THE SAT'07 CONTEST
THE FINAL RESULTS!

Daniel le Berre, Olivier Roussel, Laurent Simon
{leberre,roussel}@cril.univ-artois.fr, simon@lri.fr

May, 2007
The SAT conference

Lisboa, Portugal

Road Map

1. History, rules and Goals
2. The participants
3. The First Stage
4. The Winners
5. Conclusion and next contests
6. Certified UNSAT Special Track
7. And Inverter Graph Special Track

My Solver Did it? Did what?

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps
- My Solver did

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps
- My Solver did spend my free-time looking for bugs

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps
- My Solver did spend my free-time looking for bugs
- My Solver did

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps
- My Solver did spend my free-time looking for bugs
- My Solver did travel by email

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps
- My Solver did spend my free-time looking for bugs
- My Solver did travel by email
- My Solver did

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps
- My Solver did spend my free-time looking for bugs
- My Solver did travel by email
- My Solver did contribute to the global warming

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps
- My Solver did spend my free-time looking for bugs
- My Solver did travel by email
- My Solver did contribute to the global warming
- My Solver did

My Solver Did it? Did what?

Yes, your solvers must have done it...

- My Solver win the contest
- My Solver did compile (at least once)
- My Solver did core dumps
- My Solver did spend my free-time looking for bugs
- My Solver did travel by email
- My Solver did contribute to the global warming
- My Solver did something sometime...

Road Map

1. History, rules and Goals
   - Rules
2. The participants
3. The First Stage
4. The Winners
5. Conclusion and next contests
6. Certified UNSAT Special Track
7. And Inverter Graph Special Track

The SAT-"007" agents

Three judges that have licences to kill... solvers

Ewald Speckenmeyer Institut für Informatik, Köln

Geoff Sutcliff Department of Computer Science, University of Miami

Lintao Zhang Microsoft Research, Silicon Valley

The SAT-"007" agents

Three judges that have licences to kill... solvers

Ewald Speckenmeyer Institut für Informatik, Köln

Geoff Sutcliff Department of Computer Science, University of Miami

Lintao Zhang Microsoft Research, Silicon Valley

All the decisions were taken in agreement with the judges.

The special tracks

Certified UNSAT solvers must output a certificate of unsatisfiability. Proof format / checker provided by Allen van Gelder. See Allen's poster, wait for his presentation

And-Inverter Graphs input Using Armin Biere's And-Inverter Graphs as input format (AIGER). May be used for compact encoding of circuits. A 2-slides results will be given.

Divide (prices) to conquer (glory)

Two main divisions: Competition and Demonstration

Competition the source code of the solver must be available after the competition.

Demonstration a binary version of the solver must be available for research purpose.

Divide (prices) to conquer (glory)

Two main divisions: Competition and Demonstration

Competition the source code of the solver must be available after the competition.

Demonstration a binary version of the solver must be available for research purpose.

Ideas

Participation to the competition must benefit to the community

- By providing source code, binary or benchmarks
- By supporting the conference and the competition

Divide (prices) to conquer (glory)

Two main divisions: Competition and Demonstration

Competition the source code of the solver must be available after the competition.

Demonstration a binary version of the solver must be available for research purpose.

Ideas

Participation to the competition must benefit to the community

- By providing source code, binary or benchmarks
- By supporting the conference and the competition

Some “suspicious” solvers were moved to demonstration during the contest. Results of fixed versions of those solvers are available as well in the demonstration division, first stage.

The scoring scheme

Benchmark purse to be divided equally among the solvers able to solve it.

Speed purse to be divided unequally among the solvers able to solve a given benchmark.

Series an extra credit is given for each series solved.

Solver his score is the sum of the credits obtained per benchmarks solved.

Who can win?

Only participants that wear Contest T-Shirt can win

- Three categories: industrial, crafted and random
- Three specialties: SAT, UNSAT and SAT+UNSAT
- Three medals: gold, silver and bronze

So we have 27 awards.

Invariants rules and details

- Only 3 solvers per submitter can enter the first stage, competition division.
- Only 1 solver per submitter can enter the second stage, competition division.

Road Map

1. History, rules and Goals
2. The participants
   - Solvers
   - Benchmarks
3. The First Stage
4. The Winners
5. Conclusion and next contests
6. Certified UNSAT Special Track
7. And Inverter Graph Special Track

Local Search Solvers

adaptg2wsat0, adaptg2wsatp	Chu-Min Li, Wanxia Wei and Harry Zhang	Decreasing Variable Scoring
adaptnovelty	Holger Hoos and Dave Tompkins	SAT 2004 winner
adaptg2wsat+, FH, Hybrid1,	Wanxia Wei, Chu-Min Li and Harry Zhang	Switching between 2 or 4 heuristics, auto-adaptative
gnovelty+	Duc Nghia Pham and Charles Gretton	
SAPS, SAPSRT	Dave A. D. Tompkins, Frank Hutter, and Holger H. Hoos	Dynamic Local Search
Ranov	Duc Nghia Pham and Anbulagan	

