Automated Performance Testing and Analysis

Armin Pruessner
GAMS Development Corporation

INFORMS Annual Meeting, San Jose
Oct 17-20, 2002

Agenda

1. Performance World
   - Brief overview

2. Main Focus: Performance Tools
   - Performance Tools and Metrics
   - PAVER Web Server (Automation)
   - Examples
   - Conclusions

Performance World

Performance World

- Editorial Board
- PerformanceLib
- Performance Tools
- Performance List
- Related Links
- Search
- Contact

Welcome to the Performance World!

Performance World is a forum for discussion and dissemination of information and tools about all aspects of performance testing of solvers for mathematical programming problems. This world has been established in response to user demands for independent and reproducible performance results.

Overall performance highly depends on problem formulation, solver, and tuning parameters. Our performance tools are designed to serve the different needs of our user community. One user may be interested in finding the most reliable way to solve a proprietary or classified model. On the other hand, an academic researcher may be interested in testing a new algorithm against a set of existing test problems and competing approaches. The main features are:

- Uniform access to a comprehensive set of established and new test problems
- Automation tools for collecting performance measurements
- Tools for analyzing and visualizing test results

What's New:

- Try our online PAVER Server for automated performance analysis and visualization, batch file creation and model translation
- New tools for analyzing non-convex or discrete models
- MINLP type models from the MINLP World have been added to the PerformanceLib

Motivation for Tools

Performance Tools driven by user needs:

- Finding the most reliable way to solve a proprietary model
- Testing a new algorithm against a set of existing test problems and competing approaches
- Reproducibility of performance results

Performance World

Default Solver Options

Performance Analysis & Visualization

Custom Solver Options

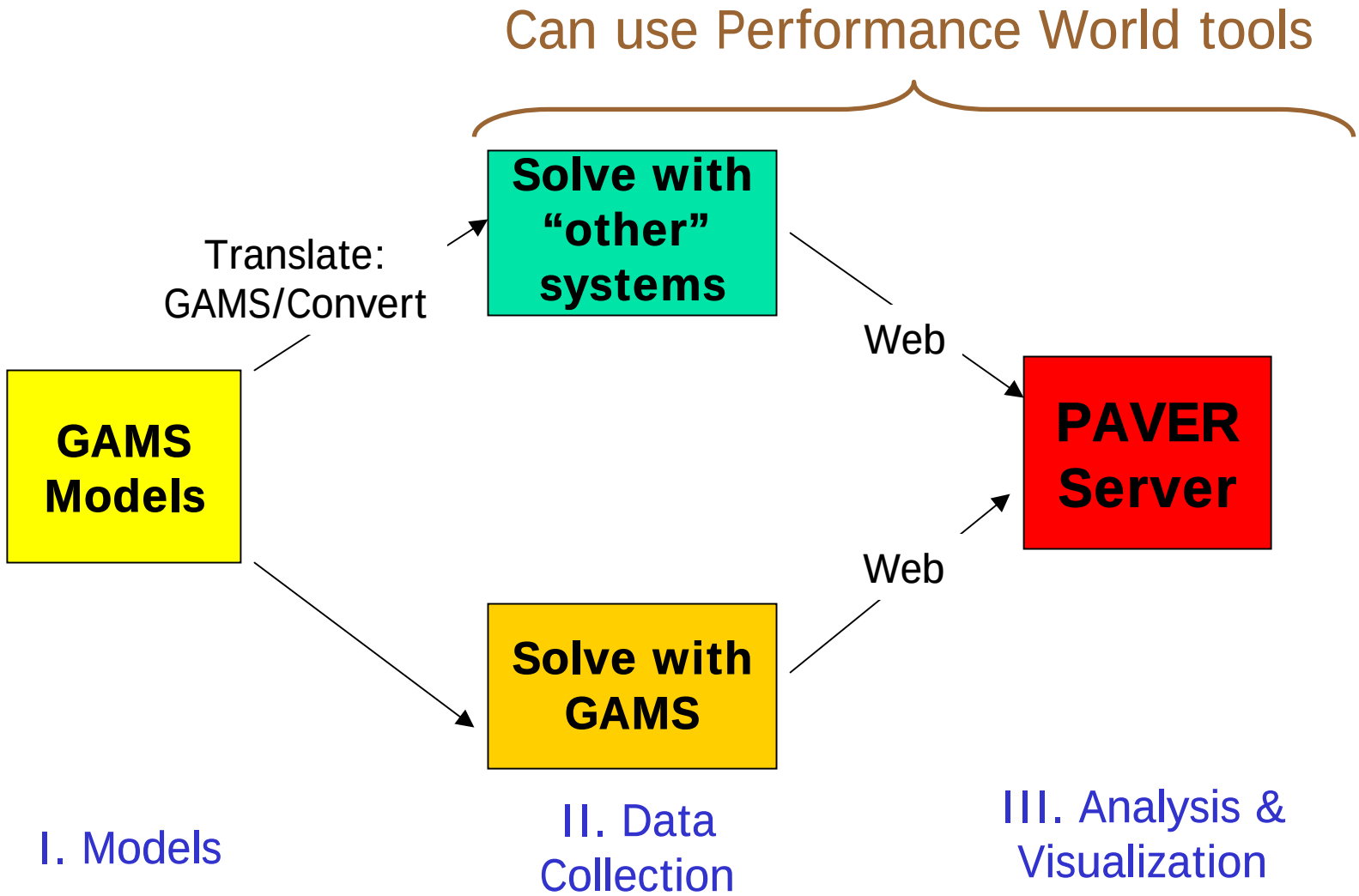
Tools: Performance Analysis

- Different objectives:
  - Solver robustness and correctness
  - Solver efficiency
  - Quality of solution (nonconvex and discrete models)

