Latest Benchmark Results

INFORMS Annual Conference

24-27 Oct 2021

H. D. Mittelmann
School of Mathematical and Statistical Sciences
Arizona State University
Services we provide

- Guide to Software: "Decision Tree"
- http://plato.asu.edu/guide.html
- Software Archive
- Software Evaluation: "Benchmarks"
- Archive of Testproblems
- Web-based Solvers (1/3 of NEOS)
Overview of Talk

• Sources, New Developments

• Current and Selected (*) Benchmarks (as of 10-22-21)
  – Benchmarks of Continuous Optimization Software
    * Simplex/Barrier QPLIB, [SOCP, SDP, NLP]
  – Benchmarks of Discrete Optimization Software
    * MILP, QPLIB, MISOCOP, MINLP

• Observations and Conclusions
Sources

- MIPLIB, QPLIB, TSPLIB, CBLIB, MINLPLIB, MPECLIB
- Own selections in LP, SDP, NLP, NETWORK

New Developments

- We cannot benchmark CPLEX and XPRESS, but indirectly
  - CPLEX in Concorde, SCIP; also used in our NEOS solvers
    * Concorde is our most used solver, SCIP is second
New Developments "China Rising"

- **Simplex**: Cardinal, AlibabaUS, Huawei

- **Barrier, Network**: Cardinal, AlibabaUS

- **MILP, SOCP**: Cardinal

**Nonlinear SCIP development at Github**

- Lead to **improvement** in QPLIB benchmarks, expected for MINLP, Implemented at NEOS
New Benchmarks

- **Infeasibility Detection** for MIPLIB problems

- Several instances replaced in Simplex and Barrier benchmarks

New NEOS solvers

- SCIP for **SOCP** (AMPL, MPS, LP input)

Next: List of Benchmarks
COMBINATORIAL OPTIMIZATION
Concorde-TSP with different LP solvers (9-16-2021)

LINEAR PROGRAMMING
* Benchmark of Simplex LP solvers (10-3-2021)
* Benchmark of Barrier LP solvers (10-3-2021)
* Large Network-LP Benchmark (commercial vs free) (10-22-2021)

SEMIDEFINITE/SQL PROGRAMMING
SQL problems from the 7th DIMACS Challenge (8-8-2002)
Several SDP codes on sparse and other SDP problems (6-6-2020)
Infeasible SDP Benchmark (10-12-2020)
Large SOCP Benchmark (10-2-2021)
* MISOCNP Benchmark (9-3-2021)

PROBLEMS WITH EQUILIBRIUM CONSTRAINTS
MPEC Benchmark (1-9-2021)
MIXED INTEGER LINEAR PROGRAMMING

* MILP Benchmark - MIPLIB2017 (10-5-2021)
MILP cases that are slightly pathological (10-7-2021)
Infeasibility Detection for MILP Problems (10-4-2021)

NONLINEAR PROGRAMMING

AMPL-NLP Benchmark (12-10-2020)

MIXED INTEGER QPs and QCPs

Non-commercial convex QP Benchmark (9-16-2021)
* Binary Non-Convex QPLIB Benchmark (8-3-2021)
* Discrete Non-Convex QPLIB Benchmark (non-binary) (9-7-2021)
* Continuous Non-Convex QPLIB Benchmark (9-16-2021)
* Convex Continuous QPLIB Benchmark (5-5-2021)
* Convex Discrete QPLIB Benchmark (8-21-2021)

MIXED INTEGER NONLINEAR PROGRAMMING

* MINLP Benchmark (8-15-2021)
Overview of Talk

- Selected Benchmarks
  - Benchmarks of Continuous Optimization Software
    * Simplex/Barrier QPLIB, [SOCP, SDP, NLP]
  - Benchmarks of Discrete Optimization Software
    * MILP, QPLIB, MISOCPP, MINLP

- Observations and Conclusions
Benchmark of Simplex LP solvers

H. Mittelmann (mittelmann@asu.edu)

This benchmark was run on a Linux-PC (i7-11700K, 3.6GHz, 64GB).