Complete, CDCL solvers

CMUSAT, CMUSAT-base	Himanshu Jain and Edmund Clarke	CDCL, Standard Template Lib.
Minisat-2007	Niklas Sorensson and Niklas Een	Who don't know Minisat?
Picosat	Armin Biere	
Rsat	Knot Pipatsrisawat and Adnan Darwiche	
MXC	David R. Bregman and David G. Mitchell	Best student solver (sat-race 06)
Mmisat	Monahov Ivan	
Tinisat, TiniSatELite	Jinbo Huang	pronounce "teeny sat"
SAT7	Christian Kern, Mohammad Khaleghi, Stefan Kugele, Christian Schallhart, Michael Tautschnig and Andreas Weis	

Complete, non-CDCL solvers

dewSatz-1a	Anbulagan	LA Saturation with restriction in satz
kcnfs-2004, kcnfs-2006, kcnfs-smp	Gilles Dequen and Olivier Dubois	
march-ks	Marijn Heule and Hans van Maaren	
tts-4.0	Ivor Spence	stands for “Ternary Tree Solver”
UnitMarch	Marijn Heule, Denis de Leeuw Duarte, and Hans van Maaren	multi-bit assignements

Other solvers

SATzilla-CRAFTED, SATzilla-RANDOM, SATzilla-ALL	Lin Xu, Frank Hutter, Holger H. Hoos and Kevin Leyton-Brown	Specialized Portfolio
MiraXTv1, MiraXTv2, MiraXTv3	Tobias Schubert, Matthew Lewis, Natalia Kalinnik and Bernd Becker	Multi-threaded solver

Demonstration Division

submitted in competition division

`barcelogic`	Robert Nieuwenhuis, Albert Oliveras and Tomas Lioret
`minimarch`	Siert Wieringa, Hans van Maaren and Marijn Heule
`dewSatz`	Anbulagan