→Tools are GAMS independent

→Results in HTML format: platform independent

Open Testing Architecture



PAVER Server

- PAVER server (**P**erformance **A**nalysis and **V**isualization for **E**ffortless **R**eproducibility)

  www.gamsworld.org/performance/paver

- Online server to facilitate performance testing and analysis/visualization

- Results sent via e-mail in HTML format
  - System independent

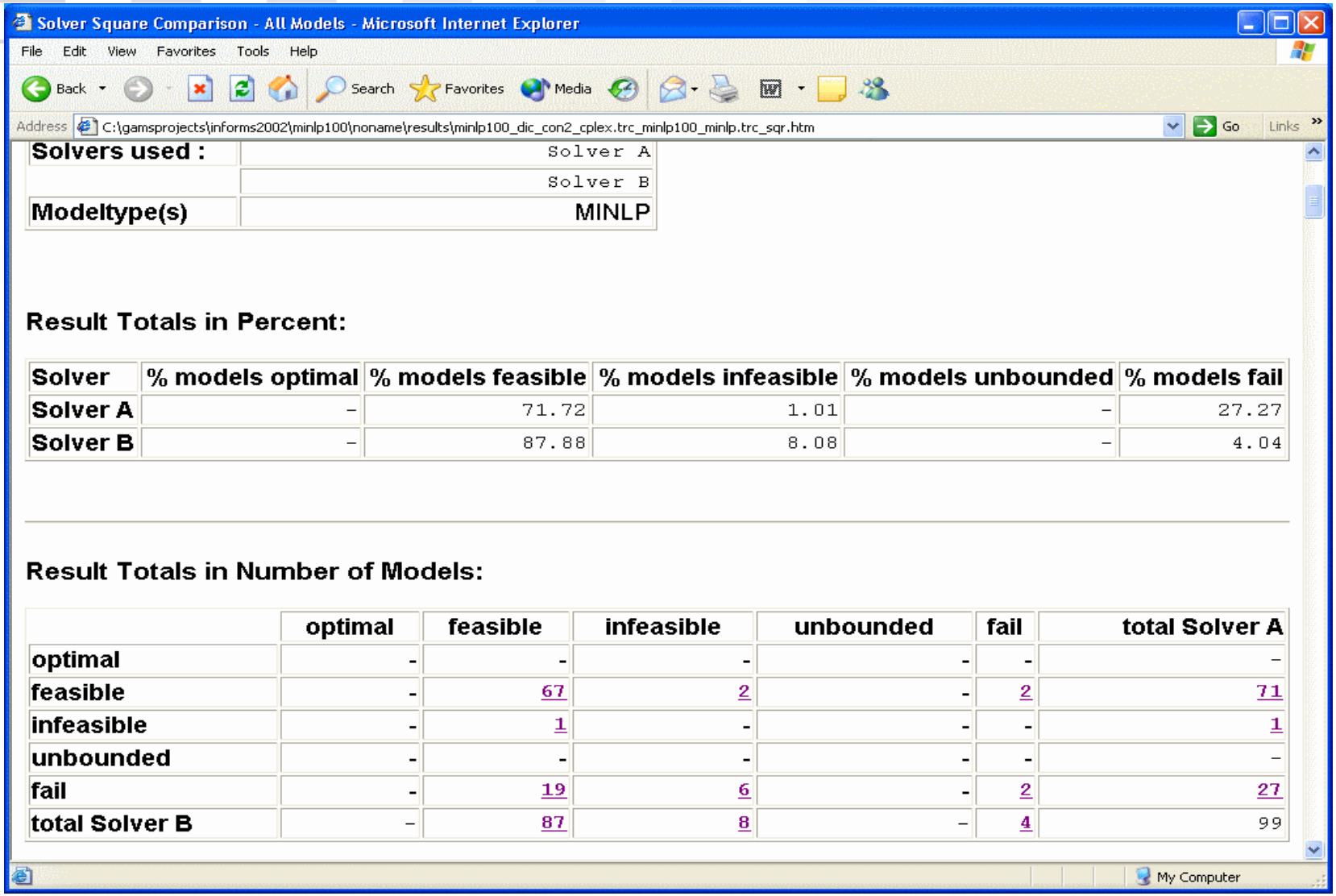
Tools: Robustness

Solver Square Utility:

- Cross comparison of solver outcomes of two solvers:
  - Optimal, integer, infeasible, unbounded, fail
- Compact tabular form for results
- Shows resource time and objective value information

