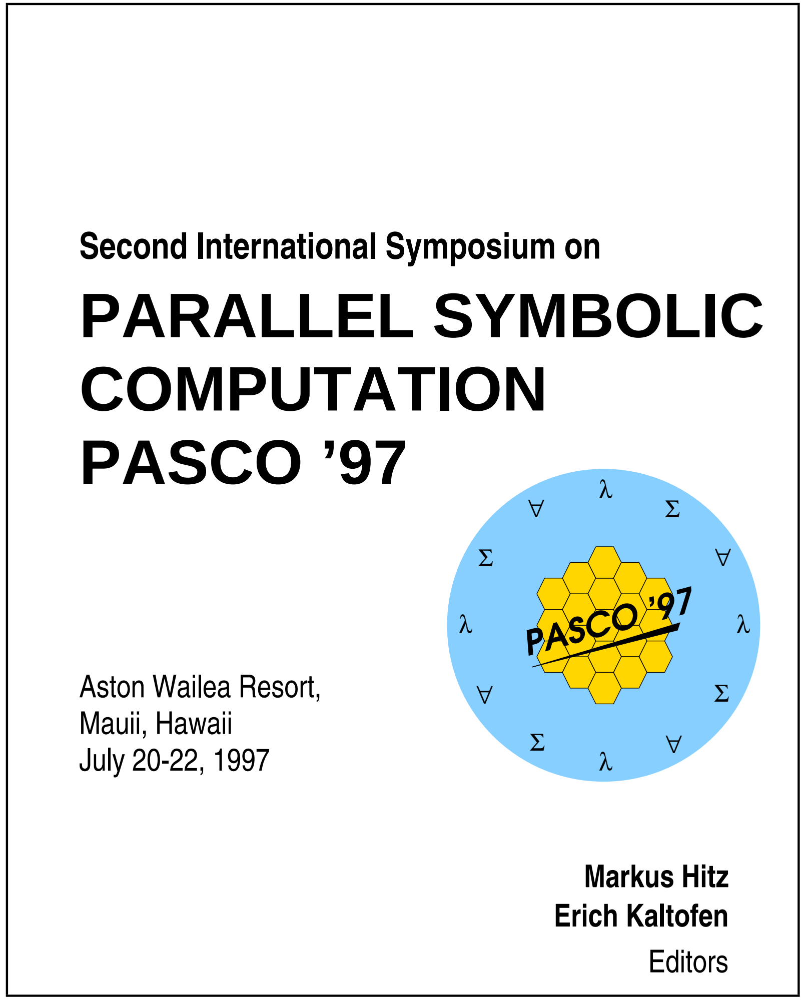
Fifteen years after DSC and WLSS2
What parallel computations I do today

Erich L. Kaltofen
North Carolina State University
`google->kaltofen`

The PASCO/Computer Algebra and Parallelism (CAP) Conference Series

1. 1988 CAP, Grenoble, France
2. 1990 CAP, Ithaca, USA
3. 1994 Hagenberg, Austria
4. 1997 Wailea, USA
5. 2007 London, Canada
6. 2010 Grenoble, France

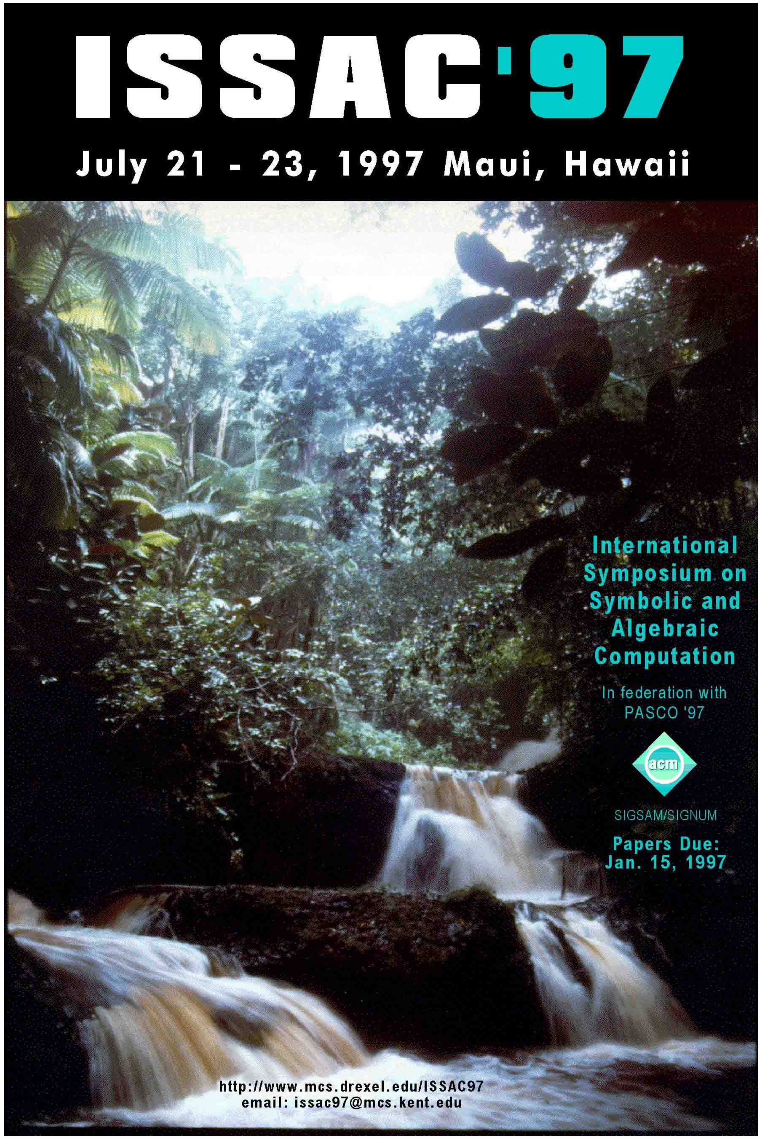
Who attended all 6?



