

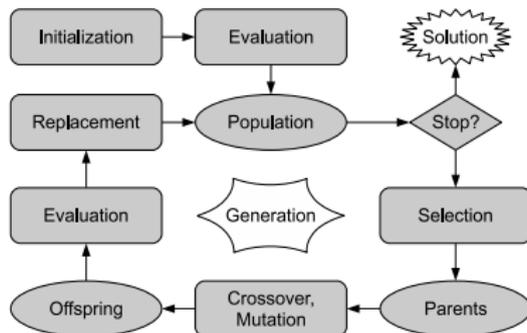
Toward Comparison-based Adaptive Operator Selection

Álvaro Fialho, Marc Schoenauer, Michèle Sebag



Orsay, France

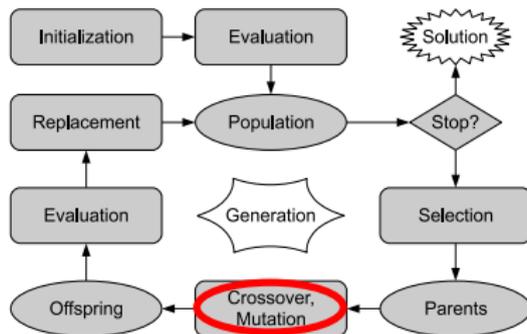
Evolutionary Algorithms – cycle and parameters



Performance very sensitive to the **parameter setting**

- Population size and number of Offspring generated
- Darwinian methods (selection and replacement) deterministic, tournament, roulette-wheel, ...
- **Variation Operators** and their parameters
 - which of them should be considered
 - and their relative probabilities of application

Evolutionary Algorithms – cycle and parameters



Performance very sensitive to the **parameter setting**

- Population size and number of Offspring generated
- Darwinian methods (selection and replacement) deterministic, tournament, roulette-wheel, ...
- **Variation Operators** and their parameters
 - which of them should be considered
 - and their relative probabilities of application

EA Variation Operators

- As in other EAs, complex and problem-dependent choice
- Several well-known existent operators
 - mutation (1 parent \rightarrow 1 offspring)
 - crossover (2+ parents \rightarrow 1 offspring)
 - problem-specific and local search operators (e.g., for TSP)
- Difficult to preview *a priori* the results of such settings

EA Variation Operators

- As in other EAs, complex and problem-dependent choice
 - Several well-known existent operators
 - mutation (1 parent \rightarrow 1 offspring)
 - crossover (2+ parents \rightarrow 1 offspring)
 - problem-specific and local search operators (e.g., for TSP)
 - Difficult to preview *a priori* the results of such settings
-
- Off-line tuning could be used to find the best one
 - Based on some statistics over several runs for each operator
 - Expensive, providing the static single best operator

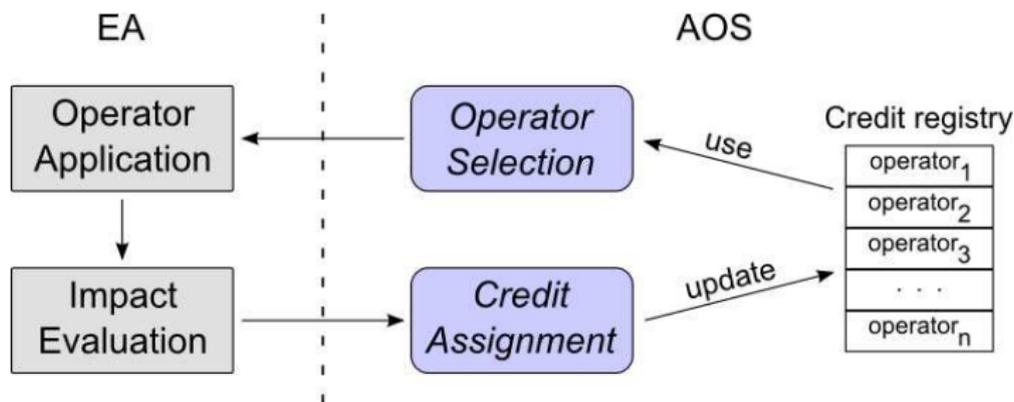
EA Variation Operators

- As in other EAs, complex and problem-dependent choice
 - Several well-known existent operators
 - mutation (1 parent \rightarrow 1 offspring)
 - crossover (2+ parents \rightarrow 1 offspring)
 - problem-specific and local search operators (e.g., for TSP)
 - Difficult to preview *a priori* the results of such settings
-
- Off-line tuning could be used to find the best one
 - Based on some statistics over several runs for each operator
 - Expensive, providing the static single best operator
-
- Best operator depends on the region of the search space
 - Should be continuously adapted, while solving the problem
 - \implies Adaptive Operator Selection

