Multi-Objective Differential Evolution with Adaptive Control of Parameters and Operators

Ke Li[1] Álvaro Fialho[2] Sam Kwong[1]

[1]Department of Computer Science, City University of Hong Kong
Hong Kong, China

[2]Laboratoire d'Informatique (LIX), École Polytechnique
Palaiseau, France

LION 5 - January 19, 2011

Outline

1. Context
2. Differential Evolution to Multi-Objective
3. Adaptive Control of Parameters
4. Empirical Validation
5. Conclusions

Context

1. Context
2. Differential Evolution to Multi-Objective
3. Adaptive Control of Parameters
4. Empirical Validation
5. Conclusions

The Differential Evolution Algorithm

- Simple but efficient population-based EA
- No Selection: each individual used to generate an offspring
  - Mutation: weighted differences between several individuals
  - Crossover: mix parts of the mutated and the original individual
- Std. Replacement: if offspring better than parent, replace it

The Differential Evolution Algorithm

- Simple but efficient population-based EA
- No Selection: each individual used to generate an offspring
  - Mutation: weighted differences between several individuals
  - Crossover: mix parts of the mutated and the original individual
- Std. Replacement: if offspring better than parent, replace it
  - **Should be adapted to Multi-Objective**

The Differential Evolution Algorithm

- Simple but efficient population-based EA
- No Selection: each individual used to generate an offspring
  - Mutation: weighted differences between several individuals
  - Crossover: mix parts of the mutated and the original individual
- Std. Replacement: if offspring better than parent, replace it
  - **Should be adapted to Multi-Objective**

Bottleneck: Parameter Setting

- Population size NP
- Mutation scaling factor F
- Crossover rate CR
- Which mutation strategies to apply?
  - and at which Application Rate?

What we propose...

1. Extend DE to multi-objective
   - New fitness evaluation
   - New replacement mechanism
2. Automate its parameter setting ("Crossing the Chasm")
   - CR, F and mutation strategy selection (NP fixed)
   - Existing adaptive parameter control methods
     - Ported to the multi-objective domain

Differential Evolution to Multi-Objective

1 Context

2 Differential Evolution to Multi-Objective

3 Adaptive Control of Parameters

4 Empirical Validation

5 Conclusions

Fitness Evaluation

1. Convergence
   - Solutions as close to the Pareto optimal front as possible
   - Existing "Pareto Dominance Strength" measure
2. Spread
   - Solutions as distributed in the Pareto front as possible
   - Novel "Tree Neighborhood Density" measure

Fitness Evaluation I: Pareto Dominance Strength

⇒ From the Strength Pareto EA (SPEA2) [Zitzler et al., 2002]

- Strength $S(i) = \#$ solutions dominated by i

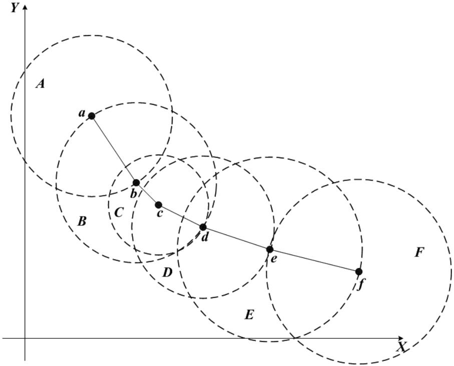
$$S(i) = ||\{j | j \in P \wedge i \succ j\}||$$

- Pareto Dominance $PD(i) =$ sum of $S(j)$, j dominates i

$$PD(i) = \sum_{j \in P, j \succ i} S(j)$$

- The smaller the better (convergence)
  - $PD(i) = 0 \rightarrow$ non-dominated solution

Fitness Evaluation II: Tree Neighborhood Density



- Connected graph $\Rightarrow$ Minimum Spanning Tree
  - number of edges = degree
  - length of edge = distance between points
  - $\Rightarrow$ neighborhood "crowdedness" of each individual

Aggregated Fitness Evaluation

$$f(i) = PD(i) + nTND(i)$$

- Only the TND measure is normalized
  1. minimizes PD (approaches the Pareto)
  2. then TND (higher diversity between non-dominated)

Replacement Mechanism

Starting from the mixed population of size $2 \times NP \ldots$

1. Pareto Dominance between parent and offspring
   - Non-dominated (if any) is maintained
2. Non-dominated sorting method (NSGA-II [Deb et al., 2002])
   - Iteratively selects non-dominated individuals for survival
3. Individuals with lowest TND (less "crowded") are maintained