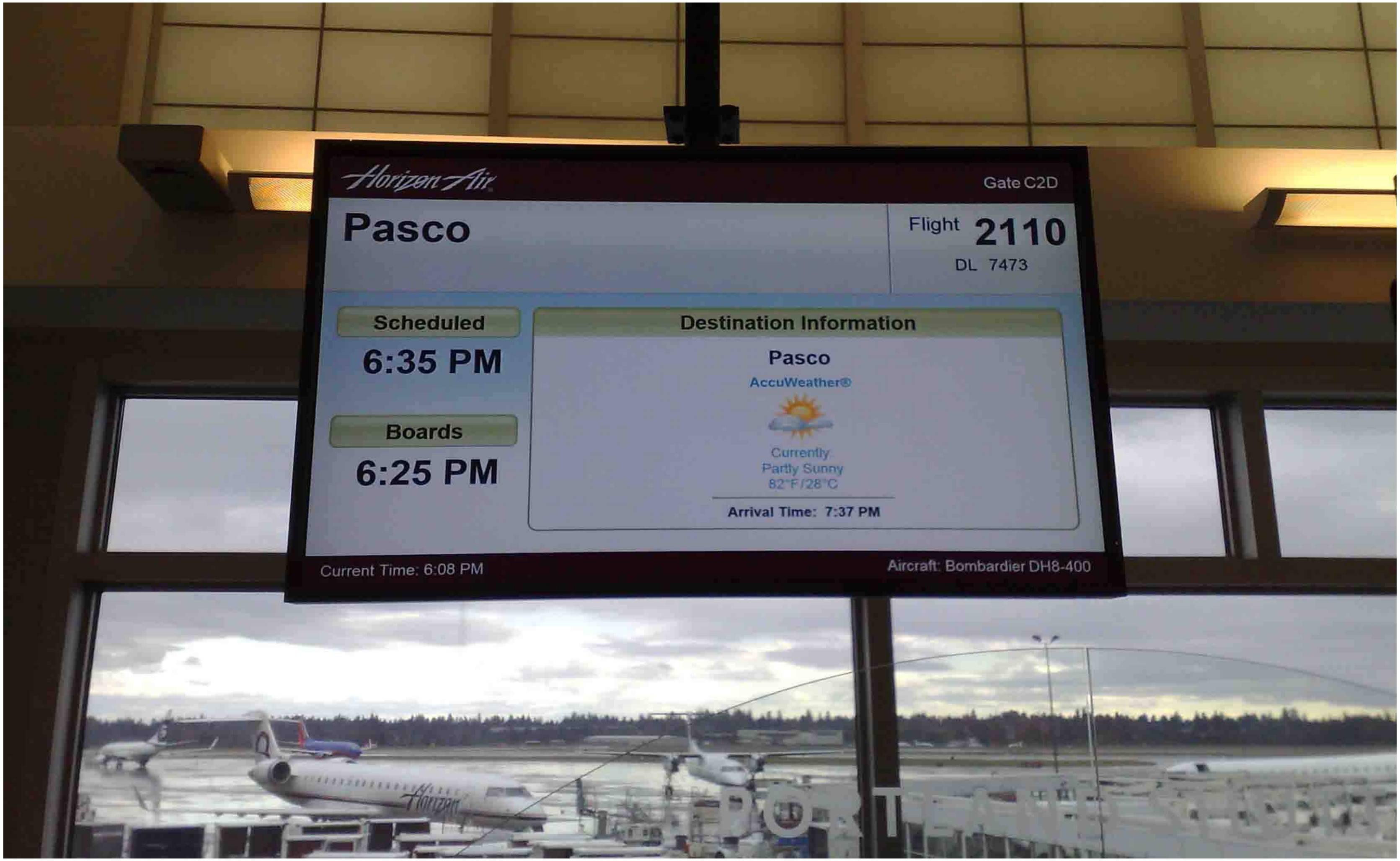
The PASCO/Computer Algebra and Parallelism (CAP) Conference Series

1. 1988 CAP, Grenoble, France
2. 1990 CAP, Ithaca, USA
3. 1994 Hagenberg, Austria
4. 1997 Wailea, USA
5. 2007 London, Canada
6. 2010 Grenoble, France

Who attended all 6?



Seattle-Tacoma Departure to Pasco



Outline

– Reminiscences

– Supersparse interpolation

– Today’s parallel hardware

– Interactive symbolic supercomputing

Shallow circuits

– ABSOLUTELY IRREDUCIBILITY $\in \mathscr{NC}$
  [Kaltofen ’85 JSC vol. 1, nr. 1]

Shallow circuits

- ABSOLUTELY IRREDUCIBILITY $\in \mathscr{NC}$
  [Kaltofen ’85 JSC vol. 1, nr. 1]

- $\mathsf{K}[x]^{n\times n}$-SMITH FORM $\in \mathscr{RNC}$
  [Kaltofen, Krishnamoorthy, Saunders SYMSAC’86 and EUROCAL’87]: Mr. Smith goes to Las Vegas

Shallow circuits

– ABSOLUTELY IRREDUCIBILITY $\in \mathscr{NC}$
[Kaltofen ’85 JSC vol. 1, nr. 1]

– $\mathsf{K}[x]^{n\times n}$-SMITH FORM $\in \mathscr{RNC}$
[Kaltofen, Krishnamoorthy, Saunders SYMSAC’86 and EUROCAL’87]: Mr. Smith goes to Las Vegas

– Parallelization of straight-line polynomials [Miller, Ramachandran, Kaltofen AWOC’86] and straight-line rational functions [Kaltofen STOC’86]

Shallow circuits

- ABSOLUTELY IRREDUCIBILITY $\in \mathscr{NC}$
  [Kaltofen ’85 JSC vol. 1, nr. 1]

- $\mathsf{K}[x]^{n \times n}$-SMITH FORM $\in \mathscr{RNC}$
  [Kaltofen, Krishnamoorthy, Saunders SYMSAC’86 and EUROCAL’87]: Mr. Smith goes to Las Vegas