Adaptive Operator Selection

Objective

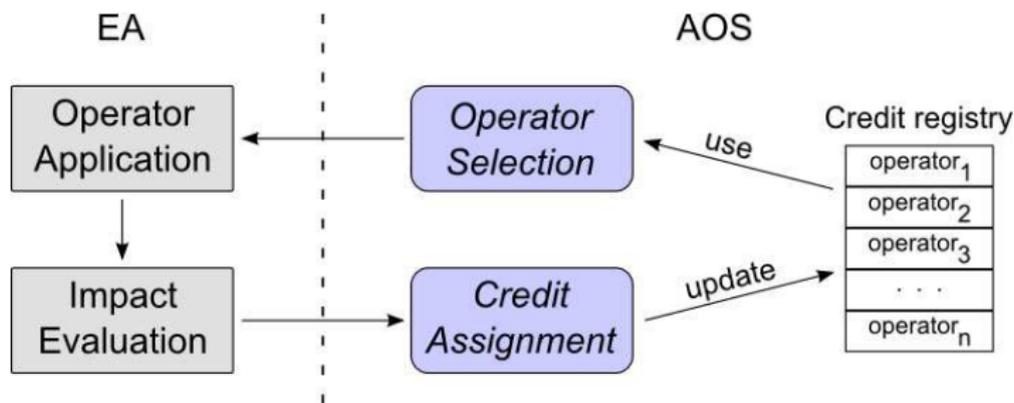
Autonomously select the operator to be applied between the available ones, based on its impact on the search up to now.



Adaptive Operator Selection

Objective

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



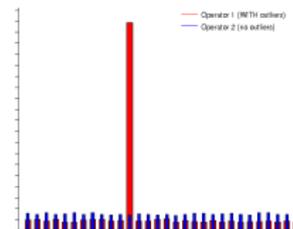
- Extreme - Dynamic Multi-Armed Bandit
- **AUC-based Multi-Armed Bandit**

Ex-DMAB Credit Assignment

- Very common: Fitness Improvement

Previous work

- Instantaneous value likely to be unstable
- Average value over a Window



Extreme Value Base (EVB)

- Outlier operators are rarely considered - smaller expectation.
- EC: Focus on extreme, rather than average events
 - Complex systems, e.g. epidemic propagation, financial markets
- Extreme Value-Based (EVB) Credit Assignment.
 - $\mathcal{R} = \text{Extreme value over a Window}$ [Fialho et al., 2008].

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

Original Multi-Armed Bandits (ML)



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

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

Goal: maximize cumulated reward

Upper Confidence Bound (UCB)

- Asymptotic optimality guarantees (static context) [Auer et al., 2002]
- At time t , choose arm j maximizing:

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

Dynamic Multi-Armed Bandit (DMAB)

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
- $n_{j,t+1} = n_{j,t} + 1$ # times used
- $\hat{p}_{j,t+1} = \hat{p}_{j,t} + \frac{1}{n_{j,t+1}} * (r_{j,t} - \hat{p}_{j,t})$ empirical estimate

DMAB = UCB + Scaling + Page-Hinkley change-detection

- Good operator is identified, no need of further exploration while it remains sufficiently good
- When it is not “good” anymore ...
 - Weight of the instant reward is inversely proportional to n_j
 - Adjusting the \hat{p}_j of an often used operator takes a long time.

⇒ Upon the detection of a change (threshold γ), **restart** the MAB.

Extreme - Dynamic MAB: Discussion

- Being fed by fitness improvements, DMAB was validated on:
 - Artificial scenarios [Da Costa et al., 2008],
 - EA benchmark problems [Fialho et al., 2009], and
 - some SAT instances [Maturana et al., 2009].
- Better results than the baseline methods
 - the common Probability Matching [Goldberg, 1990],
 - state-of-the-art Adaptive Pursuit [Thierens, 2007],
 - the naive uniform choice.
- Statistically equivalent to optimal behavior (when available)

But ...

