Computational tools for reliable global optimization

Charlie Vanaret, Institut de Recherche en Informatique de Toulouse

INRA, 10 March 2017

Short bio

Toulouse, France

- Sept 2015 - March 2017: postdoctoral fellow at IRT Saint Exupéry
  Multidisciplinary design optimization
- Nov 2014 - Aug 2015: teaching and research fellow at IRIT
- Oct 2011 - Oct 2014: PhD at ENAC
  Reliable global optimization
  Awarded two PhD prizes in 2015

Klagenfurt, Austria

- Apr 2011 - Aug 2011: research intern at Alpen-Adria-Universität Klagenfurt

- Reliable optimization
  - Interval arithmetic
  - Interval branch and bound methods
- Charibde: a hybrid cooperative solver
  - Rigorous Differential Evolution algorithm
  - Contractors
  - Exploration strategy
- Experimental results
  - Validation on COCONUT benchmark
  - New optimality results for multimodal functions
  - An open problem in molecular dynamics
- Conclusion

Reliable optimization
- Interval arithmetic
- Interval branch and bound methods

Charibde: a hybrid cooperative solver
- Rigorous Differential Evolution algorithm
- Contractors
- Exploration strategy

Experimental results
- Validation on COCONUT benchmark
- New optimality results for multimodal functions
- An open problem in molecular dynamics

Conclusion

Roundoff

$$f(x,y) = 333.75y^6 + x^2(11x^2y^2 - y^6 - 121y^4 - 2) + 5.5y^8 + x/(2y)$$

Evaluating $f(77617, 33096)$

- in simple precision: $1.172603\ldots$

Roundoff

$$f(x, y) = 333.75y^6 + x^2(11x^2y^2 - y^6 - 121y^4 - 2) + 5.5y^8 + x/(2y)$$

Evaluating $f(77617, 33096)$

- in simple precision: $1.172603\ldots$
- in double precision: $1.1726039400531\ldots$

Roundoff

$$f(x, y) = 333.75y^6 + x^2(11x^2y^2 - y^6 - 121y^4 - 2) + 5.5y^8 + x/(2y)$$

Evaluating $f(77617, 33096)$

- in simple precision: $1.172603\ldots$
- in double precision: $1.1726039400531\ldots$
- the correct result is $-\frac{54767}{66192} = -0.827396\ldots$

Rigorous enclusures[1]

Interval arithmetic extends real arithmetic

- round-off errors in numerical computations

$$\pi \in [3.141, 3.142] \quad \text{(precision: 3 digits)}$$
$$\frac{\pi}{2} \in \frac{[3.141, 3.142]}{2} \subset [1.570, 1.571]$$

- on a computer: outward rounding
- exact result not known, but rigorously bounded

[1] R. E. Moore. *Interval Analysis.* Prentice-Hall, 1966.

Interval extensions

Computing with sets

$$[a, b] + [c, d] = [a +_{\downarrow} c, b +_{\uparrow} d]$$
$$\exp([a, b]) = [\exp_{\downarrow}(a), \exp_{\uparrow}(b)]$$

Interval extension F of f

- $f(X) := \{f(x) \mid x \in X\} \subset F(X)$
- $X \subset Y \Rightarrow F(X) \subset F(Y)$
- several extensions: natural, Taylor, monotonicity-based

Interval extensions

Multiple extensions: $f(x) = 2x - \sqrt{x}$

- natural extension: $F_N(X) = 2X - \sqrt{X}$
- Taylor extension: $F_{T1}(X, c \in X) = 2c - \sqrt{c} + (2 - \frac{1}{2\sqrt{X}})(X - c)$
- centered extension: $F_C(X, z \in \mathbb{R}) = 2z - \sqrt{z} + (2 - \frac{1}{\sqrt{X}+\sqrt{z}})(X - z)$
- monotonicity-based extension: $F_M([1,4]) = [\underline{F(1)}, \overline{F(4)}]$

Overestimation of interval arithmetic

Dependency problem

- occurrence decorrelation of the variables

$$
\begin{aligned}
X - X &= [\underline{X} - \overline{X}, \overline{X} - \underline{X}] \\
&= \{x_1 - x_2 \mid x_1 \in X, x_2 \in X\} \\
&\supset \{x - x \mid x \in X\}
\end{aligned}
$$

- $F(X)$ generally overestimates $f(X)$
- enclosure is optimal when simple occurrences[2]

[2] R. E. Moore. *Interval Analysis.* Prentice-Hall, 1966.