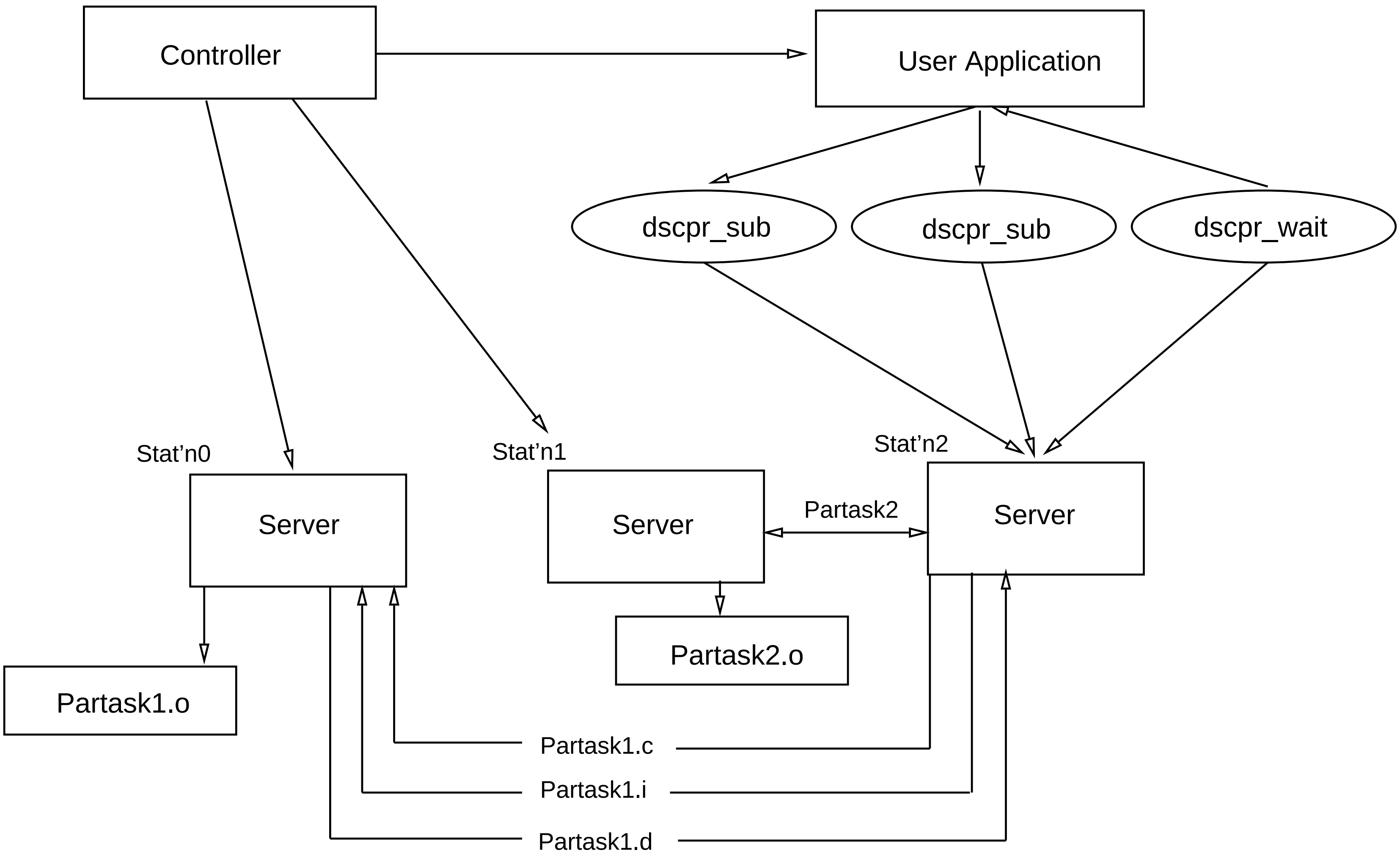
- Parallelization of straight-line polynomials [Miller, Ramachandran, Kaltofen AWOC’86] and straight-line rational functions [Kaltofen STOC’86]

- Processor-efficient parallel linear system solving [Kaltofen and Pan SPAA’91 and FOCS’92]
  Includes parallelization of automatic differentiation

