Automated Performance Analysis in the Evaluation of Nonlinear Programming Solvers

Armin Pruessner
GAMS Development Corporation

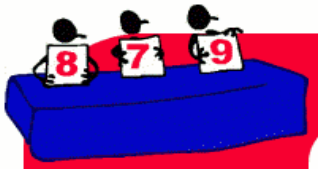
Hans Mittelmann
Arizona State University

ISMP - Copenhagen
August 18-22, 2003

Agenda

- PAVER Server (for Analysis of Benchmark Data)
  - Performance World, Motivation and Goals
  - Quality Assurance Framework
  - Description of Server
  - Performance Tools and Metrics
- Illustrative Examples
  - COPS NLP models
  - Benchmark Results
- 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

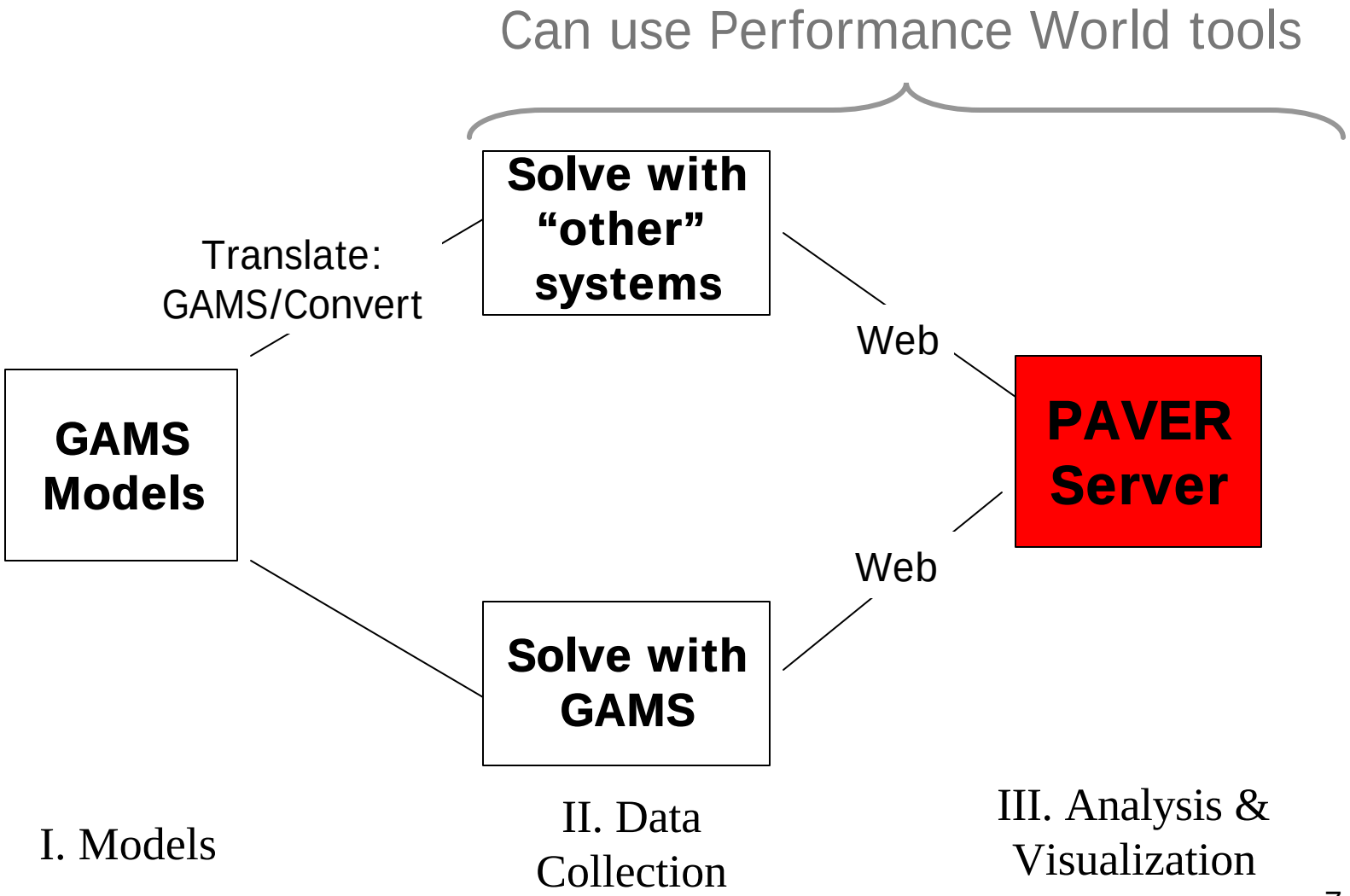
Performance
Analysis &
Visualization

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

Quality Assurance Framework



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 analysis/visualization of data
- Results sent via e-mail in HTML format
- Rely on 3 tools: solver square, resource time, performance profiles

