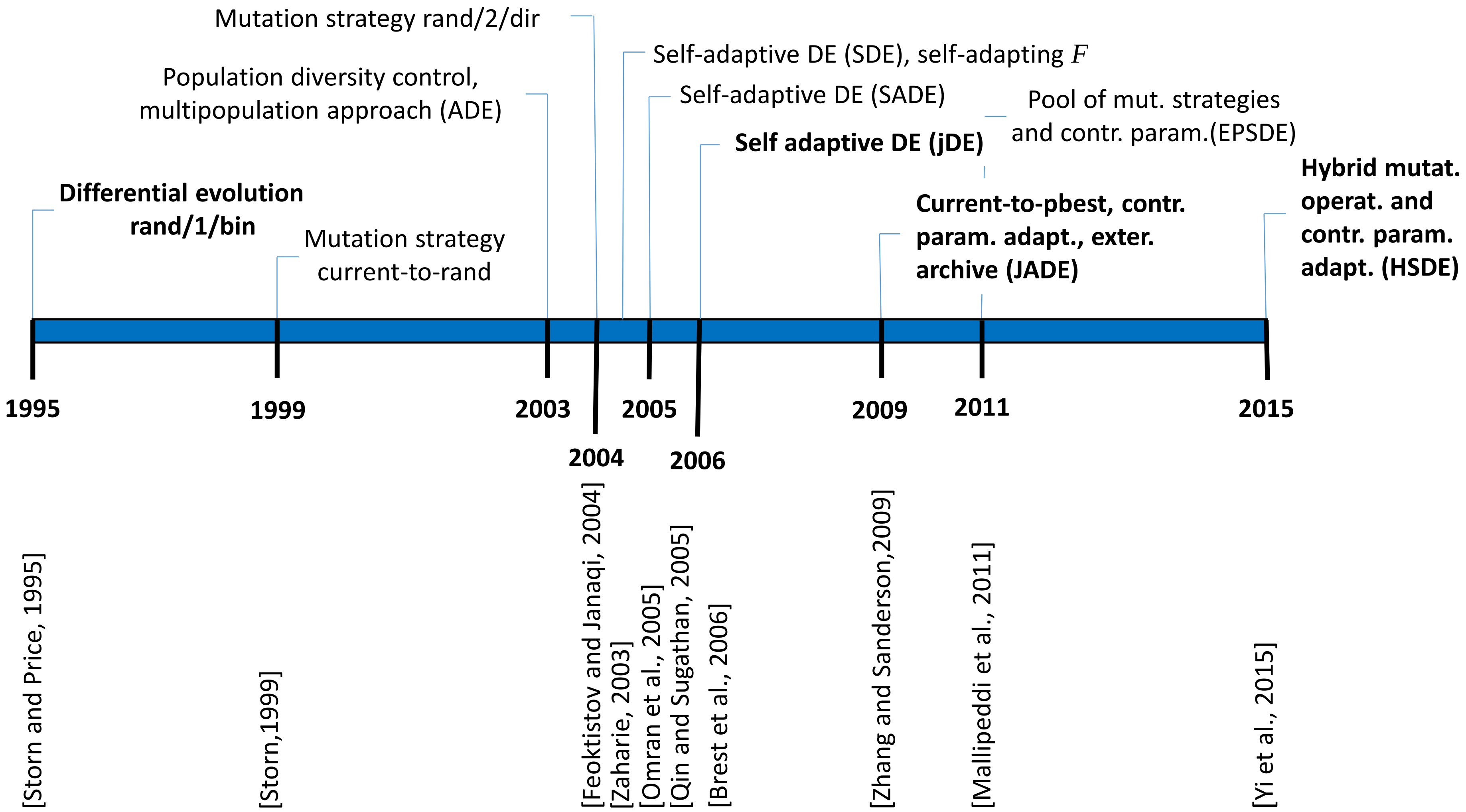
Differential Evolution

Lenka Skanderová

VSB – Technical University of Ostrava
Faculty of Electrical Engineering and Computer Science
Department of Computer Science

State of the Art (1995 - 2015)

Differential Evolution



History

- DE was described in 1995 by Storn and Price
- 1996 – it was demonstrated at the First International Contest on Evolutionary Optimization, where the DE won the third place for proposed benchmark.
- 1997 – Price presented DE at the Second International Contest on Evolutionary
- 1999 – the DE was summarized in the compendium New Ideas in Optimization
- Price and Storn described various applications of the DE, for example:
  - System design by constraint adaptation and differential evolution
  - Designing digital filters with differential evolution
- Research of Storn and Price has been early followed by many authors from all over the world

