


Lobachevsky State University of Nizhny Novgorod
National Research University

Institute of Information Technology, Mathematics and Mechanics

Solving Time-Consuming Global Optimization Problems with Globalizer Software System

Alexander Sysoyev, Konstantin Barkalov, Vladislav Sovrasov,
Ilya Lebedev, Victor Gergel

State University of Nizhny Novgorod
National Research University
Institute of Information Technology, Mathematics and Mechanics

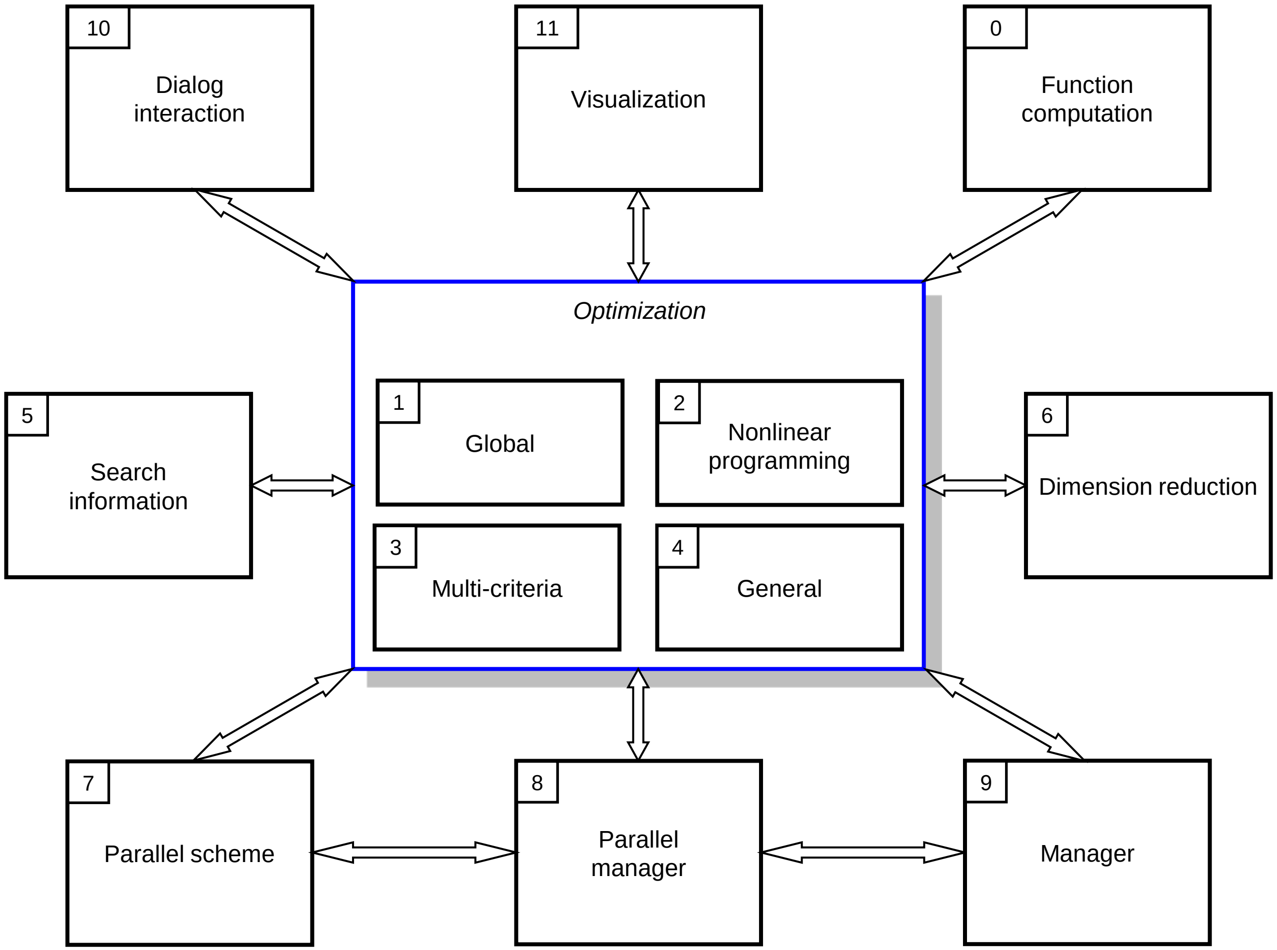
Russian Supercomputing Days, 2017

Solving Time-Consuming Global Optimization Problems with Globalizer Software System

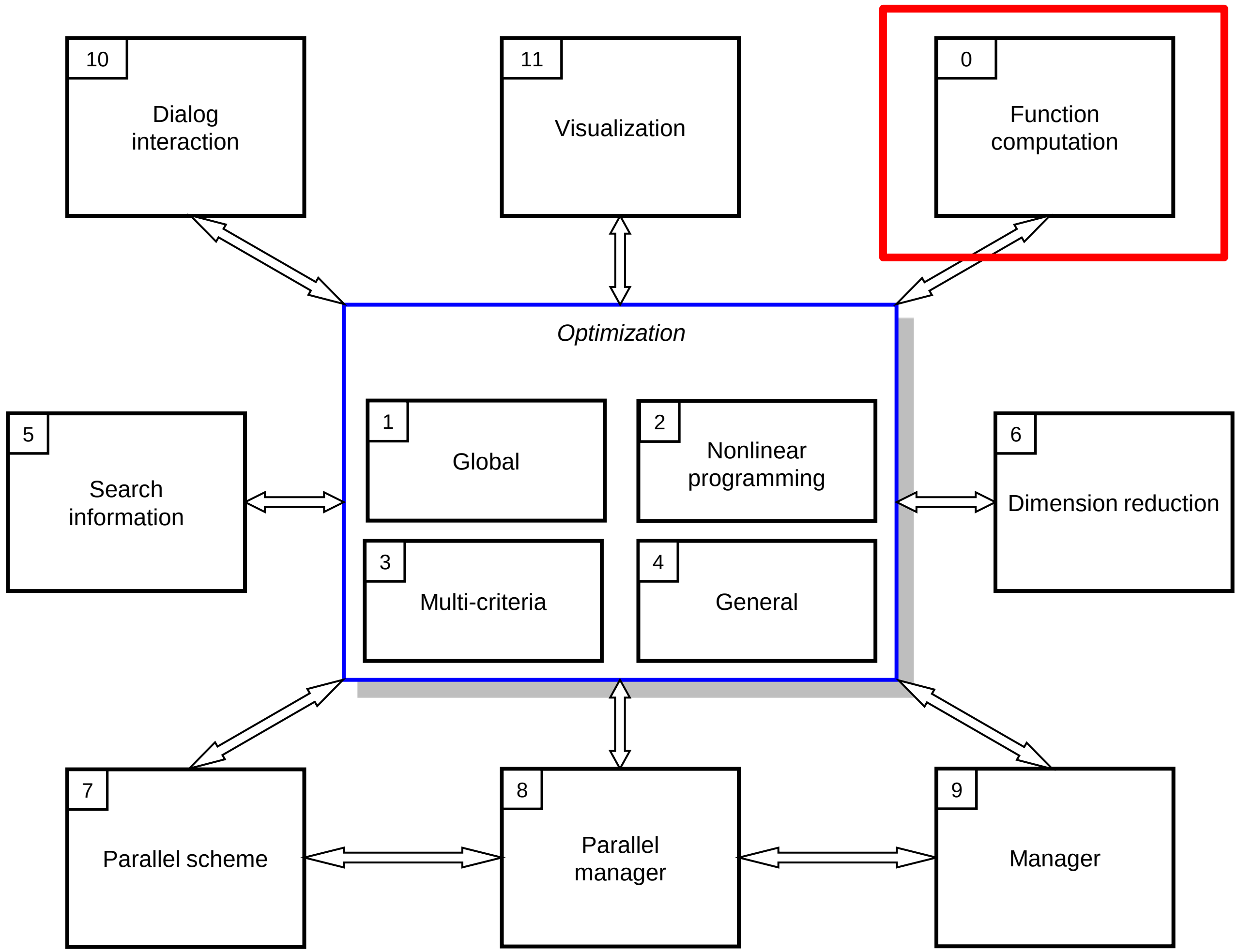
Alexander Sysoyev
Konstantin Barkalov
Vladislav Sovrasov
Ilya Lebedev
Victor Gergel

