6th GECCO Workshop on Blackbox Optimization Benchmarking (BBOB): Wrap-Up

The BBOBies
https://github.com/numbbo/coco
Numerical Blackbox Optimization

Optimize $f: \Omega \subset \mathbb{R}^n \mapsto \mathbb{R}^k$

$x \in \mathbb{R}^n \quad \rightarrow \quad f(x) \in \mathbb{R}^k$

derivatives not available or not useful
Need: Benchmarking

• understanding of algorithms
• algorithm selection
• putting algorithms to a standardized test
  • simplify judgement
  • simplify comparison
  • regression test under algorithm changes
that's where COCO and BBOB come into play

Comparing Continuous Optimizers Platform

https://github.com/numbbo/coco
Available Data Sets in COCO before and after 2016

Before 2016

- **bbob**
  - 140+ algo data sets

- **bbob-noisy**
  - 40+ algo data sets

- **bbob-biobj**
  - New in 2016
  - 15 data sets

In 2016

- 4 data sets
- 4 data sets
- 15 data sets
extension of COCO to multi-objective optimization
**bbob-biobj Testbed (new in 2016)**

- 55 functions, combining **bbob** functions
- 6 dimensions (2..40D)
- no normalization
- ideal/nadir known
- but Pareto set/front not (only refsets)

### 1. Separable Functions
- **f1** Sphere Function ✓
- **f2** Ellipsoidal Function ✓
- **f3** Rastrigin Function
- **f4** Büche-Rastrigin Function
- **f5** Linear Slope

### 2. Functions with low or moderate conditioning
- **f6** Attractive Sector Function ✓
- **f7** Step Ellipsoidal Function
- **f8** Rosenbrock Function, original ✓
- **f9** Rosenbrock Function, rotated

### 3. Functions with high conditioning and unimodal
- **f10** Ellipsoidal Function
- **f11** Discus Function
- **f12** Bent Cigar Function
- **f13** Sharp Ridge Function ✓
- **f14** Different Powers Function ✓

### 4. Multi-modal functions with adequate global structure
- **f15** Rastrigin Function ✓
- **f16** Weierstrass Function
- **f17** Schaffers F7 Function ✓
- **f18** Schaffers F7 Functions, moderately ill-conditioned
- **f19** Composite Griewank-Rosenbrock Function F8F2

### 5. Multi-modal functions with weak global structure
- **f20** Schwefel Function ✓
- **f21** Gallagher’s Gaussian 101-me Peaks Function ✓
- **f22** Gallagher’s Gaussian 21-hi Peaks Function
- **f23** Katsuura Function

---

<table>
<thead>
<tr>
<th>f1</th>
<th>f2</th>
<th>f3</th>
<th>f4</th>
<th>f5</th>
<th>f6</th>
<th>f7</th>
<th>f8</th>
<th>f9</th>
<th>f10</th>
</tr>
</thead>
<tbody>
<tr>
<td>f11</td>
<td>f12</td>
<td>f13</td>
<td>f14</td>
<td>f15</td>
<td>f16</td>
<td>f17</td>
<td>f18</td>
<td>f19</td>
<td>f20</td>
</tr>
<tr>
<td>f21</td>
<td>f22</td>
<td>f23</td>
<td>f24</td>
<td>f25</td>
<td>f26</td>
<td>f27</td>
<td>f28</td>
<td>f29</td>
<td>f30</td>
</tr>
<tr>
<td>f31</td>
<td>f32</td>
<td>f33</td>
<td>f34</td>
<td>f35</td>
<td>f36</td>
<td>f37</td>
<td>f38</td>
<td>f39</td>
<td>f40</td>
</tr>
<tr>
<td>f41</td>
<td>f42</td>
<td>f43</td>
<td>f44</td>
<td>f45</td>
<td>f46</td>
<td>f47</td>
<td>f48</td>
<td>f49</td>
<td>f50</td>
</tr>
<tr>
<td>f51</td>
<td>f52</td>
<td>f53</td>
<td>f54</td>
<td>f55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
algorithm quality =

\[ \text{normalized* hypervolume (HV) of all non-dominated solutions} \]

\[ \text{if a point dominates nadir} \]

\[ \text{closest normalized* negative distance to region of interest \([0,1]^2\)} \]

\[ \text{if no point dominates nadir} \]

* such that ideal=[0,0] and nadir=[1,1]
Bi-objective Performance Assessment

Again, as in last session's wrap-up:

• results are relative to a reference set, given as the best Pareto front approximation known (since exact Pareto set not known)
  • note: improved reference sets compared to workshop papers
• actual absolute hypervolume targets used are
  
  \[ \text{HV(refset)} - \text{targetprecision} \]

  with 51 fixed targetprecisions between 1 and \(10^{-5}\) (same for all functions, dimensions, and instances) in the displays

• all 10 instances are displayed
let's dig into the data...
all results for the bbob-biobj suites
The ECDFs are actually influenced by

① the number and set of instances and by
② the reference set and the reference hypervolume values

...so let's have a brief look behind the scenes of BBOB
the influence of the instances
BBOB-2016: Instances

• on the bbo-biobj test suite, experiments were run on 10 instances

• but all plots were based on the first 5 instances only
  • practical reason 1: we did not have enough data to produce good hypervolume reference values for all instances
  • practical reason 2: setting allows to investigate potential overfitting ("split between training and test")
BBOB-2016: Instances 1-5 (2-D)

target pairs

proportion of function

log10 of (# f-evals / dimension)
BBOB-2016: Instances 6-10 (2-D)
BBOB-2016: Instances 1-10 (same data)
now 20-D
BBOB-2016: Instances 1-10
Influence of the Instance Set

- is relatively small
- sometimes, last 5 instances harder, sometimes first 5 (depending on dimension)
- no indication of overfitting to the first 5 instances
2 the influence of the reference set
BBOB-2016: Reference Set Before

Proportion of function+target pairs

log10 of (# f-evals / dimension)
BBOB-2016: Reference Set

Bbob-biobj - f1-f55, 5-D

HMO-CMA-ES
UP-MO-CMA-ES
RM-MEDA
DEMO
RM-MEDA
SMS-EMOA-DH
SMS-EMOA-PM
RS-5
RS-4
MO-DIRECT-hv(HV)
MO-DIRECT-hv(ND)
MO-DIRECT-hv(Ran)
RS-100

Portion of function - target pairs

log10 of (# f-ev
als / dimension)
quick check: first 5 instances
BBOB-2016: Reference Set

Proportion of function

log10 of (# f-evals / dimension)
Influence of the Reference Set

• impact by the workshop algorithms the largest
• mainly on second five instances
  • which means the provided and displayed reference sets were okay
• continue with the current best in the future
  • updated reference hypervolume values will be provided in one of the next releases (this summer for sure)
• investigations on the single functions show that for some, we still do not have a good enough reference set yet
The Future of COCO

• bi-objective data will be made available online in the next days

• towards more realistic problems
  • large-scale test suite soon ready for release
  • constraints potentially ready in 2017
  • "almost real-world" problems

• online visualization of data
Your Participation it Welcome...

• ...always 😊

• benchmark your own algorithm and submit next year
• report bugs, issues, and feature requests
  • https://github.com/numbbo/coco

• contribute to the code base on github
  • issue tracker has special flag easy

• or even join us in Paris
  • as an engineer (funding for 1 year available)
  • or as postdoc, PhD student, or intern