Parameter Control Mechanisms in Differential Evolution: A Tutorial Review and Taxonomy

Tsung-Che Chiang, Cheng-Nan Chen, and Yu-Chieh Lin

Department of Computer Science and Information Engineering
National Taiwan Normal University

2013.04.18 IEEE Symposium Series on Computational Intelligence (SSCI 2013) @Singapore

Outline

- Introduction
- Proposed Taxonomy
- A Quick Review
- Research Directions

R. Storn and K. Price, "Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces," Journal of Global Optimization, vol. 11, pp. 341 – 359, 1997.

Differential Evolution (DE)

Each individual x_i serves as the target vector once.

For each target vector, a mutant vector v_i is generated by

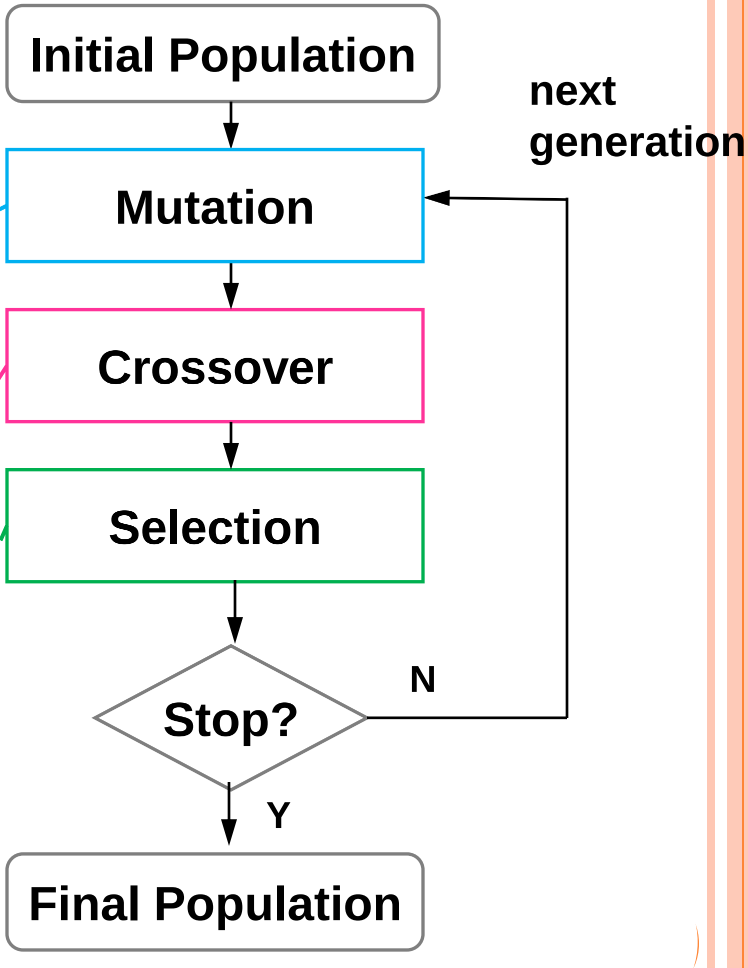
$$v_i = x_{r1} + F \cdot (x_{r2} - x_{r3})$$

A trial vector u_i is generated by

$$u_{ij} = \begin{cases} v_{ij}, & \textbf{if } U_j(0,1) \le CR \vee j = j_{rnd} \\ x_{ij}, & \textbf{otherwise} \end{cases}$$

$$j = 1,2,\ldots,D$$

The trial vector is accepted if it is not worse than the target vector.

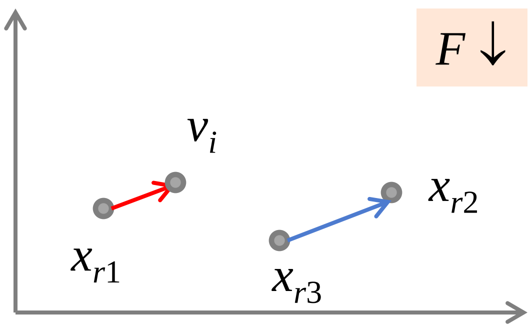


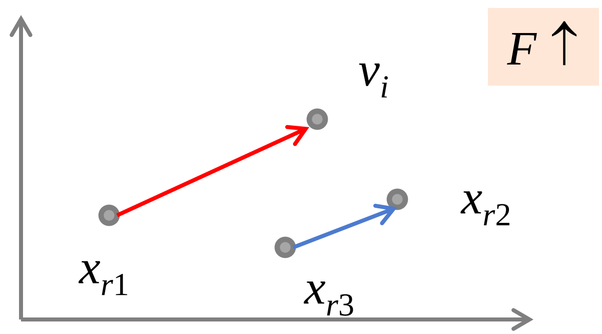
- Population size (NP)