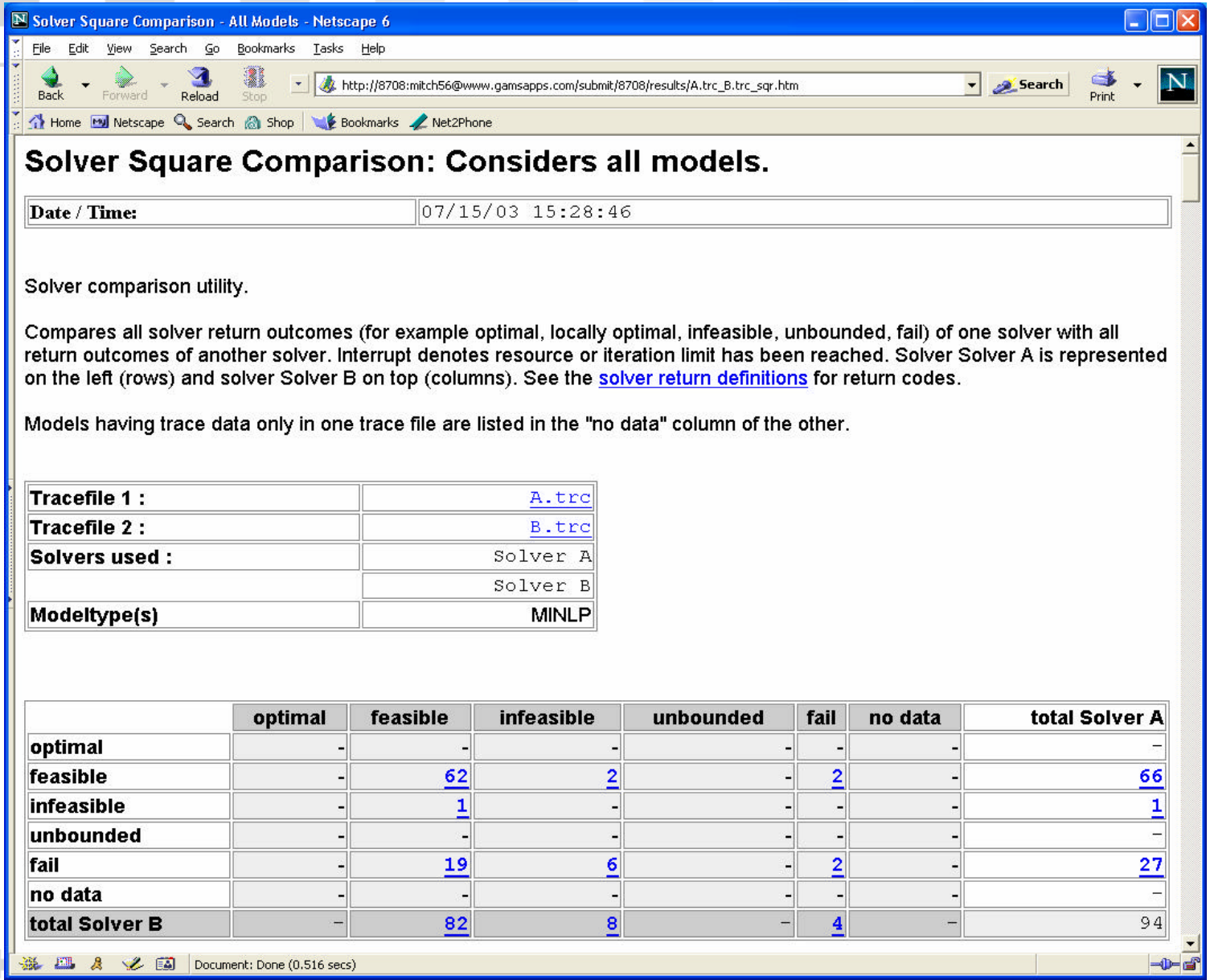
Tools: Robustness

Solver Square Utility:

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

→ Can use online using PAVER

PAVER: Solver Square



Solver Square Comparison: Considers all models.

Date / Time:	07/15/03 15:28:46

Solver comparison utility.

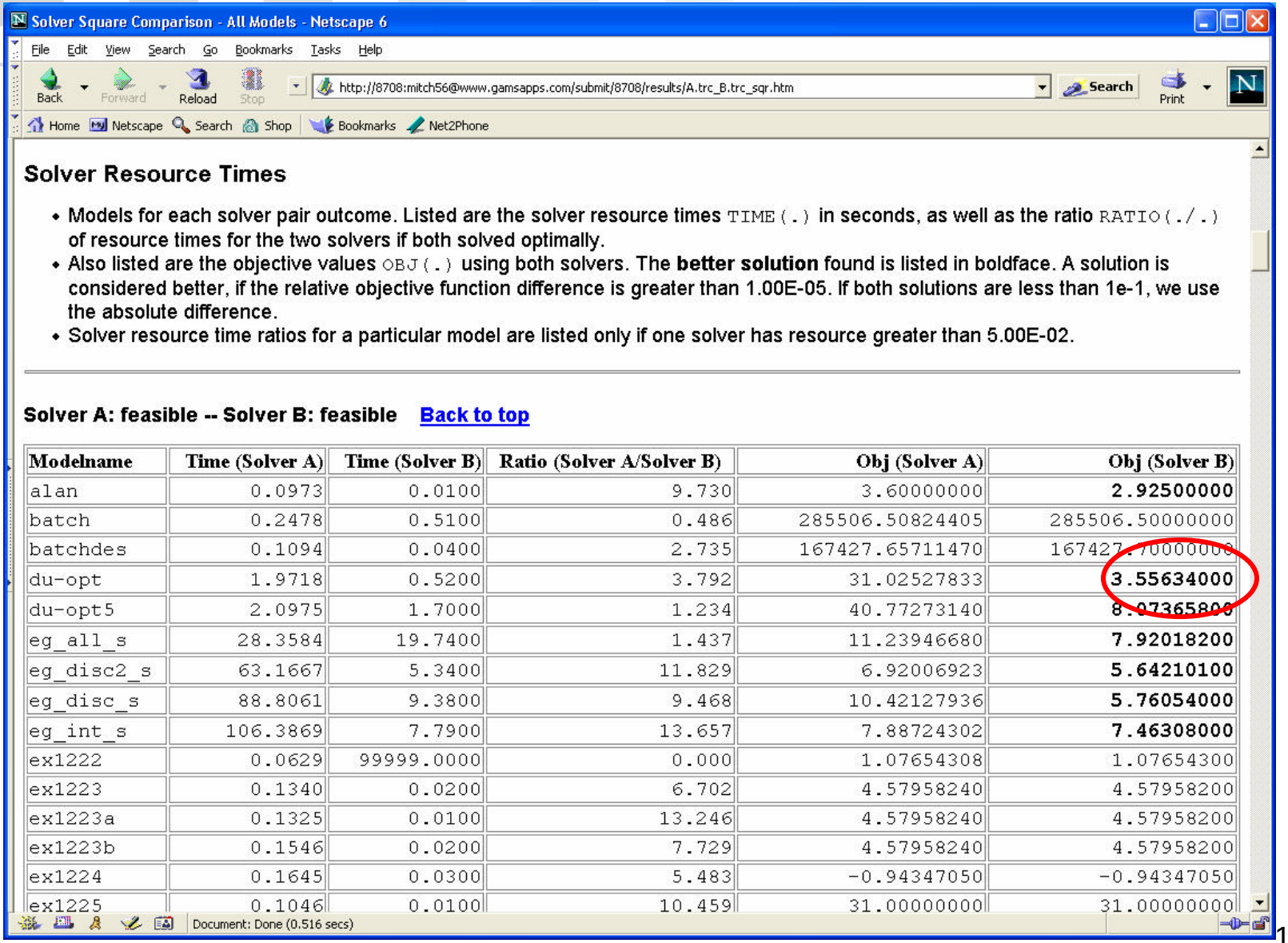
Compares all solver return outcomes (for example optimal, locally optimal, infeasible, unbounded, fail) of one solver with all return outcomes of another solver. Interrupt denotes resource or iteration limit has been reached. Solver Solver A is represented on the left (rows) and solver Solver B on top (columns). See the solver return definitions for return codes.