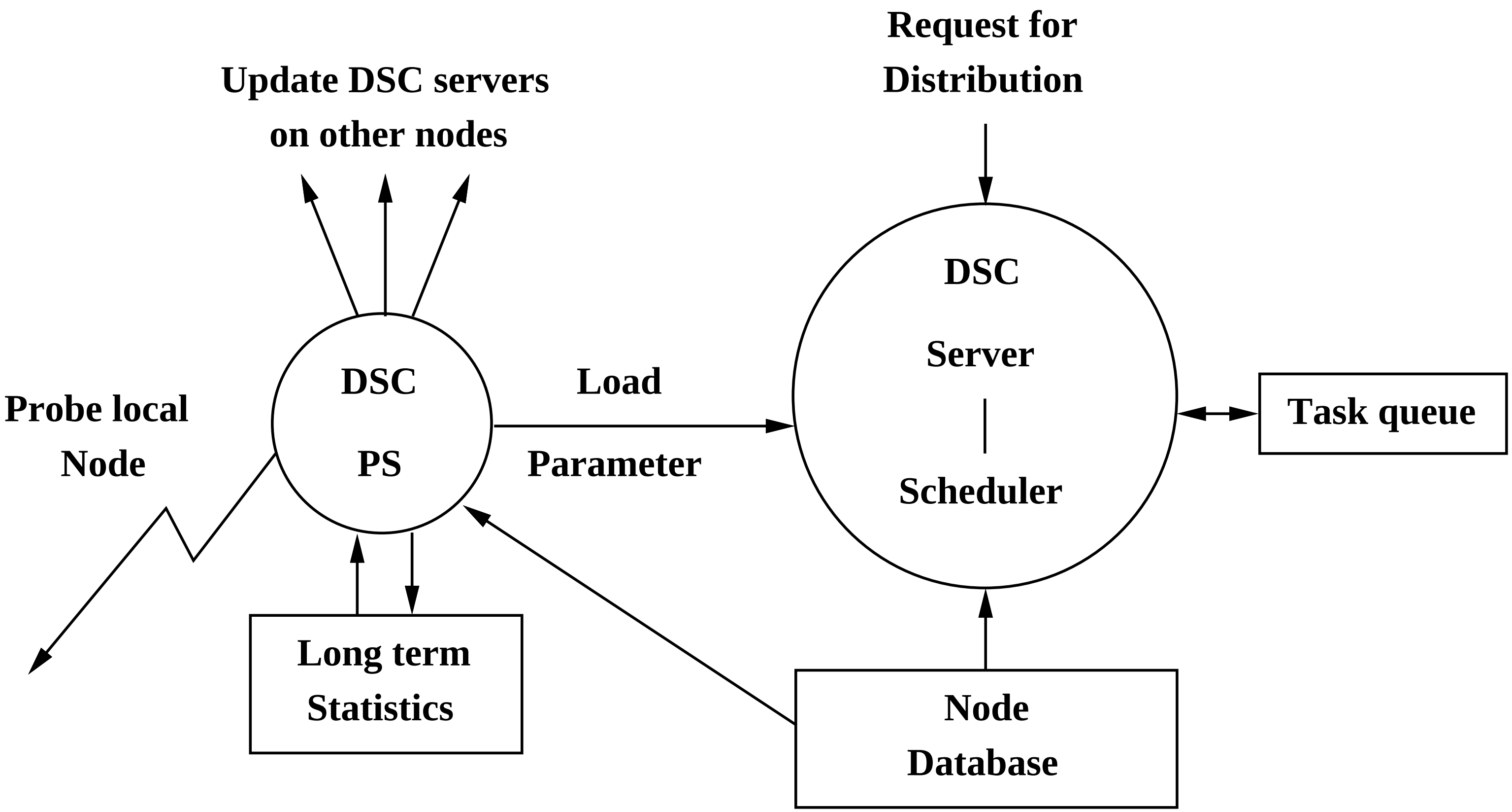
An Old Email To Ph.D. Student John F. Canny

```
From kaltofen Mon May  4 15:15:18 1987
Received: by csv.rpi.edu (5.54/1.14)
To: jfc@oz.ai.mit.edu
Subject: Your question
Dear John:
   Several articles have been written since my remark.
First, if the roots are all real, they can be isolated and
approximated in O(log(n)^2) parallel depth [Proc. STOC 1986,
340-349].
... I was told by David and Gregory Chudnovsky (8/86)
that the homotopy methods [Math. Programming 16, 159-176 (1979)]
accomplish the same for the complex root problem. ...
   Your ideas on using parallel computation to get the Tarski
problem into polynomial time have been persued as well.  There
is a paper by Ben-Or, Kozen, and Reif in a recent issue
of J. Comput. System Sci.; also one in [Proc. FOCS 1985, 515-521]
covers that area.
   A promising line of research, I think, is to investigate
parallel homotopy methods as mentioned above.  ...
                                                          Erich
```

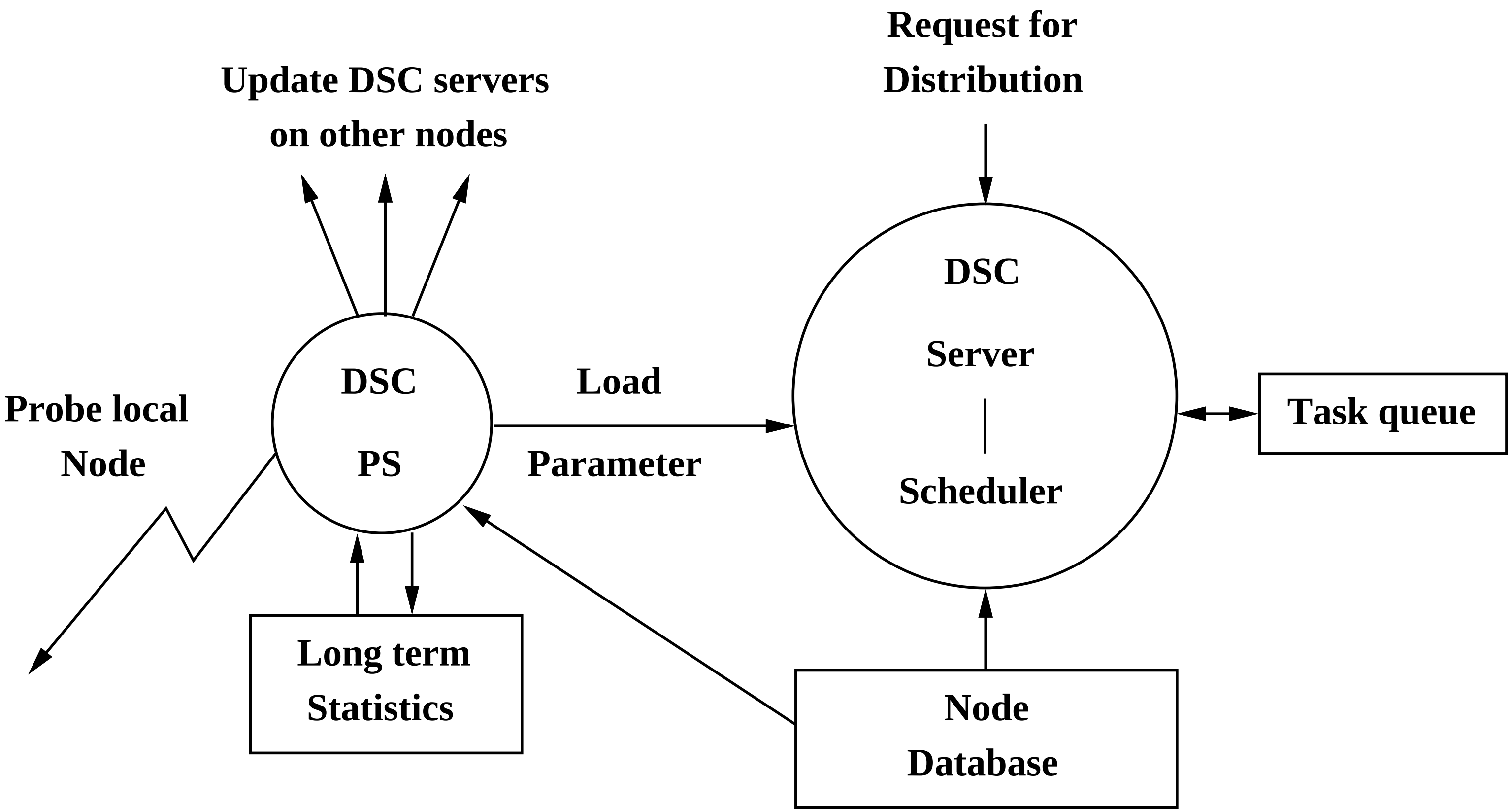
Distributed Symbolic Computation (DSC) Tool [Diaz, Kaltofen, Schmitz, Valente, ISSAC 1991]



