting the peripheral nodes overestimates the number of branches. We therefore defined the number of *connected peripheral subgraphs* as the number of connected components in the graph periphery. We plot this measure in Figure 3B as a function of time. The number of connected peripheral subgraphs jitters over time, suggesting that it is very sensitive to the small variations in morphology due to our Monte-Carlo simulation algorithm.

Interestingly, between time steps 20000 and 40000 and after time step 80000 we observed a temporary drop of the values and variability of the number of connected pe-

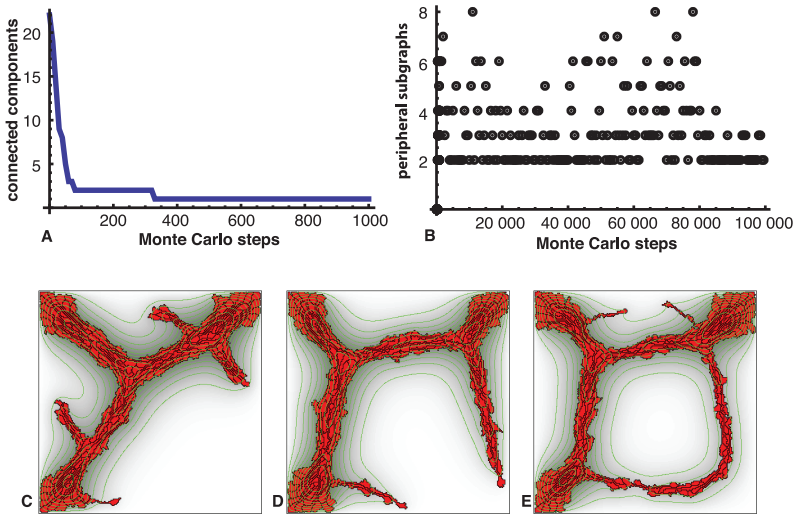


Figure 3: **A** Number of connected components of the vascular network over time (in Monte Carlo steps). **B** Number of connected peripheral subgraphs in the vascular network over time. **C** Snapshot of simulation at 20000 Monte Carlo steps, **D** at 30000 Monte Carlo steps, and **E** at 40000 Monte Carlo steps.

ripheral subgraphs. What happened in our simulation during this period? Figures 3C-E show three snapshots of our simulation at 20000, 30000, 40000 timesteps. Two large sprouts form; at around step 40000 they touch and form a large "hole" in the morphology. It steadily shrinks and disappears around step 60000. Once the two sprouts have reconnected the number of connected peripheral components increases again. A similar event occurs after timestep 80000, whereas we have not observed other, similar events in the simulation. Thus these observations suggest that the variation of the number of connected peripheral subgraphs may help to identify morphological transformations in our simulations. We are now exploring a) if other measures more reliably identify the number of branches in these networks, and b) if indeed the *time-variation* or *structural stability* of graph measures may be a useful measure for other dynamical morphological features as well.

3.3 Network complexity: community partition structure

Changes in the variability of the eccentricity-based graph measures may help to identify morphological events, as we have seen in the previous section. However, exactly because of their high variability we cannot use them for the morphological analyses of static vascular networks. We next explored the use of community structure measures [14]. A

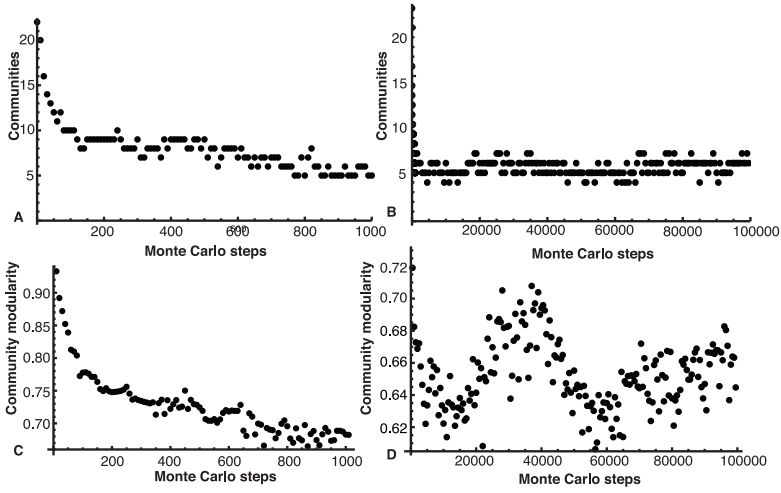


Figure 4: Community structure of vascular networks **A-B** Optimal number of community partitions detected in the vascular network over time. **A** Monte Carlo steps 0 to 1000; **B** Monte Carlo steps 0 to 100000. **C-D** Community modularity (Q) of optimal community partition of vascular networks. **C** Monte Carlo steps 0 to 1000; **D** Monte Carlo steps 0 to 100000.

community in a graph is a subset of the graph that is more strongly connected internally than with the other nodes in the graph. The *community modularity* Q of a partition of the nodes V in communities V_i, is defined as

$$Q = \sum_{i=1}^{k} (e_{ii} - a_i^2), \tag{7}$$

with e_{ii} denoting the fraction of all edges in the graph internal to community i, and a_i denoting the fraction of edges running between community i and the other nodes. The community modularity is a number smaller than 1, with larger values of Q corresponding to stronger community structure. To calculate community partitions and community modularities we use the algorithm by Clauset [2] implemented in Mathematica 8.

Figure 4A shows the number of detected communities corresponding with the highest value of Q for the first 1000 Monte Carlo steps of the vasculogenesis simulation. Panel B shows the development of the number of communities for MCS 0 to 100000. During the first 1000 Monte Carlo steps the number of communities decreases from about twenty to about five communities, corresponding with the coarsening of the network pattern. Over a longer timescale, the number of communities in the simulated vascular network fluctuates around six. Thus, the number of communities nicely characterizes the slow

coarsening of the vascular network, but once the network has settled on a characteristic size, the community structure does not capture the dynamic, topological transformations of the network.

Figures 4C-D show the community modularity Q corresponding with the optimal community partition as a function of time. During the first 1000 Monte Carlo steps of our simulation, Q drops gradually from a value of above 0.9 to a value of about 0.7, corresponding with the expected reduction in community structure during network formation and coarsening. Interestingly, over a longer time scale the community modularity continues to fluctuate. The modularity drops to a value of around 0.6 around Monte Carlo step 20000 and peaks at 0.7 near Monte Carlo step 40000, after which it drops steadily until time step 60000. Possibly this variation of the community modularity corresponds with the formation of two large sprouts between Monte Carlo steps 20000 and 40000, the reconnection of the sprouts and the subsequent shrinking and closure of the resulting hole at Monte Carlo step 60000. Thus formation and closure of holes are clearly indentified. Our initial explorations suggest that changes in optimal values of the community modularity Q identify both the initial network coarsening, and more long-term, topological transformations of the vascular network.

4 Discussion

In this contribution I presented some initial explorations on the use of discrete mathematical methodology for the analysis of biological morphogenesis. I first identified the number of connected peripheral nodes for each frame of a vascular network simulation. Although this number itself varied too much to provide useful information about the structure of the vascular networks, I found that close to major morphological events its temporal variation dropped, for example during the formation of large sprouts and the fusion of these sprouts. This suggests that the robustness of a graph measure against small modifications of the graph structure can provide useful information [1, 3], even if the graph measure itself does not. Thus, a differential graph theory might be useful for the study of multicellular morphogenesis and other dynamic networks [3].

I next looked at the evolution of community modularity in a simulated vascular network. I found that the number of communities in an optimal community partition characterizes the initial coarsening of the simulated vascular network well, whereas the subsequent formation and fusion of sprouts are not well captured. However, the community modularity associated with the optimal community partition correlates well with both the initial network coarsening and the subsequent formation and fusion of sprouts.

The presented methodology indeed is a useful extension of our existing toolbox for analyzing angiogenic sprouting, based on computational geometry and mathematical morphology [13, 7, 12]. However, the real power of using graph representations of morphogenetic topological transformations may only be unleashed once we will use it to state and solve generalized problems of biological morphogenesis in formal mathematical terms. Assuming that evolution of multicellular organisms strives for optimality, we can pose various interesting questions about optimal mechanisms of biological morphogenesis. An example: assume a two-dimensional (or n-dimensional) piece of tissue, in

which cells with diameter 1 communicate via contact-dependent ("juxtacrine") signals. That is, cells must be in contact in order to pass a signal. What is the fastest way that a signal can spread through the tissue?

If the cells do not move relative to one another (epithelial tissues, plants) the answer is trivial: if a signal spread from a cell to its neighbors in time Δt, the signal spreads over a volume of radius r in time $r\Delta t$. But what if the cells shift positions and mix? How much faster or slower will the signal spread? What is the optimal way and speed for the cells to mix in order to attain optimal spread of the signal? Suppose all cells must have been in contact with one another in order for a biological mechanism to proceed. For a finite piece of biological tissue, what is the optimal mixing strategy, under the condition that cells can only exchange positions with their neighbors? How does this depend on the topology of the tissue; *e.g.*, does it differ between branched tissues (*e.g.*, lungs, kidney, gland tissue) and flat layers of tissue? If one cell type can invade a tissue of another type, how will the mixing strategy determine the spreading speed? What is the optimal way for a flat tissue to change into a branched topology? How should they locally change their adjacency relations? Do we expect a different strategy if cells can mix, as in most animal tissues, and if they cannot, as in plants and some animal tissues? And how do these insights compare with biological observation: is biology optimal?

Using cell-based computer simulations and biological experiments it is possible to get useful insights into these questions. But what is still missing from cell-based modeling approaches in biology, is a formal differential graph theory of dynamic, discrete representations of growing and developing tissues. So, Jan Karel, if you are still wondering what to work on now you can spend more time on discrete mathematics, I hope to have convinced you that the theoretical biology of multicellular morphogenesis is full of fun and challenging problems. I look forward to discuss these with you!

5 Acknowledgments

Erik van Dijk and Margriet Palm are thanked for their suggestion to use graph representations for analyzing tissue patterns; Stefan Canzar is thanked for discussion, and András Szabó is thanked for discussion and critical reading of the manuscript. This work was cofinanced by the Netherlands Consortium for Systems Biology (NCSB), which is part of the Netherlands Genomics Initiative / Netherlands Organisation for Scientific Research, and by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research.

References

[1] J. Bryden, S. Funk, N. Geard, S. Bullock, and V. A. A. Jansen. Stability in flux: community structure in dynamic networks. *J. Roy. Soc. Interface*, 8(60):1031–1040, 2011.

[2] A. Clauset. Finding local community structure in networks. *Phys. Rev. E*, 72:026132, 2005.

[3] G. Ghoshal and A.-L. Barabasi. Ranking stability and super-stable nodes in complex networks. *Nat. Commun.*, 2:394, 2011.

[4] J. Glazier and F. Graner. Simulation of the differential adhesion driven rearrangement of biological cells. *Phys. Rev. E*, 47(3):2128–2154, 1993.

[5] F. Graner and J. A. Glazier. Simulation of biological cell sorting using a two-dimensional extended Potts model. *Phys. Rev. Lett.*, 69(13):2013–2016, 1992.

[6] R. M. H. Merks and J. A. Glazier. Dynamic mechanisms of blood vessel growth. *Nonlinearity*, 19(1):C1–C10, 2006.

[7] R. M. H. Merks, S. V. Brodsky, M. S. Goligorksy, S. A. Newman, and J. A. Glazier. Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodeling. *Dev. Biol.*, 289(1):44–54, 2006.

[8] R. M. H. Merks, Y. Van de Peer, D. Inzé, and G. T. S. Beemster. Canalization without flux sensors: a traveling-wave hypothesis. *Trends Plant Sci.*, 12(9):384–90, 2007.

[9] R. M. H. Merks and J. A. Glazier. A cell-centered approach to developmental biology. *Physica A*, 352(1):113–130, 2005.

[10] R. M. H. Merks, M. Guravage, D. Inzé, and G. T. S. Beemster. Virtualleaf: An open-source framework for cell-based modeling of plant tissue growth and development. *Plant Phys.*, 155(2):656–66, 2011.

[11] R. M. H. Merks and P. Koolwijk. Modeling morphogenesis in silico and in vitro: Towards quantitative, predictive, cell-based modeling. *Math. Mod. Nat. Phenom.*, 4(5):149–171, 2009.

[12] R. M. H. Merks, E. D. Perryn, A. Shirinifard, and J. A. Glazier. Contact-inhibited chemotaxis in de novo and sprouting blood-vessel growth. *PLoS Comp. Biol.*, 4(9):e1000163, 2008.

[13] A. Mezentsev, R. M. H. Merks, E. O'Riordan, J. Chen, N. Mendelev, M. S. Goligorsky, and S. V. Brodsky. Endothelial microparticles affect angiogenesis in vitro: role of oxidative stress. *Am. J. Physiol. Heart Circ. Physiol.*, 289(3):H1106–14, 2005.

[14] M. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(2):026113, 2004.

[15] N. Savill and P. Hogeweg. Modelling morphogenesis: From single cells to crawling slugs. *J. Theor. Biol.*, 184(3):229–235, 1997.

[16] A. Turing. The chemical basis of morphogenesis. *Phil. Trans. Roy. Soc. B*, 237:37–72, 1952.

Detecting Community*

Sander M. Bohte
Centrum Wiskunde & Informatica (CWI)
Science Park 123, NL-1098 XG Amsterdam, The Netherlands

September 5, 2011

It can be argued that most important issues are related to relative placements of certain elements versus others, and the detection of proper groupings. Typography in fact is one of the clearest examples of an application domain where proper groupings are paramount. As any sane sole knows, proper typography can make crappy research at least look good, and, conversely, can make the most excellent scientific output look quite bad (e.g. figure 1) (Leslie Lamport of course solved at least the relative placement of text and figures definitively in LaTeX, using free energy minimization between spring-embedded document elements). The issue in fact extends to the traditional managerial tasks, such as the optimal group layout, and subsequent placement within the bright new CWI wing, the distribution of NWO bonus funding, and the planning of future leisure trips vs catching up on editorial work.

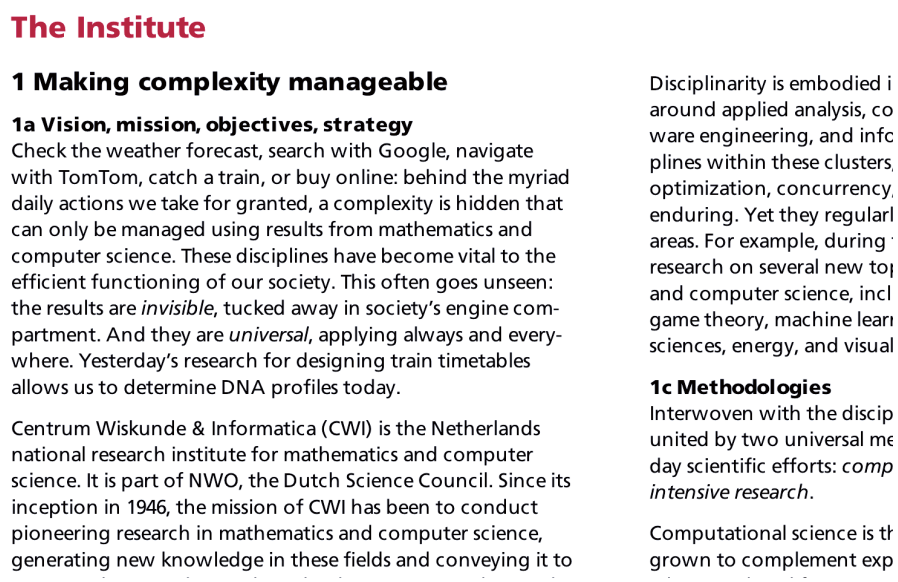
Here we consider a novel algorithmic approach for this important problem. Whereas simple Bayesian inference can solve trivial issues like wine-selection (winewinewine.com), many of the grouping and placement issues that Jan Karel wrestled with during his tenure as CWI director – and many important future decisions – can be cast into the graph-community detection domain, and as such subjected to various homeopathic solutions. For this, we turn to the vast knowledge that can be data-mined from the ~~Google~~ Bing collection of search archives – query logs.

Query Logs

Query logs contain the history of search terms entered by users into search engines such as Microsoft Bing. Since the AOL data release, it is generally well known that the most embarrassing, personal facts can be data-mined from these query-logs. From a managerial perspective, these logs thus offer many tools for management to achieve goals. In general however, old school ethical standards, combined the lack to a localized sample, prevent us from going this route. Fortunately, the usefulness of the data does not end here.

Given a large-scale query log, one of the most useful pieces of information it provides is the co-occurrence of words in different queries. The fact that two search

*The work was performed based on a Microsoft Research "Beyond Search" award, and SMB wishes to thank Microsoft Research for their support.



The Institute

1 Making complexity manageable

1a Vision, mission, objectives, strategy
Check the weather forecast, search with Google, navigate with TomTom, catch a train, or buy online: behind the myriad daily actions we take for granted, a complexity is hidden that can only be managed using results from mathematics and computer science. These disciplines have become vital to the efficient functioning of our society. This often goes unseen: the results are *invisible*, tucked away in society's engine compartment. And they are *universal*, applying always and everywhere. Yesterday's research for designing train timetables allows us to determine DNA profiles today.

Centrum Wiskunde & Informatica (CWI) is the Netherlands national research institute for mathematics and computer science. It is part of NWO, the Dutch Science Council. Since its inception in 1946, the mission of CWI has been to conduct pioneering research in mathematics and computer science, generating new knowledge in these fields and conveying it to

Disciplinarity is embodied i
around applied analysis, co
ware engineering, and infc
plines within these clusters,
optimization, concurrency,
enduring. Yet they regularl
areas. For example, during
research on several new to
and computer science, incl
game theory, machine lear
sciences, energy, and visual

1c Methodologies
Interwoven with the discip
united by two universal me
day scientific efforts: *comp*
intensive research.

Computational science is th
grown to complement exp

Figure 1: Problematic kerning. Source: 1st edition CWI Selfevaluation 2011

keywords frequently appear together in the same query is known to imply a semantic distance between them. We consider how query log data can be used to derive cluster recommendations. Such an approach should lead to answers to questions such as: What kinds of keyword combinations are useful for finding certain types of holidays? For the analysis here, we chose 50 keywords related to the tourism industry (i.e. online bookings of tickets, travel packages and such). As an illustration, this example keyword set is interesting as one's leisure time increases.

Formally, let $N(T_i, T_j)$ denote the number of times two search terms T_i and T_j appear jointly in the same query. Let $N(T_i)$ and $N(T_j)$ denote the same number of queries leading to a click, in which terms T_i, respectively T_j appear in total (regardless of other terms they co-occur with). The cosine similarity distance between terms T_i and T_j is defined as:

$$Sim(T_i, T_j) = \frac{N(T_i, T_j)}{\sqrt{N(T_i) * N(T_j)}}. \quad (1)$$

Constructing keyword correlation graphs

The most intuitive way to represent similarity distances is through a keyword correlation graph. The results from our subset of 50 travel-related terms are shown in Figure 2. In this graph, the size of each node (representing one query term) is proportional to the absolute frequency of the keyword in all queries in the log. The distances between the nodes are proportional to the similarity distance between each pair of

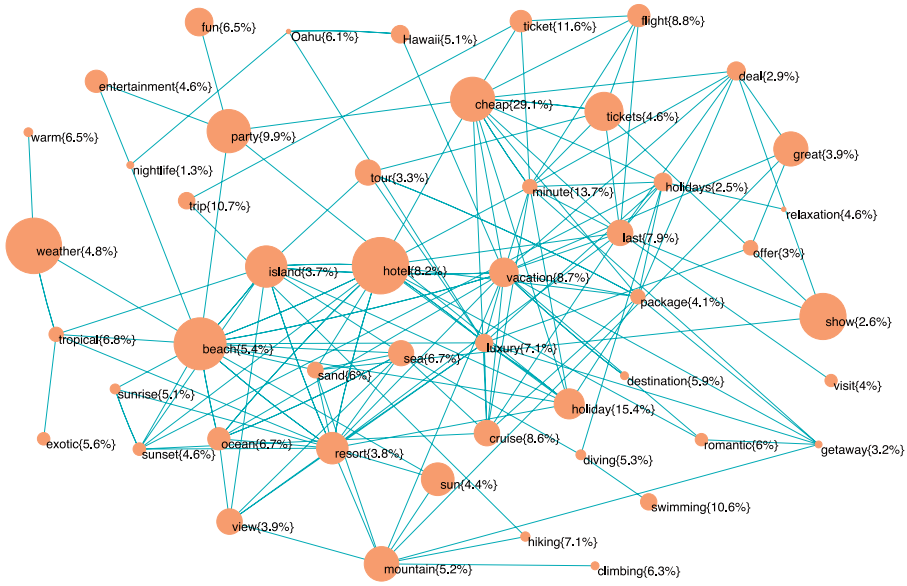


Figure 2: Visualization of a search term correlation graph, for a set of search terms related to the tourism industry. Each search term is assigned one colored dot. The size of each dot gives its relative weight (in total number of clicks received), while the distances between the dots are obtained through a spring-embedder type algorithm and are proportional to the co-occurrence of the two search terms in a query. Each dot is marked with its relative popularity with cheap-ass tour operators.

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9
beach	party	package	weather	getaway	diving	cruise	show	last
luxury	entertainment	vacation	exotic	romantic	swimming	sunrise	tickets	minute
hotel	nightlife	holidays	tropical			sunset	ticket	visit
island	fun	destination	warm				cheap	
resort	Hawaii	deal					flight	
sun	Oahu	tour						
mountain		offer						
ocean		great						
hiking								
climbing								
sea								
sand								

Figure 3: Optimal partition of the set of travel terms in semantic clusters, when the top 150 edges are considered. The partition was obtained by applying Newman's "community detection" algorithm to the graph from Fig. 2. This partition has a clustering coefficient Q=0.59.

terms, computed Equation 1, where the whole graph is drawn according to a so called "spring embedder"-type algorithm. In this type of algorithm, edges can be conceived as "springs", whose strength is inversely proportional to their similarity distance, leading to clusters of edges similar to each other to be shown in the same part of the graph. This method for instance underlies much of the LaTeX typography.

In order to visualize such query-terms in a graph (Figure 2), we used Pajek, a free and powerful academic graph drawing package. Note that not all edges are considered in the final graph. Even for 50 nodes, there are $\begin{pmatrix} 50 \\ 2 \end{pmatrix} = 1225$ possible pairwise similarities (edges), one for each potential keyword pair. Most of these dependencies represent, however, just noise in the data, and our analysis benefits from using only the top fraction, corresponding to the strongest dependencies. In the graph shown in Figure 2, only the top 150 strongest dependencies were considered in the visualization.

The most interesting effect to observe in Figure 2 are the term clusters that emerge in different parts of the graph, from the application of the spring-embedder algorithm. For example, the leftmost part of the graph has 4 terms related to weather, such as "warm", "tropical" and "exotic". On the top left part of the graph, one can find terms such as "entertainment", "nightlife", "party" and "fun", while the very bottom part includes related terms such as "climbing", "hiking" and "mountain". The central part of the graph includes terms such as "beach", "sand", "sea", "resort", "ocean", "island" etc. Additionally, pairs of terms one would naturally associate do indeed appear close together, such as "romantic" and "getaway" and "sunset" and "sunrise" and "ocean". Note that family-island-getaway is not on this map, ensuring relative quietness for this type of activity.

Automatic identification of sets of keywords

To find relevant clusters of keywords, we use so called "community detection" algorithms, also inspired by complex systems theory. In network or graph-theoretic terms, a community is defined as a subset of nodes that are connected more strongly to each other than to the rest of the network (i.e. a disjoint cluster). If the network analyzed is a social network (i.e. vertices are people), then "community" has an intuitive interpretation. However, the network-theoretic notion of community detection algorithm is broader, and has been applied to domains such as networks of Ebay items, publications on arXiv, or even food webs.

Let the network considered be represented as graph $G = (V, E)$, when $|V| = n$ and $|E| = m$. Each $v \in V$ must be assigned to exactly one group (i.e. community or cluster) $C_1, C_2, ...C_{n_C}$, where all clusters are disjoint.

In order to compare which partition is "optimal", the metric used is *modularity*, henceforth denoted by Q. Intuitively, any edge that in a given partition, has both ends in the same cluster contributes to increasing modularity, while any edge that "cuts across" clusters has a negative effect on modularity. Formally, let $e_{ij}, i, j = 1..n_C$ be the fraction of all edge weights in the graph that connect clusters i and j and let $a_i = \frac{1}{2}\sum_j e_{ij}$ be the fraction of the ends of edges in the graph that fall within cluster

Algorithm 1 *GreedyQ Partitioning*: Given a graph $G = (V, E), |V| = n, |E| = m$ returns partition $< C_1, ...C_{n_C} >$

1. $C_i = \{v_i\}, \forall i = \overline{1, n}$
2. $n_C = n$
3. $\forall i, j$, normalize e_{ij}
4. repeat
5. $\quad < C_i, C_j >= \text{argmax}_{c_i, c_j}(e_{ij} + e_{ji} - 2a_i a_j)$
6. $\quad \Delta Q = \max_{c_i, c_j}(e_{ij} + e_{ji} - 2a_i a_j)$
7. $\quad C_i = C_i \bigcup C_j, C_j = \emptyset$ *//merge* C_i *and* C_j
8. $\quad n_C = n_C - 1$
9. until $\Delta Q \leq 0$
10. $maxQ = Q(C_1, ..C_{n_C})$

i. The modularity Q of a graph $|G|$ with respect to a partition C is defined as:

$$Q(G, C) = \sum_i (e_{i,i} - a_i^2). \tag{2}$$

Informally, Q equals the fraction of edges in the network that fall within clusters, minus the expected value of the fraction of edges that would fall within the same cluster, if all edges would be assigned using a uniform, random distribution.

It is easy to see that if $Q = 0$, then the chosen partition c shows the same modularity as a random division. A value of Q closer to 1 is an indicator of stronger community structure - in real networks, however, the highest reported value is $Q = 0.75$. In practice, it was found (based on a wide range of empirical studies) that values of Q above around 0.3 indicate a strong community structure for the given network.

A greedy graph partitioning algorithm

Since we have established our framework, we can now formally define the graph partitioning algorithm. The number of possible partitions (clusters) for this problem is at least 2^{n-1} (e.g. for 50 terms: $2^{50} > 10^{15}$). Therefore, to explore all these partitions exhaustively would be clearly unfeasible. We show results for community detection using process of bottom-up successive joining of keywords, *GreedyQ*, formally specified as Algorithm 1. Initially, each of the vertices (in our case, each keyword) is assigned to its own individual cluster. Then, at each iteration of the algorithm, two clusters are selected which, if merged, lead to the highest increase in the modularity Q of the partition. The algorithm stops when no further increase in Q is possible by further merging.

The partition in figure 3, resulting directly from applying Algorithm 1, achieves a very intuitive division of keywords into relevant clusters, and, furthermore, it fits well with what can be graphically observed in Figure 2: most of the clusters obtained automatically after partitioning can be identified in different parts of the graph.

Discussion

In this short paper, we presented an approximate, greedy solution for community detection given vast data-mining resources. We believe it can be applied to much of the tough decision-making that a director is inevitably faced with. Applying it to some CWI data, we observe several strong communities, with what appears to be growing interconnections. Of course, as is well known in the algorithmic community, algorithms derive their generalizing power from the problem abstraction. For practical problems, it may be the case that additional constraints need to be considered.

Note that we by no means presented the optimal solution, as we cannot account for extreme outliers that, given the heavy tails of collective human behavior, invariably seems to lead to the Pareto principle. We do believe however that in small settings, the community detection algorithm can be applied, to such problems as holiday selection, book-topic clustering, and the combination of fine foods (again, the wine selection problem can be solved with simple Bayesian machine learning).

Acknowledgement We thank Roeland Merks for the general idea of this brief paper, and Valentin Robu for the original research and text, as well as doing the really hard work. We thank Nicole Immorlica for her transcendent contributions, and Victor Gomez for generally totally being around for us.

OROV - door Daan Crommelin

Iedere directeur in Nederland krijgt er mee te maken. De Ondernemingsraad. Dat bij wet ingestelde gezelschap. Mensen uit alle hoeken en gaten van je organisatie. Elk jaar weer nieuwe gezichten. Willen zich met van alles en nog wat bemoeien. Niet altijd ten onrechte. Bij tijd en wijle zelfs zinvol. Desalniettemin een aanslag op je kostbare tijd. Regelmatige vergaderingen (codenaam: OROV). En natuurlijk die voorzitter, die steeds weer voor je deur staat voor overleg.

Gedurende twee jaar was ik de voorzitter die Jan Karel voor zijn deur trof. Na een jaar als gewoon lid van de CWI Ondernemingsraad viel de schone taak van het voorzitterschap mij ineens toe, midden 2008. Het leek, en begon, heel overzichtelijk. Gaandeweg werd het een stuk intensiever, met reorganisatieplannen waar de OR nogal eens heel anders tegen aan keek dan Jan Karel. Er werd een hoop vergaderd en overlegd in die tijd, waarbij de sfeer soms gespannen, maar uiteindelijk altijd respectvol was.

Uit mijn begintijd als OR-lid staat me een OR-directie kennismakingssessie in het Tropenmuseum bij, onder begeleiding van een externe organisatieadviseur. De toen aanwezigen kunnen zich het rollenspel met de gekleurde petjes misschien nog herinneren. We hebben het onder andere gehad over de vraag hoe je de verhouding tussen OR en directie zag, waarbij iedereen foto's uit een stapel kon kiezen om die verhouding te verbeelden. Mij is bijgebleven dat Jan Karel op de proppen kwam met een foto van een groep mensen die gezamenlijk bezig zijn een in de modder vastgelopen bus weer op gang te krijgen. Mooi symbool. Fraaier dan het beeld van de ME'ers en demonstranten die in grimmige sfeer tegenover elkaar staan, een foto die er ook tussen zat.

Later liet Jan Karel zich een keer ontvallen dat hij bij een volgende vergelijkbare sessie wel zou willen gaan slootje springen. Helaas is dat er nooit van gekomen – de OR heeft deze kans laten lopen. Misschien een mooie activiteit om het toekomstige gemis van een OROV op te vullen? Buiten in de polder. De krachtmeting met lucht, water, modder en kroos. En, uiteraard, ieder slootje precies eenmaal over, via de kortst mogelijke route.

For JKL,

To your claim that *my domain of maths thinks Bourbaki is a pastry*, my inquiry brought to light this origin of confusion:

This dish, "Bourbaki", is from *A WSC Cook Book*. I risk a banishing by making it public in this way.

> A Num. Analyst's Cryptic Concoction: Bourbaki
>
> To start, round up:
>
> 4 axioms from Turino philosophy
> 1 work of an Oulipo madman[1]
> 10 l. unfailing logic
> a supply of intuition, physics and topology
> 1 giant portion of abstraction
> 1 shot of strong liquor (to aid in swallowing it all)
>
> Plan: Throw all intuition, physics, and topology out a window—an illustration is worth a thousand words, but a thousand words is our goal. Now, swallow all axioms on trust: a foundation for your faith. Or you can simply proclaim said axioms as truths or propositions on which your playing-card construction is built. Slowly stir in logic, until you construct a class. Go on with caution, stirring vigorously for 50 yrs, until you obtain 9-10 monographs. If you distinguish an odor of calculus on manifolds, you should quit whilst you still can.

(Following V.I. Arnol'd, who was critical of Bourbaki from a didactic standpoint, maintaining that abstract formalism must not obstruct our ability to grasp a notion. And, BTW, Bourbaki also did not know Num. Analysis.)

Much joy in your continuation of your work.

Jason Frank

[1] You might try *La Disparition*?

De JKL-2080 challenge

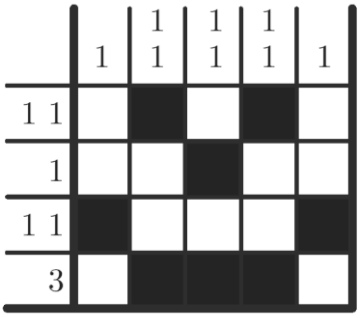
Na acht jaar aan het roer gestaan te hebben van het CWI, hoeft Jan Karel Lenstra op familiefeestjes zeker niet met een mond vol tanden te staan. Hij zou kunnen verhalen over de dijkhoogtes die *we* hebben berekend, de programmeertalen die *we* hebben ontwikkeld, of de energievraagstukken die *we* vol overgave proberen op te lossen.

Alhoewel, indruk maken op een Lenstra-feestje is wellicht toch niet zo eenvoudig. Als buitenstaander kan ik hier natuurlijk slechts naar gissen, maar ik zou me zo kunnen voorstellen dat Jan Karel daar antwoord moet geven op vragen als: "*Zeg, Jan Karel, hoe vaak ben jij eigenlijk al ver-Escherd?*" of "*En, hoeveel wereldrecords heb jij al op je naam staan?*".

Het lijkt me dan ook zaak onze vertrekkend directeur wat munitie mee te geven, zodat hij beide vragen van een passend antwoord kan voorzien.

"*Ach, Hendrik, ver-Escheren is zó 2002: ik heb mij recent nog laten Nonogrammiseren*".

Een Nonogram is een Japanse tekenpuzzel, niet te verwarren met de welbekende Sudoku. De puzzel bestaat uit een rooster van vakjes. Een aantal van die vakjes dienen zwart gekleurd te worden, zodat een afbeelding ontstaat. De zwarte vakjes in elke rij en kolom moeten voldoen aan een beschrijving: voor elke horizontale lijn is de volgorde en lengte van de segmenten van zwarte vakjes voorgeschreven. Tussen elk van die segmenten moet tenminste één wit vakjes zitten, zoals te zien is in Figuur 1. De afbeelding stelt vaak een persoon, een dier of een voorwerp voor.



Figuur 1. *Voorbeeld van een eenvoudig ingevuld Nonogram.*

Nu lijkt het wellicht eenvoudig om van een bestaande zwart-wit afbeelding een Nonogram te maken door simpelweg de bijbehorende beschrijving te berekenen. Dit leidt echter zelden tot een Nonogram met een *unieke* oplossing. In plaats van

de oorspronkelijke afbeelding zou je dus wel eens een heel ander beeld kunnen vinden met dezelfde beschrijving. Daarnaast is het niet meer mogelijk om de puzzel op basis van een reeks van logische stappen op te lossen als er meerdere oplossingen zijn.

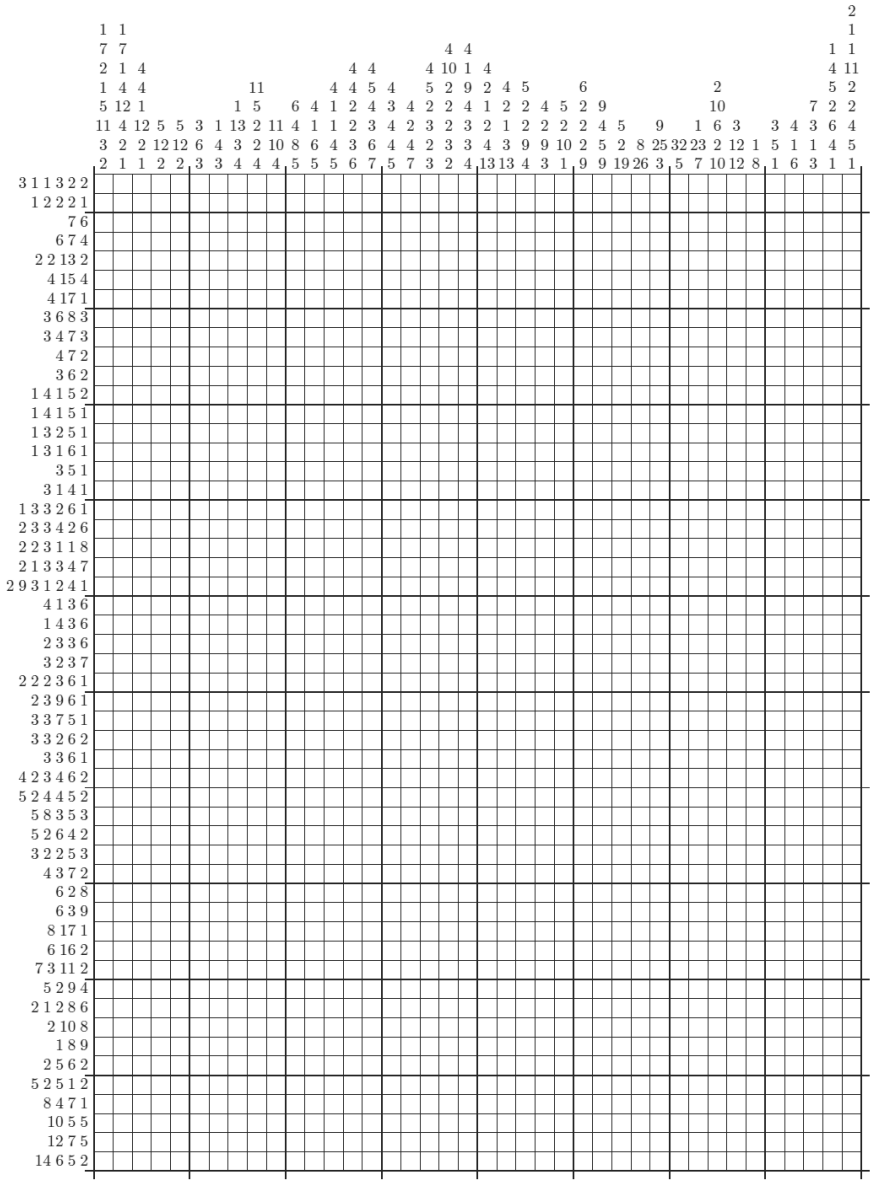
Het oplossen van Nonogrammen is niet alleen vermakelijk, er wordt ook onderzoek naar gedaan, zelfs op het CWI [1]. Dit probleem is NP-volledig en heeft relaties met verschillende plannings- en roosterproblemen. Met zijn ervaring in lokale zoek-methoden voor NP-volledige problemen, en alle tijd die nu ineens vrijkomt, is Jan Karel Lenstra dé aangewezen persoon om als eerste het Nonogram op de volgende bladzijde tot een goed einde te brengen. Al is het alleen maar om te kunnen antwoorden:

“Ja, leuk hoor Arjen, zo’n RSA-1024 challenge. Maar de JKL-2080 challenge is toch echt een véél grotere uitdaging.”

Maar wie weet is iemand Jan Karel voor: op diverse plaatsen in het CWI-gebouw is de JKL-2080 challenge vanaf heden te vinden: vind de unieke zwart-wit kleuring van de 2080 vakjes die aan de gegeven beschrijving voldoet.

Beste Jan Karel: ik wens je veel succes en plezier de komende tijd en wil je bedanken voor je alles wat je voor het CWI hebt gedaan; in het bijzonder waardeer ik ook je inspanningen rond mijn terugkeer bij het CWI.

Joost Batenburg



[1] K.J. Batenburg en W.A. Kosters, “Solving Nonograms by Combining Relaxations”, Pattern Recognition, 42(8), 1672-1683, 2009.

Recht door zee

Op de laatste CWI-werkdag voor de kerst in, ik schat, 1986 – ik was promovendus in de Afdeling Numerieke Wiskunde – had de Financiële Dienst van het CWI plotseling besloten om de budgetten voor SARA-computers aan het einde van die dag op nul te zetten en ze weer op te waarderen op de eerste werkdag in het nieuwe jaar. Vanaf de kerst tot die eerste werkdag zou er dus niet bij SARA kunnen worden gerekend, geen fijne kerstgedachte. Ik toog naar de Financiële Dienst om te vragen of het niet anders kon. Daar aangekomen, snelde chef Lenstra van de Afdeling Mathematische Besliskunde en Systeemtheorie binnen, voor hetzelfde, voor zijn afdeling. Het werd onmiddellijk geregeld! Als promovendus uit een andere afdeling profiteerde ik direct mee.

De eerste niet-toevallige ontmoeting was eind jaren '90. Onze hoofdpersoon was toen al jaren hoogleraar aan de TU Eindhoven, maar kwam nog regelmatig op het CWI. Op een dag stond hij plotseling in mijn kamer en vroeg of ik wilde meedoen met zijn plan om het Nieuw Archief voor Wiskunde (NAW) nieuw leven in te blazen. Hij wilde er een prachtblad van maken, "voor 's avonds op de bank", en zag in mij een redacteur om te helpen dit ambitieuze plan te verwezenlijken. Mijn vermoeden is dat iemand hem heeft verteld dat ik, net als hij, de afwijking heb een cursieve punt van een niet-cursieve te kunnen onderscheiden. Gegeven het kerstcadeautje van weleer had ik het liever anders gedaan, maar ik moest Jan Karel toch teleurstellen; ik had het erg druk met onderzoek en onderwijs.
Het NAW werd overigens echt een prachtblad; eveneens een mooi staaltje daadkracht.

Na zijn terugkeer op het CWI als directeur kreeg ik regelmatig met Jan Karel te maken, vanuit mijn rol als leider van de groep Scientific Computing and Control (MAC2) en later het cluster Modelling, Analysis and Computing (MAC). Daadkracht bleek een constante te zijn. MAC2 heeft hiervan de vruchten geplukt in de vorm van de aanstelling van twee nieuwe senior-onderzoekers; geen gunstgeschenk, maar een verplichting tot verder goed presteren. De andere "makkers van MAC" profiteerden ook: de groep Life Sciences werd opgericht en meerdere nieuwe senior-onderzoekers werden aangesteld.

De meest veelzeggende goede daad van Jan Karel voor het gehele CWI vind ik de oplossing van het huisvestingsprobleem. Het realiseren van de nieuwbouw en tegelijkertijd het gezond houden van de financiële positie van het CWI – ontwerp- en bouwproject hebben vast veel financiële verlokkingen en valkuilen gekend – is een prestatie van formaat!