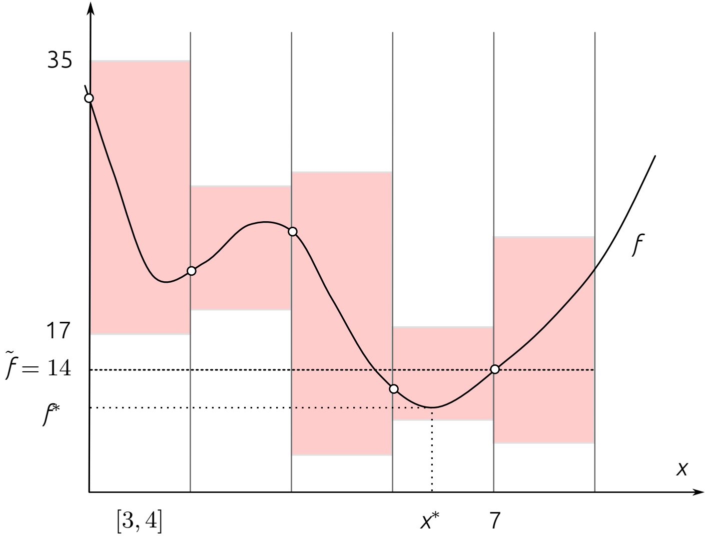
Overestimation of interval arithmetic (example)

Rewriting an analytical expression

- $$\begin{cases} f(x) = x^2 - 2x \\ g(x) = x(x-2) \\ h(x) = (x-1)^2 - 1 \end{cases}$$
- $$\begin{cases} F([1,4]) = [-7,14] & \supset f([1,4]) \\ G([1,4]) = [-4,8] & \supset g([1,4]) \\ H([1,4]) = [-1,8] & = h([1,4]) \end{cases}$$

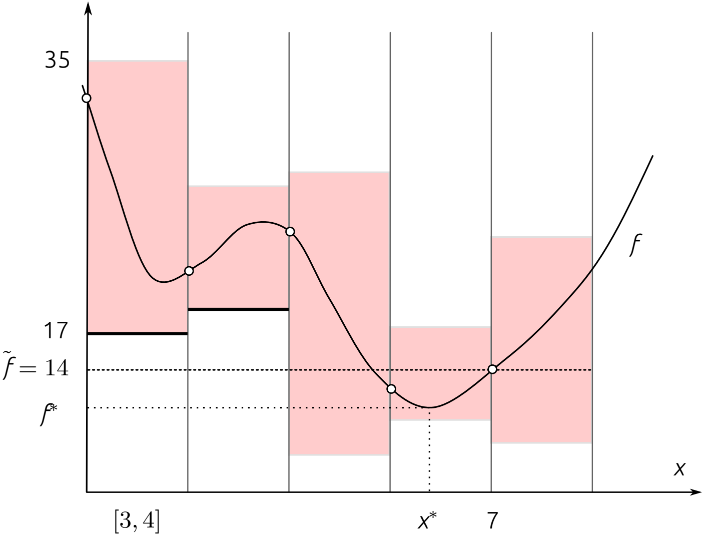
H provides an optimal enclosure (Moore)

Interval branch and bound[3]



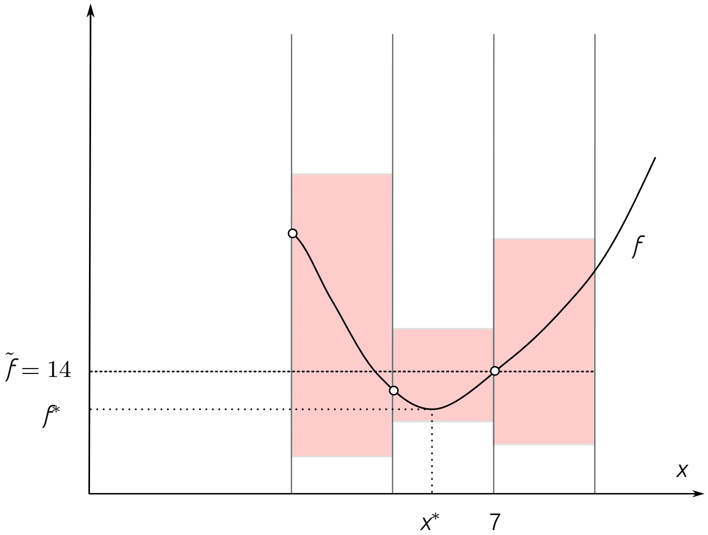
[3] Ramon E Moore. "On computing the range of a rational function of n variables over a bounded region". In: *Computing* 16.1 (1976), pp. 1–15.