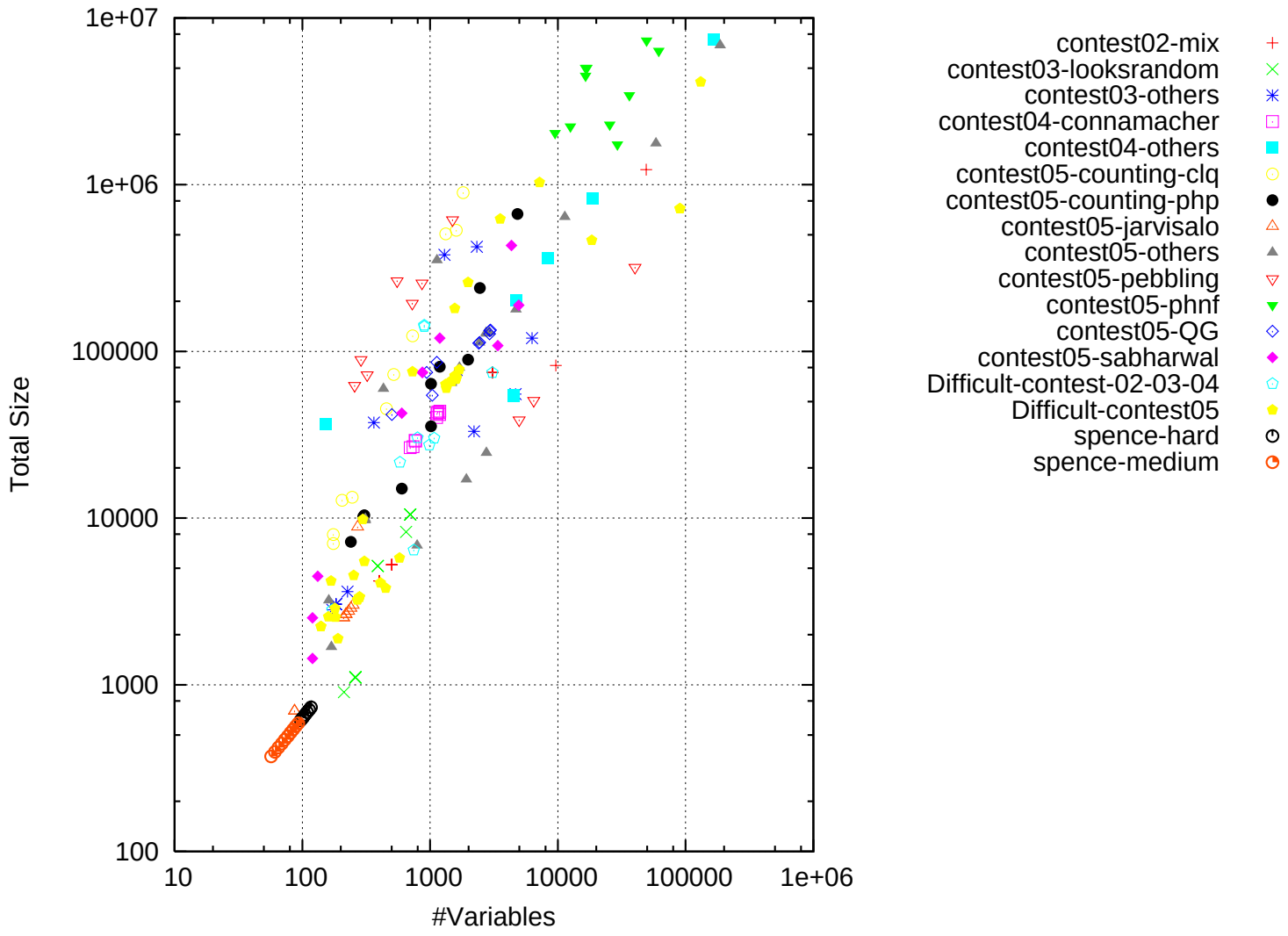
Demonstration Division

`adaptg2wsat`	Chu-Min Li, Wanxia Wei and Harry Zhang
`SAT4J-1.7`	Daniel Le Berre
`Spear, Spear-FH, Spear-FHS`	Domagoj Babic
`Ornithorynque`	Olivier Roussel

Crafted

Series	Subseries
contest02 (10)	
contest03	looksrandom (10), others (10)
contest04	connamacher (10), others (10)
contest05	counting-clq (10), counting-php (10), jarvisalo (10), others (20), pebbling (10), phnf (10), QG (10), sabharwal (10)
Difficult	contest-02-03-04 (9), contest05 (36)
spence	hard (6), medium (10)

A view of crafted benchmarks



Random

On Treshold benchmarks

5 SAT benchmarks, 5 UNSAT benchmarks per series

- 3SAT from 360 to 650 variables (7 series)
- 5SAT from 90 to 130 variables (7 series)
- 7SAT from 45 to 75 variables (7 series)