Vergaderingen kosten vaak veel tijd en zijn soms saai, maar die met Jan Karel als voorzitter zijn daarop een uitzondering. Ze zijn altijd efficiënt en ook nog onderhoudend, het laatste door de ironische, relativerende opmerkingen van de voorzitter. De één "is een lieverd" en de ander "schrijft tante betje Nederlands". Deze wiskundige "moet nog met mes en vork leren eten" en die universiteit "is rupsbanden". Hier "werken veel klapperende kunstgebitten" en daar "zit men op zijn luie . . .".
Zelf, "dames en heren, jongens en meisjes", zelf vindt hij zich "gewoon een ondersteuner".

Jan Karel heeft een enorme kennis van zaken en personen in zowel wiskunde als informatica, en heeft een fijne neus en grote waardering voor wetenschappelijke kwaliteit. Hij is een pietje precies, maar heeft lak aan bureaucratische poespas. En hij is zeer toegankelijk voor kritiek en recht door zee, ook letterlijk getuige bijgaande foto's.
Jan Karels terugtreden als directeur is een aderlating voor het CWI, maar het moment, na afronding van de evaluatie van het CWI, is heel natuurlijk.

Beste Jan Karel, je hebt veel voor het CWI gedaan en hebt vast nog veel ander nuttig werk te doen. Hartelijk dank voor je grote inzet!

Barry Koren

IJmuiden, 12 september 2006, personeelsuitstapje ter gelegenheid van 60-jarig bestaan CWI. Afvaart en half doorweekt weer op het droge.

Wi mooi ta ri lee Djaar

JKL

Kinderen en Wiskunde houden je jong!
Een fijne tijd toegewenst

Kees Oosterlee

Over het nut van opa's

Opa's kunnen soms heel nuttig zijn. Je kunt dingen van ze leren die je anders nooit zou opsteken. Zo leerde ik van mijn opa pudding maken. Geen pudding uit een pakje, maar echte. Mijn opa had het zelf, als leerlingbanketbakker, in 1898-99 weer geleerd bij *den heer F. Kottelbrink, patissier, Oude Spiegelstraat 3, te Amsterdam.* Naast opa's praktijklessen waren ook zijn kookboeken, meegenomen uit zijn Parijse tijd, een bron van nuttige kennis. Hierbij een stukje kookboek en twee recepten. Opa's verjaardagspudding is bij ons nog steeds populair en het recept is meegegroeid met de verandering van de gelatine en de andere ingrediënten.

Piet Hemker

Opa's verjaardagspudding

Ingrediënten

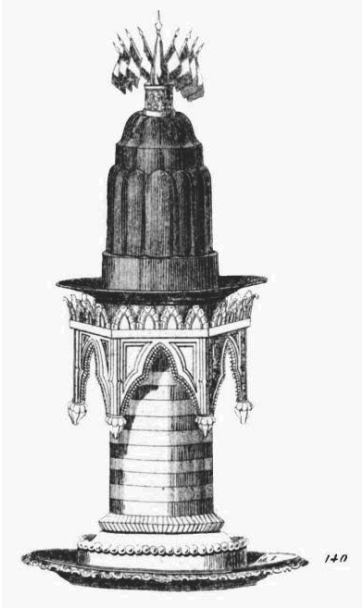
1 hele liter slagroom / altijd koud behandelen
24 g gelatine (12 blaadjes)
1 klein blikje ananas (4 à 5 schijven)
100 g zachte wener in stukjes
100 g Franse vruchten (bigarreaux)
50 g geconfijte kersen (o.i.d.)
150 g poedersuiker
2 pakjes vanillesuiker
6 bitterkoekjes
marasquin naar smaak (1/2 à 1 dl)
(Kapsel of zachte wener is een basismateriaal voor taartjes; je kunt er ook een goede kwaliteit taartbodem voor gebruiken.)

Bereiding:

- De gelatine weken in koud water.
- De slagroom stijf, maar niet te stijf kloppen met de suiker en de vanillesuiker.
- Het blikje ananas openen. Vocht afgieten en bewaren.
- De schijven ananas in kleine stukjes snijden, vermengen met de verkruimelde bitterkoekjes, het verkruimelde zachte wener, de kleingesneden bigarreaux en de gekonfijte kersen. Er zoveel marasquin op gieten als nodig is om de kruimels te doorweken.
- Het ananassap verwarmen en er de gelatine in oplossen.
- De gelatineoplossing laten afkoelen en door de slagroom scheppen. De room moet niet warm worden, maar als het mengsel te koud is stolt de gelatine al. Dat geeft vellen in de pudding. Bij twijfel kun je eerst een paar lepels slagroom door de gelatineoplossing mengen.
- Als het slagroom-gelatinemengsel begint te zetten, zodat een kers niet meer naar de bodem zakt, het gebak en de vruchten er doorheen scheppen.
- Het mengsel in een metalen vorm of in glazen gieten en enkele uren in de koelkast zetten.
- De pudding kan uit de vorm gelost worden door de vorm even in warm water te houden.

Koken als wetenschap

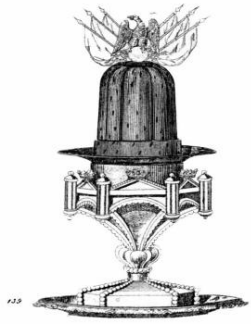
Uit het voorwoord van *La Cuisine Classique*

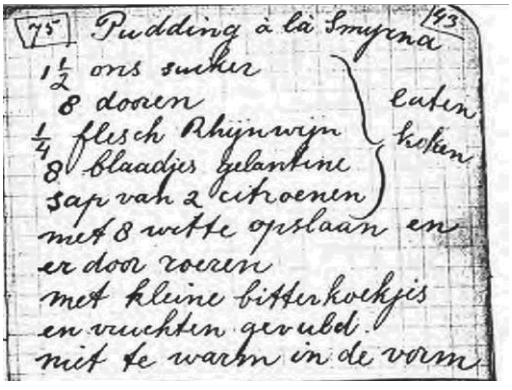


En rendant hommagè à Carème, nous voulons prouver notre profonde déférence pour le maître; mais nous n'entendons nullement affirmer qu'il ait donné le dernier mot d'une profession qui, à quelques égards, peut avoir l'importance d'une véritable science. Et d'ailleurs, personne ne voudrait soutenir que, depuis ces publications lointaines, c'est-à-dire depuis trente à quarante ans, la cuisine n'ait pas suivi le progrès, auquel elle est si intimement liée; qu'elle n'ait reçu aucune impulsion favorable, réalisé aucun perfectionnement, et que les hommes qui la pratiquent aujourd'hui avec tant d'éclat et de succès ne lui aient imprimé aucun élan de grandeur, en lui apportant des ressources et des créations utiles. Nier ce fait serait en quelque sorte méconnaitre le siècle où nous sommes et laisser croire à notre génération que cette science, jadis si active et si honorablement cultivée, a maintenant ralenti sa marche, qu'elle est engourdie ou impuissante, que les hommes ne sont plus ce qu'ils étaient, et qu'enfin nous touchons à la décadence.

Quant à nous, disciples obscurs de nos devanciers, mais imbus de leurs préceptes et comme eux dévoués à notre profession, nous avons pensé le contraire; nous avons compris qu'en raison des exigences de la situation, il restait encore quelque chose à faire.

Urbain Dubois en Émile Bernard,
Paris, mars 1856.





TRANSPARANTJES EN KRIJTBORDEN

door Ko Anthonisse

Lang geleden was ik met mijn toenmalige partner bij een voordracht door JKL. De aanwezigheid van partners geeft al aan dat het een presentatie voor breed publiek was en ongetwijfeld ter gelegenheid van een feestelijk evenement. Onlangs kwam, niet helemaal toevallig, die lezing ter sprake en ze vertelde dat ze (docente verpleegkunde) er veel van had opgestoken en dat ook haar collega docenten er profijt van hadden gehad.

In die tijd waren er nog geen laptops, laat staan hulpmiddelen als PowerPoint. Als je bij een inleiding tekst of plaatjes wilde vertonen waren er drie opties: ter plekke op het bord schrijven, gebruik maken van een overheadprojector of één en ander op papier uitdelen.

Schrijven op het bord is vertragend, een nadeel is ook dat je als spreker met je rug naar het publiek staat en bovendien vereist het een handschrift van enige kwaliteit. Het tevoren uitdelen van tekst en figuren op papier brengt het risico mee dat de luisteraars lezers worden met een half oor voor de spreker en een half oog voor de punten die pas twintig minuten later ter sprake zullen komen.

De overheadprojector (die overigens nog niet door de beamer is verdrongen) gaf de mogelijkheid om de tekst gedoseerd te vertonen. Een beschreven A4 transparant folie was een hele dosis, door het gedeeltelijk af te dekken kon je de inhoud ook regel voor regel onthullen en de nieuwsgierigheid van het publiek bespelen om zo de aandacht er bij te houden.

De transparantjes werden aanvankelijk met de hand beschreven (met kleurgebruik) de leesbaarheid was dan soms nog wel een probleem. Dat was voorbij toen tekst vanaf papier met de kopieeermachine op het folie kon worden overgebracht. Een slechte kwaliteit folie kon wel

tot het verlies van een machine leiden als de folie op de trommel smolt, maar al doende leert men.

Tijdens de voordracht door JKL bleek dat hij het werken met transparantjes tot een kunst had verheven. Ze werden niet primitief deels afgedekt, maar over elkaar gelegd om een beeld op te bouwen. En hij gebruikte kleine rechthoekjes folie (flappen) die met tape aan het eerste vel scharnierden en op het juiste moment werden omgeslagen. Zonodig zaten er flappen aan alle zijden van het eerste vel. Er waren zelfs flappen met sub-flappen, als herhaling van een creatief idee, het werd er alleen maar vrolijker op.

Die zorgvuldige voorbereiding en uitvoering van een presentatie heeft toen diepe indruk en zelfs school gemaakt.

Het bovenstaande werd enkele jaren geleden geschreven. Nu (juni 2011) wilde ik wel eens weten of de overheadprojector inmiddels al was verdrongen door digitale hulpmiddelen. Ik belde enkele leveranciers van overheadprojectoren. De eerste was wat somber, hij had nog voorraad maar die zou niet meer worden aangevuld. De tweede was enthousiast: “Meneer, ik verkoop er 5 tot 10 per week, meest tafelmodellen die mensen ook meenemen om hun presentaties te houden”. De overheadprojector is dus kennelijk nog niet dood. Wat verder pratend vertelde dezelfde man blij over 22 grote krijtborden die hij onlangs aan een universiteit had mogen leveren. Volgens hem zijn het vooral de beta-wetenschappers die er van houden om hun formules en schema’s ter plekke op te schrijven en toe te lichten. Dat degelijke handwerk wordt dus ook nog steeds beoefend.

Een wet van Moore voor software?

Jan Heering

Jan Karel heeft zijn sporen verdiend in de (discrete) algoritmiek. De constatering in het PITAC rapport [1] en elders dat betere algoritmen en betere hardware in gelijke mate hebben bijgedragen aan de fenomenale ontwikkeling van computermodellering was dan ook koren op zijn molen. Twee elkaar versterkende exponentiële trends. Een indrukwekkende onderstreping van het belang van algoritmiek en voor Jan Karel een ideale kapstok om zijn veelvuldige pitches over het belang van computational science aan op te hangen.
Sinds de eerste chips rond 1960 ten tonele verschenen heeft de evolutie van hardware de wet van Moore gevolgd. Gezien het bovenstaande geldt er blijkbaar ook een soort wet van Moore voor algoritmiek. Deze heeft geen naam en verdient nog minder een wet genoemd te worden dan die van Moore. De mogelijkheid (maar niet de praktische realisatie, laat staan het tempo van realisatie) van de wet van Moore kan herleid worden tot de schalingswetten van Dennard et al. voor MOSFETs, de op chips meest gebruikte transistors. Ik kan me vergissen maar het lijkt me kras dat een vergelijkbaar eenvoudig principe aan de basis ligt van een wet van Moore voor algoritmiek. Daarvoor is het veld te groot en te divers. Het zijn veeleer algoritmische uitvindingen en verbeteringen over een breed front en in snel tempo die samen een op de wet van Moore lijkend exponentieel verloop opleveren. De exponentiële ontwikkeling van hardware heeft een navenante ontwikkeling van software mogelijk gemaakt. Zie het groeitempo van Windows en Linux [2]. Opnieuw een soort wet van Moore, maar nu voor het aantal regels code van softwaresystemen. Het eind is niet in zicht. Onder het motto *scale changes everything* waagt het ULS rapport [3] zich aan een poging het onafwendbaar naderbij komen van ultragrootschalige systemen van miljarden regels code onder ogen te zien. Moed en optimisme kunnen de auteurs niet ontzegd worden. Geestverruimende lectuur.

Het ziet er naar uit dat de wet van Moore (de echte) zijn langste tijd gehad heeft. Maar lithografie is een kat met zeven levens. Hoeveel winst er, voorbij de grenzen van de lithografie, op moleculaire, atomaire of zelfs nucleaire (waarom niet?) schaal uiteindelijk nog te behalen zal zijn, valt niet te zeggen. Het enthousiasme voor moleculaire elektronica van mijzelf en anderen in de tachtiger jaren was, achteraf gezien, prematuur [4]. Dertig jaar later is moleculaire elektronica nog steeds niet meer dan een kandidaat-opvolger waaraan mogelijk geen behoefte zal blijken te bestaan.
Hoe dan ook, welke invloed zal een vertraging in de evolutie van hardware hebben op de algoritmiek? Het antwoord ligt voor de hand. Algoritmiek zal verder aan belang winnen naarmate de vooruitgang op hardwaregebied langzamer gaat. De versterkte inzet op parallelle algoritmen om de huidige multicoreprocessoren effectief te kunnen benutten is pas het begin. Uiteindelijk komt de algoritmiek er alleen voor te staan. Dat zal wennen zijn voor een hardware-enthousiast als ik. Maar Jan Karel zal er geen enkele moeite mee hebben. En misschien loopt het allemaal heel anders.

[1] PITAC, *Computational Science: Ensuring America's Competitiveness*, 2005. http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf
[2] James R. Larus, Spending Moore's dividend, Technical Report MSR-TR-2008-69, Microsoft Research, 2008. http://research.microsoft.com/pubs/70581/tr-2008-69.pdf
[3] Linda M. Northrop et al., *Ultra-Large-Scale Systems: The Software Challenge of the Future*, Carnegie Mellon Software Engineering Institute, 2006. http://www.sei.cmu.edu/uls/
[4] Hyungsub Choi and Cyrus C. M. Mody, The long history of molecular electronics: Microelectronics origins of nanotechnology, *Social Studies of Science* **39** 1 (February 2009). http://www.owlnet.rice.edu/~cyrus.mody/MyPubs/Molecular%20Electronics.pdf

Typography in Process Algebra

Jos Baeten, Technische Universiteit Eindhoven, and
Jan Willem Klop, Vrije Universiteit, Amsterdam

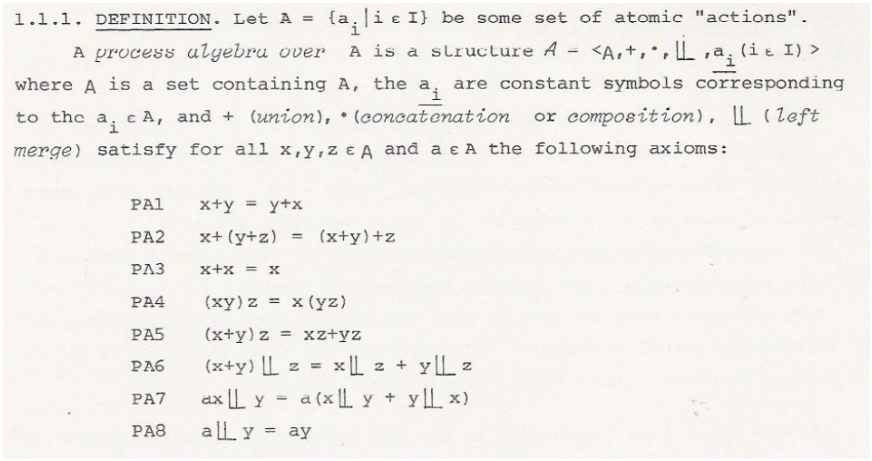
July 1, 2011

Abstract

This note is dedicated in gratitude to Jan Karel Lenstra, who reputedly is able to distinguish an ordinary full stop symbol (".") from its italicized version (".."), and who has inspired us with his knowledge and passion for typography. For his amusement we offer him a sample of typographical matters that we encountered in Process Algebra, of which the Amsterdam version originated at MC and CWI in the early eighties. The present note is written in the fonts Palatino and Euler, the latter being an upright cursive typeface, commissioned by the American Mathematical Society and designed and created by Hermann Zapf with the assistance of Donald Knuth.

1 History

The late seventies of the last century saw the advent of formal systems for communicating processes, an area then also designated as 'concurrency'. At CWI, after Jaco de Bakker introduced this new area in the Netherlands, Jan Bergstra and Jan Willem Klop endeavoured in the early eighties to develop a family of algebraic systems for various features of processes, specifically centering around ACP, Algebra of Communicating Processes. The phrase 'Process Algebra' was also coined then. Jan Willem wrote the first reports on process algebra on his trusted Olivetti with daisywheels, providing many symbols and fontsizes that could be produced on a typewriter. But not all symbols were available: the first report [17] introduced the 'left-merge' which in LaTeX can be presented as $\between$, and whose definition and employment marked our entrance in this field. Because of this auxiliary operator, process algebra is for a large part 'finitely axiomatizable', which is important for automation of the process algebra enterprise in systems such as mCRL. In [17], this symbol was written by hand, using India ink, see Figure 1.



1.1.1. DEFINITION. Let A = $\{a_i \mid i \in I\}$ be some set of atomic "actions".
A *process algebra over* A is a structure $\mathcal{A} = \langle A, +, \cdot, \between, \underline{a_i} (i \in I) \rangle$ where A is a set containing A, the $\underline{a_i}$ are constant symbols corresponding to the $a_i \in A$, and + (*union*), $\cdot$ (*concatenation* or *composition*), $\between$ (*left merge*) satisfy for all $x,y,z \in A$ and $a \in A$ the following axioms:

PA1	$x+y = y+x$
PA2	$x+(y+z) = (x+y)+z$
PA3	$x+x = x$
PA4	$(xy)z = x(yz)$
PA5	$(x+y)z = xz+yz$
PA6	$(x+y) \between z = x \between z + y \between z$
PA7	$ax \between y = a(x \between y + y \between x)$
PA8	$a \between y = ay$

Figure 1: *Olivetti and handwriting.*

For official purposes, all text written by CWI workers, either by typewriter or by hand, was converted by special typists into troff. They had to define a special symbol for the left-merge, which they called "Klop". See Figure 2, taken from [7].

Then, in 1985, Jan Bergstra and Jos Baeten moved to the University of Amsterdam, where there were no troff typists. Instead, they had Apple Macintosh computers and the typographical possibilities they offered. Some of us have fond memories of the MacPlus (Figure 3), the ImageWriter, the StyleWriter, MacWrite 5.0 and all that.

But, what to do with special process algebra symbols such as the left-merge? The group at University of Amsterdam set out to develop a special font for Process Algebra, using the Mac applications Fontastic and Fontographer. Two versions were developed: font Amsterdam, see Figure 4, based on Apple's font Times, and a sans-serif version called Brussel, based on Geneva. An example of a text in font Brussel is given in Figure 5.

The problem returned when laserwriters were introduced, then the special fonts could not be upgraded, but were still used in the books [1] and [13]. Later, the home-made process algebra typography was superseded by the superior LaTeX technology, and the rest of this note refers to that framework. One of us persisted until very recent years in using Mac typography with modern Mac applications such as Pages, but is now also converted to LaTeX.

$x+y = y+x$	A1	$a \lhd b = a$ if not $(a < b)$	P1
$x+(y+z) = (x+y)+z$	A2	$a \lhd b = \delta$ if $a < b$	P2
$x+x = x$	A3	$x \lhd yz = x \lhd y$	P3
$(x+y)z = xz+yz$	A4	$x \lhd (y+z) = (x \lhd y) \lhd z$	P4
$(xy)z = x(yz)$	A5	$xy \lhd z = (x \lhd z)y$	P5
$x+\delta = x$	A6	$(x+y) \lhd z = x \lhd z + y \lhd z$	P6
$\delta x = \delta$	A7		
$a \mid b = b \mid a$	C1		
$(a \mid b) \mid c = a \mid (b \mid c)$.	C2		
$\delta \mid a = \delta$	C3		
$x \parallel y = x \lfloor\!\lfloor y + y \lfloor\!\lfloor x + x \mid y$	CM1	$\theta(a) = a$	TH1
$a \lfloor\!\lfloor x = ax$	CM2	$\theta(xy) = \theta(x)\theta(y)$	TH2
$ax \lfloor\!\lfloor y = a(x \parallel y)$	CM3	$\theta(x+y) = \theta(x) \lhd y + \theta(y) \lhd x$	TH3
$(x+y) \lfloor\!\lfloor z = x \lfloor\!\lfloor z + y \lfloor\!\lfloor z$	CM4		
$(ax) \mid b = (a \mid b)x$	CM5		
$a \mid (bx) = (a \mid b)x$	CM6		
$(ax) \mid (by) = (a \mid b)(x \parallel y)$	CM7		
$(x+y) \mid z = x \mid z + y \mid z$	CM8		
$x \mid (y+z) = x \mid y + x \mid z$	CM9		
$\partial_H(a) = a$ if $a \notin H$	D1		
$\partial_H(a) = \delta$ if $a \in H$	D2		
$\partial_H(x+y) = \partial_H(x) + \partial_H(y)$	D3		
$\partial_H(xy) = \partial_H(x)\partial_H(y)$	D4		

Figure 2: *ACP, Algebra of Communicating Processes, with priorities.*

Figure 3: *1988 MacintoshPlus ED, still alive and well.*

Font name: Amsterdam
Font size: 12
Font number: 150
Location: System :System

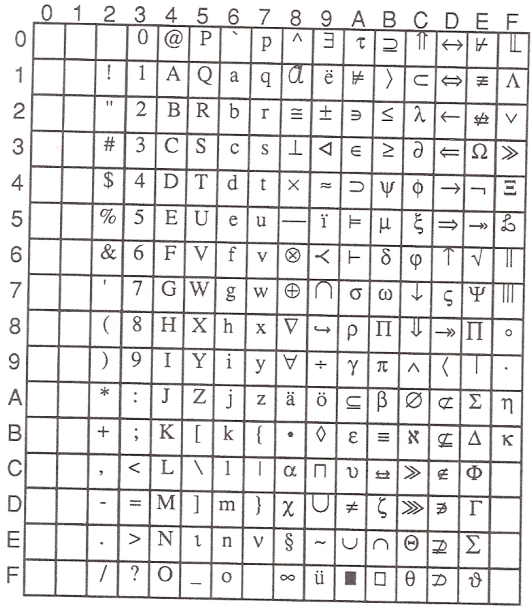
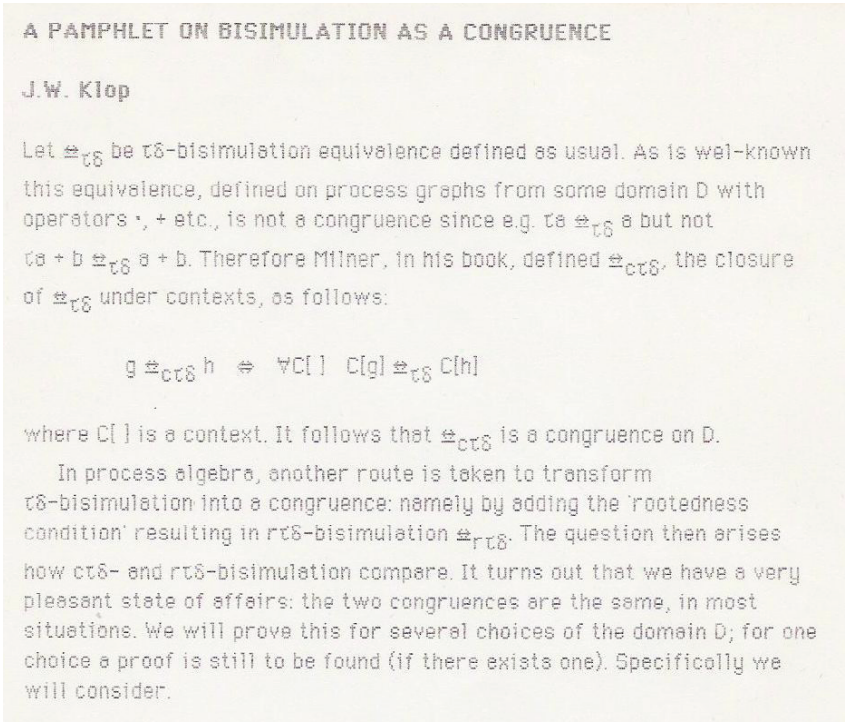


Figure 4: *process algebra font Amsterdam*.



A PAMPHLET ON BISIMULATION AS A CONGRUENCE

J.W. Klop

Let $\leftrightarrows_{\tau\delta}$ be $\tau\delta$-bisimulation equivalence defined as usual. As is well-known this equivalence, defined on process graphs from some domain D with operators $\cdot$, + etc., is not a congruence since e.g. $\tau a \leftrightarrows_{\tau\delta} a$ but not $\tau a + b \leftrightarrows_{\tau\delta} a + b$. Therefore Milner, in his book, defined $\leftrightarrows_{c\tau\delta}$, the closure of $\leftrightarrows_{\tau\delta}$ under contexts, as follows:

$$g \leftrightarrows_{c\tau\delta} h \quad \Leftrightarrow \quad \forall C[\,]\ \ C[g] \leftrightarrows_{\tau\delta} C[h]$$

where C[ ] is a context. It follows that $\leftrightarrows_{c\tau\delta}$ is a congruence on D.

In process algebra, another route is taken to transform $\tau\delta$-bisimulation into a congruence: namely by adding the 'rootedness condition' resulting in $r\tau\delta$-bisimulation $\leftrightarrows_{r\tau\delta}$. The question then arises how $c\tau\delta$- and $r\tau\delta$-bisimulation compare. It turns out that we have a very pleasant state of affairs: the two congruences are the same, in most situations. We will prove this for several choices of the domain D; for one choice a proof is still to be found (if there exists one). Specifically we will consider.

Figure 5: *A 'Pamphlet' from 1986 discussed in the PAM seminar (Process Algebra Meeting), which was held from 1985 until 2010.*

Figure 6: *Interrobang in Palatino Linotype, taken from Wikipedia.*

2 Punctuation symbols

Punctuation symbols are used in process algebra extensively. Well-known is the use of the period to denote action prefix, so e.g. $a.x$ means that action a is prefixed to term x, so first a is executed, followed by x, see CCS [23]. For sequential composition, the semicolon is sometimes used, so $x;y$ means x is followed by y, see CSP [21] or χ [15].

The exclamation mark is used for replication in the π-calculus [25], so $!x = x \between !x$, often called *bang*.

Besides this, it stands for a *send*, in CSP [21], but also in [2] (or is it [3]?), so action $c!d$ means that data element d is sent at communication port c.

A datum sent can also be received, and the question mark is used for this, so $c?d$ means that data element d is received at communication port c. In ACP-style process algebra, the simultaneous execution of matching communicating actions in parallel components is the resulting communication, so the simultaneous communication of $c!d$ and $c?d$ becomes the communication $c⁉d$, see [2] or [3].

The symbol ⁉, see Figure 6, is the *interrobang* (interrogation mark plus bang) invented by Martin K. Speckter in 1962 (see [29])).

It is used to ask a question in an excited manner, expresses excitement or disbelief in the form of a question, or asks a rhetorical question. It was popular in the 1960s, but its use declined soon after. Nowadays, we would write !? or ?! instead.

The asterisk stands for iteration or Kleene star, studied in process algebra in [18] or [10]. In x^*, either immediate termination takes place, or x is executed once, again followed by x^*. The vertical bar is used in CCS for parallel composition [23], while most other process algebras use the double vertical bar $\between$ for this. In ACP-style process algebra, the single bar is used for communication merge: the parallel composition of terms that starts with a communication action. The backslash $\backslash$ is used in CCS for restriction (also called encapsulation) [23], sometimes the 'forward' slash / is used for abstraction.

3 Greek symbols

Many Greek letters are used in process algebra. The letter α sometimes stands for the *alphabet* of a process, see [9]. Often, α is also used as ranging over a given set of actions plus some other actions. In [9], the letter β is used for an approximation of an alphabet, that is easier to calculate. The letter γ stands for the communication function, indicating which actions match in order to form a communication action together, probably because the letter occupies the same place in the Greek alphabet as the letter c in the Latin alphabet.

The letter δ is a constant standing for inaction, unsuccessful termination or deadlock in ACP-style, taking the name from the first letter of the word deadlock [19]. Since it is the neutral element of alternative composition, it was later renamed to 0 in [2]. The letter δ is also used for delay, see e.g. [24]. The capital letter Δ stands for divergence, see [20]. The letter ε is the constant standing for the empty process, successful termination or skip [30]. Since it is the neutral element of sequential composition, it was renamed to 1 in [2].

A use of the letter ζ in process algebra could not be found. The letter η was introduced for the silent step in [12], probably for no other reason than that this was an unused letter up to that time. The priority operator introduced in [8] uses the letter θ, also for no other reason than that. When in [2] a new auxiliary operator was used in the axiomatization of the priority operator, ϑ was used. The letter ι stands for an *idle* transition, e.g. in [22]. The letter κ is often used in ordinal arithmetic, but no use in process algebra was found.

The letter λ is used for the *state operator* and stems from [4]. Also a generalized version comes from here, and uses the letter Λ. The letter μ is used for a fixed point of recursion in CCS, already in [23]. The letter ν stands for now and is used for a process starting with no time delay in [5], but also stands for new and is used for a fresh variable in π-calculus [25]. In [4], the ν was used for a *localization* operator.

We have seen no uses of ξ or Ξ so far, however see the item χ further on. The omikron doesn't distinguish itself from the Latin character, so does not appear as itself. Lower case π of course names the π-calculus [25], and is used for projection in ACP-style process algebra [19]. Upper case Π is sometimes used for generalized parallel composition. The letter ρ is used for renamings in [4].

The letter σ is used for a time step, see e.g. [6] or for a time delay [11]. The capital Σ is widely used for generalized alternative composition or

sum, since [23]. The letter τ is universally used for a silent or unobservable step, dating back to [23]. As there, τ is put directly for the result of two matching communicating actions, it is the trace of a communication, and this is the origin of the name. In timed process algebra, the υ is used for time-out and time initialization, see [11].

The ϕ-calculus is a hybrid process algebra developed in [27]. The letter χ is also the name of a hybrid process algebra [15], taking its name from the initial of the word hybrid. At the moment, the language χ is being updated and renewed, with as working title Xí [14]. Reading this as Greek is the name of the capital χ, but reading it as Latin, Xí, it becomes Ξ. The X refers to the neXt operator of temporal logic. The ψ-calculus is an extension of the π-calculus with data and assertions about data, see [16]. Finally, ω is a special action denoting success in a testing semantics [26], but also is a calculus for mobile ad-hoc networks [28].

4 Other symbols

We conclude with some symbols not yet mentioned above.

left-merge We already discussed the left-merge $\between$, where $x \between y$ is the interleaving of processes x and y with the proviso that the first step has to be done by the left process x, hence the name left-merge. Typographically, the modern version of this symbol differs a bit from the ancient one, where the horizontal stroke was longer, see Figure 2. Many versions of the left-merge exist in LaTeX, such as $\between$, used in [2, 3].

encapsulation An important operator when composing processes is the *encapsulation operator* written as ∂. The reason for choosing this symbol was its association in some areas of mathematics as a boundary operator. It is parametrized by a set of actions H, so that $\partial_H(x)$ means that process x is not allowed to communicate with the environment using actions in the set H.

bisimulation Finally, there is the symbol for *bisimilarity* $\leftrightarrows$. It occurs not in the syntax of process algebra, but in its semantics, connecting the formal syntax, strings of symbols, with the real processes, mathematical entities. Processes x and y are bisimilar, $x \leftrightarrows y$, when roughly said, they can simulate each other.

References

[1] J.C.M. Baeten. *Procesalgebra*. Programmatuurkunde. Kluwer, 1986.

[2] J.C.M. Baeten, T. Basten, and M.A. Reniers. *Process Algebra (Equational Theories of Communicating Processes)*. Number 50 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2009. USA Edition.

[3] J.C.M. Baeten, T. Basten, and M.A. Reniers. *Process Algebra (Equational Theories of Communicating Processes)*. Number 50 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2010. UK Edition.

[4] J.C.M. Baeten and J.A. Bergstra. Global renaming operators in concrete process algebra. *Information and Computation*, 60(1/3):205–245, 1988.

[5] J.C.M. Baeten and J.A. Bergstra. Real time process algebra. *Formal Aspects of Computing*, 3(2):142–188, 1991.

[6] J.C.M. Baeten and J.A. Bergstra. Discrete Time Process Algebra. *Formal Aspects of Computing*, 8(2):188–208, 1996.

[7] J.C.M. Baeten, J.A. Bergstra, and J.W. Klop. Proces algebra met interrupt mechanisme. In *Proceedings NGI-SION Informatica Symposium*, pages 129–135, 1985.

[8] J.C.M. Baeten, J.A. Bergstra, and J.W. Klop. Syntax and Defining Equations for an Interrupt Mechanism in Process Algebra. *Fundamenta Informaticae*, IX(2):127–168, 1986.

[9] J.C.M. Baeten, J.A. Bergstra, and J.W. Klop. Conditional Axioms and α/β-Calculus in Process Algebra. In M. Wirsing, editor, *Formal Description of Programming Concepts - III, IFIP Conference, Proceedings*, pages 77–103. Elsevier, 1987.

[10] J.C.M. Baeten, F. Corradini, and C.A. Grabmayer. A characterization of regular expressions under bisimulation. *Journal of the ACM*, 54(2):6.1–28, 2007.

[11] J.C.M. Baeten and C.A. Middelburg. *Process Algebra with Timing*. EATCS Monographs in Theoretical Computer Science. Springer-Verlag, 2002.

[12] J.C.M. Baeten and R.J. van Glabbeek. Another look at abstraction in process algebra. In Th. Ottman, editor, *Automata, Languages and Programming, 14th Colloquium, ICALP 1987, Proceedings*, number 267 in Lecture Notes in Computer Science, pages 84–94. Springer-Verlag, 1987.

[13] J.C.M. Baeten and W.P. Weijland. *Process Algebra*. Number 18 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1990.

[14] Jos C.M. Baeten and Michel A. Reniers. Xí — a hybrid process algebra for cyber-physical systems. Draft., 2011.

[15] D.A. van Beek, K.L. Man, M.A. Reniers, J.E. Rooda, and R.R.H. Schiffelers. Syntax and consistent equation semantics of hybrid chi. *Journal of Logic and Algebraic Programming*, 68(1-2):129–210, 2006.

[16] J. Bengtson, M. Johansson, J. Parrow, and B. Victor. Psi-calculi: Mobile processes, nominal data, and logic. In *Proceedings LICS 2009*, pages 39–48, 2009.

[17] J. A. Bergstra and J. W. Klop. Fixed Point Semantics in Process Algebra. Technical Report IW 206, Mathematical Centre, Amsterdam, the Netherlands, 1982.

[18] J.A. Bergstra, W.J. Fokkink, and A. Ponse. Process Algebra with Recursive Operations. In J.A. Bergstra, A. Ponse, and S.A. Smolka, editors, *Handbook of Process Algebra*, pages 333–389. Elsevier, 2001.

[19] J.A. Bergstra and J.W. Klop. Process Algebra for Synchronous Communication. *Information and Control*, 78(3):109–137, 1984.

[20] J.A. Bergstra, J.W. Klop, and E.-R. Olderog. Failures without Chaos: a new Process Semantics for Fair Abstraction. In M. Wirsing, editor, *Formal Description of Programming Concepts - III, IFIP Conference, Proceedings*, pages 77–103. North-Holland, 1987.

[21] C.A.R. Hoare. *Communicating Sequential Processes*. Prentice Hall, 1985.

[22] A.S. Klusener. *Models and Axioms for a Fragment of Real Time Process Algebra*. PhD thesis, Eindhoven University of Technology, Department of Computer Science, 1993.

[23] R. Milner. *A Calculus of Communicating Systems.* Number 92 in Lecture Notes in Computer Science. Springer Verlag, 1980.

[24] R. Milner. Calculi for synchrony and asynchrony. *Theoretical Computer Science*, 25(3):267–310, 1983.

[25] R. Milner. *Communicating and Mobile Systems: The π-calculus*. Cambridge University Press, 1999.

[26] R. De Nicola and M. Hennessy. Testing equivalences for processes. *Theoretical Computer Science*, 34:83–133, 1984.

[27] William C. Rounds and Hosung Song. The phi-calculus: A language for distributed control of reconfigurable embedded systems. In *HSCC'03*, pages 435–449, 2003.

[28] Anu Singh, C. R. Ramakrishnan, and Scott A. Smolka. A process calculus for mobile ad hoc networks. In *Proceedings of the 10th international conference on Coordination models and languages*, COORDINATION'08, pages 296–314. Springer-Verlag, 2008.

[29] Martin K. Speckter. *Disquisition on the Composing Stick*. Typophiles, Inc., 1971.

[30] J.L.M. Vrancken. The Algebra of Communicating Processes with Empty Process. *Theoretical Computer Science*, 177(2):287–328, 1997.

TU/e-collega's

Hallo Jan Karel,

Even was ik verbaasd te horen dat je in november a.s. afscheid neemt van het CWI. Nu al? In mijn beleving ben je daar nog lang niet aan toe.

Ik heb je leren kennen in het Dommelgebouw van de Technische Universiteit Eindhoven en bij het inrichten van je kamer viel het me al op hoe goed je je drukke werkzaamheden organiseerde. Daarom vond ik het ook fijn om als secretaresse voor je te werken. Ik weet nog dat ik in 1 jaar tijd 1000 referee-brieven voor je typte in je functie als editor van een aantal belangrijke tijdschriften. (De fles champagne als dank vergeet ik niet snel.)
Wat ik ook niet vergeet: toen ik tijdens m'n zwangerschap van Inge plotseling in het ziekenhuis moest worden opgenomen, was jij de eerste die bij Hans kwam vragen hoe het met mij was.

Je zestigste verjaardag werd nog op de TU/e gevierd met een wetenschappelijke middag en afgesloten met een heerlijk diner waar ik gezellig met Catrien en Hidde heb bijgepraat en foto's van hun zoontje heb gezien.

Daarna werd het contact veel minder, maar elke keer als je weer in Eindhoven was en langs kwam voor de post, heb ik vaak teruggedacht aan de tijd dat ik je secretaresse was. Ik bewaar daar goede herinneringen aan.
Dat Karen aan de TU/e kwam werken en regelmatig even binnen kwam als zij jouw post ophaalde, om 'bij te praten' heb ik heel plezierig gevonden!

Ik stop met herinneringen ophalen. Het is tijd om vooruit te kijken. Ik wens jou, Karen en de kinderen een hele goede tijd toe. Meer tijd voor je gezin, om te lezen, te koken en naar muziek te luisteren, wens ik je van harte toe.

Hartelijke groeten, ook van Hans,

Harma Koops

Idolen en idealen

Beste Jan Karel,

Toen ik, zo kort na je zestigste verjaardag, opnieuw een uitnodiging kreeg tot het schrijven van een bijdrage aan een liber amicorum voor jou, was mijn eerste neiging te bedanken voor de eer, omdat ik slechts in herhaling zou vallen. In de weken na dat verzoek nam ik echter deel aan bijeenkomsten van de BAC voor jouw opvolging als directeur van het CWI, en realiseerde me hoe dikwijls ik jou voor ogen heb gehad als ik me afvroeg hoe iets zou moeten worden aangepakt. Ik begon na te denken over de wetenschappers die ik me als voorbeeld heb gesteld, en over de balans die ze vonden tussen onderzoek, onderwijs en onderzoeksmanagement. Dat leidde tot onderstaand klein essay.

Door de jaren heen heb ik veel grote wetenschappers leren kennen, en sommigen heb ik tot voorbeeld genomen: J.W. Cohen, Ed Coman, Alexander Rinnooy Kan, jou. Wat me telkens weer heeft getroffen is hoe totaal verschillende talenten en karakters een soortgelijke top kunnen bereiken, op totaal verschillende manieren. Cohen had het vermogen, zich als een terriër in een probleem vast te bijten. Gewapend met een groot optimisme (onontbeerlijk voor een goede onderzoeker van zijn type) kon hij na maanden een probleem kraken. Bijna iedereen zou dan al lang hebben opgegeven. Hij had een brandend onderzoeksvuur, dat zelfs niet doofde toen hij oud en erg ziek was. Hij schreef niet makkelijk, en was ook geen groot spreker. Maar zijn enorme focus, doorzettingsvermogen en werkkracht, gekoppeld aan een opmerkelijke creativiteit, maakten hem tot een prominent wetenschapper. Hij was een specialist, en zijn magnum opus *The Single Server Queue* was een huzarenstukje waarvan ik me ook nu nog nauwelijks kan voorstellen dat een mens dat allemaal bij elkaar kon schrijven. Wat zijn bestuurlijk werk betreft: ik heb begrepen dat zijn decanaat in Delft geen groot succes was, en hem heel zwaar viel. In Utrecht deed hij geen omvangrijk bestuurlijk werk. In de wereld van de prestatie-analyse had hij wel invloed, door zijn prestige, en zijn organisatie van enige congressen. Hij startte het wachtrijcolloquium, dat nog steeds op het CWI wordt gehouden, en was de motor achter de opkomst van de computerprestatie-analyse in Nederland.

