The Art of Symbolic Computation

Erich Kaltofen 寒爐
North Carolina State University
Department of Mathematics
www.kaltofen.us

Caviness's foreword to the Computer Algebra Handbook

Two ideas lie gleaming on the jeweler's velvet. The first is the calculus, the second, the algorithm. The calculus and the rich body of mathematical analysis to which it gave rise made modern science possible; but it has been the algorithm that has made possible the modern world.

—David Berlinski, *The Advent of the Algorithm*

Caviness's foreword to the Computer Algebra Handbook

Two ideas lie gleaming on the jeweler's velvet. The first is the calculus, the second, the algorithm. The calculus and the rich body of mathematical analysis to which it gave rise made modern science possible; but it has been the algorithm that has made possible the modern world.

—David Berlinski, *The Advent of the Algorithm*

So, gentle reader, I recommend this volume and all its concepts, symbols, and algorithms to you.

—Bob Caviness, *Computer Algebra Handbook*

Where it began

1960s-early 70s: MIT project MAC [Moses]

$$\int 1+(x+1)^n \mathrm{d}x = x+(x+1)^{n+1}/(n+1), \quad n \neq -1$$

S. C. Johnson, "Tricks for Improving Kronecker's Method," Bell Laboratories Report 1966.

Berlekamp/Zassenhaus's, Risch's algorithms

$$\int \frac{x+1}{x^4} e^{1/x} \mathrm{d}x = -\frac{x^2-x+1}{x^2} e^{1/x}$$

B. G. Claybrook, "A new approach to the symbolic factorization of multivariate polynomials," *Artificial Intelligence*, vol. 7, (1976), pp. 203–241.

Important algorithms: "classical" computer algebra

Euclid, Chinese remainder

Sturm chains, Seidenberg's algorithm

Gauss's distinct degree factorization, Berlekamp/Zassenhaus

Berlekamp/Massey

Gröbner, Macaulay resultants, Wu triangular sets

Risch integration and transcendence theory of special functions

FFT-based polynomial arithmetic

Gosper and Karr

Collins cylindrical algebraic decomposition

...

KUNGL. VETENSKAPSAKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Information Department, PO Box 50005, SE-104 05 Stockholm, Sweden, webbsite: www.kva.se
Tel: +46-8-673 95 95, Fax +46-8-15 56 70, e-mail: info@kva.se

THE NOBEL PRIZE IN PHYSICS 1999

PRESS RELEASE 12 OCTOBER 1999
The Prize I Further reading I The laureates

The Royal Swedish Academy of Sciences has awarded
the 1999 Nobel Prize in Physics
jointly to

Professor **Gerardus 't Hooft**, University of Utrecht, Utrecht, the Netherlands, and
Professor Emeritus **Martinus J.G. Veltman**, University of Michigan, USA, resident in Bilthoven, the Netherlands.

The two researchers are being awarded the Nobel Prize for having placed particle physics theory on a firmer mathematical foundation. **...**

The Academy's citation:
"for elucidating the quantum structure of electroweak interactions in physics."

...
One person who had not given up hope of being able to renormalize non-abelian gauge theories was **Martinus J.G.Veltman**. At the end of the 1960s he was a newly appointed professor at the University of Utrecht. Veltman had developed the *Schoonschip* computer program which, using symbols, performed algebraic simplifications of the complicated expressions that all quantum field theories result in when quantitative calculations are performed. Twenty years earlier, Feynman had indeed systematised the problem of calculation and introduced *Feynman diagrams* that were rapidly accepted by researchers. But at that time there were no computers. Veltman believed firmly in the possibility of finding a way of renormalizing the theory and his computer program was the cornerstone of the comprehensive work of testing different ideas.

Important algorithms: “middle earth”

Zippel and Ben-Or-Tiwari sparse interpolation

Singer and Kovacic differential equation solvers

Lattice basis reduction [LLL]