Interval branch and bound[3]



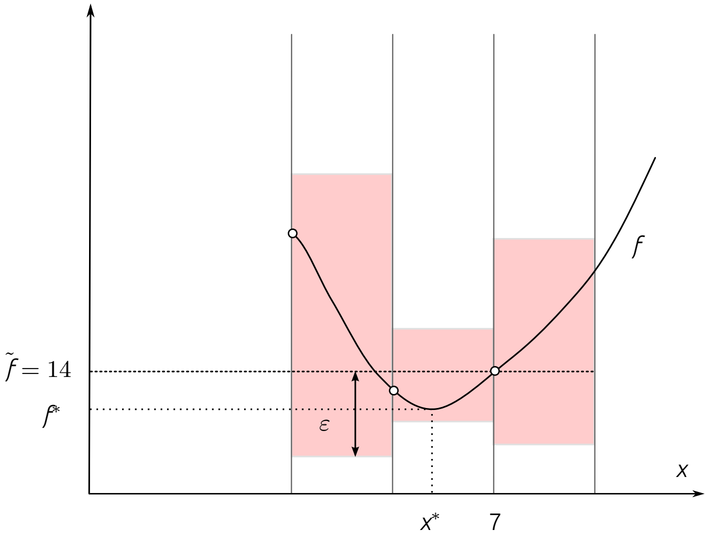
[3] Ramon E Moore. "On computing the range of a rational function of n variables over a bounded region". In: *Computing* 16.1 (1976), pp. 1–15.

Interval branch and bound[3]



[3] Ramon E Moore. "On computing the range of a rational function of n variables over a bounded region". In: *Computing* 16.1 (1976), pp. 1–15.

Interval branch and bound[3]



[3] Ramon E Moore. "On computing the range of a rational function of n variables over a bounded region". In: *Computing* 16.1 (1976), pp. 1–15.

Reliable optimization
- Interval arithmetic
- Interval branch and bound methods

Charibde: a hybrid cooperative solver
- Rigorous Differential Evolution algorithm
- Contractors
- Exploration strategy

Experimental results
- Validation on COCONUT benchmark
- New optimality results for multimodal functions
- An open problem in molecular dynamics

Conclusion

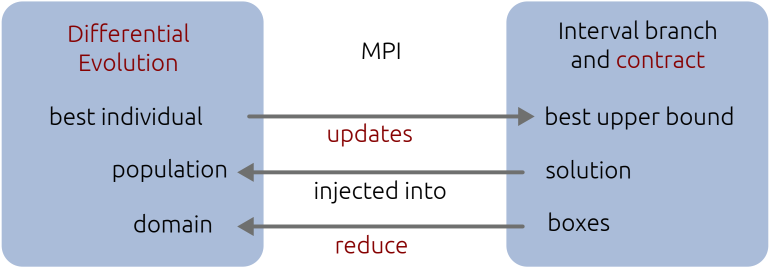
My PhD work

Devising a hybrid cooperative global reliable solver

- builds upon preliminary work at ENAC[4]
- hybrid: combines an IBB and a metaheuristic
- cooperative: MPI data exchange between parallel processes
- global: global minimum f^*
- reliable: returns a rigorous upper bound of f^*
- 10,000 lines of OCaml, uses an interval library[5]

[4] J-M. Alliot et al. "Finding and Proving the Optimum: Cooperative Stochastic and Deterministic Search". In: *20th European Conference on Artificial Intelligence* (2012).

[5] J-M. Alliot et al. "Implementing an interval computation library for OCaml on x86/amd64 architectures". In: *Proceedings of the 17th ACM SIGPLAN International Conference on Functional Programming*. 2012.

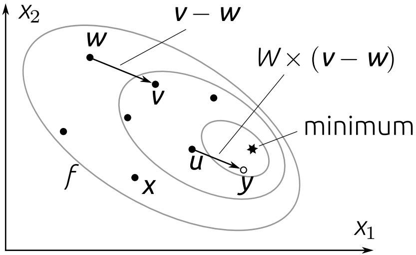


[6] C. Vanaret et al. "Preventing Premature Convergence and Proving the Optimality in Evolutionary Algorithms". In: *Lecture Notes in Computer Science* (2013), C. Vanaret et al. "Hybridization of Interval CP and Evolutionary Algorithms for Optimizing Difficult Problems". In: *21st International Conference on Principles and Practice of Constraint Programming (CP 2015)*. 2015.

Reliable Differential Evolution[7]

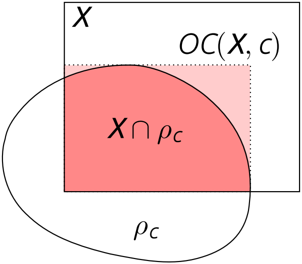
Metaheuristic = randow walk guided by heuristics

- single individual: simulated annealing
- evolutionary algorithms: GA, CMA-ES
- distributed intelligence: PSO, ACO



[7] R. Storn and K. Price. "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces". In: *Journal of Global Optimization* (1997), pp. 341–359.