Software department

Globalizer Architecture



Globalizer Architecture



Problem statement

Consider the multidimensional optimization problem

$$\varphi(y) \to \min, \quad y = (y_1, \ldots, y_n) \qquad (1)$$

the search domain is the hyperinterval

$$D = \left\{ y \in R^N : a_i \le y_i \le b_i, \quad 1 \le i \le N \right\} \qquad (2)$$

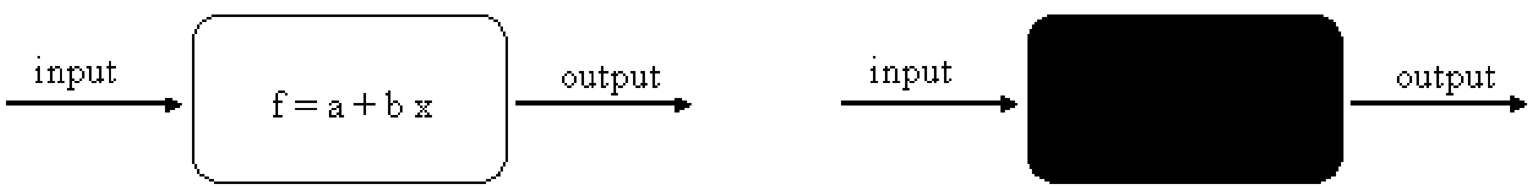
Problem statement

A priori information about the problem

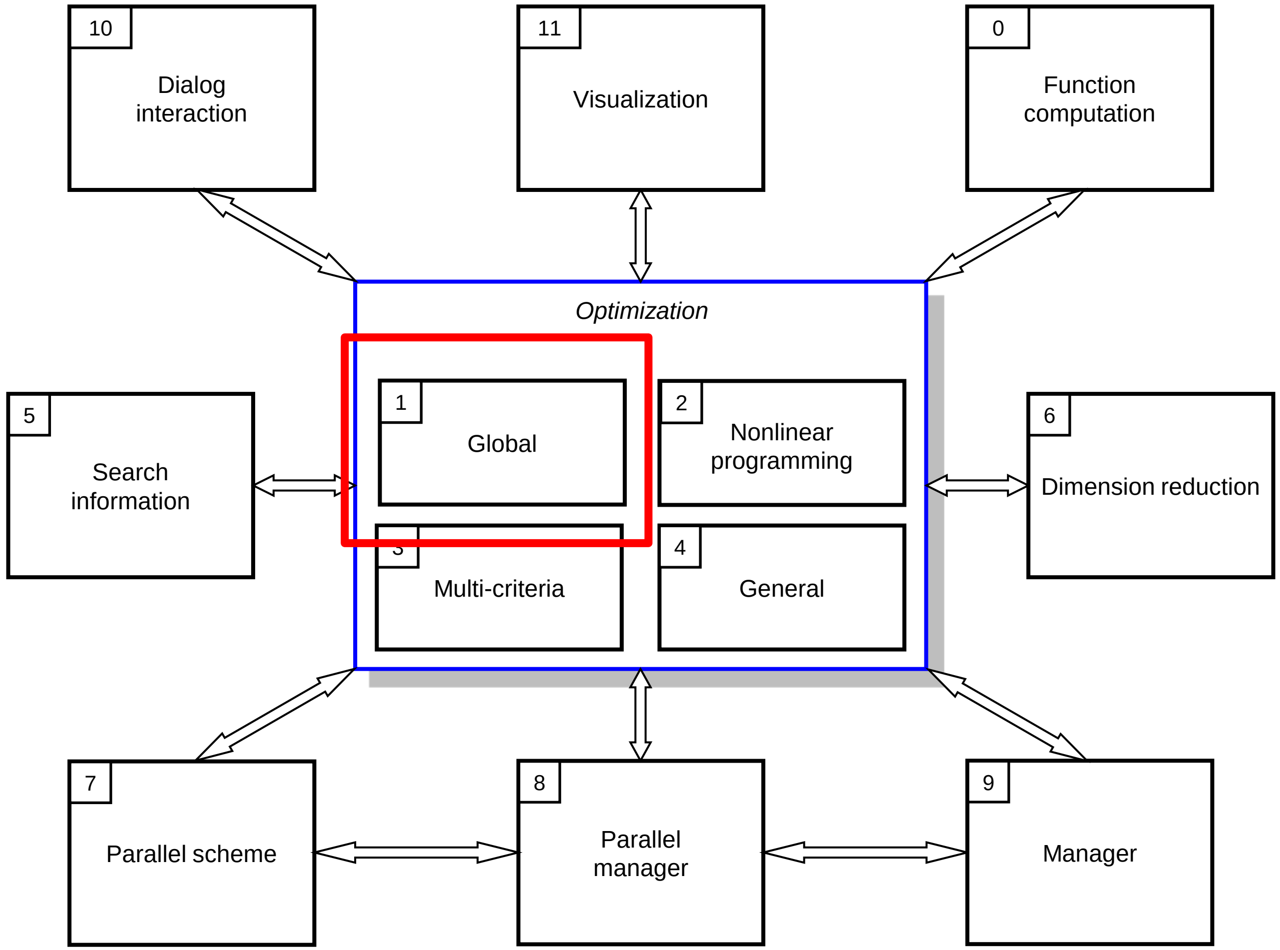
- the objective function $\varphi(y)$ satisfies the Lipschitz condition

$$\left|\varphi(y_1)-\varphi(y_2)\right| \le L\left\|y_1 - y_2\right\|,\ y_1, y_2 \in D$$

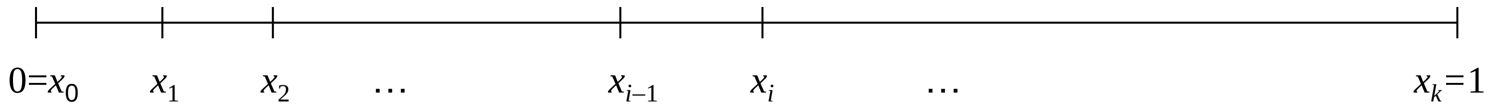
- the objective function $\varphi(y)$ may be “black box” function.



Globalizer Architecture



Algorithm of Global Search (AGS)



Let $x^0=0$, $x^1=1$.

1. For each (x_{i-1},x_i), $1\le i\le k$, calculate *characteristic* $R(i)$.
2. Find interval with maximum characteristic
   $R(t)=\max\{R(i)\colon 1\le i\le k\}$.
3. Make next trial at internal point of the interval $x^{k+1}\in(x_{t-1},x_t)$,
4. Check stop condition: $\Delta_t\le\varepsilon$

Algorithm of Global Search (AGS)