Ed Coman past misschien in sommige opzichten wat minder in het rijtje, maar Ed is een man die me dikwijls heeft geïnspireerd en geïmponeerd − misschien ook wel omdat hij zo anders dan Cohen was. Hij vroeg me bovendien voor mijn eerste editorship, met de woorden "I have served my due". De eerste kennismaking met Ed verliep slecht. In 1981 mocht ik als hulpje Cohen assisteren bij het organiseren van een IFIP Working Group 7.3 workshop aan de vooravond van het congres *Performance '81* op de VU. Hoofdtaak: het verzamelen van de titels en abstracts van de lezingen uit de workshop. Ik stuurde de sprekers een brief, en enige weken later

kwamen alle antwoorden binnen. Alle, behalve die van Coman. Licht geïrriteerd maar beleefd stuurde ik hem een herinnering. Zijn antwoord: “Look, this is a workshop. I don’t know yet what I’ll be talking about. It will be about ongoing work, and you’ll see it when I’m in Amsterdam”. Instinctief had ik een hekel aan die man die me mijn werk niet liet doen. Toen kwam de workshop. Comans lezing was veruit de beste. Twee zinnen per minuut, glashelder, en iedereen kon begrijpen waarom hij in dat bepaalde probleem geïnteresseerd was, en hoe hij het probeerde aan te pakken. Fascinerende wiskunde, maar ook een les over wat een workshop hoort te zijn, die ik mijn leven niet meer vergeten ben. Ik leerde ook van Ed door te zien hoe hij opereerde in bovengenoemde IFIP WG7.3. Hij zei niet heel veel tijdens vergaderingen van de working group. Maar als hij zijn mond open deed was het zo doordacht en duidelijk, en was de harde boodschap zo door humor verzacht, dat zijn mening dikwijls bepalend was. Nog even terugkomend op Eds stijl van lezingen geven: ik denk dat nogal eens wordt onderschat hoe belangrijk het is een goede lezing te geven, en hoe essentieel een goede voorbereiding daarbij is. Ed kan geweldige lezingen geven (maar als hij zich niet voorbereidt dan gaat het ook bij hem mis), maar er zijn zovelen die mooie voordrachten kunnen geven (en zovelen die het nooit lijken te leren). Bij lezingen heb ik dikwijls ook Weber en Mitrani voor ogen. Mijn vader en Cor Baayen neem ik, bij toespraken, het meest als voorbeeld; en natuurlijk Rinnooy Kan ...
Alexander Rinnooy Kan is een van de meest brilliante mensen die ik heb mogen leren kennen, gezegend met meer talenten dan een gewoon mens kan bevatten. In veel opzichten een tegenpool van Cohen, en samen zijn ze overtuigend bewijs van de stelling van daarnet over het op totaal verschillende manieren de top van je vakgebied bereiken. In zekere zin nam ik het Alexander zo’n twintig jaar geleden kwalijk dat hij de wetenschap vaarwel zei. Ik kon eigenlijk niet begrijpen dat iemand rector wilde worden, en daarna zelfs voorztter VNO-NCW, om maar te zwijgen over de Raad van Bestuur van de ING-bank. Mijn ideaal was − en is − mooie artikelen te schrijven en jonge mensen te helpen op een pad naar wiskundige en/of onderzoeker. Pas geleidelijk ging ik begrijpen dat als je, zoals Alexander, zoveel meer kunt dan dat (onderzoek en onderwijs), zo’n loopbaan te beperkt kan zijn.

Dat waren − naast mijn vader − mijn idolen; en nu het niet meer tegen me kan worden gebruikt wil ik wel toevoegen dat jij daar ook bij hoort. Om niet in gevlei te vervallen wil ik er een karaktereigenschap uitlichten: verantwoordelijkheidsbesef. Zelf van nature meer een probleemoplosser dan een manager, ben ik van tijd tot tijd toch in management- en bestuurstaken verzeild geraakt. Dat gebeurde om uiteenlopende redenen, zoals: (i) het is een mooie uitdaging iets goeds op te zetten (zo haalde je me in 1985 over naar het CWI te komen om de eerste Nederlandse groep in de prestatie-analyse op te zetten), (ii) iedereen moet op zijn tijd voor een klus opdraaien, en (iii) als je zelf niet bereid bent een afdeling of groep te leiden, dan

moet je maar afwachten wie je baas wordt (dat laatste was jouw argument om me in 1988 over te halen, te solliciteren naar jouw baan als afdelingschef op het CWI).

Mij overkwamen die taken min of meer, maar jij zocht ze meer actief op. Jij voelde je echt verantwoordelijk voor een CWI-afdeling, of voor het hele CWI, of voor het KWG, of voor het Nieuw Archief voor Wiskunde, of voor een tijdschrift, of voor een handboek. Jij had en hebt ook altijd een duidelijke mening over wat moet worden gedaan om tot een substantiële verbetering te komen, en dat ging je dan ook doen. Zodoende heb je heel nadrukkelijk invloed uitgeoefend op het CWI, op de Nederlandse wiskunde, op de internationale besliskunde. Ongetwijfeld zul je je wel eens afgevraagd hebben of je de juiste keuzes hebt gemaakt. Tijd stoppen in een mooi boek over combinatorische optimalisering, of in een zware bestuursfunctie?

Worden we geleid door het toeval, zijn we "children of coincidence and chance", of worden onze keuzes voor een groot deel bepaald door onze specieke aanleg en aard? Streven we heel bewust onze idealen na, of kan een enkel advies van een van onze idolen ons leven veranderen? Eigenlijk weet ik voor mijzelf het antwoord niet (al ben ik geneigd het toeval een grote rol toe te kennen, en "chance favors the prepared mind"), laat staan dat ik het voor jou kan weten. Maar zowel idealen als idolen lijken me onmisbaar voor een mooie wetenschappelijke loopbaan. Idealen houden het vuur brandend, en idolen geven je op het goede moment inspiratie of het essentiële zetje in de goede richting. Jan Karel, veel dank voor die inspiratie en die zetjes in de goede richting!

Onno Boxma

Beslissende momenten

Wil Schilders (TU Eindhoven & Platform Wiskunde Nederland)

In onze werkzame levens hebben wij dan wel niet intensief samengewerkt, wetenschappelijk danwel anderszins, maar anderzijds was je er wel op een aantal voor mij beslissende momenten. Daarom vond ik het ook leuk om een korte bijdrage te schrijven in je Liber Amicorum, en deze momenten nog even te memoreren.

Het eerste beslissende moment was eind 1999, dus nog juist in de vorige eeuw (altijd leuk om over “de vorige eeuw” te spreken, dat klinkt enigszins weemoedig, ver terug in de tijd). Ik was sinds 1980 werkzaam bij Philips, sinds 1990 op het bekende en befaamde Natuurkundig Laboratorium, nu bekend onder de naam High Tech Campus met zijn concept van Open Innovation dat jou vast heel erg aanspreekt. Wij hadden als groep altijd nauwe contacten onderhouden met de wetenschappelijke wereld, en wiskunde stond hoog in ons vaandel: wij vonden het belangrijk om de essentie van problemen bloot te leggen, te onderzoeken waar ze vandaan kwamen, en met deze wetenschap vervolgens nieuwe methoden te ontwikkelen welke bestand waren tegen de grillen van een grote variëteit van industriële problemen. Het heeft geleid tot een mooie verzameling publicaties en ook diepere inzichten in het toepassen van wiskunde binnen het bedrijfsleven. Contacten met de groep van Bob Mattheij, zowel in jaarlijkse bijeenkomsten tussen de groepen van de TU en Philips als ook via bestuurlijke werkzaamheden binnen de in 1994 opgerichte stichting Industriële en Toegepaste Wiskunde (ITW), leidden uiteindelijk tot het voorstel mij te benoemen tot deeltijdhoogleraar aan de TU/e. Jij was decaan van de faculteit W&I in die tijd, en ik herinner me nog levendig het aangename gesprek dat we toentertijd hadden. Ik herinner me ook nog dat je de toevoeging “elektronische” in de naamgeving minder geschikt vond, zodat uiteindelijk “Numerieke wiskunde voor de industrie” uit de bus kwam. Ik denk dat dit een goede beslissing was, het was ook een voorbeeld van je snelle en heldere denkwijze.

Het tweede moment ligt in een veel recenter verleden, te weten in 2010. Jij was inmiddels verhuisd naar Amsterdam en al vele jaren directeur op het CWI. Er werd al langere tijd gesproken over een andere organisatiestructuur van de wiskunde in Nederland, en jij bent hier nauw bij betrokken geweest, vooral bij de totstandkoming van rapporten over een nieuwe structuur. Maar ook door de centrale rol van het CWI in dezen nog eens te benadrukken, en huisvesting te bieden aan het kersverse Platform Wiskunde Nederland waarvan ik in september 2010 directeur werd. Op allerlei mogelijke manieren heb je hier ondersteuning aan gegeven, en ook in het eerste jaar van PWN kon ik bij je aankloppen en heb je, gevraagd en ongevraagd, adviezen gegeven en ideeën gefourneerd. Ik hoop dat PWN ook in de toekomst op je mag rekenen, en daarbij wil ik ons recente gesprek nog eens memoreren. In ieder geval wil ik je, namens PWN, hartelijk danken voor je bijdragen!

Ik wens je veel plezier in het post-CWI tijdperk!

Chains of Fame from Catherine of Siena to Jan Karel Lenstra

Jan Korst Gijs Geleijnse Verus Pronk Emile Aarts

Philips Research Laboratories, HTC 34, 5656 AE Eindhoven, The Netherlands
email: {jan.korst,gijs.geleijnse,verus.pronk,emile.aarts}@philips.com

1 Introduction

For those of us who have had the opportunity to discuss with Jan Karel Lenstra matters extending beyond mathematics will have discovered that Jan Karel takes an interest in more subjects than mathematical models and algorithms only. Most profound is his interest and deep knowledge of classical music in all its dimensions. His collection of more than 2000 CDs is quite amazing. It covers the entire time span of classical music ranging from the early anonymous middle-aged flute pieces up to work by contemporary composers such as Louis Andriessen. Also his interest in books deserves mentioning. His extensive personal collection of books reveals his large interest in a broad range of subjects. In addition to their content he also collects books for their design. He has become truly fascinated by typography and typesetting and takes great interest in various forms of information representation and visualization. Over the years he has extended this passion for shape and form from a passive to an active level by applying his insights and view to his own publications.

In addition to all these fine qualifications, which show that Jan Karel is a truly well-educated person, there is another remarkable characteristic, which has to do with famous people. One could say that he has this "thing" with fame. Using his well-developed internal reference guide, Jan Karel is quite outspoken when it comes to assessing peoples' fame, leading to disapproval of those who in his opinion do not deserve fame as well as to full-hearted support of those who have no fame but deserve it. This issue about fame inspired us to write a short contribution for this book of friends with the subject "Chains of Fame".

2 Hunting for famous dead people on the web

One of the more playful activities we carried out over the last few years is collecting information on historical figures by web information extraction - the activity that we usually refer to ourselves by 'hunting for famous dead people on the web'.

We first collected a large set of historical figures that lived during the past millennium by simply querying Google with queries of the type "(y_1 - y_2)", with y_1 and y_2 integers satisfying $1000 \leq y_1 < y_2 \leq 2005$ and $15 \leq y_2 - y_1 \leq 114$. For each of these approximately 100,000 queries, the at most 1000 text snippets that Google returns are analyzed to identify person names that directly precede the query issued. In this way, query *"(1685 - 1750)"* results in text snippets including *'Johann Sebastian Bach (1685 - 1750)'*. Candidate person names are simply sequences of one to four capitalized words (allowing additional infixes as *'van der'*). These candidate person names are next analysed to filter out non-person names (see Table 1). Next we quantified the fame of each remaining person, which is simply defined by the estimated page count that Google returns on the query that includes the last name and the years of birth and death. For example, the fame of Johann Sebastian Bach is chosen to be the estimated page count on query *"Bach (1685 - 1750)"*, assuming that this expression uniquely identifies a historical figure. Note that 'last name' is the last word in the word sequence. The above strategy resulted in a large collection of historical figures ranked by fame. In addition, we extracted additional personal details as gender, nationality, and one or more professions for each of the the top 10,000 persons as well as a corresponding image. This allowed us to make ranked sublists of, for example, Dutch painters of the 17th century.

person name	*non-person name*
Art Blakey	Art Deco
Mae West	West Virginia
Amy Beach	Miami Beach
HP Lovecraft	HP Inkjet

Table 1: Person names and non-person names are sometimes hard to distinguish.

1. Blaise Pascal (1623 - 1662)	9. John von Neumann (1903 - 1957)
2. Carl Friedrich Gauss (1777 - 1855)	10. Henri Poincaré (1854 - 1912)
3. Leonard Euler (1707 - 1783)	11. Alfred North Whitehead (1861 - 1947)
4. John Maynard Keynes (1883 - 1946)	12. Georg Cantor (1845 - 1918)
5. Pierre-Simon Laplace (1749 - 1827)	13. Joseph Fourier (1768 - 1830)
6. Alan Turing (1912 - 1954)	14. Joseph Louis Lagrange (1736 - 1813)
7. Pierre de Fermat (1601 - 1665)	15. David Hilbert (1862 - 1943)
8. George Boole (1815 - 1864)	16. Bernhard Riemann (1826 - 1866)

Table 2: Top-16 of persons that have as best scoring profession 'mathematician'.

As an example we give in Table 2 the top-16 of persons that have as best scoring profession 'mathematician'. If we would give the top of persons that have as one of their professions 'mathematician', then Rene Descartes and Isaac Newton would have been number 1 and 2 and also Kepler, Kelvin, Huygens, Lichtenberg, Husserl and Caroll would have been in the top-20. For further technical details, the reader is referred to [1,2].

For this special occasion, we decided to focus on the year of Jan Karel's birth: 1947. We first have a look at famous people that died in 1947 and next look at chains of fame ending in 1947.

3 The year 1947

Table 3 gives the top-20 persons that died in 1947. It is a remarkable list, not only containing two famous mathematicians (Whitehead and Hardy), but also two nobel-prize winning physicists (Planck and Lenard), two female writers (Cather and Huch), two French painters (Bonnard and Marquet), two composers (Hahn and Casella), two psychologists (Lewin and Janet) and two persons of infamous reputation (Crowley and Capone).

1. Henry Ford (1863 - 1947)	11. Kurt Lewin (1890 - 1947)
2. Max Planck (1858 - 1947)	12. Willa Cather (1873 - 1947)
3. Alfred North Whitehead (1861 - 1947)	13. Ricarda Huch (1864 - 1947)
4. Aleister Crowley (1875 - 1947)	14. Arthur Machen (1863 - 1947)
5. Pierre Bonnard (1867 - 1947)	15. Al Capone (1899 - 1947)
6. Godfrey Harold Hardy (1877 - 1947)	16. Philipp Lenard (1862 - 1947)
7. Reynaldo Hahn (1875 - 1947)	17. Albert Marquet (1875 - 1947)
8. Nicholas Roerich (1874 - 1947)	18. Jozef Tiso (1887 - 1947)
9. Alfredo Casella (1883 - 1947)	19. Pierre Janet (1859 - 1947)
10. Ernst Lubitsch (1892 - 1947)	20. Wolfgang Borchert (1921 - 1947)

Table 3: The 20 persons that died in 1947 with the highest fame.

4 Chains of fame ending in 1947

Having a large collection of historical figures, it is tempting to look for peculiarities such as persons that have the same birth and death years, such as Christopher Wren (1632 - 1723) and Antony van Leeuwenhoek

(1632 - 1723). Other famous pairs are Claude Debussy / Gustav Klimt (1862 - 1918), James Joyce / Virginia Woolf (1882 - 1941), Pablo Picasso / Ludwig von Mises (1881 - 1973) and Winston Churchill / William Somerset Maugham (1874 - 1965). Relating to 1947, probably the most interesting pair is Irving Fischer / Stanley Baldwin (1867 - 1947). The latter was three times prime minister of the United Kingdom. The former was a well-known American economist of high reputation. However, a few days before the great stock market crash in 1929, he famously predicted "the stock prices have reached what looks like a permanently high plateau." This and other personal details in this paper are all taken from Wikipedia.

Another peculiarity is that of perfect chains, which are defined as follows. Let M denote the collection of persons, where for each person $i \in M$ his or her birth year $b(i)$, death year $d(i)$ and fame $f(i)$ are given.

Definition 1 [chains].
A sequence $(p_1, p_2, \ldots, p_n)$ of persons from M is called a chain if and only if

$$d(p_i) \leq b(p_{i+1}) \qquad i = 1, 2, \ldots, n-1.$$

A chain is called perfect if and only if $d(p_i) = b(p_{i+1})$ for $i = 1, 2, \ldots, n-1$. □

We are interested in finding chains that have a high overall fame. Since the strength of a chain is determined by its weakest link, we choose as primary objective to maximize the minimum fame, i.e., we look for chains that maximize $\min_{i=1,\ldots,n} f(p_i)$. As a secondary objective function, when multiple chains have the same maximum minimum fame, we want to select from these a chain that maximizes the sum of fame, i.e., that maximizes $\sum_{i=1,\ldots,n} f(p_i)$.

Problem 1 [chain problem].
Given is a collection M of historical persons, where for each person $i \in M$ birth year, death year, and fame are given by integers $b(i)$, $p(i)$, and $f(i)$, respectively. Also, given are start year s, end year e, and upperbound u. Now, the problem is to find a chain $(p_1, p_2, \ldots, p_n)$ satisfying

$$\begin{array}{rcl} b(p_1) & = & s \\ d(p_n) & = & e \\ d(p_i) & \leq & b(p_{i+1}) \leq d(p_i) + u, \quad i = 1, \ldots, n-1 \end{array}$$

that maximizes as primary objective $\min_{i=1,\ldots,n} f(p_i)$, and as secondary objective $\sum_{i=1,\ldots,n} f(p_i)$. □

Again, focusing on Jan Karel's birth year, we look for chains ending in 1947, i.e., the chains for which Jan Karel can be the next link. Table 4 gives an example of an optimal chain for $s = 1547$, $e = 1947$ and $u = 0$.

Miguel de Cervantes (1547 - 1616)
John Wallis (1616 - 1703)
Jonathan Edwards (1703 - 1758)
Horatio Nelson (1758 - 1805)
Hans Christian Andersen (1805 - 1875)
Aleister Crowley (1875 - 1947)

Table 4: An optimal perfect chain from 1547 to 1947.

Of these, John Wallis has the minimum fame. Wallis was a British mathematician who is given partial credit for the development of modern calculus. Amongst others, he is still known for the Wallis' product, given by

$$\prod_{n=1}^{\infty} \frac{(2n)(2n)}{(2n-1)(2n+1)} = \frac{\pi}{2}.$$

Figure 1 gives the results for $s = 1347$, $e = 1947$, and different values of u. We observe that larger values of u generally result in higher values of $\min_{i=1,\ldots,n} f(p_i)$. However, the maximum minimum fame cannot increase beyond $\min(\max_{i \in M, b(i)=1347} f(i), \max_{i \in M, d(i)=1947} f(i))$. When this value has been reached, we can still improve the secondary objective.

It is intriguing to see that most of these chains start with saint Catherine of Siena and end with Aleister Crowley, the British occultist, writer, philosopher, and mystic, who is infamously dubbed "The Wickedest Man In the World." Catherine of Siena is the patron saint of fire prevention and of Italy. Most of Catherine's

body is currently interred in Rome (Basilica Santa Maria sopra Minerva), except for her head and right thumb (Siena), and her left foot (Venice).

5 Computational complexity

The chain problem can be solved by using dynamic programming. We assume that the persons in M are given ordered by birth year.

Let us first consider the case $u = 0$, i.e., no gaps are allowed between successive persons. By using an array of d years, with $d = e - s$, we can simply determine for all years $y \in \{s, s+1, \ldots, e\}$ the score of the best chain (if one exists) from s to y in $O(m+d)$ time, where m denotes the number of persons in M. By storing for each year, a pointer to the last person in this best chain (to backtrack to this person's birth year), we can easily reconstruct an optimal chain by backtracking from e. Note that this gives a pseudo-polynomial running time, since d is not polynomial in the size of the problem input. An alternative approach using a balanced binary tree can be shown to run in $O(m \log m)$ time.

The case $u > 0$ is somewhat more complicated. Using again an array of d years, we now can determine for each year y the score of the best chain from s to y in $O(u \cdot m + d)$ time. Here, however, we have to store for each year y two pointers, relating to the best chain ending at y with a person and the best chain ending at y with a gap. Also, for this case, an alternative approach exists that runs in $O(m \log m)$ time.

6 Conclusion

Some of us have had the opportunity to work with Jan Karel for more than two decades in a variety of roles both professional and personal. In hindsight, they have been good years with many highlights. If one would have to mention one single aspect that has played a significant role in our personal interaction and which can be marked as truly outstanding, then this would undoubtedly be Jan Karel's well-developed sense for quality. With great perseverance he applied the most strong quality measures not only to himself but also to others with whom he was working, and it is our full conviction that his persistent quality control has led to many outstanding results. Hence, we would like to conclude by expressing the wish that Jan Karel should be widely recognized for his intrinsic high-quality standard, and we sincerely hope that this will further increase his fame.

References

1. Korst, J., G. Geleijnse, N. de Jong, and M. Verschoor, Ontology-based information extraction from the world wide web, Chapter 10, in: W. Verhaegh, E. Aarts, and J. Korst (Eds.), *Intelligent Algorithms in Ambient and Biomedical Computing*, Springer, 2006.
2. Geleijnse, G., and J. Korst, Creating a Dead Poets Society: Extracting a social network of historical figures from the web, in: K. Aberer et al. (Eds.), *Proceedings of the 6th International Semantic Web Conference*, LNCS 4825, pp. 156-168.

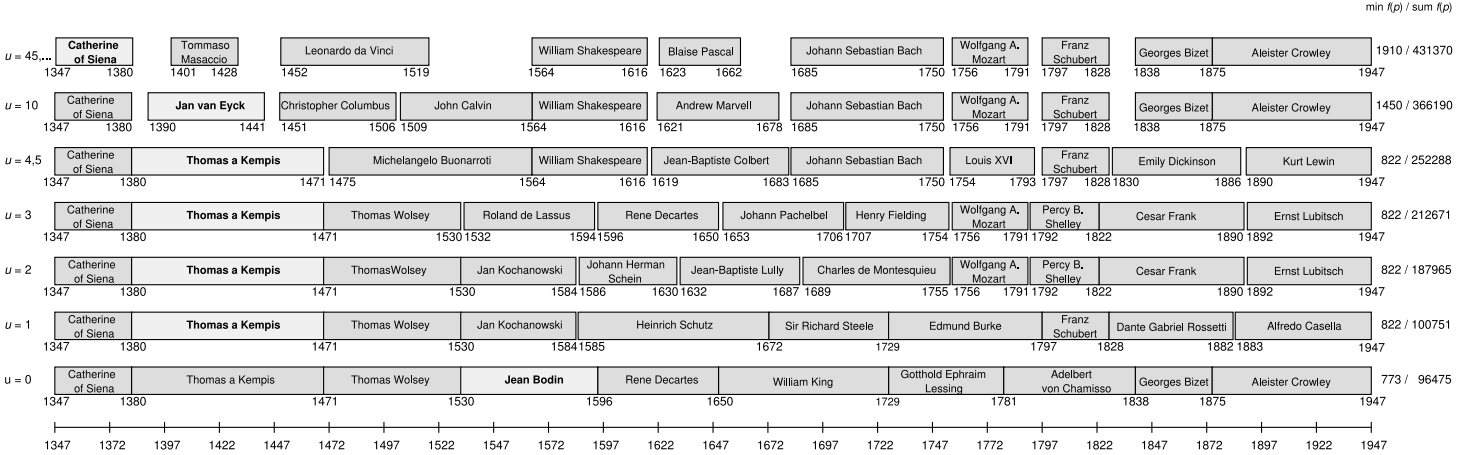


Figure 1: Optimal chains for different values of u. The weakest link is given in bold.

TU/e-collega's

Behind the screen it happens: from DSS to TSR

Kees M. van Hee
Department of Mathematics and Computer Science
Eindhoven University of Technology
k.m.v.hee@tue.nl

1. Background

In the eighties of the last century Jan Karel and I where very much interested in Decision Support Systems (DSS). We both had a background in operations research, Jan Karel in combinatorial models and I in stochastic models. We both had a strong interest in applying OR techniques in practice. Before 1980 computers where used by OR experts to find an optimal solution to some planning problem, but only the output was communicated to the planners. This was fine for strategic planning problems with a long planning horizon, but for operational planning the goals and constraints where to volatile. Due to the appearance of personal computers the concept of a Decision Support System was born: a system where the planner could update the goal and the constraints and find a new plan himself. This were in fact "do-it-yourself planning systems". Jan Karel and I where pioneers in this area. Jan Karel in his role as head of the OR group of CWI and I as managing director of the consultancy firm AKB at Rotterdam. We had a lot of experience in building DSS. We preferred the term Interactive Planning Systems (IPS) because the term DSS was misused too often according to our opinion.

In 1988 we were together involved in a DSS project at the International Institute for Applied System Analysis (IIASA) in Vienna. Jan Karel, Ko Anthonisse and I defined and published (see [1]) specifications for an IPS. It was a Resource-constrained Project Scheduling system. Later in 1994 Jan Karel and I became guest editor for a special issue of EJOR (see [2]) where the results of seven of these systems were evaluated. In the eighties each of us published several papers about DSS. One of the papers of Jan Karel e.a. (see [3]) I will never forget because of its title: Behind the Screen: DSS from an OR Point of View. I use the same ambiguity of "behind" in this paper. Paper [3] summarized the best practices of building DSS at that time. In the nineties we lost interest in DSS and we moved to other research areas.

In 2007 a common friend and a colleague in DSS, Jo van Nunen and I discussed the system of research funding: big programmes with appealing names collecting groups of scientific friends, who would do their own preferred research, independent of each other, under the umbrella of the programme. So we were thinking of a programme as appealing as a-man-on-the-moon, that could be explained in an elevator pitch, but that would contain a set of strongly related research and development projects with high potential economical value. Our solution of this problem is described in the rest of this paper. The elevator-pitch description is:

Make a system of robots that can perform daily life activities that a human can do in a normal environment, but controlled by a human at a distance.

Such a technology would have revolutionary consequences. For instance production activities can be performed at the location of the consumers or where the raw material is produced instead of moving the production to the location where labour costs are minimal.

In fact this "great idea" did not disappear and I started in 2009 a programme for Teleoperated Service Robots (TSR's). A TSR consists of two subsystems: a *slave robot* that acts in the real world and a *cockpit* from which the operator or user controls the slave robot. In our programme which takes only two years, we are building three prototypes of a TSR (called Rose-x, x=0,1,2) to perform experiments. The project is *experiment-driven*, which means that our specifications and designs are based on *use cases* and that we test them with prototypes in an *experience laboratory* to learn from experiments. The hypothesis is that you only discover problems and their solutions by performing experiments. Actually that hypothesis seems to hold. In the programme ten organizations from the Eindhoven region work closely together. One partner is ZuidZorg, which is the largest organization for home care in the Eindhoven region. They offer the experience laboratory. We have chosen home care as application domain because there is a great need for productivity improvement in this domain.
We try to use as much as possible available components and software. We specifically use open source Robot Operating System (ROS), originally developed at Stanford University and now in hands of the spin-off company Willow Garage (see [4]). The status of our project with movies of the experiments can be found in [5]).

TSR is a challenging field and there are several similarities with the DSS field of the past:

- The operator should plan the activities of the slave robot and be able to adapt an existing plan to new circumstances. There are hard mathematical problems and soft goals and constraints. For instance the operator is not able to specify exact coordinates of places to move to.
- Planning requires mathematical models, even optimization methods from OR.
- The user interaction is extremely important: the operator should obtain all relevant state information in an easy-to-read format, should be able to enter plans himself, to generate plans automatically, to evaluate plans and to update existing plans in execution.

So actually the cockpit contains a DSS! There is also an important difference: The slave robot acts in the real world: actions are not printed on paper or screen like in a DSS, but are sent to actuators that execute them. So there really happens something behind the screen. There is feed back from sensors, showing that the slave robot did not do exactly what was planned. So updating the plan is necessary. This all happens in real time, which means that the cockpit has to react fast. Since the slave robot is acting in a daily life environment there are several safety issues.

2. What are Teleoperated Service Robots?

Telemanipulation is probably the oldest form of robotics (see [6]). Telemanipulation enables a person, called operator, to act remotely as if the operator was on the spot, by copying the manipulations of the operator at a distance. Telemanipulation is in fact restricted to grasping, moving and releasing physical objects remotely. In this way it is for instance possible to write a letter at a distance by grasping a pen and paper, fixing the paper and moving the pen over the paper. A natural extension of telemanipulation is to transform the human movements, by extending or shrinking the distances, or to increase or decrease the forces executed by the operator. In this way the operator can not only copy his actions at a *distance*, he is also able to act at a different *scale*, e.g. with higher precision, or at larger reach or with more force. Probably the oldest form of telemanipulation is a pantograph, with was used to copy images at a different scale. Typical examples of telemanipulation are found in aerospace and in medical surgery.

Telemanipulation is one of the basic functions of a TSR. But a TSR has more advanced functionality. In order to compete with a human on the spot, the TSR system should be able to perform tasks fast which means that the operator should be able to give a simple command to perform a complex task. In fact a TSR needs a lot of autonomy in task performance. Before we study a TSR in more detail, we remark that the TSR field differs

fundamentally from the field of *industry robots*. The field of industry robots is mature and the most well-known applications are in the automotive industry. Compared to the TSR the industry robots look much more advanced. The big difference is that industry robots are operating in a completely controlled environment that is often designed for them. To program them, only the *kinematics* of the system are important: the control is completely determined by the coordinates of a position of the robot or its arm. For example the robot arm moves fast (often in an optimal way) exactly to a given position. An industry robot can perform an arbitrary sequence of complex tasks as many times as we like. But a TSR moves in an unknown and an adapted environment and the operator is not able to give coordinates.
For example consider the movement to a door that must be opened: the operator sees the door via a camera on a screen and he has to give a command to move to the door, to grip the door handle, to move that handle downwards and to pull (or push) the door. This (complex!) task can not be commanded by providing coordinates. There could be obstacles in the room that were not there when a similar task was performed before, so we can not replay the same set of instructions from the past, as is done with industry robots. So industry robots are *numerically* controlled by a *program* for very specific tasks in a completely known environment and they operate *autonomously*, while a TSR acts in a completely unknown environment, on the fly controlled by an operator, who provides *non-numerical* commands. Therefore the field of TSR differs fundamentally from the more classical field of industry robots.

A TSR consists of a cockpit and a slave robot. The slave robot consists of three components: a *mobile platform*, a set of *arms* (one, two or even more) each one equipped with a *gripper* and a *vision system*. The cockpit is an integrated set of devices that enables the operator to control the slave robot. Seen from a different perspective the cockpit is also a robot, but one with a "human behind the screen" or formulated differently a "human in the loop". In fact, the slave robot is like in telemanipulation, only able to grasp, move and release physical objects.
In a basic TSR the operator has to demonstrate the actions to be executed by the service precisely, maybe at a different scale. In advanced TSR's the operator has a high-level *command language* in which he can order a complex task for the service robot with a simple command. Such a command can be given to the master by means of advanced input devices such as gloves, joysticks with haptic feedback or by voice recognition. An even more advanced TSR is able to learn behaviour from past behaviour, *programming by example,* and by operator training which is in fact *supervised learning*.

We assume however that our service robot acts in unknown environments, so the system has no map of the environment and there is neither an external navigation system with beacons, nor cameras in the environment that pass information to the TSR. We also do not assume "common knowledge" in our system, so the system does not know that a door turns on hinges and how it should be opened with a door handle.
We use the term *service robot* because the TSR replaces the human by performing tasks supposed to be done by humans, sometimes in environments where humans do not like to work, such as in dangerous environments.
There is another class of robots, called *humanoids,* that differs from TSR. Humanoids are used to *imitate* humans, while TSR's are used to *replace* humans. A humanoid should look like a human and it should be able to imitate emotional contact with humans, e.g. by producing a smile, while TSR's only try to take over human actions at a distance and possibly at a different scale, with the least possible effort of the operator.
We expect that TSR's need to have some autonomous behaviour for specific domains. On the other hand it seems to be an illusion to be able to program autonomous robots to perform all types of tasks a human can do. One reason for this is that a human can perform many tasks governed by the unconsciousness, like moving the pedals of a bicycle or applauding with two hands. This is tacit knowledge. Trying to make this knowledge explicit and transfer it into programs for an autonomous robot, seems to be infeasible. Therefore human control will be necessary for all service robots that act in an unknown environment. Hence we foresee that the gap between autonomous robots (like industry robots and humanoids) and TSR's will disappear in the long run: it will always be needed to control robots with a master.

3. What is the economical impact?

There are in principle many application domains for TSR's:

- *Care* for elderly and disabled people. Here we can have two modes of application: the person who needs the care is operating the system himself, or there is another caretaker at a distance who controls the system. Of course it is possible to switch between these two modes. An example of the first mode is some user that is not able to walk and the service robot picks up the news paper from the mail box. An example of the second mode is that a caretaker at a distance helps in waking up a person with a cup of tea.
- *Cure* such as surgery has two potential modes of use. One is telemedicine, where the surgeon is at a distance, for instance on a cruise ship without a surgeon but with a service robot on board. Also the high

precision TSR is important for care, for instance in heart, eye or brain surgery.

- *Maintenance* of equipment or installations requires service engineers to be on the spot. Preventive maintenance can be planned, but in many cases equipment breakdowns require fast repair and for these cases the travelling time to the spot is a bottleneck that can be diminished using TSR. Also maintaining equipment in difficult places, like clean rooms, and in dangerous environments can be improved using TSR.
- *Security* in buildings or compounds is done mainly by means of personal inspection and security cameras. There are great opportunities for TSR here: both to observe objects from unforeseen positions as well as for preventing people to do something or even to arrest them. Also the disarming of explosives is an example in this class of applications. (Today's robots are only used to inspect explosive bombs.)
- *Manufacturing* is done these days at locations where labour is relatively cheap and not at the location where the final products are needed or where the raw materials are won. With TSR it is in principle possible to have factories "manned" with TSR and operators at a great distance. In a 24 hours economy we could have operators on several locations on earth, each working in its own day time. Also in places where manufacturing is done in an unpleasant environment, we could use TSR, e.g. in a slaughterhouse. A TSR is able to work on a different scale, which might be a specific advantage here.
- *Agriculture* is often very laborious and the margins are so small that guest workers are imported from low-labour-cost countries. In principle it is possible to make a dedicated machine for each kind of harvesting, but economically that is infeasible. With TSR it is possible to let them stay in their own living environment and let them work at a distance, like in manufacturing. Since the TSR are general purpose devices, they can be used for other tasks by other people after the harvesting period. So TSR rental could be a profitable business.

TSR's are a special class or robots, different from industry robots and humanoids which both are autonomous robots. One of the unique features of TSR's is that the operator is in the loop which enlarges the scope of applications dramatically compared to autonomous robots. The operator in the loop does not mean that the operator has to be at the same location as the slave. In fact by teleoperations we have *decoupled* the location where a human resides and the location where he performs a task. This is a fundamental forward in technology, comparable with the decoupling of the location where energy is produced and where energy is consumed: electricity is produced in a power plant and we are using it in our homes. This

decoupling in energy production and consumption created an industrial revolution. We expect that the decoupling of the location of creating an activity and the location where the activity is performed could create a similar *industrial revolution*.
Another important feature of TSR's is acting at a different *scale* as the operator. So at a larger scale the TSR might have longer arms, move faster or it can carry more weight. At a smaller scale the TSR can manipulate objects at the micro level with high precision.
Since a TSR is a system that can in principle act in all situations where a human can act, TSR's are multi-purpose devices. This means that they are in principle suitable for *mass production*, which implies that they can become affordable for a large variety of applications.
We envision new industrial activities when TSR's will really make a breakthrough. First of all there will be companies that produce the hardware components of TSR's, such as motion platforms, arms, vision systems, and interface devices for the master. Secondly a lot of software has to be built and maintained. Thirdly the integration of all components has to be done by system integrators. So they will together produce the complete TSR systems. Then there will arise a new type of *services* that might be called consultancy. This concerns the deployment of TSR's in the various application domains. This requires training of users (operators), adapting business processes and also programming the TSR's for specific tasks. Last but not least there will be companies that maintain TSR's and companies to let out TSR's.

4. Challenges

In this programme there are two kinds of challenges: scientific challenges and engineering challenges. Since many techniques and components are available today, one of the main challenges is to integrate them effectively and efficiently in order to reach our goal. Combining existing pieces into a new product is a creative activity. In fact authors arrange existing words in a new way to produce a novel and chemical engineers arrange existing atoms to create new molecules.
We emphasize the *scientific* challenges here. For the state-of-the-art of techniques see [5]. We discovered by our experiments open problems. We list some of them.

- *Computer vision.* The operator has to navigate the slave robot through an environment and he has to position the gripper to grasp, shift or release objects. The operator can see by means of cameras on the slave robot where the slave is moving and in particular where the arm with the gripper is moving. It can also see objects in the environment. But the problem is that the camera images are projected on a 2-dimensional screen and so we do not see depth, which is a serious handicap in

grasping objects. It is possible to use stereo cameras and to display an image such that with special goggles it is possible to see a 3 dimensional picture. However it is not very practical to use these goggles and the precision is also not very good: if you see that the gripper is right on top of the object that should be grasped, you are not sure that this also in reality. There is a huge amount of algorithms for computer vision mainly for autonomous systems. So there are algorithms to make a 3-dimensional computer image of an environment using stereo cameras, maybe augmented with laser scanners. But these images are not easy to interpret by human operators. So the challenge here is to combine the direct 2-dimensional images with the computed 3-dimensional images in order to give the operator fast and accurate insight in the position of the slave in its environment. The combination of real and computed images is called *augmented reality*.

- *Simultaneous control*. For industry robots there are all kinds of control algorithms to move to grippers to the right position. But our first problem is that we do not have precise positioning information and in many cases we have feedback via the operator. A more challenging problem is that for certain tasks we need to control both the arm and the mobile platform to perform a certain task. For example for opening a door it could be necessary to move the arm the gripper of which is holding the door handle and simultaneously to move the platform backwards to give the door the space to open. Another example of simultaneous control is when we have to move a plate (for instance with some cups on it) with two hands. Then we do not want to move both arms (grippers) independently, but we want to direct the centre of the plate to a particular position and the control algorithms should take care of the fact that the plate moves exactly horizontally in de chosen direction. In both cases we have to combine the control of two different systems simultaneously.
- *Command language and user interaction*. The operator should be able to give his commands in an easy to learn and efficient way. The first challenge is to determine the right level of commands, which is actually the same as the right level of task complexity and the parameters for the command. The second challenge is to determine the right user interaction device to express the command and its parameters. Here we can profit from the developments in the world of computer games. A particular aspect of the commands is that they have not only input parameters but they may also produce output, which is actually the feedback from the slave robot. Of course we have visual feedback but we can also have *haptic* feedback where the operator can feel forces. We

are looking for a good (to avoid the term "optimal") combination of these elements in order to let the TSR be as fast as a human on the spot.

- *Software architecture*. We are developing a *component-based* software architecture that makes it easy to add new functions, even at run time. So when we develop for instance a new method for navigation then it should be easy to add this function to the system such that it can work in combination with the other functions. So this is the required *plug-and-play* functionality. Further we would like the software to be applicable for different configurations of the slave robot. For instance if we add an extra arm or we replace an arm by a very long or very strong arm then the adaptation of the software should be minimal. In this way we can use the software for a whole family of TSR's. So this is the required *configurability* functionality. Last but not least we require that the software architecture enables the formal proof of safety of the robots system. It is of course essential that the TSR will never damage objects or hurt persons when it is in operation.
- *Learning*. We discussed already the need for learning capabilities. For industry robots that is easy: it is just a capture-and-replay functionality. But for TSR's this is much more difficult. Consider the task of picking balls form the table and putting them in a basket. The operator is controlling this task with five red balls on the table. What to do if the next time the five balls are in a bit different position? And what to do if there are six red balls or five red ones and a blue one? The big challenge here is not to determine *how* the TSR should learn but *what* exactly should be learnt.
- *Object recognition*. It must be possible to give the TSR the command to find a coffee cup. This means that it should be able to recognize a shape that fits the characteristics of a coffee cup. There is a lot of research in this area today, but it will take some time to let the results be useful in the TSR. We may expect that in the future most daily life objects will have a 3-dimensional image in a public library. Since most objects are designed with CAD (Computer Aided Design) systems, the data are already available.

There are all kinds of *engineering* challenges. One of them is the *speed* and *reliability* of the communication between the master and the slave. Another one is to design the slave and the master in such a way that it is relatively easy to make a *family* of TSR's with different scale properties (bigger, stronger, smaller, more precise). This requires a component structure that allows assembling different variants of components in an easy way. Of course the greatest challenge is to make the TSR for a reasonable price!

5. References

Note: Jo van Nunen passed away suddenly in 2010. Parts of this text appeared also in my contribution to his remembrance book.