Contractors[8]

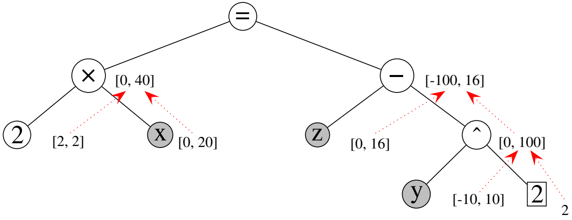


- constraints
  - $f \leq \tilde{f} - \varepsilon$
  - $g \leq 0, h = 0$ (constrained)
  - $\nabla f = 0$ (unconstrained) via automatic differentiation
- contractors: HC4, Mohc, X-Newton, 3B, CID

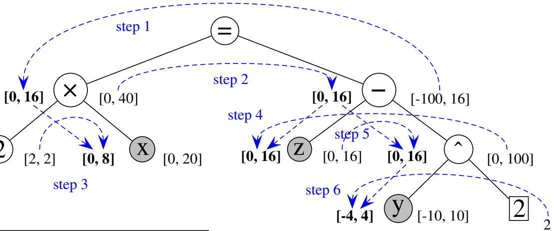
[8] G. Chabert and L. Jaulin. "Contractor programming". In: *Artificial Intelligence* 173 (2009), pp. 1079–1100.

HC4Revise[9] ($2x = z - y^2$)

- bottom-up phase: evaluation



- top-down phase: constraint propagation using inverse functions



[9] F. Benhamou et al. "Revising Hull and Box Consistency". In: *International Conference on Logic Programming*. MIT press, 1999, pp. 230–244.

MaxDist: an exploration strategy

Standard strategies

- ”depth-first” search: gets often stuck in local minima
- ”best-first” search: subject to dependency
- ”largest first” search: does not give advantage to promising regions

MaxDist

Extract the box that is the farthest from the current solution

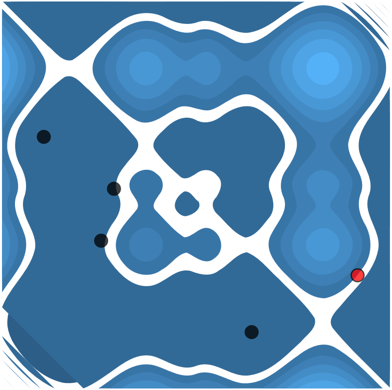
- explore the neighborhood of x^* only when the best possible upper bound is available
- explore regions that are hardly accessible by the DE

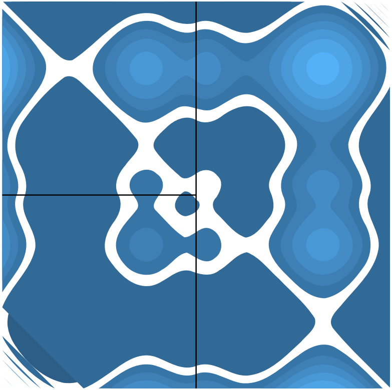
MaxDist: an exploration strategy

Experimentally

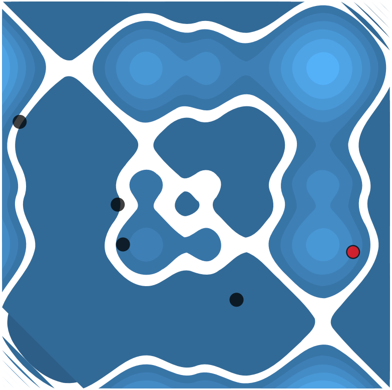
- the maximum size of the queue is very low
- the remaining boxes may be sent to DE at low cost
- their convex hull is cheap and conservative
- DE individuals may be restarted in the new contracted domain

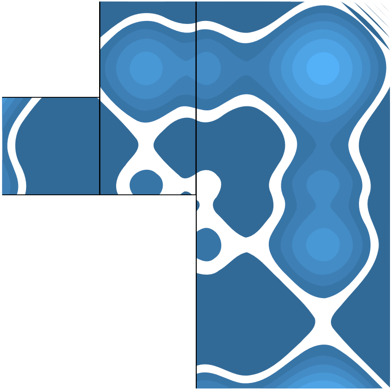
Example on Schwefel function ($n = 2$)



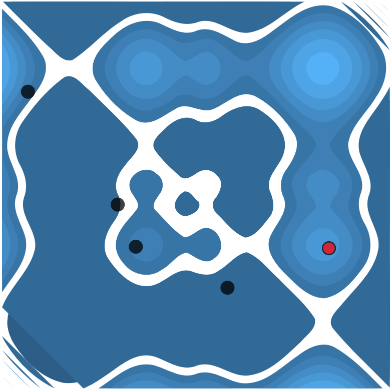


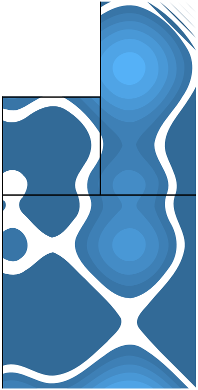
Example on Schwefel function ($n = 2$)