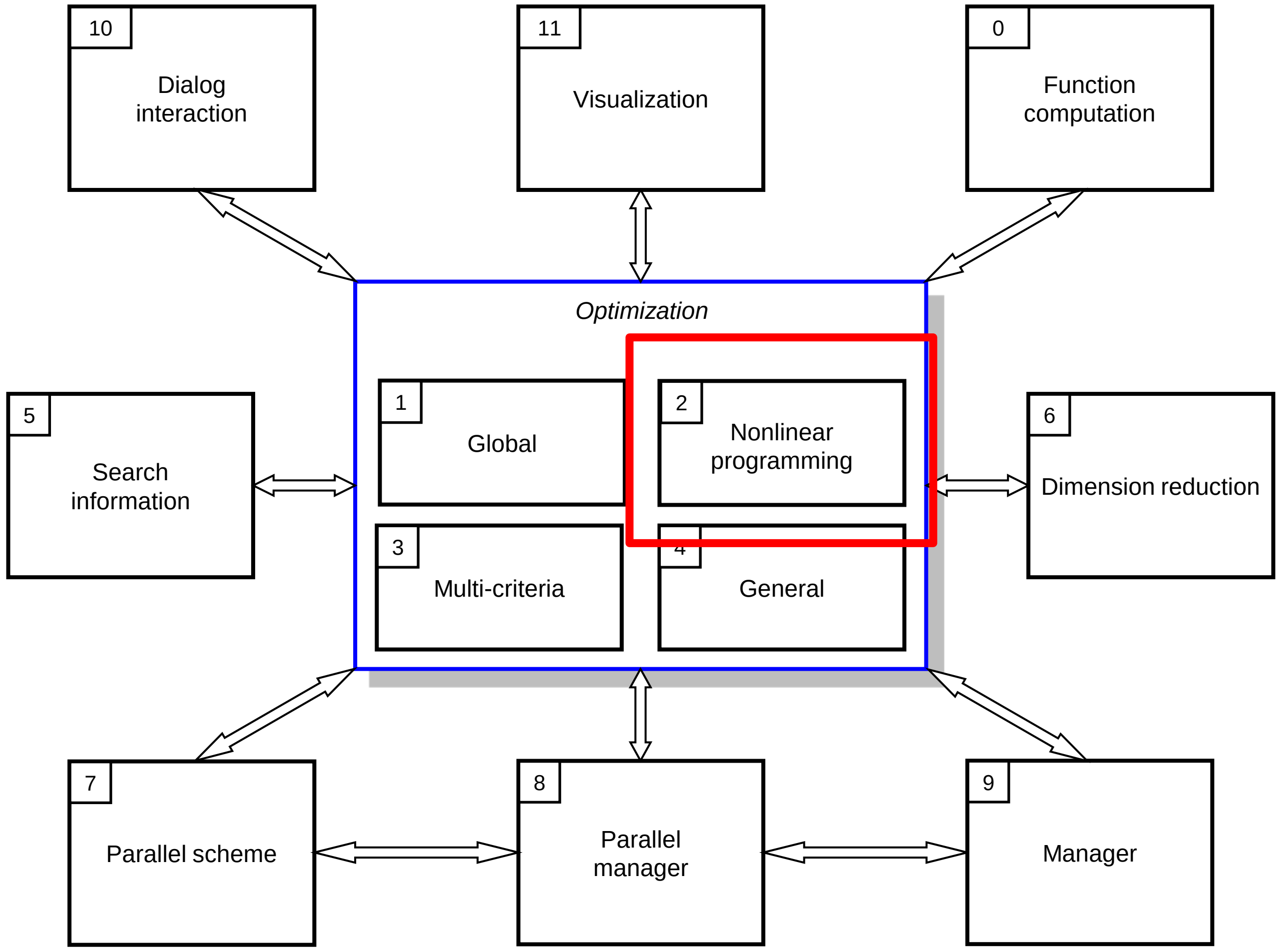
- Characteristic $R(i) = \Delta_i + \frac{(z_i - z_{i-1})^2}{r^2 \mu^2 \Delta_i} - 2\frac{(z_i + z_{i-1})}{r\,\mu},$

where $\mu = \max\left\{ \frac{|z_i - z_{i-1}|}{\Delta_i}, \quad i = 1,\dots,k \right\}$ is adaptive estimation of Lipschitz constant L, $r > 1$ – parameter.

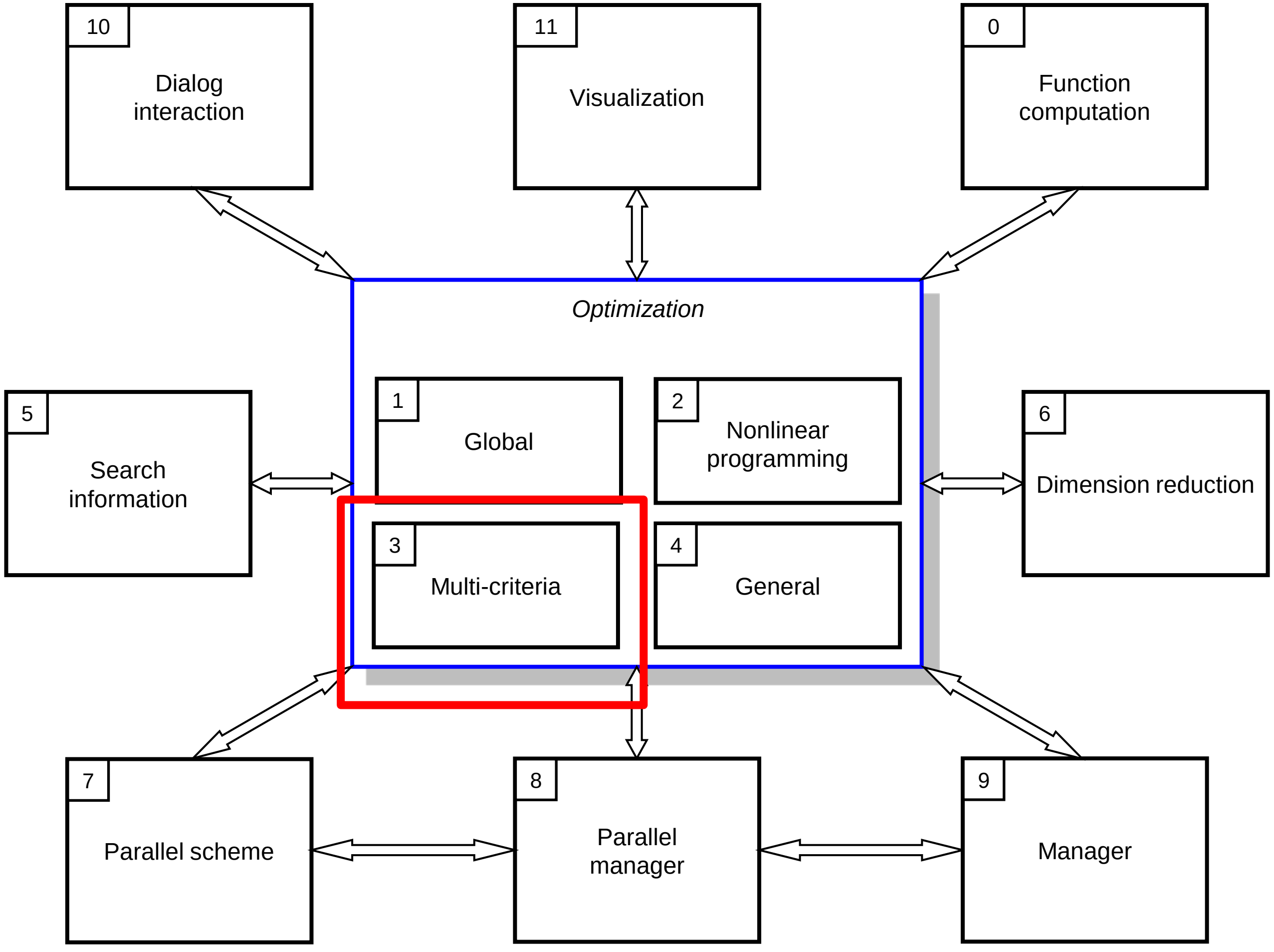
- New point $x^{k+1} = \frac{x_t + x_{t-1}}{2} - \text{sign}(z_t - z_{t-1})\frac{1}{2r}\left[\frac{|z_t - z_{t-1}|}{\mu}\right]^N$

Theory of convergence of AGS presented in [Strongin, Barkalov].