Random

On Treshold benchmarks

5 SAT benchmarks, 5 UNSAT benchmarks per series

- 3SAT from 360 to 650 variables (7 series)
- 5SAT from 90 to 130 variables (7 series)
- 7SAT from 45 to 75 variables (7 series)
- 2+0.7SAT from 3500 to 6500 variables (4 series)
- 2+p0.8SAT from 1295 to 2405 variables (4 series)
- 2+p0.9SAT from 630 to 1170 variables (4 series)

Random

On Treshold benchmarks

5 SAT benchmarks, 5 UNSAT benchmarks per series

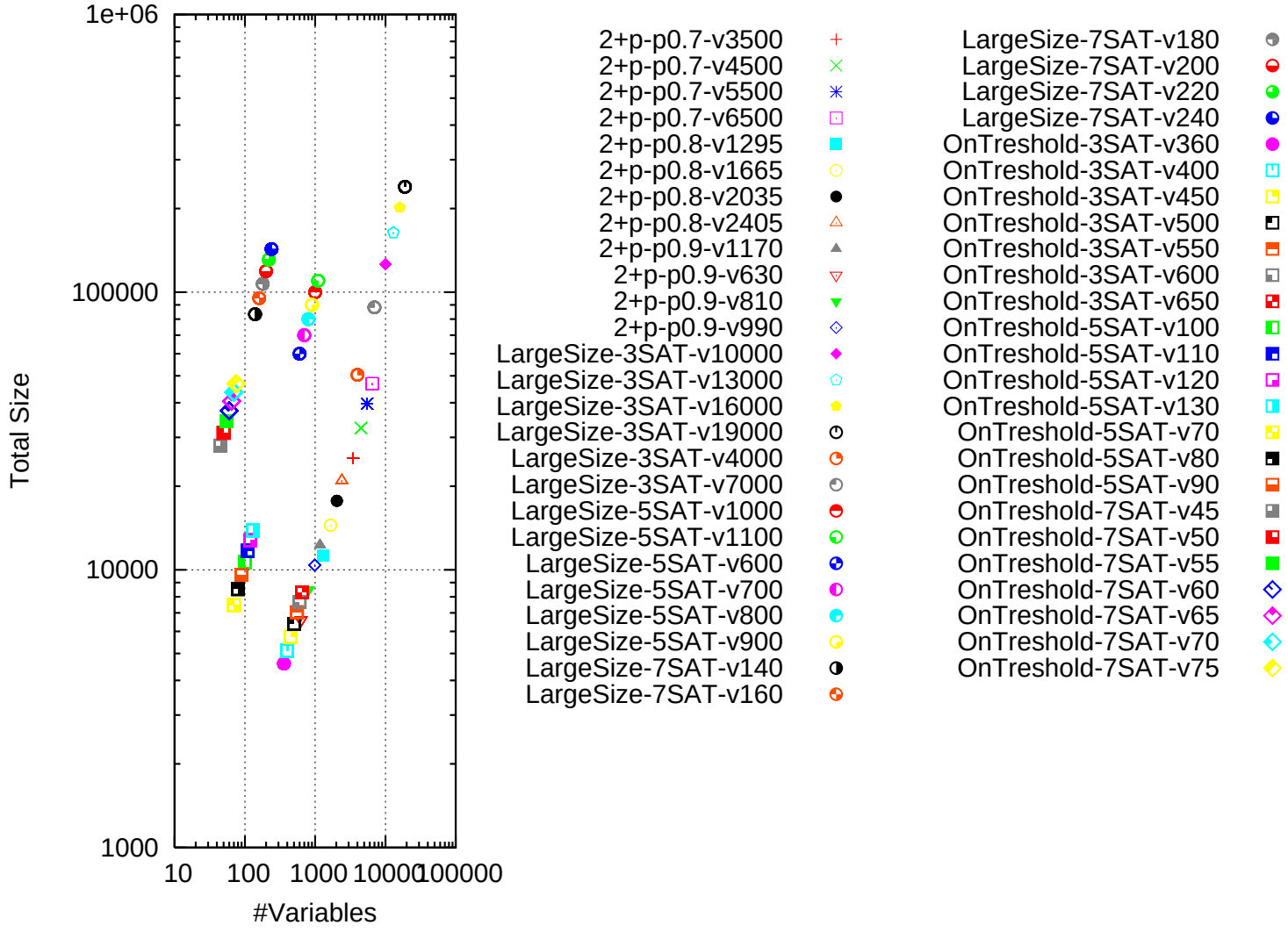
- 3SAT from 360 to 650 variables (7 series)
- 5SAT from 90 to 130 variables (7 series)
- 7SAT from 45 to 75 variables (7 series)
- 2+0.7SAT from 3500 to 6500 variables (4 series)
- 2+p0.8SAT from 1295 to 2405 variables (4 series)
- 2+p0.9SAT from 630 to 1170 variables (4 series)