- The PH test is triggered just in case of an abrupt change
- While in (EA) reality, changes tend to happen more gradually
- C and γ very sensitive, as **reward is problem-dependent**

Rank-based Rewarding: AUC

Area Under ROC Curve (AUC)

- ML: comparison between 2 binary classifiers
- In AOS, 1 operator versus others [Fialho et al., 2010]
- Credits received by the operators are ranked.
- Position r is assigned rank-value $D^r(W - r)$
- Size of the segment = assigned rank-value



AUC-MAB

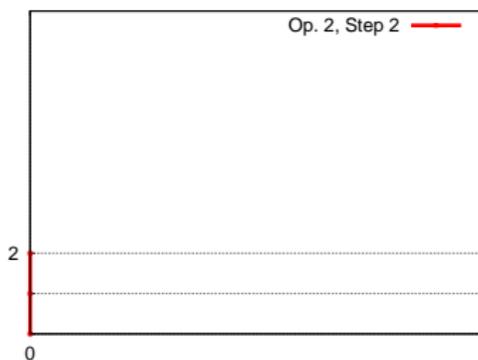
- Uses the ranks of the fitness improvements (ΔF)

F-AUC-MAB: Comparison-based Rewarding

- Ranks of the fitness values (F), rather than ΔF
- **Invariant** with respect to monotonous transformations

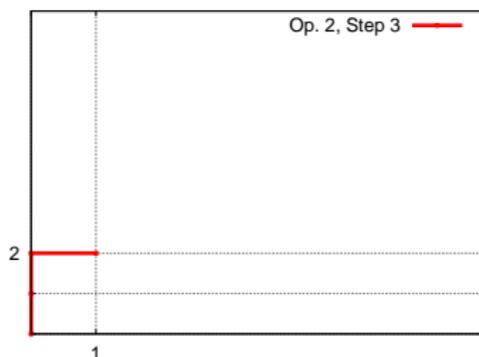
Rank-based Rewarding: Area Under ROC Curve

| # | O |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



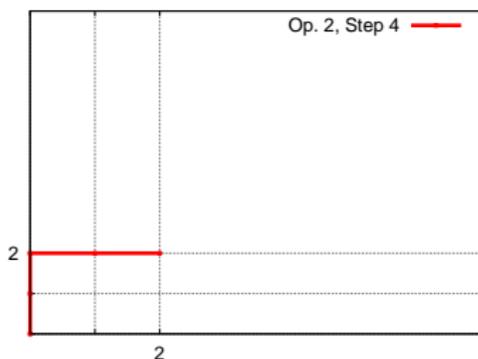
Rank-based Rewarding: Area Under ROC Curve

| # | O |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



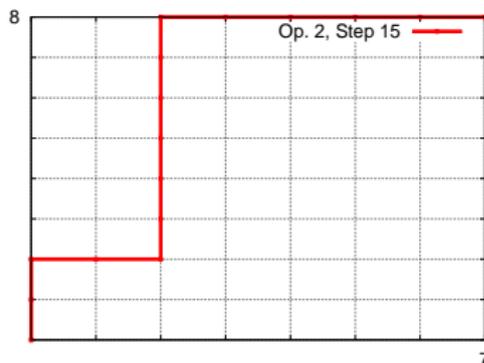
Rank-based Rewarding: Area Under ROC Curve

| # | O |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



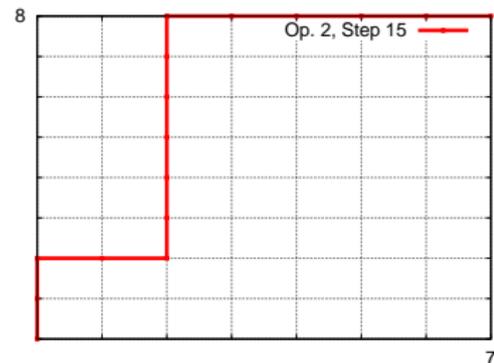
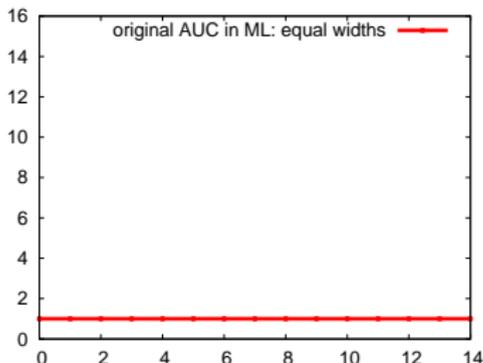
Rank-based Rewarding: Area Under ROC Curve