→ Can use online using PAVER

PAVER: Solver Square



Solvers used :	Solver A
	Solver B
Modeltype(s)	MINLP

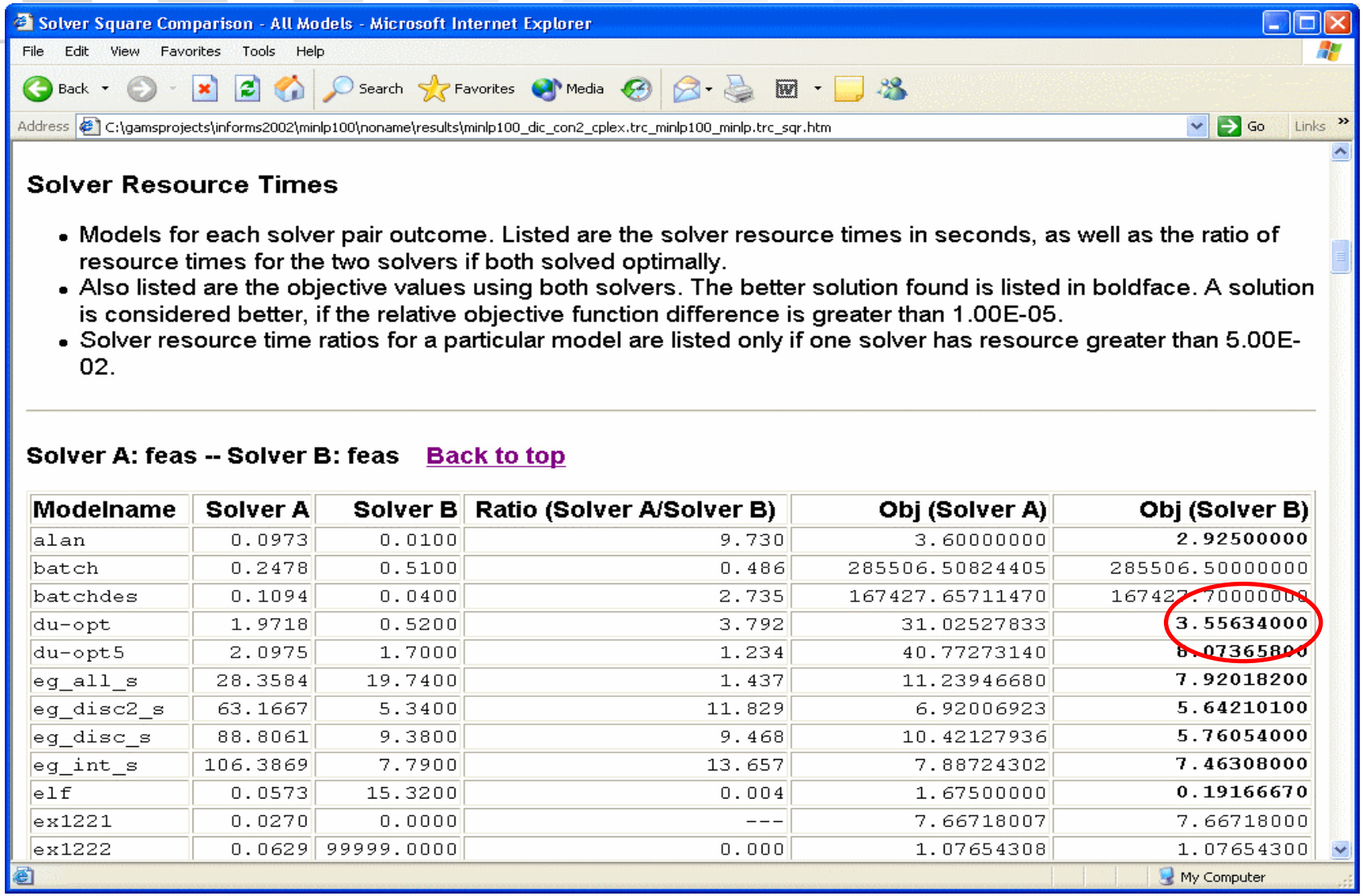
Result Totals in Percent:

Solver	% models optimal	% models feasible	% models infeasible	% models unbounded	% models fail
Solver A	-	71.72	1.01	-	27.27
Solver B	-	87.88	8.08	-	4.04

Result Totals in Number of Models:

	optimal	feasible	infeasible	unbounded	fail	total Solver A
optimal	-	-	-	-	-	-
feasible	-	67	2	-	2	71
infeasible	-	1	-	-	-	1
unbounded	-	-	-	-	-	-
fail	-	19	6	-	2	27
total Solver B	-	87	8	-	4	99

PAVER: Square (cont.)



Modelname	Solver A	Solver B	Ratio (Solver A/Solver B)	Obj (Solver A)	Obj (Solver B)
alan	0.0973	0.0100	9.730	3.60000000	**2.92500000**
batch	0.2478	0.5100	0.486	285506.50824405	285506.50000000
batchdes	0.1094	0.0400	2.735	167427.65711470	167427.70000000
du-opt	1.9718	0.5200	3.792	31.02527833	**3.55634000**
du-opt5	2.0975	1.7000	1.234	40.77273140	**8.07365800**
eg_all_s	28.3584	19.7400	1.437	11.23946680	**7.92018200**
eg_disc2_s	63.1667	5.3400	11.829	6.92006923	**5.64210100**
eg_disc_s	88.8061	9.3800	9.468	10.42127936	**5.76054000**
eg_int_s	106.3869	7.7900	13.657	7.88724302	**7.46308000**
elf	0.0573	15.3200	0.004	1.67500000	**0.19166670**
ex1221	0.0270	0.0000	---	7.66718007	7.66718000
ex1222	0.0629	99999.0000	0.000	1.07654308	1.07654300