Zeilenberger

Wiedemann, block Wiedemann/Lanczos, matrix Padé

Straight-line and black box polynomial factorization

Baby steps/giant steps algorithms for lin. and polynomial algebra

Tellegen’s principle

Real roots of polynomial systems

Noda-Sasaki approximate GCD, Sasaki approx. factorization

Corless et al. SVD methods

…

Important algorithms: "modern" symbolic computation

Sparse resultants, A- and J-resultants

Giesbrecht/Mulders-Storjohann diophantine linear solvers

Fast bit complexity in linear algebra over the integers

Black box matrix preconditioners, early termination

Sasaki/van Hoeij power sums, Bostan et al. logarithmic derivatives

Sparsest shift of polynomials

Villard-Jeannerod optimal polynomial matrix inverse

Skew, Ore and differential polynomial factorization

Approximate polynomial factorization via differential equations

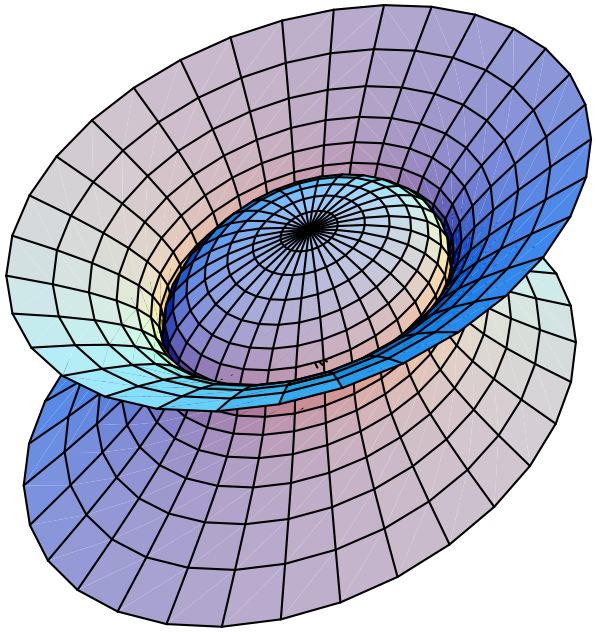
Barvinok-Woods and De Loera et al. short rational functions

Lenstra/Kaltofen-Koiran lacunary polynomial factorization

...

Factorization of "noisy" polynomials over the complex numbers

$$81x^4+16y^4-648z^4+72x^2y^2-648x^2-288y^2+1296=0$$



$$(9x^2+4y^2+18\sqrt{2}z^2-36)(9x^2+4y^2-18\sqrt{2}z^2-36)=0$$

$$81x^4+16y^4-648.003z^4+72x^2y^2+.002x^2z^2+.001y^2z^2$$
$$-648x^2-288y^2-.007z^2+1296=0$$

The Approximate Factorization Problem [Kaltofen, Sasaki 1989]

Given $f \in \mathbb{C}[x,y]$ irreducible, find $\tilde{f} \in \mathbb{C}[x,y]$ such that
$\deg \tilde{f} \leq \deg f$, $\tilde{f}$ factors, and $\|f - \tilde{f}\|$ is minimal.

The Approximate Factorization Problem [Kaltofen, Sasaki 1989]

Given $f \in \mathbb{C}[x,y]$ irreducible, find $\tilde{f} \in \mathbb{C}[x,y]$ such that $\deg \tilde{f} \leq \deg f$, $\tilde{f}$ factors, and $\|f - \tilde{f}\|$ is minimal.

Problem depends on choice of distance norm $\|\cdot\|$, and notion of degree.

We use Euclidean-norm, and multi-degree: $\mathrm{mdeg}\ f = (\deg_x f, \deg_y f)$

The Approximate Factorization Problem [Kaltofen, Sasaki 1989]

Given $f \in \mathbb{C}[x,y]$ irreducible, find $\tilde{f} \in \mathbb{C}[x,y]$ such that
$\deg \tilde{f} \leq \deg f$, $\tilde{f}$ factors, and $\|f - \tilde{f}\|$ is minimal.

Problem depends on choice of distance norm $\|\cdot\|$, and
notion of degree.

We use Euclidean-norm, and multi-degree: $\operatorname{mdeg} f = (\deg_x f, \deg_y f)$

Degree bound is important:
$(1+\delta x)f$ is reducible but for $\delta < \varepsilon/\|f\|$,

$$\|(1+\delta x)f - f\| = \|\delta x f\| = \delta\|f\| < \varepsilon$$

Status of the Approximate Factorization Problem

- No polynomial time algorithm (except for constant degree factors [Hitz, Kaltofen, Lakshman '99])

Status of the Approximate Factorization Problem

- No polynomial time algorithm (except for constant degree factors [Hitz, Kaltofen, Lakshman '99])

- Several algorithms and heuristics to find a nearby factorizable $\hat{f}$ if f is "nearly factorizable"
  [Corless et al. '01 & '02, Galligo and Rupprecht '01,
  Galligo and Watt '97, Huang et al. '00, Sasaki '01,...]