| # | O |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



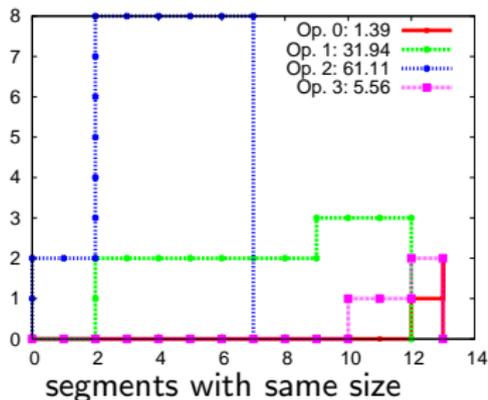
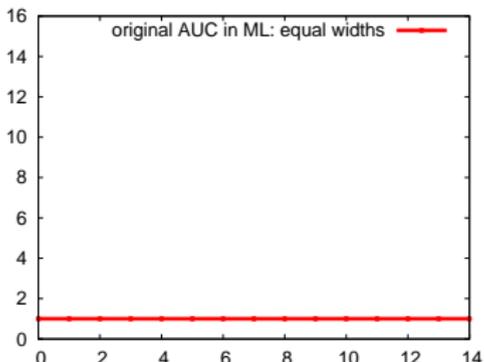
Rank-based Rewarding: Area Under ROC Curve

| # | 0 |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



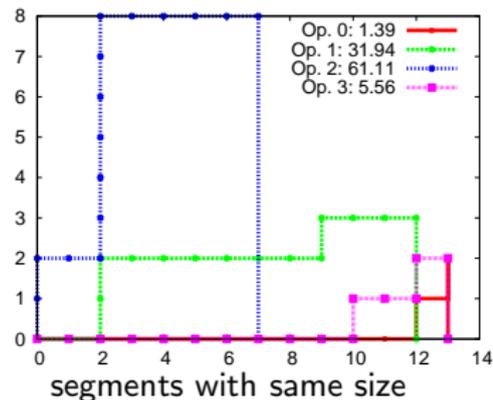
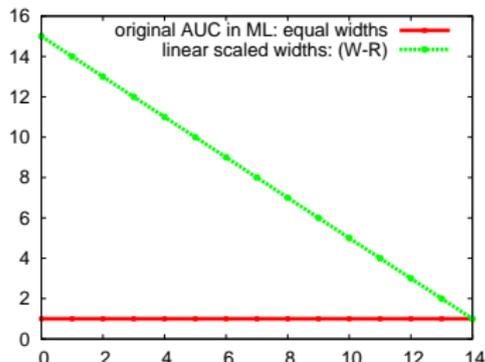
Rank-based Rewarding: Area Under ROC Curve

| # | O |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



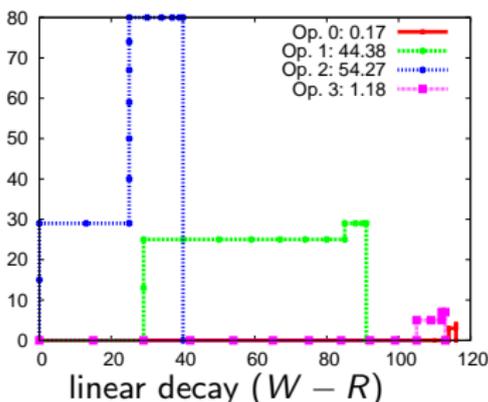
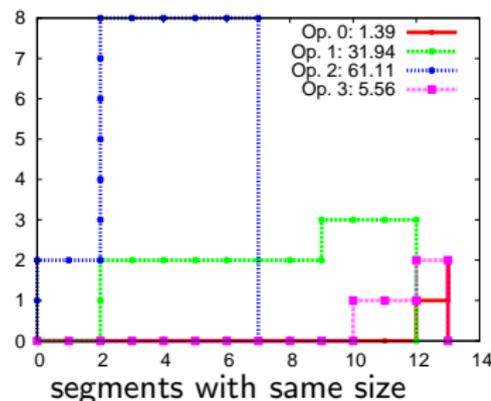
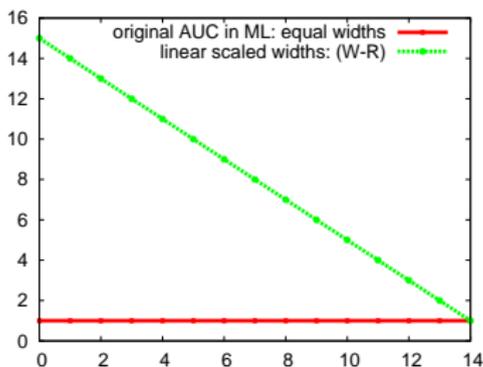
Rank-based Rewarding: Area Under ROC Curve

