What now?

Open issues in benchmarking

1. What we are doing
2. What should we be doing?
What we are doing:
Benchmarking Network

Consolidate + stimulate activities on benchmarking iterative optimization heuristics

- Organising events
- Sharing resources
- Place for discussion

<table>
<thead>
<tr>
<th>Software</th>
<th>Topic</th>
<th>Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td>COCO</td>
<td>Performance Assessment, Single-Objective, Multi-Objective, Test Problems, Continuous Optimization</td>
<td>C, C++, Java, Python, MATLAB, Octave</td>
</tr>
<tr>
<td>DEAP</td>
<td>Single-Objective, Multi-Objective, Continuous Optimization, Discrete Optimization, Performance Assessment, Test Problems</td>
<td>Python</td>
</tr>
<tr>
<td>ECJ</td>
<td>Single-Objective, Multi-Objective, Continuous Optimization</td>
<td>Java</td>
</tr>
<tr>
<td>eor</td>
<td>Single-Objective, Multi-Objective, Performance Assessment, Continuous Optimization, Discrete Optimization</td>
<td>R</td>
</tr>
<tr>
<td>EO</td>
<td>Single-Objective, Continuous Optimization, Discrete Optimization</td>
<td>C++</td>
</tr>
<tr>
<td>fLacco</td>
<td>Features, Single-Objective, Continuous Optimization, Single-Objective</td>
<td>R, Python, GUI</td>
</tr>
</tbody>
</table>
GECCO 2020

- Workshop: Good Benchmarking Practices for Evolutionary Computation (now)
- Tutorial: Benchmarking and Analyzing Iterative Optimization Heuristics with IOHprofiler (July 09 [16:10 - 18:00])
- Poster: Towards Realistic Optimisation Benchmarks: Early Results from a Questionnaire on the Properties of Real-World Problems (July 10 [17:40 - 19:20])
Questionnaire on Real-World Optimization Problems

- Goal: Create new test problems for benchmarking optimization algorithms with properties of real-world problems
- Question: What are the properties of real-world optimization problems?

Optimization algorithms need to work well on real-world problems, but are developed and benchmarked using artificial ones.

Participate in the survey at https://tinyurl.com/opt-survey
PPSN 2020

- Workshop: Good Benchmarking Practices for Evolutionary Computation
- Workshop: (Multimodal) Multi-Objective Optimization
- Workshop: Understanding Machine Learning Optimization Problems
- Tutorial: Exploratory Landscape Analysis
- Competition: Game Benchmark Competition
- Competition: Open Optimization Competition
Ongoing Initiatives (and who to contact)

- COST Action ImAppNIO Benchmarking working group (Pietro S. Oliveto): https://imappnio.dcs.aber.ac.uk/working-groups/working-group-3
- COST Action Proposal incl. benchmarking working group (Carlos M. Fonseca)
- Thematic Seminars - Dagstuhl, Lorentz Center (Carola Doerr, Pascal Kerschke)
16th ACM/SIGEVO Workshop on Foundations of Genetic Algorithms

FOGA 2021
September 6-8, 2021

Premier event to discuss

- Theoretical foundations of randomized search heuristics
- Suitable analysis frameworks
- Benchmarking aspects
- Connections between search heuristics and machine learning

Location:
Vorarlberg University of Applied Sciences (FHV)
Dornbirn, Austria

Website: https://www.fhv.at/foga2021

Submission deadline: April 30, 2021
Author rebuttal phase: June 1-7, 2021
Early-registration until: July 14, 2021
What should we be doing?
Open Issues: Topics

- **Sharing** and **reproducing** benchmarking resources, tools and results
- **Characterising** and understanding problems (including unknown properties)
- Problem **coverage** / Overfitting to benchmarks
- **Analysis** of obtained results (including visualisation and statistical tests)
- **Relevance** for real-world problems + (theoretical) research questions
- **Performance measures** (especially multi-objective and unknown optima)
- **Usage** of Benchmarks
Open Issues: Topics

- **Sharing** and **reproducing** benchmarking resources, tools and results
- **Characterising** and understanding problems (including unknown properties)
- Problem **coverage** / Overfitting to benchmarks
- **Analysis** of obtained results (including visualisation and statistical tests)
- **Relevance** for real-world problems + (theoretical) research questions
- **Performance measures** (especially multi-objective and unknown optima)
- **Usage** of Benchmarks