DSC Predictive (ARIMA) Scheduler [Hitz et al. DISCO'93, Kaltofen and Samadani PODC'95]

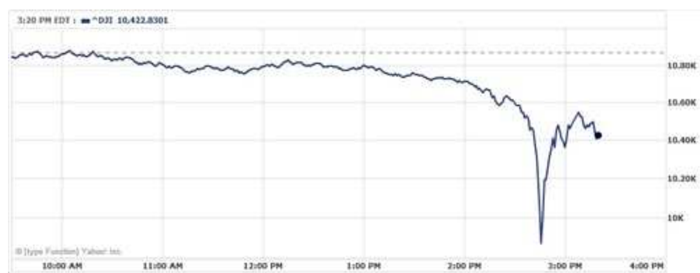


DSC Predictive (ARIMA) Scheduler
[Hitz et al. DISCO'93, Kaltofen and Samadani PODC'95]



DJIA May 6, 2010:

Multiple prediction programs?



Goldwasser-Kilian-Atkin Primality Prover in DSC

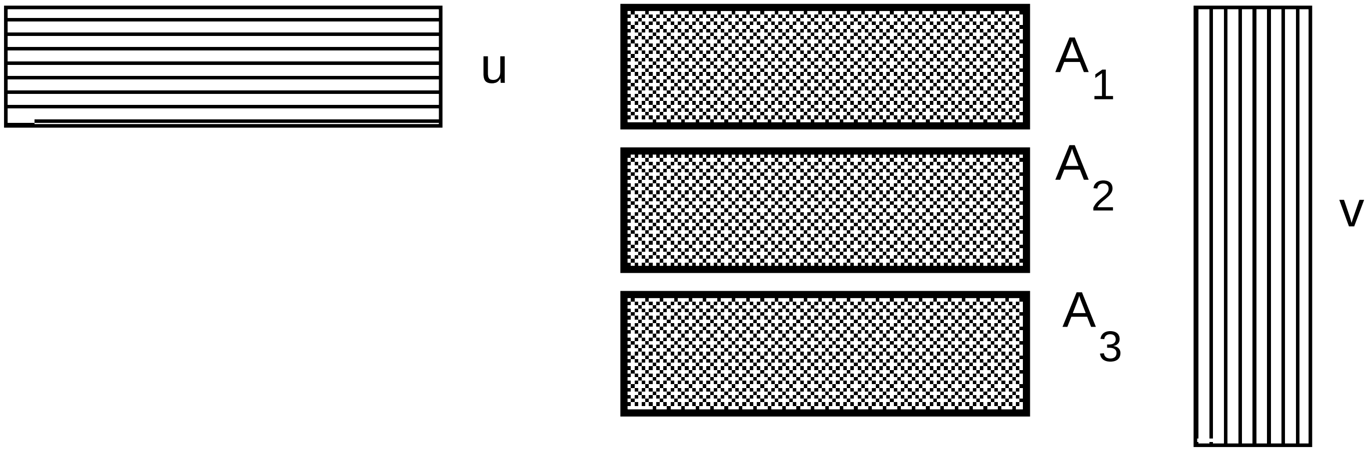
Digits	# parallel subtasks 1	2	4	8	16
22	0:52(2)	0:54(5)	0:36(3)	1:22(5)	1:18(2)
	0:45(2)	0:54(5)	0:44(3)	1:00(3)	2:17(3)
	0:50(2)	0:54(5)	0:46(3)	1:08(3)	1:12(3)
	0:52(2)	0:54(5)	0:47(3)	0:51(3)	3:08(5)
43	31:37(7)*	7:43(6)	1:27(5)	3:07(9)	7:09(7)
	31:08(7)*	7:21(7)	2:40(8)	3:31(6)	7:47(9)
	33:39(7)*	7:20(7)	2:39(8)	2:25(7)	6:03(11)
		7:34(7)	2:46(8)	2:55(8)	4:13(5)
99			38:28(18)	43:53(19)	25:21(19)
			44:38(21)	31:51(16)	27:27(21)
			68:14(17)	32:15(19)	23:31(21)
			46:19(21)	30:58(15)	27:22(21)

Timings for GKA (first stage, first decent strategy).

Times shown are minutes:seconds (# of descents).

* First descent with 110th discriminant on list.

Wiedemann Linear System Solver for the IBM SP-2 (WLSS2) [Kaltofen and Lobo HPC'96]



N	Grain	Sequence	Minpoly	Evaluation	Total
52,250	32	$0^h10'$	$0^h42'$	$0^h04'$	$0^h57'$
252,222	32	$7^h15'$	$15^h24'$	$3^h51'$	$26^h30'$

Parallel CPU Time (hours^{h}minutes$'$) for finding 128 solutions with optimized WLSS2 package on 4 nodes of an SP-2 multiprocessor.