Adaptive Control of Parameters

1. Context
2. Differential Evolution to Multi-Objective
3. Adaptive Control of Parameters
4. Empirical Validation
5. Conclusions

DE Mutation Strategies

- Around a dozen of well-known existent strategies
- As in other EAs, complex and problem-dependent choice

- Off-line tuning could be used to find the best one
- Based on some statistics over several runs for each strategy
- <u>Expensive</u>, providing the <u>static</u> single best strategy

DE Mutation Strategies

- Around a dozen of well-known existent strategies
- As in other EAs, complex and problem-dependent choice

- Off-line tuning could be used to find the best one
- Based on some statistics over several runs for each strategy
- Expensive, providing the static single best strategy

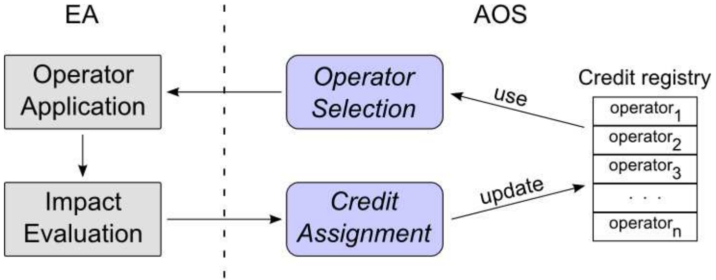
- Best strategy depends on the region of the search space
- Should be continuously adapted, while solving the problem

$\Rightarrow$ Adaptive Operator/Strategy Selection $\Leftarrow$

Adaptive Operator/Strategy Selection

Objective

Autonomously select the operator to be applied between the available ones, based on its impact on the search up to now.



Adaptive Operator/Strategy Selection

(inspired from [Gong et al., 2010])

Credit Assignment

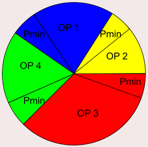
- Impact of strategy application: *Normalized Relative ΔFitness*

$$R = \frac{|f_{parent} - f_{offspring}|}{|f_{best} - f_{worst}|}$$

- Credit assigned (once every generation): *Average*

$$r_a(g) = \sum_{i=1}^{|R_a|} \frac{R_a(i)}{|R_a|}$$

Operator Selection: Probability Matching



$$\hat{Q}_{j,t+1} = (1-\alpha)\cdot\hat{Q}_{j,t} + \alpha \cdot r_{j,t}$$

$$s_{i,t+1} = p_{min} + (1 - K\cdot p_{min})\cdot\frac{\hat{Q}_{i,t+1}}{\sum_{j=1}^{K}\hat{Q}_{j,t+1}}$$

Adaptive Control of CR and F

- Crossover rate CR
- Mutation scaling factor F

⇒ Slightly modified from the JADE algorithm [Zhang and Sanderson, 2009]

- After each generation:
  - CR_i^a = crossover rate for individual i using operator a

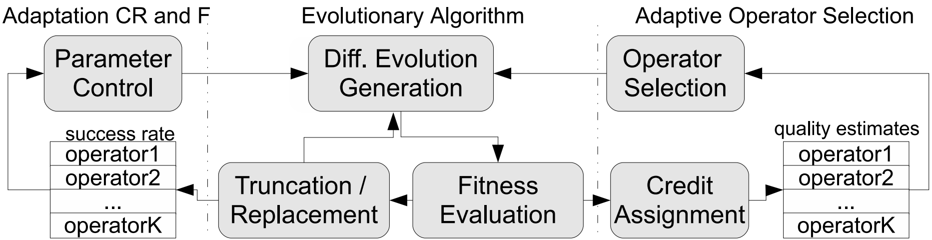
$$CR_i^a = norm(\mu_{CR}^a, 0.1)$$

  - S_{CR}^a = set of successful CR values at current generation

$$\mu_{CR}^a = (1 - c) \cdot \mu_{CR}^a + \mathbf{c} \cdot mean(S_{CR}^a)$$

- The same is done to adapt F

Combined Adaptive DE for Multi-Objective Optimization



Empirical Validation

1 Context
2 Differential Evolution to Multi-Objective
3 Adaptive Control of Parameters
4 Empirical Validation
5 Conclusions

Experimental Settings

Experimental Aims

- Efficiently controls CR/F and selects between
  1. rand/1/bin
  2. current-to-rand/1/bin
  3. rand/2/bin
  4. rand-to-best/2/bin
- $NP = 100$; other parameters as in the original references