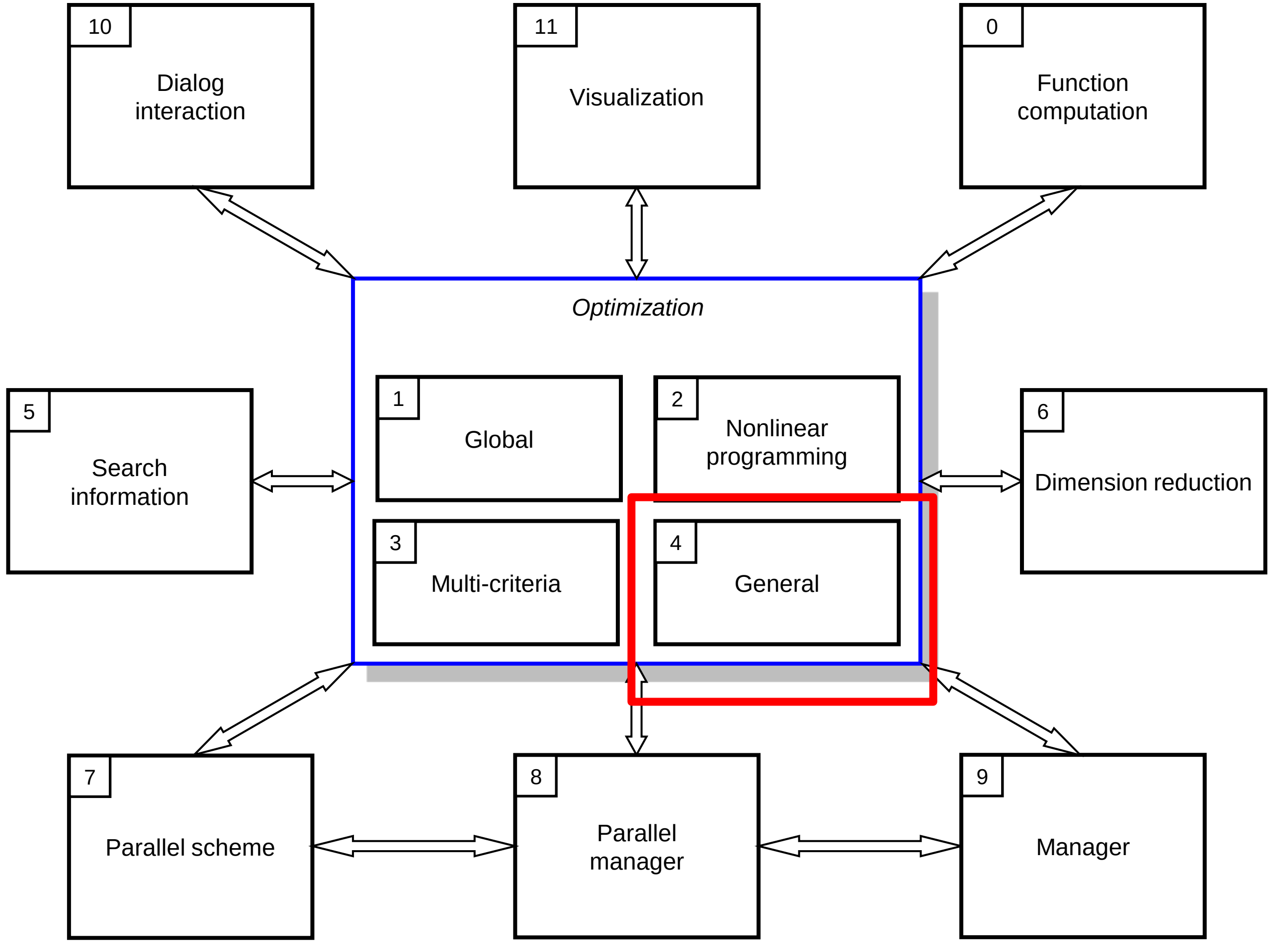
Globalizer Architecture



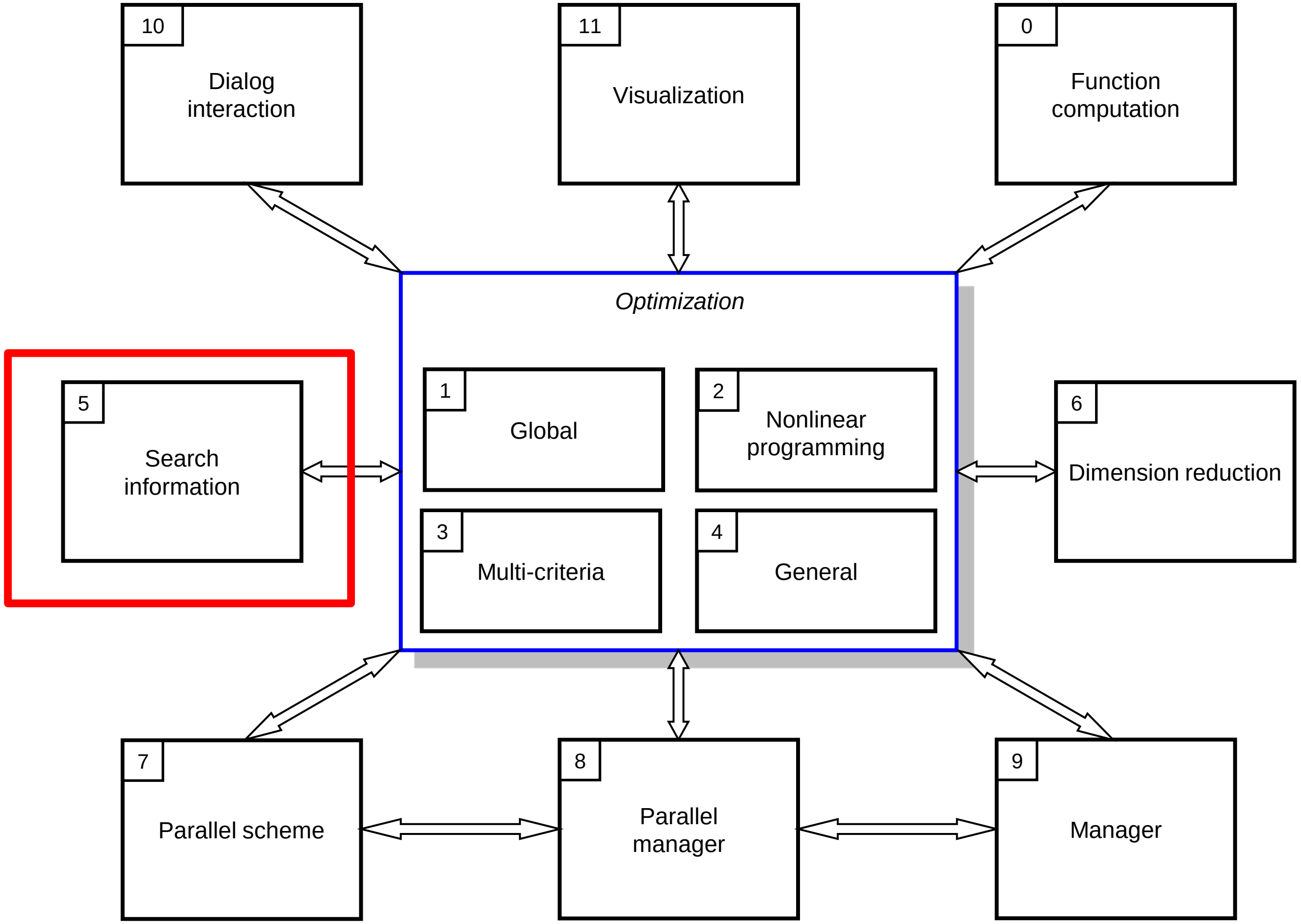
Globalizer Architecture



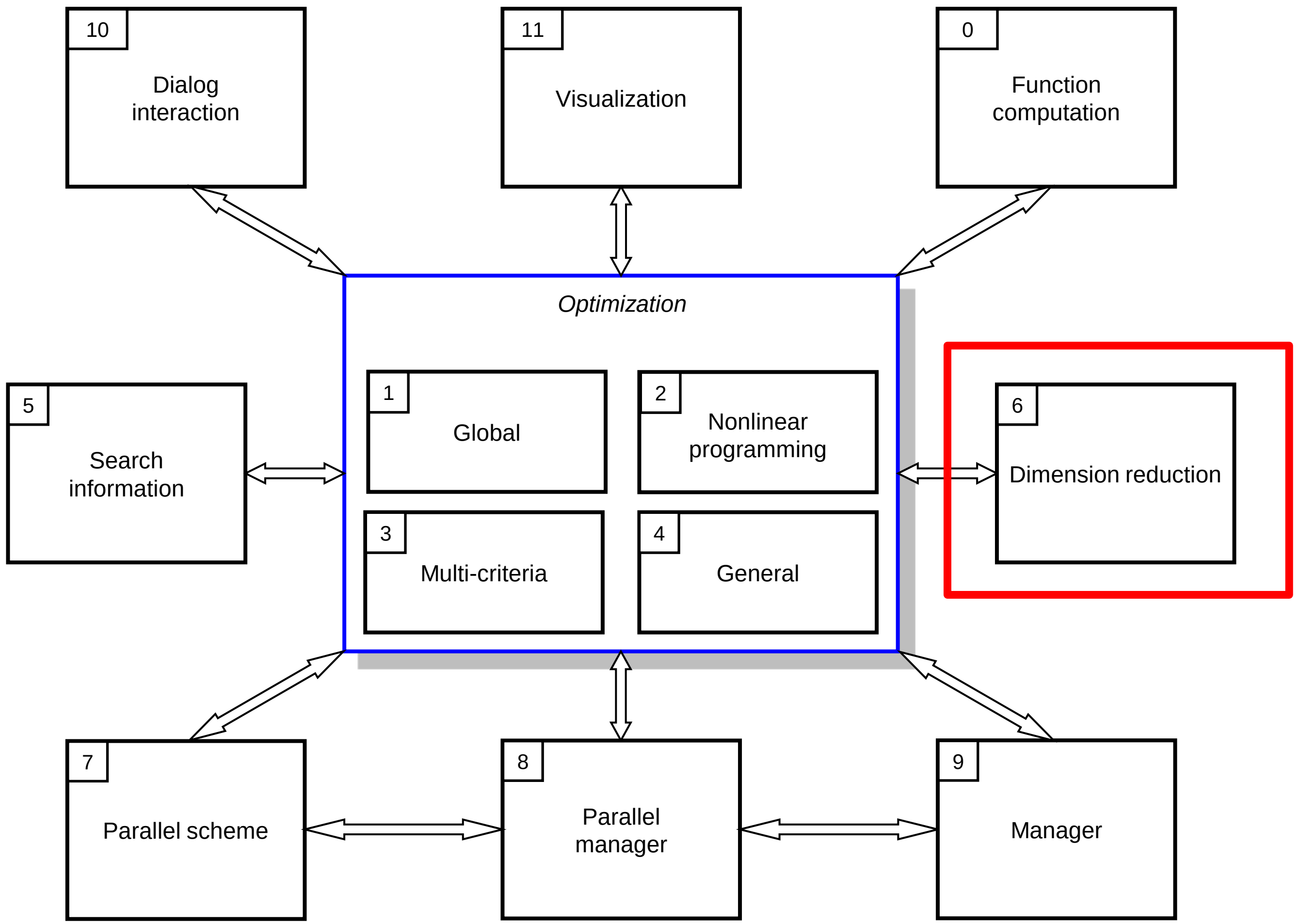
Globalizer Architecture



Globalizer Architecture



Globalizer Architecture



Reduction to one-dimensional problem

• The Globalizer implements a block multistage scheme of dimension reduction

• Initial vector y is represented as a vector of the «aggregated» macro-variables

$$y = (y_1, y_2, \dots, y_N) = (u_1, u_2, \dots, u_M)$$

where

$$u_1 = (y_1, y_2, \dots, y_{N_1})$$

$$u_2 = (y_{N_1+1}, y_{N_1+2}, \dots, y_{N_1+N_2})$$

$$u_M = (y_{N-N_M+1}, y_{N-N_M+2}, \dots, y_N)$$

and

$$N_1 + N_2 + \dots + N_M = N$$