Ben-Or/Tiwari '88 Interpolation

Main idea: Let $f(y) = c_1 y^{e_1} + \cdots + c_t y^{e_t} \in \mathsf{K}[y]$
$g \in \mathsf{K}$ with $g^{e_i} \neq g^{e_j}$ for $i \neq j$

Then $a_k = f(g^k)$, $k = 0, 1, \ldots$ is minimally linearly generated by

$$\Lambda(z) = (z - g^{e_1}) \cdots (x - g^{e_t})$$

Ben-Or/Tiwari ’88 Interpolation

Main idea: Let $f(y) = c_1 y^{e_1} + \cdots + c_t y^{e_t} \in \mathsf{K}[y]$
$g \in \mathsf{K}$ with $g^{e_i} \neq g^{e_j}$ for $i \neq j$

Then $a_k = f(g^k),\ k = 0, 1, \ldots$ is minimally linearly generated by

$$\Lambda(z) = (z - g^{e_1}) \cdots (x - g^{e_t})$$

My 1988 manuscript fragment on deterministic rational function recovery:

> “In order to avoid large intermediate integral results one can work modulo a sufficiently large prime $\bar{p}$ such that $\bar{p} - 1$ is smooth. Then discrete logarithms can be found efficiently and ...”

1988 Remark cont.: Kronecker substitution

For $F(x_1,\ldots,x_n) \in \mathsf{K}[x_1,\ldots,x_n]$ interpolate

$$f(y) = F(y, y^{d_1+1}, y^{(d_1+1)(d_2+1)}, \ldots), d_j \geq \deg_{x_j}(F)$$

1988 Remark cont.: Kronecker substitution

For $F(x_1,\ldots,x_n) \in \mathsf{K}[x_1,\ldots,x_n]$ interpolate

$$f(y) = F(y, y^{d_1+1}, y^{(d_1+1)(d_2+1)}, \ldots), d_j \ge \deg_{x_j}(F)$$

$p \ge (d_{\max}+1)^n \quad \Longrightarrow \log(p) = n\, \mathbf{log}(d_{\max})$

$\Longrightarrow$ algorithm is polynomial in $\log(\deg F)$

$\Longrightarrow$ term degrees of, e.g., 2^{500} allowed

Supersparse (lacunary) interpolation in 1988!

Pohlig-Hellman 1978 Primes Via Dirichlet

$p = \mu Q + 1$ is prime for:

$\mu < \infty$ [Dirichlet 1837]

$\mu = O(Q^{L-1})$, L constant [Linnik 1944]

$\mu = O(Q^{4.5})$ [Heath-Brown 1992]

$\mu = O(Q\,(\log Q)^2)$ under GRH

$\mu = O((\log Q)^2)$ conjectured in [Heath-Brown 1979]:
$L = 2$ "may presumably be reduced to $[P(n)] \ll n\,(\log n)^2$"

Pohlig-Hellman 1978 Primes Via Dirichlet

$p = \mu Q + 1$ is prime for:

$\mu < \infty$ [Dirichlet 1837]

$\mu = O(Q^{L-1})$, L constant [Linnik 1944]

$\mu = O(Q^{4.5})$ [Heath-Brown 1992]

$\mu = O(Q\,(\log Q)^2)$ under GRH

$\mu = O((\log Q)^2)$ conjectured in [Heath-Brown 1979]:
$L = 2$ “may presumably be reduced to $[P(n)] \ll n\,(\log n)^2$”

Example: $37084 = \max_{m \le 12300} (\operatorname{argmin}_{\mu \ge 1}(\mu 2^m + 1 \text{ is prime}))$.

Linear Generators Via Block Generators

Example: $t = 18$

Compute minimal matrix generator $\Gamma(z) \in \mathsf{K}[z]^{2\times 2}$ of

$$\begin{bmatrix} a_i & a_{9+i} \\ a_{9+i} & a_{18+i} \end{bmatrix}, \quad i = 0, 1, \ldots, 18.$$

by matrix Berlekamp/Massey algorithm.

Then $\Lambda(z) = \det(\Gamma(z))$

Linear Generators Via Block Generators

Example: $t = 18$

Compute minimal matrix generator $\Gamma(z) \in \mathsf{K}[z]^{2\times 2}$ of

$$\begin{bmatrix} a_i & a_{9+i} \\ a_{9+i} & a_{18+i} \end{bmatrix}, \quad i = 0, 1, \ldots, 18.$$

by matrix Berlekamp/Massey algorithm.

Then $\Lambda(z) = \det(\Gamma(z))$

Why? 1. if $f(g^{k+1})$ is easier to compute with $f(g^k)$
("giant-steps/baby steps")

2. Locality, locality, locality [Moreno Maza 2010]

Siegfried Rump's 2006 Model Problem

For $n = 1, 2, 3, \ldots$ compute the global minimum μ_n:

$$\mu_n = \min_{P,Q} \frac{\|PQ\|_2^2}{\|P\|_2^2 \|Q\|_2^2} \qquad \text{(rational function)}$$

$$\text{s. t. } P(Z) = \sum_{i=1}^{n} p_i Z^{i-1}, Q(Z) = \sum_{i=1}^{n} q_i Z^{i-1} \in \mathbb{R}[Z] \setminus \{0\}$$

Siegfried Rump's 2006 Model Problem

For $n = 1, 2, 3, \ldots$ compute the global minimum μ_n:

$$\mu_n = \min_{P,Q} \frac{\|PQ\|_2^2}{\|P\|_2^2 \|Q\|_2^2} \qquad \text{(rational function)}$$

$$\text{s. t. } P(Z) = \sum_{i=1}^{n} p_i Z^{i-1}, Q(Z) = \sum_{i=1}^{n} q_i Z^{i-1} \in \mathbb{R}[Z] \setminus \{0\}$$

$$\Updownarrow$$

$$\mu_n = \min_{P,Q} \|PQ\|_2^2$$

$$\text{s. t. } \|P\|_2 = \|Q\|_2 = 1, \deg(P) \leq n-1, \deg(Q) \leq n-1$$

Local Minimum By Lagrangian Multipliers

Siegfried Rump's 2006 Model Problem

For $n = 1, 2, 3, \ldots$ compute the global minimum μ_n :

$$\mu_n = \min_{P,Q} \frac{\|PQ\|_2^2}{\|P\|_2^2 \|Q\|_2^2} \qquad \text{(rational function)}$$