MOSEK-9.2/3  www.mosek.com
CLP-1.17.6  projects.coin-or.org/Clp
Google-GLOP  LP with Glop
SOPLEX-5.0.2  soplex.zib.de/
Gurobi-9.1.2  gurobi.com
GLPK-5.00  www.gnu.org/software/glpk/glpk.html
MATLAB-R2020b  mathworks.com (dual-simplex)
COPT-2.0.1  Cardinal Optimization
MindOpt-0.14.0  alibabaUS
Optverse-0.2.1  huawei.com
HiGHS-1.1.0  HiGHS@github

Unscaled and scaled shifted (by 10 sec) geometric mean of runtimes

<p>| | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>653</td>
<td>270</td>
<td>1008</td>
<td>1465</td>
<td>69.9</td>
<td>3758</td>
<td>1415</td>
<td>38.8</td>
<td>30.9</td>
<td>42.6</td>
</tr>
<tr>
<td>40 probs</td>
<td>21.1</td>
<td>8.73</td>
<td>32.6</td>
<td>47.3</td>
<td>2.26</td>
<td>121</td>
<td>45.7</td>
<td>1.25</td>
<td>1</td>
<td>1.38</td>
</tr>
<tr>
<td>solved</td>
<td>35</td>
<td>37</td>
<td>27</td>
<td>31</td>
<td>40</td>
<td>25</td>
<td>29</td>
<td>40</td>
<td>40</td>
<td>39</td>
</tr>
</tbody>
</table>
This benchmark was run on a Linux-PC (i7-11700K, 3.6GHz, 64GB).

The barrier methods were tested of:

- MOSEK-9.2/3 www.mosek.com
- MATLAB-R2020b mathworks.com (interior-point, NO CROSSOVER)
- Gurobi-9.1.2 gurobi.com
- CLP-1.17.6 projects.coin-or.org/Clp
- Tulip-0.8.0 [Tulip at Github] (NO CROSSOVER)
- COPT-2.0.1 [Cardinal Optimization]
- MindOpt-0.14.0 MindOpt
- KNITRO-12.4.0 www.artelys.com/knitro/ (NO CROSSOVER)
- HiGHS-1.1.0 HiGHS@github

Unscaled and scaled shifted (by 10 sec) geometric mean of runtimes

<table>
<thead>
<tr>
<th></th>
<th>90.8</th>
<th>708</th>
<th>30.1</th>
<th>1324</th>
<th>1078</th>
<th>19.3</th>
<th>44.9</th>
<th>275</th>
<th>293</th>
</tr>
</thead>
<tbody>
<tr>
<td>47 probs</td>
<td>4.71</td>
<td>36.7</td>
<td>1.56</td>
<td>68.6</td>
<td>55.9</td>
<td>1</td>
<td>2.33</td>
<td>14.3</td>
<td>15.2</td>
</tr>
<tr>
<td>solved</td>
<td>42</td>
<td>35</td>
<td>47</td>
<td>35</td>
<td>31</td>
<td>47</td>
<td>47</td>
<td>37</td>
<td>41</td>
</tr>
</tbody>
</table>

===============================================================================

<table>
<thead>
<tr>
<th>problem</th>
<th>MOSEK</th>
<th>MATLAB</th>
<th>Gurobi</th>
<th>CLP</th>
<th>TULIP</th>
<th>COPT</th>
<th>MDOPT</th>
<th>KNITRO</th>
<th>HiGHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>----------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>------</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>3 Oct 2021</td>
<td>======</td>
<td>======</td>
<td>========</td>
<td>======</td>
<td>=======</td>
<td>======</td>
<td>=======</td>
<td>========</td>
<td>=======</td>
</tr>
</tbody>
</table>
Large Network-LP Benchmark (commercial vs free)

H. Mittelmann (mittelmann@asu.edu)

Logfiles of these runs at: plato.asu.edu/ftp/net_logs/

This benchmark was run on an Intel i7-11700K (3.6 GHz, 64GB, Linux, 8 cores). The following codes were tested with one thread and in default mode except where indicated:

- www.mosek.com/ MOSEK-9.3.1
- projects.coin-or.org/Clp/ Clp-1.17.6 (-network -dualsimplex)
- www.math.uwaterloo.ca/~bico/qsopt/ QSopt-1.01 (option -d 9)
- mathworks.com MATLAB-R2020b (dual-simplex)
- soplex.zib.de/ SOPLEX-5.0.2
- COPT COPT-3.0.2
- Gurobi Gurobi-9.1.2
- MindOpt MindOpt-0.15.1
- HiGHS HiGHS-1.1.1

Times are elapsed times. 1 hour time limit.