Reduction to one-dimensional problem

- Using the macro-variables, the main relation of the well-known multistage scheme can be rewritten in the form

$$\min_{y\in D}\varphi(y) = \min_{u_1\in D_1}\min_{u_2\in D_2}\ldots\min_{u_M\in D_M}\varphi(y)$$

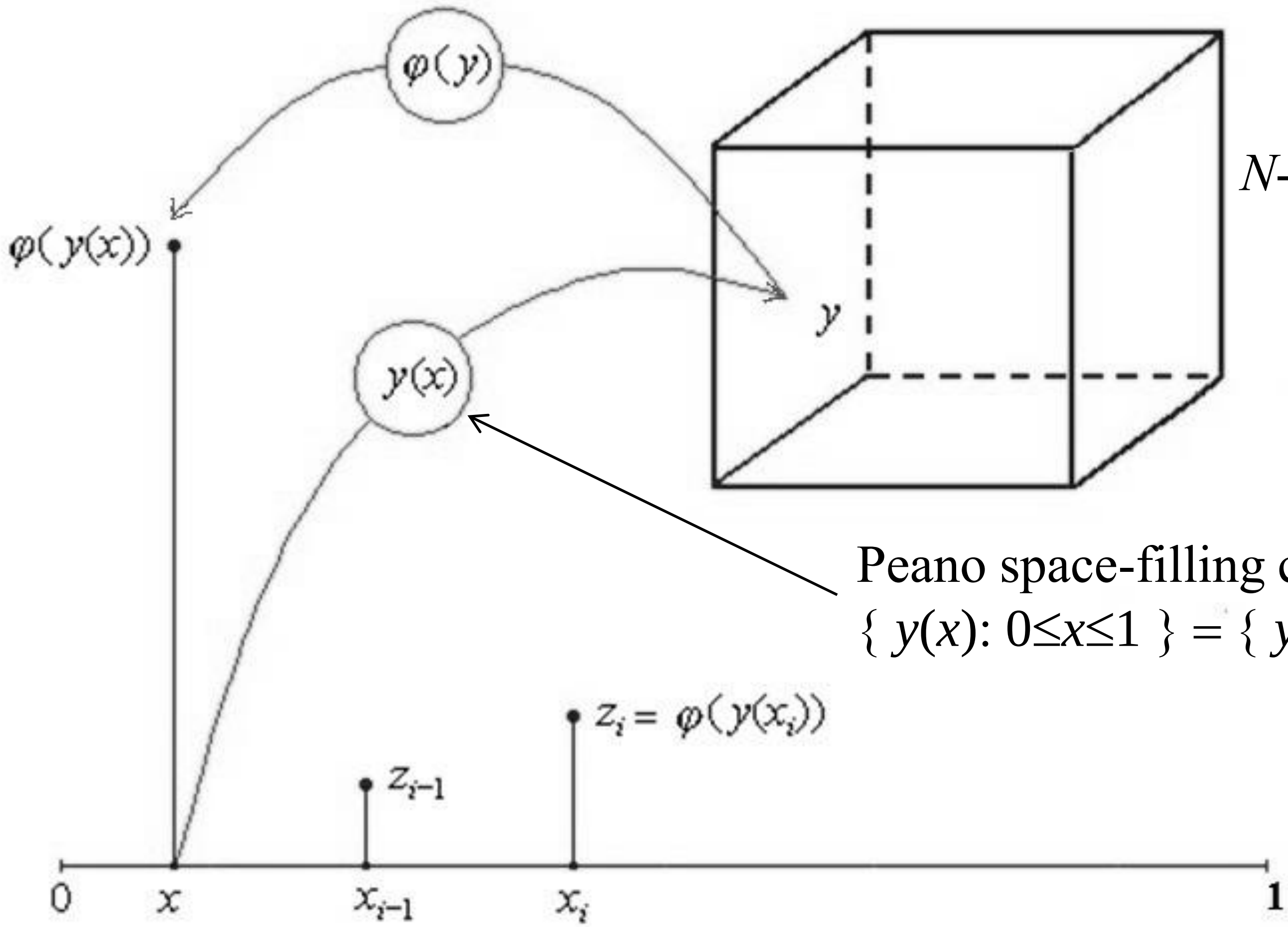
where D_i, $1 \le i \le M$, are the projections of D onto the subspaces corresponding to the macro-variables u_i, $1 \le i \le M$.

- The nested subproblems

$$\varphi_i(u_1,\ldots,u_i) = \min_{u_{i+1}\in D_{i+1}} \varphi_{i+1}(u_1,\ldots,u_i,u_{i+1}),\ 1 \le i \le M-1$$

are multidimensional and for their solution we can use the dimension reduction on the basis of the Peano curves