1. Anthonisse J.M., Van Hee K.M. and Lenstra J.K. (1988) Resource-constrained project scheduling: An international exercise in DSS development. Decision Support Systems 4.
2. Van Hee, K.M. and Lenstra J.K. (1994) A comparative study in DSS development. European Journal of Operations Research vol. 79, nr 2.
3. Anthonisse J.M., Lenstra J.K. and Savelsbergh M.W.P (1988) Behind the screen: DSS from an OR point of view. Decision Support Systems 4.
4. Willow Garage (2008), http://www.willowgarage.org.
5. Rose (2010), http://www.robot-rose.nl.
6. Siciliano, B. and Khatib, O. (2008) Handbook of robotics. Springer-Verlag Heidelberg. ISBN: 978-3-540-30301-5.

Promovendi

Een uitnodiging tot samenwerken

Leen Stougie

In een project over computational biology met Gunnar Klau en Stefan Canzar, kwam ik het volgende probleem tegen. Aangezien jij de groep Life Sciences op het CWI hebt neergezet, en specifiek de algoritmische kant daarvan, mag je je een beetje verantwoordelijk voor dit probleem voelen.

Gegeven zijn een graaf $G = (V, E)$ met gewichten op de knopen $w : V \to \mathbb{R}$ (zowel positief als negatief) en een constante C. Bestaat er een partitie van de graaf in ten hoogste k samenhangende componenten, zodanig dat de som over de componenten van de absolute waarde van het totale gewicht van de knopen per component ten hoogste C is?

Wat is de complexiteit van dit probleem? Is het NP-volledig? Het is lang geleden dat ik me met zo'n vraag zonder aarzeling tot jou zou hebben gewend. Als het al bekend was zou jij dat geweten hebben, of je zou met iemand contact opnemen die het zou kunnen weten, of, als beide opties niet opgaan, zou je het antwoord zelf snel vinden, vooral als het probleem NP-volledig zou blijken te zijn. Jij was de meester van de reductie.

Naast het feit dat je recent geen onderzoek gedaan hebt, tenzij in het geheim, 's nachts in bed, en dan weinig succesvol, heb ik de laatste jaren een alternatief orakel gevonden in Google. Helaas leverde dat voor het bovenstaande probleem niets op. We hebben zelf een bewijs voor NP-volledigheid gevonden via een reductie van Exact cover by 3-sets. De grafen die in de reductie ontstaan hebben cykels. Nu willen we ook weten wat de complexiteit is van ons probleem indien de graaf een boom is. Dit zou wel eens efficiënt oplosbaar kunnen zijn. Aangezien je terugtreedt als directeur en ik niet gehoord heb of er al direct een volgende klus op je ligt te wachten, bestaat er een mogelijkheid dat we ons met dit soort vragen weer tot jou kunnen wenden.

Het biologische probleem achter dit fraaie graaf-partitie probleem heeft enige gelijkenis met het Test cover probleem. Voor het Test cover probleem heb je me beloofd je definitieve pensionering uit te stellen en het artikel over de Branch-and-Bound algoritmen te schrijven, waarvoor we de basis in 1984 gelegd hebben.

Het zou echt heel fraai zijn als je weer terug zou kunnen keren in het onderzoek. We hebben net MAPSP achter de rug en velen daar zullen je van harte welkom terug thuis wensen. Combinatorische optimalisering is nog steeds een gebied met heel veel vragen. Essentieel inzicht ontbreekt nog steeds, en niet alleen met

betrekking tot de P versus NP vraag. Er bestaan veel trucs die toegepast kunnen worden, maar bijna zonder uitzondering ontbreken voldoende voorwaarden voor toepasbaarheid. Het lijkt erop dat iedere stap voorwaarts slechts bestaat uit het oplossen van een probleem en op zijn best een nieuw stuk gereedschap aan het arsenaal toevoegt waarmee toekomstige problemen te lijf gegaan kunnen worden. Ik zal aan de hand van een recent onderzoekssucces, waar ik zelf bij betrokken ben, aangeven wat ik hiermee bedoel, daarmee een beetje bijdragend aan wetenschappelijke inhoud van dit vriendenboek.

Onlangs [1] ben ik, samen met Sylvia Boyd, René Sitters, je wettelijke wetenschappelijke zoon, en Suzanne van der Ster, je toekomstige wetenschappelijke kleindochter, zo gelukkig geweest om een flinke duw te geven tegen het onderzoek naar een verbetering van het resultaat van Christofides voor het Metrisch-TSP, ook wel Δ-TSPgenoemd. De grens op de approximatie-ratio van 3/2 voor dit prachtige algoritme uit 1976 is nog steeds het beste dat bestaat voor polynomiale tijd algoritmen. Iedereen is er van overtuigd dat een betere benaderbaarheid mogelijk zou moeten zijn, maar 35 jaar onderzoek heeft tot nu toe niets opgeleverd.

Recent is deze grens verbroken voor wat algemeen als een belangrijke deelklasse van Metrisch-TSP wordt gezien. Het betreft de Graaf-TSP, waarbij de afstand tussen twee knopen gegeven wordt door het aantal kanten op het kortste pad er tussen. Deze deelverzameling van Metrisch-TSP instanties bevat alle traditionele, klaslokaal, slechtste gevallen voor Christofides' algoritme. In 2010, hebben wij ratio's 4/3 en 7/5 weten te bereiken voor Graaf-TSP beperkt tot respectievelijk kubieke grafen en grafen van graad ten hoogste 3 (merk op dat ook deze verdere beperking de traditionele slechtste gevallen voor Christofides bevat). In het begin van dit jaar, 2011, was er een ϵ van de ratio 3/2 afgesnoept [2] voor algemene Graaf-TSP. In het voorjaar is, deels gebaseerd op ideeën uit ons artikel, de grens naar 1.461 gebracht [3]. Ons artikel is inmiddels verschenen in de proceedings van IPCO 2011. Laatstgenoemd artikel wordt, met kans 1, geselecteerd voor FOCS. En ik voorspel dat voor het eind van 2011 de 4/3 voor Graaf-TSP zal zijn bereikt. Dit is opwindend onderzoek in de combinatorische optimalisering!

Zo zeker als de acceptatie [3] voor FOCS zal zijn, zo zeker zal dit onderzoek geen wiskundig inzicht verschaffen dat het oplossen van combinatorische optimaliseringsproblemen in de praktijk wezenlijk zal vergemakkelijken. In het beste geval levert het een nieuw inzicht om ook approximatie-ratio's voor andere problemen te verbeteren. Voor de vooruitgang van het vakgebied Operations Research zal het, zeker op korte termijn, slechts van beperkte waarde blijven. Daarvoor zou het veel interessanter zijn om erachter te komen in welke situaties technieken als simulated annealing, genetische algoritmen, die jij in het verleden aanduidde als homeopatische algoritmen, gebruikt kunnen worden. Er is tot nu toe amper enige structuur in combinatorische optimaliseringsproblemen ontdekt die antwoorden op dergelijke vragen binnen handbereik brengt.

De meest interessante vraagstukken binnen de OR liggen dus nog wijd open en zij zouden een enorme uitdaging moeten vormen voor onderzoekers die het zich kunnen permitteren om slechts spaarzaam resultaten te krijgen. Jij hebt nu deze status bereikt. Ik zal je over een jaar of 10 volgen. In een bijdrage aan het vriendenboek ter gelegenheid van je 60ste verjaardag heb ik benadrukt hoe belangrijk je in een vroeger deel van mijn leven bent geweest. Het zou me verheugen als we op wat gelijkwaardiger niveau een rol in de rest van elkanders leven kunnen spelen.

Referenties

[1] S. Boyd, R. Sitters, S. van der Ster, and L. Stougie, TSP on cubic and subcubic graphs, in: *Proc. of the 15th Conference on Integer Programming and Combinatorial Optimization (IPCO)*, 2011.

[2] S.O. Gharan and A. Saberi and M. Singh: A Randomized Rounding Approach to the Traveling Salesman Problem, *manuscript*, 2011.

[3] T. Mömke and O. Svensson: Approximating graphic TSP by matchings, *manuscript*, 2011.

Statistiek, minimal music, en optimalisering

Jan Karel Lenstra heb ik leren kennen in de jaren ’80, meer dan 20 jaar geleden. In die tijd werkte ik op het CBS in Heerlen als wetenschappelijk medewerker. Ik wilde promoveren, en was op zoek naar een promotor. Via een student van de KUB (tegenwoordig de Universiteit van Tilburg) die een doctoraal stage bij ons deed, en wiens begeleider ik was op het CBS, kwam ik in contact met Arie Kapteyn, Theo Nijman en Lex Schrijver. Toen ik Lex Schrijver vertelde over het onderwerp waarop ik wilde promoveren, vertelde hij dat zijn interesses meer theoretisch waren. Hij verwees mij door naar Jan Karel. Deze bleek geïnteresseerd in het onderwerp. Dat gold ook voor Kapteyn en Nijman. Het onderwerp van mijn proefschrift had te maken met het project op het CBS waar ik toen bij betrokken was, het data controle en correctie project o.l.v. Wouter Keller. Dit project was belangrijk voor het CBS en heeft tot een omwenteling geleid van de verwerking van statistische gegevens. Een ander product dat zijn oorsprong vond in dit project is het Blaise systeem, voor de verzameling verwerking van enquêtegegevens. Het systeem is ontstaan in een tijd dat de meeste CBS-gegevens verzameld werden op steekproefbasis. Innovatief was toen dat dat via Blaise met behulp van elektronische vragenlijsten kon die op laptops van enquêtrices konden worden ‘afgespeeld’ tijdens interviews. De Blaise vragenlijst bepaalde zelf de vragen en de volgorde waarin deze aan een respondent werden gesteld, vanaf een vaste beginvraag n vervolgens op basis van gegeven antwoorden. Tijdens het interview kon de vragenlijst ook combinaties van gegeven antwoorden checken. Als er een inconsistentie geconstateerd werd, kon die ‘in het veld’ meteen worden opgelost door de respondent ermee te confronteren, en deze de gelegenheid te geven één of meerdere van de gegeven antwoorden te herzien. Dat was met papieren vragenlijsten, die tot dan toe gebruikelijk waren om enquêtegegevens te verzamelen, volstrekt onmogelijk. Een ander voordeel van het gebruik van elektronische vragenlijsten was dat de verzamelde antwoorden in de vorm van een databestand beschikbaar waren. Daardoor waren ze meteen geschikt voor verdere elektronische verwerking. Bij de traditionele papieren vragenlijsten moesten de gegevens eerst vertoetst worden. Dit gebeurde door zalen vol ponstypistes. Door de komst van Blaise zijn deze in vrij korte tijd verdwenen van het CBS. Het Blaise systeem bestaat nog steeds, en het wordt in veel statistische bureaus in de wereld gebruikt.

Het werk voor mijn proefschrift had te maken met enerzijds het testen van de correctheid van vragenlijsten, waarvoor een ingewikkelde routing- en editstructuur gedefinieerd kan zijn, en anderzijds met de verwerking van gegevens die met een (eventueel) ingewikkelde vragenlijst (elektronisch of m.b.t. een papieren vragenlijst) verzameld konden zijn. De routingstructuur definieert hoe men van vraag naar vraag springt, op basis van eerder gegeven antwoorden (vaak het laatst gegeven antwoord). De editstructuur bestaat uit een verzameling controles op combinaties van waarden, die kunnen zijn toegevoegd aan een elektronische vragenlijst, om de consistentie van combinaties van antwoorden te controleren.
Het eerste deel van mijn proefschrift had betrekking op de formele structuur van vragenlijsten, en behandelde de complexiteit van een aantal consistentiechecks. Het tweede deel ging over de controle en correctie (data editing) van antwoorden gegeven op een vragenlijst, die elektronisch of van papier kon zijn. Voor het eerste deel was Jan Karel speciaal verantwoordelijk, en voor het tweede, dat meer statistisch van aard is, Theo Nijman en Arie Kapteyn. Met Jan Karel heb ik veruit de meeste contacten gehad in de tijd waarin het proefschrift werd geschreven. Doordat onze werkkringen nogal ver van elkaar verwijderd waren zagen we elkaar niet dagelijks. Wel hadden we geregeld telefonisch contact. Meestal belde Jan Karel me op als hij over een probleem had nagedacht, en vertelde dan dat de oplossing voor dat probleem hem onder de douche te binnen was geschoten. Hij schetste dan de oplossing, waarna ik aan de slag ging om een en ander uit te werken. Ook bezocht ik, zo mogelijk, Jan Karel om de paar weken, ofwel op zijn werk of bij hem thuis in Amsterdam. Deze bezoeken heb ik altijd bijzonder aardig en aantrekkelijk gevonden. Ze gaven me vaak niet alleen de mogelijkheid om Jan Karel te spreken, en ook enkele van zijn medewerkers, maar ook om een universiteitsbibliotheek te bezoeken, een soort Walhalla voor mij destijds.

Het leuke van de persoonlijk bezoeken vond ik dat ze niet alleen over het proefschrift gingen, maar ook over zaken die in de persoonlijke sfeer lagen. Ook was muziek steevast een onderwerp waar we over praatten, en naar luisterden. Jan Karel hield vooral van het klassieke repertoire, en ik zelf was in die tijd erg vol van de minimal music, die ik ongeveer in die tijd ontdekt had. Die ontdekking was een LP van John Adams met de Harmonielehre erop, in de stadsbibliotheek van Heerlen. Dit werk heeft een diepe indruk op me heeft gemaakt en heeft een wereld aan schitterende muziek voor me geopend, met werk van componisten als, naast Adams, Steve Reich, Philip

Glass, Simeon ten Holt en (in mindere mate) John Cage. Jan Karel kon mijn enthousiasme voor die muziekstijl niet helemaal appreciëren. Ik herinner me dat hij me bij het feest ter gelegenheid van mijn promotie een CD als cadeau heeft gegeven met de Goldberg variaties van Bach, gespeeld door Glenn Gould, met als commentaar daarbij dat dit echte minimal music was. Dat was voor mij aanleiding om me meer in het werk van Bach te verdiepen. De interesse voor de muziek van Bach is tot de huidige dag gebleven (evenals die voor minimal music, trouwens).

Voor mij was het ook een periode om mezelf bij te scholen in de complexiteitstheorie van algoritmen en de combinatorische OR. Ik ben afgestudeerd in de wiskunde, met als hoofdvak waarschijnlijkheidsrekening. Ik heb het me verdiepen in die nieuwe disciplines als een bijzonder leuke en boeiende uitstap ervaren. Daardoor ben ik me ook meer gaan interesseren voor OR toepassingen binnen de officiële statistiek. Als gevolg hiervan hebben toepassingen op het gebied van de combinatorische optimalisering binnen onze afdeling ook een aantal jaren in de belangstelling gestaan. Toepassingen waren er bijvoorbeeld op het gebied van data editing (controleren en corrigeren van gegevens) en statistische beveiliging (het aanpassen van gegevens zodat individuele bijdragen niet herkend of nauwkeurig berekend kunnen worden als ze gepubliceerd worden). Zelf heb ik me daarna een jaar of tien met dit laatste onderwerp bezig gehouden, vanuit het CBS in Voorburg, waar ik na mijn promotie ben gaan werken. Deze activiteit van een klein groepje medewerkers heeft tot een flink aantal nota's, rapporten, publicaties en twee boeken geleid. Diverse afstudeerders hebben een afstudeerscriptie kunnen schrijven op het gebied van de OR toegepast op onderwerpen uit de officiële statistiek, in het bijzonder de data editing en de statistische beveiliging. Gedurende al die jaren is er geregeld contact geweest tussen onze groep en die van Jan Karel, in het bijzonder ook met Cor Hurkens. Het lag dan ook voor de hand om Jan Karel te vragen om te participeren in een Europees project op het gebied van de statistische beveiliging, toen de mogelijkheid daarvoor zich aandiende. Niet alleen deed Jan Karel mee, maar ook heeft hij mij in contact gebracht met twee andere OR specialisten, Matteo Fischetti uit Italië en Juan José Salazar uit Spanje (Tenerife), die bereid bleken mee te werken aan dit project, het SDC project. Met een goed gevoel denk ik ook terug aan de bezoeken in Eindhoven in het kader van het SDC project. Die waren een voortzetting van de eerdere bezoeken bij hem thuis in Amsterdam, toen ik met mijn proefschrift bezig was. Net als toen ging het niet alleen over vakinhoudelijke zaken. Tijdens

zo'n bezoek bij Jan Karel thuis luisterden we steevast naar klassieke muziek. Hij had een indrukwekkende kast vol met CD's waaruit hij dan een voor de gelegenheid passende selectie maakte, om af te spelen tijdens ons gesprek, onder het genot van een biertje. Ik herinner me nog goed dat ik bij één zo'n bezoek uit een flinke stapel afgedankte, maar nog in zeer goede staat verkerende LP's, die exemplaren mocht meenemen die mij interesseerden. Dat aanbod heb ik toen dankbaar aanvaard. De LP's heb ik nog steeds, en heb ze destijds regelmatig afgedraaid. Dat is nu echter al jaren niet meer het geval. Ook bij mij hebben de CD's op een gegeven moment de LP's verdrongen.

Het SDC project bleek niet alleen een zeer vruchtbaar, maar ook een zeer leerzaam, en nuttig project. Het heeft mij in staat gesteld, als projectleider, om diverse interessante reizen te maken, met inspirerende en aardige mensen in contact te komen, op congressen of bij andere gelegenheden. Het heeft ook tot een software pakket geleid, ARGUS, dat tot op heden gebruikt wordt op het CBS, en ook andere statistische bureaus. Op zijn gebied heeft het een standaard gezet. Sergey Tiourine, die destijds als promovendus in Jan Karels groep werkte, heeft zijn bijdrage aan het SDC verwerkt tot een hoofdstuk in zijn proefschrift. Het was een leuke ervaring Sergey enige vragen te mogen stellen tijdens zijn promotie.

Het einde van het SDC project betekende voor mij persoonlijk ook het einde van mijn intensieve betrokkenheid bij de statistische beveiliging op het CBS. Hoewel ik nog een tijd op een ander gebied op het CBS werkzaam ben geweest waar optimalisering werd gebruikt (optimaal alloceren van steekproefadressen aan enquêtrices), hield na enige jaren ook mijn bemoeienis met de OR op. Ik ben in de loop der jaren naar andere gebieden afgedreven. Dit had tot gevolg dat ook de contacten met Jan Karel allengs minder werden. Dat heb ik wel jammer gevonden vanwege de prettige persoonlijke contacten tijdens de hierboven beschreven periode van samenwerking.

Ik dank de initiatiefnemers van dit liber amicorum voor de uitnodiging een bijdrage te leveren. Dit bood een mooie gelegenheid om herinneringen op te halen aan een periode die ik als zeer aangenaam en stimulerend heb ervaren.

Tot slot wil ik me rechtstreeks tot Jan Karel richten.

Jan Karel, ik hoop dat je nog vele jaren voor de boeg zult hebben, in goede gezondheid, lichamelijk zowel als geestelijk. Ik dank je voor de vriendschap en de hulp die ik van je heb mogen ontvangen. Ik ben er nog steeds trots op dat jij één van mijn promotoren bent geweest. Achteraf bezien ben ik een passant in het vakgebied van de OR geweest, dat jouw geestelijke habitat is. Ik zie het als een voorrecht dat ik met een grootheid op dit terrein als jij een tijd heb mogen optrekken en je mij liet profiteren van je kennis en wijsheid. Ook je gevoel voor stijl, je muzieksmaak en je humor heb ik altijd zeer gewaardeerd. Dit geldt ook voor je efficiënte manier van werken. Ik heb je op dit punt altijd beschouwd als een lichtend, en nastrevenswaardig voorbeeld.

Leon Willenborg

Speciale editie: JKL Boulevard

Jan Karel aan de zijlijn (en soms als speler)

Door: Marjan van den Akker en Han Hoogeveen

In het onderstaande willen we twee onderbelichte facetten uit het leven van Jan Karel Lenstra belichten. Het gaat hierbij om de toch wel belangrijke onderwerpen: relaties en sport. We beginnen met de relaties.

'Het groeide en bloeide in het Dommelgebouw'

Wiskundigen staan bekend als kamergeleerden die in een ivoren toren leven, maar Jan Karel was een econometrist en kon zich daarom altijd bezighouden met mensen en hun beweegredenen. De verkregen informatie (en alles wat daarvoor doorging) werd dan op gezette tijden uitgewisseld resulterend in roddelpartijen met Emile Aarts. Voor onwetende buitenstaanders leek het er soms op dat Jan Karel bijkluste als relatiebemiddelaar, waar hij zelf ook van profiteerde. Deze speciale verslaggevers willen op gepaste wijze een paar feiten/roddels presenteren die deze roddels/feiten ondersteunen.

Han Hoogeveen werd, zonder zijn promotie op zak te hebben, aangesteld als postdoc in Eindhoven. Net op dat moment dreigde de faculteit een vrouwelijke AIO uit de groep op te zadelen met de onaangename taak om werkcolleges Statistiek te geven (met als verantwoordelijke docent Nico L.). Jan Karel bood Han de kans om een heldenrol te spelen door zich schijnbaar op te offeren om deze werkcolleges op zich te nemen. Een investering in de toekomst zou later blijken.

Met enige regelmaat verscheen er in Eindhoven een Scandinavische dame die werkzaam was in het vakgebied. Zij werd ook, met de hele groep, uitgenodigd op het afscheidsfeest bij Jan Karel thuis. 'Men' vroeg zich natuurlijk af waarom zij werd uitgenodigd. Enkele jeugdige UD's uit de groep meenden dat de dame in kwestie voor hen kwam, maar later kwam de aap uit de mouw, al moest de groep het uit de krant halen dat de tortelduifjes waren getrouwd.

Jan Karel zag meer mogelijkheden binnen de groep (er waren per slot van rekening nog twee vrijgezelle dames). Zo stuurde hij Marjan van den Akker samen met Cor Hurkens naar de conferentie 'Netflow'. Het ingestuurde paper had weinig met dit onderwerp te maken, maar de conferentie vond plaats in een

goede ambiance in Toscane. Dit leidde via een omweg tot resultaat, want bij terugkomst maakte Han Hoogeveen er werk van.

Over de eventuele betrokkenheid van Jan Karel bij het tot stand komen van de relatie tussen Marc Wennink en Cleola van Eijl wordt wel gefluisterd, maar dit zijn niet meer dan vage roddels. In ieder geval kan niet worden ontkend dat Jan Karel heeft bijgedragen aan het creëren van de benodigde gunstige omstandigheden.

'Piet Keizer'

Het tweede thema is sport. Sport kun je onderverdelen in actief en passief. Het actieve sportleven van Jan Karel kan in één woord worden samengevat: afwezig (in ieder geval in de laatste 25 jaar). Het enige dat enigszins in de buurt komt is het hanteren van de keu in Leeds (met zeer beperkt succes); vermoedelijk was dit te wijten aan een combinatie van alcohol en de aanwezigheid van een bepaald iemand (zie boven). Het passieve sportleven is duidelijk uitgebreider. Hoewel je van een gepassioneerde liefde niet echt kunt spreken, is er toch één punt dat eruit springt: voetbal (en dan met name Ajax). Zo is Jan Karel ooit eens met zijn AIO van het eerste uur, Anton Kolen, en een latere generatie AIO naar Ajax – Porto geweest. Een duel om de supercup, dat Ajax mocht spelen omdat het in 1987 de Europa Cup II had gewonnen. Bij de voorbeschouwing liet Jan Karel zijn grote eruditie op het gebied Ajax versus Portugese clubs in de Europa Cup blijken door te vertellen over de strijd van Ajax tegen Benfica in het jaar

1968. Bij de thuiswedstrijd lag het veld onder de sneeuw, maar de Portugezen konden beduidend beter overweg met deze, voor hen onbekende omstandigheden, en wonnen met 1-3, waarbij een van de Benfica-spelers uit pure vreugde een berg sneeuw op zijn hoofd deed. Vol vuur betoogde Jan Karel dat Ajax in de uitwedstrijd uit deze bijna kansloze situatie ontsnapte door in Lissabon met 1-3 te winnen, waarna Ajax de beslissingswedstrijd in Parijs met 3-0 won. Helaas bleek al snel dat Jan Karels kennis van het Ajax uit 1987 niet zo uitgebreid was. Zijn beste commentaar was dat linksbuiten Alistair Dick 'duidelijk niet hetzelfde in zijn benen had als Piet Keizer'. Ajax hield, in vergelijking met 1968, de schade beperkt door kansloos met 1-0 te verliezen, maar in tegenstelling tot 1968 werd uit niet met dezelfde cijfers gewonnen maar verloren, zodat de supercup voor Porto was.

Kennis van sport is natuurlijk geen must, toch is een zekere basiskennis wel nuttig. Hoewel Jan Karel daarover beschikt (hij weet iets over het Ajax uit de gouden jaren, en het Ajax-team uit 1972 is volgens een Italiaanse computer-simulatie het beste team allertijden), moet hij hoognodig worden bijgeschoold (je kunt dit beschouwen als inburgering in het Nederlandse sporterfgoed). We doelen hierbij natuurlijk op een wedstrijd van bijna 60 jaar geleden, waarin Ajax tegen Heerenveen speelde. In Heerenveen speelde toen niet alleen Jan Lenstra, maar ook de iets bekendere Abe Lenstra, die als de beroemdste Lenstra allertijden wordt gezien. Veel leesplezier toegewenst.

Abe Lenstra wordt na de wedstrijd op de schouders genomen.
Foto LC

In een zinderende wedstrijd wint Heerenveen van Ajax. Bij een 5-1 achterstand krijgt Abe Lenstra het op zijn heupen. De verbluffte Ajacieden worden in het laatste half uur weggespeeld door een swingend Heerenveen.

Abe verslaat Ajax in voetbalkraker

Neen, dit is met geen ook nog zo vlotte pen te beschrijven, dit hypersensationele voetbal-gevecht dat zich daar gisteren in het Heerenveense Sportpark heeft afgespeeld. Werkelijk, woorden schieten jammerlijk tekort, om dat laatste opwindende halfuur van deze kampioenskamp weer te geven. Van die weergaloze, zenuwslopende minuten waarin de nimmer versagende pompebêdden-dragers het hebben bestaan een 5 – 1 achterstand op een technisch verder superieur Ajax niet alleen in te lopen, maar daarenboven zelfs in een 6 – 5 overwinning om te zetten !

Een zege, behaald temidden van het tot ongekende hoogte opgezweepte enthousiasme der golvende rijen supporters die een wedstrijd-van-één-maal-in-je-leven meemaakten. Geloof ons, nimmer was een overwinning zo zoet als die van Heerenveen gisteren. En nog nooit is een overwinning zo zuur geweest, als die de Amsterdamse rood-witten zondagmiddag leden. (..)

Toen Michels in de vijftiende minuut een fout van Voolstra radicaal afstrafte en de stand op 5 – 1 bracht, toen kon niemand meer hopen op een zege van Heerenveen. Ze waren geslagen; dit wàs niet meer in te halen meende men.

Goed de geestdrift ontvlamde weer, toen Abe onmiddellijk na dit vijfde Ajax – doelpunt een corner-bal van Hofma door de Ajax keeper voor z'n voeten geworpen gestompt kreeg en daarna het net deed trillen (2 – 5). De Ajaxieden namen dit echter onverstoord op. Wat kon hun immers gebeuren? En de Heerenveense supporters zagen nog steeds het zwerk boven de Friese kampioenen drijven.

Maar Jonkman zette evenwel z'n ploeg om. Zelf ging hij op de midvoorplaats. Brandsma speelde rechtsbuiten en Ploegh rechtsbinnen. En Jan Lenstra verruilde z'n backplaats tegen de half van Molenaar. Reeds enkele minuten later wierp deze opstelling vruchten af. In de 25ste minuut gaf Jan Lenstra een zuiver afgestemde verre pass naar Brandsma die onthoudbaar inkopte (3 – 5).

Pas daarna sloeg de vlam der hoop weer in de duizenden Heerenveen aanhangers. Nu beseften zij, dat een gelijk spel nog mogelijk was. Nog twintig minuten te spelen. Heerenveen speelde nu sterk aanvallend, elke kans, elke welfeling van de Ajax defensie benuttend. Drie minuten later werd Abe in het strafschop gebied door Potharst onder de voet gelopen en scheidrechter Nijs wees onverbiddelijk naar de witte stip. En terwijl Abe kalm wegwandelde, knalde Jonkman de penalty onhoudbaar langs de Ajax-doelman.

Golven van geestdrift sloegen door de toeschouwers. Even leek het er op, dat Ajax z'n vorm van voor de rust zou hervinden, maar dra liepen de Friezen al weer storm op het Amsterdamse doel. De Ajaxieden werden geheel overspoeld. Alles verdedigde bij hen, maar voor dit laaiende enthousiasme van de Heerenveeners was geen kruit meer gewassen.

Vijf minuten voor het einde besliste Jonkman een duel met Potharst in zijn voordeel, Hofma snelde toe en koppend werd de gelijkmaker geboren. En luttele minuten daarna kwam dat winnende zesde doelpunt van Heerenveen ontstaan uit een combinatie van Ploegh, Abe en Brandsma, dat met een hard schot afsloot.

En toen stond het Sportpark op z'n kop. Iedereen stond van z'n plaats op. Een oorverdovend gejuich, gebrul en gekrijs steeg op uit de duizenden kelen. Armen zwaaiden wild door de lucht. Petten en hoeden werden de lucht ingegooid. Men vergat zichzelf en iedereeen in een wild en onstuimig enthousiasme, dat tot de laatste minuut voortduurde. (8-5)

De lokaal zoekende Jan Karel

Tjark Vredeveld

Beste Jan Karel,

Nu je afscheid neemt als directeur van het CWI, en daarmee ook van je bestuurlijke carrière in de wetenschap, zal je weer tijd krijgen voor niet-bestuurlijke zaken, bijvoorbeeld voor onderzoek. Misschien wil je wel weer eens nadenken over de theoretische aspecten van lokaal zoeken, de richting waarin je mij hebt gestuurd tijdens mijn tijd als Onderzoeker-in-Opleiding. Ik wil je hierbij door middel van een kleine selectie van artikelen vertellen wat er in de laatste jaren zoal gebeurd is in de doorsnede van dit gebied met het terrein van scheduling. Bij mijn eerste artikel hierover [9, 10], dat ik samen met Petra heb geschreven, stond je nog toe dat we de door jou in [6] geïntroduceerde $(\alpha|\beta|\gamma)$-notatie gebruikten – nu hadden we in dit artikel ook alleen maar standaard problemen bekeken en was deze notatie niets anders dan een eenvoudige afkorting. Enkele jaren geleden, tijdens een Dagstuhl workshop, werd je hier iets strikter in: voor iedere keer dat tijdens een presentatie de $(\alpha|\beta|\gamma)$-notatie gebruikt werd, moest de spreker een boete van € 10 betalen. Als je belooft dat het geld dat je hiermee ophaalt naar een goed doel gaat, zal ik bij de eerst volgende gelegenheid een aantal keer zondigen.

Zowel tijdens als ook na mijn tijd als OiO heb ik veel plezier beleefd aan mijn onderzoek naar het slechtste geval gedrag van lokaal zoeken. Dat wij niet de eerste waren die lokaal zoek methoden vanuit een slechtste geval analyse onderzochten, wist ik natuurlijk wel. Jij hebt tijdens ons onderzoek Petra en mij toen al gewezen op een artikel uit 1998 [8], maar tijdens het schrijven van de inleiding van mijn proefschrift was ik toch enigszins onaangenaam verrast toen ik in de kelders van de bibliotheek een artikel van zo'n 20 jaar daarvoor uit het stof haalde. Onaangenaam, omdat Finn en Horowitz [5] namelijk toen al de eerste twee stellingen uit het IPCO artikel van Petra en mij hadden bewezen. Je vroeg je toen af of we konden schrijven dat we het onafhankelijk van hun hadden bewezen, wat inderdaad het geval was, maar al snel kwamen we tot de conclusie dat dit resultaat toch een van de kleinere resultaten was uit mijn proefschrift. Het artikel van Finn en Horowitz heeft uiteindelijk nog wel tot een artikel voor Cor en mij geleid: Zij beweerden namelijk dat een zogenaamde jump-optimale oplossing voor

$P||C_{\max}$[1] door een iteratief verbeteringsproces in $\mathcal{O}(n)$ jumps gevonden kan worden, waarbij n staat voor het aantal taken dat uitgevoerd moet worden. Hiervoor gaven zij een argument dat mij niet geloofwaardig over kwam. In ons artikel [7] hebben Cor en ik aangetoond dat zij het inderdaad fout hadden en dat de analyse in [1] de goede was en dat deze niet te verbeteren viel.

Hoewel wij dus niet de eerste waren die een slechtste geval analyse van lokaal zoek methoden onderzochten, denk ik dat ons werk wel een nieuwe impuls aan dit type onderzoek heeft geleverd. Gerhard heeft in Twente nog een Aio project besteed aan dit onderzoek. Tijdens MAPSP 2005 heb ik nog met deze Aio en Johann Hurink over het werk dat zij gedaan hadden [2] gepraat en een uitnodiging gekregen om eens in Twente langs te komen om samen verder te werken aan dit type onderzoek. Natuurlijk heb ik deze uitnodiging aangenomen en we hebben als uitgangspunt voor ons werk een vraag die Gerhard mij tijdens mijn verdediging gesteld had, genomen. Zijn vraag was of een bepaalde buurruimte die exponentieel veel buur-oplossingen kan bevatten een betere garantie zou kunnen geven dan de eenvoudige jump-buurruimte. Tijdens mijn verdediging had ik al aangetoond dat dit niet het geval was. Aangezien het slechtste voorbeeld voor deze zogenaamde split buurruimte alles behalve jump-optimaal is en het slechtste voorbeeld van de jump buurruimte zeker niet split-optimaal is, rees de vraag of een gecombineerde buurruimte een betere prestatie garantie zou kunnen geven. Dit bleek inderdaad het geval te zijn, maar als het aantal machines erg groot zou zijn, was het verschil tussen de twee garanties verwaarloosbaar. Het opmerkelijke was dat als we de split buurruimte met een wat intelligentere versie van de jump buurruimte combineren, we dan wel de garantie van ongeveer 2 naar 3/2 konden krijgen [3].

Een kleine tien jaar na mijn werk met Petra, heb ik twee van mijn promovendi, één uit Maastricht, de ander uit Quito, aan het werk gezet op basis van een resultaatje dat Petra en ik uiteindelijk niet opgenomen hebben in ons artikel. De vraag was hier hoe goed lokaal zoeken werkt voor scheduling problemen waarbij een taak slechts op een beperkte deelverzameling van de aanwezige machines uitgevoerd mag worden. Daar waar we een constante garantie konden bewijzen voor het standaard probleem op identieke machines, bleek dat voor het probleem met machine restricties, de garantie voor jump optimale oplossingen groeit met de wortel van het aantal machines.

[1] Graag hoor ik van je naar welk goed doel ik het bedrag van € 10 over moet maken

Tijdens de verdediging van mijn proefschrift kreeg ik – ik geloof van Kees Roos – een vraag over of slechtste geval analyse de goede maatstaf is om het theoretisch gedrag van lokaal zoek methoden te analyseren. Hierbij werd voorgesteld om, in plaats van het slechtste geval, te kijken naar het tweede slechtste lokale optimum. Ik zag aan je gezicht – of althans ik dacht toen aan je gezicht te zien – dat je dat geen goed idee vond en ook ik zag hier niet echt het heil van in. Ik denk dat ik toen heb geantwoord dat we dan beter een gemiddelde analyse konden maken. In het afgelopen jaar, heb ik hier dan eindelijk werk van gemaakt, samen met een aantal collega's [4]. Voor de scheduling problemen waarin de jump buurruimte een prestatie garantie heeft die groter wordt naarmate het aantal machines groter is, hebben we een zogenaamde smoothed analyse uitgevoerd – helaas heb ik hier geen goede Nederlandse vertaling voor. *Smoothed* analyse ligt ergens tussen slechtste geval en gemiddelde analyse in: voor de slechtste instantie wordt de verwachte prestatie geanalyseerd als deze instantie onderhevig is aan een aantal toevallige veranderingen. Het interessante hier is dat als de taken op alle machines uitgevoerd mogen worden, de verwachte prestatie garantie van $\Theta(\sqrt{m})$ of $\Theta\,(\log m / \log\log m)$ in het slechtste geval naar een constante in het gemiddelde geval gaat, terwijl in het geval van gerestricteerde machines, de slechtste gevallen robuust blijken te zijn en dat de prestatie garantie niet veranderd in de orde van grootte.

Ik denk dat de smoothed analyse een interessante nieuwe methode is om lokaal zoekmethoden te onderzoeken. Interessant zou het zijn om te weten of het voorbeeld waarmee we bewezen hebben dat het gegeneraliseerde graafkleuringsprobleem exponentieel veel verbeteringsstappen nodig heeft om tot een flip optimale oplossing te komen [11], stand houdt als het een aantal toevallige veranderingen ondergaat, waarbij ik voorstel dit toch echt eerst te bestuderen voor het max-cut probleem, aangezien dat vele malen interessanter is. Tenslotte, is er ook nog veel onderzoek te doen naar de berekenbaarheid van lokale optima, aangezien hier weinig over bekend is. Zelfs voor de swap buurruimte voor de scheduling problemen die we onderzocht hebben, weten we niet of we in polynomiale tijd een lokaal optimale oplossing kunnen vinden of dat dit misschien PLS-compleet is.

References

[1] P. Brucker, J.L. Hurink, and F. Werner. Improving local search heuristics for some scheduling problems II. *Discrete Applied Mathematics*, 72:47–69, 1997.

[2] T. Brueggemann, J.L. Hurink, and W. Kern. Quality of move-optimal schedules for minimizing total weighted completion time. *Operations Research Letters*, 34(5):583–590, 2006.

[3] T. Brueggemann, J.L. Hurink, T. Vredeveld, and G.J. Woeginger. Very large-scale neighborhoods with performance guarantees for minimizing makespan on parallel machines. In C. Kaklamins and M. Skutella, editors, *Approximation and Online Algorithms* (WAOA 2007), volume 2909 of LNCS, pages 41–55. Springer, Berlin, 2008.

[4] T. Brunsch, H. Röglin, C. Rutten, and T. Vredeveld. Smoothed performance guarantees for local search. In *Proceedings of the 19th European Symposium on Algorithms (ESA 2011)*, 2011. To appear.

[5] G. Finn and E. Horowitz. A linear time approximation algorithm for multiprocessor scheduling. *BIT*, 19:312–320, 1979.

[6] R.L. Graham, E.L. Lawler, J. K. Lenstra, and A.H.G. Rinnooy Kan. Optimization and approximation in deterministic sequencing and scheduling: a survey. *Annals of Discrete Mathematics*, 5:287–326, 1979.

[7] C.A.J. Hurkens and T. Vredeveld. Local search for multiprocessor scheduling: How many moves does it take to a local optimum? *Operations Research Letters*, 31:137–141, 2003.

[8] M.R. Korupolu, C.G. Plaxton, and R. Rajaraman. Analysis of a local search heuristic for facility location problems. In *Proceedings of 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1–10, 1998.

[9] P. Schuurman and T. Vredeveld. Performance guarantees of local search for multiprocessor scheduling. In K. Aardal and B. Gerards, editors, *Proceedings of 8th Integer Programming and Combinatorial Optimization Conference*, volume 2081 of *Lecture Notes in Computer Science*, pages 370–382. Springer, Berlin, Germany, 2001.

[10] P. Schuurman and T. Vredeveld. Performance guarantees of local search for multiprocessor scheduling. *Informs Journal on Computing*, 19(1):52–63, 2007.

[11] T. Vredeveld and J.K. Lenstra. On local search for the generalized graph coloring problem. *Operations Research Letters*, 31:28–34, 2003.

Vakgenoten

EEN CURSIEF EINDPUNT

Het is een nostalgisch genoegen om een bijdrage te leveren aan het Liber Amicorum van Jan Karel Lenstra. Onze eerste gezamenlijke publicaties waren voornamelijk bestemd voor vergelijkbare feestbundels, gekoppeld aan hoogtepunten als afstuderen, huwelijk en promotie – ik herinner mij vooral 'Een pak met korte broek', een titel die verwees naar de feestkleding ooit toegedacht aan zijn briljante jongere broertje Hendrik. En nu, zo'n dertig jaar later, beginnen we aan hoogtepunten als afscheid en pensioen.

Het kan verkeren. Jan Karel en ik hebben ooit een generiek en universeel bruikbaar In Memoriam opgesteld voor de prototypische nukkige geleerde ('Voor wie zich de niet geringe moeite getroostte om hem beter te leren kennen werd hij een goede vriend, al bleef menigeen zich verkijken op zijn af en toe wel heel aparte gevoel voor humor. Maar waar hij zulke hoge, vaak onhaalbaar hoge eisen stelde aan anderen, maakte hij het ook zichzelf niet gemakkelijk.') . Ik zie, met enige zorg, de toepasbaarheid daarvan aan de horizon opdoemen.

Een Liber Amicorum biedt actieve onderzoekers de mogelijkheid om een enkel niet eerder gepubliceerd pareltje alsnog een onderkomen te gunnen in de literatuur. Gene Lawler, Jan Karel en ik schreven in dat kader ooit het artikel 'An algorithm to find this paper', niet een van onze meest geciteerde publicaties. Maar helaas: mijn publicaties van de laatste tien jaar, voorzover al pareltjes, spelen zich af op terreinen waar Jan Karel zich in volle beleefdheid van distantieert, en echt plezier zou ik hem er al helemaal niet mee doen.