Criteria

- Uniform Assessment (UA): diversity of final solutions
- Hyper-Volume (HV): comprehensive performance indicator

Comparative Results

		Str.1	Str.2	Str.3	Str.4	Adap-MODE	S
ZDT1	UA	7.905E-1±1.93E-2	7.999E-1±1.94E-2	7.436E-1±2.63E-2	4.078E-1±5.18E-2	**8.080E-1±1.62E-2**	†
	HV	3.66191±3.39E-5	3.66192±3.44E-5	3.65588±2.23E-3	1.90258±3.97E-1	**3.66193±3.15E-5**	
ZDT2	UA	7.880E-1±1.54E-2	8.046E-1±1.90E-2	7.197E-1±2.72E-2	3.612E-1±6.35E-2	**8.069E-1±1.89E-2**	
	HV	3.32858±3.46E-5	**3.32861±3.88E-5**	3.31965±4.58E-3	1.90504±2.46E-1	3.32853±4.19E-5	‡
ZDT3	UA	**7.689E-1±2.13E-2**	7.560E-1±2.68E-2	5.148E-1±8.01E-2	3.406E-1±2.69E-2	7.660E-1±1.98E-2	
	HV	**4.81519±6.38E-5**	4.81427±1.58E-3	4.77590±1.35E-2	1.78116±3.45E-1	4.81463±4.81E-4	‡
ZDT4	UA	**8.133E-1±1.69E-2**	8.026E-1±1.63E-2	8.073E-1±1.97E-2	3.340E-1±3.94E-2	8.055E-1±1.85E-2	
	HV	3.63669±1.06E-1	3.66195±3.81E-5	3.64949±8.60E-2	0.00000±0.00E+0	**3.66201±5.33E-4**	
ZDT6	UA	7.994E-1±1.94E-2	8.172E-1±2.00E-2	**8.198E-1±1.94E-2**	7.641E-1±4.63E-2	7.896E-1±2.27E-2	‡
	HV	**3.04183±1.77E-5**	**3.04183±2.42E-5**	**3.04183±1.55E-5**	3.04131±3.16E-3	**3.04183±1.62E-5**	
DTLZ1	UA	**8.280E-1±2.09E-2**	8.241E-1±1.68E-2	8.254E-1±1.52E-2	4.708E-1±4.28E-2	8.246E-1±1.48E-2	
	HV	9.70350±1.04E-3	9.70342±5.64E-4	9.69684±7.15E-4	0.00000±0.00E+0	**9.73582±2.75E-4**	†
DTLZ2	UA	8.171E-1±1.79E-2	8.011E-1±2.00E-2	8.003E-1±1.99E-2	7.897E-1±2.54E-2	**8.236E-1±1.84E-2**	
	HV	7.34877±1.41E-2	7.33683±1.40E-2	7.33509±7.86E-3	7.30348±8.85E-3	**7.40523±1.14E-2**	†
DTLZ3	UA	8.071E-1±1.83E-2	3.524E-1±3.93E-2	3.454E-1±3.27E-2	4.010E-1±3.97E-2	**8.304E-1±1.72E-2**	†
	HV	6.538E+0±2.06E+0	0.00000±0.00E+0	0.00000±0.00E+0	0.00000±0.00E+0	**7.32465±5.76E-1**	†
DTLZ4	UA	2.541E-1±3.81E-2	2.491E-1±3.29E-2	2.570E-1±3.05E-2	2.363E-1±3.05E-2	**2.654E-1±2.99E-2**	†
	HV	5.58000±1.10E+0	6.63999±4.80E-1	6.35973±7.69E-1	5.97162±1.13E+0	**7.02943±5.46E-1**	†
DTLZ5	UA	7.388E-1±2.19E-2	7.200E-1±2.36E-2	7.219E-1±2.16E-2	7.336E-1±2.79E-2	**7.866E-1±1.82E-2**	
	HV	6.07366±3.43E-3	6.06755±3.47E-3	6.06536±3.91E-3	6.05228±5.33E-3	**6.10548±4.40E-3**	†
DTLZ6	UA	7.900E-1±2.00E-2	7.914E-1±2.09E-2	**7.970E-1±1.74E-2**	7.600E-1±2.57E-2	7.759E-1±2.18E-2	‡
	HV	6.10701±4.43E-3	6.10657±4.26E-3	**6.10802±5.39E-3**	5.76474±1.03E+0	6.10732±4.88E-3	
DTLZ7	UA	7.642E-1±1.81E-2	7.692E-1±1.67E-2	7.434E-1±2.19E-2	5.176E-1±1.35E-1	**7.723E-1±1.86E-2**	
	HV	13.41285±5.65E-2	13.42783±4.80E-2	13.34603±7.35E-2	7.73571±3.70E+0	**13.4648±7.43E-2**	†