Reduction to one-dimensional problem



N-d space

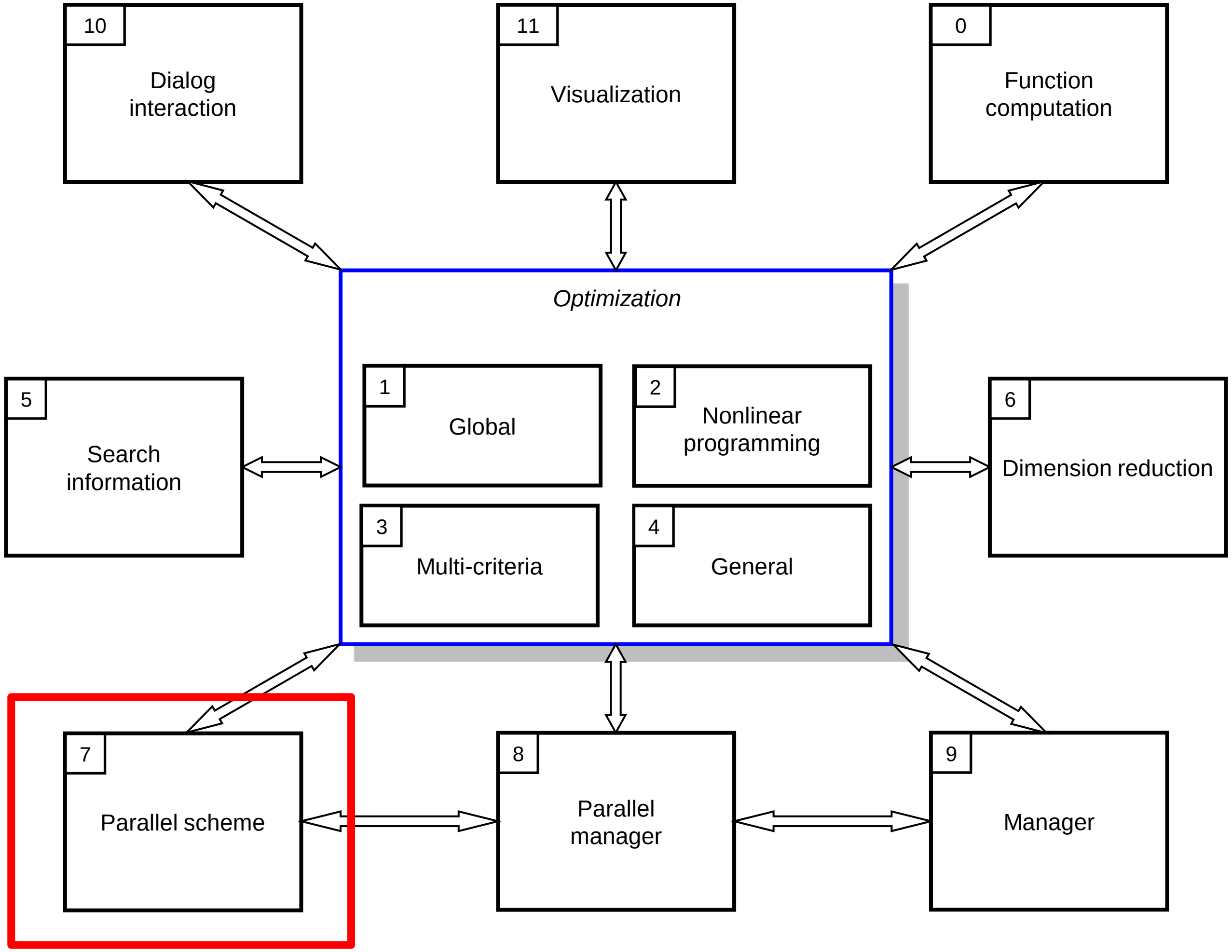
Peano space-filling curve

$\{ y(x): 0 \le x \le 1 \} = \{ y \in R^N: y \in D \}$

1-d scale

$$\varphi(y^*) = \min \{ \varphi(y): y \in D, g_j(y) \le 0, 1 \le j \le m \}$$
$$= \min \{ \varphi(y(x)): x \in [0,1], g_j(y(x)) \le 0, 1 \le j \le m \}$$

Globalizer Architecture



Parallel scheme

Options for parallelization:

- to carry out search domain decomposition;
- to parallelize calculation of the problem functions;
- to parallelize implementation of the algorithm computing rules for selection of the next trial point;
- **to change the algorithm for the purpose of carrying out several trials in parallel.**

Parallel scheme

- Consider a vector of parallelization degrees

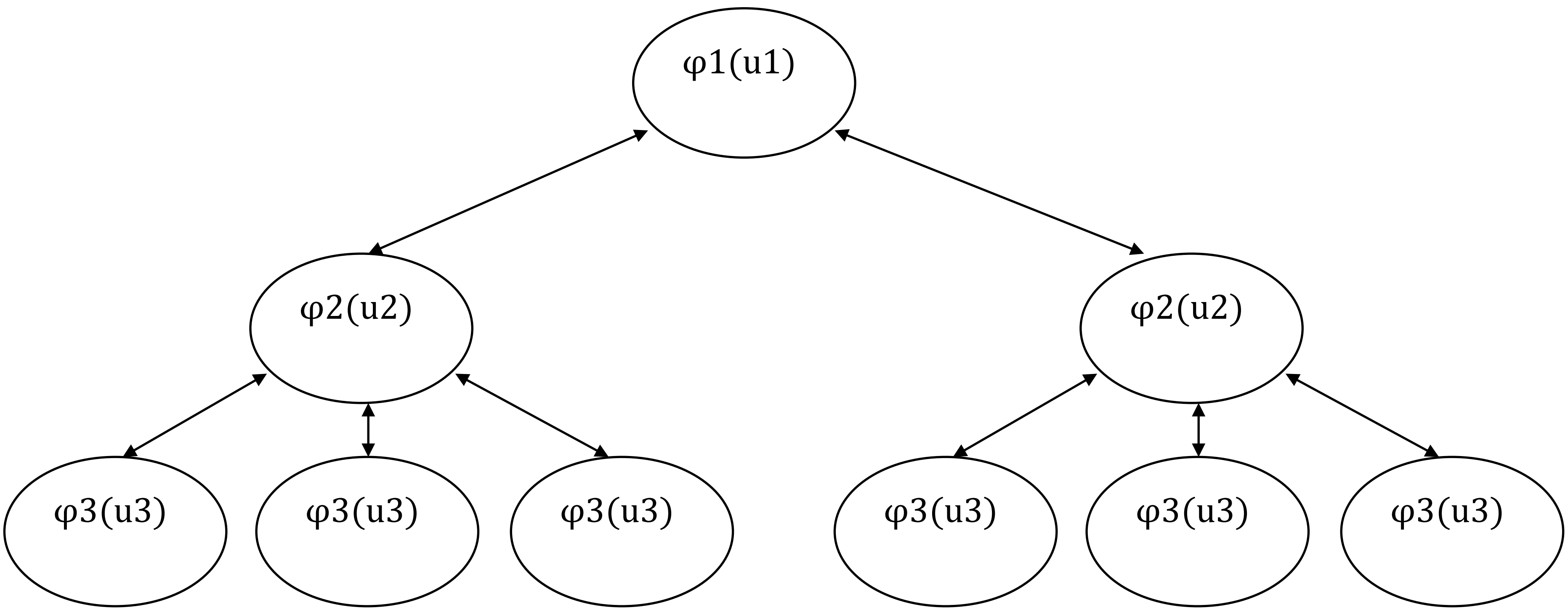
$$\pi = (\pi_1, \pi_2, \ldots, \pi_M)$$

- For the macro-variable u_i, the number π_i means the number of parallel trials at i-th level
- The total number of processors used will be

$$\Pi = 1 + \sum_{i=1}^{M-1} \prod_{j=1}^{i} \pi_j$$

Parallel scheme

- Process Tree (example, M = 3)



Parallel computing for distributed memory

Let's consider the *multiple mapping* $Y(x)$

$$Y(x)=\{y^1(x), y^2(x), \ldots, y^L(x)\}.$$

where $y_i(x)$ is transformed Peano curve.

Use of multiple mapping forms the set of L problems

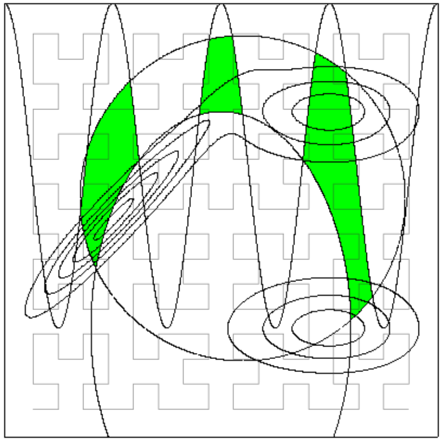
$$\min\{\varphi(y^l(x)): x\in[0,1], g_j(y^l(x))\le 0, 1\le j\le m\}, 1\le l\le L.$$

They can be solved in parallel.

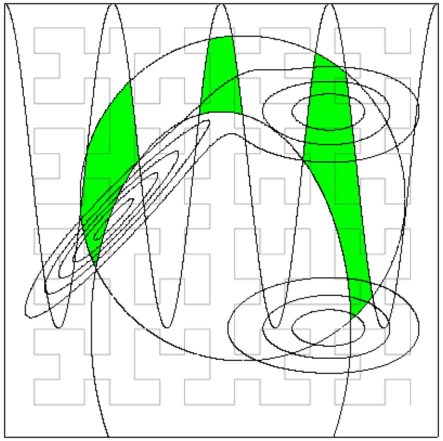
Each one-dimensional problem is solved on a separate processor/node.