Large benchmarks

: 10 (hopefully) SAT benchmarks per series

- 3SAT from 4000 to 19000 variables (6 series)
- 5SAT from 600 to 1100 variables (6 series)
- 7SAT from 140 to 240 variables (6 series)

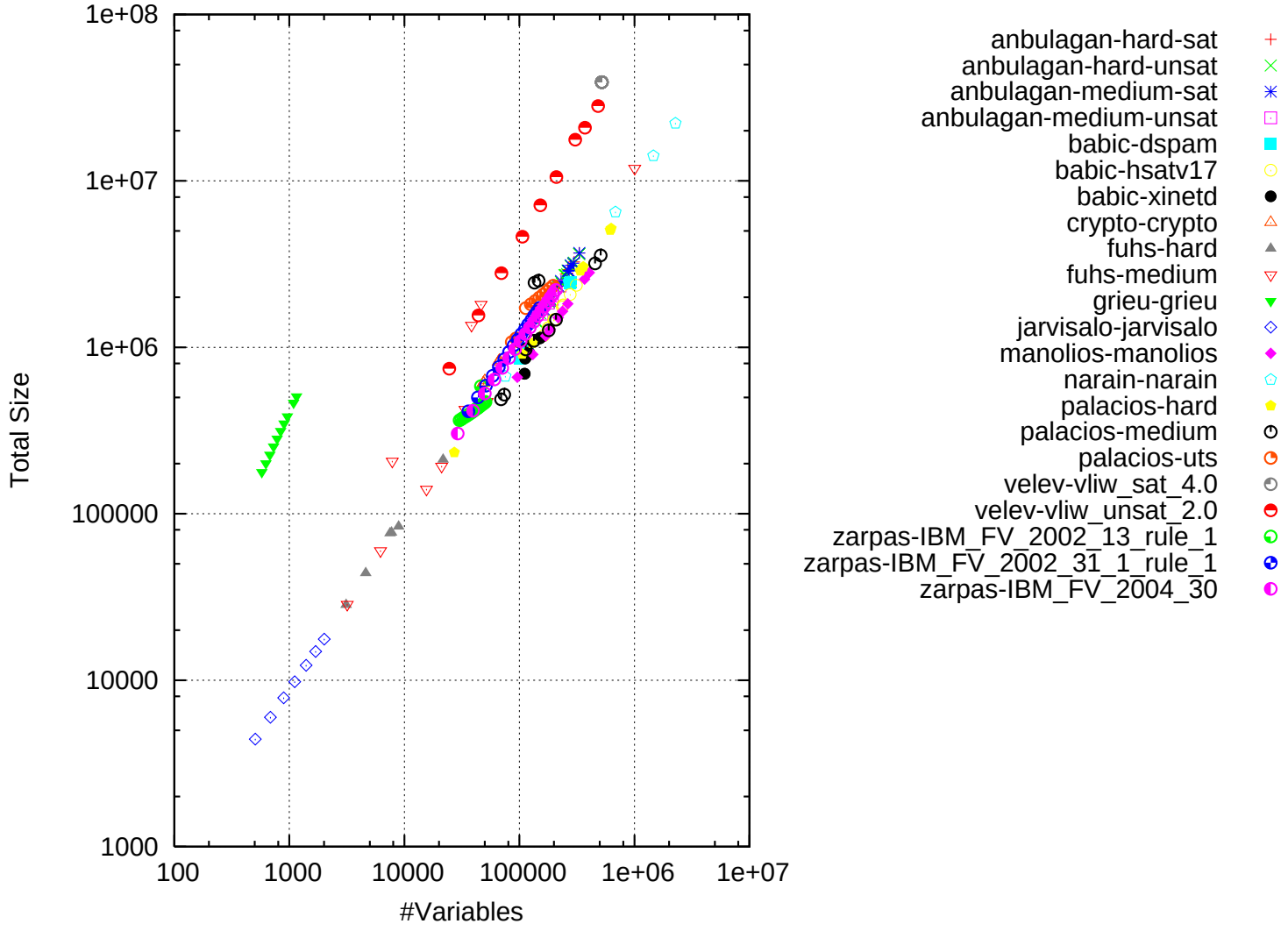
A view of random benchmarks



Industrial

Series	Subseries
anbulagan	hard-sat (7), hard-unsat (24), medium-sat (10), medium-unsat (6)
babic	dspam (10), hsatv17 (10), xinetd (10)
crypto (10)	
fuhs	hard (6), medium (10)
grieu (10)	
jarvisalo (7)	
manolios (10)	
narain (5)	
palacios	hard (7), medium (10), uts (10)
velev	vliw-sat-4.0 (10), vliw-unsat-4.0 (10)
zarpas	IBM-FV-2002-13-rule-1 (19), IBM-FV-2002-31-1-rule-1 (16), IBM-FV-2004-30 (18)