Models having trace data only in one trace file are listed in the "no data" column of the other.

Tracefile 1 :	A.trc
Tracefile 2 :	B.trc
Solvers used :	Solver A
	Solver B
Modeltype(s)	MINLP

	optimal	feasible	infeasible	unbounded	fail	no data	total Solver A
optimal	-	-	-	-	-	-	-
feasible	-	62	2	-	2	-	66
infeasible	-	1	-	-	-	-	1
unbounded	-	-	-	-	-	-	-
fail	-	19	6	-	2	-	27
no data	-	-	-	-	-	-	-
total Solver B	-	82	8	-	4	-	94

PAVER: Square (cont.)



Solver Resource Times

- Models for each solver pair outcome. Listed are the solver resource times `TIME(.)` in seconds, as well as the ratio `RATIO(./.)` of resource times for the two solvers if both solved optimally.
- Also listed are the objective values `OBJ(.)` using both solvers. The **better solution** found is listed in boldface. A solution is considered better, if the relative objective function difference is greater than 1.00E-05. If both solutions are less than 1e-1, we use the absolute difference.
- Solver resource time ratios for a particular model are listed only if one solver has resource greater than 5.00E-02.

Solver A: feasible -- Solver B: feasible Back to top

Modelname	Time (Solver A)	Time (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**
ex1222	0.0629	99999.0000	0.000	1.07654308	1.07654300
ex1223	0.1340	0.0200	6.702	4.57958240	4.57958200
ex1223a	0.1325	0.0100	13.246	4.57958240	4.57958200
ex1223b	0.1546	0.0200	7.729	4.57958240	4.57958200
ex1224	0.1645	0.0300	5.483	-0.94347050	-0.94347050
ex1225	0.1046	0.0100	10.459	31.00000000	31.00000000

Document: Done (0.516 secs)