Parameters of DE

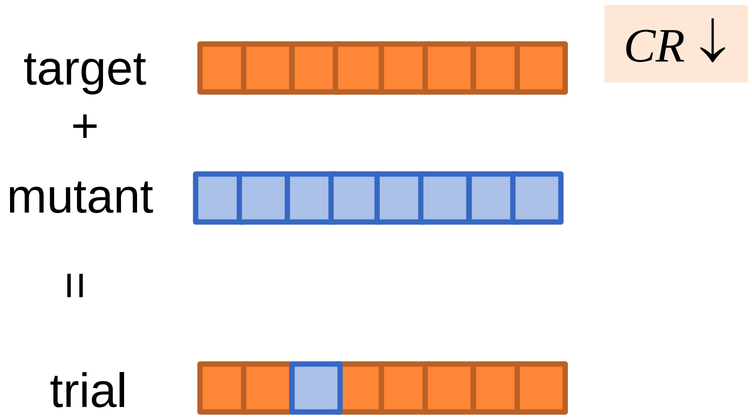
Scaling Factor (F)

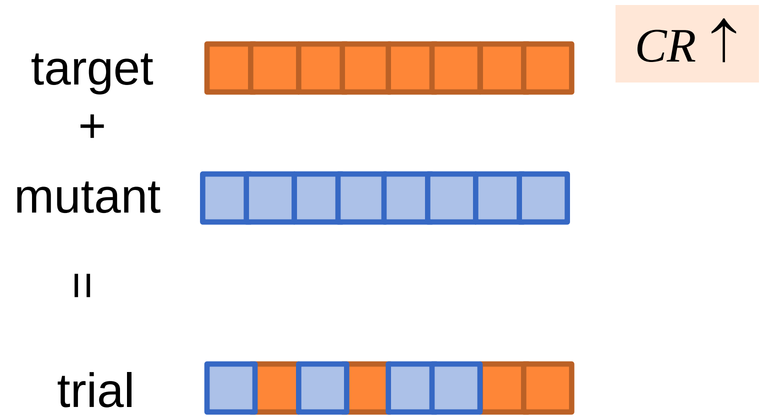
$$v_i = x_{r1} + \boldsymbol{F} \cdot (x_{r2} - x_{r3})$$





Crossover Rate (CR)





Parameters of DE

- **Suggestions on parameter values**

F	CR	Reference
[0.4, 1.0]	0.1, 0.9	Storn & Price1997
0.9	0.9	Liu & Lampinen 2002
	0.5	Ali & Törn 2004
0.9	0.9	Rönkkönen et al. 2005
0.5	0.5	Kaelo & Ali 2006
≥ 0.6	≥ 0.6	Zielinski et al. 2006
[0.35, 0.37]	0.5	Salman et al. 2007