Comparative Results

		CR1/F.5 & AOS	S	CR/Fcontrol & Unif.OS	S	Adap-MODE
ZDT1	UA	7.860E-1±2.08E-2	†	7.851E-1±2.42E-2	†	**8.080E-1±1.62E-2**
	HV	3.66162±2.97E-4	†	3.66066±2.69E-4	†	**3.66193±3.15E-5**
ZDT2	UA	7.809E-1±2.02E-2	†	7.793E-1±1.71E-2	†	**8.069E-1±1.89E-2**
	HV	3.32840±3.13E-4	†	3.32612±5.27E-4	†	**3.32853±4.19E-5**
ZDT3	UA	7.538E-1±2.83E-2	†	7.487E-1±1.52E-2	†	**7.660E-1±1.98E-2**
	HV	4.81448±1.18E-3		4.81228±1.18E-3	†	**4.81463±4.81E-4**
ZDT4	UA	**8.127E-1±2.30E-2**		7.486E-1±6.12E-2		8.055E-1±1.85E-2
	HV	3.64150±1.43E-1	†	3.65409±4.26E-2	†	**3.66201±5.33E-4**
ZDT6	UA	7.626E-1±2.34E-2	†	**8.078E-1±2.34E-2**	‡	7.896E-1±2.27E-2
	HV	3.04179±3.22E-5	†	**3.04183±4.93E-5**		**3.04183±1.62E-5**
DTLZ1	UA	**8.247E-1±1.80E-2**		8.200E-1±1.73E-2		8.246E-1±1.48E-2
	HV	9.69925±5.41E-4	†	9.17842±1.25E-1	†	**9.73582±2.75E-4**
DTLZ2	UA	8.096E-1±2.01E-2	†	8.224E-1±1.56E-2		**8.236E-1±1.84E-2**
	HV	7.33762±1.10E-2	†	7.40368±9.20E-3		**7.40523±1.14E-2**
DTLZ3	UA	6.365E-1±1.44E-1	†	8.289E-1±1.42E-2		**8.304E-1±1.72E-2**
	HV	7.13704±3.70E-1	†	4.59535±2.92E+0	†	**7.32465±5.76E-1**
DTLZ4	UA	2.092E-1±3.32E-2	†			**2.654E-1±2.99E-2**
	HV	6.78321±6.03E-1	†	4.66216±1.09E+0	†	**7.02943±5.46E-1**
DTLZ5	UA	7.334E-1±2.26E-2	†	**7.792E-1±1.95E-2**		7.866E-1±1.82E-2
	HV	6.07005±3.69E-3	†	**6.10649±3.78E-3**		6.10548±4.40E-3
DTLZ6	UA	7.739E-1±2.32E-2		**7.876E-1±2.00E-2**	‡	7.759E-1±2.18E-2
	HV	**6.10841±5.67E-3**		6.10640±4.14E-3		6.10732±4.88E-3
DTLZ7	UA	7.621E-1±1.83E-2	†	7.634E-1±1.70E-2	†	**7.723E-1±1.86E-2**
	HV	13.42436±6.19E-2	†	13.43145±7.25E-2	†	**13.46486±7.43E-2**