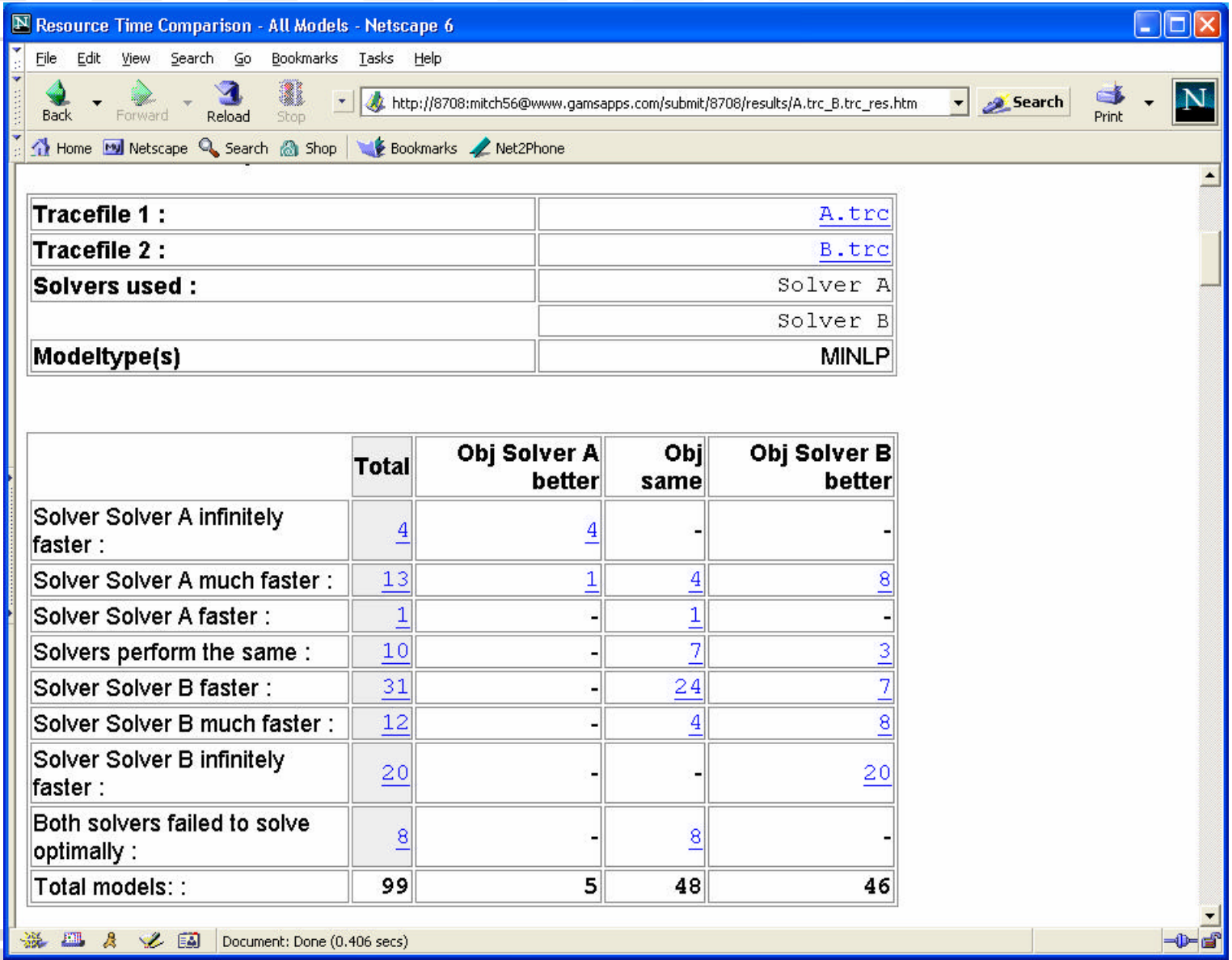
Tools: Efficiency

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

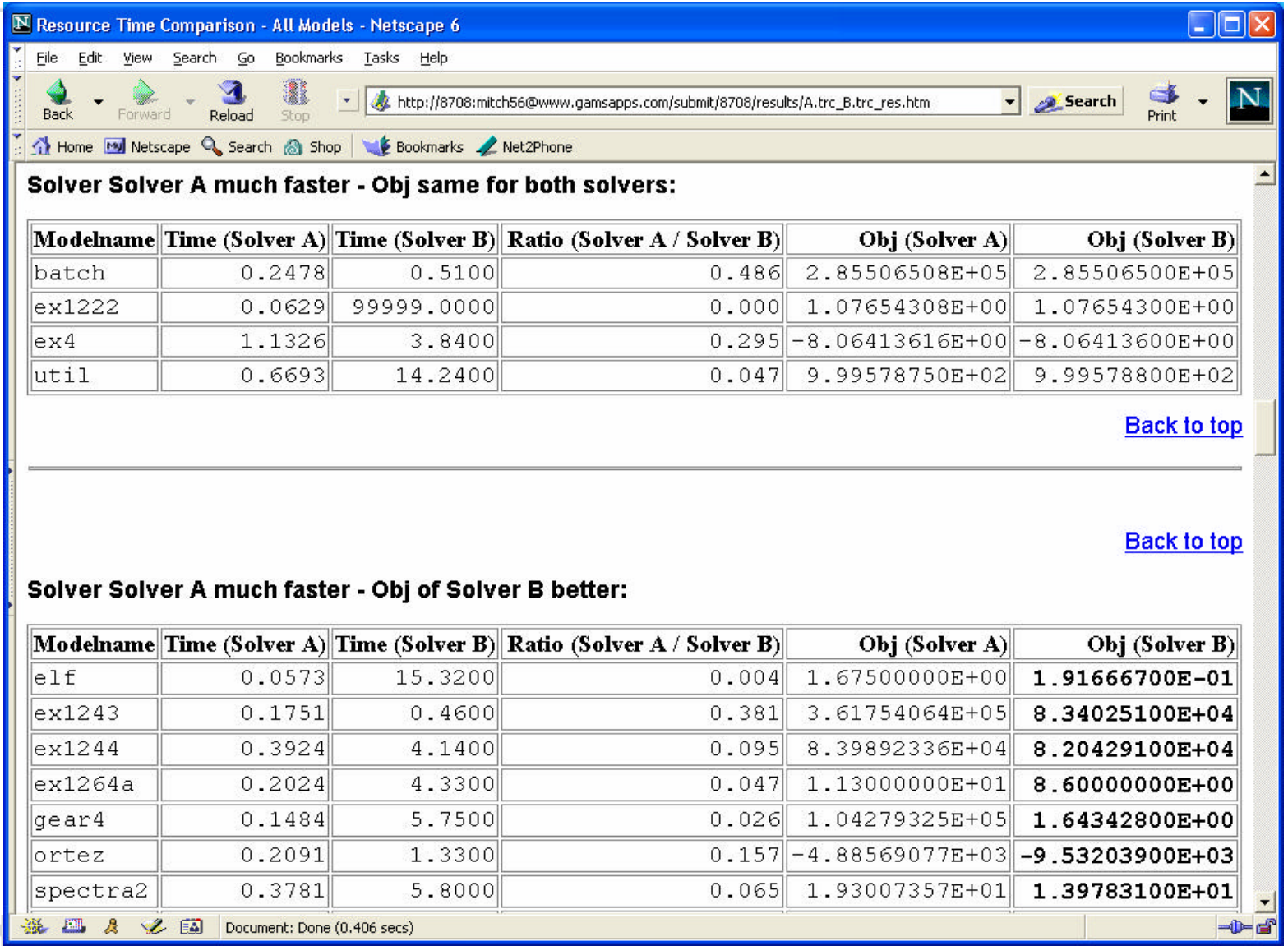
PAVER: Solver Resource Time



Tracefile 1 :	A.trc
Tracefile 2 :	B.trc
Solvers used :	Solver A
	Solver B
Modeltype(s)	MINLP

	Total	Obj Solver A better	Obj same	Obj Solver B better
Solver Solver A infinitely 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	-	24	7
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	48	46

PAVER: Resource Time (cont.)



Modelname	Time (Solver A)	Time (Solver B)	Ratio (Solver A / Solver B)	Obj (Solver A)	Obj (Solver B)
batch	0.2478	0.5100	0.486	2.85506508E+05	2.85506500E+05
ex1222	0.0629	99999.0000	0.000	1.07654308E+00	1.07654300E+00
ex4	1.1326	3.8400	0.295	-8.06413616E+00	-8.06413600E+00
util	0.6693	14.2400	0.047	9.99578750E+02	9.99578800E+02

Modelname	Time (Solver A)	Time (Solver B)	Ratio (Solver A / Solver B)	Obj (Solver A)	Obj (Solver B)
elf	0.0573	15.3200	0.004	1.67500000E+00	**1.91666700E-01**
ex1243	0.1751	0.4600	0.381	3.61754064E+05	**8.34025100E+04**
ex1244	0.3924	4.1400	0.095	8.39892336E+04	**8.20429100E+04**
ex1264a	0.2024	4.3300	0.047	1.13000000E+01	**8.60000000E+00**
gear4	0.1484	5.7500	0.026	1.04279325E+05	**1.64342800E+00**
ortez	0.2091	1.3300	0.157	-4.88569077E+03	**-9.53203900E+03**
spectra2	0.3781	5.8000	0.065	1.93007357E+01	**1.39783100E+01**

Tools: Visualization

Performance Profiles (Dolan and Moré, 2002):