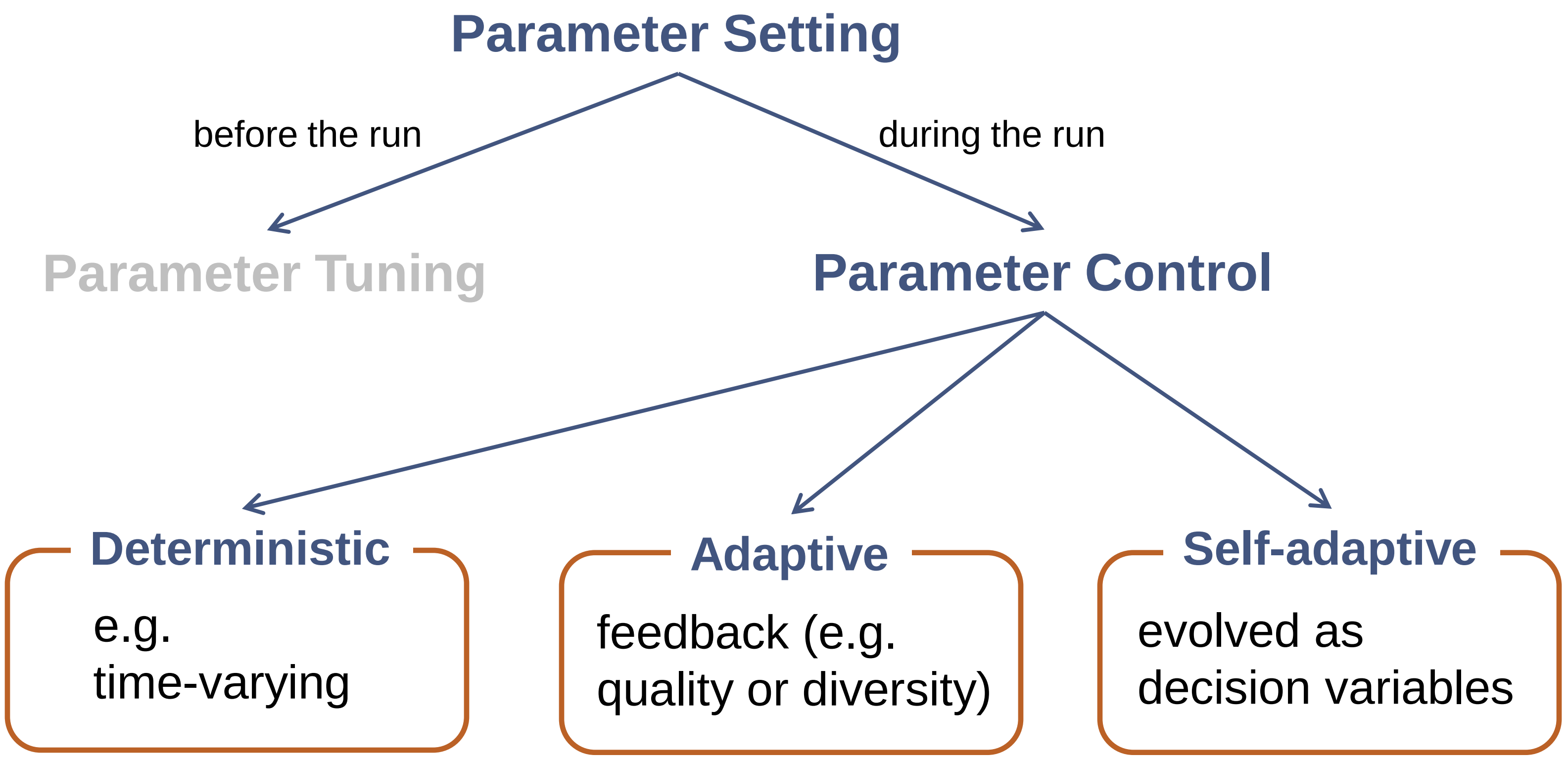
Parameters of DE

- Different parameter values are required for different
  - problem instances
  - mutation strategies
  - search stages
  - search regions

- We need to determine the parameter values dynamically (online).

