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

The BBQBies!
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                                           In 2016
• bbob                      140+ algo data sets    4 data sets
• bbob-noisy                 40+ algo data sets    4 data sets
• bbob-biobj                new in 2016           15 data sets
extension of COCO to multi-objective optimization
Testbed (new in 2016)

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

algorithm quality =

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

if a point dominates nadir

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

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

  HV(refset) \ 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 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

als / dimension

log10 of (# f-ev)
BBOB-2016: Instances 1-10 (2-D)
BBOB-2016: Instances 1-10 (same data)
now 20-D
BBOB-2016: Instances 1-5

Proportion of function+target pairs

log10 of (# f-evals / dimension)
BBOB-2016: Instances 6-10

- bbob-biobj - f1-f55, 20-D
- 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5 instances

log10 of (# f-evals / dimension)
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
the influence of the reference set
BBOB-2016: Reference Set Before
quick check: first 5 instances
BBOB-2016: Reference Set Before
BBOB-2016: Reference Set

Proportion of function values

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

Proportion of function+target pairs

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 is 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