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

Tools: Performance Profiles

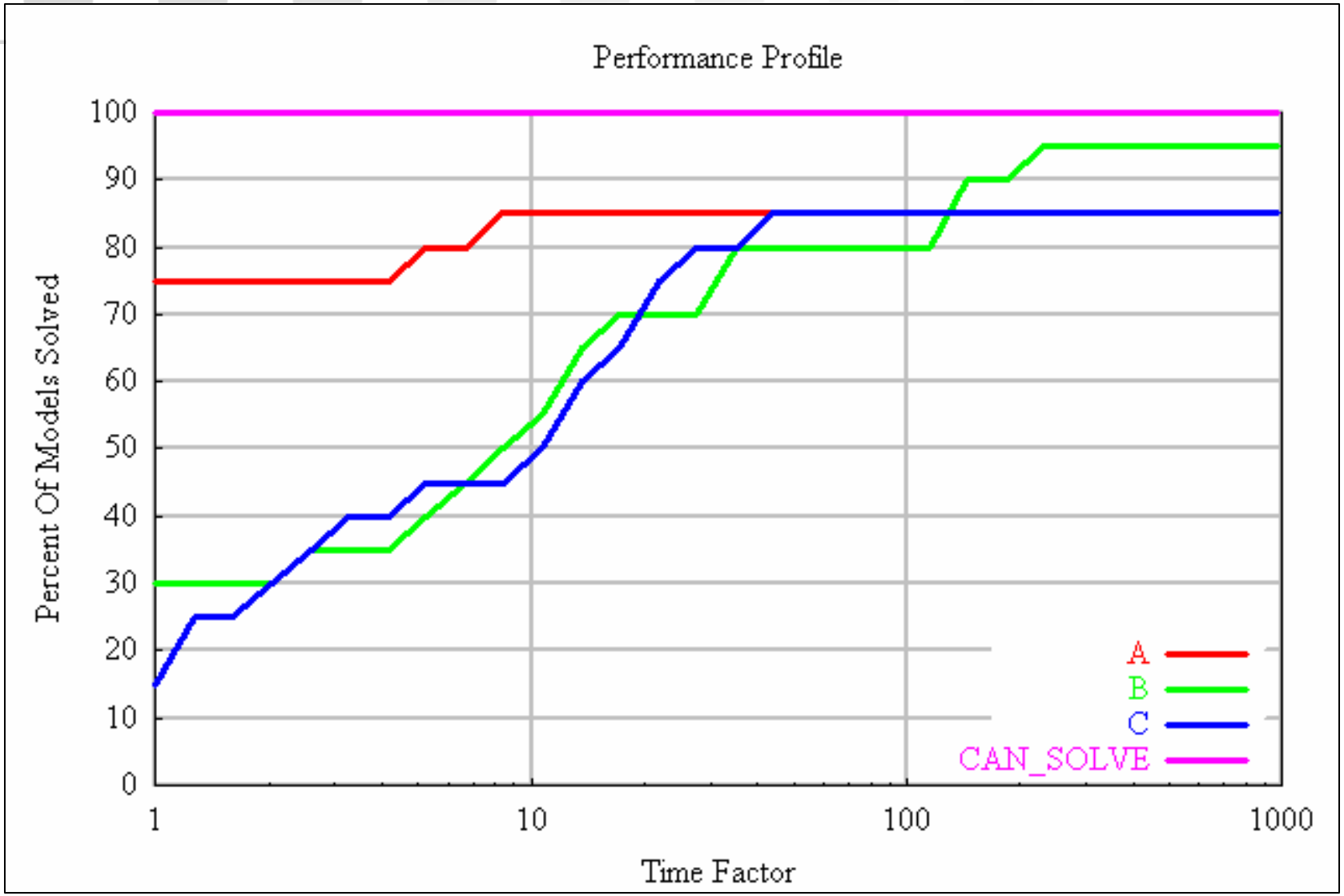
Interpretation (for t=*ratio*, P=*profile*):

- Efficiency: $P(t)$ for $t = 1$
- Probability of success:

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

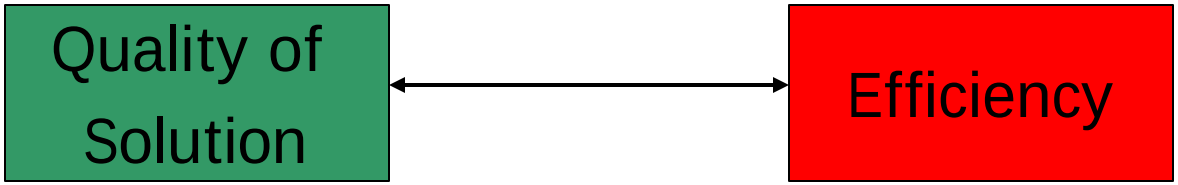
- Compact graphs summarize all information

Profiles (best resource time)



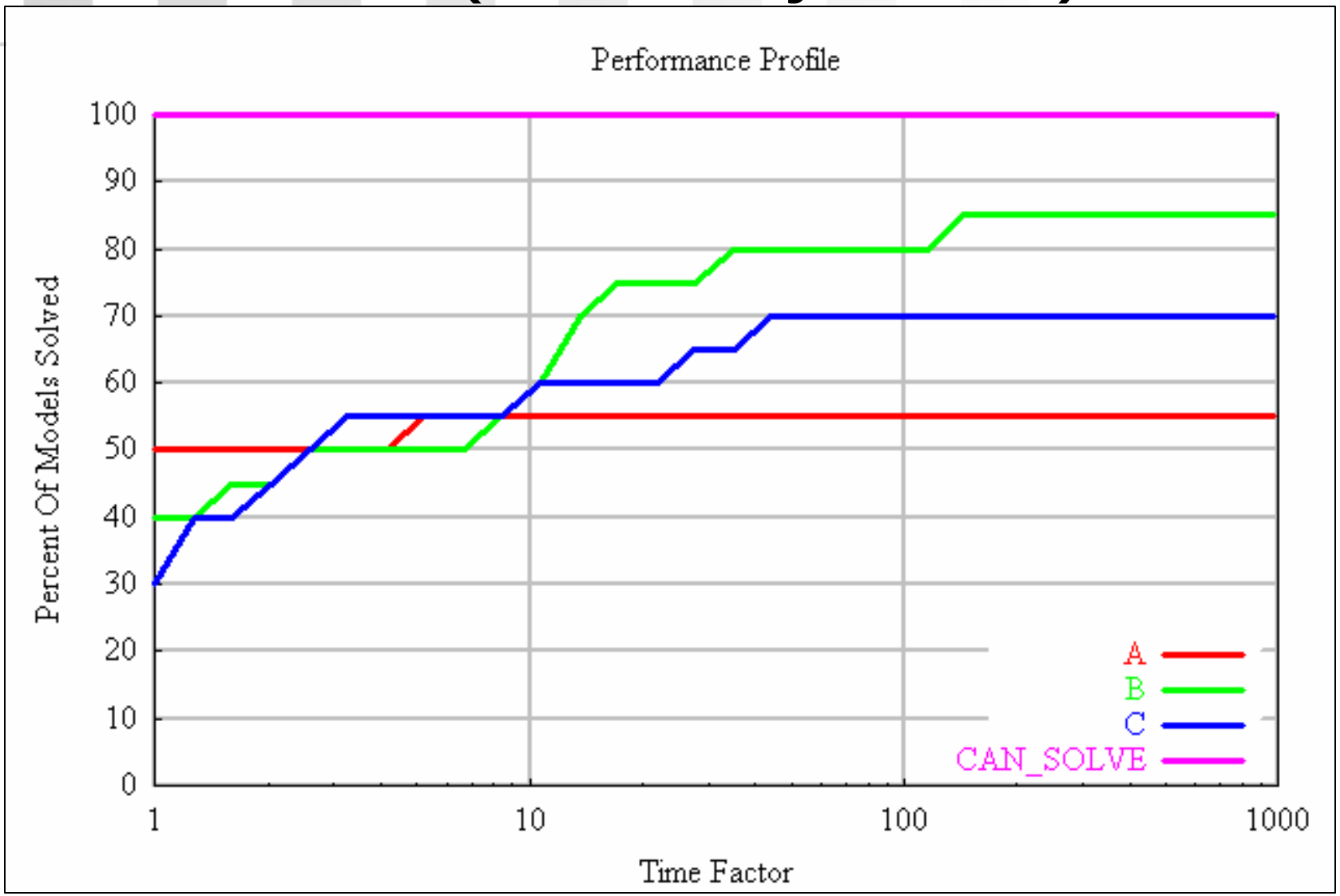
Tools: Visualization

Performance Profiles: considers both



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

Profiles (best objective)



