Constrained global optimization

Arnold Neumaier

University of Vienna

Vienna, Austria

In a **constraint satisfaction problem (CSP)**,
one asks about **existence**, and one or several examples:

Can 13 *balls of radius r touch a ball of radius R?*

For $r = R$, No! (van der Waerden, $\sim$ 1953)

In a **global optimization problem (GOP)**,
one asks about an **extremum**,
and a configuration where it is achieved.

What is the smallest possible R?

It is $R \approx 1.09r$.

In a **constraint projection problem (CPP)**,
one asks about **exhaustion** of the solution set,
and display of a suitable low-dimensional projection of it.

What is the set of possible (r, R)?

$\Sigma = \{(r, R) \mid R \geq 1.09r\}$ *(apart from roundoff).*

In a competitive world, only the best (safest, cheapest, . . .) is good enough. This is why optimization (and often global optimization) is very frequent in application.

Global optimization is one of the oldest of sciences, part of the art of successful living.

maximize	**service (or money? or happiness?)**
s.t.	**gifts and abilities**
	hopes and expectations (ours; others)
	bounded stress

Thousands of years of experience . . .

. . . resulted in the following algorithmic framework recommended by St. Paul (ca. 50 AD):

"Consider everything. Keep the good. Avoid evil whenever you recognize it." (*1 Thess. 5:21–22*)

In modern terms, this reads:

Do global search by branch and bound!

My personal global optimization problem is a never ending challenge:

"Be perfect, **as our father in heaven is perfect."** **(Jesus, ca. AD 30)**

On the mathematical level, the quest for perfection is

rigorous global optimization

... though if we can't do it we must be content with heuristics.

Why global optimization?

There are a number of problem classes where it is indispensable to do a complete search.

- **Hard feasibility problems (e.g., robot arm design), where local methods do not return useful information since they generally get stuck in local minimizers of the merit function, not providing feasible points**

- **Computer-assisted proofs (e.g., the recent proof of the Kepler conjecture by HALES), where inequalities must be established with mathematical guarantees**

- **Semi-infinite programming, where the optimal configurations usually involve global minimizers of auxiliary problems**

- **Safety verification problems, where treating nonglobal extrema as worst cases may severely underestimate the true risk**

- **Many problems in chemistry, where often only the global minimizer (of the free energy) corresponds to the situation matching reality**