Status of the Approximate Factorization Problem

- No polynomial time algorithm (except for constant degree factors [Hitz, Kaltofen, Lakshman '99])

- Several algorithms and heuristics to find a nearby factorizable $\hat{f}$ if f is "nearly factorizable"
  [Corless et al. '01 & '02, Galligo and Rupprecht '01,
  Galligo and Watt '97, Huang et al. '00, Sasaki '01,...]

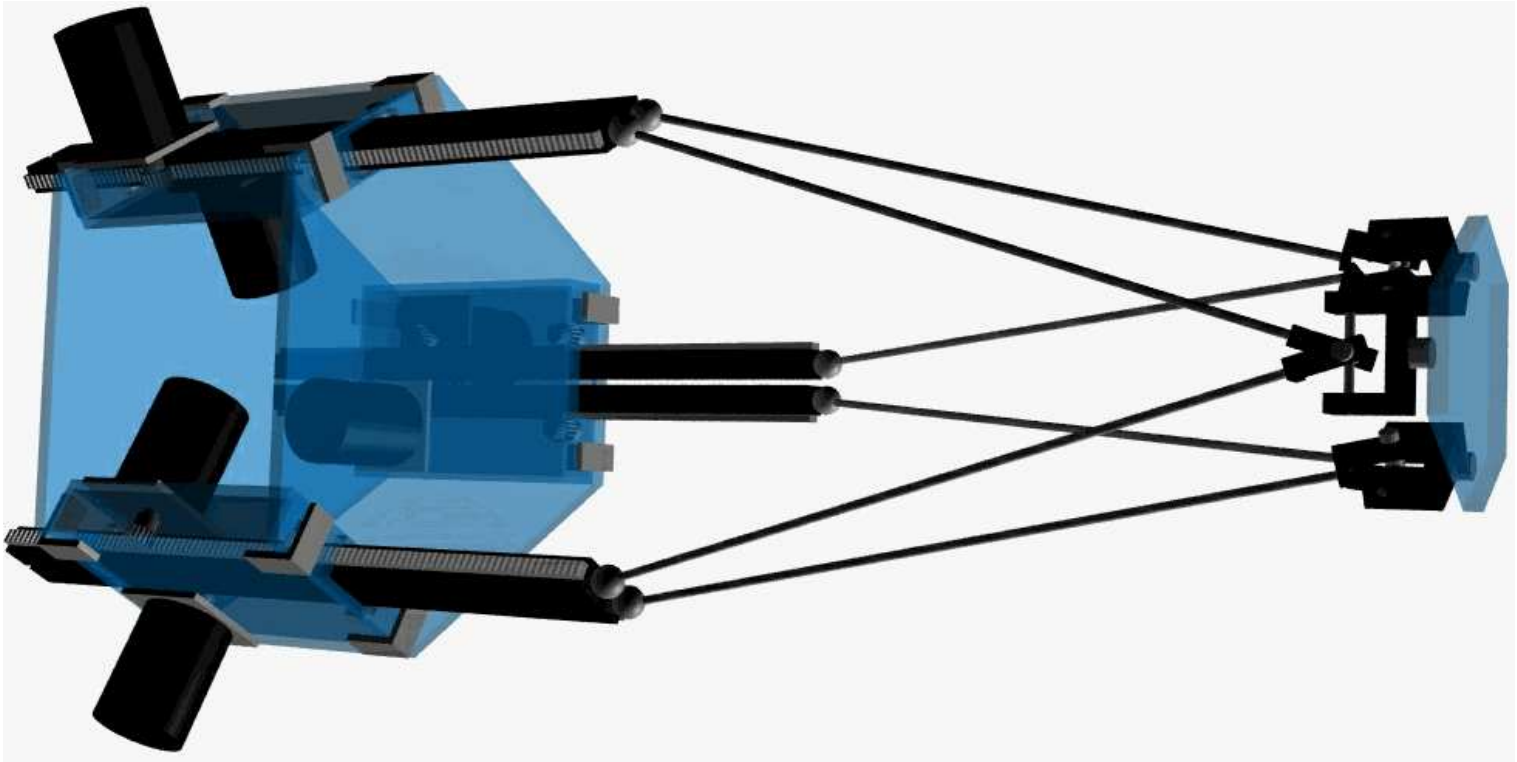
- There are lower bounds for $\min \|f - \tilde{f}\|$ ("irreducibility radius")
  [Kaltofen and May ISSAC 2003; Nagasaka CASC 2004, 2005]

Our ISSAC'04 Results [Gao, Kaltofen, May, Yang, Zhi]

An algorithmically-engineered practical algorithm to find the factorization of a nearby factorizable polynomial given any f.

especially "noisy" f:
Given $f = f_1 f_2 + f_{\text{noise}}$,
we find $\hat{f}_1, \hat{f}_2$ such that $\|f_1 f_2 - \hat{f}_1 \hat{f}_2\| \approx \|f_{\text{noise}}\|$
even for large noise: $\|f_{\text{noise}}\| / \|f\| \geq 10^{-3}$

Verschelde's Stewart-Gouch Platform Benchmarks



Josh Targownik's bypass surgery motorized manipulator

What is an algorithm?