Overview of Tools

Primary performance metric employed by each tool:

Metric / Tool	Robustness	Efficiency	Solution Quality
Square	x		
Resource Time		x	x
Profiles	x	x	x

Other Uses of PAVER

- Quality assurance
  - Integral part of GAMS quality control
- Fine tuning a new solver version
  - Comparison to previous versions
  - Want new version to be better w.r.t. various performance metrics
- Determining default solver options
  - Run solver using different options and compare

Benchmarking process

Two components:

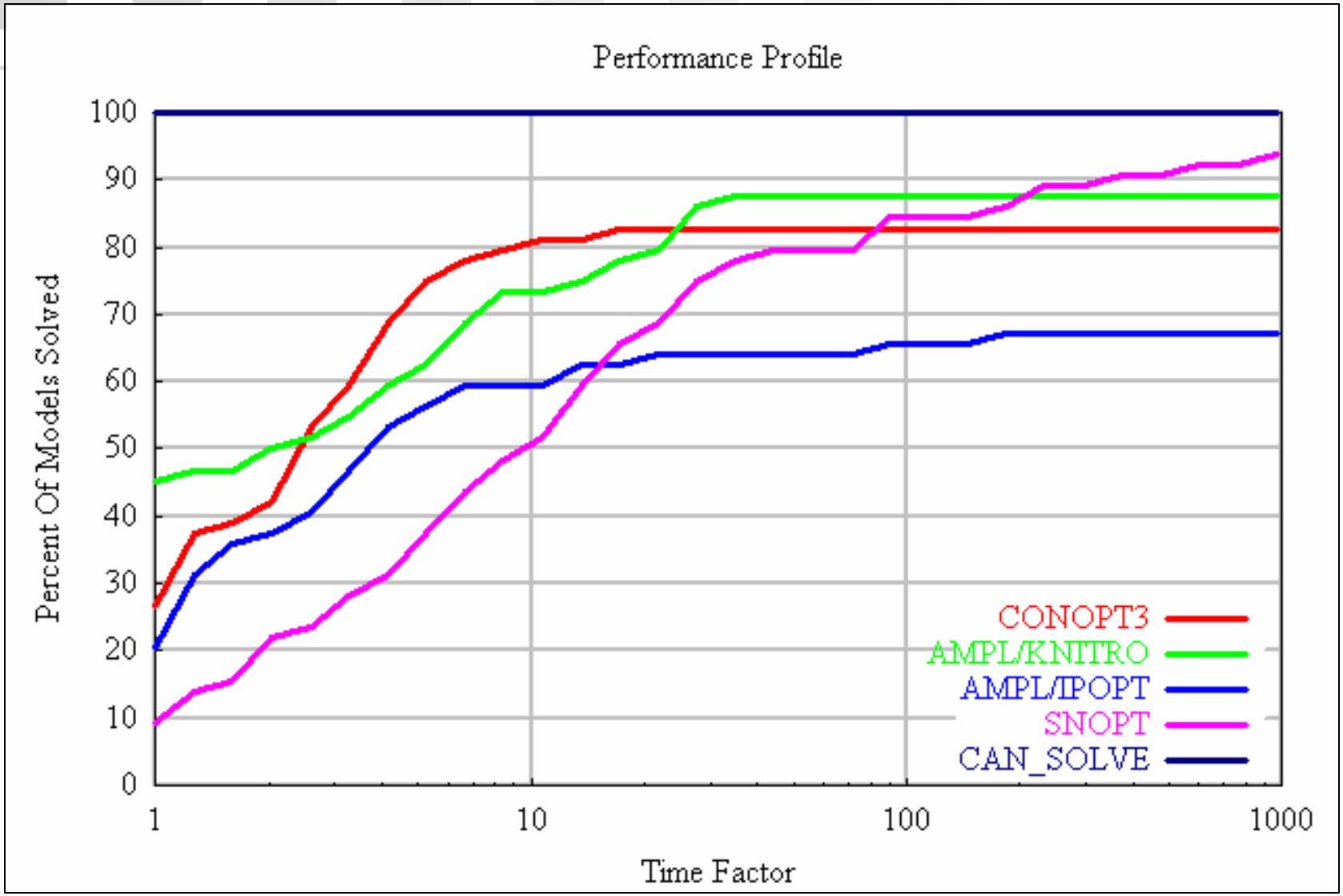
- Subjective Component
  - Choice of models
  - Choice of solvers
  - Choice of solver options
- Non-subjective component:
  - Obtaining performance data
  - Performance analysis and visualization