Laat ik dan de gelegenheid maar benutten om nog eens te getuigen van mijn bewondering voor wat hij in de loop van de afgelopen zestig jaar heeft bijgedragen aan zijn (en ooit ons) vakgebied. Nauwgezette toewijding en inzet, onuitputtelijke energie, hoge standaarden ('hoge eisen…'), het zijn evenzovele kwalificaties die daarvan iets vastleggen en die een in de tijd constante factor vormen. Als ik stuit op een van de schaarse foto's die onze gezamenlijke momenten in de jaren zeventig vastleggen, dan is hij onmiddellijk te vinden, meestal onverstoorbaar een pijpje rokend. Het ontbreken van dat laatste is het enige wezenlijk waarneembare verschil met recenter beeldmateriaal.

Jan Karel heeft mij opgescheept met een levenslange norm van perfectionisme in taal en typografie. Eerste artikelversies van mijn hand waren gedoemd op ongeveer elk woord door hem gecorrigeerd te worden, en

vanaf de pre-word processor periode was hij de enige bij wie het typewerk in goede handen geacht werd te zijn. Zijn fenomenale pietluttigheid ter zake leverde uiteraard onberispelijke eindresultaten op; de enkele keer dat ik noodgedwongen op eigen benen stond (bij mijn proefschrift, bijvoorbeeld – onze promotor, Gijs de Leve, had ons het duo-proefschrift sterk ontraden) was het eerste wat ik zag bij openslaan van het allereerste gedrukte exemplaar een hinderlijke en alleszins vermijdbare drukfout.

Hij was zich wel bewust van deze gekte. Zijn grote held was de legendarische eerste directeur van het Mathematisch Centrum, A. van Wijngaarden, die een keer een tekst had afgekeurd omdat de punt aan het eind van een cursief getypte bijzin in een verder rechtop getypte hoofdzin cursief was uitgevoerd. Wie een dergelijke leerschool achter de rug heeft, hoeft zich niet te verontschuldigen voor bij vlagen excentriek gedrag.

En van dat laatste was af en toe zeker sprake. Zijn publieke optredens – uiteraard opgesierd met onberispelijke meer-kleuren ‘transparantjes’ – werden voorafgegaan en gevolgd door een noest doorwerken tijdens de voordrachten van andere sprekers, voornamelijk aan het in minuscuul handschrift bijwerken van een uitgebreide lijst lopende verplichtingen, in het bijzonder die welke voortvloeiden uit zijn onuitputtelijke reeks redacteur- en hoofdredacteur-schappen. Desalniettemin wist hij na afloop de inhoud van deze voordrachten uitstekend samen te vatten en van een kort, bij voorkeur dodelijk (‘hoge eisen…’) commentaar te voorzien. Het vakgebied mag zich gelukkig prijzen een dergelijke Arbiter Elegantiarum in zijn midden te hebben gehad en te houden.

Binnenkort gaat hij met pensioen. Eerder dan ik, dankzij de onverbiddelijke logica van de anciënniteit. Gaat hij met een minuscuul schaartje de geraniums onderhouden? Of mogen we ons verheugen op de uiteindelijke publicatie van het definitieve standwerk over scheduling theorie, waar wij samen met Gene Lawler toe besloten op de Lange Leidsedwarsstraat in Amsterdam, 35 jaar geleden? Ik hoop op het laatste, en zet mij bij voorbaat schrap voor zijn commentaar op mijn eerste versie van het voorwoord.

Alexander Rinnooy Kan

Liber Amicorum for Jan Karel Lenstra

George L. Nemhauser
Georgia Institute of Technology
Atlanta, USA

Jan Karel was my colleague at Georgia Tech for nearly two years, far too short of a time. CWI is to blame for his unfortunate departure from Georgia Tech. I remember Jan Karel coming to my office and telling me that he just got offered his “dream job” and from the twinkle in his eye I know that his time at Tech was over. Of course, our loss was CWI’s gain.

I first met Jan Karel at the Center for Operations Research and Econometrics (CORE) in Leuven, Belgium in 1976. I had heard of the reputation of this young Dutch combinatorial optimizer and invited him to give a seminar at CORE. We quickly became friends and when I was elected as Chair of the Mathematical Programming Society I asked Jan Karel to stay on as Chair of the Executive Committee. He then succeeded me as Chair of the Society. Later Jan Karel joined me as Series co-Editor of the Handbooks of Operations Research and Management Science and as co-Editor-in-Chief of Operations Research Letters.

Jan Karel is the most impeccably organized person I have ever known. His attention to detail is amazing. With his tiny handwriting, he could take more notes on one page of a small pad than I could on the three legal-sized pages. And he could put his information to work. With his exquisite knowledge of people in the field and his excellent taste and judgment, he would choose the best young researchers to invite as Associate Editors. I’m sure this carried over to his leadership position at CWI.

Many of my colleagues and I at Tech had hoped he would become the Chair of our School and I know it would be a far better place now if Jan Karel had been its leader.

I have nothing but great praise for Jan Karel's professional accomplishments, but my favorite thing about him is his wry sense of humor.

I can't imagine Jan Karel retiring. He is far too energetic and his wisdom will be greatly missed. Jan Karel, now that you have more time, please come visit us in Atlanta again and be sure to bring Karen and the kids.

Best Wishes,

George

A randomized variation on a theorem of Lenstra

David B. Shmoys *

Lenstra (1976) states the following theorem (in a rough translation from the original Dutch):

> "At least one half of the references to the article by the traveling salesman quartet of Little, Murty, Sweeney, and Karel contains at least one error.
>
> J.D.C. LITTLE, K.G. MURTY, D.W. SWEENEY, J. KAREL (1963) An algorithm for the traveling salesman problem. *Operations Res. 11*, 972–989."

Lenstra does not elaborate on the proof technique used to obtain this result, but one might presume, given the state of the art when this work was published, not to mention the title of the thesis, that an enumerative method was employed. Enumerative procedures, especially complete enumeration, are often impractically slow methods, and in this setting there is the added complication of the seemingly exponential increase in the number of papers published.

The purpose of this brief note is to provide a significantly faster procedure to derive such a theorem, based on randomization. Suppose that you would like to prove a Lenstra-type theorem for a new publication X. Rather than checking each published article that contains a citation of X to verify its correctness, suppose that one has the ability to identify one such published article uniformly at random from among the universe U of papers that cite X. Let n denote the number of papers in this universe U. We can use such a "black box oracle" as follows: draw k independent samples by using the oracle, and check each of them to determine if X is cited correctly. Then, depending on the specific values of k and n, one can assert, with a particular success probability that is easy to calculate, a lower bound on the fraction of papers that contains an error. In particular, we have relied on a standard calculation, using the so-called Chernoff-Hoeffding bound in order to obtain the following theorem.

Theorem 1 *With probability at least .95, at least one half of the references to the book by the traveling salesman quartet of Lawler, Lenstra, Rinnooy Kan, and Shmoys contains at least one error.*

E.L. LAWLER, J. LENSTRA, A.H.G. RINNOOY KAN, D.B. SHMOYS (eds.) (1985) *The traveling salesman problem.* Wiley, Chichester.

In the spirit of the work of Lenstra, we will not give the details of the proof of this new theorem. However, Google Scholar provides us with the means for generating the random samples, and we have estimated n to be 2450; furthermore, we could rely on relatively few samples (in our case, 100), since the fraction that contains an error appears to be significantly greater than one half. We have estimated each parameter conservatively, so as to ensure the correctness of the theorem.

*shmoys@cs.cornell.edu. School of Operations Research & Information Engineering and Department of Computer Science, Cornell University, Ithaca, NY 14853. A preliminary version of this appeared in the Proceedings of the 60th Annual Meeting of the FoJKL. Research supported in part by NSF grants CCF-0832782 and CCF-1017688.

Many different errors were observed. Not surprisingly, “travelling” is substituted quite often. Of the 2^4 possible subsets of editors, a significant fraction are listed as the supposed editors in at least one citation. Since the sample contained quite a number of pairs and singletons of authors, it would be interesting to see if the full enumeration would also yield the empty set. The claimed publication date is as early as the previous millennium. Each editor has incorrect initials in some citation. And, with the exception of Lenstra, every editor has his family name misspelled at least once (at least in the random sample examined). One reassuring fact is that in no instance selected was one of any of Lenstra’s siblings cited instead.

Many significant open questions remain in this line of research. In particular, will the new traveling salesman quartet fare better, or is their TSP book similarly fated? For this question, there is some preliminary evidence that it will fare similarly, though of the roughly 400 citations found by Google Scholar, many actually cite a preliminary technical report, and hence are not part of the true data set. And more generally, perhaps one can prove that there exist values j, k, and ℓ, such that for each at least j-authored publication on the traveling salesman problem, in which the authors have at least k initials in total, and for which there are at least ℓ distinct citations, such a Lenstra-type theorem holds.

Bibliography

D.L. APPLEGATE, R.E. BIXBY, V. CHVÁTAL, W.J. COOK (2006) *The traveling salesman problem.* Princeton University, Princeton.

E.L. LAWLER, J.K. LENSTRA, A.H.G. RINNOOY KAN, D.B. SHMOYS (eds.) (1985) *The traveling salesman problem.* Wiley, Chichester.

J.K. LENSTRA *Sequencing by enumerative methods.* (1976) Mathematisch Centrum, Amsterdam.

To Jan Karel

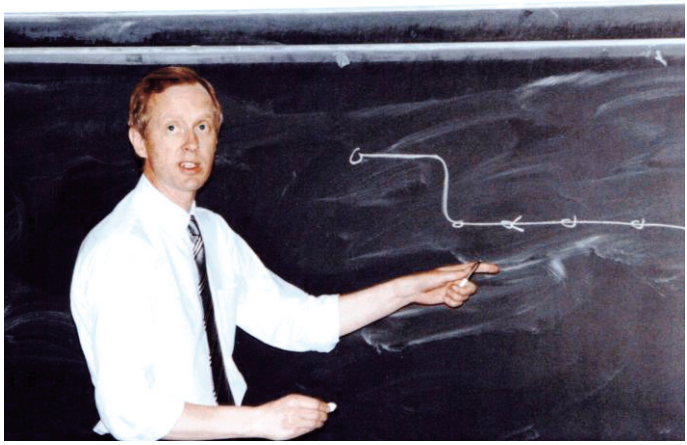
Jan Karel and I have known each other for a long time. The first time we met must have been in Bonn in the early eighties, when I still lived in the Netherlands. He was already famous, while I just got acquainted with scheduling and complexity theory. His work together with Alexander Rinnooy Kan and Gene Lawler fascinated me, and I remember that Franz Josef Radermacher and I convinced Jan Karel to give a course on scheduling at the Technical University of Aachen. I volunteered to be his teaching assistant and thus got first class knowledge about the area.

Jan Karel became my “Dutch connection” when I moved away from Aachen. We worked together for the Mathematical Programming Society (MPS) in the early nineties, jointly participated in PhD committees in Eindhoven and Berlin (Albert Wagelmans, Markus Schäffter, Petra Schuurman, Marc Uetz), set up the MAPSP series of conferences and encouraged mutual visits of members of our groups. I still remember the first MAPSP conference in Villa Vigoni in 1992, where he introduced Karen to the scheduling community. Six year later, at MAPSP IV in Renesse, they proudly presented their daughter Elisabeth.

Two characteristics of Jan Karel that I find most distinctive are his talent for organizational matters and his ability to attract young and talented people to our area. He was the memory of MPS with his small booklet filled with meticulous notes, raised enormous amounts of money for the conferences that he organized and could give wonderful lectures on wonderfully designed handwritten slides that fascinated his audience. He came to Berlin many times as an invited speaker, at the Berlin Algorithm Day, in our Graduate Program, and also as a lecturer at the influential 1996 Summer School on “Approximation Algorithms and Combinatorial Optimization”. Seven participants became a professor in that area: Thomas Erlebach, Arie Koster, Ekkehard Köhler, Jörg Rambau, Matthias Müller-Hannemann, Andreas Schulz, Martin Skutella.

The pictures below are from one such occasion, the Berlin Algorithm Day in 1991. He hasn't changed that much ever after.

Rolf Möhring

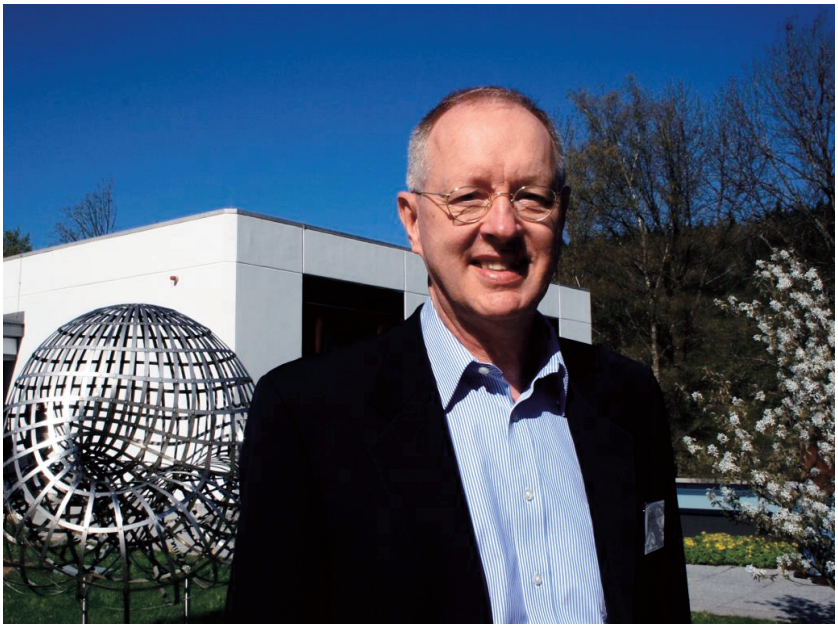


Berlin Algorithm Day (1991)

Berlin Algorithm Day (1991)

Berlin Algorithm Day (1991)

Jan Karel Lenstra



And the Mathematisches Forschungsinstitut Oberwolfach

To Jan Karel at the occasion of his resignation as General Director of CWI

Gert-Martin Greuel
Mathematisches Forschungsinstitut
Oberwolfach
Schwarzwaldstr. 9-11
D-77709 Oberwolfach-Walke

Dear Jan Karel,
This is a small picture book, documenting some of your activities in connection with Oberwolfach. I hope you enjoy the pictures and the documentation of your first participation at a workshop in Oberwolfach, more than 30 years ago.

The Institute

According to our files, a Jan Karel Lenstra participated at the following Oberwolfach workshops:

* 1981: Mathematische Optimierung
* 1986: Combinatorial optimization and its relations to other mathematical areas
* 1988: Mathematische Optimierung
* 1989: Computational Aspects of Combinatorial Optimization
* 1990: Mathematische Optimierung
* 1991: Combinatorial Optimization
* 1992: Mathematische Optimierung
* 1993: Combinatorial Optimization
* 1996: Combinatorial Optimization
* 1997: Mathematische Optimierung
* 1999: Combinatorial Optimization

Moreover you have been invited to further workshops where you did not attend.

As Chair of ERCOM you attended the inauguration of the library extension in Oberwolfach in 2007 and were a scientific member of the evaluation committee of the Leibniz Association, which evaluated Oberwolfach in March 2009.

Your first visit in Oberwolfach was in January 1981, where you participated at the meeting "Mathematische Optimierung" organized by Heinz König, Bernhard Korte, and Klaus Ritter:

Math. Forschungsinstitut
Oberwolfach
E 20/01737

Mathematisches Forschungsinstitut Oberwolfach

Tagungsbericht 6/1981

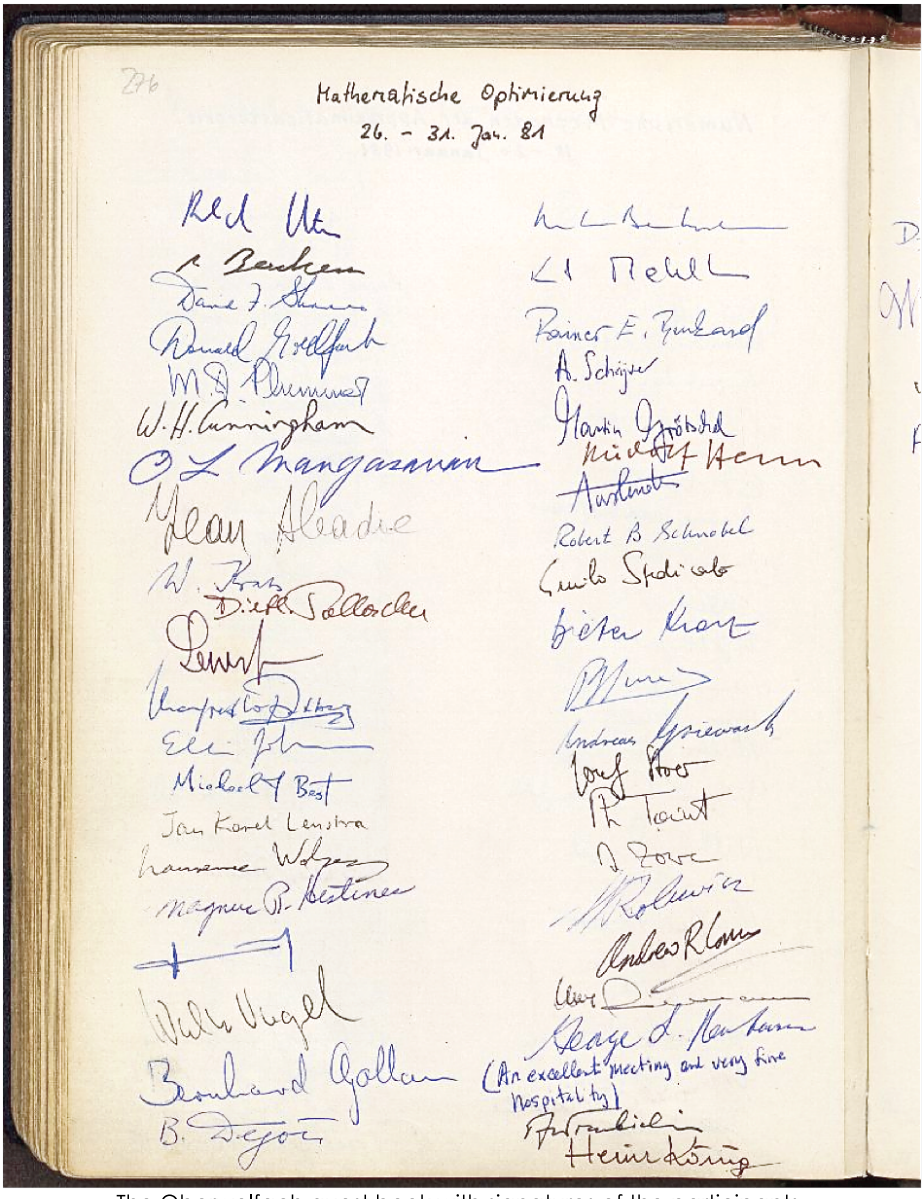
Mathematische Optimierung

25.1. bis 31.1. 1981

Leitung: Heinz König (Saarbrücken)
Bernhard Korte (Bonn)
Klaus Ritter (Stuttgart)

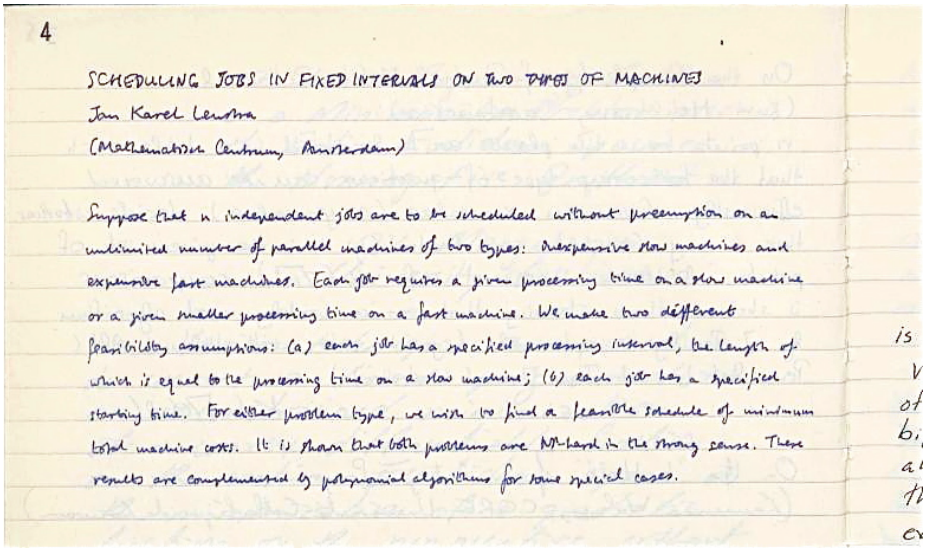
Diese zweite Tagung über Mathematische Optimierung in Oberwolfach hat wiederum einen sehr großen Anklang in der Fachwelt gefunden, was sich in der relativ hohen Zahl von 68 Tagungsteilnehmern aus 11 Ländern wiederspiegelt. 20 Teilnehmer reisten aus den USA und Kanada an.

Die 45 Vorträge der Tagung überdeckten das gesamte Gebiet der Mathematischen Optimierung, wobei hervorzuheben ist, daß die Interaktion zwischen verschiedenen Teilgebieten im Vordergrund stand, was sich besonders befruchtend auf die weitere Entwicklung der Mathematischen Optimierung auswirken wird. Insbesondere wurde über die Beziehung zwischen submodularen Funktionen und der konvexen Analysis, die algorithmische Äquivalenz von Separation und Optimierung, allgemeine Dualitätstheorie, stochastische lineare Programmierung, diskrete Optimierung mit nichtlinearen Funktionen und über die Entwicklung von polynomialen Algorithmen mit Hilfe von Newton- und Strafkosten-Verfahren berichtet. Darüber hinaus wurden neue Ergebnisse der Graphen- und Polyedertheorie, der Charakterisierung von Facetten, der Komplexitätstheorie, Algorithmen für Scheduling- und Sequencing-Probleme, sowie Varianten des Simplex-Verfahrens für spezielle kombinatorische Fragestellungen vorgestellt. Einen weiteren Schwerpunkt bildeten Modifikationen des Newton-Verfahrens für spezielle C_1-Funktionen und in Punkten mit singulären Ableitungen, Line-Search-Verfahren, Optimalitätsbedingungen für nicht-glatte Optimierungsprobleme, quadratische Optimierung, und die Verfeinerung von Lösungsverfahren mittels Parallelprojektion und variabler Metrik.



The Oberwolfach guest book with signatures of the participants

You gave a talk on "Scheduling Jobs in Fixed Intervals on Two Types of Machines". The handwritten abstract and the typed workshop report of this talk are shown below.



4

SCHEDULING JOBS IN FIXED INTERVALS ON TWO TYPES OF MACHINES

Jan Karel Lenstra

(Mathematisch Centrum, Amsterdam)

Suppose that n independent jobs are to be scheduled without preemption on an unlimited number of parallel machines of two types: inexpensive slow machines and expensive fast machines. Each job requires a given processing time on a slow machine or a given smaller processing time on a fast machine. We make two different feasibility assumptions: (a) each job has a specified processing interval, the length of which is equal to the processing time on a slow machine; (b) each job has a specified starting time. For either problem type, we wish to find a feasible schedule of minimum total machine costs. It is shown that both problems are NP-hard in the strong sense. These results are complemented by polynomial algorithms for some special cases.

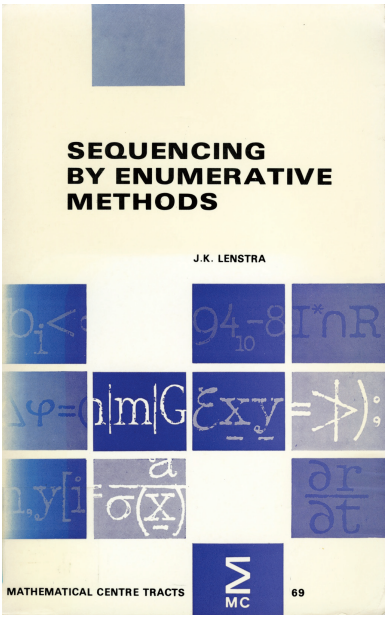
J. K. Lenstra:

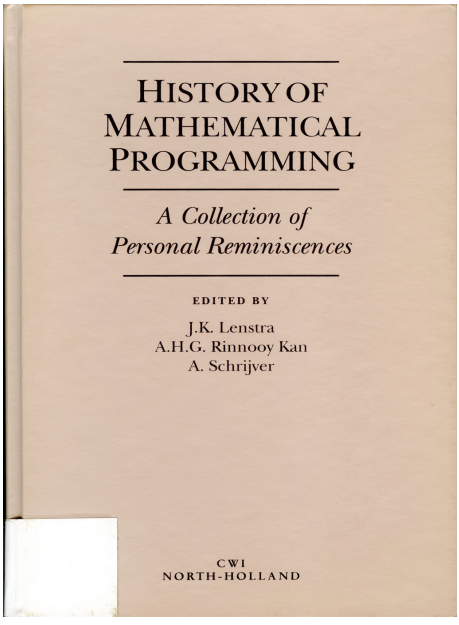
Scheduling Jobs in Fixed Intervals on Two Types of Machines

Suppose that n independent jobs are to be scheduled without preemption on an unlimited number of parallel machines of two types: inexpensive slow machines and expensive fast machines. Each job requires a given processing time on a slow machine or a given smaller processing time on a fast machine. We make two different feasibility assumptions: (a) each job has a specified processing interval, the length of which is equal to the processing time on a slow machine; (b) each job has a specified starting time. For either problem type, we wish to find a feasible schedule of minimum total machine costs. It is shown that both problems are NP-hard in the strong sense. These results are complemented by polynomial algorithms for some special cases.

Library at Oberwolfach

The library of Oberwolfach contains several books by J. K. Lenstra, where you are either an author, a co-author or an editor.





We met many times at ERCOM meetings. Here are two photos from these meetings; the first was taken during the ERCOM meeting at the Mittag-Leffler Institute in May 2009 and the second during the meeting in Oberwolfach in April 2011.

Best Wishes Jan Karel!

Oberwolfach, July 2011

An Exact Dynamic Programming Algorithm for Regression Trees

Jacqueline J. Meulman, Elise Dusseldorp and Bart Jan van Os *

Achtergrond: Mijn belangstelling voor algoritmes voor combinatorische optimalisatie komt op de eerste plaats door mijn jarenlange samenwerking met Lawrence J. Hubert (University of Illinois at Urbana-Champaign) en Phipps Arabie (Rutgers University, Newark). Toen ik in 1994 een NWO-PIONIER subsidie kreeg, werd Bart Jan van Os een van mijn promovendi. Omdat hij een uitzonderlijk talent voor het ontwerpen en implementeren van algoritmes bleek te hebben, richtte het onderwerp van zijn proefschrift zich al gauw op globale optimalisatie d.m.v. dynamisch programmeren. Elise Dusseldorp, ook uit het PIONIER-project, promoveerde op het onderwerp regressiebomen. Omdat ik mijzelf met toepassingen van ensembles van kleine regressiebomen had beziggehouden, kwam het idee om gezamenlijk aan globaal optimale regressiebomen te werken bijna als vanzelf.

Tijdens het schrijven van zijn proefschrift heb ik Bart Jan aangeraden om een kursus bij Jan Karel Lenstra, toen nog in Eindhoven, te volgen. Jan Karel vertelde mij bij een latere gelegenheid dat hij nogal verbaasd was geweest om een promovendus uit de Faculteit Sociale Wetenschappen te Leiden in zijn collegezaal te zien. Nu vormden wij inderdaad een bijzondere groep binnen de FSW: de Facultaire vakgroep Datatheorie (het geesteskind van John P. van de Geer, de vader van Sara). Nadat een verregaande integratie van de Datatheoriegroep binnen een FSW departement mislukt was, kwam ik zelf tot mijn grote vreugde bij het Mathematisch Instituut terecht. Het artikel over optimale regressiebomen was echter nooit afgekomen. Na de uitnodiging om een bijdrage te leveren aan dit Liber Amicorum, kreeg ik het idee om het artikel ter ere van Jan Karel af te maken. Alleen bleek het toen als bijdrage aan het Liber veel te lang te zijn geworden. Daarom hier nu een sterk ingekorte impressie van ons werk, dat hopelijk te zijner tijd elders in volledige vorm gepubliceerd zal worden (maar in ieder geval al te lezen is op www.math.leidenuniv.nl/∼jmeulman). Wij danken hierbij Jan Karel voor zijn inspiratie.

*Jacqueline Meulman is Professor of Applied Statistics, Department of Mathematics, Leiden University, The Netherlands (e-mail: jmeulman@math.leidenuniv.nl), Elise Dusseldorp is Statistician, TNO Quality of Life Research, Leiden, The Netherlands (e-mail: elise.dusseldorp@tno.nl) and Bart Jan van Os is Statistician, 4os Consultancy, The Netherlands (e-mail: advies@4os.nl). This paper is based on work first presented in the PhD thesis Dynamic Programming for partitioning in multivariate data analysis by Bart Jan van Os, Leiden 2001, under supervision of Jacqueline J. Meulman and Lawrence J. Hubert.

1 Introduction

Although trees were introduced for the first time by Morgan and Sonquist (1963) in a method called Automatic Interaction Detection, the methodology was only fully developed in Breiman, Friedman, Olshen, and Stone (1984). Their CART algorithm complemented regression trees for the prediction of a continuous outcome variable, with classification trees that deal with a categorical (nominal) outcome variable. Since then a number of different methods and software packages have become available (for an overview, see Hand (1997)), especially after trees obtained renewed interest by research in data mining methods. The methodology has become popular for several reasons: trees automatically select predictors and find complicated interaction effects in the presence of noise variables, they deal with missing data without having to resort to imputation methods, and they can handle a large number of predictors of different measurement level (numeric, ordinal, binary and categorical variables are all handled with equal ease). Trees are immune to the effects of extreme outliers among the predictor variables, and are invariant to the scaling of predictors or any monotone transformation of them. Finally, a fitted tree mimics the way a hierarchically nested set of decision rules can be represented, and therefore has a clear interpretation.

Despite their success, trees are also known to have a major drawback: their limited prediction accuracy in applications, due to their instability. Instability means that a small change in the data can result in a very different tree, and the latter makes the interpretability unreliable. Because of the hierarchical nature of the process, the effect of an error in the top split is propagated down to all of the splits below it, and this constitutes an inherent instability (Hastie, Tibshirani, and Friedman (2001), p. 274). Therefore various approaches have tried to extend greedy algorithms, with successful advancements obtained by combining multiple tree structures in ensembles, obtaining a stronger predictor or classifier. Examples are Boosting (Freund & Schapire, 1996; Drucker & Cortes, 1996; Friedman, 2001), Bagging (Breiman, 1996), Random Forests (Breiman, 2001) and Rule Ensembles (Friedman & Popescu, 2008), which are competitive with other predictive learning methods such as Support Vector Machines (Vapnik, 1996), Kernel methods and local regression (see Loader, 1999), and Neural Networks (see Ripley, 1996).

Unfortunately, the result of an ensemble method is deprived of one of the most important advantages of a single tree, i.e., its high interpretability. At the expense of higher prediction accuracy, we lose the representation of the prediction rules in a simple tree structure. Because prediction using ensemble methods works more or less as a black box, considerable efforts have been invested to obtain secondary tools for interpretation, such as the importance of a variable by looking at partial dependence (for example, Friedman 2001). However, with respect to interpretability, the overview by looking at many partial dependence plots cannot compete with looking at a single tree. Conse-

quently, adapting greedy tree methods remains an active domain of research, also because of their ability to handle large datasets that arise in data mining. Research in this area includes the introduction of more complex splits that combine several predictors through a discriminant analysis in each node (Loh & Vanichesetakul, 1988), finding more advanced split rules to overcome variable selection bias (Loh, 2002), and combining the tree model with standard regression models (Alexander & Grimshaw, 1996; Dusseldorp & Meulman, 2001, 2004; Dusseldorp, Conversano, & Van Os, 2010).

The present paper aims at strengthening tree methods at their heart, i.e., at the way trees are constructed (fitted). A major way to remove the inherent instability, and still to obtain a single tree representation, is to use an algorithm that finds a *globally optimal* tree. However, from the early days on, simultaneous optimization over all possible splits in the complete tree structure has been considered computationally infeasible. Nowadays, the problem is believed to be most likely intractable, because parts of the problem and the related problem of constructing minimum size regression trees from decision tables are proven to be NP-complete (Hyafil & Rivest, 1976; Naumov, 1994). The effort of finding a globally optimal tree immediately raises a crucial question: is optimizing the performance of prediction rules on the training data actually worth the effort, or does it induce over-fitting so that the expected prediction error (obtained for the test data) actually increases? We believe that the latter is not the case, because the algorithm aims at improving the accuracy without increasing the complexity, or possibly even decreases the complexity.

2 Regression Trees

In regression problems, one has a system consisting of a random "outcome", "response", or "dependent" variable Y and a set of random "explanatory" ,"predictor", or "independent" variables $X = \{X_1, \cdots, X_J\}$, where J denotes the number of predictors. The problem defines a "training" sample, $\{y_i, \mathbf{x}_i\}_1^N$ of known values for Y and $\mathbf{X}$, where $(y_i, \mathbf{x}_i)$ links the predictor variables of the ith object with the ith value of the outcome variable, and where $i = 1, \ldots, N$. The space spanned by the predictor variables, denoted by $\mathbf{X}$, is called the *measurement space*. The regression tree problem consists of finding a set of rules that partition the measurement space into regions such that within each region R_k we are able to predict a measurement $\hat{y}_i$ as close as possible to y_i for all objects that are included into that region (for an example, see Hastie et al. (2001), p. 268). The measurement space representation has the advantage that it provides a concise description of the regression tree problem as a partitioning task, and immediately suggests relations to other problems that were solved globally, such as hierarchical clustering (Hubert, Arabie, & Meulman, 2001; Van Os, 2001; Van Os & Meulman, submitted). Consequently, the measurement space representation will be used to present the global optimization approach and the proposed algorithm.

2.1 Formal Definition of the Optimal Tree Problem

In the sequel we will show how regression trees can be formulated as a combinatorial optimization problem. Without loss of generality we will restrict ourselves to binary trees.

Each tree results in a partition $\mathcal{P}$ of the measurement space $\mathbf{X}$ into a number of $|\mathcal{P}|$ regions $R_k \in \mathcal{P}$ that correspond to the induced terminal nodes. Let $\Omega_d(\mathbf{X})$ denote the set of all admissible partitions of the measurement space $\mathbf{X}$ with binary trees of at most K regions (terminal nodes) and depth d, where d denotes the maximum number of layers from the top to any terminal node in the tree. Note that the definition of $\Omega_d(\mathbf{X})$ may include additional admissibility restrictions that are widely applied in tree methods, such as the required minimum number of observations in a terminal node, or a minimum impurity criterion that stops further splitting if the value of the loss function (that measures the discrepancy between the training data and the representation) is close (or equal) to zero.

We now can formulate regression trees as special cases of the mathematical problem:

$$\begin{cases} f(\mathbf{y}, \mathbf{X}) = \underset{\mathcal{P}}{\text{opt}} \sum\limits_{R_k \in \mathcal{P}} h(\mathbf{y}, \mathbf{X}, R_k) \\ \text{subject to} \\ \mathcal{P} \in \Omega_d(\mathbf{X}), \\ d \leq D, \\ |\mathcal{P}| \leq K. \end{cases} \tag{1}$$

The *impurity* function $h(\cdot)$ can be defined in many ways without violating the Dynamic Programming principles explained below. For regression trees that minimize the sum of squared error by estimating a constant a_k in each region R_k, $h(\cdot)$ is defined by

$$h(\mathbf{y}, \mathbf{X}, R_k) = \sum_i (y_i - a_k)^2 I(\mathbf{x}_i \in R_k), \tag{2}$$

but $h(\cdot)$ also includes other loss functions such as absolute deviations

$$h(\mathbf{y}, \mathbf{X}, R_k) = \sum_i |y_i - a_k| I(\mathbf{x}_i \in R_k), \tag{3}$$

or even so-called Treed Regression (Alexander & Grimshaw, 1996) that estimates a *simple* linear regression model in each region by defining impurity as

$$h(\mathbf{y}, \mathbf{X}, R_k) = \sum_i (y_i - (a_k + \mathbf{b}_k' \mathbf{x}))^2 I(\mathbf{x}_i \in R_k), \tag{4}$$

where $\mathbf{b}_k$ only has one non-zero element corresponding to the best linear predictor in that region, excluding predictors that are used for splits up to the particular terminal node. Note that in all these cases the required additional parameters can be estimated *independently* from other regions.

The mathematical problem formulation also explicitly includes two types of size constraints. The first type of constraint ($d \leq D$) imposes the maximum depth D on the tree, and corresponds to the maximum order of interaction between predictors (Friedman, 2001). Let an *np*-split (a new predictor split) be a split that is defined on a predictor that has not been used in previous splits, if existing. The maximum order of interaction is then equal to the largest number of np-splits from the top to any terminal node. The second type of constraint ($|\mathcal{P}| \leq K$) imposes a maximum K on the number of terminal nodes ($|\mathcal{P}|$ in the tree. These two restrictions are obviously related by $K \leq 2^D$, and only one of the two may be in effect for a given problem. Both constraints define the admissible *maximum* size rather than the required maximum size, because additional admissibility conditions, such as the minimum number of observations in a terminal node, may actually prevent larger trees in some cases. Because in general the best prediction accuracy obtained by the least complexity of the tree is preferred, the smallest tree that optimizes (1) under the size constraints is accepted.

Note that in our approach these size constraints do not necessarily have to be chosen a priori. The mathematical problem (1) is solved for a series of values for K and D, and fortunately, the algorithm introduced below will provide such a series of solutions with relatively little additional effort.

3 The Dynamic Programming recursion

To formulate our Dynamic Programming algorithm some more definitions are needed that take into account the hierarchical nesting of the splits. Let v be any node (including the root node) of a regression tree $T(R)$ of the complete measurement space R, and let R' and $T(R')$ denote the region ($R' \subseteq R$) and the (sub)tree associated with v; let $l(R', s, j)$ and $r(R', s, j)$ be a pair of left and right functions that split the region R' into a left and right measurement space using split $s \in \mathcal{S}_j(R')$, where $\mathcal{S}_j(R')$ denotes the set of admissible splits of R' by predictor X_j. Furthermore, let $h(R')$ be a shorthand notation for $h(\mathbf{y}, \mathbf{X}, R')$, and let $t(R')$ denote the value of the objective function for obtaining a (sub)tree $T(R')$ for region R', such that

$$t(R') = \sum_{R_k \in \mathcal{P}(R')} h(R_k). \tag{5}$$

Finally, let $T^*(R')$ be an *optimal* (sub)tree and $t_d(R')$ the associated (optimal) value of the objective function. We have the following *principle of optimality* (Bellman, 1957):

Proposition 1 *Any subtree $T(R')$ starting at some node v of an optimal tree $T^*(R)$ is itself an optimal tree of the measurement space R' associated with its root node v.*

PROOF. The objective function can be split into the sum of two parts

$$t^*(R) = \sum_{R_k \in \mathcal{P}(R)} h(R_k) = \sum_{R_{k_1} \in \mathcal{P}(R')} h(R_{k_1}) + \sum_{R_{k_2} \in \mathcal{P}(R-R')} h(R_{k_2}), \qquad (6)$$

where the first part contains the regions that belong to the subtree $T(R')$, and the other part contains all other regions. Hence, we can always replace any subtree of an optimal tree with a better subtree if that would exist, but this would contradict the optimality of the original tree T.
This particular splitting of the objective function is often referred to as the *separability* of the objective function. In this case, it stems from the fact that we can evaluate the part of the objective function for the subtree independently from the way in which the tree partitions the rest of the measurement space.

This principle quite naturally leads to a recursion for optimal regression trees. In fact, the transition from a recursion to a tree and vice versa is so transparent that trees are often used to explain DP algorithms. Moreover, these type of recursions for trees are well-known in computer science, where the classical example is the use of DP for the minimal size "Optimal Binary Search Tree", a problem that is not NP-hard, and were DP is very efficient. In contrast, DP for NP-hard problems is often not considered because of the explosion of computer resources needed for increasing problem sizes, with the major exception of Hubert et al. (2001), Hubert, Arabie, and Meulman (2006), who present DP solutions (and software) for finding exact solutions to a number of combinatorial data analysis problems.

In formulating the recursion, however, we have to take into account the size restrictions of the mathematical problem (1) that are implicit in this principle of optimality. In other words, the principle of optimality only holds if the size constraints for the subtree conform to the overall size constraints of the complete tree. To define practical recursions that can readily be converted into algorithms, it is more convenient to explicitly include the size constraints. For maximum-depth trees, where only the depth-constraint is imposed, we have the following recursion:

$$\begin{cases} t_d^*(R') = & \underset{s,j}{\text{opt}} \left[t_{d-1}^*(l(R', s, j)) + t_{d-1}^*(r(R', s, j) \right] \\ & \text{subject to } s \in \mathcal{S}_j(R'); j = 1, \ldots, J; \\ t_d^*(R') = & h(R') \\ & \text{if } d = 1 \text{ or} \\ & \text{other admissibility restrictions prevent further splitting of } R'. \end{cases} \qquad (7)$$

In words, the process of finding an optimal tree can be described by: i. search through all possible splits at the current node; ii. for each split, find independently an optimal subtree at the left and the right node; iii. a. if a node is allowed to be split further, the search process is recursively started for this new node; iv. b. if a node is an end node, we simply use the objective function value

for this node; v. the objective function value for the current split at the current node is the sum of the objective function values of the optimal subtrees; vi the split that results in the best combination of subtrees gives the optimal (sub)tree at the current node and is returned; vii if the current node is the root node, we have found the complete optimal tree; otherwise we return to the previous level in the search.

