Differential Evolution with Rank-based Adaptive Strategy Selection

Álvaro Fialho, Marc Schoenauer, Michèle Sebag

MICROSOFT RESEARCH INRIA JOINT CENTRE

Orsay, France

Differential Evolution

- Population-based EA
- Each individual is used to generate a new offspring
- Mutation: weighted differences between several individuals
- Crossover: mix parts of the mutated and the original individual
- Mutation always applied, Crossover applied with rate (1-CR)
- Replacement: 1x1, if offspring better than parent, replace it

User-defined parameters

- Population size NP
- Mutation scaling factor F
- Crossover rate CR
- **Which mutation strategies to apply?**

DE Mutation Strategies

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

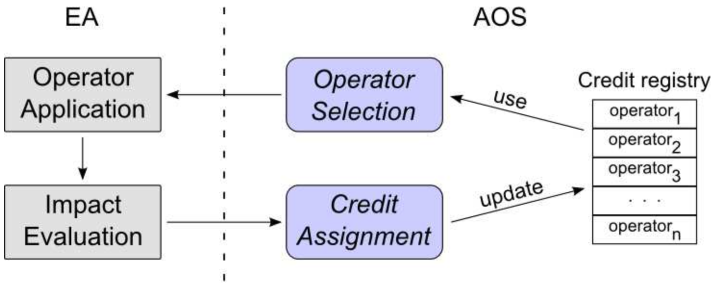
- Off-line tuning could be used to find the best one
- Based on some statistics over several runs for each strategy
- Expensive, providing the static single best strategy

- Best strategy depends on the region of the search space
- Should be continuously adapted, while solving the problem
- $\Longrightarrow$ Adaptive Strategy Selection

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.

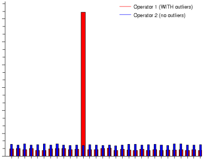


How to Measure the Impact of an Operator Application?

- Very common: Fitness Improvement

Which statistics to use?

- Instantaneous value likely to be unstable
- Average value over a Window
- Extreme value over a Window [Fialho et al., 2008]



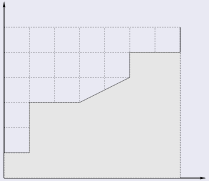
But...

- Ranges of rewards depend on the problem
- Some normalization methods were proposed [Fialho et al., 2009, Gong et al., 2010]
- Problem-dependent while fitness values are considered...
- Consequently, AOS becomes also problem-dependent

Rank-based Rewarding

Area Under ROC Curve (AUC)

- ML: comparison between 2 binary classifiers
- In AOS, 1 operator versus others [Fialho et al., 2010]
- Position r is assigned rank-value $D^r(W-r)$
- Parameter D is the decay factor, fixed at .5
- Size of the segment = assigned rank-value
- Example without decay factor:
  $(+ - + + - - [- - +] + - - +)$



Comparison-based Rewarding

- Ranks of the fitness values, rather than fitness improvements
- Invariant with respect to monotonous transformations

Op. Selection: A (kind of) Multi-Armed Bandit problem

Multi-Armed Bandits

Several "arms"; at time t, *gambler* plays arm j

$$\textit{reward at } t \;:\; r_{j,t} = \begin{cases} 1 & \text{with some prob,} \\ 0 & \text{otherwise.} \end{cases}$$

Goal: maximize cumulated reward

Upper Confidence Bound (UCB)

- Assimptotic optimality guarantees [Auer et al., 2002]
- At time t, choose arm j maximizing:

$$\hat{p}_{j,t} + \mathcal{C}\sqrt{\frac{2\log\sum_k n_{k,t}}{n_{j,t}}}, \text{ where } \begin{cases} \hat{p}_{j,t} & \text{,empirical estimate arm } j \\ n_{j,t} & \text{,chosen times arm } j \end{cases}$$

Op. Selection with Multi-Armed Bandits: the true story

From the original UCB

- $\hat{q}_{j,t} = \hat{p}_{j,t} + C\sqrt{\frac{2\log\sum_k n_{k,t}}{n_{j,t}}}$ score
- $\hat{p}_{j,t+1} = (n_{j,t} * \hat{p}_{j,t} + r_{j,t})/(n_{j,t} + 1)$ empirical estimate
- $n_{j,t+1} = n_{j,t} + 1$ # times used

Dynamics

- MAB: reward distribution assumed to be stationary
- AOS: depends on the current region of the search space
- UCB would take too long to adapt to a new best operator
- How to deal with such dynamics?