→ reproducible

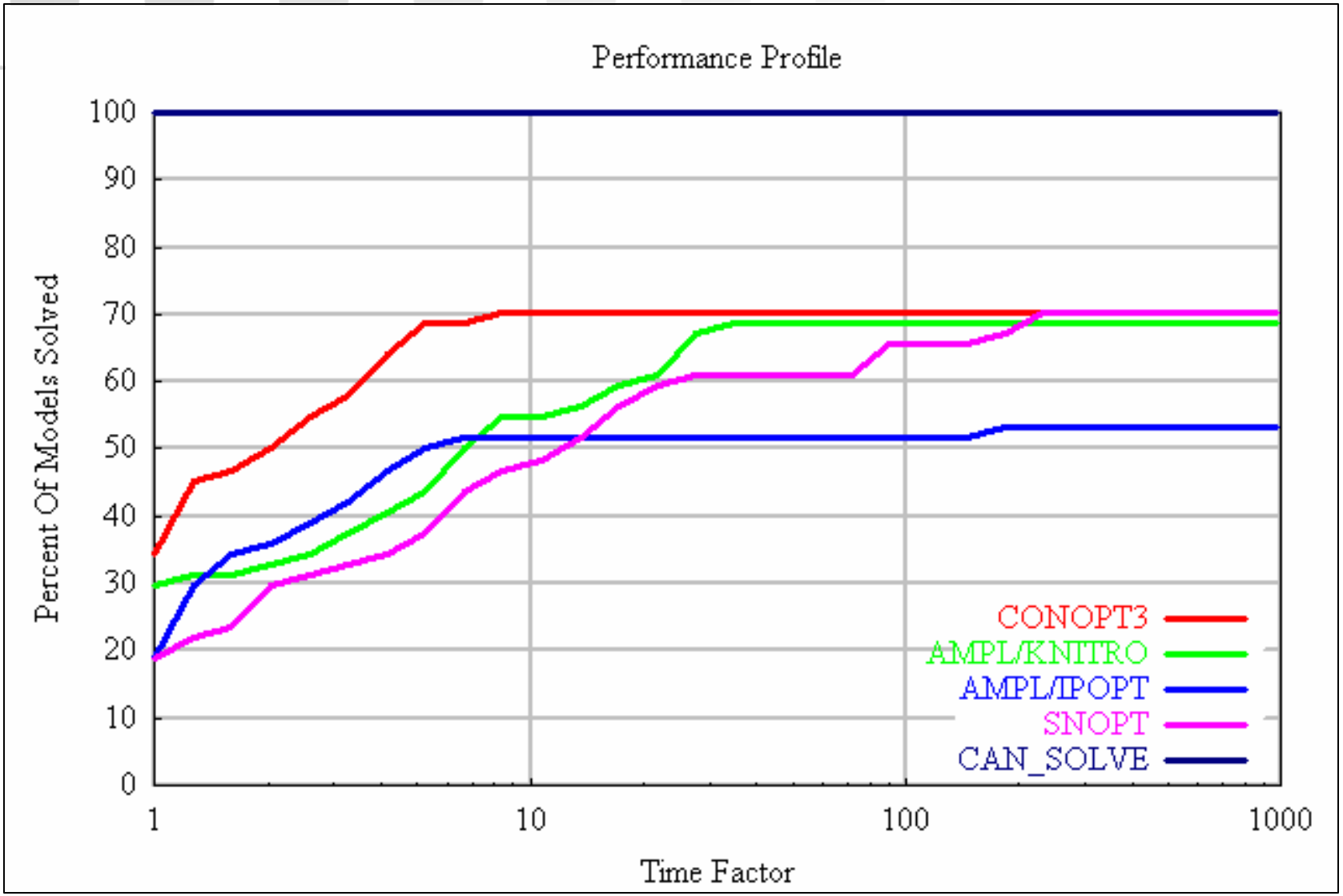
Illustrative Example: COPS

- Choose well-known COPS test set for NLP
  - Large scale nonlinear optimization test set
  - Models come from various application areas (fluid dynamics, population dynamics, optimal design, optimal control)
- Instance comes from GLOBALlib
  - Choose 16 models/4 sizes each → Total of 64 models
- Solvers used
  - AMPL: KNITRO (version 2.1 08/01/02) IPOPT (version 2.0.1)
  - GAMS: CONOPT3 (Library301F) SNOPT (version 5.3-5(2))
- Experimental Conditions:
  - Dual processor 450MHz, 2Gb memory, Sun/Solaris 8
  - 3600 sec resource time limit (CPU)

Results: Profiles (Efficiency)



Results: Profiles (Solution Quality)



Subjectivity in Benchmarking

Performance Tools:

- Partly takes care of non-subjective component in data analysis phase

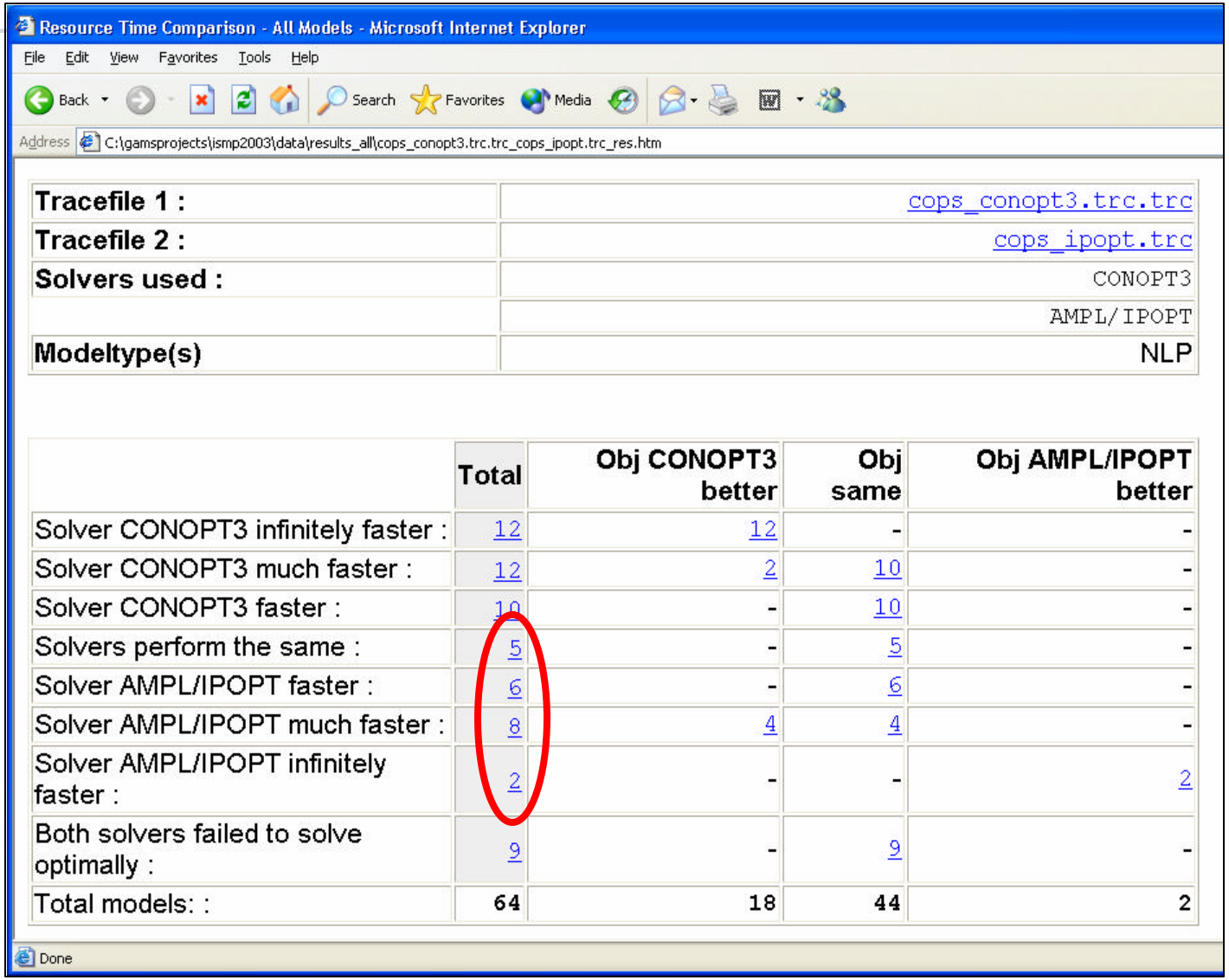
- Choices during data collection phase:
  - Models?
  - Solver and solver options?
  - Uniform stopping criteria?