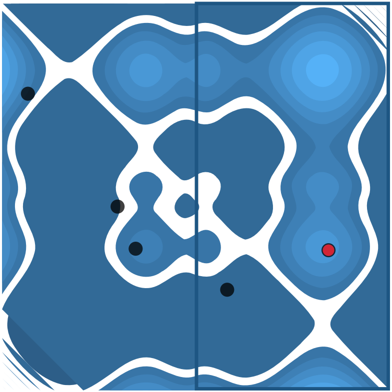


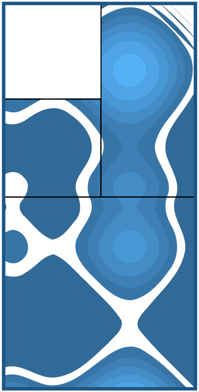


Stuck in a local minimum

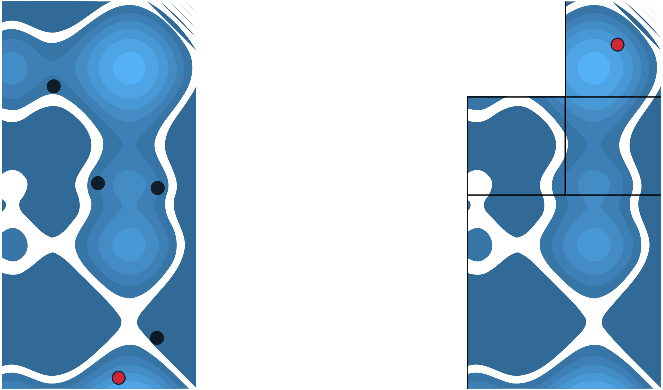




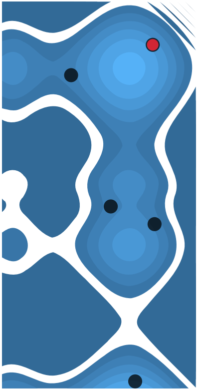


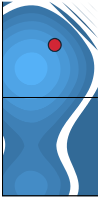


Restart in contracted domain

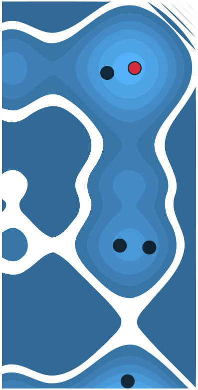


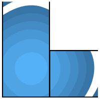
Solution injected into population





Proof of optimality





Reliable optimization

- Interval arithmetic
- Interval branch and bound methods

Charibde: a hybrid cooperative solver

- Rigorous Differential Evolution algorithm
- Contractors
- Exploration strategy

Experimental results

- Validation on COCONUT benchmark
- New optimality results for multimodal functions
- An open problem in molecular dynamics

Conclusion

Comparison with

- reliable interval solvers
  - GlobSol[10]
  - IBBA[11]
  - Ibex[12]
- non-reliable MINLP solvers
  - BARON
  - Couenne

[10] R Baker Kearfott. *Rigorous global search: continuous problems.* Springer, 1996.

[11] Frédéric Messine. "Méthodes d'optimisation globale basées sur l'analyse d'intervalles pour la résolution de problemes avec contraintes". PhD thesis. INPT-ENSEEIHT, Toulouse, 1997.

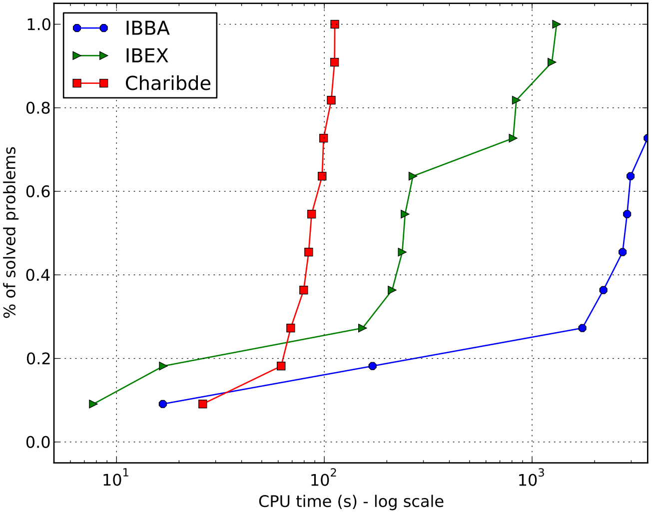
[12] Gilles Trombettoni et al. "Inner Regions and Interval Linearizations for Global Optimization." In: *AAAI*. 2011.

Validation on COCONUT benchmark

Table: CPU time (in s)

Problem	n	m	GlobSol	IBBA	Ibex	Charibde
ex2_1_7	20	10		16.7	**7.74**	26
ex2_1_9	10	1		154	**9.07**	36
ex6_2_6	3	1	306	1575	136	**6.92**
ex6_2_8	3	1	204	458	59.3	**10.66**
ex6_2_9	4	2	463	523	25.2	**4.54**
ex6_2_11	3	1	273	140	7.51	**2.76**
ex6_2_12	4	2	196	112	22.2	**10.75**
ex7_2_3	8	6		TO	544	**1.59**
ex7_3_5	13	15		TO	28.91	**8.8**
ex14_1_7	10	17		TO	406	**4**
ex14_2_7	6	9		TO	66.39	**0.3**
Sum			> 1442	TO	1312.32	**112.32**