$$\text{s. t. } P(Z) = \sum_{i=1}^{n} p_i Z^{i-1}, Q(Z) = \sum_{i=1}^{n} q_i Z^{i-1} \in \mathbb{R}[Z] \setminus \{0\}$$

$$\Updownarrow$$

$$\frac{1}{\mu_n} = \max_{P,Q} B_{n-1}$$

$$\text{s. t. } \|P(Z)\|_2^2 \cdot \|Q(Z)\|_2^2 = B_{n-1} \|P(Z) \cdot Q(Z)\|_2^2$$

$$P, Q \in \mathbb{R}[Z] \setminus \{0\}, \deg(P) \leq n-1, \deg(Q) \leq n-1$$

Mignotte's factor coefficient bound: $\frac{1}{\mu_n} \leq \binom{2n-2}{n-1}^2$

Rational Function Detail

Minimize the rational function $\frac{f(\mathbf{X})}{g(\mathbf{X})}$ with

$$f(\mathbf{X}) = \|PQ\|_2^2 = \sum_{k=2}^{2n} \Big(\sum_{i+j=k} p_i q_j \Big)^2,$$

$$g(\mathbf{X}) = \|P\|_2^2 \|Q\|_2^2 = \Big(\sum_{i=1}^{n} p_i^2\Big)\Big(\sum_{j=1}^{n} q_j^2\Big)$$

where

$$\mathbf{X} = \{p_1, \ldots, p_{\lceil n/2 \rceil}\} \cup \{q_1, \ldots, q_{\lceil n/2 \rceil}\},$$

because P, Q achieving μ_n must be symmetric or skew-symmetric

[Rump and Sekigawa 2006]

Rational Function Optimization By Sparse SOS

A (positive) lower bound of $\mu_n = \min \dfrac{f}{g}$, g positive, is obtained by solving the sparse block semidefinite program:

$$
\begin{aligned}
\mu_n^* := \sup_{r\in\mathbb{R},W} \quad & r \\
\text{s. t.} \quad & f(\mathbf{X}) = m_{\mathscr{G}}(\mathbf{X})^T \cdot W \cdot m_{\mathscr{G}}(\mathbf{X}) + rg(\mathbf{X}) \\
& \left(f(\xi_1,\ldots,\xi_n) = \mathrm{SOS} + rg(\xi_1,\ldots,\xi_n) \ge rg(\xi_1,\ldots,\xi_n)\right) \\
& W \succeq 0, W^T = W, \quad r \ge 0
\end{aligned}
$$

where $m_{\mathscr{G}}(\mathbf{X})$ is the term vector restricted to $p_i q_j$

For $n = 14$: $W \in \mathbb{R}^{49\times 49}$, 784 equality constraints

Certified Rump Model Lower Bounds [with Bin Li, Zhengfeng Yang, Lihong Zhi 2009]

n	k	#iter	prec.	secs/iter	lower bound r_n	relative Δ_n	$\Delta_n^{[ISSAC'08]}$	#sq	logH
7	1	60	10×15	0.27	3.418506980e−05	2.048e−14	2.018e−14	16	2485
8	2	80	6×15	0.24	3.905435600e−06	2.561e−15	7.681e−11	16	1563
9	1	280	10×15	1.75	4.360016539e−07	3.784e−14	6.881e−08	25	3919
10	2	280	12×15	1.89	4.783939568e−08	4.517e−13	8.361e−07	25	4660
11	1	510	13×15	9.62	5.178700000e−09	9.481e−06	1.931e−04	36	7201
12	2	210	5×15	8.79	5.545390000e−10	8.869e−05	5.439e−03	36	2881
13	1	270	5×15	41.93	5.881019273e−11	9.639e−04	1.728e−02	49	4271
14	2	440	25×15	33.68	6.100000000e−12	1.679e−02	9.368e−01	49	3121
15	1	1070	25×15	162.84	6.000000000e−13	8.239e−02	—	64	5751
16	2	640	25×15	153.94	6.000000000e−14	1.273e−01	—	64	5312
17	1	1650	10×15	**504.10**	1.000000000e−15	6.011e+00	—	81	12984
17	1	4200	10×15	380.75	6.000000000e−15	1.685e−01	—	81	13029
18	2	6440	10×15	344.75	1.000000000e−16	6.238e+00	—	81	12570
18	2	8800	10×15	352.62	3.000000000e−16	1.413e+00	—	81	12571
18	2	26800	10×15	330.36	7.000000000e−16	3.406e−02	—	81	12578

2007 MacPro: 4 Cores, 1 Memory Bus = Bus Contention

2009 MacPro: 16 Cores, “Nehalem” Bus = Less Contention

Contention Reduction By Software?

Google's *TCmalloc* [see paper by Dumas, Gautier, Roch]

Garbage collection vs. TCmalloc
Since the Lisp Machine days, I became a sceptic of GC for large computations (one reason why no Java LinBox)

Contention Reduction By Software?

Google's *TCmalloc* [see paper by Dumas, Gautier, Roch]

Garbage collection vs. TCmalloc
Since the Lisp Machine days, I became a sceptic of GC for large computations (one reason why no Java LinBox)

Moreno Maza: Could TCmalloc be used to deal with memory contention when several Maple sessions are to be run concurrently on a multicore?
Darin Ohashi: Interesting. When running multiple sessions, each session has its own independent kernel.

MMM: Since these Maple sessions are running the same algorithm with different input, would the Thread Package be appropriate instead?
DO: The problem could be that some library routines may not be thread safe.