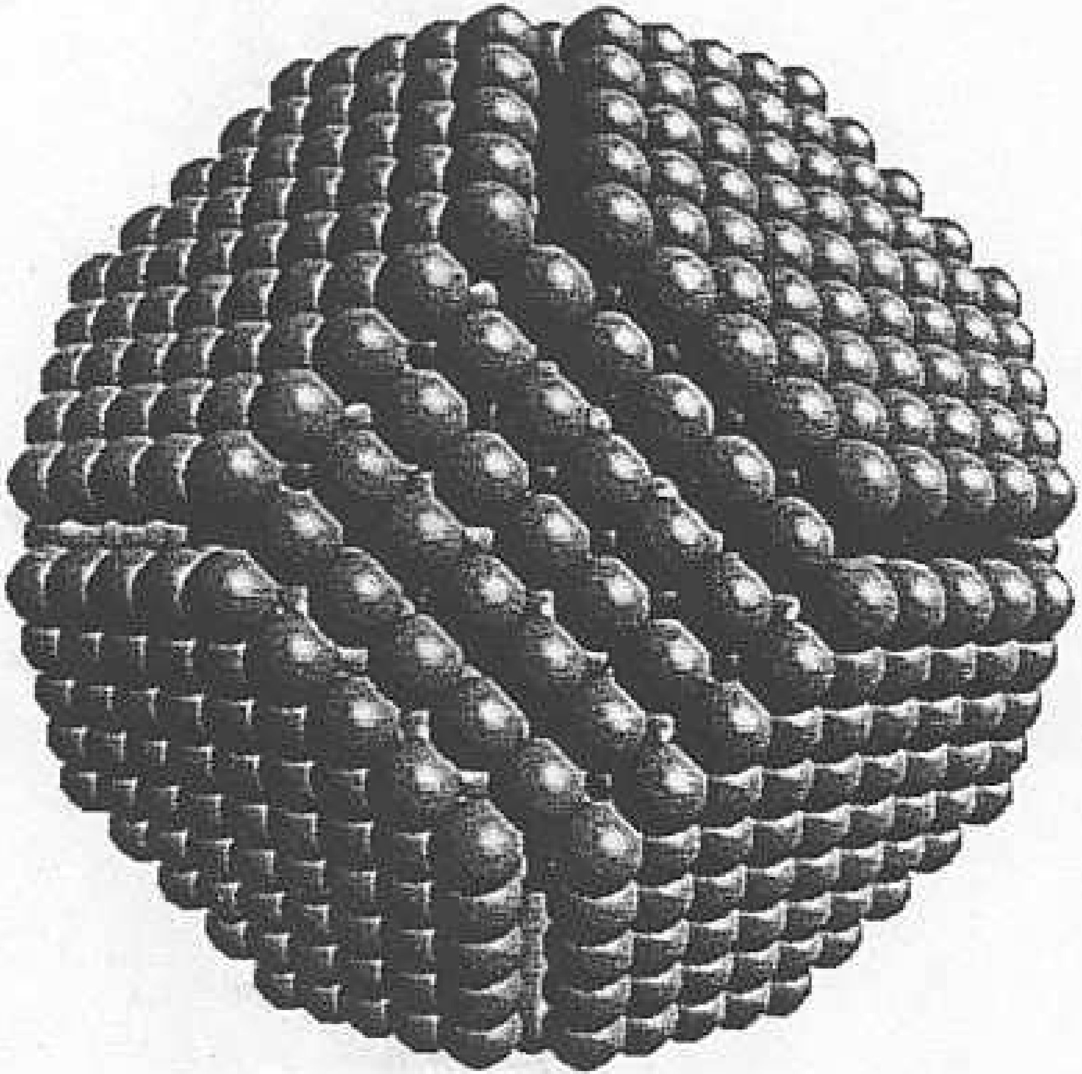

Figure 6. FCC configuration at $N = 2142$, lower in energy than the best MIC conformation of the same size.

This talk uses slides made for various occasions, including joint work with

- HERMANN SCHICHL **(Vienna, Austria)**
- OLEG SHCHERBINA **(Vienna, Austria)**
- WALTRAUD HUYER **(Vienna, Austria)**
- TAMAS VINKO **(Szeged, Hungary)**

within the COCONUT project

(COCONUT = Continuous Constraints – Updating the Technology)

`www.mat.univie.ac.at/coconut`

funded 2000–2004 by the European Community

Dichotomies

- **heuristic $\Leftrightarrow$ guaranteed**
  ... important for reliability
- **linear $\Leftrightarrow$ nonlinear**
  ... important for realistic modeling
- **local $\Leftrightarrow$ global**
  ... important for jumps in quality
- **black box $\Leftrightarrow$ structured access**
  ... important for efficient analysis

Degrees of rigor

- incomplete (heuristics; e.g., smoothing techniques)
- asymptotically complete (no means to know when a global minimizer has been found; e.g., multiple random start, pure branching)
- complete (knows after a finite time that an approximate global minimizer has been found to within prescribed tolerances, assuming exact calculation).
- rigorous (complete, with full rounding error control)

(Often, the label deterministic is used to characterize the last two categories of algorithms; however, this label is slightly confusing since many incomplete and asymptotically complete methods are deterministic, too.)

Complexity

Already in the case where only bound constraints are present, global optimization problems and constraint satisfaction problems are

- **undecidable on unbounded domains (Wenxing Zhu 2004), and**
- **NP-hard on bounded domains.**

This implies natural limits on the solvability of such problems in practice.

**In particular, methods which work with direct attack (analytic transformations without problem splitting) lead to transformed problems of exponentially increasing size,
while branch-and-bound methods split the problems into a worst case exponential number of subproblems.**

It is very remarkable that in spite of this, many large-scale problems can be solved efficiently.

This is achieved by carefully balancing the application of the available tools and using the internal structure which realistic problems always have.

The great success of the current generation of complete global solvers is due mainly to improvements in our ability to analyze global optimization problems mathematically.