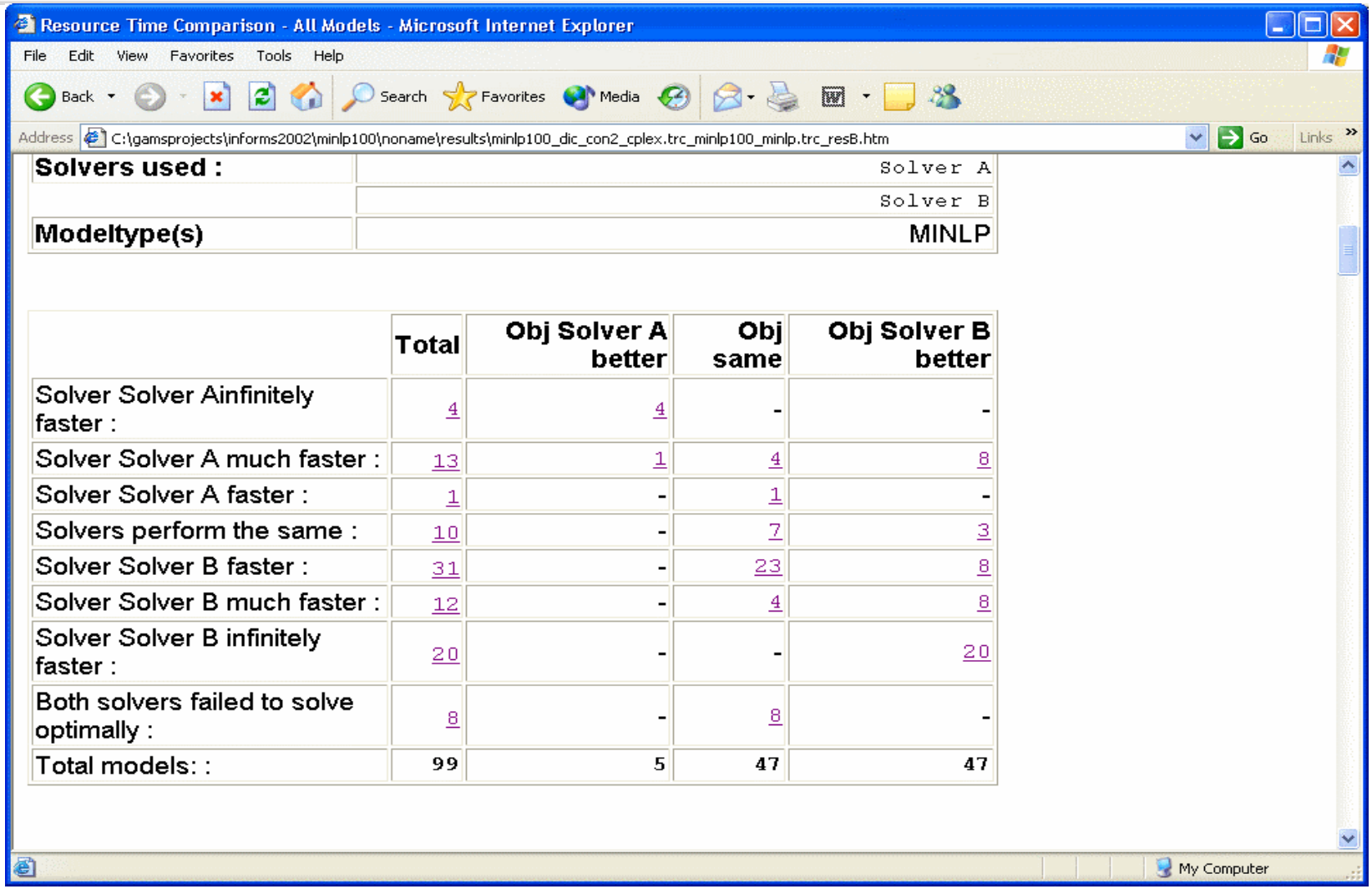
Tools: Efficiency

Resource Time Utility:

- Cross comparison of solver resource times of two solvers
- Further disaggregation by objective function
- Ratios of resource times