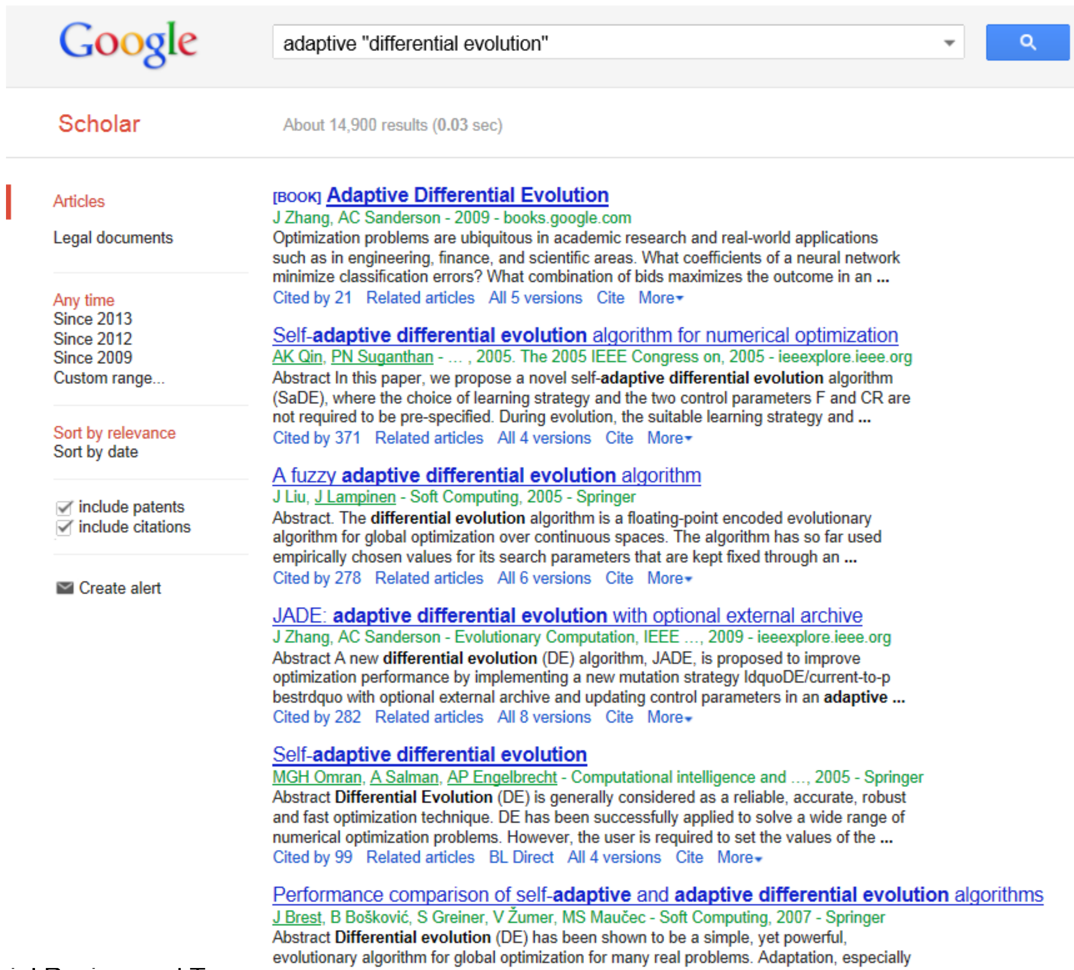
Eiben, Hinterding, and Michalewicz, “Parameter control in evolutionary algorithms,” IEEE Transactions on Evolutionary Computation, vol. 3, no.2, pp. 124 – 141, 1999.

Parameter Control Paradigms



Parameter Control Paradigms

- **There are many "adaptive" or "self-adaptive" DE algorithms.**



R. Storn and K. Price, “Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces,” Journal of Global Optimization, vol. 11, pp. 341 – 359, 1997.

Proposed Taxonomy

- **3-field notation for the DE variants**

the number of difference vectors

the vector to be mutated

the crossover scheme

x / y / z

rand/1/bin

best/2/bin

current-to-rand/1/bin

rand/1/exp

rand/2/exp

Proposed Taxonomy

- 3-field notation for the parameter control mechanisms

x / y / z

- **the number of candidate values of a parameter**
  - **dis** (discrete) **/ con** (continuous)
- **the number of parameter values used in a single generation**
  - **1** / **mul** (multiple) / **idv** (individual) / **var** (variable)
- **information used to adjust parameter values**
  - **rnd** (random) / **pop** (population) / **par** (parent) / **idv** (individual)

Proposed Taxonomy

- **We classify 23 recent studies into nine categories.**

Category	Algorithms
con / **1** / **pop**	DEPD, FADE, ADEA
con / **mul** / **rnd**	NSDE
con / **mul** / **pop**	SaDE, SaNSDE, JADE, JADE2, SaJADE
con / **idv** / **rnd**	jDE, jDE-2, CSDE, MOSADE
con / **idv** / **pop**	RADE, ISADE
con / **idv** / **par**	SPDE, SDE, DESAP, DEMOsWA
con / **idv** / **idv**	SFLSDE, SspDE
con / **var** / **pop**	APDE
dis / **mul** / **pop**	DEBR

M. M. Ali and A. Tőrn, "Population set-based global optimization algorithms: some modifications and numerical results," Computers & Operations Research, vol. 31, no. 10, pp. 1703–1725, 2004.