Scaled and shifted geometric means of runtimes.

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>38.6</td>
<td>3.04</td>
<td>22.9</td>
<td>12.6</td>
<td>41.4</td>
<td>1</td>
<td>3.64</td>
<td>1.35</td>
</tr>
</tbody>
</table>

problem nodes arcs MOSEK CLP QSOPT MATL SOPLX COPT GUR MDOPT HGHS
16 Sep 2021

Continuous Non-Convex QPLIB Benchmark

H. Mittelmann (mittelmann@asu.edu)

Logiles at plato.asu.edu/ftp/cnconv_logs/

Baron-21.1.13   BARON
ANTIGONE-1.1    ANTIGONE
SCIP-7.0.3.5    SCIP/SOPLEX
COUENNE-0.5     COUENNE [projects.coin-or.org/Couenne]
MINOTAUR-0.2.1  MINOTAUR
Octeract-3.4.1   OCTERACT
Gurobi-9.1.0    GUROBI
RAPOSA-2.0.2    RAPOSa

The above solvers were run on an Intel Xeon E5-4657L (48 cores, 512GB) on the continuous non-convex problems (102 total) from QPLIB.
Times given are elapsed times in seconds. Time limit 3hrs.
Only those instances are shown for which at least one solver succeeded.
Shifted and scaled geometric mean of runtimes:

<table>
<thead>
<tr>
<th>mean</th>
<th>1.59</th>
<th>1.83</th>
<th>4.39</th>
<th>4.63</th>
<th>3.55</th>
<th>2.47</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>solved</td>
<td>29</td>
<td>25</td>
<td>8</td>
<td>6</td>
<td>12</td>
<td>16</td>
<td>28*</td>
</tr>
</tbody>
</table>

==================================================================
prob#  ANTIGONE   BARON   COUENNE   MINOTAUR   SCIP   OCTERACT   GUROBI
==================================================================
The above solvers were run on a 3 GHz Intel i7-5960X (8 cores, 48GB) on
the 32 continuous convex problems from QPLIB. Times given are elapsed
times in seconds; time limit 2hrs, 8 threads
Shifted and scaled geometric mean of runtimes:

<table>
<thead>
<tr>
<th></th>
<th>prob#</th>
<th>MOSEK</th>
<th>KNITRO</th>
<th>IPOPT</th>
<th>Gurobi</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>solved</td>
<td>1</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.74</td>
<td>3.61</td>
<td>1.27</td>
<td></td>
</tr>
</tbody>
</table>

Logiles at plato.asu.edu/ftp/cconvex_logs/
Overview of Talk

- Selected Benchmarks
  - Benchmarks of Continuous Optimization Software
    * Simplex/Barrier QLIB, [SOCP, SDP, NLP]
  - Benchmarks of Discrete Optimization Software
    * MILP, QPLIB, MISOCOCP, MINLP

- Observations and Conclusions
The MIPLIB2017 Benchmark Instances

H. Mittelmann (mittelmann@asu.edu)

The following codes were run with a limit of 2 hours on the MIPLIB2017 benchmark set on two platforms.

1 thread: Intel i7-4790K, 4 cores, 32GB, 4Ghz;
8 thread: Intel i7-11700K, 8 cores, 64GB, 3.6GHz;

CBC-2.10.5:   projects.coin-or.org/Cbc
COPT-2.0.1:   Cardinal Optimization
GLPK-5.0:     www.gnu.org/software/glpk/glpk.html
LP_SOLVE-5.5.2: lpsolve.sourceforge.net/
MATLAB-2020a: MATLAB (intlinprog)
(F)SCIP/spx]-7.0.0: FiberSCIP (SCIP+SOPLEX on 1 thread)
Gurobi-9.1.0:  Gurobi
HiGHS-1.1.1:   HiGHS@github
The third line lists the number of problems (240 total) solved.