For history, a thorough introduction, a comprehensive overview over the techniques in use, and extensive references see my survey

A. Neumaier **Complete search in continuous global optimization and constraint satisfaction,** **pp. 1-99 in: Acta Numerica 2004,** **Cambridge Univ. Press 2004.**

Complete search techniques I

a) Direct attack is feasible for polynomial systems of moderate degree (up to about 20)

- semidefinite relaxations
- Gröbner basis methods
- resultant-based techniques

Complete search techniques II

Branch-and-bound methods are the choice for larger problems. Basic approaches use

- **constraint propagation**
- **outer approximation (linear, convex, conic, semidefinite)**
- **DC (difference of convex function) techniques**
- **interval Newton and related methods**

Interval techniques became initially known mainly as a tool to control rounding errors. However, it is much less known – and much more important – that their real strength is the ability to control nonlinearities in a fully satisfying algorithmic way.

A full account of the theoretical background for interval techniques in finite dimensions is available in my book

A. Neumaier
Interval methods for systems of equations
Cambridge Univ. Press 1990.

The book is still up to date (with a few minor exceptions). While it is officially out of print, if you order it at Cambridge University Press, they'll print an extra copy especially for you. (Apparently, this is still profitable for them.)

Example: Nonlinear systems $F(x) = 0$

Newton's method approximates

$$0 = F(x) \approx F(z) + F'(z)(x - z) \quad \Rightarrow \quad x \approx z - F'(z)^{-1} F(z).$$

Interval slopes control errors by using instead

$$0 = F(x) = F(z) + F[x, z](x - z) \quad \Rightarrow \quad x = z - F[x, z]^{-1} F(z).$$

If we look for zeros in a box x and know

$$F[z, x] \in \mathbf{A} \quad \textbf{for all } x \in \mathbf{x}$$

with a regular interval matrix A (computable by automatic differentiation like techniques) then

$$x \in \mathbf{x}' := z - \mathbf{A}^{-1} F(z).$$

Properties of the computed box $\mathbf{x}'$:

$\mathbf{x}' \cap \mathbf{x} = \emptyset \quad \Rightarrow \quad$ **there is no zero in x**

$\mathbf{x}' \in \text{ int } \mathbf{x} \quad \Rightarrow \quad$ **there is a zero in x**

In any case, any zero in x is also in $\mathbf{x}' \cap \mathbf{x}$

Uniqueness proofs are also possible.

Similar properties hold for systems of equations and inequalities, and for optimization problems.

Complete search techniques III

Further efficiency is gained by

- **use of optimality conditions**
- **multiplier techniques (duality, Lagrangian relaxation)**
- **cut generation**
- **adding redundant constraints**
- **graph decomposition techniques**
- **certificates of optimality/infeasibility**

Necessary and sufficient global optimality conditions

We consider the polynomial optimization problem

$$\begin{aligned} &\min f(x) \\ &\text{s.t. } \; C(x) \geq 0, \quad F(x) = 0, \end{aligned} \tag{1}$$

with $f \in \mathbb{R}[x_{1:n}]$**,** $C \in \mathbb{R}[x_{1:n}]^m$**, and** $F \in \mathbb{R}[x_{1:n}]^r$**, where** $\mathbb{R}[x_{1:n}]$ **is the algebra of real polynomials in the variables** $x_1, \ldots, x_n$**.**

Write SOS **for the set of sums of squares of polynomials in** $x_1, \ldots, x_n$**. Write** B **for the vector with components** $B_0(x) = f(\hat{x}) - f(x)$ **and** $B_k = C_k$ **for** $k > 0$**, and** B_S **for the vector whose entries are the** 2^{m+1} **polynomials** $\prod_{i \in I} B_i$**, where** I **runs over all subsets of** $\{0 : m\}$**.**

Theorem (NEUMAIER & SCHICHL **2005)**

For a feasible point $\hat{x}$ of the polynomial optimization problem (1)**, the following are equivalent:**

(i) The point $\hat{x}$ is a global minimum of (1)**.**

(ii) There are a polynomial $y_0 \in SOS$, polynomial vectors $Y \in SOS^{2^{m+1}}$, $Z \in \mathbb{R}[x_{1:n}]^r$, and a positive integer e such that

$$B_0(x)^e + y_0(x) + Y(x)^T B_S(x) + Z(x)^T F(x) = 0 \tag{2}$$

identically in x.