| # | O |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



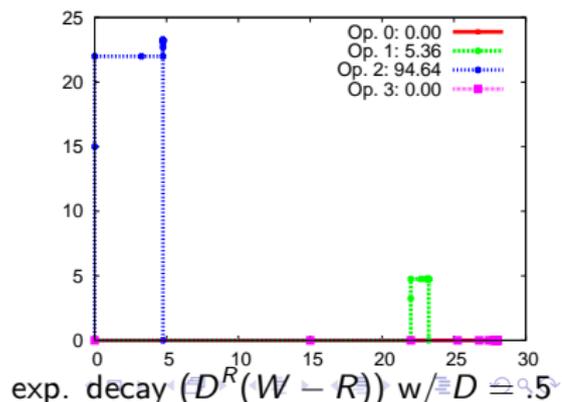
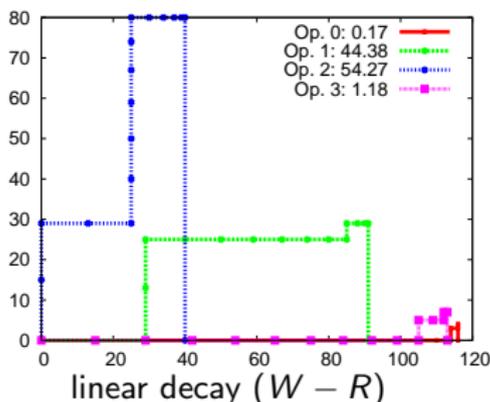
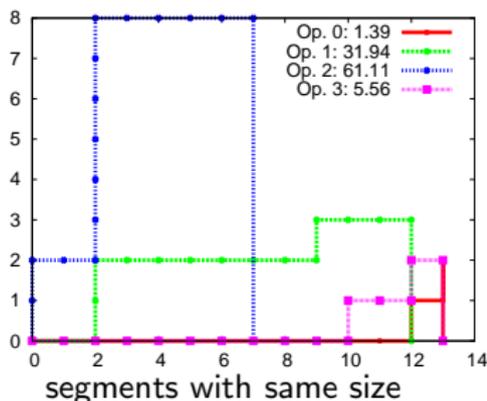
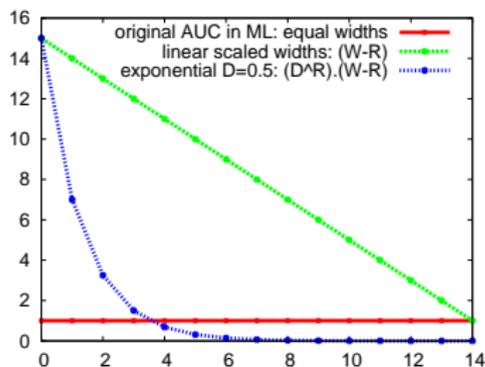
Rank-based Rewarding: Area Under ROC Curve

| # | O |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



Rank-based Rewarding: Area Under ROC Curve

| # | 0 |
|----|---|
| 1 | 2 |
| 2 | 2 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |
| 6 | 2 |
| 7 | 2 |
| 8 | 2 |
| 9 | 2 |
| 10 | 2 |
| 11 | 3 |
| 12 | 1 |
| 13 | 0 |
| 14 | 3 |
| 15 | 0 |



AUC - Multi-Armed Bandit

- For all MAB techniques, C (and γ for DMAB) is very sensitive
 - The reward is problem dependent
 - Motivation to propose the rank-based rewards
- AUC as reward for MAB/DMAB is ineffective
 - In MAB/DMAB, \hat{p} is the average of all received rewards
 - AUC is already an aggregation
 - AUC incorporates the behavior of all operators
- → directly use AUC in bandit equation