When we also impose the overall size constraint $|\mathcal{P}| \leq K$, the recursion becomes more elaborate, because we generally do not know in advance the size of the left and the right subtree for any split:

$$\begin{cases} t^*_{dk}(R') = & \underset{s,j,k_l}{\text{opt}} \left[t^*_{d-1,k_l}(l(R',s,j)) + t^*_{d-1,k-k_l}(r(R',s,j) \right] \\ & \text{subject to } s \in \mathcal{S}_j(R'); j = 1,\ldots,J; k_l = 1,\ldots,k-1; \\ t^*_{dk}(R') = & h(R') \\ & \text{if } d = 1 \text{ or } k = 1 \text{ or} \\ & \text{other admissibility restrictions prevent further splitting of } R'. \end{cases} \tag{8}$$

The search process is now extended to not only include an optimal choice of the split s and the predictor j, but also the optimal split of the size constraint k at the node into two size constraints k_l and $k_r, k = k_l + k_r$ for the left and right subtrees. Consequently, imposing the size constraint $|\mathcal{P}| \leq K$ additionally to imposing the depth constraint increases the workload of completely searching the recursion. For an algorithm that finds optimal trees for a range of size constraints, however, the recursion can be reformulated such that the number of possible nodes visited for finding a range of depth-constrained optimal trees is of the same order of magnitude as for finding a range of both depth and size constrained trees.

4 Algorithm Implementation

Given the recursions (7) and (8), we can immediately formulate recursive algorithms. For example, Algorithm 4.1 gives a DP algorithm for finding an optimal maximum-depth tree. This algorithm explicitly uses a heap to store and retrieve previously searched nodes, and only gives the (optimal) objective function value of an optimal tree. To retrieve the optimal tree itself, the tuple $(R', (s, j))$ that corresponds to the optimal split must be stored in the heap as well. When the algorithm is finished, a simple recursive routine can restore the optimal tree by retrieving all optimal $(R', (s, j))$ from the heap, starting at the root node for R', applying the split (s, j) giving $R'^{(l)}$ and $R'^{(r)}$, retrieving the optimal splits for those measurement spaces, and so forth.

The current algorithm constrains the tree size by maximum depth. For finding optimal trees with a maximum number of terminal nodes K, we have to extend the algorithm. Taken the above recursion, a straightforward implementation is to induce an additional loop that circles over all possible sizes of

Algorithm 4.1 $h \leftarrow$OptimalTree(R', d)
Find the minimal objective function value h of the best tree, springing from a node defined by region R', with depth d (RECURSIVE)

```
1: if ∃(R', h) then
2:    retrieve (R', h)
3: else if (d = 1) ∨ (S_j(R') = ∅, ∀j) ∨ (R' is homogeneous enough) then
4:    h ← h(R')
5: else
6:    h ← inf
7:    for j = 1 to J do
8:       for s ∈ S_j(R') do
9:          h^(l) ← OptimalTree(l(R', s, j), d − 1)
10:         h^(r) ← OptimalTree(r(R', s, j), d − 1)
11:         h ← min(h, h^(l) + h^(r))
12:      end for
13:   end for
14:   store (R', h)
15: end if
```

Ensure: $\exists h$

(sub)trees before stepping into a lower level. However, a more efficient approach is given by Algorithm 4.2 that extends the search by simultaneously searching at each node for all optimal trees of size one up to $K-1$ rather than *the* optimal tree. To do so, we have to search for all optimal combinations of subtrees such that $2 \leq |P(R'^{(l)})| + |P(R'^{(r)})| \leq K$, and store the optimal combinations for all sizes $1, \ldots, K$. (The 'optimal combination' of size one is no split). The practical details as well as the limits for using a (partial heap) are discussed in the full paper. Effective implementations have so far mostly not implemented a heap, but instead simply fully exploit the recursive algorithm.

5 Discussion

The present paper proposes an exact Dynamic Programming algorithm for finding optimal regression trees, where optimality is defined as the quality of the predictions at the leaves. DP is often dismissed a priori as a candidate for such hard problems, because of the enormous computational resources involved. Practical limitations are indeed imminent, but due to the particular implementation, our algorithm can find (guaranteed) optimal trees for sample sizes of about 500 observations and 15 predictors, with a maximum size of the tree, the *depth*, of seven. Note that these numbers applied at the time the current algorithm was implemented, which was some time ago, and a depth of seven is usually more than sufficient. So ExactTree can find sufficiently large optimal trees for moderately sized datasets, and thus it can deal with nontrivial statis-

Algorithm 4.2 $H \leftarrow$OptimalSCTrees(R', d, K)
Find the minimal objective function values $H = \{h_1, \ldots, h_K\}$ of all Size Constrained trees with number of terminal nodes not exceeding $2, \ldots, K$, springing from a node defined by region R', with depth d (RECURSIVE)

if $\exists(R', H)$ **then**
  retrieve (R', H)
else if $(d = 1) \vee (\mathcal{S}_j(R') = \emptyset, \forall j) \vee (R'$ is homogeneous enough$)$ **then**
  $h_k \leftarrow h(R'), k = 1, \ldots, K$
else
  $H \leftarrow \inf$
  $h_1 \leftarrow h(R')$
  for $j = 1$ to J **do**
    for $s \in \mathcal{S}_j(R')$ **do**
      $H^{(l)} \leftarrow$ OptimalSCTrees$(l(R', s, j), d-1, K-1)$
      $H^{(r)} \leftarrow$ OptimalSCTrees$(r(R', s, j), d-1, K-1)$
      $h_k \leftarrow \min(h_k, h_{k'}^{(l)} + h_{k-k'}^{(r)}), k = 2, \ldots, K,$ and $k' = 1, \ldots, k$
    end for
  end for
  $h_k \leftarrow \min(h_{k'}), k = 2, \ldots, K,$ and $k' = 1, \ldots, k$
  store (R', H)
end if
Ensure: $\exists H$

tical applications in a lot of domains. Moreover, ExactTree can do this while using a plethora of optimality criteria that are beyond the reach of alternative strategies such as linear programming and branch and bound; it uses a simple interface and does not require special expertise of the user (a Matlab version of the software is available from the third Author).

Our applications so far have shown that using the ExactTree algorithm indeed improves upon the prediction of greedy methods; compared to the greedy algorithm, ExactTree both finds smaller trees with similar prediction error, and larger trees with smaller prediction error (while having a smaller standard error of the prediction error). These results might indicate that using an exact algorithm within the bounds of size constraints on the tree structure, indeed leads to a more stable and accurate estimation of the tree structure without over-fitting on the training data. Moreover, restricting the search space of splits for each variable by optimally digitizing the predictors into a moderate number of categories appears to improve the stability of the prediction (and this seems to be also true for greedy methods). Further study to confirm these results is obviously needed.

ExactTree cannot deal with the size of datasets that are common in large-scale data mining, nor can it be used to replace automatic variable selection in the presence of a large number of variables. In the latter case, however, it is

worthwhile to use ExactTree to estimate an optimal tree for given predictors that were selected by another method. Another application of ExactTree is as a benchmark tool, for truly evaluating the performance of more elaborate heuristic approaches as compared to greedy methods. Finally, an obvious application of ExactTree would be as a part of a so-called Boosting Algorithm, such as MART, where the ensemble is made up by many *small* trees, to obtain a really strong predictor.

References

Alexander, W. P., & Grimshaw, S. D. (1996). Treed regression. *Journal of Computational and Graphical Statistics.*, *5*, 156–175.

Bellman, R. E. (1957). *Dynamic programming.* Princeton, NJ: Princeton University Press.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *26*, 123–140.

Breiman, L. (2001). Random forrest. *Machine Learning*, *45*, 5–32.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees.* Belmont, CA: Wadsworth.

Drucker, H., & Cortes, C. (1996). Boosting decision trees. In D. Touretzky, M. Mozer, & M. Hasselmo (Eds.), *Advances in neural information processing systems 8* (pp. 479–485). MIT Press.

Dusseldorp, E., Conversano, C., & Van Os, B. J. (2010). Combining an additive and tree-based regression model simultaneously: Stima. *Journal of Computational and Graphical and Statistics*, *19*, 514–530.

Dusseldorp, E., & Meulman, J. J. (2001). Prediction in medicine by integrating regression trees into regression analysis with optimal scaling. *Methods of Information in Medicine*, *5*, 403–409.

Dusseldorp, E., & Meulman, J. J. (2004). The regression trunk approach to discover treatment covariate interaction. *Psychometrika*, *69*, 355–374.

Freund, Y., & Schapire, R. (1996). Experiments with a new boosting algorithm. In *Proc. 13th international conference on machine learning* (pp. 148–156). Morgan Kaufman.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, *29*, 1189–1232.

Friedman, J. H., & Popescu, B. E. (2008). Predictive learning via rule ensembles. *Annals of Applied Statistics*, *2*, 916–954.

Hand, D. J. (1997). *Construction and assessment of classification rules.* Chichester: Wiley.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning: Data mining, inference, and prediction.* New York: Springer.

Hubert, L. J., Arabie, P., & Meulman, J. J. (2001). *Combinatorial data analysis: Optimization by dynamic programming.* Philadelphia: SIAM.

Hubert, L. J., Arabie, P., & Meulman, J. J. (2006). *The structural representation of proximity matrices with matlab.* Philadelphia: ASA-SIAM Series on Statistics and Applied Probability.

Hyafil, L., & Rivest, R. L. (1976). Constructing optimal binary decision trees is np-complete. *Information Processing Letters*, *5*, 15–17.

Loader, C. (1999). *Local regression and likelyhood.* New York: Springer-Verlag.

Loh, W.-Y. (2002). Regression trees with and unbiased variable selection and interaction detection. *Statistica Sinica*, *12*, 361–386.

Loh, W.-Y., & Vanichesetakul, N. (1988). Tree structured classification via generalized discriminant analysis. *Journal of the American Statistical Association*, *83*, 715–725.

Morgan, J. N., & Sonquist, J. A. (1963). Problems in the analysis of survey data,and a proposal. *Journal of the American Statistical Association.*, *58*, 415–434.

Naumov, G. E. (1994). Np-completeness of problems of construction of optimal decision trees. *Soviet Physics, Doklady*, *36*, 270–271.

Ripley, B. D. (1996). *Pattern recognition and neural networks.* Cambridge: Cambridge University Press.

Van Os, B. J. (2001). *Dynamic programming for partitioning in multivariate data analysis.* Unpublished doctoral dissertation, Leiden University.

Van Os, B. J., & Meulman, J. J. (submitted). *Dynamic programming for globally optimal hierarchical cluster analysis.*

Vapnik, V. (1996). *The nature of statistical learning theory.* New York: Springer-Verlag.

Bestuurscollega's

HET AFSCHEID

Ik voel mij heden diep bedroefd.
Ik laat geen drank meer staan.
’t Is Jan Karel die mij zo beproeft
door van ons heen te gaan.

Ruim 25 jaar gelee
Mocht ik zijn promovendus zijn.
Met teksten niet zo snel tevree
verscheen dan “lelik” in de kantlijn.

Eindhoven, Utrecht en Amsterdam,
de paden kruisten vaak.
En als het op muziek aankwam
hadden we dezelfde smaak.

Informatica en wiskunde,
of P gelijk is aan NP,
de directeur die CWI runde,
gelijk JK zijn er geen twee.

Een somb’re kilte drukt mij neer.
Een leegte grijnst mij aan.
Och, zie ik Jan Karel ooit nog weer?
Ik pers – er komt een traan.

Peter van Laarhoven

Elke seconde telt

Nederland heeft de ambitie om wereldwijd structureel tot de top vijf van kenniseconomieën te behoren. Innovatie- en concurrentievermogen zijn in een mondialiserende economie daarvoor bepalend. Informatie- en communicatietechnologie (ICT) is niet alleen de belangrijkste aanjager van innovatie, de toepassing van ICT verbetert ons concurrentievermogen en ook de kwaliteit van onze samenleving.

Kijkend met een vluchtige blik, lijkt het erop dat de Nederlandse bijdrage aan belangrijke ICT-innovaties met wereldwijde impact slechts beperkt is. Dat beeld klopt niet. De helaas overleden Martin Rem heeft in zijn boek "Tegen de stroom in" op boeiende wijze laten zien dat wij op een aantal zeer belangrijke terreinen een cruciale rol hebben vervuld. Nederlandse wetenschappers stonden aan de wieg van fundamentele doorbraken. Bluetooth maakt bellen in de auto mogelijk, Wifi is een meer dan welkome aanvulling op de netwerken voor mobiele datacommunicatie. Beide technologieën vinden hun oorsprong in Nederland. Het zijn voorbeelden, en zeker niet de enige, van innovaties waar wereldwijd van wordt geprofiteerd. Al kort na de oorlog werd op de voorloper van het Centrum voor Wiskunde en Informatica, een van de eerste computers gebouwd. Wie weet er dat de programmeertaal Python, die daaruit is voortgekomen, noodzakelijk is om Google te laten draaien? We zouden in Nederland eigenlijk best wel wat trotser mogen zijn op de bijzondere ICT-prestaties die Nederland heeft geleverd. Niet alleen in het verleden, maar ook vandaag de dag.

ICT wordt inmiddels alom gezien als Innovatie As in alle sectoren van de economie. Het is een belangrijke enabler voor een excellente kenniseconomie. Eén geslaagde ICT-technologie of één geslaagd ICT-systeem is drager van innovatie in vele sectoren. Voorbeelden hiervan zijn cloud computing, (open)data en -platformen en sensornetwerken. Deze verbindende kracht die ICT biedt, stelt sectoren in staat te interacteren waardoor de waarde die gecreëerd wordt binnen één sector ook door andere sectoren kan worden benut. Daarnaast heeft ICT ook een sterk transformerende kracht die sectoren in staat stelt zich beter te organiseren. Om de potentie van ICT volledig te kunnen benutten is het daarom noodzakelijk voor iedere sector invulling te geven aan de rol van ICT als Innovatie As. Dit wordt ook zo erkend door het kabinet. De valorisatie van wetenschappelijke kennis staat hoog op de beleidsagenda.

Valorisatie is in feite het verzilveren van wetenschappelijke kennis en techniek. Valorisatie is dan ook onlosmakelijk verbonden met innovatie. Wetenschappelijke kennis wordt waardevol wanneer deze ook buiten het wetenschappelijke circuit gebruikt wordt en resulteert in nieuwe

verrassende combinaties. Het gaat immers om de kennis, en om hetgeen je er uiteindelijk mee kunt doen. Juist in de ICT luistert dit nauw. In het verleden duurde de ontwikkeling van nieuwe technologieën en producten vaak jaren. Door de toegenomen concurrentie en de opkomst van nieuwe spelers uit landen zoals China en India zijn deze processen in een geweldige stroomversnelling gekomen. De time-to-market laat zich nu in maanden uitdrukken. Nieuwe releases volgen elkaar in een geweldig hoog tempo op.
Dit tempo wordt als gezegd mondiaal bepaald. Daar zullen we in Nederland dan ook op in moeten spelen. Gechargeerd gesteld: er is geen tijd te verliezen, want elke seconde telt!
Succesvolle valorisatie is het omzetten van wetenschappelijke kennis in kansrijke producten en toepassingen. Henk Broeders, de vorige voorzitter van ICT~Office, verwoordde het op treffende wijze: "Innovatie is op een nieuwe manier geld verdienen." En dat lukt alleen als je sneller en eerder bent dan je concurrenten.

Minister Verhagen van Economische Zaken, Landbouw en Innovatie heeft gekozen om de komende jaren met name te investeren in negen topsectoren. De toepassing van ICT zorgt voor vernieuwing en economische groei in de topsectoren van onze economie (water, agrofood, tuinbouw, hightech, life sciences, chemie, energie, logistiek en creatieve industrie). Daarom heeft ICT~Office ook een convenant gesloten met NWO. ICT~Office en NWO maken een speerpunt van valorisatie van wetenschappelijk onderzoek. Zodoende kan de rol van ICT als Innovatie As worden verstevigd binnen de economische topsectoren.
Met een continue stroom van doorbraken op ICT-gebied kan de Nederlandse economie telkens weer nieuwe producten en diensten creëren. Dat geldt met name voor de topsectoren. Topsectoren zijn vrijwel zonder uitzondering gevoelig voor de opname van de meest moderne ICT-technologie omdat ze alleen daarmee hun concurrentie kracht kunnen handhaven. Doorsluizing van ICT-onderzoek naar de topsectoren is dus een must. Het Nederlandse ICT-onderzoek is hoogwaardig en er is aan de kant van de onderzoekers een bereidheid tot publiek-private valorisatie. De ICT-onderzoekssector heeft blijk gegeven haar rol te willen oppakken om de topgebieden nog effectiever te maken in hun innovatie.

Telkens blijkt dat één hoogwaardig onderzoek in de ICT impact heeft in meerdere, geheel verschillende sectoren. ICT-planning kan worden gebruikt voor de landbouw, de verkorting van de ligduur in ziekenhuizen en voor vervoer. Sensor netwerken kunnen worden gebruikt in high-tech systemen, in water en in energie. Multimedia data-exploratie is waardevol in de creatieve industrie, in energie en in high-tech systemen. Computer systemen zijn noodzakelijk voor high-tech systemen en voor logistiek.

Nauwere en betere samenwerking tussen wetenschap en bedrijfsleven is dringend gewenst. Want daar kunnen we het verschil maken. Bij de kennisinstellingen is voor onze lidbedrijven veel te halen, wat vertaald kan worden naar de klantensituatie. Natuurlijk gebeurt er al veel goeds op dit gebied. Er wordt immers steeds meer vraaggericht onderzoek verricht. Andere voorbeelden zijn het vermarkten door universiteiten van uitvindingen en patenten, het verlenen van licenties en de creatie van zogenaamde spin-off's. Dat is valorisatie in optima forma. Kennis wordt omgezet in *business*, door anderen in de kennis te laten delen. Dit vraagt om een maatschappij-gerichte, open houding van wetenschappers, onderzoeksinstellingen en universiteiten. Ondernemende wetenschappers zouden we moeten koesteren. Zij zijn het die werk maken van valorisatie. En valorisatie en innovatie gaan hand in hand. Datzelfde geldt voor valoriseren en innoveren. Als werkwoords-vorm zijn het misschien nog wel krachtiger begrippen. Het is immers vooral een kwestie van doen!

Sylvia Roelofs, algemeen directeur ICT~Office

Het Centrum voor Wiskunde en Informatica, een instituut met een indrukwekkend verleden, een grote staat van dienst, en een briljante toekomst. Zo zie ik het, maar mocht iemand hier ook maar *epsilon* op willen afdingen: laat hem met Jan Karel Lenstra praten. NWO's institutenbeleid, NWO's strategienota, strategische onderzoeksthema's, valorisatie, topsectorenbeleid: onder leiding van Jan Karel Lenstra vindt het CWI moeiteloos zijn plaats. En het knappe is: zonder ook maar iets te hoeven forceren. Het CWI anticipeert, het CWI reageert, het CWI handelt. En altijd vanuit het door Jan Karel uitgedragen zelfbewustzijn, gebaseerd op de observatie: wiskunde is overal, informatica is overal en het CWI opereert aan het front van beide disciplines. En doet dat in samenwerking met partners in binnen- en buitenland: zo'n instituut zou iedereen wel willen leiden!

Jan Karel heeft besloten op zijn hoogtepunt te stoppen. Als directeur dan, hij zal als hoogleraar en onderzoeker zijn vak blijven beoefenen. Ik weet niet of hij zijn mooie kamer in het nieuwe CWI gebouw blijft gebruiken: hij was daar in elk geval erg trots op. Prachtig uitzicht op de dijk (met een beetje goede wil dan) – en op het lange gebouw waar ooit elektronen tot (bijna) 900 MeV versneld werden. Nieuwe of oude kamer: ik wens Jan Karel een omgeving toe waarin hij de inspiratie blijft vinden om de wiskunde te beoefenen en als haar ambassadeur naar buiten te treden. En natuurlijk wens ik hem ietsje meer vrije tijd toe dan hij heeft in zijn huidige functie. Daarbij hoop en verwacht ik dat wij elkaar regelmatig blijven ontmoeten om van gedachten te wisselen over de stand van het land!

Jos Engelen

Actions for creating an attractive research climate in Europe

Peter Nijkamp

1. Optimizing conditions for excellence

Europe has the ambition to be a world knowledge leader. But for several decades already, Europe is a net exporter of research talent. Talented people in research appear to monitor and assess their opportunities on a global scale and often choose to go abroad to pursue their research and to develop their careers in a foreign environment which offers greater opportunities to perform the research they want, to gain desired and prestigious academic positions, to build up and lead important and vital research groups, to make use of excellent scientific infrastructure, etc. As a result of this global competition in the science domain, many European researchers seek for a research career outside Europe (and in particular move to the USA), while also many talented Asian researchers find their research base there. In the light of the needs of a society which is critically dependent on advancements in science and technology and needs talented people to carry knowledge forward, Europe needs to redress the imbalanced outflow of its own talents and moreover it even needs to attract talented people from other parts of the world.

For a further diagnosis, we need to examine the causes of the current situation. After a long period of scientific progress in Europe in the past centuries, other parts of the world, most notably the USA, have in the post war period come up with science systems which offer higher benefits to and more recognition for the best scientists. Larger investments and available funds, a higher drive for promoting talented researchers to important academic positions and a higher concentration of qualified people at top universities pose rather attractive working conditions, so that researchers from all over the world often aspire to work in the academic environment of the USA. In Europe, relatively smaller investments in science, fragmentation in the research system (and thus less critical mass in area where this is required for advancing the frontiers in science), barriers for career movements of the best people, to mention a few causes, have all contributed to a less attractive research climate, thus eroding Europe's competitive position vis-à-vis the rest of the world.

Moreover, other parts of the world outside the USA (e.g. Japan, Korea, China, India) are stepping up their efforts and investments in research and development. These processes will finally result in strong emerging science nations like China and India, which will compete for the best people and make it more attractive for their citizens to perform sophisticated research domestically. The speed with which Europe is making progress is too slow in order to gain real competitiveness, despite current initiatives both of individual countries and of the European Union. More is possible however, as Europe's potential is underdeveloped and needs to be unlocked.

As the awareness grows that Europe needs to work harder and as the sense of urgency is growing, many initiatives are undertaken at local, national, international and European level. Current initiatives cover many domains, but approaches often are too fragmented and have too little scope and mass. Also, within Europe itself, a significant imbalance still exists between countries and universities in terms of investments and appropriate conditions. The endeavour of building up the European Research Area is of the utmost urgency and needs to be continued full force. Moreover, it is of the utmost importance that the pace of progress in building up the ERA is accelerated. Thus, the quality of the research climate (see below) is both a European and a national and local concern.

2. Opportunities

High quality research depends on many factors next to the human brain factor: the presence of good scientific infrastructure and facilities, the provision of sufficient and well allocated funding and adequate funding instruments, excellent access to information, a favourable social capital, good possibilities for exchange and interaction, attractive career structures, clear and visible leaders, etc. Together, all such factors determine the attractiveness and quality of the operating context for researchers, thus constituting what we could call an attractive *research climate*. It is this climate which provides – and crucially sometimes fails to provide – opportunities for people to perform qualitatively good research.

In order to create an attractive research climate, a wide range of opportunities needs to be generated, for example, by generous funding and appropriate career structures, by freedom and responsibility for talented young researchers to choose and direct their own research, by the possibility to address challenging, novel issues and high risk research

topics, and by access to world-class facilities. The more and better such opportunities are created in Europe, the more it will succeed in retaining and attracting talented people, resulting in a higher quality of research in science and technology, which is key to a successful European economy and society. However, in order to achieve this goal, important improvements are still necessary in a number of critical aspects, of which we regard the following to be particularly important:

- **careers and funding** Positions in the science and knowledge domain often lack enough attractiveness in terms of career opportunities for talented people (unclear perspectives, lack of autonomy) and labour market status (worse job conditions and social provisions than positions outside academia). Attracting young people starts with good doctoral education (high quality, open structure, professional supervision, competitiveness with other sectors in the labour market). Subsequently, attractive conditions exist if talented young people have the opportunity to build their career and leadership profile, in academia, but also in industry. Furthermore, adequate funding providing greater autonomy to individual investigators is still in short supply, whereas progress towards the creation of a better career structure, e.g. tenure track systems, is too slow.

- **facilities and access** Concentrations of outstanding research facilities and activities are very important to create challenging and attractive working conditions and opportunities for talented people. Such centres foster high quality of research through the larger scale of competition and greater possibilities, which form the seedbed for creative thinking. Also contributing to the enhancement of the attractiveness of research is the facilitation of access to and the use of (new) technologies and sophisticated research infrastructure facilities.

- **mobility** (International) mobility of researchers is vital for broadcasting ideas and enhancing creativity. Still many barriers to mobility exist, thus impairing the inflow and exchange of highly skilled workers. Complicated and prohibitive legal, administrative and entry regulations[1] obstruct free

[1] On 23 October 2007, the Commission adopted a general proposal for a Directive on a single application procedure for a single permit for third-country nationals to reside and work lawfully in the territory of a Member State (‘blue card’) and on a common set of rights for third-country workers legally residing in a Member State. Its aim is to simplify and streamline application procedures. Nevertheless, there are still many barriers in a free exchange of knowledge workers into the EU.

entry to research opportunities and prevent exchange of scientists and scientific ideas. Diminishing (or removing) barriers offers great potential benefits to the careers of researchers and to their research efforts. Europe needs to improve mobility both inside and outside its borders on a larger scale (Marie Curie is already a good step in that direction). Furthermore, negative financial implications for researchers (e.g. the lack of pension portability[2]) unnecessarily inhibit mobility, whereas intersectoral mobility between academia and industry is too low, damaging both economy and society. In conclusion, the research climate in Europe needs to be made much more attractive in order to re-establish very competitive working conditions for talented people in the research and technology domain. As mentioned above, significant improvements have to be made in various core domains in order to remedy Europe's main weaknesses, in particular:

- entrance conditions to Europe: tearing down formal, legal, practical barriers for knowledge workers;
- gearing funding instruments towards the attraction of qualified people on a global scale;
- creating merit-based research career structures;
- investing more in scientific research infrastructure;
- diminishing inequities in science opportunities across Europe;
- promoting opportunities for undertaking creative research in Europe worldwide.

3. Action needed: Europe, the place to be

The domains mentioned above are diverse and cannot be taken up by just one party, but require efforts from many parties. For example, the improvement of career opportunities for researchers often depends on regulations at the local level (universities, institutions), at the national level (labour market regulations and conditions) and at the transnational level (attractiveness of working conditions differ among countries in Europe as well as compared to the rest of the world). It is therefore important to discern and outline what should be done, by whom and at what level in practical action terms.

At the local level and the national level many variables influence the working environment, e.g. job conditions, career paths, etc. Shortcomings

[2] On 9 October 2007, the Commission adopted an amended proposal for a Directive on improving the portability of supplementary pension rights is aimed at reducing obstacles found within some supplementary pension schemes in order to facilitate worker mobility.

here will have to be remedied by changing local or national circumstances by providing adequate chances to promising young researchers. On the other hand, the European level is required to act with respect to conditions which impact the broader aspects of pursuing careers in science and technology throughout Europe. Such conditions concern especially mobility, European-wide supporting frameworks (e.g. for patenting), efforts concerning large, European-wide research endeavours (big science facilities, trans-European networks, certain domains of science, e.g. climate issues), etc. Joint efforts here at the European level can exploit opportunities offered by the scale of the European economy, while complementing local and national efforts.

The endeavour of the European Research Area (ERA) is a relatively young, but important new concept which can provide an answer to various challenges in the domain of European research, provided it offers the right solutions to the right selection of problems, i.e., if it delivers benefits that are not otherwise accessible to its member states and if it supports common needs and exploits the opportunities of scale. In order to create an attractive research climate and opportunities in such a manner that highly qualified researchers aspire to work in European science, we envisage that at the European level the following three issues deserve serious consideration by all stakeholders:

• **Improve the inflow and exchange of qualified people**
-Create critical mass in terms of increasing the opportunities of research grants and positions which contribute to the independent career of the best scientists. In this respect, the grants of the European Research Council may prove to be seminal for creating more attractive opportunities for researchers by enabling them to start an independent career. At the European level, the number of grants and the budget of the ERC should still be extended.

-Allow/increase the portability of granting instruments between universities, industry and countries without bureaucratic or financial obstacles (i.e. implementation of the principle of money follows researcher, but also practical solutions to social security portability).

-Stimulate the opening up of national regulations and programmes in order to enable portability of and access to research opportunities and grants. This implies running down formal and informal barriers which form impediments

for highly skilled knowledge workers (e.g. legal issues concerning visa, job security, pensions, etc.).

• **Exploit the scale offered by Europe**
Using the larger European scale especially applies to the area of investments in expensive facilities and scientific infrastructure. Such facilities are important, as they drive science forward at the edges through fundamental discoveries, but they also attract talented people and foster innovative research, thus creating a dynamic and inspiring environment for people to work in. Europe should drive forward the formation of research concentrations at the European level by, for example, coordinating the provision and the development of funding of major, expensive infrastructure. The European Strategy for Research Infrastructure (ESFRI) can offer an effective means to engage the relevant research communities in developing initiatives and agenda setting, but this has to be followed up by national and local initiatives.

• **Improve the career perspectives for (young) researchers**
Last, but not least, there is a great need to establish better career paths for researchers and to attract young people to a science career. A number of things are needed in this domain:

-Improvements in the structure of doctoral education in order to increase the quality and attractiveness for young people to pursue careers in science and technology. Improvements include better structure of education (not only at the tertiary level; conditions at the secondary level have to be improved too, in order to create interest in pursuing science careers), as well as better coordination across systems (enhancing exchange possibilities). In this area, the Bologna process is very important.

-Also, better conformity to industry requirements in order to attract qualified researchers to work in industry is needed as well as an open bidirectional flow of researchers between industry and academia.
-The most talented young researchers need to have a clear career perspective, and a system of tenure track has to be put in place. The career perspective should also involve the opportunity to establish a group and/or laboratories under their leadership.

-Funding of different sources (research councils, ministries) has to be adapted to support suitable career paths for young researchers. Experiences at many levels show that granting career and research autonomy improves both quality of research as well as science career opportunities and science career attractiveness. Therefore, funding instruments which support this have to be developed and stimulated. Europe can help to bring different instruments and initiatives in line with each other.

-Finally, more and better coordinated efforts in terms of communication and dialogue with society are required to promote the visibility and the attractiveness of European science. In this respect, the recent plans for creating a European Innovation Partnership are important.

4. Practical proposals

To be effective, it is necessary to focus policy actions specifically on a few selected items, which can be relatively easily achieved and which are likely to have relatively immediate effects. The creation of an appealing research climate is a joint responsibility which must lead to a dedicated operational strategy. In doing so, the support and cooperation of the EU is needed, as it is necessary to join efforts at the national, but especially at the European level. The priority initiatives and actions comprise the following:

- A joint initiative together with the EU is needed to **attract talented people** from outside its borders to Europe. The possibilities of attracting qualified people to Europe stand to benefit enormously from a Europe-wide and broadly recognizable initiative which makes available financial resources for salaries, scientific equipment, personnel costs etc. at the disposal of outstanding scientists with the proven ability to drive science forward. On the basis of a proven track record, national parties (science councils, universities, research institutes) should select and invite candidates of their choice, in order to stay in Europe for an extended period (e.g. min. 5 years) on a significant budget (e.g. 5-10 M€). The selected researchers would be given the opportunity to perform research activities of their choice, at institutes of their choice, without bureaucratic and other limitations. Of course, quality based on competition is critical here. The national parties might share their costs with the EU, which at the European level would select initiatives to match.

- In a knowledge based society an intensive **publicity and promotion of attractive European science positions** in the rest of the world is needed, most notably in the US, as we believe that better visibility of the opportunities in Europe is necessary and opportune. Therefore, we suggest in cooperation with the EU, to take the initiative to organise events in the US to promote Europe science positions and job openings. For example, this could be set up alongside important science conferences or gatherings of science communities (e.g. AAAS).

- An intra-European initiative is necessary, financed by the EU, to **build up capacity in science excellence**, by getting researchers from e.g. recent member countries involved in high-level research environments in other European countries. In this way, scientists from new accession countries build up competencies and experiences, to be used in their countries of origin when they return to important positions there.

- **Easier access** from third countries (e.g. US) to EU, national and multinational research programmes (e.g. ERA-Nets) and vice versa is needed in order to intensify collaboration and to enhance possibilities for joint research cooperation across continents.

5. Epilogue

Europe needs and deserves an attractive research environment, not only for scientific reasons (important as these are), but also for economic and societal reasons. The advancement of human knowledge through the development of intellectual capital – in particular, the development and acquisition of skills through innovative research support systems – is a sine qua non for a healthy future of Europe. Excellent knowledge will create favourable benefits in the business and commercial sector and for society at large. It will also enhance our opportunities to improve quality of life (e.g., environmental protection, health schemes, social integration, cultural advances), while it will favour balanced policy-making in the public and private sector as well. That is why highly innovative, novel and frontier-shifting research is so badly needed. This will stimulate effective capacity building in new promising domains, lead to training and education of highly skilled people and develop new modi operandi for government and business. The value added of excellent research is invaluable! And Europe has the task to invest in the best it has, viz. scientific talent.

Finally, it is our firm belief that a joint action agenda for improving the attractiveness of the research climate for talented scientists will make **Europe the place to be**.

Selected list of relevant documents

- Academy of Finland: Methods for Evaluating the Impact of Basic Research Funding, Helsinki, 2006.
- ALLEA: Challenging of the Future: Reflections of ALLEA on ERA, 2007.
- EC Green Paper "The European Research Area: New Perspectives", 2007.
- EC Green Paper "From Challenges to Opportunities: Towards a Common Strategic Framework for EU Research and Innovation Funding", 2011.
- EUA: Doctoral Programmes in Europe's Universities: Achievements and Challenges, 2007.
- EUA: Lisbon Declaration. Europe's Universities beyond 2010: Diversity with a Common Purpose, 2007 -EURAB: Increasing the Attractiveness of Science, Engineering & Technology Careers, 2007.
- EUROHORCs-ESF response to the Commission Green paper, 2007 - ESLF: Careers in the Life Sciences, 2003.
- LERU: Universities and Innovation: the challenge for Europe, 2006.
- LERU: The Future of the European Research Area, 2007 - Research Councils UK, Excellence with Impact, Swindon, 2007.

Beste Jan Karel,

Onze eerste kennismaking gaat terug naar mijn studententijd in Tilburg. Wij – studenten econometrie – kregen een nieuwe docent Analyse. Dat zou jij worden, maar uiteindelijk werd het Lex Schrijver. Iets met ruilen of zo. Ondanks dat ik geen les van jou heb mogen ontvangen, ben ik toch nog goed terecht gekomen, denk ik.

Bij onze tweede kennismaking was jij directeur CWI en ik de nieuwe NWO-waarnemer in het bestuur. Op subtiele wijze – via een anekdote waar jouw stoel was – en dus niet mijn stoel, kreeg ik mijn plaats aan de bestuurstafel.

Die vergaderingen blonken uit in efficiency, lage frequentie en heldere besluitvorming. Hoe goed jij het CWI runde, bleek ook tijdens de evaluatie van het instituut. Zowel in 2005 als in 2011 kwam er zonder omwegen een predicaat 'excellent' uit de bus.

Hadden we dan nooit onenigheid? Natuurlijk wel. Ik herinner me de discussies over huisstijl en schreefloze letters en schuine puntjes op de i nog als de dag van gisteren. Maar ook daar kwamen we uit en CWI heeft een strakke huisstijl, inclusief onderdelen waar NWO nog niet eens aan gedacht had.

Je betrokkenheid met je medewerkers en toewijding aan het instituut werd letterlijk zichtbaar toen je tijdens de verbouwing jouw directiekamer inruilde voor een bezemkast. Als iedereen moest inschikken, dan ook de directeur!

Inmiddels is de verbouwing klaar en staat er een prachtig instituut. Letterlijk en figuurlijk.

Ze zeggen wel 'ns: je moet weggaan op het hoogtepunt. En dat doe je. Maar of het een lokaal of globaal maximum is, of misschien wel een zadelpunt, zal de tijd leren.

Ik waag me er niet aan om een wiskundige of informaticus te citeren, maar heb voor jou wel een toepasselijke van een scheikundige:
'Imagination, as well as reason, is necessary to perfection in the philosophic mind. A rapidity of combination, a power of perceiving analogies, and of comparing them by facts, is the creative source of discovery.'
En ik wens je nog veel ontdekkingen toe.

Ron Dekker
Directeur Instituten, Financiën en Infrastructuur
NWO

Prima la matematica, dopo la musica

Wim Hutter

In de jaren dat ik als adviseur bij het CWI was betrokken, heb ik met Jan Karel Lenstra natuurlijk over tal van bestuurlijke vragen gediscussieerd. Vaak echter vonden wij dan ook nog de gelegenheid om over onze gemeenschappelijke passie – de muziek – te praten. De parafrase van een zinsnede uit de opera Capriccio van Richard Strauss die ik als titel van dit stukje heb gebruikt was hier beslist van toepassing!
Er is een verband tussen wiskunde en muziek. Leibniz merkte al op dat muziek het genoegen is dat de menselijke geest ervaart uit verborgen sommen, zonder te beseffen dat het om sommen gaat. Er zijn fascinerende verbindingen tussen de wiskunde, de wetenschap die patronen en structuren bestudeert, en de muziek, die gecomponeerd is binnen een structuur en (ritmische en harmonische) patronen vertoont.
Ritme als periodiek element in de muziek is al heel oud; in de muziek van "primitieve" volken is het ritme overheersend terwijl toonhoogte vrijwel kan ontbreken. In moderne tijden zijn sommige componisten hiernaar teruggekeerd: de "clapping games" van Steve Reich zijn daar een mooi voorbeeld van.

Pythagoras
De Grieken rekenden de muziekleer evenals de meetkunde, rekenkunde en sterrenkunde tot de mathematische wetenschappen; in latere tijden werden deze wetenschappen gezamenlijk wel het "quadrivium" genoemd. De muziek werd door hen kwantitatief onderzocht, al in de 6^{e} eeuw v.Chr. diepgaand door Pythagoras. Een belangrijke notie is dat de tonen van het octaaf in de Griekse muziek (met een toonladder van vijf tonen) kunnen worden uitgedrukt in *fracties*. Als A de grondtoon is dan is de eerstvolgende toon met een snaarlengte van 4/5 (ofwel 5/4 maal de frequentie van de A) een C, 3/4 is een D, 2/3 een E en 3/5 een F, terwijl 1/2 een A, een octaaf hoger is. Het spel met deze fracties (en fracties van fracties) kan worden voortgezet waardoor de in de westerse muziek zo vertrouwde diatonische toonladder van 12 tonen ontstaat. De Grieken hechtten grote waarde aan het feit dat alle harmonieën op deze wijze met eenvoudige getallen (2, 3, 4 en 5) zijn te beschrijven. Strikte toepassing van de Pythagoreïsche ratio's leidt evenwel tot een toonladder waarin bepaalde intervallen heel consonant zijn terwijl andere juist "schuren". Een probleem dat voor toetsinstrumenten pas

in de 18e eeuw is opgelost door de “gelijkzwevende stemming”, waarbij de “pijn” gelijkelijk tussen de tonen van de toonladder is uitgesmeerd; alle intervallen klinken vals, maar zo weinig dat dat vrijwel onhoorbaar is. Als ik mijn clavecimbel in deze stemming stem, zal ik dat doen in kwinten, waarbij elke kwint een minieme zweving heeft. De tekortkoming in het Pythagoreïsche stemstelsel blijkt ook uit het feit dat een stemming in zuivere kwinten die na 12 keer weer de grondtoon, maar dan zeven octaven hoger, zou moeten opleveren, iets te hoog klinkt (ongeveer 74/73e te hoog), een verschil dat de *komma* van Pythagoras wordt genoemd en dat door alle kwinten iets te krap te stemmen wordt “weggemasseerd”.

Componisten en getallen

Getallen kunnen in muziekstukken een belangrijke rol spelen als symbolen. In Mozarts Zauberflöte bijvoorbeeld speelt het maçonnieke getal drie een belangrijke rol. De ouverture opent met drie akkoorden (die de bij de vrijmetselaren gebruikelijke drie kloppen op de deur laten horen – in dit geval natuurlijk Tamino's ambitie om de tempel van Sarastro te betreden) en heeft drie voortekens (Es gr.t.). Voorts wordt het allegro-gedeelte van deze ouverture nog een keer onderbroken door drie plechtige akkoorden. Denk voorts aan de *drie* dames en de *drie* knapen die verderop in de opera zingend optreden.

Bij Bach speelt het getal drie ook een grote rol als symbool van de drie-eenheid (triniteit). De grote orgelmis (het derde deel van de Clavir-Übung) wordt bijvoorbeeld besloten met een tripelfuga (d.w.z. een fuga met drie thema's) die duidelijk in drie episodes is opgezet en die drie voortekens heeft (Es gr.t.).

Omdat Bachs muziek zulke overzichtelijke patronen vertoont is wel verondersteld dat de muziek ook allerlei verborgen boodschappen en symbolen bevat; om een voorbeeld te geven: in het credo van de Hohe Messe komt in de delen Credo in unum Deum en Patrem omnipotentem het woord *credo* 7 x 7 keer voor en de frase *in unum deum* 12 x 7 [1]).

