Not all parents are equal for MO-CMA-ES

Ilya Loshchilov[1,2], Marc Schoenauer[1,2], Michèle Sebag[2,1]

[1]TAO Project-team, INRIA Saclay - Île-de-France

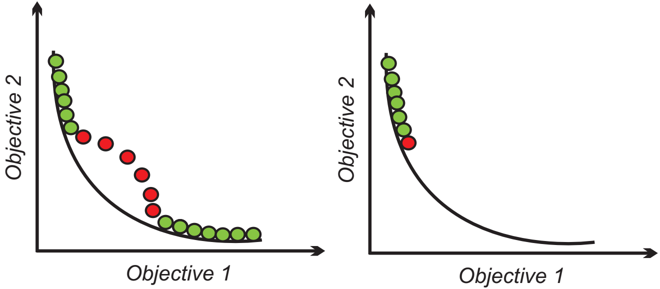
[2]and Laboratoire de Recherche en Informatique (UMR CNRS 8623)
Université Paris-Sud, 91128 Orsay Cedex, France

Kindly presented by Jan Salmen, Universität Bochum
with apologies from the authors

Evolutionary Multi-Criterion Optimization - EMO 2011

Content

Why some parents are better than other ?



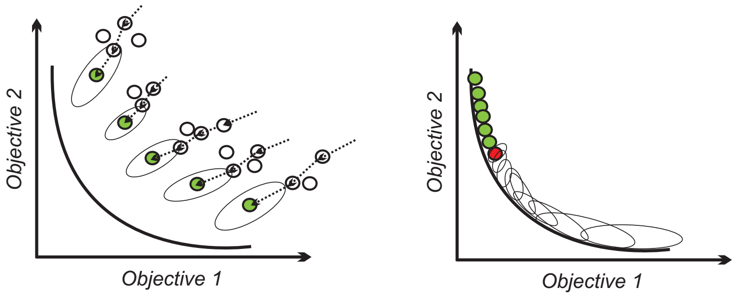
- Hard to find segments of the Pareto front
- Some parents improve the fitness faster than other

Why MO-CMA-ES ?

μ_{MO}-(1+1)-CMA-ES $\equiv \mu \times$ (1+1)-CMA-ES + global Pareto selection

- CMA-ES excellent on single-objective problems (e.g., BBOB)
- In μ_{MO}-(1+1)-CMA-ES, each individual is a (1+1)-CMA-ES

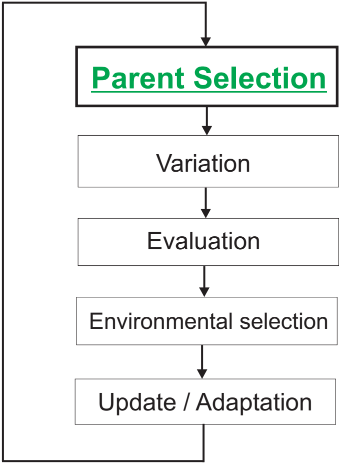
[1]



Two typical adaptations for MO-CMA-ES

[1] C. Igel, N. Hansen, and S. Roth (2005). "The Multi-objective Variable Metric Evolution Strategy"

Evolution Loop for Steady-state MO-CMA-ES



Tournament Selection (quality-based)

A total preorder relation $\prec_X$ is defined on any finite subset X:

$$x \prec_X y \Leftrightarrow \quad PRank(x, X) < PRank(y, X) \qquad \text{// lower Pareto rank}$$

or // same Pareto rank and higher Hypervolume Contribution

$$PRank(x, X) = PRank(y, X) \text{ and } \Delta H(x, X) > \Delta H(y, X)$$

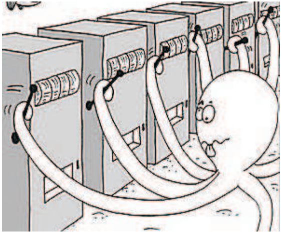
Tournament $(\mu +_t 1)$ selection for MOO

Input: tournament size $t \in \mathbb{N}$; population of μ individuals X
Procedure: uniformly select t individuals from X
Output: the best individual among t according to $\prec_X$ criterion

With $t = 1$: standard steady-state MO-CMA-ES with random selection [a]

[a] C. Igel, T. Suttorp, and N. Hansen (EMO 2007). "Steady-state Selection and Efficient Covariance Matrix Update in the Multi-objective CMA-ES"

Multi-armed Bandits



Original Multi-armed Bandits (MAB) problem

A gambler plays arm (machine) j at time t

and wins reward : $r_{j,t} = \begin{cases} 1 & \text{with prob. p} \\ 0 & \text{with prob. (1-p)} \end{cases}$