Validation on COCONUT benchmark



Optimal results for highly multimodal functions

Few (putative) solutions known in the literature

Function	Contraints	Domain	Literature	Charibde
Michalewicz	bounds	$[0, \pi]^n$	up to $n = 50$	$n = 70$
Sine Envelope	bounds	$[-100, 100]^n$	$n = 2$	$n = 5$
Eggholder	bounds	$[-512, 512]^n$	$n = 2$	$n = 10$
Keane	inequality	$[0, 10]^n$	up to $n = 100$	$n = 5$
Rana	bounds	$[-512, 512]^n$	$n = 2$	$n = 7$

How multimodal are we talking about?

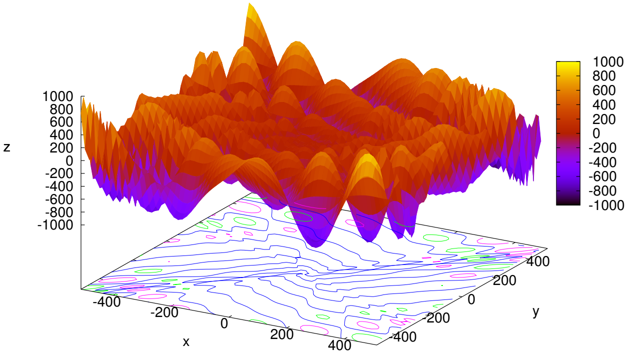


Figure: Eggholder function[13]

[13] S. K. Mishra. *Some new test functions for global optimization and performance of repulsive particle swarm method.* Tech. rep. University Library of Munich, Germany, 2006.

Performance profile

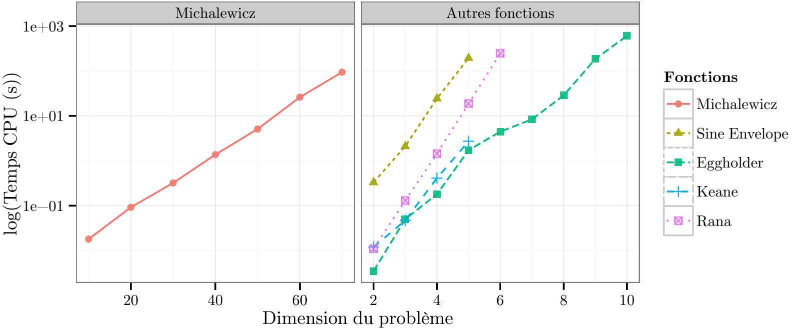


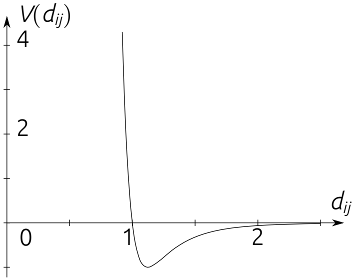
Figure: CPU time wrt dimension n

Lennard-Jones potential[14]

Describes the interactions for a pair of atoms

$$V(d_{ij}) = 4\left(\underbrace{\frac{1}{d_{ij}^{12}}}_{\text{repulsion}} - \underbrace{\frac{1}{d_{ij}^{6}}}_{\text{attraction}} \right)$$

where $d_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2$



[14] J. E. Jones. "On the Determination of Molecular Fields. I. From the Variation of the Viscosity of a Gas with Temperature". In: *Proceedings of the Royal Society of London. Series A* 106.738 (1924), pp. 441–462.

Lennard-Jones clusters

Most stable spatial configuration of a cluster with N atoms

$$\sum_{i<j}^{N} V(d_{ij}) = 4 \sum_{i<j}^{N} \left(\frac{1}{d_{ij}^{12}} - \frac{1}{d_{ij}^{6}} \right)$$

Very combinatorial problem

- nonconvex and highly multimodal ($O(e^N)$ local minima)
- invariant wrt translation and rotation
- putative minima gathered in Cambrige base, optimal for $N \leq 4$

Open problem for $N \geq 5$: a real challenge

Table: Minima and CPU time

Solver	Minimum	Search	Proof	Status
BARON	-9.103853464	0.23s	0.23s	local
Couenne	-9.103870326	41.94s	61.7s	global

- Best known solution[15]: -9.103852415708
- Interval solvers time out

[15] N.J.A. Sloane et al. "Minimal-energy clusters of hard spheres". In: *Discrete & Computational Geometry* 14 (1 1995), pp. 237–259.