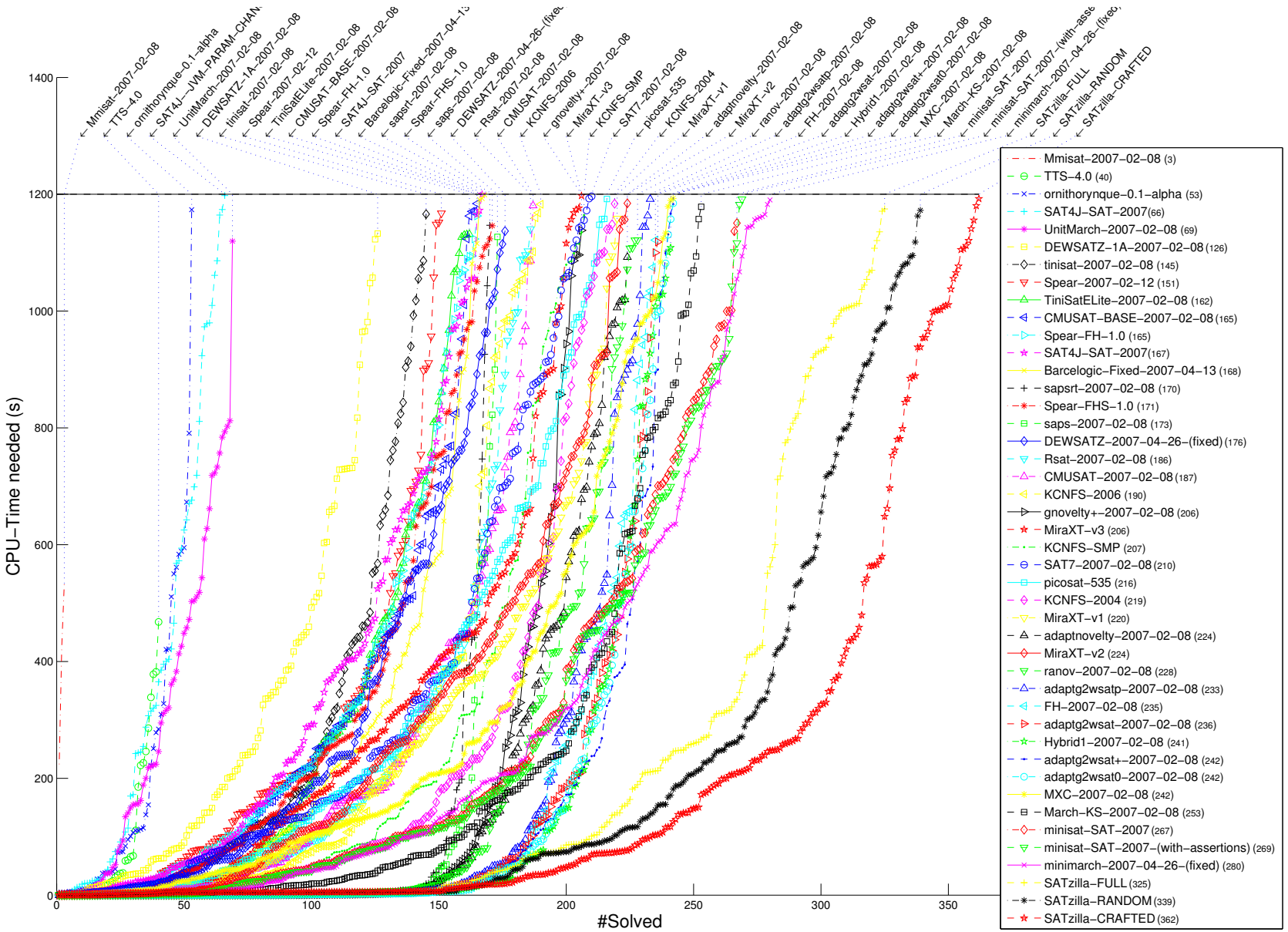
A view of industrial benchmarks



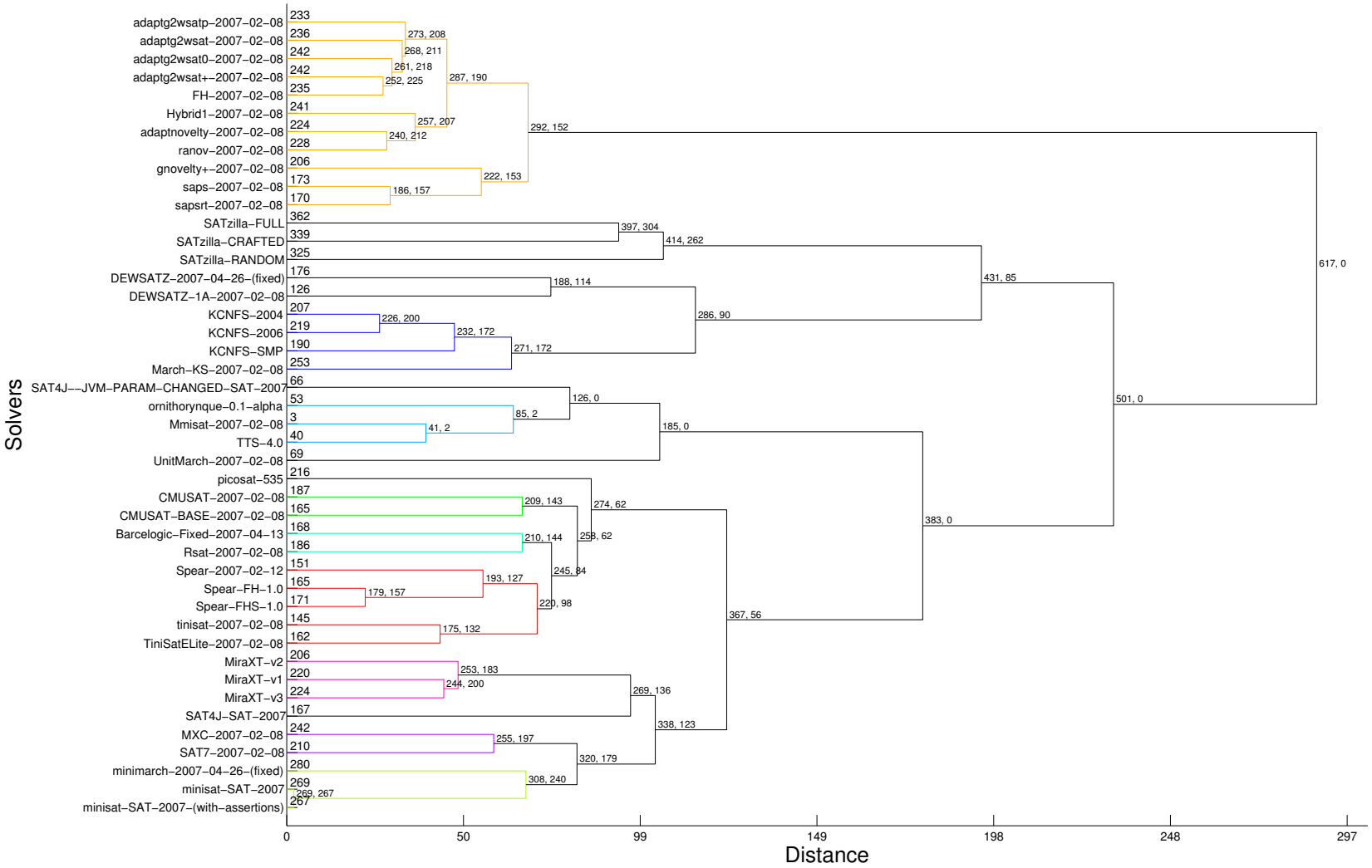
Road Map

1. History, rules and Goals
2. The participants
3. The First Stage
   - Overall Pictures
   - Remaining benchmarks
4. The Winners
5. Conclusion and next contests
6. Certified UNSAT Special Track
7. And Inverter Graph Special Track