– **finite** unambiguous list of steps ("control, program")

– computes a function from $D \longrightarrow E$ where D is **infinite** ("infinite Turing tape")

Ambiguity through randomization

– Monte Carlo (BPP): "always fast, probably correct".
Examples: `isprime`

Lemma *[DeMillo&Lipton'78, Schwartz/Zippel'79]*
Let $f, g \in \mathbb{F}[x_1, \ldots, x_n], f \neq g, S \subseteq \mathbb{F}$.

$$\text{Probability}(f(a_1, \ldots, a_n) \neq g(a_1, \ldots, a_n) \mid a_i \in S) \geq 1 - \max\{\deg(f), \deg(g)\}/\text{cardinality}(S)$$

sparse polynomial interpolation, factorization, minimal polynomial of a sparse matrix

Do we exactly know what the algorithm computes? E.g., in the presence of floating point arithmetic?

Do we exactly know what the algorithm computes? E.g., in the presence of floating point arithmetic?

– Las Vegas (RP): “always correct, probably fast”.
  Examples: polynomial factorization in $\mathbb{Z}_p[x]$, where $p \gg 2$.
  Determinant of a sparse matrix

Do we exactly know what the algorithm computes? E.g., in the presence of floating point arithmetic?

– Las Vegas (RP): "always correct, probably fast".
Examples: polynomial factorization in $\mathbb{Z}_p[x]$, where $p \gg 2$.
Determinant of a sparse matrix

De-randomization: conjectured slow-down is within polynomial complexity.

M. Agrawal, N. Kayal, N. Saxena, "PRIMES is in P," 2002.

Zeev Dvir and Amir Shpilka, "Quasi-polynomial polynomial identity testing for depth-3 circuits with bounded top fan-in," 2005.

Kabanets and Impagliazzo [STOC 2003]
If Schwartz/Zippel **can be** de-randomized (subexponentially), then there **do not** exist polynomial-size circuits for NEXP or the permanent.

Kabanets and Impagliazzo [STOC 2003]
If Schwartz/Zippel **can be** de-randomized (subexponentially), then there **do not** exist polynomial-size circuits for NEXP or the permanent.

Efficiency dilemma: the higher the confidence in the result, the more time it takes to compute it.

Black box polynomials

$f \in \mathbb{F}[x_1, \ldots, x_n]$

$\mathbb{F}$ an arbitrary field, e.g., rationals, reals, complexes

Perform polynomial algebra operations, e.g., factorization with

$(n \cdot \deg(f))^{O(1)}$ { black box calls, arithmetic operations in $\mathbb{F}$ and randomly selected elements in $\mathbb{F}$