Moreover, any solution of (2) **satisfies**

$$y_0(\hat{x}) = 0, \quad \inf\{Y(\hat{x}), B_S(\hat{x})\} = 0, \quad F(\hat{x}) = 0, \tag{3}$$

$$\delta_{e1} f'(\hat{x})^T = B_S{}'(\hat{x})^T Y(\hat{x}) + F'(\hat{x})^T Z(\hat{x}). \tag{4}$$

Complete search techniques IV

Efficiency and reliability also require the use of

- **local optimization for upper bounds**
- **clever box selection heuristics**
- **adaptive splitting heuristics**
- **reliable stopping criteria**
- **combination heuristics**
- **safeguarding techniques**

Benchmarking

Solver	Minos	LGO	BARON	ICOS	GlobSol
access language	GAMS	GAMS	GAMS	AMPL	Fortran90
optimization?	+	+	+	−	+
integer constraints	−	+	+	−	−
search bounds	−	required	recommended	−	required
black box eval.	+	+	−	−	−
complete	−	(−)	+	+	+
rigorous	−	−	−	+	+
local	++	+	+	+	(+)
CP	−	−	+	++	+
other interval	−	−	− +	++	++
convex/LP	−	−	++	+	−
dual	+	−	+	−	−
available	+	+	+	+	+
free	−	−	−	+	+

Solver	Premium Solver	LINGO Global	αBB	GloptiPoly	OQNLP
access language	Visual Basic	LINGO	MINOPT	Matlab	GAMS
optimization?	+	+	+	(+)	+
integer constraints	+	+	+	−	+
search bounds	+	−	?	−	+
black box eval.	−	−	−	−	+
complete	+	+	+	+	−
rigorous	(+)	−	−	−	−
local	+	+	+	−	+
CP	+	+	−	−	−
interval	++	+	+	−	−
convex	+	++	++	+	−
dual	−	+	−	++	−
available	+	+	−	+	+
free	−	−	−	+	−

The COCONUT test set

Number of variables	$1-9$	$10-99$	$100-999$	≥ 1000	
	size 1	size 2	size 3	size 4	total
Library 1 (GLOBALLIB from GAMS)	84	90	44	48	266
Library 2 (CUTE from Vanderbei)	347	100	93	187	727
Library 3 (CSP from EPFL)	225	76	22	6	329
total	656	266	159	241	1322

The test set actually used is slightly smaller, and we didn't test size 4.

Detailed test results are available on the COCONUT homepage (`www.mat.univie.ac.at/coconut`).

Reliability analysis for MINOS	
	global minimum found/accepted
size 1	**464/627 ≈ 74%**
size 2	**162/245 ≈ 66%**
size 3	**72/154 ≈ 47%**
all	**698/1026 ≈ 68%**
	claimed infeasible/accepted and feasible
size 1	**39/619 ≈ 6%**
size 2	**19/238 ≈ 8%**
size 3	**16/151 ≈ 11%**
all	**74/1008 ≈ 7%**

Reliability analysis for OQNLP	
	global minimum found/accepted
size 1	**551/619 ≈ 89%**
size 2	**207/242 ≈ 86%**
size 3	**91/130 ≈ 72%**
all	**847/993 ≈ 86%**
	claimed infeasible/accepted and feasible
size 1	**5/611 ≈ 1%**
size 2	**3/235 ≈ 1%**
size 3	**9/124 ≈ 8%**
all	**17/944 ≈ 2%**

Reliability analysis for BARON 7.2	
	global minimum found/accepted
size 1	**524/579 ≈ 91%**
size 2	**210/229 ≈ 92%**
size 3	**88/140 ≈ 63%**
all	**821/950 ≈ 86%**
	correctly claimed global/accepted
size 1	**450/579 ≈ 78%**
size 2	**151/229 ≈ 66%**
size 3	**43/140 ≈ 31%**
all	**644/950 ≈ 68%**
	wrongly claimed global/claimed global
size 1	**14/464 ≈ 3%**
size 2	**7/158 ≈ 4%**
size 3	**8/51 ≈ 16%**
all	**29/675 ≈ 4%**
	claimed infeasible/accepted and feasible
size 1	**3/571 ≈ 1%**
size 2	**1/222 ≈ 0%**
size 3	**0/128 = 0%**
all	**4/921 ≈ 0.4%**