→ Can use online using PAVER

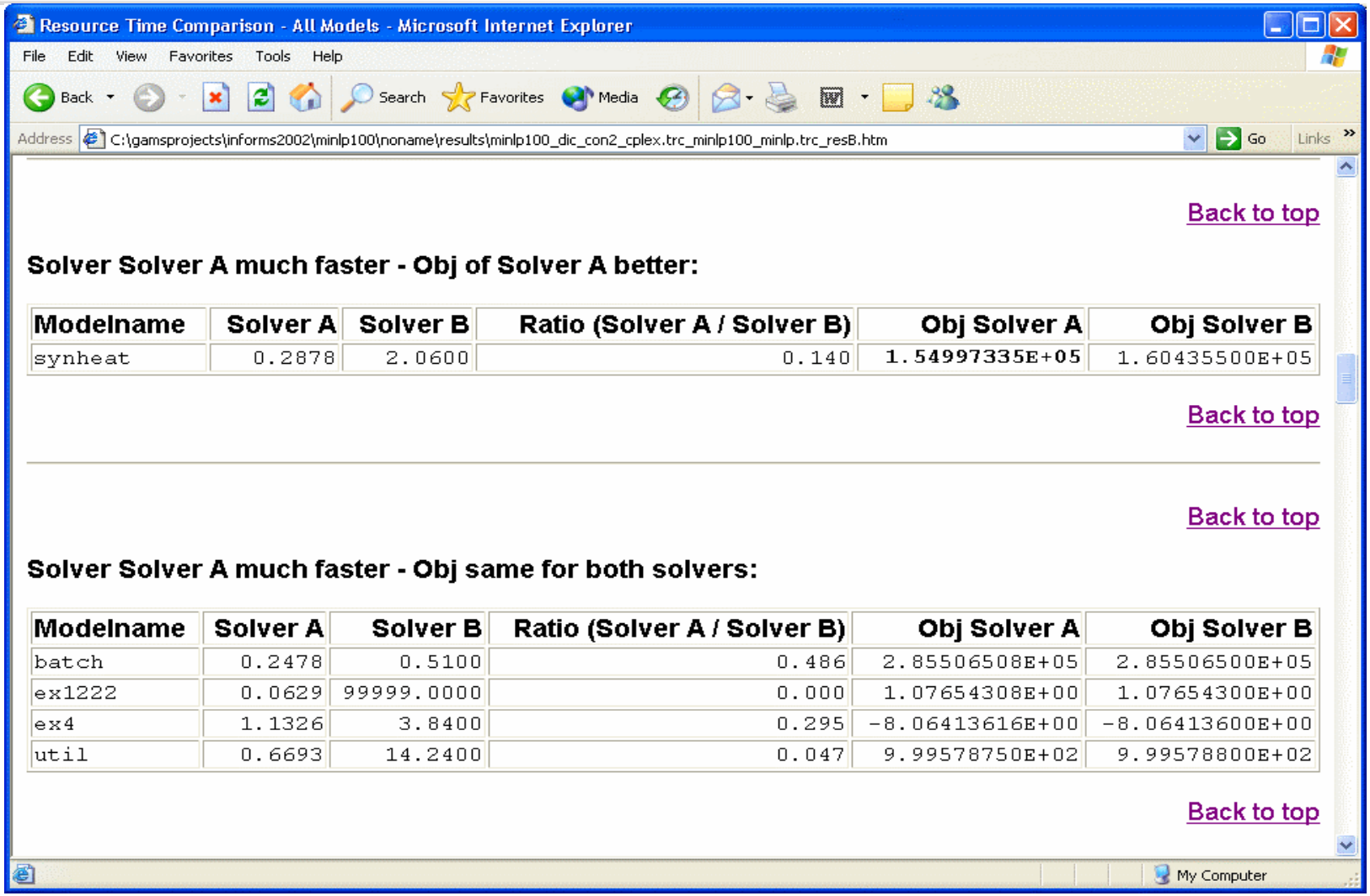
PAVER: Solver Resource Time



Solvers used :	Solver A
	Solver B
Modeltype(s)	MINLP

	Total	Obj Solver A better	Obj same	Obj Solver B better
Solver Solver Ainfinitely faster :	4	4	-	-
Solver Solver A much faster :	13	1	4	8
Solver Solver A faster :	1	-	1	-
Solvers perform the same :	10	-	7	3
Solver Solver B faster :	31	-	23	8
Solver Solver B much faster :	12	-	4	8
Solver Solver B infinitely faster :	20	-	-	20
Both solvers failed to solve optimally :	8	-	8	-
Total models: :	99	5	47	47

PAVER: Resource Time (cont.)



Tools: Visualization

Performance Profiles (Dolan and More, 2002):

- Cumulative distribution function for a performance metric
- Performance metric: ratio of current solver time over best time of all solvers
- Intuitively: probability of success if given τ times fastest time (τ=ratio)

Tools: Performance Profiles

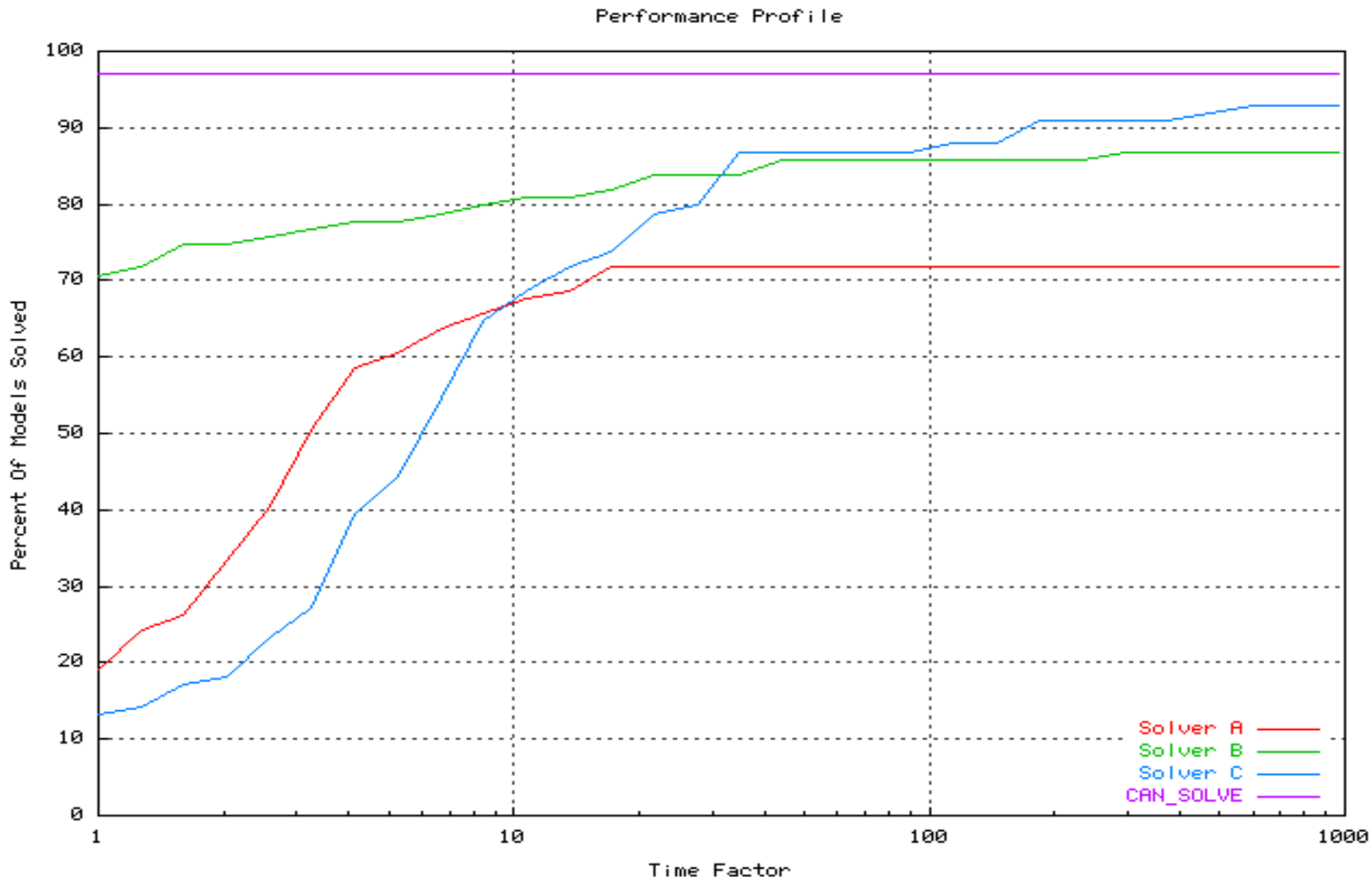
Interpretation (for τ=*ratio*, *P*=*profile*):