Nonlinear Diophantine Optimization

- Maximization of the **single factor coefficient bound** for integer polynomials w.r.t. infinity norm.

$$c_n = \max_{F,G} \chi_n$$

$$\text{s. t.} \quad \min(\|F(z)\|_\infty, \|G(z)\|_\infty) = \chi_n \|F(z)\cdot G(z)\|_\infty$$

$$F, G \in \mathbb{Z}[z] \textit{ irreducible}, \deg(F) + \deg(G) = n$$

Fact: $\forall n:\ c_n$ is a rational number.
Conjecture $\lim_{n\to\infty} c_n = \infty$

Nonlinear Diophantine Optimization

- Maximization of the **single factor coefficient bound** for integer polynomials w.r.t. infinity norm.

$$c_n = \max_{F,G} \chi_n$$

$$\text{s. t.} \quad \min(\|F(z)\|_\infty, \|G(z)\|_\infty) = \chi_n \|F(z) \cdot G(z)\|_\infty$$

$$F, G \in \mathbb{Z}[z] \textit{ irreducible}, \deg(F) + \deg(G) = n$$

Fact: $\forall n : c_n$ is a rational number.
Conjecture $\lim_{n\to\infty} c_n = \infty$

- Boyd's 1997 bound $\geq$ 2.7339
  Collins's 2004 bound $\geq$ 2.2005
  Our 2008 record $\geq$ 3.4334
  Abbott's 2005/'09 bound $\geq$ 13.7500

Lehmer's Mahler measure problem

$$f(z) = a \cdot \prod_{i=1}^{n} (z - \alpha_i) \in \mathbb{Z}[z],\ \alpha_i \in \mathbb{C},$$

$$M(f) = |a| \cdot \prod_{i=1}^{n} \max\{|\alpha_i|, 1\}$$ (the Mahler measure).

For Lehmer's 1933 polynomial we have

$L(z) = z^{10} + z^9 - z^7 - z^6 - z^5 - z^4 - z^3 + z + 1$

$M(L) = 1.1762808\ldots$

Is there a polynomial in $\mathbb{Z}[z]$ with $1 < M(f) < M(L)$?

Michael Mossinghoff's Top 100

MM's	deg	Mahler measure	count
1.	10	1.176280818260	2248
2.	18	1.188368147508	—
3.	14	1.200026523987	1
4.	18	1.201396186235	8804
5.	14	1.202616743689	105
6.	22	1.205019854225	10
7.	28	1.207950028412	—
8.	20	1.212824180989	4
9.	20	1.214995700776	—
10.	10	1.216391661138	198
11.	20	1.218396362520	1598
12.	24	1.218855150304	—
13.	24	1.219057507826	—
14.	18	1.219446875941	—
15.	18	1.219720859040	—
16.	34	1.220287441693	—
17.	38	1.223447381419	—
18.	26	1.223777454948	—
19.	16	1.224278907222	1779
20.	18	1.225503424104	35
21.	30	1.225619851977	—
22.	30	1.225810532354	—
23.	26	1.226092894512	17
24.	36	1.226493301473	—
25.	20	1.226993758166	194

MM's	deg	Mahler measure	count
26.	12	1.227785558695	77
27.	30	1.228140772740	—
28.	36	1.229482810173	—
29.	22	1.229566456617	1
30.	34	1.229999039697	—
31.	38	1.230263271363	—
32.	42	1.230295468643	—
33.	10	1.230391434407	27995
34.	46	1.230743009076	—
35.	18	1.231342769993	—
36.	48	1.232202952743	—
37.	20	1.232613548593	133
38.	28	1.232628775929	—
39.	38	1.233672001767	—
40.	52	1.234348374876	—
41.	24	1.234443834873	—
42.	26	1.234500336789	—
43.	16	1.235256705642	72
44.	46	1.235496042193	—
45.	22	1.235664580390	—
46.	42	1.235761099712	—
47.	32	1.236083368052	—
48.	32	1.236198469859	—
49.	32	1.236227922245	—
50.	40	1.236249557349	—

MM's	deg	Mahler measure	count
51.	16	1.236317931803	8
52.	54	1.236566917569	—
53.	34	1.236579223637	—
54.	44	1.236674812187	—
55.	28	1.236808305865	—
56.	26	1.237504821217	—
57.	46	1.237634830280	—
58.	58	1.237684127894	—
59.	48	1.238040100176	—
60.	56	1.238431627359	—
61.	56	1.238708978554	—
62.	42	1.239505770490	—
63.	54	1.239747974875	—
64.	48	1.239861326360	—
65.	60	1.240061859037	—
66.	26	1.240254178706	—
67.	50	1.240379074717	—
68.	28	1.240699637594	—
69.	12	1.240726423653	471
70.	18	1.240770634960	17
71.	54	1.240983403460	—
72.	68	1.241372531335	—
73.	70	1.241422024928	—
74.	58	1.241541076676	—
75.	72	1.241788568356	—

MM's	deg	Mahler measure	count
76.	58	1.241902161200	—
77.	50	1.241974375265	—
78.	58	1.242217045566	—
79.	52	1.242362139933	—
80.	58	1.242610289442	—
81.	76	1.242775741593	—
82.	50	1.242878658278	—
83.	32	1.242940115199	—
84.	70	1.242979209676	—
85.	60	1.243027765980	—
86.	30	1.243128704866	—
87.	72	1.243210037398	—
88.	16	1.243477618690	38
89.	60	1.243486256656	—
90.	46	1.243564793293	—
91.	46	1.243682745689	—
92.	28	1.243878801656	—
93.	56	1.243935933206	—
94.	60	1.244271476892	—
95.	70	1.244273644932	—
96.	40	1.244414501983	—
97.	62	1.244598861428	—
98.	18	1.244617058976	1
99.	76	1.244729172444	—
100.	22	1.244802445450	—

With Leading Coefficient 2

EK's	deg	(Mahler measure)/2	count
1.	22	1.014537415604	3
2.	22	1.023188691065	3
3.	18	1.023467897775	3
4.	22	1.025607305528	2
5.	20	1.026063881400	18
6.	18	1.027106023710	61
7.	20	1.027522935540	1
8.	18	1.028643802187	16
9.	22	1.028656563641	1
10.	18	1.029862344505	22

Interactive Symbolic Supercomputing

Idea of Edelman's *Star-P* in Matlab of Microsoft's INTER*CTIVE: Use overloading for remote supercomputer execution. Sessions are **indistinguishable** from local sessions.

Proposal: overload Maple's *LinearAlgebra* package as a supercomputer LinBox.

Merci!