Challenge accepted

Lennard-Jones problem: a new hope

Symmetry reduction ($3N - 6$ variables)

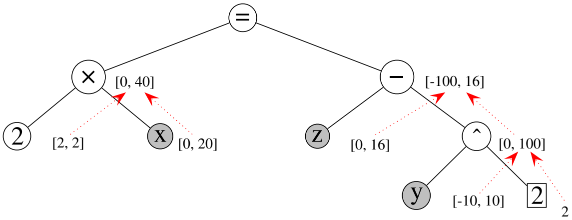
$$\begin{cases} x_1 = y_1 = z_1 = 0 \\ x_2 \geq 0, y_2 = z_2 = 0 \\ x_3 \geq 0, y_3 \geq 0, z_3 = 0 \\ x_4 \geq 0, y_4 \geq 0, z_4 \geq 0 \end{cases}$$

Dependency reduction

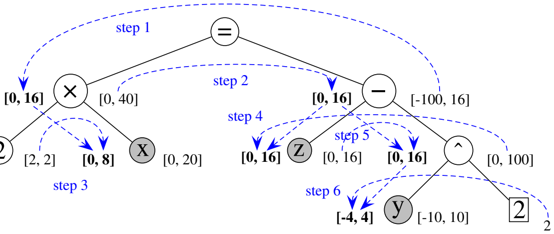
$$\sum_{i<j}^{N} V(d_{ij}) = 4\sum_{i<j}^{N}\left(\frac{1}{d_{ij}^{12}} - \frac{1}{d_{ij}^{6}}\right) = \sum_{i<j}^{N}\left[4\left(\frac{1}{d_{ij}^{6}} - \frac{1}{2}\right)^{2} - 1\right]$$

- each d_{ij} occurs once
- nevertheless, coordinates occur multiple times

- bottom-up phase: evaluation



- top-down phase: constraint propagation using inverse functions



Lennard-Jones problem: return of the contractor

Key idea

Perform the automatic differentiation in reverse mode after the constraint propagation[16]

- nodes of the syntax tree are shared
- refutation tests with more accurate derivatives (monotonicity tests, convexification)

[16] Hermann Schichl and Arnold Neumaier. "Interval analysis on directed acyclic graphs for global optimization". In: *Journal of Global Optimization* 33.4 (2005), pp. 541–562.

The first rigorous proof for $N = 5$

Table: Minima and CPU time

Solver	Minimum	Search	Proof	Status
BARON	-9.10385$\underline{3464}$	0.23s	0.23s	local
Couenne	-9.1038$\underline{70326}$	41.94s	61.7s	global
Charibde	-9.103852416	0.11s	1436s	certified ($\varepsilon = 10^{-9}$)

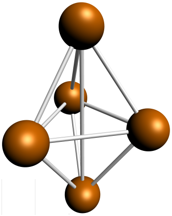


Figure: Optimal configuration: triangular bipyramid

- Reliable optimization
  - Interval arithmetic
  - Interval branch and bound methods
- Charibde: a hybrid cooperative solver
  - Rigorous Differential Evolution algorithm
  - Contractors
  - Exploration strategy
- Experimental results
  - Validation on COCONUT benchmark
  - New optimality results for multimodal functions
  - An open problem in molecular dynamics
- Conclusion

Conclusion

Charibde: a reliable hybrid solver

- reconciles mathematical programming, numerical analysis and artificial intelligence
- highly competitive against state-of-the-art solvers on difficult benchmarks
- rigorous > exhaustive (BARON, Couenne)
- first numerical proof of optimality for an open problem in molecular dynamics

Computational tools for reliable global optimization

Charlie Vanaret, Institut de Recherche en Informatique de Toulouse

INRA, 10 March 2017

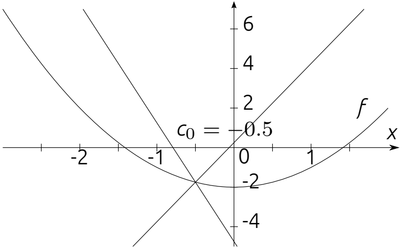
- Reliable global optimization
    - Contractors
    - Experimental results

Reliable global optimization

- Contractors
- Experimental results

Interval Newton: $f = 0$

- Newton operator: $N(X_k, c_k \in X_k) = c_k - \frac{F(c_k)}{F'(X_k)}$
- Zeros of f in X_k bounded by $X_{k+1} = X_k \cap N(X_k, c_k)$
- If $N(X_k, c_k) \subseteq X_k$, existence of a unique zero in X_k



$$f(x) = x^2 - 2, X_0 = [-3, 2], c_0 = -0.5$$

$$N(X_0, c_0) = -0.5 - \frac{(-0.5)^2 - 2}{2[-3, 2]}$$
$$= [-\infty, -\frac{19}{24}] \cup [-\frac{1}{16}, +\infty]$$

$$X_0 \cap N(X_0, c_0) = [-3, -\frac{19}{24}] \cup [-\frac{1}{16}, 2]$$

Propagation loop (fixed point)