Opmerkelijk is dat de naam BACH in het Duitse taalgebied geheel in noten kan worden omgezet: Bes – A – C – B. Dit sterk chromatische motief komt in veel van Bachs werken voor als een soort muzikale handtekening, bijvoorbeeld in de laatste – onvoltooide – tripelfuga uit de Kunst der Fuge. Verder verschijnt het in getransponeerde vorm in tal van Bachs werken, zoals als eerste thema in de vierde, vijfstemmige, fuga in boek 1 van het Wohltemperierte Klavier (WTK). De artistieke kracht van het Bach-motief

wordt wel geïllustreerd door het feit dat tal van latere componisten het hebben gebruikt, vaak als expliciete hommage aan de grote componist: Schumann, Liszt, Reger, Webern, Schönberg en anderen [2]).
Velen zijn op zoek geweest naar het vóórkomen van het Bach-getal 14 (dat ontstaat als Bachs naam in cijfers wordt omgezet met B=2, A=1, C=3 en H=8). Een voorbeeld: de eerste fuga waarmee het eerste boek van het WTK opent heeft 14 noten. Het tweede deel van deze fuga (na het halfslot van maat 13) omvat 14 maten waarin het thema in een dicht netwerk van stretti 14 keer voorkomt. Muzikaal lijkt mij het belangrijkste dat dit stuk, ook zonder deze numerieke analyse of welke andere analyse dan ook, als fantastische muziek wordt ervaren.

Maar ook andere componisten hebben op deze wijze hun handtekening afgegeven. Alban Berg meende dat het getal 23 een grote betekenis had in zijn leven[3]) en in tal van zijn composities speelt het getal 23 een rol. In de Lyrische suite waarin Bergs platonische liefde voor Hanna Fuchs-Robettin zo'n rol speelt, is het aantal maten van de zes delen steeds een veelvoud van 23 en/of van 10, het getal van de geliefde Hanna. En in het derde deel wordt in maat 23 een belangrijk thema geïntroduceerd. In het vioolconcert van Berg treedt het hoofdritme voor het eerst op in maat 23 en heeft het tweede deel 230 maten. In de opera Lulu lijkt in sommige delen (akte 2 m.n.) een grote voorkeur te bestaan voor metronoomwaarden die een veelvoud van 23 zijn.

Een componist met een uitgesproken obsessie voor ritmische patronen en structuren is Olivier Messiaen. Zo bevat het slotdeel van het Livre d'orgue (“les soixante-quatre durées”) notenwaarden van 64 verschillende lengtes (van tweeëndertigsten tot breves) die in twee reeksen verlopen waarbij de laagste een kreeftgang is van de hoogste, en waarbij de langste notenwaarden geleidelijk worden verkort, terwijl gelijktijdig de noten met de kortste waarden steeds worden verlengd, – een nauwkeurig mathematisch procédé dat ook nog eens fantastisch klinkt. De componist gebruikt in diverse werken complexe uit de Hindoeïstische muziek afkomstige ritmische patronen (“tala's”) en smeedt deze samen tot ingenieuze constructies. Niet ten onrechte beschouwde Messiaen zichzelf vooral als “rythmécien”.

Gulden snede
Een ander interessant discussiepunt is de vraag of componisten in hun werk de gulden snede hebben toegepast. In de Muziek voor snaren, slagwerk en

celesta van Béla Bartók begint het derde deel met een solo van de xylofoon op een repeterende hoge f die de progressie 1:2:3:5:8:5:3:2:1 vertoont. Ook in werken van Debussy (Reflets dans l'eau uit de Six Images) en Satie is het gebruik van de gulden snede zichtbaar.

Tot besluit

Tellen, meten, patronen ontdekken, structuren analyseren – het kan allemaal met muziek. Soms geven deze aktiviteiten een imponerend inzicht in de wijze waarop de componist zijn muzikale expressie heeft georganiseerd. Maar voorop blijft natuurlijk altijd het klinkend resultaat staan. De muziek moet voor zich zelf spreken en zijn kwaliteiten bij het beluisteren bewijzen. Jan Karel Lenstra is een kritische en nieuwsgierige luisteraar naar muziek; moge hij in de nu aangebroken levensfase nog vele nieuwe muziek ontdekken – prima la musica!

Noten

1. Kees van Houten en Marinus Kasbergen maakten het wel erg bont in hun boek “Bach en het getal” uit 1985, waarin zij tal van “Bach-getallen” weten te “ontdekken” en daaraan allerlei symbolische betekenissen verbinden.
2. Andere voorbeelden van dergelijke muzikale handtekeningen zijn frequent te vinden in het werk van Dmitri Sjostakovitsj (in de Duitse spelling begint zijn naam met de letters DSCH leidend tot het muzikaal sterke motief D-Es-C-H) of in het Kammerkonzert van Alban Berg waarin de namen van Arnold Schönberg, Anton Webern en hemzelf figureren. In dit concert hebben de eerste twee delen elk 240 maten en het derde (laatste) deel 480 maten.
3. Dominique Jameux “Berg”, Solfège/Seuil 1980 waarin een brief van Berg aan Schönberg over dit onderwerp wordt geciteerd.

Een cryptokromme voor Jan Karel

Het Mathematisch Instituut in Leiden, dat volgens sommigen gaarne voor een dubbeltje op de eerste rang zit, kent een lange traditie in het aanbieden van exquisiete immateriële geschenken. Bij je afscheid van het CWI wil ik je dan ook, geheel in stijl, een gepersonaliseerde *elliptische kromme* meegeven.

Deze kromme, $\mathcal{E}$ geheten, heeft Weierstrass-gedaante

$$Y^2 = X^3 + 26266794886933499972341(X - 1).$$

Precies zoals de politiek dat dezer dagen graag ziet, is $\mathcal{E}$ een toonbeeld van *valorisatie*, dat laat zien dat ook de zuiverste wiskunde heel toepasbaar kan zijn. Immers, neem je je kromme modulo het priemgetal $10^{60} + 38187$, dan krijg je een puntengroep waarvan de orde een *priemgetal* van zestig cijfers is. De discrete logaritme in zo'n puntengroep zal nog wel tot na mijn afscheid in Leiden ondoenlijk blijven, dus $\mathcal{E}$ is uitermate geschikt voor alle huis-, tuin- en keukencryptografische toepassingen die je de komende jaren nog zult hebben. Met zo'n kromme kun je bij vrouw en kinders thuiskomen.

Neem je $\mathcal{E}$ echter modulo $p = 194712192003938259582209$, een veel kleiner priemgetal, dan blijkt $\mathcal{E} \bmod p$ je persoonlijke CWI-kromme te zijn, met puntenaantal

$$19471219\ 20031001\ 20111104.$$

Met dank voor vele jaren van prettige samenwerking,

Peter Stevenhagen

Beste Jan Karel,

Zonder dat we intens hebben samengewerkt, hebben onze wegen elkaar jarenlang gekruist. Meer dan 50 jaar geleden mocht ik tijdens een sporttoernooi in Groningen je fiets lenen, rond 40 jaar geleden bleken we allebei geïnteresseerd in het handelsreizigersprobleem, waarbij mijn daadwerkelijke interesse overigens aanmerkelijk korter duurde en beperkter was dan de jouwe, ongeveer 15 jaar geleden maakten we beiden deel uit van het bestuur van de Akademie Raad voor Wiskunde en enkele jaren geleden zaten we samen in een commissie van de KNAW waarin we advies uitbrachten over het nationale rekenonderwijs. Daar tussendoor vroeg ik je raad bij het samenstellen van een commissie en vroeg ik je aandacht voor een kwestie die mij verontrustte. Ik zal ongetwijfeld nog wat vergeten zijn.

Wat ik zeer in je waardeer en me af en toe jaloers maakt is de rustige, vastberaden wijze waarop je je taken vervult. Goed inschattend wat haalbaar is, je doel in de gaten houdend en contacten leggend waar dat nuttig is. Je kunt je bedoelingen ook goed verwoorden en ziet ze in perspectief. Nadat een commissie-verslag onder je voorzitterschap gemaakt is, ga je het ook gebruiken om het beoogde voorstel te realiseren. Wiskundigen hadden nogal de neiging om te denken dat hun ideeën zo goed waren, dat ze direct omarmd en uitgevoerd zouden worden. In de tijd dat je directeur van het CWI was, is dezelfde lijn gevolgd. Het CWI is betrekkelijk onopvallend en met een duidelijk beleid bestuurd en verschillende vacatures zijn door uitstekende wetenschappers opgevuld. Daarbij had je ook oog voor de bijzondere plaats die het CWI in de Nederlandse wiskunde inneemt.

Ik wens jou en je vrouw, die ik eveneens waardeer, nog vele gelukkige jaren toe. Onze wegen zullen vast elkaar af en toe nog kruisen. Dat zal me een genoegen zijn.

Met hartelijke groet,
Rob Tijdeman

Synopsis van een filmscript

Scene 1
Een functioneel vergaderzaaltje. De kwestie: wie wordt de volgende directeur van het CWI? Kandidaten presenteren zich, na elkaar en elkaar niet ontmoetend. Eén komt uit de USA, maar is door en door Nederlands. Doet zijn verhaal, nauwkeurig en voorzichtig formulerend. Is een hele precieze. Weet wat-ie wil. Na zijn vertrek wordt er stevig gediscussieerd. Hij wordt uiteindelijk benoemd.

Scene 2
Twee mannen in de trein. Tweede klasse. Ze discussiëren over iets wat klinkt als 'ieceetee'. Na een tijdje wordt het stil en schrijft de één minutieus zeer kleine tekentjes (cijfers?) in een notitieboekje. De ander leest in 'n rapport.

Scene 3
Sfeervolle vergaderzaal in een antiek pand. Een bont gezelschap van acht mannen en twee vrouwen. Het gaat alweer over 'ieceetee'. Iemand probeert voortdurend aan het woord te zijn, een ander doet zijn mail op de laptop, een vrouw kiest een verrassende invalshoek. Een luistert aandachtig, weet op het juiste moment de aandacht te krijgen van de voorzitter en brengt zorgvuldig zijn standpunt. Dat niet door iedereen wordt gedeeld. Maar dat hindert hem niet zichtbaar.

Scene 4
Beelden van perspublicaties over het rekenonderwijs in Nederland. Uitgesproken opinies over wat er fout is en hoe het moet, alleen niet convergerend. Flitsen van sprekers met uiteenlopende opinies. Hectiek. Overvloeier naar presentatie van een rapport aan de Minister van Onderwijs. Applaus. De aanbieder heeft het schier onmogelijke gepresteerd: een unaniem advies dat breed wordt gedragen.

Scene 5
Een diner. Hetzelfde gezelschap als in scene 3. De laatst-genoemde spreker blijkt ook hilarische anekdotes te kunnen presenteren, uit de boeiende academische wereld. Bovendien weet hij alles van Italiaans koken en benadrukt het belang van het element tijd (slow food?).

Scene 6
Op de gevel van het CWI verschijnt een nieuw logo. Grafisch, geen opsmuk, passend in een lange Nederlandse traditie op het gebied van grafische vormgeving. De directeur heeft er zich persoonlijk mee bemoeid. Hij blijkt een grote belangstelling te hebben voor typografie en weet er veel van. Beelden van zijn gesprekken met de ontwerper.

Scene 7
Alweer het gezelschap van scene 3 en 4. Het is de laatste vergadering. Er worden cadeautjes uitgedeeld. De voorzitter krijgt mooie woorden, en van één lid een heel bijzonder boekwerkje over typografie, uit de privébibliotheek van de schenker. De voorzitter is er heel erg blij mee. Want ook hij houdt van typografie (en van Italiaans koken).

Scene 8
Feestelijke opening van een nieuw gebouw. Vanaf een bordes spreekt de directeur de menigte toe. Historisch gedegen, helder van boodschap, met visie en humor. Daar kan de ingehuurde acrobate niet tegenop.

Scene 9
Overvloeier van de naar beneden spiralende acrobate naar een luchtopname van een fjord. Inzoomen naar een huis in Noorwegen. Gedekte tafel, smörgåsbord, huiselijkheid. Een blij gezin met kleine kinderen. Een vader die spelletjes met ze doet. En af en toe minuscule karaktertjes neerpent in een klein notitieboekje.

Paul ‘t Hoen
25 mei 2011

Beste Jan Karel,

25 januari 2005 moet in jouw geheugen gegrift staan als een belangrijke dag in je leven. Als je nu twijfelt of vreest iets groots gemist te hebben dan klopt dat. Dat was de datum waarop je ingeschreven bent bij de KvK als directeur van WTCW N.V. Dat is al weer veel langer geleden dan je dacht. En dat kan dan alleen maar wijzen op twee dingen: 1. je vond het geweldig leuk om te doen en dan geldt: *time flies.* 2. deze functie was zo op je lijf geschreven dat je als rasbestuurder iedereen en alles naar je hand zette.

Ik mocht je in mijn rol als secretaris van nabij volgen in deze functie en ik denk dat beide conclusies enige nuancering behoeven. Ik denk niet dat je het geweldig leuk vond om te doen, maar soms behoor je dingen te doen, die komen met de *'job'* en als directeur van het CWI had je min of meer ook de schone taak (‘*noblesse oblige'*) om deel uit te maken van deze directie. En ben je een rasbestuurder die zich als een vis in het water voelt bij besturen? Ik weet het niet zeker, maar heb het gevoel dat je inhoudelijke zaken nog steeds interessanter vindt dan bestuurlijke.

Maar is het directeurschap van WTCW N.V. dan een grote opgave geweest, een lijdensweg, doorbijten en afzien? Dat nou ook weer niet. Het waren op zich geen onplezierige vergaderingen met je collega-directieleden, want ze stonden er allemaal min of meer hetzelfde in als jij. Allemaal verantwoordelijk voor een wetenschappelijk instituut, allemaal veel ervaring met zelf onderzoek doen en allemaal behept met een gezonde afkeer en relativering van het belang van bestuurders. Dat schept al meteen een band en is een goede basis om met elkaar van gedachten te wisselen over hoe onderzoek op het Science Park (beter) op elkaar afgestemd kan worden en hoe er in totaliteit meer onderzoek op het Park gerealiseerd kan worden.

Het heeft je ook in staat gesteld bestuurlijk carrière te maken. Als voorzitter van het NISB mocht jij de directie aansturen en begeleiden op het glibberige

pad van de systeembiologie. Het heeft geleid tot verruiming van je horizon, inhoudelijk en bestuurlijk. Het heeft je gebracht tot nieuwe uitdagingen, bestuurlijke overleggen en het benutten van al je creativiteit. Zo'n verrijking van je leven had je nooit willen missen, toch?

Jan Karel, ik vond het heel plezierig om de afgelopen jaren met je te mogen samenwerken, mijn dank daarvoor. Binnenkort zal ik je uitschrijven bij de Kamer van Koophandel als statutair directeur van WTCW N.V. Ik zal zorgen dat je daar persoonlijk bericht van krijgt, want dan weet je wat de precieze datum is waarop dit avontuur officieel voor je afgelopen is en dan is ook deze cirkel weer rond. Het gaat je goed!

Johan Vos

IHÉS
Institut des Hautes Études Scientifiques
Le Directeur

Bures-sur-Yvette, June 30, 2011

Dear Jan Karel,

You told me already that you would beat me in retiring before me. Now, this is it. I must concede my defeat.

We met rather late in our lives but let me say that I greatly appreciated your straight way of looking at things, as well as your vision about the need for mathematicians to get better organized and not afraid at all to confront other disciplines, and the world at large. Our community needs many more entrepreneurial type research mathematicians like you, and for sure France is not very good at that.

Enjoy many other aspects of life as a new world is born to you.

Most sincerely,

Jean-Pierre

INSTITUT DES HAUTES ÉTUDES SCIENTIFIQUES
Fondation reconnue d'utilité publique
Le Bois-Marie • 35, route de Chartres • 91440 Bures-sur-Yvette
T +33 1 60 92 66 70 F +33 1 60 92 66 69 M jpb@ihes.fr I www.ihes.fr

Informatics Challenges in Future Internet

Keith G Jeffery
Director International Relations, Science and Technology Facilities Council, UK
President-Elect ERCIM
keith.jeffery@stfc.ac.uk

Introduction

A great advantage of age is the wealth of experience gained over the years. Commonly mature researchers dismiss the latest research as a re-hash of something done before. In fact usually modern research is a re-use or re-interpretation with additional elaboration on consolidated research results obtained earlier. For many the concepts of Future Internet (FI), personal mobile devices, GRIDs, CLOUDs, Web 2.0, social networking or even WWW itself are all seen as 'hype' or old concepts or even products wrapped in 'new clothes'. There is some truth in this, but each of these technological components of FI either integrated previous technologies in a new way, extended pre-existing technologies to make them more usable or allied the technology to a new business model so increasing accessibility / availability.

Unfortunately, below the 'hype' level of enthusiastic discussion on FI and its benefits, lie many unsolved technological problems (social, legal and economic problems are not addressed here except occasionally when co-related to a technological challenge). A discussion of those challenges forms the core of this paper.

The paper first discusses and characterises FI, then surveys the current state of ICT and finally deals with the technological challenges.

Future Internet

The term 'future internet' (FI) is basically the idea of a society where people, objects, services and information sources are completely connected and available to end-users through any appropriately-configured device. It may be argued there is no FI since this is a continuing evolution of technology and societal demands. Those of us who remember attempts in the 1960s and 1970s to connect computers to each other, connect them to detector devices for data gathering and to connect interactive end-user devices, will note that many of the problems identified then still exist. The difference is the

magnitude with huge numbers of devices or appliances having IP addresses in the FI. However FI is a useful label for the emerging configuration of ICT.
In a European context the EC has initiated a 600m€ Future Internet PPP (Public-Private Partnership) activity with a General Assembly and funded projects; however the activity has been dominated by commercial companies pursuing agendas for their own purposes and in many ways is lacking the innovation required to ensure that the expected societal benefits especially in the quality of life (security, health environment, education, entertainment, transport) are realised alongside the wealth-creation benefits.

In many ways FI is an integration or packaging of the results of earlier research and development.

At the base is networking. Provision of a 100Gb or faster backbone, with widely available fast inexpensive wireless zones, is a prerequisite and the increasing number of devices will require a greater number of IP addresses and better scheduling and utilisation of bandwidth. IPv6 is intended to accomplish this. Connections to the network devices are standardised with appropriate physical and logical connections for devices of different capabilities. However, is should be noted that – over the last 20 years – improvements in performance and in cost-effectiveness of networking have lagged by many orders of magnitude those achieved in data storage and processors. It would appear that the problem is not technology, but the willingness of telecommunications suppliers to make available appropriate bandwidth.

Connected to the network are various devices: for end-users, for detection / data collection, for control, for processing and for storage. There have been great improvements in performance, miniaturisation and packaging, power consumption, flexibility and adaptability and mobility. There has been less progress on heat output and its management including utilisation. There is little standardisation in the devices themselves leading to a rich diversity and choice but – in general – required hardware interfaces are standardised allowing some degree of 'plug and play'.

The devices depend on operating systems which control the resources available and provide the interconnections. Here there are strong proprietary influences at the user device level (Microsoft, Apple, Symbian, Android etc for mobiles with Windows and various versions of Linux for laptops) necessarily due to the differing characteristics of the hardware being controlled. Interestingly, tablet devices are classified with mobile phones rather than laptops – essentially because they lack backing (disk) storage.

However the operating systems generally adopt communication protocols that are widely used so providing a base-level of interoperation. There is a real challenge in cooperative control of heterogeneous devices in a distributed environment because of the varying characteristics of the operating system – and particularly its local configuration - of each device.

Software standards exist but commonly proprietary features are added by suppliers - for commercial advantage - making interoperation a task of identifying the most feature-rich subset of characteristics available on all platforms.

Data standards are in many ways the key. The formal standardisation of Unicode for character sets was a significant step (although many legacy systems exist which are not Unicode-compatible). Above this there are many standards at both syntactic and semantic levels, usually application-dependent. However some general standards – such as those developed by W3C (World Wide Web Consortium) of which XML is an example – provide a general basis for syntactic interoperability. Semantic interoperability requires facilities to enable understanding – by computer systems – of the meaning of terms. W3C has standardised OWL and SKOS but, in fact, any system capable of storing triples provides the logic basis for semantic interoperability.

In many ways FI is based on the concept of e-Science articulated first in 1999 in UK [Je99] (itself an integration of the results of much earlier thinking on systems and e-infrastructure) and developed much further leading to a summary as 'the fourth paradigm' [HeTaTo2009]. The key principles of (a) virtualisation to the degree where the end-user neither knows nor cares where the data is stored, where the processing is done, how the resources are obtained and (b) how the connection is maintained as long as appropriate service level is achieved were developed within the e-Science concept utilising first GRIDs and later CLOUDs as configurations to provide this virtualisation.

The key technological demands of FI include (a) very high speed access; (b) multiple functional devices for processing, storage, detection, control, user access requirements with standard network interconnects and appropriate performance; (c) interoperation of software; (d) interoperation of data and information; (e) security and privacy appropriate to the user and the application. In this paper we concentrate on key technologies required to realise the FI and particularly (1) metadata; (2) state; (3) data representativity; (4) data quality, veracity and permanence; (5) trust, security

and privacy; (6) management of service levels and quality of service; (7) systems design, development, operation and decommissioning. It is no surprise that in examining these technologies we challenge conventional, existing ICT and the underlying computer science.

Existing ICT

Existing ICT depends dominantly on the concepts of Alan Turing (the Turing machine) [Tu36] which implies sequential processing and of John von Neumann (stored programme hardware / operating system architecture) [vonNe45] which optimises for expensive CPU and memory resources. These concepts were developed in an era when interconnection of computing devices was not envisaged.

Much modern existing ICT falls into the following main categories:

1. Large scale transaction processing for line-of-business applications (including production industry) with associated data warehouses for analytics and decision support. This is typical of most industry, commercial and government organisations and major products are available from e.g. ORACLE and SAP;
2. CAD/CAM (Computer-aided design, computer-aided manufacturing) systems for production industry especially including the use of production robots;
3. Control and optimisation systems for production industry (e.g. oil refinery, chemical plant) or for management of air traffic, water supply, electricity grid, transport logistics etc;
4. Scientific / engineering systems involving data collection from observation or experiment, simulation and complex analysis and visualisation. Data collection commonly requires managing streamed data. This is typical of enterprises in science, engineering and research;
5. Systems to access, update, retrieve from multiple multimedia information sources (sometimes resolving linguistic, syntactic and semantic heterogeneity) with general or subject domain scale often interlinked with the above;
6. Systems for multimodal human intercommunication with voice, e-mail, social networking, games and education, documents (creation, editing, sharing and retrieval), audio, video, task lists, calendaring and associated push technologies usually connected with the above but sometimes used independently;

The majority of these kinds of systems utilise WWW at the user interface and structured relational databases 'behind the scenes' commonly in B2C

(Business to Customer) applications whether within a company or public-facing (such online purchasing or online banking). Additionally B2B (Business to Business) utilising relational databases and web services is used commonly for many applications especially supply chain environments. Many examples of such systems are meeting problems in performance, reliability / business continuity, security, privacy and – where choice is involved – take-up by end-users.

Examples of systems envisaged in the FI – and differentiated from existing ICT systems which will continue and evolve - include:

a) Systems for healthcare at home; monitoring, measuring, advising and administering appropriate drugs;
b) Systems integrating home refrigerator stock control with online delivery of foodstuffs;
c) Systems assisting a traveller in completing their journey whether by public or private car, train, boat, plane with appropriate factors such as time, cost, safety, convenience, environmental cost;
d) Systems for self-paced life-long education and training (including interactions);
e) Systems for managing energy use in homes and industrial / commercial buildings with appropriate factors such as cost, safety, legislation protecting rights of humans, health advice;
f) Systems for politically democratic access to information, debate, discussion and voting;
g) Systems for advanced distributed multiplayer games;
h) Systems for barter and exchange of goods and services;
i) Systems for integrated urban planning including services as well as buildings and physical infrastructure;

These examples require high performance, adaptable, user-friendly systems integrating communications, detectors, processing and data storage within an appropriate framework of service levels covering performance, reliability, trust, security, privacy and legislation.

Put simply the requirement is - instead of imposing ERP (Enterprise Resource Planning)-type systems on end-users (whether within an organisation or as customers or suppliers outside it) - constructing user-centric, flexible systems dynamically to meet the current (and about to change) end-user needs.

The Challenges

To move with success from the existing ICT environments to the FI requires some key developments, most of which do not yet have solutions (or adequate solutions) and most of which require extensive research to reach a suitable state of development.

Metadata

Metadata is data about data. What is metadata to one application is data to another. Thus metadata provides a logical view for utilisation of some data. Since metadata is critically important for interoperation and semantic understanding, there is a requirement for precise and formal representation of metadata to allow automated processing. A basic classification was proposed [Je00] distinguishing schema, navigational and associative metadata, the latter partitioned into descriptive (like a library catalogue card), restrictive (covering e.g. copyright or charges for use) and supportive (relating to the whole domain and providing thesauri, domain ontologies). Research is required into the metadata representation language expressivity in order to represent the real world accurately. For example, the existing Dublin Core Metadata standard [DC] is machine-readable but not machine-understandable, and furthermore mixes navigational, associative descriptive and associative restrictive metadata. A formal version has been proposed [Je99a]. However, metadata is now a 'hot' topic and there is a plethora of 'standards' in the area with consequent problems for interoperation and utilisation of already-collected datasets.

State

Computer systems are expected to portray accurately the state of the real world of interest. Air traffic control systems are expected to provide – on a second-by-second basis - an accurate picture of the location, height, speed and type of planes. Nuclear reactor control systems similarly should portray the state of the reactor including all pertinent variables. In a conventional database management system state is maintained in near-real time by ACID (Atomicity, Consistency, Isolation, Durability) transactions which lock areas of the database to prevent update (or repeatable read) conflicts. As transactions become more complex in such an environment (more instructions related to the update, more database tables affected) the duration of locking – with subsequent performance degradation - becomes unacceptable and so other techniques are used. Roll-back assumes that conflicts occur rarely and so to improve performance allows transactions to proceed without locking with the possibility that a conflict occurs leaving the database in an inconsistent state. The roll-back re-schedules the updates to run sequentially so restoring – with a temporal delay – the database state.

Compensation is based on a similar philosophy and – on detection of a conflict - initiates a transaction which restores the database to a correct state. In a distributed database environment a transaction may affect tables geographically remote from each other with communications latency. Two-phase commit protocols are designed to ensure consistent state across the distributed database – even if this state no longer corresponds with the real world. Similarly compensating transactions can restore a distributed database to an internally-consistent state which may not represent the real world. Use of such transaction controls typically takes from microseconds to minutes. In an environment when millions of nodes are self-updating locally from sensors (including audio, video) and human input it is clear that conventional database technology does not work. The best we can achieve is local state internally consistent and consistent with (a small slice of) the world of interest, and reconciliation across the distributed database as and when required (i.e. when inconsistent parts of the database are referenced for retrieval or update) using lazy techniques. Alternatively, we re-think the notions of state and transactions.

Data representativity

The problem is to represent the real world inside the database system so that changes caused by events triggered by external factors - humans or machines in the real world - are reflected in the database. The database has to represent values of attributes of entities. However, these attributes can range from simple characters (even then there is the problem of representation in legacy systems while today solved by Unicode) through numeric representations with precision and accuracy to whole text (with language variants) or multimedia objects (with different encodings) or even a binary representation impenetrable without more knowledge of the representation. It is thus necessary to describe the attributes by descriptive metadata, ensure integrity using schema metadata, locate using navigational metadata and restrict usage as appropriate using restrictive metadata.

The structure of the information constructed from the data is a key aspect. Commonly humans structure ‘things’ into hierarchies – for example groups within departments within faculties or schools within a university. However, few universities have such a structure and groups may ‘belong’ to > 1 department, faculties or schools may not exist and research centres may either exist independently of any department or ‘owned’ by several. This would indicate that hierarchies do not represent the real world, and in fact the relationships between hierarchic levels in a data structure have semantics, temporal duration and probability – the latter especially important when dealing with incomplete and inconsistent information.

This should not surprise us; we are all familiar with the non-hierarchic '2 parents one child' graph (which has semantics, temporal duration and probability) and the generalization to many-to-many cardinality. Another example is classification of species and subspecies: simple hierarchies do not match the real world state and again the relationships have semantics, temporal duration and probability. Much of the problem with relational database implementations is because people try to force a real world into a hierarchic representation (using primary and foreign keys) whereas using base tables and 'linking relations' allows naturally a n:m relationship (with semantics, temporal and probability attributes) to be expressed. One data model with these properties for the domain of CRIS (Current Research Information Systems) is [CERIF]. In fact this data model has been used in conventional organisational settings as well. As a by-product such an information structure maps directly to hypermedia structures (semantic web and linked open data) such as WWW.

Such a structure has many advantages; for example it is common in some applications for the base tables to be relatively static (append not update) whereas the 'linking relations' are frequently updated – indicating rapid changes in inter-relationships of 'things'. Is this not exactly what is required for the FI? Alternatively, the base tables may be updated frequently whereas the 'linking relations' may remain static – examples include detector arrays where streams of data are collected according to a plan – and the plan is encapsulated in the structural relationships between the data streams – at least for a given temporal duration.

Such a representation allows the 'time machine' to operate: the state can be recreated at any time slice by retrieval across the 'linking relations'. How many times do we wish to know the 'state of the world of interest at a given time or over a given interval? As a side effect storing the temporal information in 'linking relations' is much less expensive than storing temporal information with base attributes as in the conventional temporal relational model [Sn94] and allows standard SQL processing without recourse to the temporally-extended TSQL. As a by-product such a representation assists greatly with provenance.

Such a representation allows for probability on the relations between 'things' (entities or objects) to be expressed within the 'linking relations'. This is particularly helpful in many applications where relationships are deduced by humans, inferred by discussion or speculative in a scientific research environment. Of course the 'intelligence' of processing the probabilities

depends on the capabilities of the system from simple relational calculus through to full fuzzy logic capabilities.

Interoperation – usually to provide a homogeneous view of information to an end-user from heterogeneous sources – clearly requires rich metadata about the attributes and syntax (structure) to be able to match schemas (and more detailed metadata) and to map them to each other. Invariably, additional knowledge processing is required: this may be supplied by humans but progressively more of this task is done by computer systems, using a variety of techniques including graph theory (matching structures) [SkKoBeJe99] or lexical matching with thesauri and knowledge-based reasoning (using domain ontologies related to the attributes and syntax of the domain of discourse) [JeHuKaWiBeMa94]. The current 'half-way-house' of part-machine, part-human reconciliation of databases where only the required resources are integrated and then on a 'pay as you go' basis is the research area commonly referred to as 'dataspaces' [FrHaMa05]. To improve the machine support of this process it is thus necessary to describe the attributes by descriptive metadata, ensure integrity using schema metadata, locate using navigational metadata and restrict usage as appropriate using restrictive metadata – exactly as stated in the first paragraph of this section.

Data quality, veracity and permanence

The purpose of data, especially when structured in context as information, is to represent the world of interest. There are real research issues in ensuring that this is true – especially when the data is incomplete or uncertain, when the data is subject to certain precision, accuracy and associated calibration constraints or when only by knowing its provenance can a user utilise it confidently. Clearly metadata is required to provide the contextual information within which the primary data may be processed. However, there are particular problems in representing incompleteness and uncertainty and in processing such data. Research on modal logics [GaGu03] is providing some hope of solutions.

Validation of data on input is of critical importance to ensure veracity and accuracy. This is best done by validation against constraints stored in the (schema) metadata. However, this is time-consuming and may degrade significantly performance, especially in database update.

Once the data are stored there remains the problem of ensuring they are available and accessible over time. There is a need to preserve digital data from media fading, media migration and loss of ability to process (usually stored in the metadata associated with the dataset or in the associated

software). Recently interest in digital preservation has increased, not least due to the cost of collecting (especially scientific) data. A recent report from the EC [Wo10] highlights the opportunities and difficulties with large-scale scientific research data where they are encountered earlier than – but surely will be followed by – commercial data systems. Work is ongoing on standards for digital preservation, technical solutions and business models to sustain digital preservation. A comprehensive resource is available at [APA].

Trust, security and privacy

Security is an issue in any system, and particularly in a distributed system. It becomes even more important if the system is a common marketplace with great heterogeneity of purpose and intent. The security takes the forms:
a) prevention of unauthorised access: this requires identification of the user, authentication of the user, authorisation of the user to access or use a source or resource and provision or denial of that access. The current heterogeneity of authentication and authorisation mechanisms provides many opportunities for deliberate or unwitting security exposure;
b) ensuring availability of the source or resource: this requires techniques such as replication, mirroring and hot or warm failover supported by digital preservation. There are deep research issues in transactions and rollback/recovery and optimisation;
c) ensuring continuity of service: this relates to (b) but includes additional fallback procedures and facilities and there are research issues concerning the optimal (cost-effective) assurance of continuity.

In the case of interrupted communication there is a requirement for synchronisation of the end-user's view of the system between that which is required on the PDA and / or laptop and the servers. There are particular problems with wireless communications because of interception. Encryption of sensitive transmissions is available but there remain research issues concerning security assurance.

The privacy issues concern essentially the tradeoff of personal information provision for intelligent system reaction. There are research issues on the optimal balance for particular end-user requirements. Furthermore, data protection legislation in countries varies and there are research issues concerning the requirement to provide data or to conceal data. A very common security exposure is through 'social engineering' which tricks the end-user into providing information that can be used to compromise a system.

When any end-user purchases online (e.g. a book from www.amazon.com using B2C) there is a trust that the supplier will deliver the goods and that the purchaser's credit card information is valid. This concept requires much extension in the case of contracts for supply of engineered components for assembly into e.g. a car using B2B. The provision of an e-marketplace brings with it the need for e-tendering, e-contracts, e-payments, e-guarantees as well as opportunities to re-engineer the business process for effectiveness and efficiency. This is currently a very hot research topic since it requires the representation in an IT system of artefacts (documents) associated with business transactions and software to reconcile policies so declared.

Management of service levels and quality of service

In the FI the end-user expects services to be delivered with a given quality defined by a service level agreement. The end-user neither knows nor cares how the service is delivered – the underlying data, software, servers, communications systems are virtualised and thus hidden. Thus there is a requirement for service policies offered to be monitored and enforced through recording of parameters against the restrictive metadata including performance levels and costs. There is a need for service level negotiation using the policies in processes related to trust (see above).

Systems design, development, operation and decommissioning

Conventional systems design has typically relied on models and techniques developed in the sixties and seventies. In particular the approaches of Hierarchic Input Process Output from IBM [HIPO], Jackson Structured Programming (JSP) [Ja75] and Jackson Systems Development (JSD) [Ja83]. These data stream led approaches contrasted strongly with the algorithmic approaches of Wirth [Wi71] and Dijkstra [DaDiHo72]. The latter were much better suited to formal proofs using environments such as those of Z, B and VDM although the high cost of formal verification has led to such techniques being restricted to safety-critical applications such as aircraft control. With the arrival of the WWW and the need for additional functionality in the browser and in the Web Services layer, the concept of mobile code was revived and the Java implementation [GoJoStBr05] became dominant with the associated Java Virtual machine platform.

More recently the re-use of software has become topical, analogous to the re-use of artefacts in mechanical or electrical engineering. This has led to the Service-Oriented Architecture Concept [SOA] with variants of static composition (a system developer composes services together to provide a

complete application) or dynamic composition (the system environment itself discovers and integrates the services). It is claimed this reduces maintenance and development costs and provides more reliable software. It also circumvents the need for developing new and decommissioning old systems since an evolutionary approach of service replacement using 'best of breed' is advocated.

Thus we reach a stage of systems development where we provide services that can be discovered, validated as suitable, composed (dynamically) and executed to meet the requirements of an end-user. Such services require detailed metadata to permit the above operations and there is currently much research in this area. Following on from the work within the GRIDs community and onwards into the CLOUD community, the concept of SOKU (service oriented knowledge utility) has been proposed as a metadata-rich encapsulated serviced which meets the above requirements. SOKUs also have some characteristics found in autonomous agents in that they are loosely coupled to each other and to the operating system environment. One can envisage an ecosystem of autonomous SOKUs interacting dynamically to meet the end-user requirement. One analogy is with a colony of ants or bees; some local autonomy but a commonality of purpose. This links neatly with some approaches for 'future computing' which are bio-inspired.

Conclusion

The Future Internet challenges many long-held hypotheses, theories and practice in ICST. The scale and numbers of objects, the heterogeneity and the demands for appropriate security, privacy, trust, permanence and performance all impinge upon traditional ICST engineering practices. However, the challenges are being approached by research both in narrow specialised fields and across a broad front. There is plenty of stimulating research to be done to meet the challenges.

Jan Karel Lenstra

Much of the above benefited from discussions with Jan Karel in the Strategy Task Group of ERCIM. Although his background is more in mathematics and algorithms, he criticised constructively the architectural ideas and informatics challenges presented here. This work is much richer for his contributions as a colleague and friend.

References

[APA] http://www.alliancepermanentaccess.org/

[CERIF] http://eurocris.org/Index.php?page=CERIF2008&t=1

[DaDiHo72] O.-J. Dahl, E.W. Dijkstra, C.A.R. Hoare Structured Programming, Academic Press, London, 1972 ISBN 0-12-200550-3.

[DC] http://dublincore.org/documents/dcmi-terms/

[FrHaMa05] Michael Franklin, Aron Halevy, David Maier: From Database to Dataspaces: A New Abstraction for Information Management SIGMOD Record Vol 34 No 4 December 2005 pp 27-33 available at http://www.cs.washington.edu/homes/alon/files/dataspacesDec05.pdf

[GaGu03] D.M. Gabbay, F. Guenther (Eds): Handbook of Philosophical Logic 2nd Edition. Kluwer Academic Publishers, Dordrecht, Netherlands., ISBN 1-4020-1644-1.

[GoJoStBr05] Gosling, James; Joy, Bill; Steele, Guy; Bracha, Gilad (2005). The Java Language Specification (3rd ed.). Addison-Wesley. ISBN 0-321-24678-0. http://java.sun.com/docs/books/jls/index.html.

[HIPO] IBM Corporation (1974).*HIPO—A Design Aid and Documentation Technique*, Publication Number GC20-1851, IBM Corporation, White Plains, NY, 1974.

[Ja75] M.A. Jackson (1975). *Principles of Program Design*. Academic Press. 1975.

[Ja83] Michael A. Jackson (1983). *System development*, Prentice Hall, Englewood Cliffs, New Jersey, 1983.

[Je99] Jeffery, K.G 'Knowledge, Information and Data' September 1999 Paper submitted to Director General of Research Councils proposing the programme that became e-Science; available at http://epubs.cclrc.ac.uk/work-details?w=29021

[Je99a] Jeffery, K G: 'An Architecture for Grey Literature in a R&D Context' Proceedings GL'99 (Grey Literature) Conference Washington DC, October 1999.

[Je00] Jeffery, K G: 'Metadata': in Brinkkemper, J; Lindencrona, E; Solvberg, A (Eds): 'Information Systems Engineering' Springer-Verlag, London 2000. ISBN 1-85233-317-0.

[JeHuKaWiBeMa94] Jeffery, K G; Hutchinson, E K; Kalmus, J R; Wilson, M D; Behrendt, W; Macnee, C A: 'A Model for Heterogeneous Distributed Databases' Proceedings BNCOD12 July 1994; LNCS 826 pp. 221-234 Springer-Verlag 1994.

[SkKoBeJe99] Skoupy, K; Kohoutkova, J; Benesovsky, M; Jeffery, K G: 'Hypermedata Approach: A Way to Systems Integration' Proceedings Third East European Conference, ADBIS'99, Maribor, Slovenia, September 13-16, 1999, Published: Institute of Informatics, Faculty of Electrical Engineering and Computer Science, Smetanova 17, IS-2000 Maribor, Slovenia,1999, ISBN 86-435-0285-5, pp. 9-15.

[Sn94] ftp://ftp.cs.arizona.edu/tsql/tsql2/spec.pdf

[SOA] http://www.soa-manifesto.org/

[Tu36] Turing, A.M. (1936). "On Computable Numbers, with an Application to the Entscheidungsproblem". *Proceedings of the London Mathematical Society*. 2 **42**: 230–65. 1937. doi:10.1112/plms/s2-42.1.230. (and Turing, A.M. (1938). "On Computable Numbers, with an Application to the Entscheidungs-problem: A correction". *Proceedings of the London Mathematical Society*. 2 **43**: 544–6. 1937. doi:10.1112/plms/s2-43.6.544.). Reprinted in many collections, e.g. in *The Undecidable* pp. 115–154; available on the web in many places, e.g. at Scribd.

[vonNe45] John von Neumann (1945) First Draft of a Report on the EDVAC; avilable on the web at
http://qss.stanford.edu/~godfrey/vonNeumann/vnedvac.pdf
[Wo10] John Wood (chair) Riding the Wave: Final Report from the High Level Expert Group on Scientific Data 2010 available at http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

[Wi71] Nicklaus Wirth (1971) Program Development by Stepwise Refinement *Communications of the ACM*, Vol. 14, No. 4, April 1971, pp. 221-227.

Jan Karel LENSTRA: Chairman of the Digiteo Scientific Council

In 2007, we contacted Jan Karel to propose him to chair the Scientific Council of Digiteo, the academic research cluster in ICT located South of Paris. As a renowned scientist in our field, Director of a famous institution, CWI, and involved in a number of European ventures, he had the ideal profile. The Scientific council is a body of crucial importance for Digiteo as the "outside eye" on what we are doing or intending to do, and it is really external since it includes only scientists working outside France.

We are very happy that Jan Karel accepted an additional burden being already involved in so many tasks elsewhere. He has taken this chairmanship knowing that he was bringing help and encouragement to a new scientific undertaking which will not give him a lot of benefits! This generosity is one of the many qualities of Jan Karel.

Some courage was also needed: Jan Karel was not afraid of the increasing complexity of the French system which creates new structures at a speed larger than anywhere else. Actually, he manages to understand quite well the context in which Digiteo is developing ("very dynamic" as he said with his usual humour). A possible explanation is that it is not the first time he had to deal with the French system…

So, he has conducted the Scientific Council since 2007 with precision, rigor, and with the wide vision he has about what's going on, Europe-wide, in ICT. He is carefully following the activities of Digiteo and has attended very regularly our annual forum, even when the transportation was problematic last year because of a strike!