Subjectivity in Model Selection

Choice of models, solvers, options is subjective!

- Models used can skew data
  - several models of same structure may exist with different data
  - can choose subset of models where each solver is superior!
  - use resource time comparison to choose subset where IPOPT is most efficient!

Results: IPOPT Best (Efficiency)

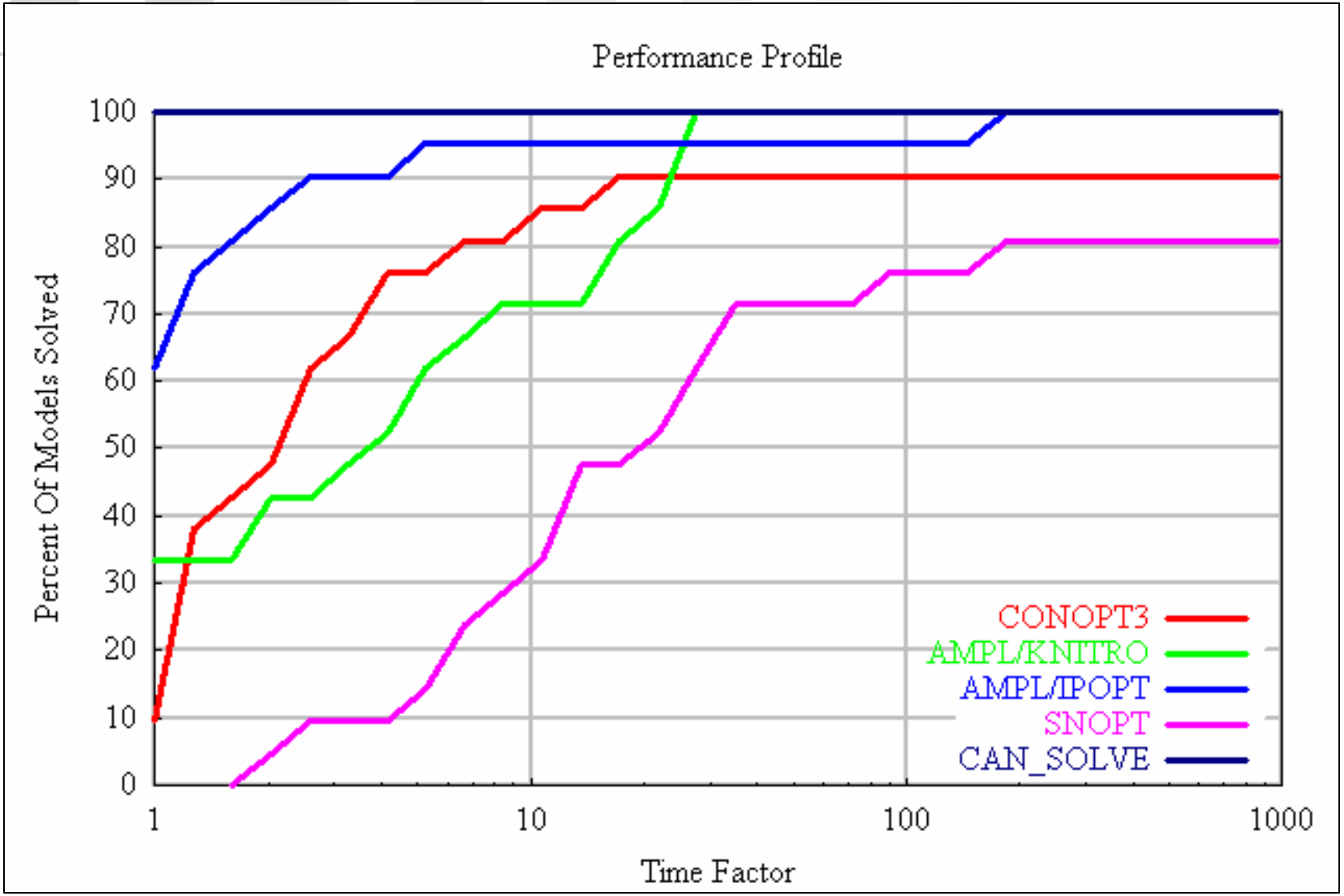


Tracefile 1 :	cops_conopt3.trc.trc
Tracefile 2 :	cops_ipopt.trc
Solvers used :	CONOPT3
	AMPL/IPOPT
Modeltype(s)	NLP

	Total	Obj CONOPT3 better	Obj same	Obj AMPL/IPOPT better
Solver CONOPT3 infinitely faster :	12	12	-	-
Solver CONOPT3 much faster :	12	2	10	-
Solver CONOPT3 faster :	10	-	10	-
Solvers perform the same :	5	-	5	-
Solver AMPL/IPOPT faster :	6	-	6	-
Solver AMPL/IPOPT much faster :	8	4	4	-
Solver AMPL/IPOPT infinitely faster :	2	-	-	2
Both solvers failed to solve optimally :	9	-	9	-
Total models: :	64	18	44	2

- Choose subset of models where IPOPT is fastest!
- Use the resource time comparison table!

Results: IPOPT Best (Efficiency)



Stopping criteria

Comparisons of solutions:

- Independent verification of feasibility and optimality
- Issues:
  - Optimality criterion is measured differently by solvers
    - tweak solvers until tolerance is reached
  - How do we choose the proper solver options to satisfy independent verification tolerances?
    - one solver may work unduly hard to reach tolerance

Solution Verification: Examiner

- Independent check of solution accuracy using GAMS/Examiner
- Check feasibility and optimality
  - Can check a subset of these
  - User can set tolerances used, etc.
- Runs in between GAMS and a solver
  - Acts as a silent observer

Conclusions

- Online 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

References

- E. D. Dolan and J. J. Moré (2000). Benchmarking optimization software with performance profiles, *Math Programming*, **91** (2), 201-213.
- Examiner (2003), *GAMS: The Solver Manuals*. available online at http://www.gams.com/solvers/allsolvers.pdf
- H. D. Mittelmann and A. Pruessner (2003). A server for automated performance analysis of benchmarking data, *working paper*.
- This presentation is available at http://www.gams.com/presentations