Any computed value $z_i=g_\nu(y^l(x_i))$ for the problem l can be transformed to the value $z_j=g_\nu(y^k(x_j))$ for the problem k without time-consuming computing of the function $g_\nu(y)$.

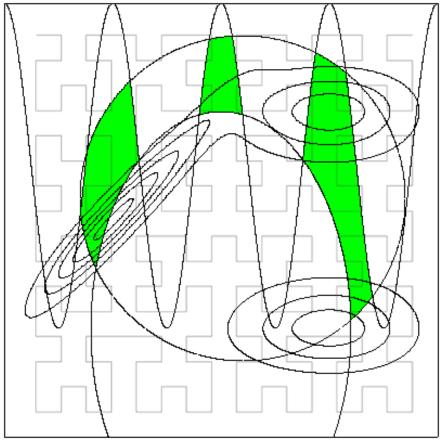
Parallel computing for distributed memory



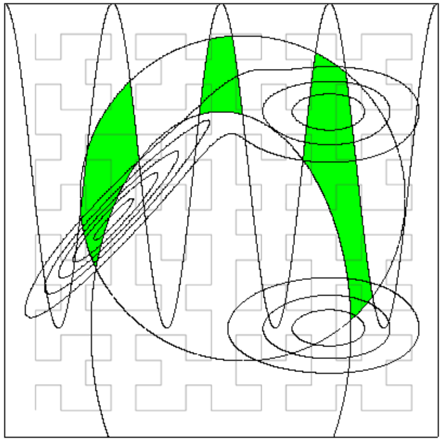
$\min\{\varphi(y^1(x)):x\in[0,1],\ g_j(y^1(x))\le 0,\ 1\le j\le m\}$