Black box matrices

$A \in \mathbb{F}^{n \times n}$ singular

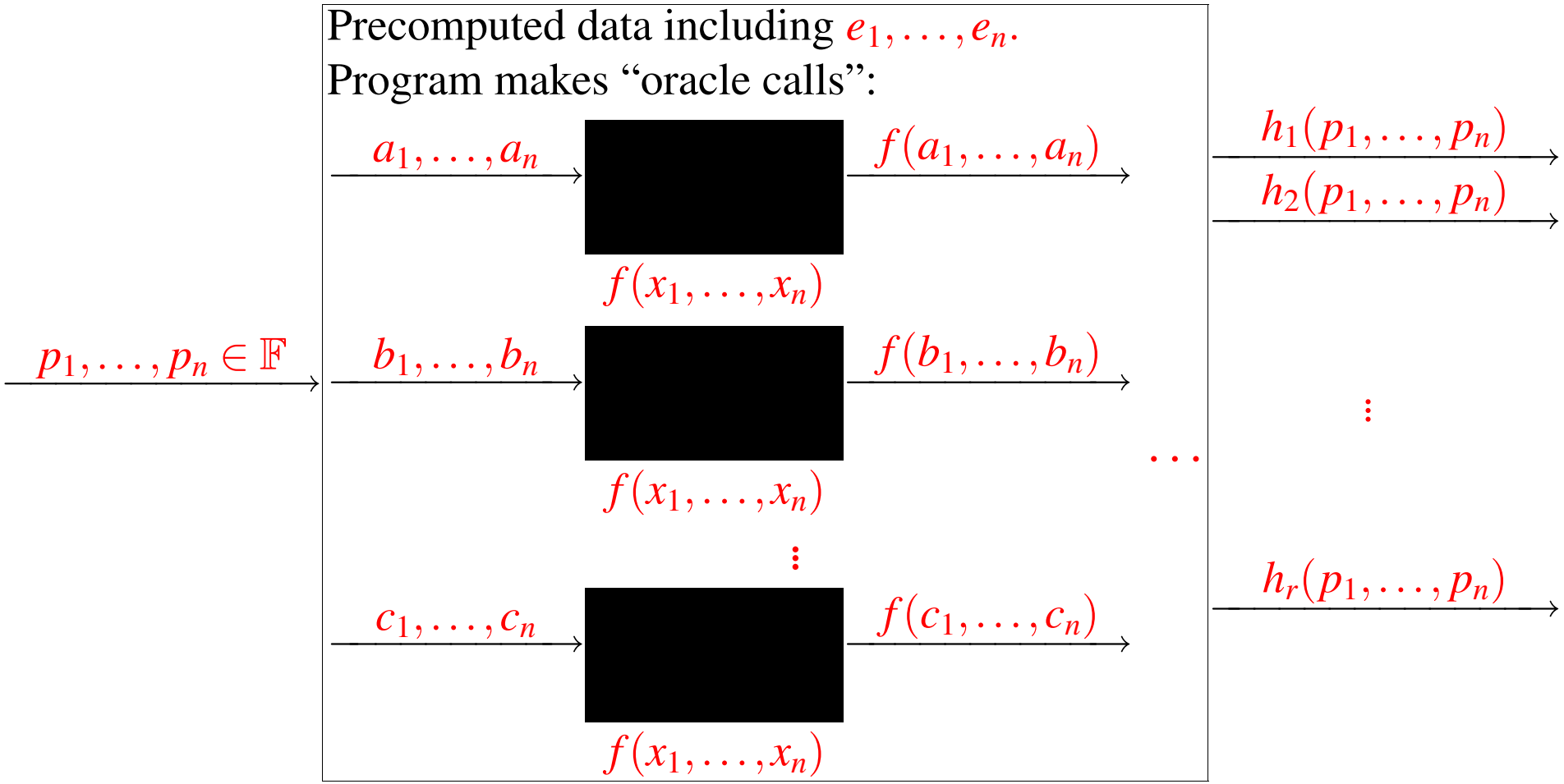
$\mathbb{F}$ an arbitrary, e.g., finite field

Perform linear algebra operations, e.g., $A^{-1}b$ [Wiedemann 86] with

$O(n)$	black box calls and
$n^2(\log n)^{O(1)}$	arithmetic operations in $\mathbb{F}$ and
$O(n)$	intermediate storage for field elements

LinBox Release 1.0 [`www.linalg.org`]: an exact Matlab

Black box manipulation ("functional programming"): Factorization [Kaltofen and Trager 1988]



$$f(x_1, \ldots, x_n) = h_1(x_1, \ldots, x_n)^{e_1} \cdots h_r(x_1, \ldots, x_n)^{e_r}$$

$h_i \in \mathbb{F}[x_1, \ldots, x_n]$ irreducible.

Given a black box

$f(x_1, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$

$\mathbb{F}$ a field

compute by multiple evaluation of this black box the sparse representation of f

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{t} a_i x_1^{e_{i,1}} \cdots x_n^{e_{i,n}}, \quad a_i \neq 0$$

Many algorithms that are polynomial-time in $\deg(f), n, t$:

Zippel 1979, 1988; Ben-Or, Tiwari 1988

Kaltofen, Lakshman, Wiley 1988, 1990

Grigoriev, Karpinski, Singer 1988

Kaltofen, Lee, Lobo 2000, 2003

Mansour 1992; Giesbrecht, Lee, Labahn 2003: numerical method

Show Wen-shin Lee's Maple worksheet

FoxBox [Díaz, Kaltofen 1998] example: determinant of symmetric Toeplitz matrix

$$\det\left(\begin{bmatrix} a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \\ a_1 & a_0 & \cdots & a_{n-3} & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-2} & a_{n-3} & \cdots & a_0 & a_1 \\ a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{bmatrix}\right)$$

$$= F_1(a_0, \ldots, a_{n-1}) \cdot F_2(a_0, \ldots, a_{n-1}).$$

over the integers.