con / 1 / pop

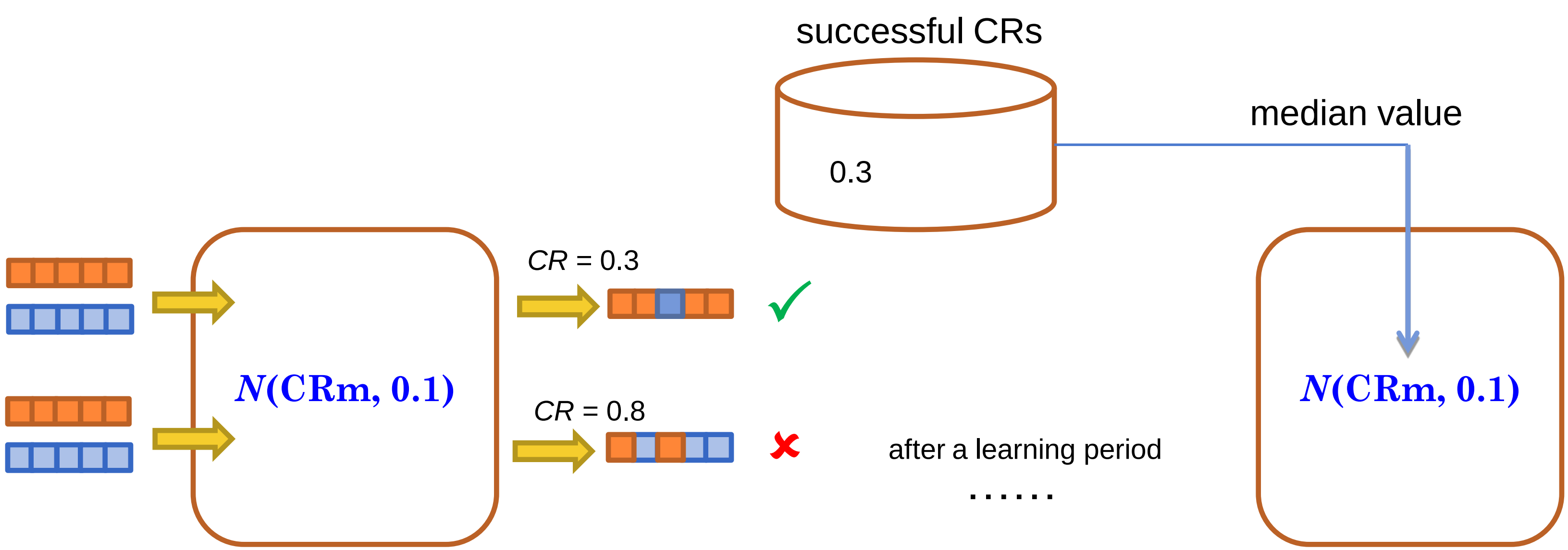
- **Example: DEPD (Ali and Tőrn, 2004)**
  - **Fixed *CR* value**
  - **Increase *F* as the difference of fitness between the best and worst individuals decreases**

$$F = \begin{cases} \max\{F_{\min}, 1 - \left| f_{\max} / f_{\min} \right|\}, & \text{if } \left| f_{\max} / f_{\min} \right| < 1, \\ \max\{F_{\min}, 1 - \left| f_{\min} / f_{\max} \right|\}, & \text{otherwise.} \end{cases}$$

- **Other population-based information**
  - **diversity (genotypic or phenotypic distance) within a generation or between generations (e.g. FADE)**
  - **distribution along the Pareto front (e.g. ADEA)**

con / mul / pop

- **Example: SaDE (Qin et al. 2009)**
  - **Generate *F*/*CR* values by normal distribution**
    - ***F*~*N*(0.5, 0.5)**
  - **Record the successful *CR* values in a learning period**
    - **successful: the trial vector is accepted**



con / mul / pop

- **Design issues**
  - **random distributions for generating the parameter value**
    - **normal / Cauchy (e.g. SaDE / SaNSDE & JADE)**
  - **learning period (*LP*)**
    - ***LP* = 1 / *LP* > 1**
  - **calculation of distribution parameters**
    - **median / weighted sum (e.g. SaDE / SaNSDE & JADE)**

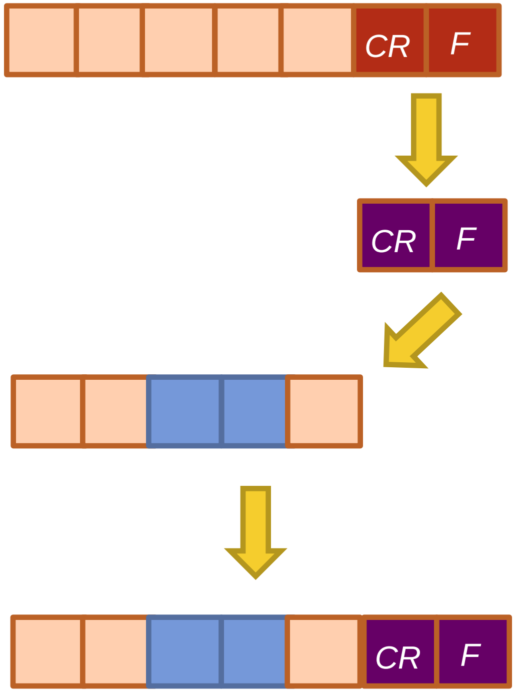
J. Brest, S. Greiner, B. Bošković, M. Mernik, and V. Žumer, "Self-adapting control parameters in differential evolution: a comparative study on numerical benchmark problems," IEEE Transactions on Evolutionary Computation, vol. 10, no. 6, pp. 646–657,

con / idv / rnd

- Example: **jDE** (Brest et al. 2006)
  - Each individual records its parameter values.
  - The values undergo random perturbation probabilistically.