AUC - Multi-Armed Bandit

AUC-MAB

- Rank-based rewarding, no problem-dependency
- Comparison-based if ranks over fitness values
- Original MAB, $\hat{p}$ is the avg of all received rewards
  - Takes too long to adapt to changes
- AUC is already a continuously up-to-date aggregation
- $\Longrightarrow$ directly use AUC in bandit equation

$$\hat{q}_{j,t} = AUC_{j,t} + \mathcal{C}\sqrt{\frac{2\log\sum_k n_{k,t}}{n_{j,t}}}$$

- AUC incorporates the behavior of all operators (dynamics)

Experimental Settings

Differential Evolution

- Population size $NP = 10 \times DIM$
- Mutation scaling factor $F = 0.5$
- Crossover rate $CR = 1.0$, i.e., no crossover
- No tuning of these parameters, focus on AOS

Mutation Strategies

1. rand/1: $\mathbf{v}_i = \mathbf{x}_{r_1} + F \cdot (\mathbf{x}_{r_2} - \mathbf{x}_{r_3})$
2. rand/2: $\mathbf{v}_i = \mathbf{x}_{r_1} + F \cdot (\mathbf{x}_{r_2} - \mathbf{x}_{r_3}) + F \cdot (\mathbf{x}_{r_4} - \mathbf{x}_{r_5})$
3. rand-to-best/2:
   $\mathbf{v}_i = \mathbf{x}_{r_1} + F \cdot (\mathbf{x}_{best} - \mathbf{x}_{r_1}) + F \cdot (\mathbf{x}_{r_2} - \mathbf{x}_{r_3}) + F \cdot (\mathbf{x}_{r_4} - \mathbf{x}_{r_5})$
4. current-to-rand/1: $\mathbf{v}_i = \mathbf{x}_i + F \cdot (\mathbf{x}_{r_1} - \mathbf{x}_i) + F \cdot (\mathbf{x}_{r_2} - \mathbf{x}_{r_3})$

Baseline Methods

Dynamic Multi-Armed Bandit (DMAB) [Da Costa et al., 2008]

- Original UCB MAB algorithm
- Dynamics: Page-Hinkley change-detection test (threshold γ)

Adaptive Pursuit (AP) [Thierens, 2005]

- Winner-take-all strategy to update the operators rates
- Best operator with rate p_{max}, others p_{min}

- DMAB and AP rewarded by Extreme values

Probability Matching (PM-AdapSS-DE [Gong et al., 2010])

- Operators rates are proportional to their qualities
- Average of normalized fitness improvements

Meta-Parameters

Summary

- Strategy selection:

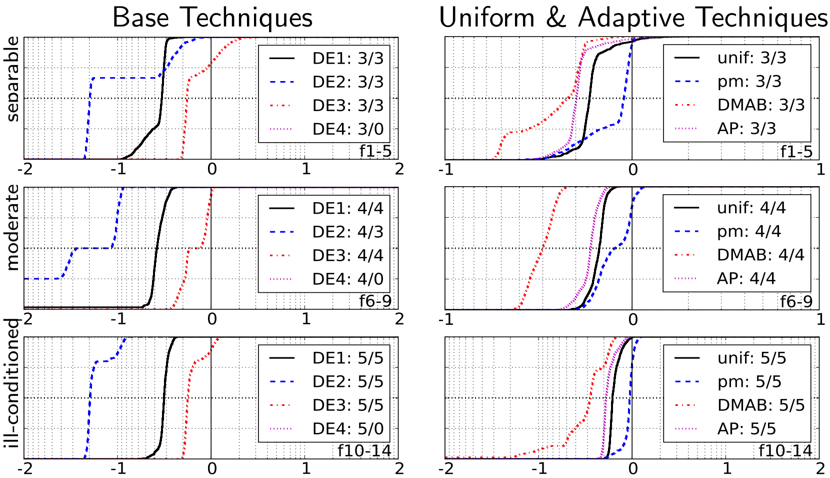
PM	: minimal probability p_{min}, learning rate α
AP	: minimal probability p_{min}, learning rate α, adaptation rate β
MAB	: scaling factor $\mathcal{C}$
DMAB	: scaling factor $\mathcal{C}$, PH threshold γ
AUC-B	: scaling factor $\mathcal{C}$, decay factor $\mathcal{D}$ (fixed at .5)

- Credit Assignment:
  sliding window size $\mathcal{W}$ and type (Avg, Extreme, AUC)