Reliability analysis for LINGO9	
	global minimum found/accepted
size 1	**533/632 ≈ 84%**
size 2	**179/248 ≈ 72%**
size 3	**71/158 ≈ 45%**
all	**783/1038 ≈ 75%**
	correctly claimed global/accepted
size 1	**491/632 ≈ 78%**
size 2	**142/248 ≈ 57%**
size 3	**36/158 ≈ 23%**
all	**669/1038 ≈ 64%**
	wrongly claimed global/claimed global
size 1	**44/535 ≈ 8%**
size 2	**33/175 ≈ 19%**
size 3	**17/53 ≈ 32%**
all	**94/763 ≈ 12%**
	claimed infeasible/accepted and feasible
size 1	**1/624 ≈ 0%**
size 2	**1/241 ≈ 0%**
size 3	**1/143 ≈ 0%**
all	**3/1008 ≈ 0.3%**

Reliability analysis for ICOS (on pure CSPs only)	
	global minimum found/accepted
size 1	**145/207 ≈ 70%**
size 2	**34/63 ≈ 54%**
size 3	**5/20 ≈ 25%**
all	**184/290 ≈ 63%**
	correctly claimed global/accepted
size 1	**68/207 ≈ 33%**
size 2	**12/63 ≈ 19%**
size 3	**0/20 = 0%**
all	**80/290 ≈ 28%**
	wrongly claimed global/claimed global
size 1	**0/68 = 0%**
size 2	**0/12 = 0%**
all	**0/80 = 0%**
	claimed infeasible/accepted and feasible
size 1	**0/201 = 0%**
size 2	**0/59 = 0%**
size 3	**0/18 = 0%**
all	**0/278 = 0%**

New, promising solvers

using new techniques

(not yet competitive with BARON)

- **COCOS** (Hermann Schichl, **Vienna**)

  `http://www.mat.univie.ac.at/coconut`

- **LaGO** (Ivo Nowak, **Berlin**)

  `http://www.mathematik.hu-berlin.de/~eopt/LaGO/documentation/`

As we have seen, complete, but nonrigorous mehods are already today superior to the best general purpose heuristic methods in small and medium dimensions (< 1000).

But incomplete methods are currently still the only choice available for difficult large-scale problems such as protein folding, radiation therapy planning, optimal design and packing.

Even this might change in the near future.

Rounding errors

Lack of reliability due to rounding errors is still a problem in current codes. In the optimization community, awareness of this problem is only slowly growing.

When I wrote in 1986 my first branch and bound code for covering solution curves of algebraic equations, every now and then a pixel was missing in the pictures. It turned out that the reason was a loss of solutions due to poor handling of roundoff problems.

Geometric visualization software had to cope with the same problem: Today they all use carefully designed algorithms with safe rounding in critical computations.

Not so in optimization codes, unfortunately!

Even high quality mixed integer linear programming (MILP) codes which have already a long commercial tradition may fail due to roundoff!

CPLEX 8.0 and all but one MILP solver from NEOS failed in 2002 to handle a simple 20 variable MILP problem with small integer coefficients and solution, claiming that no solution exists.

But cheap rigorous safeguards based on interval arithmetic are now available (NEUMAIER & SHCHERBINA 2004, JANSSON 2004, 2005), and will probably be soon part of commercial packages.

The 'innocent' MILP is the case $s = 6$ of the 20 variable integer linear problem

$$\begin{array}{ll} \min & -x_{20} \\ \text{s.t.} & (s+1)x_1 - x_2 \geq s-1, \\ & -sx_{i-1} + (s+1)x_i - x_{i+1} \geq (-1)^i(s+1) \ \textbf{for}\ i = 2:19, \\ & -sx_{18} - (3s-1)x_{19} + 3x_{20} \geq -(5s-7), \\ & 0 \leq x_i \leq 10 \ \textbf{for}\ i = 1:13, \\ & 0 \leq x_i \leq 10^6 \ \textbf{for}\ i = 14:20, \\ & \textbf{all}\ x_i\ \textbf{integers} \end{array}$$

FortMP found the solution $x = (1,2,1,2,...,1,2)^T$, but...