Principle I

- Population of individuals

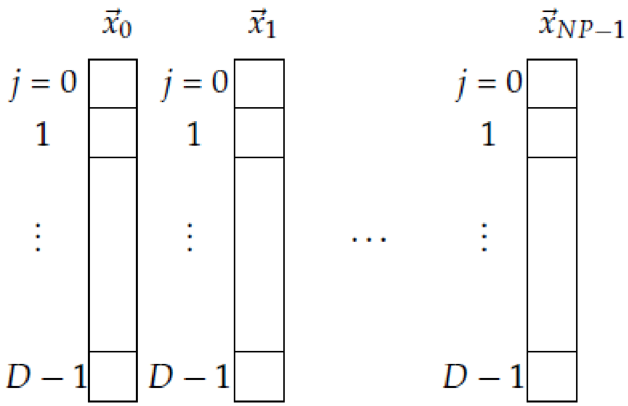
- Mutation, crossover, natural selection
  - Individuals are mutated before the crossover operator is applied -> difference between DE and GA
  - Two types of crossover – exponential and **binomial**

- Three control parameters:
  - Scaling factor F
  - Crossover rate CR
  - Population size NP
  - **Control parameters settings is problem dependent task** -> for this reason adaptation and self-adaptation of control parameters

Principle II

- Population of individuals

$$\vec{x}_i^G = \{x_{i,0}^G, \ldots, x_{i,D-1}^G\}$$
$$P_x^G = \{\vec{x}_0^G, \ldots, \vec{x}_{NP-1}^G\}$$



Initial population is generated randomly within the space of possible solutions

Principle III

- For each individual (target vector) $\vec{x}_i^G$, three random individuals are selected as follows: $\vec{x}_{r_1}^G \neq \vec{x}_{r_2}^G \neq \vec{x}_{r_3}^G \neq \vec{x}_i^G$
- Mutation operation:

scaling factor

$$\vec{v}_i^G = \vec{x}_{r_1}^G + F \cdot (\vec{x}_{r_2}^G - \vec{x}_{r_3}^G)$$

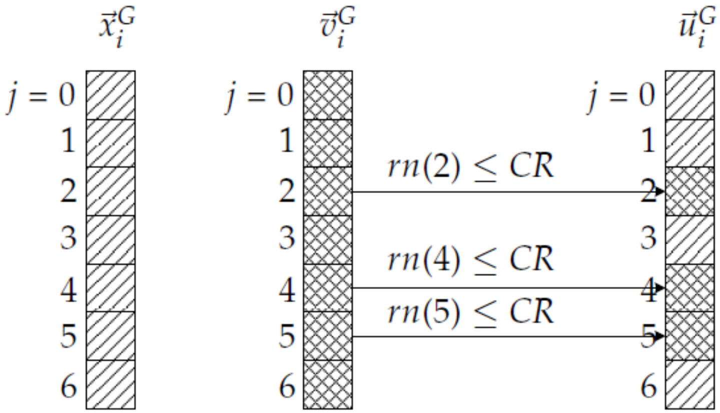
- Crossover operation (binomial):

crossover rate

Random number from interval [0,1]

$$u_{i,j}^G = \begin{cases} v_{i,j}^G & \text{if } rn(j) \leq CR \text{ or } j = j_{rn} \\ x_{i,j}^G & \text{otherwise}. \end{cases}$$

random position



Principle IV

- Natural selection:

trial vector

$$\vec{x}_i^{G+1} = \begin{cases} \vec{u}_i^G & \text{if } f(\vec{u}_i^G) \leq f(\vec{x}_i^G) \\ \vec{x}_i^G & \text{otherwise}\,, \end{cases}$$

target vector

Better individual survives to the next generation

We accept a trial vector even in the case that its fitness is the same as the fitness of a target vector (because of the flat surface)

Different Mutation Strategies

- DE/best/1:

individual with the best fitness

$$\vec{v}_i^G = \vec{x}_{best}^G + F \cdot (\vec{x}_{r_1}^G - \vec{x}_{r_2}^G)\,,$$

- DE/target-to-best/1:

$$\vec{v}_i^G = \vec{x}_i^G + F \cdot (\vec{x}_{best}^G - \vec{x}_i^G) + F \cdot (\vec{x}_{r_1}^G - \vec{x}_{r_2}^G)\,,$$

- DE/best/2:

$$\vec{v}_i^G = \vec{x}_{best}^G + F \cdot (\vec{x}_{r_1}^G - \vec{x}_{r_2}^G) + F \cdot (\vec{x}_{r_3}^G - \vec{x}_{r_4}^G)\,,$$

- DE/rand/2:

$$\vec{v}_i^G = \vec{x}_{r_1}^G + F \cdot (\vec{x}_{r_2}^G - \vec{x}_{r_3}^G) + F \cdot (\vec{x}_{r_4}^G - \vec{x}_{r_5}^G)$$