Comparative Results

		NSGA-II	GDE3	Adap-MODE	S
ZDT1	UA	0.45141±4.40e-2	**0.51370±4.49e-2**	0.48543±2.34e-2	‡
	HV	3.66029±2.72e-4	3.66116±9.23e-4	**3.66154 ± 3.11e-4**	
ZDT2	UA	0.44849±4.53e-2	0.50960±4.68e-2	**0.52161±4.52e-2**	†
	HV	3.32678±2.12e-3	3.31997±1.33e-3	**3.32069 ± 4.43e-3**	
ZDT3	UA	0.43172±4.17e-2	**0.51792±4.25e-2**	0.40185±2.41e-2	‡
	HV	4.81509±8.04e-5	4.81459±1.45e-4	**4.81521 ± 2.41e-4**	
ZDT4	UA	0.43397±4.83e-2	0.43868±1.23e-1	**0.49862±2.11e-1**	†
	HV	3.65876±1.76e-3	3.23512±3.27e-1	**3.66082±5.21e-3**	†
ZDT6	UA	0.30441±5.06e-2	0.16473±4.20e-2	**0.45004±3.16e-2**	†
	HV	2.99080±7.78e-3	**3.03849±3.26e-4**	3.03021 ± 3.78e-4	
DTLZ1	UA	0.37779±3.58e-2	0.43224±4.48e-2	**0.44561±4.13e-2**	†
	HV	**0.94094±9.31e-3**	0.36739±3.29e-1	0.87849±5.49e-2	‡
DTLZ2	UA	0.37248±4.33e-2	0.38791±4.78e-2	**0.42588±3.44e-2**	†
	HV	7.16923±6.77e-2	7.29626±5.42e-2	**7.30181±2.94e-2**	†
DTLZ3	UA	0.36240±4.04e-2	0.45105±4.29e-2	**0.45749±3.58e-2**	
	HV	7.13655±7.92e-2	7.09485±5.93e-2	**7.15978±2.91e-2**	†
DTLZ4	UA	**0.51266±5.83e-2**	0.24955±3.32e-2	0.49989±5.39e-2	
	HV	**6.55645±2.28e-1**	5.66181±3.84e-1	6.55593 ± 2.03e-1	
DTLZ5	UA	**0.40438±4.41e-2**	0.27342±7.53e-2	0.38946±6.55e-2	
	HV	6.10036±1.87e-3	**6.10215±2.59e-3**	6.10129 ± 5.48-3	
DTLZ6	UA	0.33096±4.58e-2	0.15052±3.46e-2	**0.33151±4.21-2**	
	HV	5.92389±6.88e-2	**6.10113±6.34e-3**	6.01598±8.52e-2	‡
DTLZ7	UA	**0.41141±4.16e-2**	0.29847±5.70e-2	0.36802±1.55e-2	‡
	HV	12.82917±6.82e-1	12.79845±5.45e-1	**13.23157±3.45E-1**	†

Conclusions

1. Context
2. Differential Evolution to Multi-Objective
3. Adaptive Control of Parameters
4. Empirical Validation
5. Conclusions

Conclusions - Summary

1. Extended DE to multi-objective
   - New fitness evaluation
     - Convergence: Pareto dominance strength from [Zitzler et al., 2002]
     - Spread: novel Tree Neighborhood Density estimator
   - New replacement mechanism
     - Three-step comparison procedure
2. Automated setting of most parameters (except NP)
   - Ported existing adaptive control methods to MOO
     - Adaptive Operator Selection from [Gong et al., 2010]
     - Adaptive Param. Control of CR/F from [Zhang and Sanderson, 2009]

- Achieves better convergence and diversity in most problems
  - Compared to MODE with a single strategy;
  - Each of the adaptive modules alone;
  - And to other well-known algorithms for MOO

Conclusions - Perspectives for Further Work

- Mixture of methods showed to be efficient
  - But where are the benefits mostly coming from?
- Analyze sensitivity and robustness of parameters
- Try more sophisticated schemes for adaptive control
  - e.g., AOS based on multi-armed bandits
  - and other alternatives for controlling CR and F
- Current scheme is very expensive
  - Specially the calculation of the Tree Neighborhood Density

References I

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002).
A fast and elitist multiobjective genetic algorithm: NSGA-II.
IEEE Trans. Evol. Comput., 6:182–197.

Gong, W., Fialho, A., and Cai, Z. (2010).
Adaptive strategy selection in differential evolution.
In J. Branke et al., editor, *Proc. GECCO*. ACM.

Zhang, J. and Sanderson, A. C. (2009).
JADE: Adaptive differential evolution with optional external archive.
IEEE Trans. Evol. Comput., 13:945–958.

Zitzler, E., Laumanns, M., and Thiele, L. (2002).
SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization.
In K.C. Giannakoglou et al., editor, *Evolutionary Methods for Design, Optimization and Control with Application to Industrial Problems*, pages 95–100. CIMNE.

Multi-Objective Differential Evolution with Adaptive Control of Parameters and Operators

Ke Li[1] Álvaro Fialho[2] Sam Kwong[1]

[1]Department of Computer Science, City University of Hong Kong
Hong Kong, China

[2]Laboratoire d'Informatique (LIX), École Polytechnique
Palaiseau, France

LION 5 - January 19, 2011