It is therefore a pleasure to have this opportunity to pay a tribute to the role of Jan Karel. The good news is that he will continue to chair our Scientific Council and we are very grateful to him for his contribution.

Maurice ROBIN
Director Digiteo

Rekenonderwijscollega's

Jan Karel en het rekenonderwijs

Beste Jan Karel,

Het begin van dit verhaal ken ik alleen uit overlevering maar ik mag het toch graag vertellen omdat het de start is geweest van onze intensieve en leuke samenwerking. Op een zekere oudejaarsavond, al terugdenkend zal het die van 2007 zijn geweest, besloten Robbert Dijkgraaf, Alexander Rinnooy Kan en jij dat het zo toch niet verder kon met het rekenonderwijs in Nederland. Het niveau van het rekenonderwijs daalde alarmerend en voor- en tegenstanders van het 'realistische' rekenonderwijs buitelden in de media over elkaar heen. De discussie hierover was emotioneel en persoonlijk geworden. Beide kampen hadden zich ingegraven en van een inhoudelijke discussie was allang geen sprake meer.

Robbert zou in het nieuwe jaar president van de KNAW worden en Alexander was heel druk als voorzitter van de SER, dus tsjaa…, waarschijnlijk werd enkele oliebollen later geconcludeerd dat het voor de hand lag dat jij deze klus moest gaan klaren. Binnen het bureau van de KNAW hou ik me bezig met de technische wetenschappen, natuurkunde, scheikunde én wiskunde. Wiskunde is net zoiets als rekenen, vonden mijn collega's, dus was ik de meest aangewezen persoon om jou daarin als secretaris te ondersteunen. Zoals je zelf altijd hebt volgehouden had je geen

enkel verstand van rekenonderwijs, en dat gold voor mij zeker net zo sterk. Twee voetballers in een team die allebei niet weten wat buitenspel is, moest dat nou een sterrenteam worden?
Van het bestuur van de KNAW kregen we – gesteund door de staatssecretaris – de opdracht mee om in kaart te brengen wat er bekend is over de relatie tussen rekendidactiek en rekenvaardigheid op grond van bestaande inhoudelijke inzichten en empirisch feitenmateriaal. Als we daarbij ook nog eens aan zouden kunnen geven hoe ouders en leraren meer te kiezen zouden krijgen, zou dat helemaal mooi zijn. En zo gingen we aan de slag.

Gezien de strijd tussen de verschillende geloofsrichtingen voorwaar een ambitieuze opdracht. Achteraf kunnen we stellen dat het heel erg goed heeft uitgepakt. Het advies is door alle partijen geroemd, de toenmalige staatssecretaris heeft direct na ontvangst aangegeven alle aanbevelingen over te nemen en misschien wel het meest belangrijk, de publieke discussie gaat sindsdien niet meer alleen over de nadelen van realistisch rekenonderwijs en de onmisbaarheid van de staartdeling. Daarmee zijn natuurlijk nog niet alle problemen met het rekenonderwijs verholpen, dat vergt een veel langere tijd en het is geweldig dat jij je daar nog steeds op verschillende manieren voor inzet. Jouw rol in het proces is cruciaal geweest. Geheel in jouw stijl zette je direct de toon door aan te kondigen dat het een ‘stevig advies zonder franje moest worden’. Dat is je gelukt, waarbij misschien nog de mooiste prestatie is dat de hele commissie, waarin de aanhangers van de verschillende stromingen duidelijk waren vertegenwoordigd, trots was op het uiteindelijke advies.
Dat geldt zeker ook voor mij. Ik kijk met bijzonder veel genoegen terug op dit advies en vooral op de manier waarop we hier samen vorm aan hebben kunnen geven. Ik hoop dat we in de toekomst nog eens zo’n project samen mogen doen.
Heel veel plezier met al je toekomstige activiteiten. Het ga je goed!

Arie Korbijn
Secretaris KNAW-commissie rekenonderwijs op de basisschool

De wondere wereld van het rekenonderwijs op de basisschool

Kees van Putten en Marian Hickendorff
(Universiteit Leiden)

Eind 2008 ontstonden de eerste plannen om een adviescommissie van de KNAW in te stellen naar het rekenonderwijs op de basisschool, met Jan Karel als voorzitter. De aanleiding was onenigheid tussen professoren over de vernieuwingen in het rekenonderwijs, die met grote felheid in de media werd uitgevochten. Het doel van de commissie was deze ruzies te sussen door in kaart te brengen in hoeverre de verschillende uitspraken over het rekenonderwijs nou eigenlijk wetenschappelijk gefundeerd waren. Niet alleen voor de KNAW was het onderwijs op de basisschool onontgonnen terrein, ook voor Jan Karel was dit een nieuw wetenschappelijk veld. Vol goede moed werd begonnen met het samenstellen van de commissie, en wij waren daar onderdeel van.

Het rekenonderwijs op de basisschool bleek al snel een complex veld. We bogen ons over vragen als 'wat is cijferen?', 'wat is realistisch rekenen, en wat is juist traditioneel rekenen?' en 'waar moet een empirisch onderzoek aan voldoen wil het betrouwbare resultaten leveren?' Af en toe duizelde het de relatieve nieuwkomers in het veld zoals Jan Karel. Maar hij wist zich kranig overeind te houden en is inmiddels ook een expert op het gebied van het rekenonderwijs geworden.

Eén van de grote verdiensten van Jan Karel is dat hij de commissieleden — die soms nogal tegengestelde meningen hadden — bij elkaar wist te brengen en te houden. Daarbij hield hij de sfeer gemoedelijk en constructief. Ook hebben we meerdere malen mogen genieten van anekdotes over het 'jaren vijftig rekenen': daar moeten we toch maar niet meer naar terug willen…

Al met al heeft de commissie een mooi rapport (KNAW, 2009) kunnen afleveren, op grond waarvan tevens nieuwe gelden voor onderzoek naar het rekenonderwijs beschikbaar zijn gesteld. Een mooie prestatie, die niet hetzelfde was geweest zonder Jan Karel! Het siert hem bovendien dat hij allerlei ontwikkelingen die uit het rapport zijn voortgekomen nog steeds actief volgt en becommentarieert: het rekenonderwijs en de opleiding van leraren gaan hem echt ter harte.

Hoofd- versus cijferrekenen in Vlaanderen. Resultaten van een pilootstudie met de "choice/no-choice" methode

Lieven Verschaffel[1] en Joke Torbeyns[1,2]

[1] Centrum voor Instructiepsychologie en –Technologie, Katholieke Universiteit Leuven

[2] GROEP T – Leuven Education College

1. Inleiding

In Nederland is sinds enkele jaren een brede en hevige maatschappelijke discussie aan de gang over het soort van wiskundeonderwijs dat kinderen en jongeren nodig hebben.

Aan de ene kant is er het door het Freudenthal Instituut ontwikkelde en gepropageerde realistische onderwijsmodel, dat uitgaat van "wiskunde als menselijke activiteit", (1) waarin veel aandacht wordt besteed aan het begeleid "her-uitvinden" van wiskundige begrippen en structuren op basis van al aanwezige intuïtieve inzichten en informele strategieën van leerlingen; (2) waarin veel belang wordt gehecht aan betekenisvolle contexten, zowel bij de vorming van de beoogde wiskundige begrippen en structuren als bij het leren toepassen ervan; (3) waarin het wiskundeleren wordt opgevat als het geleidelijk voortbewegen langs verscheidene niveaus van internalisering, abstractie, en formalisering; (4) waarin een centrale plaats wordt ingenomen door reflectie op het eigen handelen, gestimuleerd door interactie met de leerkracht en de medeleerlingen, en (5) waarin ten slotte zorgvuldig gelet wordt op de onderlinge afstemming van verschillende leerlijnen (Treffers, 1987). Deze visie op wiskundeonderwijs, die tot ontwikkeling en bloei is gekomen vanuit een kritiek op het toentertijd als (te) mechanistisch, empiristisch of structuralistisch ervaren wiskundeonderwijs (Treffers, 1987), heeft de voorbije veertig jaar een grote groep Nederlandse wiskundedidactici en -practici geïnspireerd, en heeft er - volgens de voorstanders van deze visie op wiskundeonderwijs - rechtstreeks toe bijgedragen dat het begrip van wiskunde, de vaardigheid in het toepassen ervan, en het vertrouwen en plezier in wiskunde is toegenomen bij leerlingen. Ook internationaal kan deze visie op ruime erkenning, waardering en navolging rekenen (zie bijv. Becker & Selter, 1996; De Corte, Greer, & Verschaffel, 1996; Verschaffel, Greer, & De Corte, 2007).

Daartegenover staat een groep van tegenstanders van het realistische onderwijsmodel - met als bekendste protagonisten Van de Craats (2007) in Nederland en Feys (2008) in Vlaanderen - die stellen dat Nederlandse kinderen – mede door toedoen van dat realistisch onderwijsmodel – minder goed kunnen rekenen dan voorheen, dat de parate kennis van rekenfeiten en de vlotte beheersing van standaardprocedures zijn verdwenen, dat de door de realisten gepropageerde instructieprincipes het rekenonderwijs onnodig chaotisch maakt en wiskundige structurering en abstractie in de weg staan, en dat er dus dringend nood is aan een aanzienlijke koerswijziging weg van dat "realistische rekenen". Ook hun kritiek is overigens te kaderen binnen een ruimere internationale (tegen)beweging die het vernieuwingsgezinde (in het Engels: "reform-based") gedachtegoed, waartoe we de realistische visie kunnen rekenen, op z'n zachtst gezegd niet genegen is.

De meningsverschillen over deze kwesties in de Nederlandse media waren anno 2010 zo heftig, liepen zo uiteen en zorgden voor zoveel verwarring onder beleidsmensen en practici, dat de gerenommeerde Koninklijke Nederlandse Akademie voor Wetenschappen (KNAW) het initiatief nam om helderheid in de discussie te proberen scheppen door de instelling van een commissie met als opdracht te komen tot (empirisch) gefundeerde uitspraken over de relatie tussen onderwijsmethode en rekenvaardigheid en tot daaruit voortvloeiende aanbevelingen over welke richting het Nederlandse wiskundeonderwijs best verder uitgaat.

Een moedig en belangrijk initiatief volgens de enen, een arrogante en ronduit gevaarlijke onderneming volgens anderen. Hoe dan ook, deze commissie is er gekomen, en zij heeft – onder leiding van Jan Karel Lenstra – een goed onderbouwd, evenwichtig en behulpzaam rapport uitgebracht (Koninklijke Nederlandse Akademie van Wetenschappen, 2009), dat zowel bij de vertegenwoordigers van beide stromingen als bij de onzekere middengroep, met belangstelling en waardering is ontvangen (zie http://www.knaw.nl/Pages/DEF/27/161.bGFuZz1OTA.html). Als lid van deze commissie heeft de eerste auteur van deze bijdrage van dichtbij meegemaakt hoe de voorzitter van deze commissie, door middel van zijn rijke intellectuele en wetenschappelijke bagage, zijn strategisch inzicht, zijn uitstekende pen, en zijn grote sociale en persoonlijke vaardigheden, op een zeer kort tijdsbestek en met beperkte mankracht en middelen, deze delicate onderneming tot een goed einde wist te brengen.

De tweestrijd die tot de installering van deze commissie aanleiding gaf, spitste zich steeds meer toe op het strijdpunt van de (noodzaak van de)

beheersing van de standaardalgoritmen voor de vier basisbewerkingen. Met aan de ene kant de voorstanders van realistisch rekenonderwijs die het belang van de (perfecte) beheersing van die algoritmen, inz. van de staartdeling, relativeren, die pleiten voor meer aandacht voor alternatieve rekenwijzen gebaseerd op hoofd- en schattend rekenen, en voor een leerlijn waarbij het cijferrekenen actief, stapsgewijs en inzichtelijk wordt opgebouwd vanuit die andere rekenwijzen. En aan de andere kant de voorstanders van de terugkeer naar "het rekenen van opa", waarin de beheersing van de standaardalgoritmen nog een koninklijke plaats innam in het curriculum, waarin de perfecte beheersing van die eindalgoritmen nog als een onbetwistbaar algemeen einddoel gold, en waarin het leertraject er naartoe gekenmerkt was door een sterk leraargestuurde, prestatiegerichte aanpak met weinig plaats voor inbreng van de leerlingen en weinig zorg om inzicht en flexibiliteit.

In deze korte bijdrage gaan wij ons niet ten gronde uitspreken over de moeilijke vakdidactische kwesties rond hoofd- versus cijferrekenen. Maar de aanbevelingen van de commissie indachtig dat er nood is aan bijkomend empirisch onderzoek (Koninklijke Nederlandse Akademie van Wetenschappen, 2009, aanbeveling 4.1), rapporteren we de resultaten van een kleinschalige studie die wij recent rond dit thema hebben verricht. Hoewel de voornaamste bedoeling ervan was de bruikbaarheid van een bepaalde uit de experimentele psychologie afkomstige onderzoeksmethode – namelijk het "choice/no-choice" paradigma (Siegler & Lemaire, 1997) – voor dit thema te testen, leverde deze studie toch ook een aantal interessante, en enigszins verrassende, initiële bevindingen op, die relevant kunnen zijn voor de discussie rond de efficiëntie en waarde van cijferrekenen.

Met die belangrijke voorafgaande bedenking evenwel, dat de resultaten verzameld zijn in de Vlaamse context, waarvan algemeen bekend is dat daar – ook ten aanzien van het thema van de cijferalgoritmen - globaal genomen een "gematigde" rekendidactiek wordt gevolgd, in die zin dat daar de beheersing van de vier standaardalgoritmen voor alle leerlingen een communaal doel is (Ministerie van de Vlaamse Gemeenschap, 1998), dat er in de meest gehanteerde methoden nog veel aandacht wordt besteed aan de verwerving ervan, en dat in de meeste methoden een aanpak wordt gehanteerd die afwijkt van de "progressieve schematisering" doch eerder als een strakke, inzichtelijke aanpak te typeren valt (Van Biervliet, 1994), zonder dat hoofdrekenen en schattend rekenen evenwel verwaarloosd worden.

2. Onderzoekskader

Zoals hoger toegelicht, kan deze studie gesitueerd worden binnen de context van de actuele en (soms) hevige discussies over de plaats en de waarde van de cijferalgoritmen in het huidige wiskundeonderwijs. Waar, volgens de tegenstanders van de vernieuwingsbeweging, de perfecte beheersing van de cijferalgoritmen nog steeds een belangrijk doel is van, en dus een centrale plaats moet krijgen binnen, het wiskundeonderwijs in de 21ste eeuw, vragen de aanhangers van het realistisch onderwijsmodel veeleer bijzondere aandacht voor alternatieve rekenstrategieën (hoofdrekenen, schattend rekenen) en een geleidelijke en stapsgewijze introductie van de cijferalgoritmen, voortbouwend op deze eerder verworven strategieën. Door eerst uitvoerig aandacht te besteden aan de verwerving van diverse en voor de leerlingen betekenisvolle (hoofd)rekenstrategieën, kunnen zij de noodzakelijke kennis, vaardigheden en attituden opbouwen om deze strategieën ook later, na de introductie van de cijferalgoritmen, efficiënt en flexibel te blijven toepassen. Zo wordt verondersteld dat leerlingen sommen als 534 + 299 ook na de introductie van de cijferalgoritmen blijven oplossen met de compensatiestrategie via 534 + (300-1) = 834-1 = 833, en niet met het (voor deze sommen complexe en met fouten behepte) standaardalgoritme, mits aan deze handige strategie gedurende de eerste jaren van het wiskundecurriculum voldoende aandacht is besteed. Als belangrijk argument voor deze centrale plaats van alternatieve rekenstrategieën verwijzen de vernieuwers naar de onmisbare rol die deze hoofd- en schattend rekenstrategieën spelen bij de geleidelijke, actieve en inzichtelijke opbouw van de cijferalgoritmen (Van den Heuvel-Panhuizen, Buys, & Treffers,2000).

Het aantal empirische studies dat bouwstenen kan aanreiken tot het beantwoorden van de meest prominente vragen in deze erg moeilijke discussie is tot op heden echter beperkt (Koninklijke Nederlandse Akademie van Wetenschappen, 2009). Het beschikbare (schaarse) onderzoek aansluitend bij dit thema bracht aan het licht dat basisschoolleerlingen - in beperkte dan wel in sterke mate geïnstrueerd op basis van het realistische gedachtegoed - op een weinig efficiënte en weinig flexibele manier gebruik maken van hoofdrekenen en cijferen bij het beantwoorden van elementaire rekentaken. Zo toonde Selters (2001) studie aan dat Duitse leerlingen van het derde en vierde leerjaar basisonderwijs (resp. groep 5 en groep 6) de traditionele cijferalgoritmen erg frequent, weinig efficiënt en allerminst flexibel toepassen van zodra deze zijn geïntroduceerd in het onderwijs - en dat zij nog amper de eerder verworven (handige) hoofdrekenstrategieën hanteren, zelfs niet bij oefeningen als 701 - 698 en 199 + 198 die zich uitermate lenen tot het efficiënt en handig strategiegebruik. Daarnaast

bracht het onderzoek van Hickendorff, Heiser, Van Putten en Verhelst (2009) bij Nederlandse basisschoolleerlingen aan het licht dat deze leerlingen erg frequent gebruik maken van hoofdrekenstrategieën bij meercijferige getalsopgaven - en (dus) weinig frequent van de traditionele cijferalgoritmen - maar dat deze verandering in strategiekeuze niet leidt tot betere prestaties. Integendeel, de hogere frequentie van hoofdreken-strategieën gaat gepaard met een sterke daling in de prestaties van de leerlingen bij deze opgaven. Deze tegenvallende onderzoeksresultaten worden door tegenstanders van de vernieuwingsbeweging aangehaald als argumenten voor een terugkeer naar het traditionele curriculum, met een centrale plaats van de cijferalgoritmen van bij de start van het wiskunde-onderwijs. Tegen de achtergrond van deze maatschappelijke en empirische context, beoogden wij in deze pilootstudie de strategiekenmerken van Vlaamse leerlingen bij het optellen en aftrekken met meercijferige getallen te analyseren, met bijzondere aandacht voor het gebruik van hoofdrekenen en cijferen.

Daarbij maakten we gebruik van het “model of strategy change” (Lemaire & Siegler, 1995) en de “choice/no-choice” methode (Siegler & Lemaire, 1997) om de strategiekenmerken van de leerlingen te analyseren.

In het “model of strategy change” wordt onderscheid gemaakt tussen vier kenmerken van strategiegebruik, namelijk (1) strategierepertoire, of de verschillende strategieën die een leerling gebruikt bij het beantwoorden van een rekentaak; (2) frequentie van strategiegebruik, verwijzend naar de frequentie waarmee de leerling deze verschillende strategieën toepast; (3) efficiëntie van strategie-uitvoering, of de accuratesse en de snelheid waarmee de strategieën worden toegepast; en (4) de adaptiviteit of flexibiliteit van strategiekeuzen, gedefinieerd als het kiezen voor die strategie die het snelst leidt tot een accuraat antwoord op elke oefening.

De “*choice/no-choice” methode* komt in essentie neer op het aanbieden van oefeningen in twee soorten van condities, namelijk een keuzeconditie en één of meerdere geen-keuzecondities. In de keuzeconditie kunnen de leerlingen zelf kiezen welke strategieën ze gebruiken om elk van de aangeboden oefeningen te beantwoorden. In de geen-keuzecondities krijgen de leerlingen de opdracht om alle oefeningen op te lossen via één bepaalde, door de onderzoeker opgelegde strategie. Aangezien de leerlingen in de geen-keuzecondities alle oefeningen oplossen via één bepaalde strategie en (dus) geen selectief gebruik kunnen maken van hun voorkeursstrategie bij elke oefening, is het voor de onderzoeker mogelijk om in deze geen-keuzecondities voor elke strategie betrouwbare en valide

efficiëntiedata te verzamelen. Het vergelijken van de efficiëntiedata verzameld in de geen-keuzecondities met de strategiekeuzen van de leerlingen in de keuzeconditie, maakt het verder mogelijk om de adaptiviteit van deze strategiekeuzen in kaart te brengen: maken de leerlingen bij elke oefening daadwerkelijk gebruik van die strategie die hen het snelst leidt tot een accuraat antwoord (zoals kan worden afgeleid uit de efficiëntiedata van de geen-keuzecondities)?

3. Methode

3.1 Deelnemers

Aan het onderzoek namen 21 leerlingen van het vierde leerjaar (groep 6) in Vlaanderen deel (13 jongens; gemiddelde leeftijd: 9 jaar 10 maanden). Alle leerlingen hadden vanaf het tweede leerjaar (groep 4) instructie ontvangen in hoofdrekenen bij het optellen en aftrekken met meercijferige getallen. Daarbij hadden de leerkrachten bijzondere aandacht besteed aan het rijgen als basisstrategie om optel- en aftrekopgaven in het getallengebied tot 1000 te beantwoorden (vb. 614 + 199 = ?; 614 + 100 = 714; 714 + 90 = 804; 804 + 9 = 813). Naast het rijgen, waren ook de splitsstrategie (vb. 614 + 199 = ?; 600 + 100 = 700; 10 + 90 = 100; 4 + 9 = 13; 700 + 100 + 13 = 813) en de compensatiestrategie (vb. 614 + 199 = ?; 614 + 200-1 = 813) aan bod gekomen. De instructie in de cijferalgoritmen voor optellen en aftrekken was gestart in het derde leerjaar (groep 5). Zoals in de meeste Vlaamse rekenmethoden het geval is, waren de cijferalgoritmen slechts beperkt opgebouwd vanuit de (hoofd)rekenstrategieën die de leerlingen al verworven hadden en hadden de leerkrachten deze algoritmen van bij de start van de instructie in cijferen in hun finale vorm aangereikt. Na intensieve oefening gedurende het hele schooljaar, werden de leerlingen verwacht deze algoritmen te beheersen aan het einde van het derde leerjaar. De eerste maanden van het vierde leerjaar (groep 6) hadden de leerkrachten aandacht besteed aan het verder oefenen en herhalen van het hoofdrekenen en de cijferalgoritmen bij het optellen en aftrekken in het getallendomein tot 1000.

3.2 Materialen

Alle leerlingen kregen drie reeksen van telkens vier optel- en vier aftrekoefeningen in het getallendomein tot 1000 aangeboden. We maakten onderscheid tussen twee typen oefeningen (met in elke reeks telkens twee optel- en twee aftrekoefeningen per type). Het eerste type oefeningen, *hoofdreken- of HR-oefeningen*, zijn oefeningen die zich bij uitstek lenen tot het efficiënt gebruik van de compensatiestrategie (vb. 498 + 263; 601-126). We definieerden deze oefeningen op basis van twee kenmerken, namelijk

(1) één van beide termen kan eenvoudig worden afgerond tot het meest nabije honderdtal; de af te ronden term is slechts 1 of 2 eenheden groter of kleiner dan het meest nabije honderdtal; (2) de andere (dan de eenvoudig af te ronden) term is minstens 26 eenheden groter of kleiner dan het meest nabije honderdtal (d.i., TE waarde 26 tot en met 74). Het tweede type oefeningen, *cijfer- of C-oefeningen*, zijn oefeningen die niet uitnodigen tot het gebruik van de compensatiestrategie of een andere handige hoofdrekenstrategie (vb. 456 + 267; 632-164). Deze oefeningen werden gedefinieerd op basis van de twee volgende kenmerken: (1) zowel de eerste als de tweede term uit de opgave is minstens 26 eenheden groter of kleiner dan het meest nabije honderdtal (d.i., TE waarde 26 tot en met 74); (2) geen van beide termen heeft als eenheid de waarde 5, 8 of 9. Alle oefeningen bevatten een brug over zowel de tientallen als de eenheden.

De oefeningen werden geordend op basis van vijf criteria: (1) elke reeks start met de (vier) opteloefeningen; (2) de verschillende typen van oefeningen worden in een niet-voorspelbare volgorde door elkaar aangeboden; (3) de honderdtallen, tientallen en eenheden van de eerste term worden niet herhaald bij twee opeenvolgende trials; (4) de honderdtallen, tientallen en eenheden van de tweede term worden niet herhaald bij twee opeenvolgende trials; (5) het antwoord op de oefening is niet eenvoudig af te leiden uit de vorige trial.

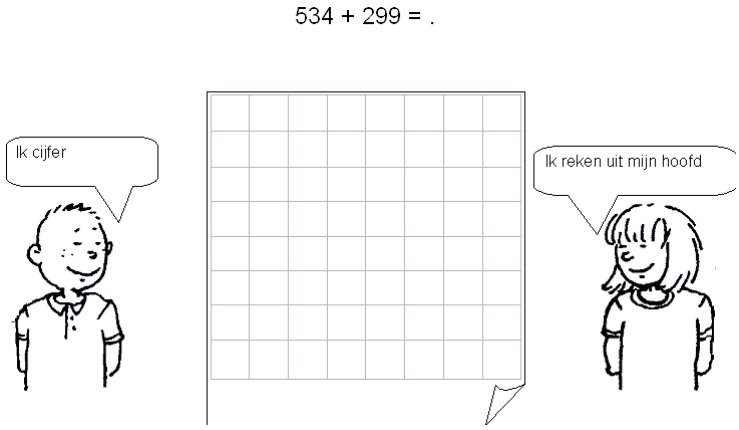
3.3 Condities

Alle leerlingen losten de drie reeksen van oefeningen op in één keuzeconditie en twee geen-keuzecondities. In de *keuzeconditie* konden de leerlingen bij elke oefening kiezen tussen hoofdrekenen en cijferen. Deze strategieën werden gerepresenteerd door twee verschillende figuren, een jongen en een meisje (zie Figuur 1). De leerlingen kregen de opdracht om de oefeningen op te lossen zoals hetzij de jongen hetzij het meisje; zij werden gevraagd om bij het beantwoorden van de oefeningen de (belangrijkste) tussenstappen en resultaten in hun oplossingsproces schriftelijk te noteren op het kladblaadje.[1]

In de *geen-keuzecondities* kregen de leerlingen de opdracht alle oefeningen te beantwoorden via resp. hoofdrekenen en cijferen. Zij werden daartoe verplicht via de instructie en de vormgeving van de toetsboekjes:

[1] Hoofdrekenen werd in deze studie gedefinieerd als rekenen *met* het hoofd (en niet als rekenen *in* het hoofd) (cf. Van den Heuvel, Buys, & Treffers, 2000). Het gebruik van een hoofdrekenstrategie sluit het maken van schriftelijke notities van de (belangrijkste) stappen in het oplossingsproces dus niet uit.

de leerlingen kregen enkel het figuurtje dat de beoogde strategie representeerde aangeboden in de geen-keuzecondities.



Figuur 1. Voorbeeld van een oefening aangeboden in de keuzeconditie

3.4 Procedure

Alle leerlingen losten de drie reeksen van oefeningen individueel op in een rustig lokaal tijdens de schooluren. Alle leerlingen startten met de oefeningen in de keuzeconditie. De orde van de geen-keuzecondities en de representatie van de strategieën (door jongen dan wel meisje) werden gecontrabalanceerd over de leerlingen. De proefleider registreerde in elke conditie het antwoord, de snelheid van antwoorden en het strategiegebruik per leerling en per oefening. De snelheid van antwoorden werd geregistreerd met behulp van een chronometer (tot op 0,01 seconde nauwkeurig). Indien een leerling geen of onduidelijke tussenstappen noteerde tijdens het oplossingsproces, werd hij/zij gevraagd de gevolgde strategie nadien verbaal te rapporteren aan de proefleider, die daarvan letterlijk nota maakte.

4. Resultaten

De resultaten worden gepresenteerd in termen van de vier parameters van het “model of strategy change”. Voorafgaande analyses toonden aan dat de orde van de geen-keuzecondities (eerst hoofdrekenen vs. eerst cijferen) en de representatie van de strategie (jongen vs. meisje) geen invloed hadden op de prestaties van de leerlingen in de verschillende condities, *p*s > .05. Deze variabelen werden daarom niet mee opgenomen in de verdere analyses.

4.1 Repertoire en frequentie van strategiegebruik

De meeste leerlingen (57%) beantwoordden de set van 8 oefeningen in de keuzeconditie via zowel hoofdrekenen als cijferen. Ongeveer één vierde van de leerlingen (24%) loste alle oefeningen cijferend op. De overige leerlingen (19%) maakten enkel gebruik van hoofdrekenen.

Hoewel de meeste leerlingen de oefeningen in de keuzeconditie beantwoordden via zowel hoofdrekenen als cijferen, werd slechts 19% van de oefeningen opgelost met behulp van een hoofdrekenstrategie. De frequentie van hoofdrekenen verschilde niet tussen de verschillende typen oefeningen, $F(1, 20) = 0.00$, $p > .05$, wat er op wijst dat zowel C- als HR-oefeningen minder via hoofdrekenen dan via cijferen werden beantwoord.

Kwalitatieve analyse van de hoofdrekenstrategieën die de leerlingen toepasten in de keuzeconditie, bracht aan het licht dat de leerlingen het frequentst gebruik maakten van rijgen, splitsen of een mengvorm van deze strategieën (meer dan 80% van de oefeningen die hoofdrekenend werden beantwoord, werden opgelost via één van deze drie strategieën). Daarnaast maakten ze ook in beperkte mate gebruik van de compensatiestrategie en indirect optellen (resp. bij 15% en minder dan 5% van de oefeningen die al hoofdrekenend werden opgelost).

4.2 Efficiëntie van strategie-uitvoering

Tabel 1 beschrijft de accuratesse en snelheid van strategie-uitvoering in de geen-keuzecondities per conditie en per type oefening. Zoals kan worden afgeleid uit Tabel 1, resulteerde het verplicht gebruik van de cijferalgoritmen in de geen-keuzecondities in een hogere accuratesse van antwoorden dan het verplicht gebruik van een hoofdrekenstrategie, $F(1, 60) = 25.20$, $p < .01$. Deze hogere accuratesse van cijferen werd geobserveerd voor beide typen van oefeningen, $F(1, 60) = 0.28$, $p > .05$. Met andere woorden, zowel C- als HR-oefeningen werden meer accuraat beantwoord via cijferen dan via hoofdrekenen in de geen-keuzecondities.

Het verplicht gebruik van de cijferalgoritmen resulteerde ook in een grotere snelheid van antwoorden in de geen-keuzecondities dan het verplicht gebruik van een hoofdrekenstrategie, $F(1, 60) = 13.06$, $p < .01$. Ook dit effect werd geobserveerd bij beide typen oefeningen, $F(1, 60) = 0.12$, $p > .05$; het verplicht gebruik van de cijferalgoritmen leidde dus zowel bij C- als bij HR-oefeningen sneller tot een antwoord dan het verplicht gebruik van een hoofdrekenstrategie.

Kwalitatieve analyse van de hoofdrekenstrategieën die de leerlingen toepasten in de geen-keuzeconditie hoofdrekenen bracht aan het licht dat de leerlingen ook in deze geen-keuzeconditie overwegend gebruik maakten van weinig efficiënte splitsstrategieën en amper de beoogde compensatiestrategie hanteerden.

Tabel 1. Accuratesse (in proportie correct) en snelheid (in seconden) van antwoorden in de geen-keuzecondities

	Hoofdrekenen		Cijferen	
	Accuratesse	Snelheid	Accuratesse	Snelheid
HR-oefeningen	0.67	38.50	0.87	27.40
C-oefeningen	0.67	35.13	0.92	25.98
Totaal	0.67	36.81	0.89	26.69

4.3 Adaptiviteit van strategiekeuzen

We bepaalden de adaptiviteit van de strategiekeuzen in de keuzeconditie door, voor elke leerling, de samenhang te berekenen tussen de frequentie van gebruik van hoofdrekenen in de keuzeconditie en het verschil in accuratesse en snelheid tussen hoofdrekenen en cijferen in de geen-keuzecondities. Indien de leerlingen adaptief gebruik maakten van hoofdrekenen en cijferen in de keuzeconditie, verwachtten we een positieve samenhang tussen enerzijds de frequentie van strategiegebruik in de keuzeconditie en anderzijds de accuratesse- en snelheidsverschillen tussen beide strategieën in de geen-keuzecondities (er op wijzend dat de leerlingen het meest gebruik maken van die strategie die ze het best beheersen). Deze analyses brachten aan het licht dat de leerlingen in hun strategiekeuzen flexibel rekening hielden met de snelheid ($r = 0.67$, $p < .01$) maar niet met de accuratesse ($r = 0.12$, $p > .05$) waarmee ze de verschillende strategieën beheersten.

5. Discussie

Concluderend kan worden gezegd dat Vlaamse vierdeklassers (groep 6), na één jaar intensieve oefening van de cijferalgoritmen voor de optelling en de aftrekking, deze niet enkel zeer frequent maar ook erg efficiënt toepassen en daarbij zelfs in zekere mate rekening houden met hoe goed ze de verschillende soorten rekenwijzen beheersen. De beheersing van de hoofdrekenstrategieën was minder goed dan bij cijferen en het gebruik

ervan relatief gering, met name bij de opgaven die sterk om handig strategiegebruik vroegen. Deze enigszins verrassende – en wat het hoofdrekenen betreft eerder teleurstellende - resultaten kunnen vermoedelijk worden verklaard door de sterke focus op de cijferalgoritmen vanaf het derde leerjaar en de daarmee gepaard gaande daling in aandacht voor alternatieve, handige hoofdrekenstrategieën zoals compenseren. Zoals hoger toegelicht, besteedden de leerkrachten in het tweede leerjaar (groep 4) en tijdens de eerste maanden van het derde leerjaar (groep 5) uitvoerig aandacht aan het efficiënt gebruik van hoofdrekenstrategieën bij het optellen en aftrekken in het getallengebied tot 1000, maar stond vanaf medio derde leerjaar het toepassen van de cijferalgoritmen centraal. Deze sterke focus op de cijferalgoritmen, samen met de daarbij gehanteerde didactiek, resulteerde vermoedelijk niet alleen in een goede beheersing van deze algoritmen, maar ook in een klascultuur die het flexibel gebruik van verschillende soorten (handige) hoofdrekenstrategieën (zoals compenseren) niet bevorderde (Yackel & Cobb, 1996). Aangenomen dat dit resultaat – namelijk dat Vlaamse leerlingen vanaf tienjarige leeftijd optel- en aftreksommen met grote getallen (waarvan een aantal zich bij uitstek leent voor handig strategiegebruik) bij voorkeur cijferend aanpakken, en daarbij bovendien correct en snel te werk gaan, terwijl de frequentie en de efficiëntie van hoofdrekenen relatief laag is - in grootschaliger vervolgonderzoek bevestigd wordt, is dit dan een positief of een zorgwekkend gegeven? Het antwoord op deze vraag hangt, zoals voor zoveel vragen ui t de wiskundedidactiek, af van wat men denkt over wiskunde en wiskundeonderwijs en van welke aspecten ervan men maatschappelijk en persoonlijk belangrijk en waardevol acht. Allicht zullen voorstanders van het traditionele rekenen onze resultaten beschouwen als een bijkomende aanwijzing dat we best al onze beschikbare tijd en middelen investeren in het doelgericht onderwijzen van de vier basisalgoritmen, zonder ons daarbij al te veel te bekommeren om de “overbodige” en “moeilijk aan te leren” alternatieve rekenwijzen (hoofd- en schattend rekenen); realistisch geïnspireerde vakdidactici daarentegen zullen er allicht eerder een aansporing in zien om de didactiek van het hoofd- en cijferrekenen verder bij te schaven in realistische zin, zodat leerlingen de dispositie ontwikkelen om, ook nadat ze hebben leren cijferen, (1) de handige hoofdrekenstrategieën als waardevolle werkwijzen te (blijven) beschouwen, en (2) deze strategieën inzichtelijk en handig in te zetten vooral in die kwantitatieve (probleem)situaties waarin hun computationele winst ten opzichte van het standaardalgoritme manifest is.

Referenties

Becker, J. P., & Selter, C. (1996). Elementary school practices. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 511-564). Dordrecht, The Netherlands: Kluwer.

De Corte, E., Greer, B., & Verschaffel, L. (1996). Learning and teaching mathematics. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 491-549). New York: Macmillan.

Feys, R. (Ed.). (2008). Mad Math en Math War (Themanummer*), Onderwijskrant*, nr. 146.

Hickendorff, M., Heiser, W. J., Van Putten, C. M., & Verhelst, N. D. (2009). Solution strategies and achievement in Dutch complex arithmetic: Latent variable modeling of change. *Psychometrika, 74,* 331-350.

Koninklijke Nederlandse Akademie van Wetenschappen (2009). *Rekenonderwijs op de basisschool: analyse en sleutels tot verbetering*. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General, 124,* 83-97.

Ministerie van de Vlaamse Gemeenschap (1998). *Ontwikkelingsdoelen en eindtermen. Informatiemap voor de onderwijspraktijk. Gewoon basisonderwijs*. Brussel: Ministerie van de Vlaamse Gemeenschap, Departement Onderwijs, Afdeling Informatie en Documentatie.

Selter, C. (2001). Addition and subtraction of three-digit numbers: German elementary children's success, methods, and strategies. *Educational Studies in Mathematics, 47,* 145-173.

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General, 126*, 71-92.

Treffers, A. (1987). *Three dimensions. A model of goal and theory description in mathematics education. The Wiskobas project*. Dordrecht, The Netherlands: Reidel.

Van Biervliet, P. (1994). Cijferen: een kunst? Vier aanpakken voor het cijferrekenen. *Onderwijskrant*, nr. 84, 25-37.

Van de Craats, J. (2007). Waarom Daan en Sanne niet kunnen rekenen. *Nieuw Archief voor Wiskunde. 5/8*(2), 132-136.

Van den Heuvel-Panhuizen, M., Buys, K., & Treffers, A. (Red.) (2000). *Kinderen leren rekenen. Tussendoelen annex leerlijnen: hele getallen, bovenbouw basisschool*. Utrecht: Freudenthal Instituut.

Verschaffel, L., Greer , B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 557-628). Greenwich, CT: information Age Publishing.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27,* 458-477.

Omslagontwerpen

Cover voor liber amicorum jkl

door kitty molenaar

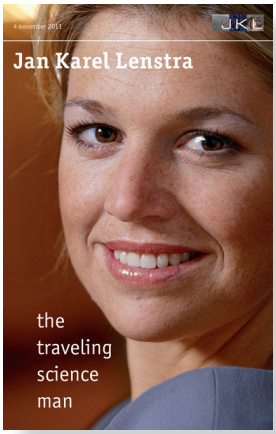
Diverse omslagontwerpen zijn aan de redactie voorgelegd. Doel was een omslag te selecteren dat zich onderscheidt in vormgeving, typografie, beeldbehandeling en grafisch-technische productie. Hier volgen zes ontwerpen die de revue zijn gepasseerd.

vrolijk, levendig en actueel

Het nonogram van Joost Batenburg, iets meer herkenbaar gemaakt, is op de binnenzijde van het omslag gedrukt en subtiel zichtbaar aan de buitenzijde. Het omslagmateriaal is in de ogen van de redactie: 'Mooi, maar kwetsbaar'.

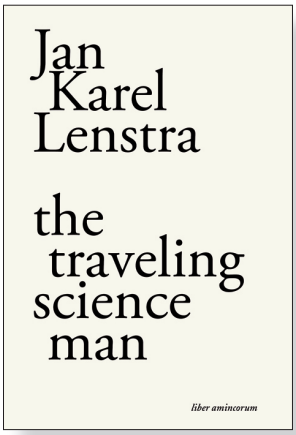
opvallend door zijn eenvoud

Er is hier optimaal gebruik gemaakt van het schutblad, dat een verrassende tegenstelling vormt met de ingetogen voorzijde. De voorgestelde integraalband is flexiber dan een hardcover. Een redactielid over het ontwerp: 'Mooi, maar wel wat braaf'.



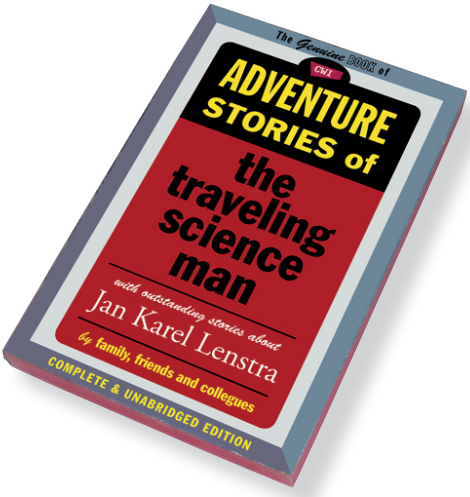
hollands en herkenbaar

Dit ontwerp is geïnspireerd door de bijdrage 'Chains of Fame'. Toepasselijk portret op mooie eenvoudige cover met herkenbare elementen uit de omslag van het Nieuw Archief voor Wiskunde.



terughoudend en betrokken

De passie voor boeken en typografie van jkl is bekend. Er is getracht een evenwichtig typografisch omslag te ontwerpen. Het lettertype is de Garamond.



prettig afwijkend en lichtvoetig

De uitstraling van deze cover doet gedateerd aan (jaren '50). Het is waarschijnlijk als grap bedoeld.

oog voor detail

Steeds wordt verwezen naar het oog voor detail van jkl, vooral op het gebied van de punt. Dat is uitgewerkt in het uiteindelijke ontwerp. Het lettertype op de voorzijde is de Bembo, de lettertypes op de achterzijde corresponderen met de punten op de voorzijde. De keuze voor goud spreekt voor zich.