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

Meta-Parameters

Summary

- Operator selection:
 - MAB : scaling factor \mathcal{C}
 - DMAB : scaling factor \mathcal{C} , PH threshold γ
 - AUC-B : scaling factor \mathcal{C} , decay factor \mathcal{D} ($\equiv 0.5$)
- Credit Assignment:
 - sliding window size \mathcal{W} and **type** (Avg, Extreme, AUC)

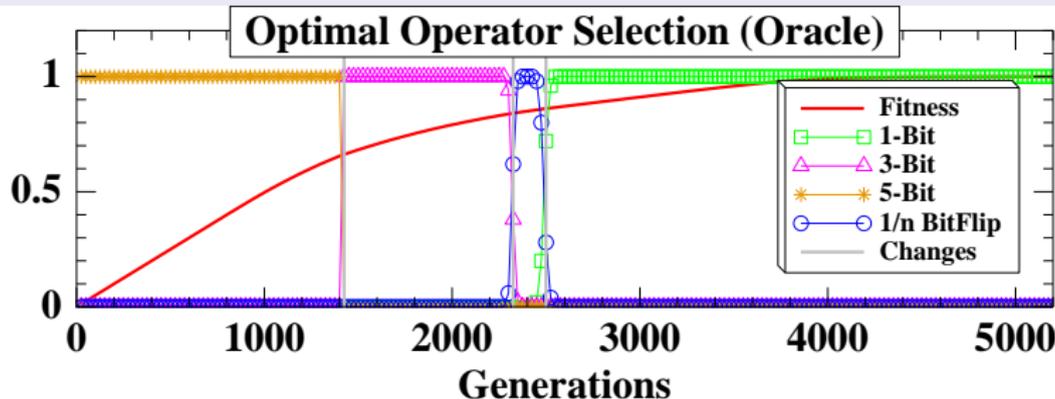
Tuning

- Tuned off-line by F-RACE [\[Birattari et al., 2002\]](#)
 - 11 initial runs, up to 50 runs or 1 configuration left
 - Friedman's test at 95% confidence level
- Comparisons of techniques at their best

The One-Max Problem

- (1+50)-GA
- 10^4 bits, fitness is the number of “1”s in the bitstring
- Very simple unimodal problem, the “drosophila of EC”

The One-Max “Oracle”



Comparative Results on the One-Max

| AOS | #Gens. |
|--|----------------------------------|
| Naive ($\{1,3,5\}$ -bit and $1/n$ bit-flip) | 7813 \pm 708 |
| Best Naive ($\{1,5\}$ -bit) | 6793 \pm 625 |
| Optimal | 5059 \pm 203 |
| AUC - MAB | 5069 \pm 292 |
| Ext - DMAB | 5254 \pm 313 |
| Avg - DMAB | 7494 \pm 611 |
| Ext - MAB | 9044 \pm 840 |
| Avg - MAB | 8198 \pm 683 |

Analysis of Invariance w.r.t. Monotonous Transformations

- Original OneMax: $\mathcal{F} = \sum_{i=1}^n b_i$ (with 2 op., 5-Bit and 1-Bit)
- 3 monotonous transformations: $\log(\mathcal{F})$, $\exp(\mathcal{F})$ and $(\mathcal{F})^2$

| (max – min) | $Avg(all)$ | $\mathcal{F} = \sum b_i$ | AOS tech. |
|-------------|-------------|--------------------------|--------------------|
| 95 | 5120±258 | 5088±252 | AUC - MAB |
| 0 | 6147± 458 | 6147± 458 | F-AUC - MAB |
| 2099 | 6367± 1416 | 5096± 240 | Ext - DMAB |
| 1758 | 7066± 1788 | 6492± 370 | Avg - DMAB |
| 8226 | 11019± 4012 | 7307± 499 | Avg - MAB |
| 8267 | 11263±3811 | 8322±652 | Ext - MAB |