Serialization of **factors box** of 8 by 8 symmetric Toeplitz matrix modulo 65521

15,8,-1,1,2,2,-1,8,1,7,1,1,20752,-1,1,39448,33225,984,17332,53283,
35730,23945,13948,22252,52005,13703,8621,27776,33318,2740,
4472,36959,17038,55127,16460,26669,39430,1,0,1,4,20769,16570,
58474,30131,770,4,25421,22569,51508,59396,10568,4,20769,16570,
58474,30131,770,8,531,55309,40895,38056,34677,30870,397,59131,
12756,3,13601,54878,13783,39334,3,41605,59081,10842,15125,
3,45764,5312,9992,25318,3,59301,18015,3739,13650,3,23540,44673,
45053,33398,3,4675,39636,45179,40604,3,49815,29818,2643,16065,
3,46787,46548,12505,53510,3,10439,37666,18998,32189,3,38967,
14338,31161,12779,3,27030,21461,12907,22939,3,24657,32725,
47756,22305,3,44226,9911,59256,54610,3,56240,51924,26856,52915,
3,16133,61189,17015,39397,3,24483,12048,40057,21323

Serialization of **checkpoint** during sparse interpolation

28, 14, 9, 64017, 31343, 5117, 64185, 47755, 27377, 25604, 6323, 41969, 14, 3, 4, 0, 0, 3, 4, 0, 1, 3, 4, 0, 2, 3, 4, 0, 3, 3, 4, 0, 4, 3, 4, 1, 0, 3, 4, 1, 1, 3, 4, 1, 2, 3, 4, 1, 3, 3, 4, 2, 0, 3, 4, 2, 1, 3, 4, 2, 2, 3, 4, 3, 0, 3, 4, 3, 1, 14, 59877, 1764, 59012, 44468, 1, 19485, 25871, 3356, 2, 58834, 49014, 65518, 15714, 65520, 1, 2, 4, 4, 1, 1

Numerical	*Randomized (Monte Carlo)*
more efficiency, but approximate result	more efficiency, but uncertain result
ill-conditionedness near singular inputs	unfavorable inputs: pseudo-primes, $\sum_i \prod_j (x_i - j)$, Coppersmith's "pathological" matrices
convergence analysis	probabilistic analysis
try algorithms on unproven inputs	try algorithms with limited randomness

Numerical + *randomized*, e.g., approximate factorizer:
all of the above (?)

Hallmarks of a good heuristic

- Is algorithmic in nature, i.e., always terminates with a result of possibly unknown validity

Hallmarks of a good heuristic

- Is algorithmic in nature, i.e., always terminates with a result of possibly unknown validity

- Is a proven complete solution in a more stringent setting, for example, by restricting the inputs or by slowing the algorithm

Hallmarks of a good heuristic

- Is algorithmic in nature, i.e., always terminates with a result of possibly unknown validity

- Is a proven complete solution in a more stringent setting, for example, by restricting the inputs or by slowing the algorithm

- Has an experimental track record, for example, works on 50% of cases

Ron Rivest's lessons learned (2002 Turing Award lecture)

- Try to solve "real-world" problems
- Moore's law (#transistors/in^2 doubles every year) matters
- Theory matters
- Organizations matter

Moore's law and asymptotically fast algorithms

Strassen matrix multiplication
Knuth/Schönhage half GCD
baby-steps/giant-steps polynomial factorization
Tellegen's principle,. . .

are practical on today's problem sizes

Moore's law and asymptotically fast algorithms

Strassen matrix multiplication
Knuth/Schönhage half GCD
baby-steps/giant-steps polynomial factorization
Tellegen's principle,...

are practical on today's problem sizes

Abstract model vs. actual computer:
PRAM has $\Omega(\sqrt[3]{n})$ memory access time [D. Bernstein]
Turing algorithms have no cache faults [A. Schönhage]
$O(\log\log n)$ factors are model specific [E. Kaltofen]

Organizations

WRI and Maplesoft
ACA/ISSAC
JSC/AAECC
SIGSAM/Fachgruppe/JSSAC
W3C MathML committee

Real-world problems

Macdonald's Microwave antenna system: 4 equations in 4 variables of total degree 6 and 146 terms in total

The many medium size computations are real world applications

有難う