$\min\{\varphi(y^2(x)):x\in[0,1],\ g_j(y^2(x))\le 0,\ 1\le j\le m\}$

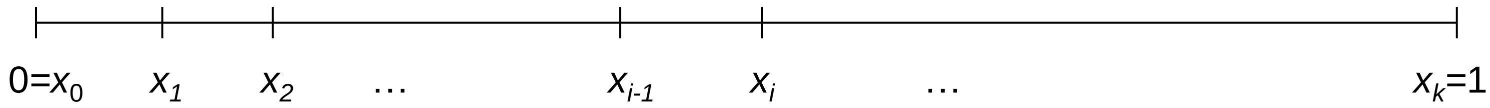


$\min\{\varphi(y^3(x)):x\in[0,1],\ g_j(y^3(x))\le 0,\ 1\le j\le m\}$



$\min\{\varphi(y^3(x)):x\in[0,1],\ g_j(y^3(x))\le 0,\ 1\le j\le m\}$

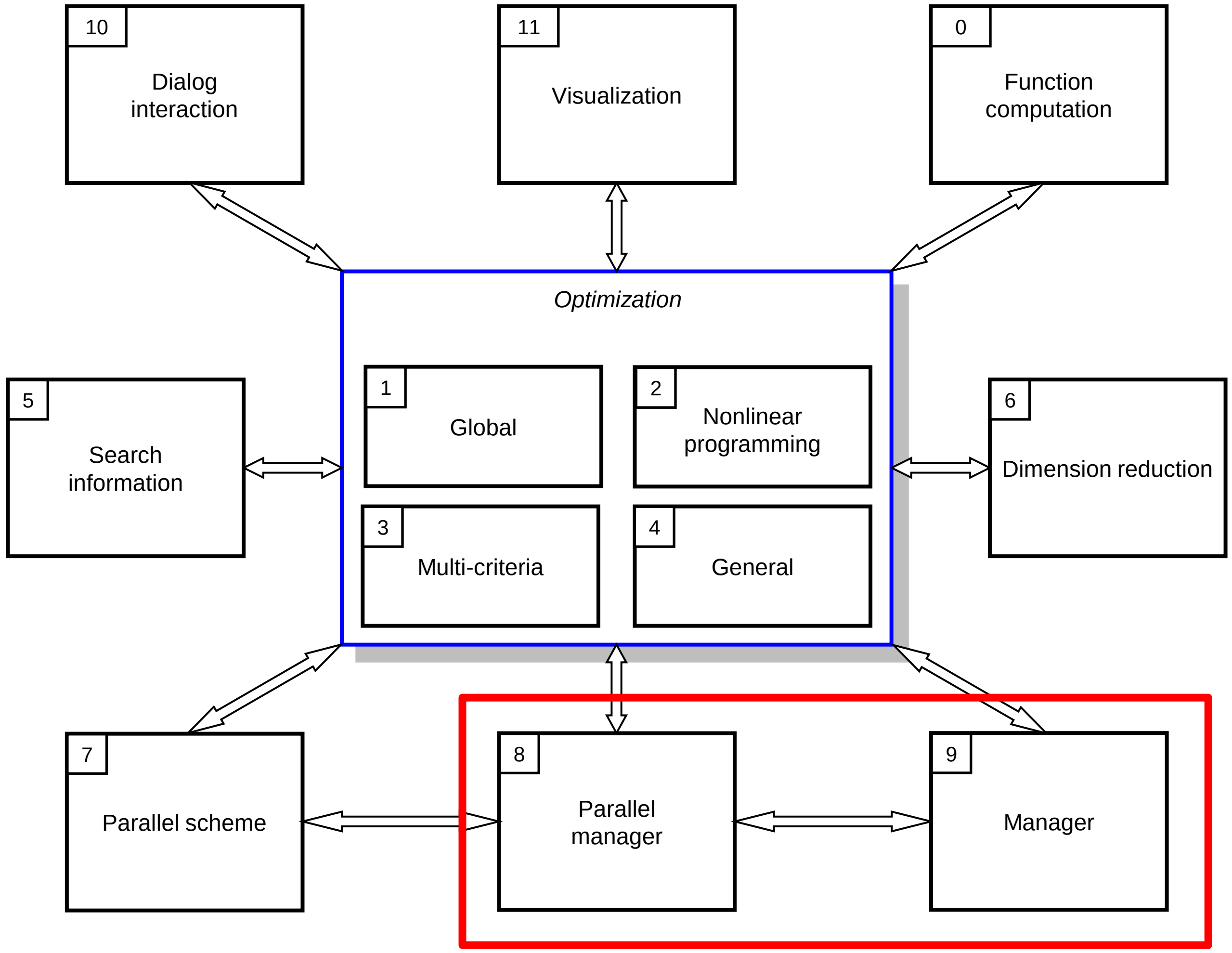
Parallel computing for shared memory



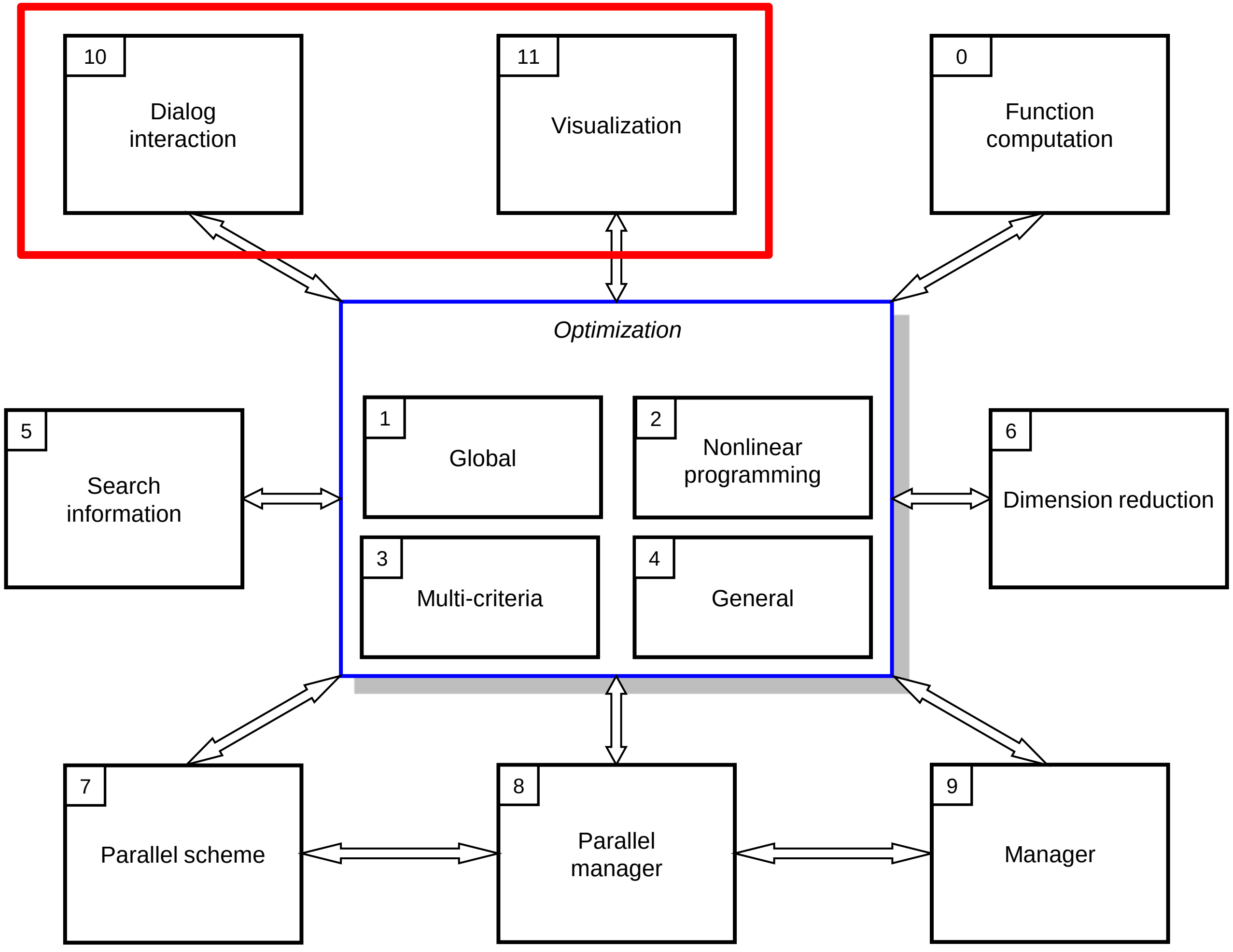
Let $x^0=0, \ldots, x^p=1$.

1. For each (x_{i-1}, x_i), $1 \le i \le k$, calculate *characteristic* $R(i)$.
2. **Sort intervals by characteristic, take p intervals with largest characteristics $R(t_1) > R(t_2) > \ldots > R(t_p)$**
3. **Make next p trials in parallel at internal points of the intervals (x_{t_1-1}, x_{t_1}), (x_{t_2-1}, x_{t_2}),… (x_{t_p-1}, x_{t_p})**
4. Check stop condition: $\Delta_{t_i} \le \varepsilon$

Globalizer Architecture



Globalizer Architecture



Numerical Results

- Vibration Isolation Problem

$$
\begin{aligned}
&\ddot{\xi}_1 = -\beta(\dot{\xi}_1 - \dot{\xi}_2) - \xi_1 + \xi_2 + u + v, \\
&\ddot{\xi}_2 = -\beta(\dot{\xi}_2 - \dot{\xi}_1) - \xi_2 + \xi_1 + v, \\
&\xi_1(0) = \xi_2\,(0) = 0, \qquad \dot{\xi}_1(0) = \dot{\xi}_2\,(0) = 0.
\end{aligned}
\tag{1}
$$

ξ_1 and ξ_2 are coordinates of the material points,

v is the base acceleration up to sign (the external excitation),

u is the control force,

β is a positive damping parameter.

Numerical Results

- Let's rewrite (1) in the standard form

$$
\begin{aligned}
&\dot{x_1} = x_3, \\
&\dot{x_2} = x_4, \\
&\dot{x_3} = -x_1 + x_2 - \beta x_3 + \beta x_4 + v + u, \qquad\qquad (2) \\
&\dot{x_4} = x_1 - x_2 + \beta x_3 - \beta x_4 + v, \\
&x_1(0) = x_2(0) = x_3\,(0) = x_4\,(0) = 0.
\end{aligned}
$$

- This model can describe the typical situations of vibration isolation for devices, apparatuses and humans located on moving vehicles.

Numerical Results

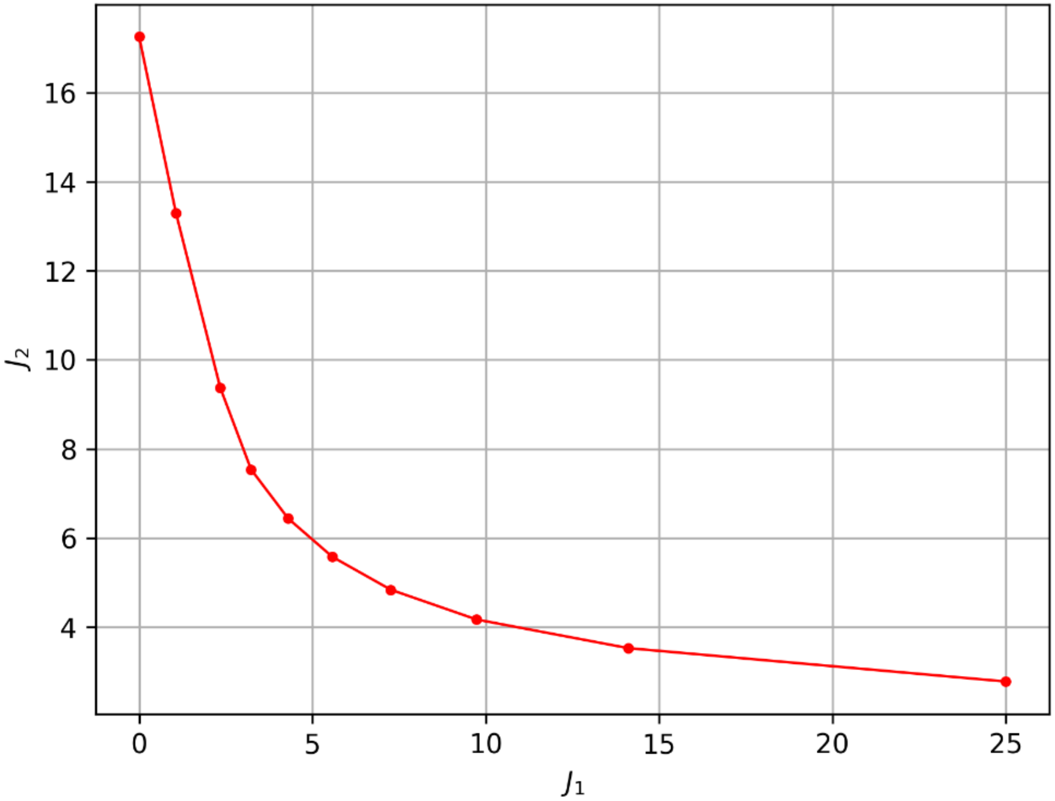
- Choose two criteria for this system to characterize the process of vibration isolation

$$J_1(u) = \sup_{v \in L_2} \frac{\sup_{t \geq 0} |x_1(t)|}{\|v\|_2},$$

$$J_2(u) = \sup_{v \in L_2} \frac{\sup_{t \geq 0} |x_2(t) - x_1(t)|}{\|v\|_2}.$$

Numerical Results

- Consider two-objective control problem for state-feedback case.
- The Pareto optimal front computed by Globalizer



Thank you for attention

References

1. R.G. Strongin,Ya.D. Sergeyev (2000) Global optimization with non-convex constraints. Sequential and parallel algorithms. Dordrecht: Kluwer Academic Publishers.
2. Barkalov, K.A., Strongin, R.G.: A global optimization technique with an adaptive order of checking for constraints (2002) *Comput. Maths. Math. Phys.* 42(9), pp. 1289–1300.
3. Ya.D. Sergeyev, R.G.Strongin, D. Lera (2013) Introduction to global optimization exploiting space-filling curves. Springer.
4. K. Barkalov, V. Gergel, I. Lebedev. Use of Xeon Phi Coprocessor for Solving Global Optimization Problems // V. Malyshkin (ed.): PaCT 2015, Lecture Notes in Computer Science, vol. 9251, pp. 307-318 (2015)
5. K. Barkalov, V. Gergel. Parallel global optimization on GPU // Journal of Global Optimization. vol. 66 (1), pp. 3-20 (2016)