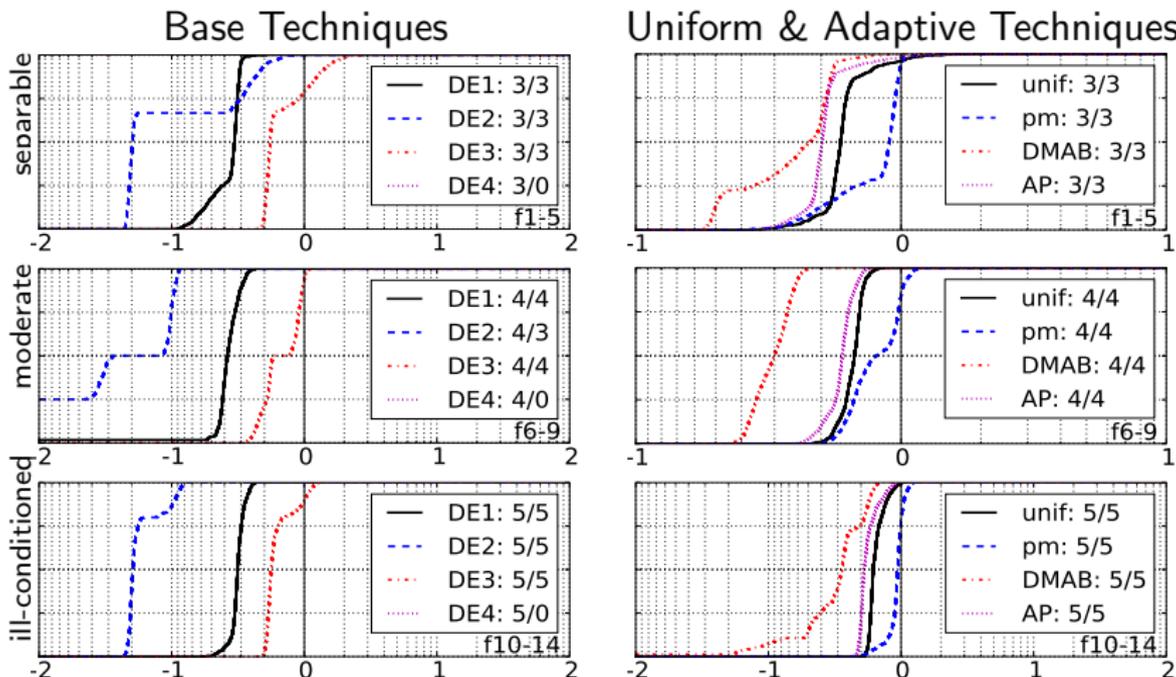
Black-Box Optimization Benchmark (BBOB)

- Standardize the benchmarking of continuous optimizers
 - Workshops at GECCO'09 and GECCO'10 [Hansen et al., 2010]
- 24 single-objective **continuous functions** divided in 5 classes
- Several problem dimensions (2, 3, 5, 10, 20, 40)
- Post-processing scripts provide standard tables and plots for comparisons

Adaptive Strategy Selection in Differential Evolution

- A completely different evolutionary algorithm
- With 4 possible mutation strategies
- Preliminary results presented at **BBOB'10**, extension to be presented at **PPSN'10**.

Parwise comparisons of AUC-Bandit with ...



X-axis: $\log\left(\frac{\text{effort}(\text{AUC-B})}{\text{effort}(\text{AOS})}\right)$ – Y-axis: Proportion of runs

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 $\hat{p} = \text{Area Under the Curve}$

- 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
- F-AUC-MAB **comparison-based**

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/naive/known adaptive approaches**

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

Thank you for your attention!

Got interested into Adaptive Operator Selection?

“Adaptive Strategy Selection in Differential Evolution”

Wenyin Gong, Álvaro Fialho, Zhihua Cai

- A different, much simpler, although efficient technique:

PM-AdapSS-DE

- Combined with DE, assessed on BBOB problem-set
- Collaboration with China University of Geosciences

ESEP: Best Paper Nominees
Today at 14h50, room Salon A
(session starts at 14h00)

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.



Goldberg, D. (1990).
Probability Matching, the Magnitude of Reinforcement, and Classifier System Bidding.
Machine Learning, 5(4):407–426.

References II



Hansen, N., Auger, A., Finck, S., and Ros, R. (2010).

Real-parameter black-box optimization benchmarking 2010: Experimental setup.
Technical Report RR-7215, INRIA.



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. (2007).

Adaptive strategies for operator allocation.

In Lobo, F., Lima, C., and Michalewicz, Z., editors, *Parameter Setting in Evolutionary Algorithms*, pages 77–90. Springer Verlag.

Toward Comparison-based Adaptive Operator Selection

Álvaro Fialho, Marc Schoenauer, Michèle Sebag



Orsay, France