<table>
<thead>
<tr>
<th>1 thr</th>
<th>CBC</th>
<th>GLPK</th>
<th>LP_SOL</th>
<th>MATLAB</th>
<th>SCIP</th>
<th>Gurobi</th>
<th>COPT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unscl</td>
<td>2107</td>
<td>5044</td>
<td>5335</td>
<td>3301</td>
<td>1100</td>
<td>245</td>
<td>1029</td>
</tr>
<tr>
<td>scaled</td>
<td>8.59</td>
<td>20.5</td>
<td>21.7</td>
<td>13.5</td>
<td>4.48</td>
<td>1</td>
<td>4.19</td>
</tr>
<tr>
<td>solved</td>
<td>89</td>
<td>23</td>
<td>20</td>
<td>63</td>
<td>125</td>
<td>201</td>
<td>132</td>
</tr>
</tbody>
</table>

Due to wide availability of multi-cores this benchmark will not be updated.

<table>
<thead>
<tr>
<th>8 thr</th>
<th>CBC</th>
<th>FSCIP</th>
<th>Gurobi</th>
<th>COPT</th>
<th>SCIP</th>
<th>HiGHS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unscl</td>
<td>1328</td>
<td>794</td>
<td>100</td>
<td>450</td>
<td>955</td>
<td>1195</td>
</tr>
<tr>
<td>scaled</td>
<td>13.3</td>
<td>7.98</td>
<td>1</td>
<td>4.52</td>
<td>9.60</td>
<td>12.0</td>
</tr>
<tr>
<td>solved</td>
<td>107</td>
<td>146</td>
<td>225</td>
<td>176</td>
<td>134</td>
<td>134</td>
</tr>
</tbody>
</table>
Binary Non-Convex QPLIB Benchmark

H. Mittelmann (mittelmann@asu.edu)

Logiles at plato.asu.edu/ftp/qplib_logs/

Baron-21.1.13 BARON
(F)SCIP-7.0.0 (Fiber)SCIP-SOPLEX (only open source code included)
ANTIGONE-1.1 ANTIGONE
MINOTAUR-0.2.1 MINOTAUR
OCTERACT-3.4.1 OCTERACT
Gurobi-9.R1.2 Gurobi

The above solvers were run on a 3 GHz Intel i7-5960X (8 cores, 48GB) on the binary nonconvex problems (128 total) from QPLIB. Times given are elapsed times in seconds. Mipgap is zero, time limit 1hr; 8 threads, except SCIP&Minotaur. Only those instances are shown for which at least one solver succeeded. Shifted and scaled geometric mean of runtimes:

mean 13.4 44.3 53.3 71.1 21.0 1 37.6
solved 42 26 23 7 36 85 34

.prob# BARON SCIP ANTIGONE MINOTAUR OCTERACT GUROBI FSCIP
The above solvers were run on an Intel Xeon E5-4657L (48 cores, 512GB) on the discrete non-convex problems (160 total) with not only binary variables from QPLIB. Times given are elapsed times in seconds. Time limit 3hrs. Only those instances are shown for which at least one solver succeeded. Shifted and scaled geometric mean of runtimes:

<table>
<thead>
<tr>
<th>prob#</th>
<th>ANTIGONE</th>
<th>BARON</th>
<th>COUENNE</th>
<th>MINOTAUR</th>
<th>SCIP</th>
<th>OCTERACT</th>
<th>GUROBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>23.3</td>
<td>20.7</td>
<td>51.5</td>
<td>59.1</td>
<td>10.8</td>
<td>26.0</td>
<td>1</td>
</tr>
<tr>
<td>solved</td>
<td>29</td>
<td>27</td>
<td>8</td>
<td>4</td>
<td>37</td>
<td>20</td>
<td>66*</td>
</tr>
</tbody>
</table>
Convex Discrete QPLIB Benchmark

H. Mittelmann (mittelmann@asu.edu)

Logiles at plato.asu.edu/ftp/convex_logs/

<table>
<thead>
<tr>
<th>prob#</th>
<th>MOSEK</th>
<th>KNITRO</th>
<th>BARON</th>
<th>Bonmin</th>
<th>SCIP</th>
<th>ANTIGONE</th>
<th>MINOTAUR</th>
<th>Gurobi</th>
<th>Shot</th>
<th>OCTERACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>10.3</td>
<td>16.0</td>
<td>5.46</td>
<td>15.7</td>
<td>12.6</td>
<td>41.3</td>
<td>26.0</td>
<td>1</td>
<td>48.3</td>
<td>34.4</td>
</tr>
</tbody>
</table>