- Efficiency: $P(\tau) \text{ for } \tau = 1$
- Probability of success:

$$\lim P(\tau) \text{ as } \tau \to \infty$$

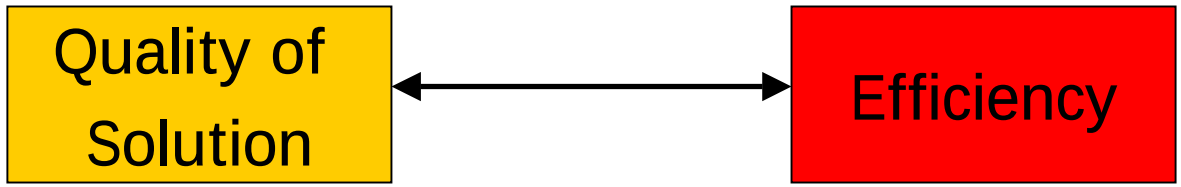
- Compact graphs summarize all information

Profiles (best resource time)



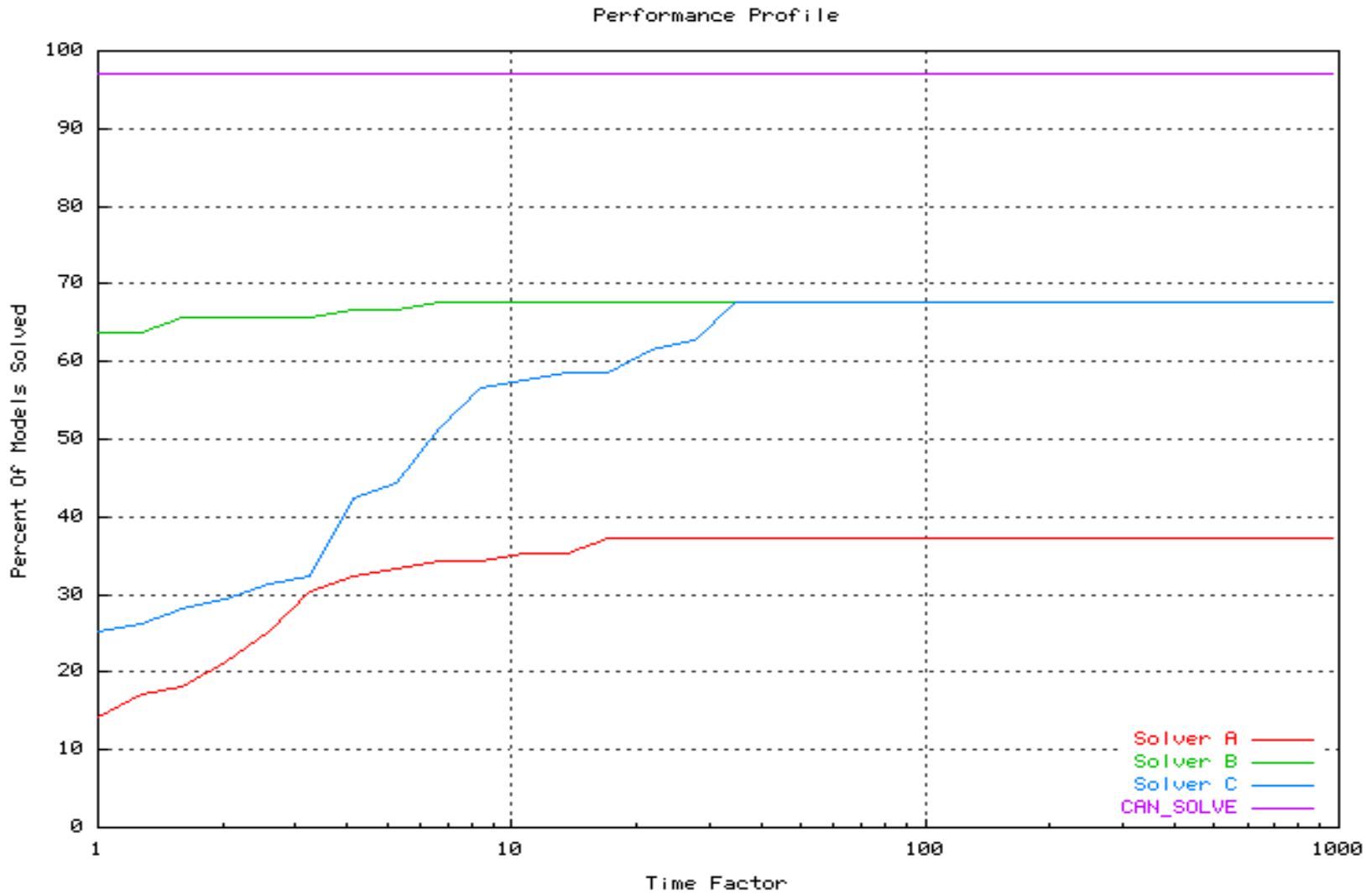
Tools: Visualization

Performance Profiles: considers both



- Further disaggregation by objective function:
- Success only if best solution (over all solvers) found

Profiles (best objective)



Benchmarking process

Two components:

- Subjective component:
  - Choice of models
  - Choice of solvers
  - Choice of solver options
- Non-subjective component:
  - Obtaining performance data
  - Performance analysis and visualization

→ reproducible

Subjectivity in Benchmarking

Performance Tools:

- Takes care of non-subjective component

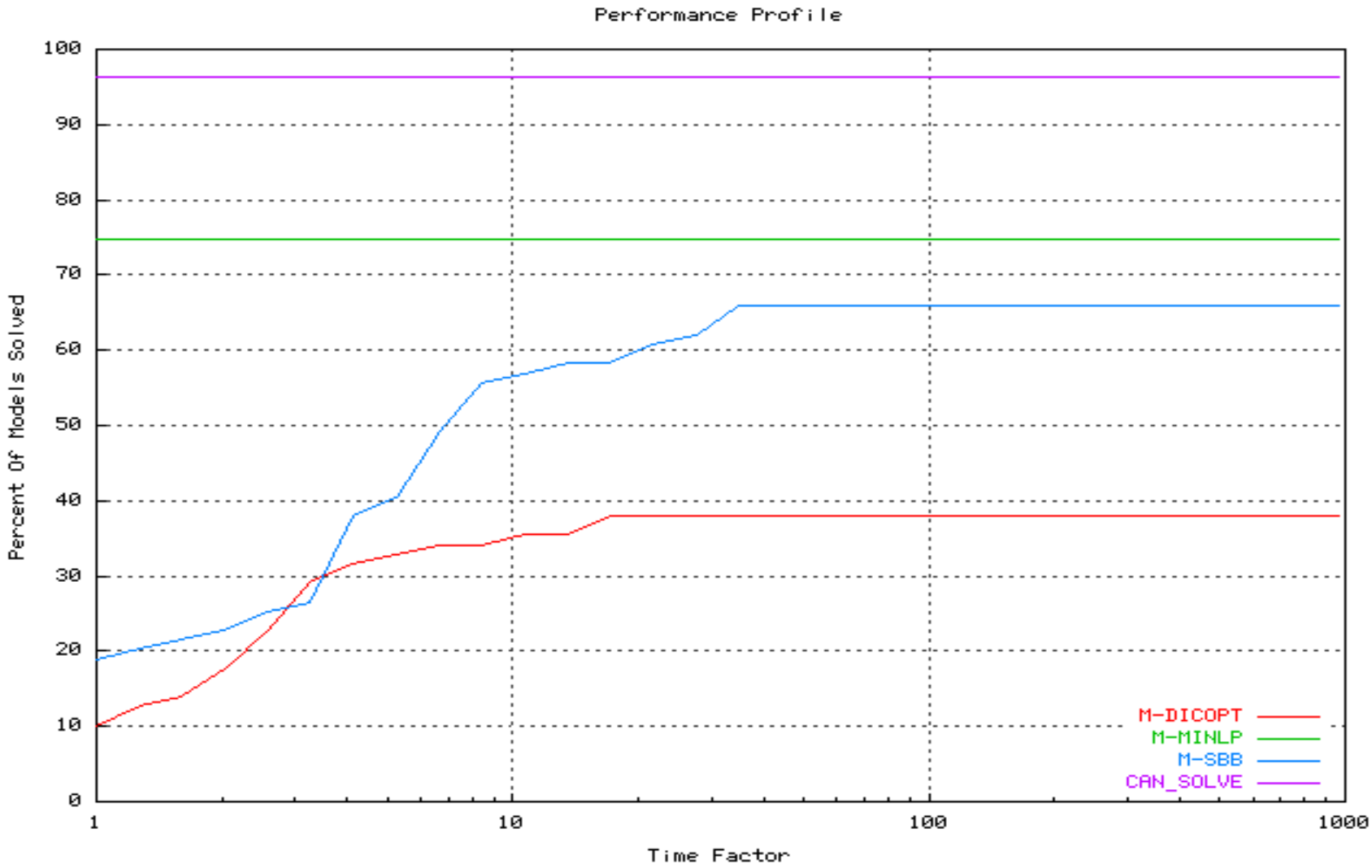
Choices:

- Models?
- Solver and solver options?

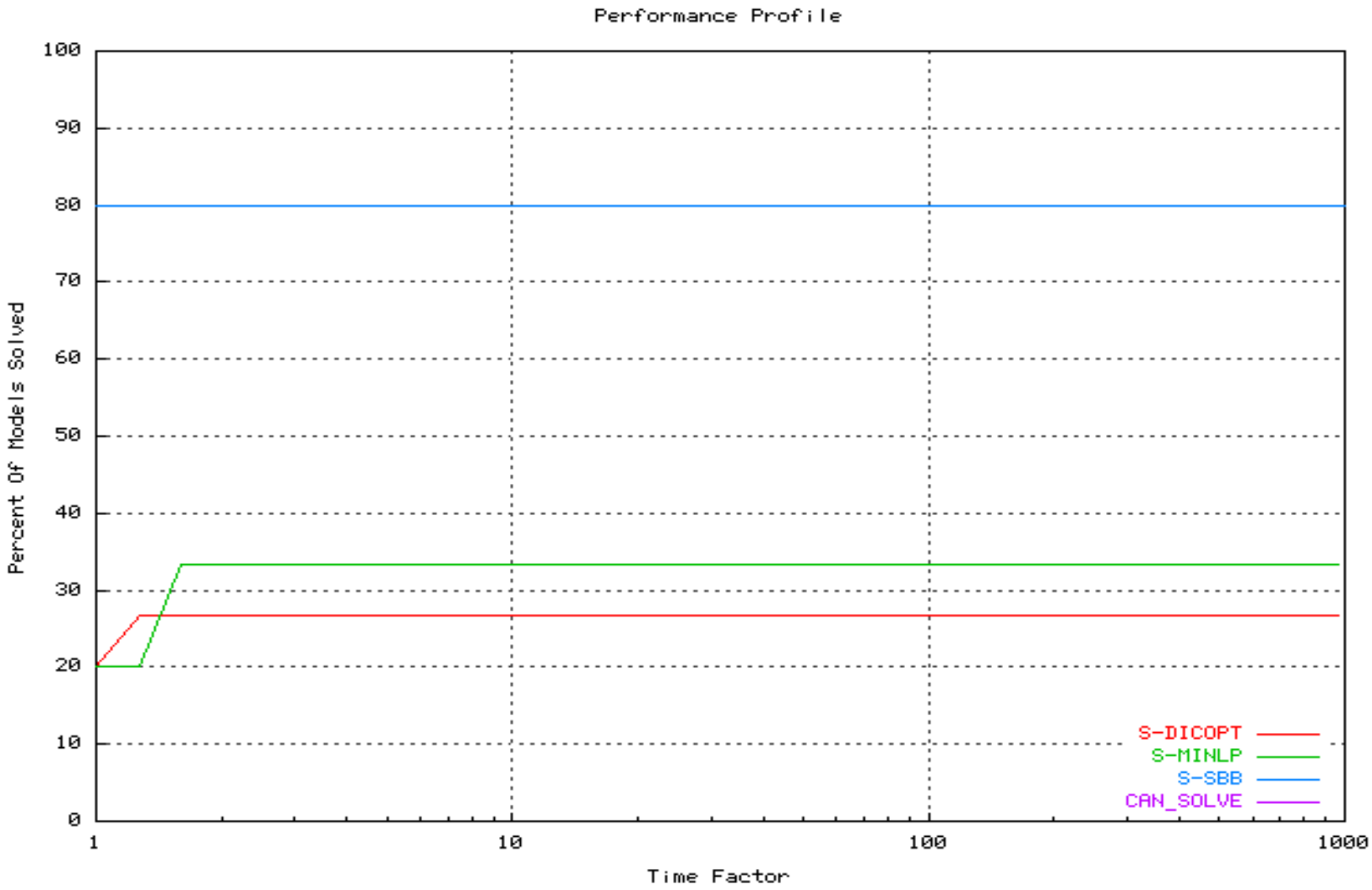
Example:

- Can choose a set of models where each solver is best:

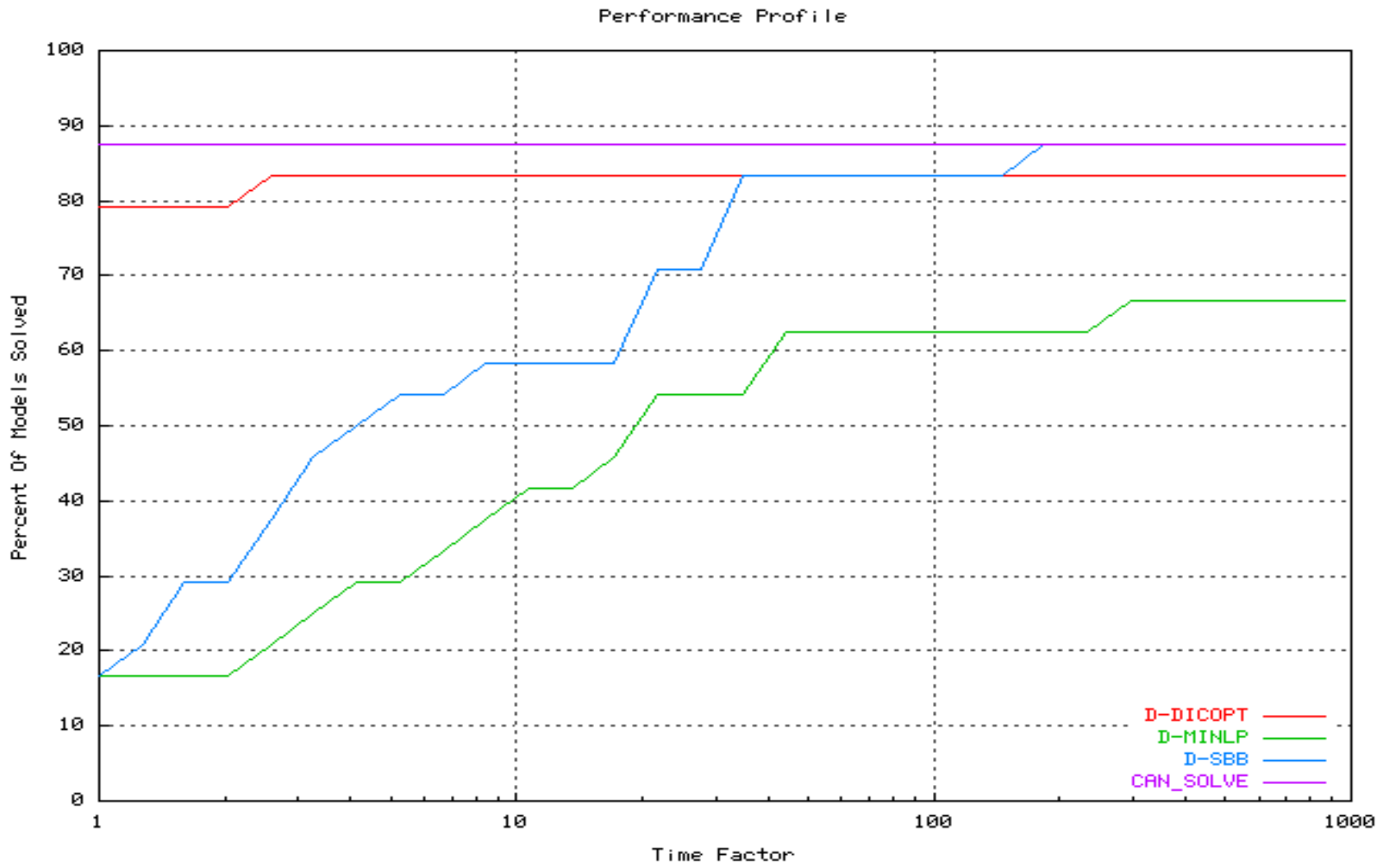
Profile: MINLP/Filter (best obj.)



Profile: SBB (best obj.)



Profile: DICOPT (best time)



Issues

Choice of models, solvers, options is subjective!

- Models used can skew data
  - several models of same structure may exist with different data
- Default or custom solver options
- Platform dependence:
  - Different resource times on different platforms

Other Issues

Timings:

- How is resource time measured (dependence on solver)?
- How are resource time limits enforced?
- Intermediate results if resource time limit reached

Conclusions

- Automation tools for collecting performance measurements
- Tools for analyzing and visualizing test results
  - Solver efficiency, robustness
  - Profiles and profile plots
- Enable users to reproduce performance results
- Automated performance analysis using the PAVER Server:

  www.gamsworld.org/performance/paver