Goal: maximize the sum of rewards

Upper Confidence Bound (UCB) [Auer, 2002]

Initialization: play each arm once

Loop: play arm j that maximizes:

$$\bar{r}_{j,t} + C\sqrt{\frac{2\ln \sum_k n_{k,t}}{n_{j,t}}},$$

where $\begin{cases} \bar{r}_{j,t} & \text{average reward of arm } j \\ n_{j,t} & \text{number of plays of arm } j \end{cases}$

Reward-based Parent Selection

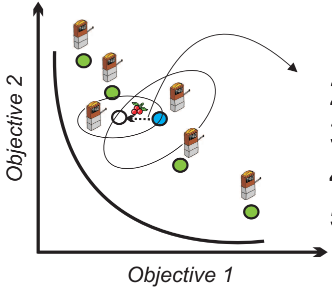
MAB for MOO

$\bar{r}_{j,t}$ is the average reward along a time window of size w

$n_{i,t} = 0$ for new offspring or for an individual selected w steps ago

$$\text{select parent } i = \begin{cases} i \text{ with } n_{i,t} = 0 \text{ if exist,} \\ i = Argmax\left\{\bar{r}_{j,t} + C\sqrt{\frac{2\ln\sum_k n_{k,t}}{n_{j,t}}}\right\} & \text{otherwise} \end{cases}$$

At the moment, $C = 0$ (exploration iff $n_{i,t} = 0$)



1. ● = **ParentSelection()**
2. ○ = ●+mutation
3. Is ○ successful ?
4. Update ● and ○
5. **ComputeRewards()**

Defining Rewards I: $(\mu + 1_{succ})$, $(\mu + 1_{rank})$

Parent a from the population Q^g generates offspring a'.
Both the offspring and the parent receive reward r:

$(\mu + 1_{succ})$

If a' becomes member of new population $Q^{(g+1)}$:

$$r = 1 \text{ if } a' \in Q^{(g+1)}, \text{ and } 0 \text{ otherwise}$$

$(\mu + 1_{rank})$

A smoother reward is defined by the rank of a' in $Q^{(g+1)}$:

$$r = 1 - \frac{rank(a')}{\mu} \text{ if } a' \in Q^{(g+1)}, \text{ and } 0 \text{ otherwise}$$

Defining Rewards II: $(\mu + 1_{\Delta H_1})$, $(\mu + 1_{\Delta H_i})$

$(\mu + 1_{\Delta H_1})$

Set the reward to the increase of the total Hypervolume contribution from generation g to $g+1$:

$$r = \begin{cases} 0 \text{ if offspring is dominated} \\ \sum_{a \in Q^{(g+1)}} \Delta H(a, Q^{(g+1)}) - \sum_{a \in Q^{(g)}} \Delta H(a, Q^{(g)}) \text{ otherwise} \end{cases}$$

$(\mu + 1_{\Delta H_i})$

A relaxation of the above reward, involving a rank-based penalization:

$$r = \frac{1}{2^{k-1}} \left(\sum_{ndom_k(Q^{(g+1)})} \Delta H(a, ndom_k(Q^{(g+1)})) - \sum_{ndom_k(Q^{(g)})} \Delta H(a, ndom_k(Q^{(g)})) \right)$$

where k denotes the Pareto rank of the current offspring, and $ndom_k(Q^{(g)})$ is k-th non-dominated front of $Q^{(g)}$.

Experimental validation

Algorithms

- The steady-state MO-CMA-ES with modified parent selection:
  - 2 tournament-based: $(\mu +_2 1)$ and $(\mu +_{10} 1)$;
  - 4 MAB-based: $(\mu + 1_{succ})$, $(\mu + 1_{rank})$, $(\mu + 1_{\Delta H_1})$ and $(\mu + 1_{\Delta H_i})$.
- The baseline MO-CMA-ES:
  - steady-state $(\mu + 1)$-MO-CMA , $(\mu_{\prec} + 1)$-MO-CMA and generational $(\mu + \mu)$-MO-CMA.

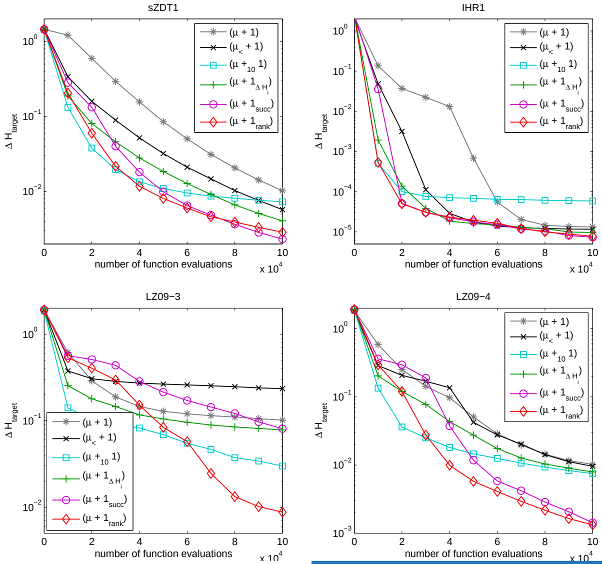
Default parameters ($\mu = 100$), 200,000 evaluations, 31 runs.