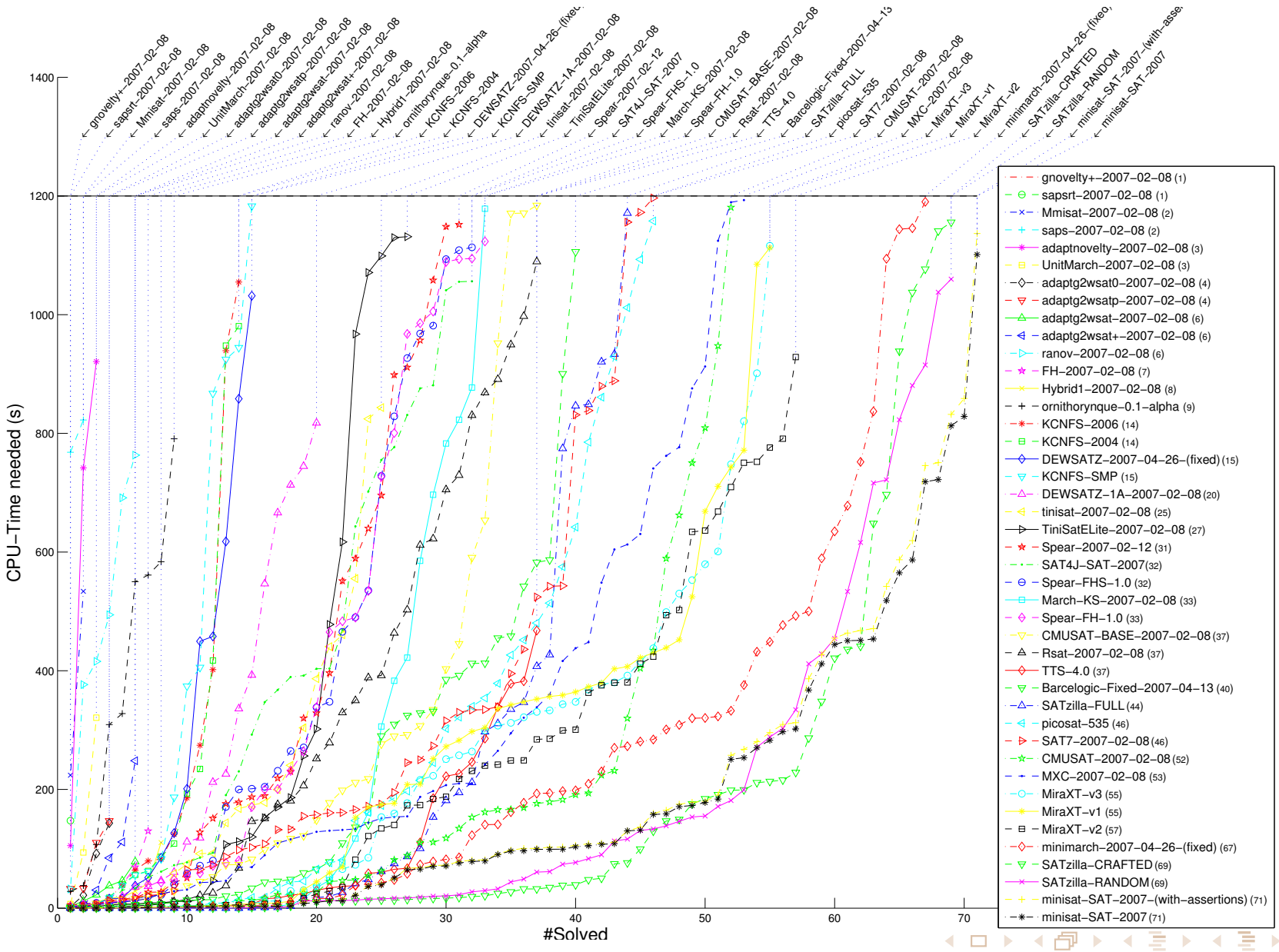
All solvers, all benchmarks



Clustering of solvers



All solvers on Crafted benchmarks



Scores of solvers on Crafted benchmakrs

Solver	#Solv.	Sc. All	Sc. SAT	Sc. UNS
TTS 4.0	37 (2 35)	36458	252	32300
SATzilla CRAFTED	69 (24 45)	30337	8239	22054
March-KS 2007-02-08	33 (15 18)	29134	9621	14048
minisat SAT-2007	71 (22 49)	24601	8652	15905
MXC 2007-02-08	53 (15 38)	15504	2133	13596
MiraXT v3	57 (18 39)	14739	6012	8683
MiraXT v1	55 (15 40)	13793	5318	8431
CMUSAT 2007-02-08	52 (15 37)	13728	5779	9405
SATzilla RANDOM	44 (22 22)	13113	8117	5667
MiraXT v2	55 (15 40)	10923	2837	8543
picosat 535	46 (20 26)	10514	4838	5841
Rsat 2007-02-08	37 (11 26)	9164	2583	6574
SAT7 2007-02-08	46 (16 30)	7353	2398	5216
CMUSAT-BASE 2007-02-08	37 (14 23)	5276	1694	4074
DEWSATZ-1A 2007-02-08	20 (4 16)	4609	434	4258
TiniSatELite 2007-02-08	27 (9 18)	3574	1135	2854
tinisat 2007-02-08	25 (10 15)	3134	1131	2418
KCNFS SMP	15 (2 13)	2222	272	2015

Scores of solvers on Crafted benchmakrs

Solver	#Solv.	Sc. All	Sc. SAT	Sc. UNS
TTS 4.0	37 (2 35)	36458	252	32300
SATzilla CRAFTED	69 (24 45)	30337	8239	22054
March-KS 2007-02-08	33 (15 18)	29134	9621	14048
minisat SAT-2007	71 (22 49)	24601	8652	15905
MXC 2007-02-08	53 (15 38)	15504	2133	13596
MiraXT v3	57 (18 39)	14739	6012	8683
MiraXT v1	55 (15 40)	13793	5318	8431
CMUSAT 2007-02-08	52 (15 37)	13728	5779	9405
SATzilla RANDOM	44 (22 22)	13113	8117	5667
MiraXT v2	55 (15 40)	10923	2837	8543
picosat 535	46 (20 26)	10514	4838	5841
Rsat 2007-02-08	37 (11 26)	9164	2583	6574
SAT7 2007-02-08	46 (16 30)	7353	2398	5216
CMUSAT-BASE 2007-02-08	37 (14 23)	5276	1694	4074
DEWSATZ-1A 2007-02-08	20 (4 16)	4609	434	4258
TiniSatELite 2007-02-08	27 (9 18)	3574	1135	2854
tinisat 2007-02-08	25 (10 15)	3134	1131	2418
KCNFS SMP	15 (2 13)	2222	272	2015

The TTS case

Scoring not so strange behavior

How can TTS have a so great score while solving so few benchmarks?