$$F_{i,g+1} = \begin{cases} U_1(F_{\min}, F_{\max}), & \text{if } U_2(0,1) < \tau_1, \\ F_{i,g}, & \text{otherwise.} \end{cases}$$

$$CR_{i,g+1} = \begin{cases} U_3(0,1), & \text{if } U_4(0,1) < \tau_2, \\ CR_{i,g}, & \text{otherwise.} \end{cases}$$



  - Good parameter values lead to good trial vectors and survive together with the trial vectors.

J. Brest, S. Greiner, B. Bošković, M. Mernik, and V. Žumer, “Self-adapting control parameters in differential evolution: a comparative study on numerical benchmark problems,” IEEE Transactions on Evolutionary Computation, vol. 10, no. 6, pp. 646–657,

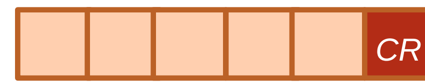
con / idv / rnd

- **Design issues**
  - **random distributions for perturbation**
    - **uniform / Cauchy (e.g. jDE / CSDE)**
  - **population information (⇒ con / idv / pop)**
    - **change the parameter values only for low-quality individuals (e.g. RADE)**
    - **lower the parameter values for high-quality individuals (e.g. ISADE)**
  - **individual information (⇒ con / idv / idv)**
    - **use local search to enhance the values (e.g. SFLSDE)**
    - **record the history of the use of parameter values for each individual (e.g. SspDE)**

H. A. Abbass, “The self-adaptive Pareto differential evolution algorithm,” IEEE Congress on Evolutionary Computation, pp. 831–836, 2002.

con / idv / par

- **Example: SPDE (Abbass 2002)**
  - **Each individual records its parameter values.**
  - **Parameter values evolve as the decision variable values evolve.**



$$CR_i = CR_{r1} + N(0,1) \cdot (CR_{r2} - CR_{r3})$$

- **Design issues**
  - **the way to calculate the new parameter values based on parents’ values (e.g. DEMOwSA)**

$$CR_i = \frac{CR_i + CR_{r1} + CR_{r2} + CR_{r3}}{4} \cdot e^{\tau \cdot N(0,1)}$$

con / var / pop

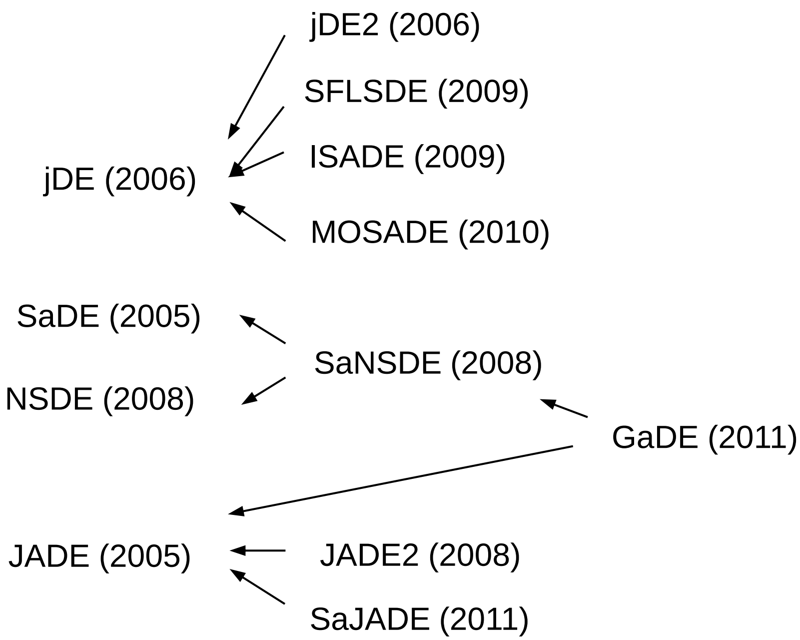
- **Example: APDE** (Zaharie and Petcu 2004)
  - **Parameter values are associated with variables.**
  - **Relationships between the parameter values and the variance of values of decision variables are investigated.**
  - **Values of *CR* and *F* are adjusted alternately.**