Benchmark Problems:

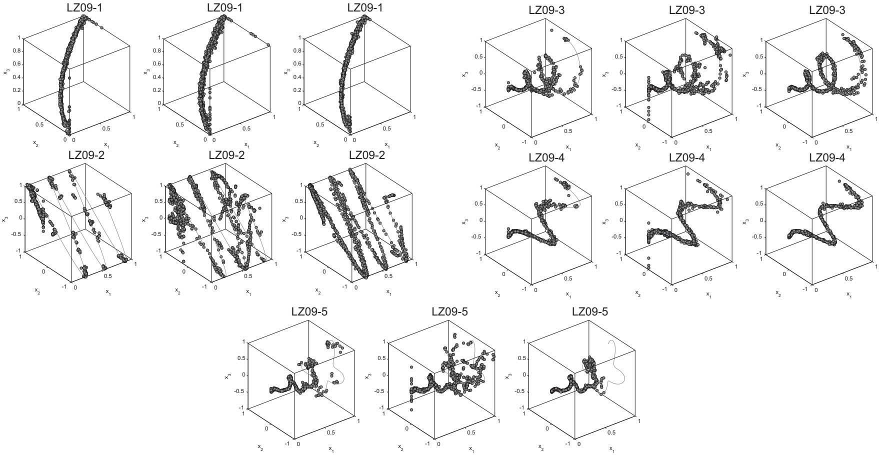
- sZDT1:3-6 with the true Pareto front shifted in decision space: $x'_i \leftarrow |x_i - 0.5|$ for $2 \leq i \leq n$
- IHR1:3-6 rotated variants of the original ZDT problems
- LZ09-1:5 with complicated Pareto front in decision space [2]

[2] H. Li and Q. Zhang. "Multiobjective Optimization Problems With Complicated Pareto Sets, MOEA/D and NSGA-II." IEEE TEC 2009

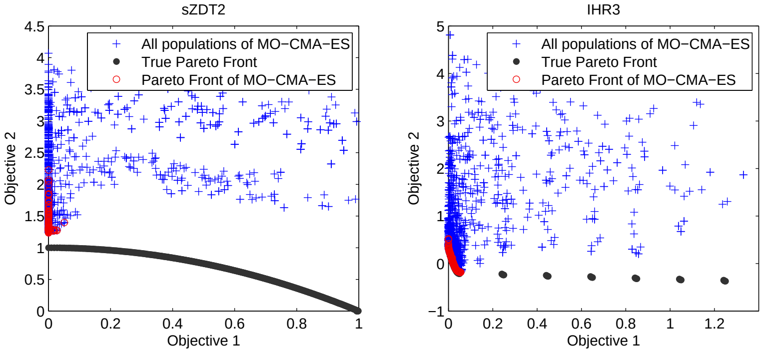
Results: sZDT1, IHR1, LZ3 and LZ4



Results: $(\mu + 1)$, $(\mu +_{10} 1)$, $(\mu + 1_{rank})$ on LZ problems



Loss of diversity



Typical behavior of $(\mu + 1_{succ})$-MO-CMA on sZDT2 (left) and IHR3 (right) problems after 5,000 fitness function evaluations.

Summary

- Speed-up (factor 2-5) with MAB schemes: $(\mu + 1_{rank})$ and $(\mu + 1_{\Delta H_i})$.
- Loss of diversity, especially on multi-modal problems. Too greedy schemes: $(\mu + 1_{succ})$ and $(\mu +_t 1)$.

Perspectives

- Find "appropriate" C for **exploration term** $\bar{r}_{j,t} + C\sqrt{\frac{2\ln \sum_k n_{k,t}}{n_{j,t}}}$.
- Allocate some **budget** of evaluations for **dominated arms** (individuals) generated in the past to preserve the diversity.
- Integrate the reward mechanism and update rules of (1+1)-CMA-ES, e.g. **success rate** and average reward in $(\mu + 1_{rank})$.
- Experiments on Many-objective problems.

Thank you for your attention !

Please send your questions to
{Ilya.Loshchilov,Marc.Schoenauer,Michele.Sebag}@inria.fr