The above solvers were run on a 3 GHz Intel i7-5960X (8 cores, 48GB) on the 31 discrete convex problems from QPLIB. Times given are elapsed times in seconds. Mipgap zero, time limit 2hrs, 8 threads; SCIP, Minotaur 1 thread

Shifted and scaled geometric mean of runtimes:

<table>
<thead>
<tr>
<th>prob#</th>
<th>MOSEK</th>
<th>KNITRO</th>
<th>BARON</th>
<th>Bonmin</th>
<th>SCIP</th>
<th>ANTIGONE</th>
<th>MINOTAUR</th>
<th>Gurobi</th>
<th>Shot</th>
<th>OCTERACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>10.3</td>
<td>16.0</td>
<td>5.46</td>
<td>15.7</td>
<td>12.6</td>
<td>41.3</td>
<td>26.0</td>
<td>1</td>
<td>48.3</td>
<td>34.4</td>
</tr>
</tbody>
</table>
Mixed-integer SOCP Benchmark  
Hans D. Mittelmann (mittelmann@asu.edu)

Logfiles for these runs are at: plato.la.asu.edu/ftp/misocp_logs/

MOSEK-9.3.1 MOSEK  
SCIP-7.0.3.5 SCIP+Cplex  
Gurobi-9.1.2 Gurobi

These codes were tested on a selection of the MISOCP problems from CBLIB2014 and from here. Given are total CPU seconds.

The codes were run in default mode (except mipgap=0) on an Intel i7-11700K (3.6 GHz, 64GB). Time limit 2 hrs.

Scaled shifted geometric means of runtimes (t counted as maxtime)

<table>
<thead>
<tr>
<th>Problem solved of 47</th>
<th>MOSEK</th>
<th>SCIP</th>
<th>GUROBI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>21.2</td>
<td>26.2</td>
<td>1</td>
</tr>
</tbody>
</table>
The following codes were run through GAMS-36.1.0 with a limit of 2 hours on these instances from MINLPLIB and with one thread on an Intel i7-11700K, 64GB, 3.6GHz.

Description of selection process of benchmark instances. Statistics of the instances.

ANTIGONE, BARON, LINDO, SCIP

Table for all solvers, Result files per solver, Log files per solver, Trace files per solver, Error files per solver

Scaled and shifted geometric means of run times

Feasibility tolerance set to 1e-6. All non-successes are counted as max-time.

The shifted geometric mean is computed on the 69 instances for which no solver failed.

<table>
<thead>
<tr>
<th>87 probs</th>
<th>ANTIGONE</th>
<th>BARON</th>
<th>LINDO</th>
<th>SCIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>geom mean</td>
<td>4.17</td>
<td>1</td>
<td>5.93</td>
<td>2.06</td>
</tr>
<tr>
<td>solved</td>
<td>53</td>
<td>63</td>
<td>29</td>
<td>55</td>
</tr>
</tbody>
</table>
Overview of Talk

- Selected Benchmarks
  - Benchmarks of Continuous Optimization Software
    * Simplex/Barrier QPLIB, [SOCP, SDP, NLP]
  - Benchmarks of Discrete Optimization Software
    * MILP, QPLIB, MISOCP, MINLP

- Observations and Conclusions
"TOP PERFORMERS"

LP-Simplex: MindOpt, COPT, OptVerse, Gurobi
LP-Barrier: COPT, Gurobi, MindOpt, MOSEK
Network-LP: COPT, MindOpt, Gurobi
MILP: Gurobi, COPT, (F)SCIP
SOCP: MOSEK, Gurobi, COPT
MISOCP: Gurobi, MOSEK, SCIP
nonc BQCP: Gurobi, BARON, FSCIP
nonc DQCP: Gurobi, SCIP, BARON
nonc CQCP: Gurobi, ANTIGONE, BARON
conv DQCP: Gurobi, BARON, MOSEK
conv CQCP: MOSEK, Gurobi, KNITRO
MINLP: BARON, SCIP, ANTIGONE
Thank you!

slides at: http://plato.asu.edu/talks/