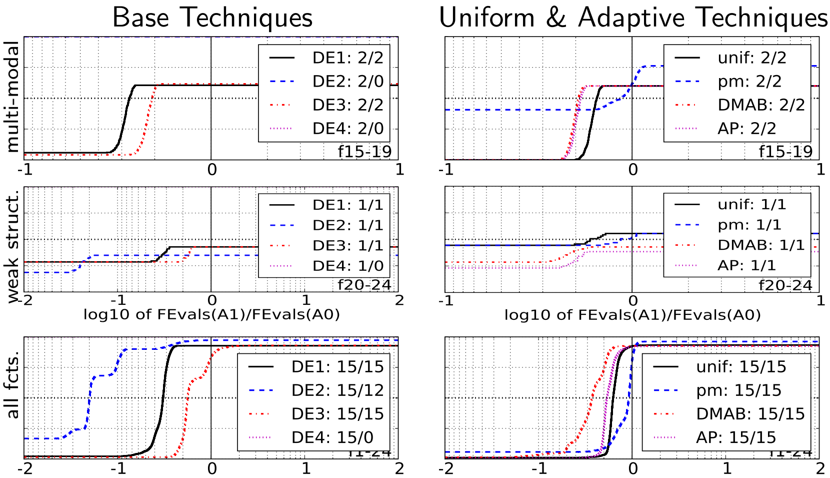
Tuning

- Tuned off-line by F-RACE [Birattari et al., 2002]
  - 1 racing lap = 1 run over all functions
  - Elimination after each lap, based on Friedman test at 95%
- Comparisons of AOS algorithms using best configurations

Parwise comparisons of AUC-Bandit with . . .



Parwise comparisons of AUC-Bandit with ...



Summary

MAB — Multi-Armed Bandit

- provides guarantees for optimal EvE in a static setting

DMAB — MAB + Page-Hinkley change-detection test

- Very strong ... if PH parameter γ is well-tuned
- C and γ extremely problem-dependent (reward = fitness improvement).

AUC-MAB — MAB with $\hat{p}$ =*Area Under the Curve*

- AUC rank-based: much more robust w.r.t. C
- One other parameter: window size W (decay factor $D \equiv 0.5$)
- $C.5D.5W50$ best configuration on very different situations
- AUC-MAB comparison-based with F instead of ΔF

Discussion and Perspectives

Discussion

- Fixed number of hyper-parameters, while user-defined parameters grow w.r.t. number of used operators
  - Operator type, application rate, and underlying parameters.
- In real problems, optimal behavior is not known
  - X-MAB better than fixed and known adaptive approaches
- Assess the AOS techniques, rather to "compete" on BBOB
  - Better tuning and enhanced versions of DE could be used

Further Work

- Further assessment
  - Use within other meta/hyper-heuristics (GA, DE, ??)
  - SAT, real-world problems, ...
- Real-world problems are often multi-modal
  - Diversity should also be considered for the reward

References I

Auer, P., Cesa-Bianchi, N., and Fischer, P. (2002).
Finite-time analysis of the multiarmed bandit problem.
Machine Learning, 47(2/3):235–256.

Birattari, M., Stützle, T., Paquete, L., and Varrentrapp, K. (2002).
A racing algorithm for configuring metaheuristics.
In *Proc. GECCO'02*.

Da Costa, L., Fialho, A., Schoenauer, M., and Sebag, M. (2008).
Adaptive operator selection with dynamic multi-armed bandits.
In M. Keijzer et al., editor, *Proc. GECCO'08*, pages 913–920. ACM Press.

Fialho, A., Da Costa, L., Schoenauer, M., and Sebag, M. (2008).
Extreme value based adaptive operator selection.
In *Proc. PPSN'08*.

Fialho, A., Schoenauer, M., and Sebag, M. (2009).
Analysis of adaptive operator selection techniques on the royal road and long k-path problems.
In G. Raidl et al., editor, *Proc. Genetic and Evolutionary Computation Conference*, pages 779–786. ACM.

Fialho, A., Schoenauer, M., and Sebag, M. (2010).
Toward comparison-based adaptive operator selection.
In J. Branke et al., editor, *Proc. GECCO*. ACM Press.

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

References II

Maturana, J., Fialho, A., Saubion, F., Schoenauer, M., and Sebag, M. (2009).
Extreme compass and dynamic multi-armed bandits for adaptive operator selection.
In *Proc. CEC'09*.

Thierens, D. (2005).
An adaptive pursuit strategy for allocating operator probabilities.
In Beyer, H.-G., editor, *Proc. GECCO'05*, pages 1539–1546. ACM Press.

Differential Evolution with Rank-based Adaptive Strategy Selection

Álvaro Fialho, Marc Schoenauer, Michèle Sebag

MICROSOFT RESEARCH INRIA JOINT CENTRE

Orsay, France