Control Parameters Settings

- Problem dependent task
- NP:
  - Storn and Price suggested values between $5\mathrm{D} - 10\mathrm{D}$
- F:
  - is strictly greater than zero and the value 0.5 is considered to be a good initial choice
  - If F is set to larger values, the probability of escaping from the local optimum will be increased, however, if the scale factor value is greater than 1.0 the convergence speed will decrease
- CR:
  - depends among the others on the objective function
  - for separable functions, a value of CR between 0.0 and 0.2 is the best choice
  - for multi-modal parameter dependent problems CR in the range [0.9, 1.0] is suggested

Adaptive and Self-adaptive DE

jDE

- Brest et al. **jDE** – Self-adaptive DE
- Each individual has its own scale factor F and crossover rate CR
- New control parameters before mutation of a target vector:

$$F_i^{G+1} = \begin{cases} F_l + r_1 F_u & \text{if } r_2 < \tau_1 \\ F_i^G & \text{otherwise}, \end{cases}$$

set to 0.1

$$CR_i^{G+1} = \begin{cases} r_3 & \text{if } r_4 < \tau_2 \\ CR_i^G & \text{otherwise}, \end{cases}$$

random real numbers from interval [0,1] (uniform distr.)

- At the beggining F and CR generated randomly: $F \in [0.1, 1.0]$ $CR \in [0,1]$

DE Using More Mutation Strategies I

CoDE

- EPSDE, CoDE, HSDE

- **CoDE:**
  - Three mutation strategies:
    - DE/rand/1/bin
    - DE/rand/2/bin
    - DE/current-to-rand/1
  - Pool of three control parameters settings:
    - [F = 1.0, CR = 0.1]
    - [F = 1.0, CR = 0.9]
    - [F = 0.8, CR = 0.2]
  - Three trial vectors generated:
    - The best trial vector replaces the target vector if its fitness is equal or better

DE Using More Mutation Strategies II

HSDE

- **HSDE:**
  - Each individual has its own control parameters F and CR
  - For each target vector five mutually different individuals (solution vectors) are randomly selected
  - Mutation operator:

$$\vec{v}_i^G = \begin{cases} \vec{x}_{r_3}^G + F_i^G \cdot (\vec{x}_{r_4}^G - \vec{x}_{r_5}^G) & \text{if } f(\vec{x}_i^G) > f(\vec{x}_{r_1}^G) \text{ and } f(\vec{x}_i^G) > f(\vec{x}_{r_2}^G) \\ \vec{x}_i^G + F_i^G(\vec{x}_{best}^G - \vec{x}_i^G) + F_i^G \cdot (\vec{x}_{r_4}^G - \vec{x}_{r_5}^G) & otherwise\,, \end{cases}$$

exploration

exploitation

  - Control parameters adaptation:

$$F_i^{G+1} = \begin{cases} F_l + rand_1 \cdot F_u, & \text{if } f(\vec{u}_i^G) > f(\vec{x}_i^G) \\ F_i^G & otherwise\,, \end{cases}$$

$$CR_i^{G+1} = \begin{cases} rand_2, & \text{if } f(\vec{u}_i^G) > f(\vec{x}_i^G) \\ CR_i^G & otherwise\,, \end{cases}$$

DE Using Novel Mutation Strategies II

JADE I

- **JADE:**
  - Each individual has its own control parameters F and CR
  - External archive A
  - Novel mutation strategy:
    - *Current-to-pbest/1* , where p is usually 5% best individuals of a population (explitation on more than one place)

$$\vec{v}_i^G = \vec{x}_i^G + F_i \cdot (\vec{x}_{best}^{p,G} - \vec{x}_i^G) + F_i \cdot (\vec{x}_{r_1}^G - \vec{x}'^G_{r_2})$$

  - where $\vec{x}'^G_{r_2}$ is selected from the union $A \cup P$ and P denotes a population
  - At the beginning of the algorithm, the archive A is empty. Then, after each generation G, parent solutions that failed in the selection process are added to the archive. If the archive size exceeds a certain threshold, some solutions are randomly selected and removed

DE Using Novel Mutation Strategies II

JADE II

- **JADE:**
  - Scale factor is recomputed as:

$$F_i = randc_i(\mu_F, 0.1)$$

where $\mu_F = (1 - c) \cdot \mu_F + c \cdot mean_L(S_F)$ and $mean_L(S_F) = \frac{\sum_{F \in S_F} F^2}{\sum_{F \in S_F} F}$

  - Crossover rate is recomputed as:

$$CR_i = randn_i(\mu_{CR}, 0.1)$$

where $\mu_{CR} = (1 - c) \cdot \mu_{CR} + c \cdot mean_A(S_{CR})$

Archives of successful control parameters from actual generation

arithmetic mean

DE Using More Populations

Clustering, k-means etc.

- Population divided into sub-populations
- For each subpopulation, another mutation strategy is used.
- Currently, trend that each individual has its own mutation strategy and its control parameters are self adapted automatically
- Mutation strategies competition:
  - The most successful mutation strategy will get the most computation resources