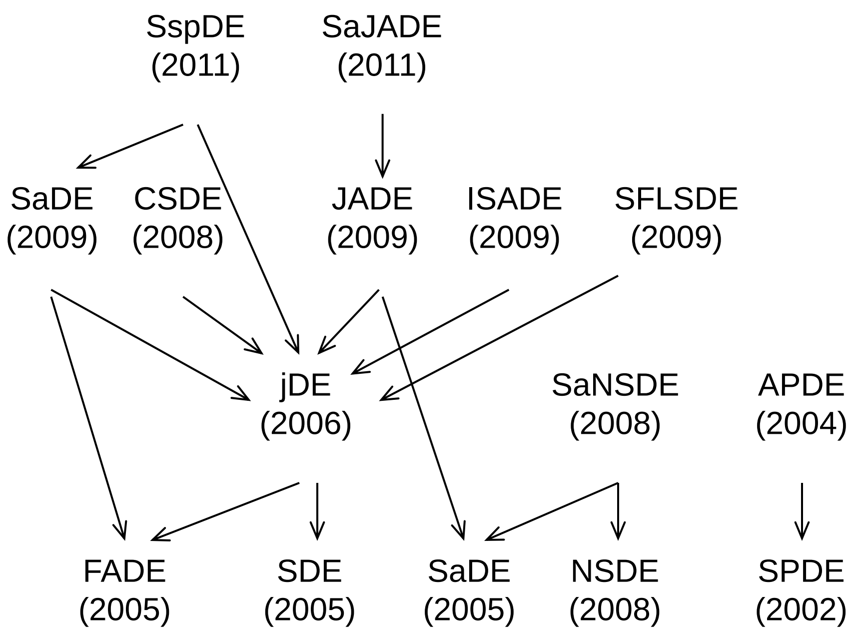
$$CR_i = \begin{cases} -(F_i^2 N - 1) + \sqrt{(F_i^2 N - 1)^2 - N(1 - c_i)}, & \text{if } c_i \geq 1 \\ CR_{\min} & \text{otherwise} \end{cases}$$

$$c_i(g+1) = \gamma \frac{Var(x^i(g))Var(f(g))}{Var(x^i(g+1))Var(f(g+1))}$$

$$Var(f(g)) = \frac{1}{M}\sum_{j=1}^{M} Var(f_j(g))$$

Some Relationships





Design Inheritance

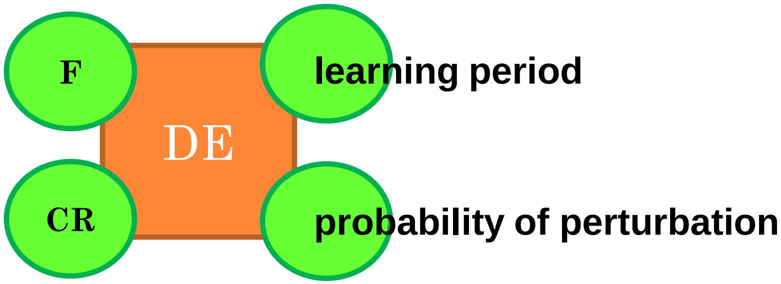
A ← B (derived)

Performance Comparison

A ← B (better)

Research Directions

- **Making the algorithm simpler**
  - **avoid extra parameters like learning period or probability.**

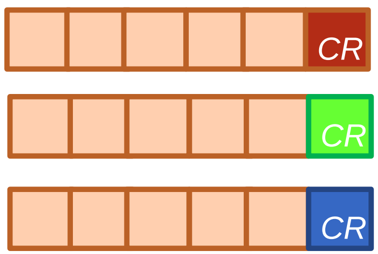


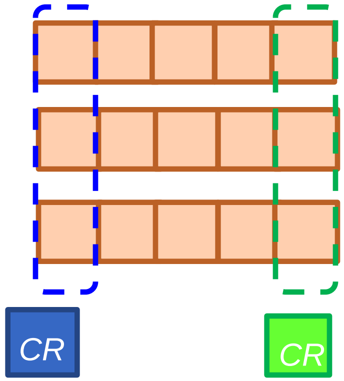
- **Considering problem-oriented information**
  - **Quality & diversity information have been used.**
  - **Problem information (unimodal/multimodal, separable/non-separable, etc.) has rarely been considered.**
  - **$\cdot/\cdot/\texttt{pop}$, $\cdot/\cdot/\texttt{idv}$, then $\cdot/\cdot/\texttt{prb}$?**