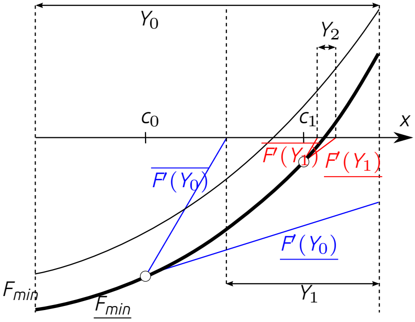
function FixedPoint($X, \mathcal{C}, OC$)
  $\mathcal{Q} \leftarrow \mathcal{C}$
  repeat
    Pick a constraint c_i in $\mathcal{Q}$
    $X' \leftarrow OC(X, c_i)$
    if $X' \neq X$ **then**
      $\mathcal{Q} \leftarrow \mathcal{Q} \cup \{c_j \mid c_j \in \mathcal{C} \wedge \exists v_k \in var(c_j), X'_k \neq X_k\}$
      $X \leftarrow X'$
    end if
    $\mathcal{Q} \leftarrow \mathcal{Q} \setminus \{c_i\}$
  until $\mathcal{Q} = \varnothing$
end function

Mohc

$$f(X) \subset F_M(X) = [\underline{F(X^-)}, \overline{F(X^+)}]$$

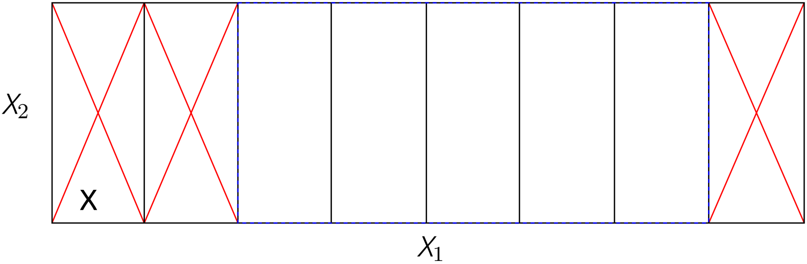
MohcRevise

- HC4Revise$(F(X^-) \leq 0)$
- HC4Revise$(F(X^+) \geq 0)$
- MonotonicBoxNarrow: Newton on $\underline{F(X^-)}$ and $\overline{F(X^+)}$



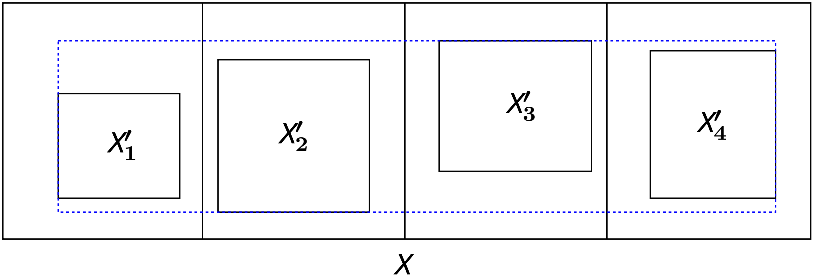
3B

- refutation of small portions of the domain
- quasi fixed point (propagation loop on the variables)



CID

- convex hull on contracted portions
- preserves information about contraction of other variables
- quasi fixed point (propagation loop on the variables)



Lower bound of a constrained problem

- affine arithmetic [6, 14]
- Corner Taylor: X-Newton [3]

$$(\mathcal{P}_{lb}) \quad \min \quad \sum_{i=1}^{n} \underline{\frac{\partial F}{\partial x_i}(X)} \cdot x_i$$

$$\text{s.c.} \quad g_j(\underline{X}) + \sum_{i=1}^{n} \underline{\frac{\partial G_j}{\partial x_i}(X)} \cdot (x_i - \underline{X_i}) \leq 0, \quad \forall j \in \{1, \dots, m\}$$

$$\underline{X_i} \leq x_i \leq \overline{X_i}, \quad \forall i \in \{1, \dots, n\}$$

LJ5: hyperparameters

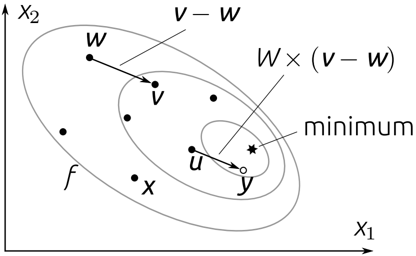
Hyperparameter	Value
ε (precision)	10^{-9}
NP (population size)	40
W (amplitude)	0.7
CR (crossover rate)	0.4
Bisection strategy	Largest
Exploration strategy	MaxDist
η (fixed point ratio)	0

LJ5: statistics

Table: Average results over 100 runs

CPU time (s)	1436
Max CPU time (s)	1800
Max size of $\mathcal{L}$	46
Evaluations of F (IBC)	7,088,758
Evaluations of ∇F (IBC)	78,229,737
Evaluations of f (DE)	483,642,320
Evaluations of F (DE)	132

Differential Evolution algorithm



repeat
    for x in the population **do**
        Pick individuals (u, v, w) in the population
        Combine x and $u + W \times (v - w)$
        if new individual better than x **then**
            It replaces x in the population
        end if
    end for
until termination criterion met