Five other MIP solvers from NEOS (June 2002), namely GLPK, XPRESS-MP, MINLP, BONSAIG, XPRESS and the solver CPLEX 8.0 (not available through NEOS) claimed no solution exists.

Most solvers suffer from rounding errors introduced through ill-conditioning. (The solution is a nearly degenerate nonoptimal vertex of the linear programming relaxation.)

WARNING: A high proportion of real life linear programs (72% according to Ordóñez & Freund, **and still 19% after preprocessing) are ill-conditioned!**

Primal linear program:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \underline{b} \le Ax \le \overline{b}, \end{aligned} \tag{5}$$

Corresponding dual linear program:

$$\begin{aligned} \max \quad & \underline{b}^T y - \overline{b}^T z \\ \text{s.t.} \quad & A^T(y - z) = c,\ \ y \ge 0,\ \ z \ge 0. \end{aligned} \tag{6}$$

Introduce boxes:

$$\mathbf{b} := [\underline{b}, \overline{b}] = \{\tilde{b} \in \mathbb{R}^n \mid \underline{b} \le \tilde{b} \le \overline{b}\},$$

Assume

$$Ax \in \mathbf{b} \quad \Rightarrow \quad x \in \mathbf{x} = [\underline{x}, \overline{x}].$$

From an approximate solution of the dual program we calculate an approximate multiplier $\lambda \approx z - y$, and a rigorous interval enclosure for

$$r := A^T\lambda - c \in \mathbf{r} = [\underline{r}, \overline{r}].$$

Since $c^Tx = (A^T\lambda - r)^Tx = \lambda^TAx - r^Tx \in \lambda^T\mathbf{b} - \mathbf{r}^T\mathbf{x}$,

$$\mu := \inf(\lambda^T\mathbf{b} - \mathbf{r}^T\mathbf{x})$$

is the desired rigorous lower bound for c^Tx.

In well-conditioned cases, the bound is quite accurate, while in ill-conditioned cases, it is so poor that it warns the user (or the algorithm) that something went wrong and needs special attention.

Safeguarding MILP solutions is more involved but can also be done.

Challenges for the future I

- **Ensuring reliability (safe bounds, finite termination analysis, certificates)**
- **Integrating MIP and SDP techniques into a branch-and-bound framework**
- **unbounded variables**
- **unconstrained/bound constrained problems (the more constraints the easier the problem! $\Rightarrow$ bounded residual estimation preferable to least squares)**

Challenges for the future II

- **Problems with severe dependence (volume preserving recurrences imply heavy wrapping)**
- **Problems with symmetries (optimal design of experiments, chemical cluster optimization)**
- **sensitivity analysis**
- **parametric global optimzation**
- **constraint projection problems**

Challenges for the future III

- **Differential constraints**
  (optimal control of chemical plants;
  space mission design)

- **Integral constraints**
  (expectations; value at risk,
  engineering safety factors)

- **Other desirables**
  (black box functions; expensive functions;
  nonsmoothness; noise; small discontinuities;
  uncertain domain of definition; SNOBFIT

A. Neumaier
Complete search in continuous global optimization and constraint satisfaction,
pp. 1-99 in: Acta Numerica 2004,
Cambridge Univ. Press 2004.

A. Neumaier, O. Shcherbina, W. Huyer and T. Vinko
A comparison of complete global optimization solvers
Math. Programming B 103 (2005), 335-356.

H. Schichl and A. Neumaier
Transposition theorems and qualification-free optimality conditions, submitted
http://www.mat.univie.ac.at/~neum/papers.html#trans

Global (and Local) Optimization site

`www.mat.univie.ac.at/~neum/glopt.html`

COCONUT homepage

`www.mat.univie.ac.at/coconut`

A. Neumaier
Interval methods for systems of equations
Cambridge Univ. Press 1990.
(individual copies are reprinted upon request)