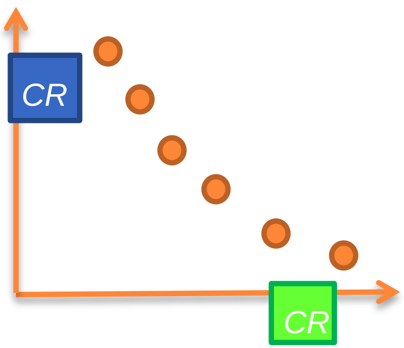
$\Rightarrow$ **LMDE (Takahama and Sakai 2012)**

Research Directions

- Adapting with respect to multiple objectives





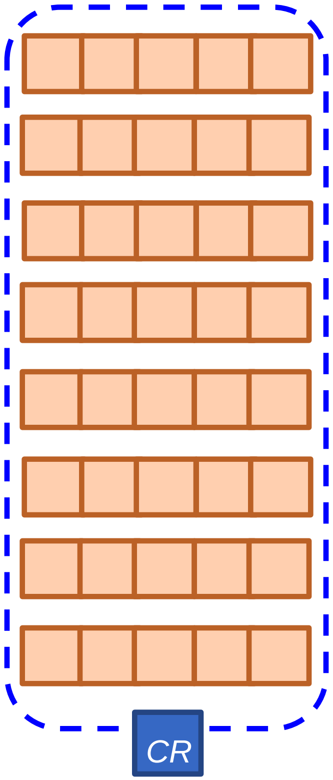


- ·/idv/· ·/var/·, then ·/obj/· ?

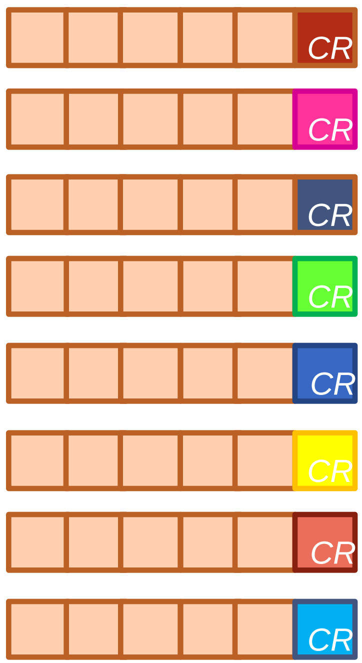
⇒ **OW-MOSaDE** (Huang et al. 2009)

Research Directions

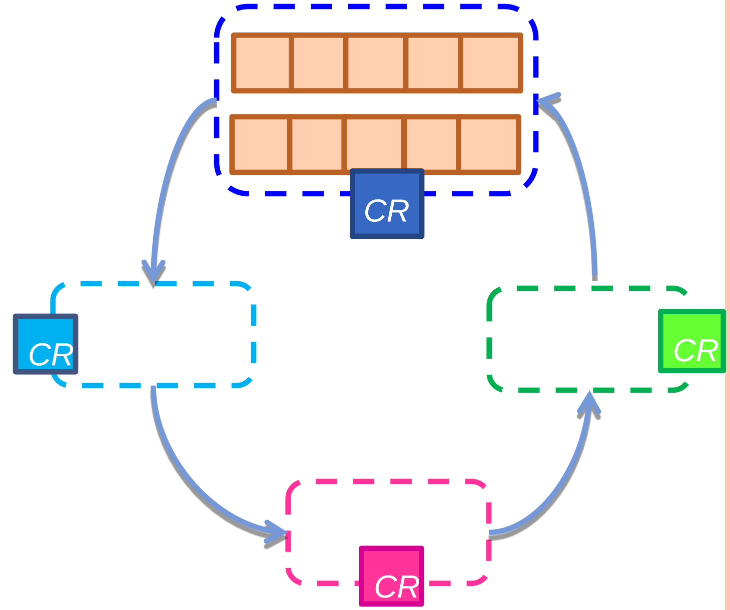
- **Doing parameter control through distributed DE**



·/1/·



·/idv/·,



then ·/sub-pop/· ?

⇒ **FACPDE** (Weber et al. 2010)

Thank you very much for your attention!