The TTS case

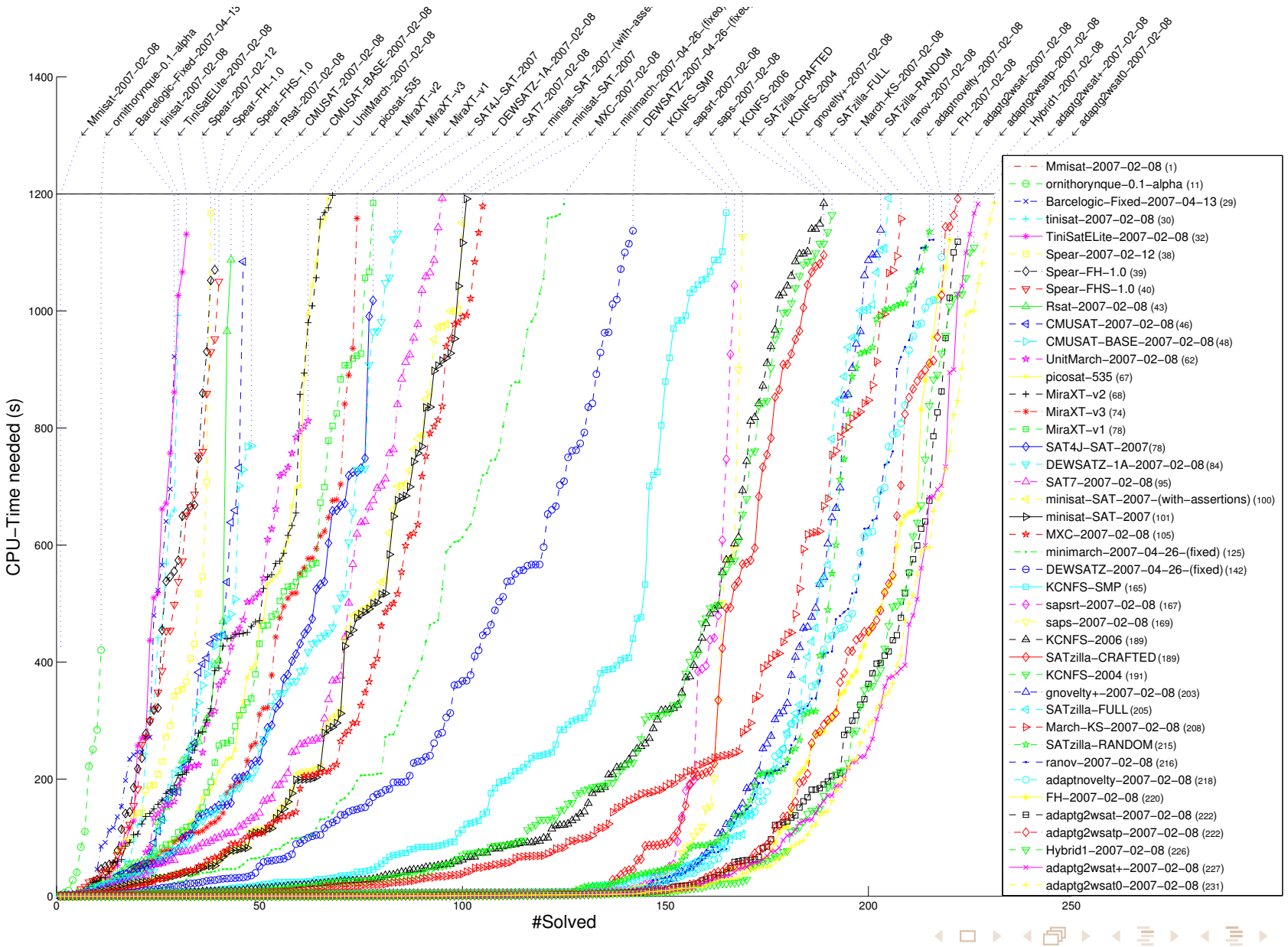
Scoring not so strange behavior

How can TTS have a so great score while solving so few benchmarks?

TTS is the only one to solve the following benchmarks

spence/hard/s101-100
spence/hard/s97-100
Hard/contest05/counting-php/harder-fphp-016-015
Hard/contest05/counting-php/easier-fphp-020-015
Hard/contest05/counting-php/harder-fphp-018-017
Difficult/contest05/sabharwal/counting-easier-fphp-014-012
Difficult/contest05/sabharwal/counting-harder-php-018-017
Difficult/contest05/sabharwal/counting-easier-php-018-014
Difficult/contest05/sabharwal/counting-harder-php-014-013
Difficult/contest05/jarvisalo/mod2c-3cage-unsat-10-3
Difficult/contest05/jarvisalo/mod2c-3cage-unsat-10-2

All solvers on Random benchmarks



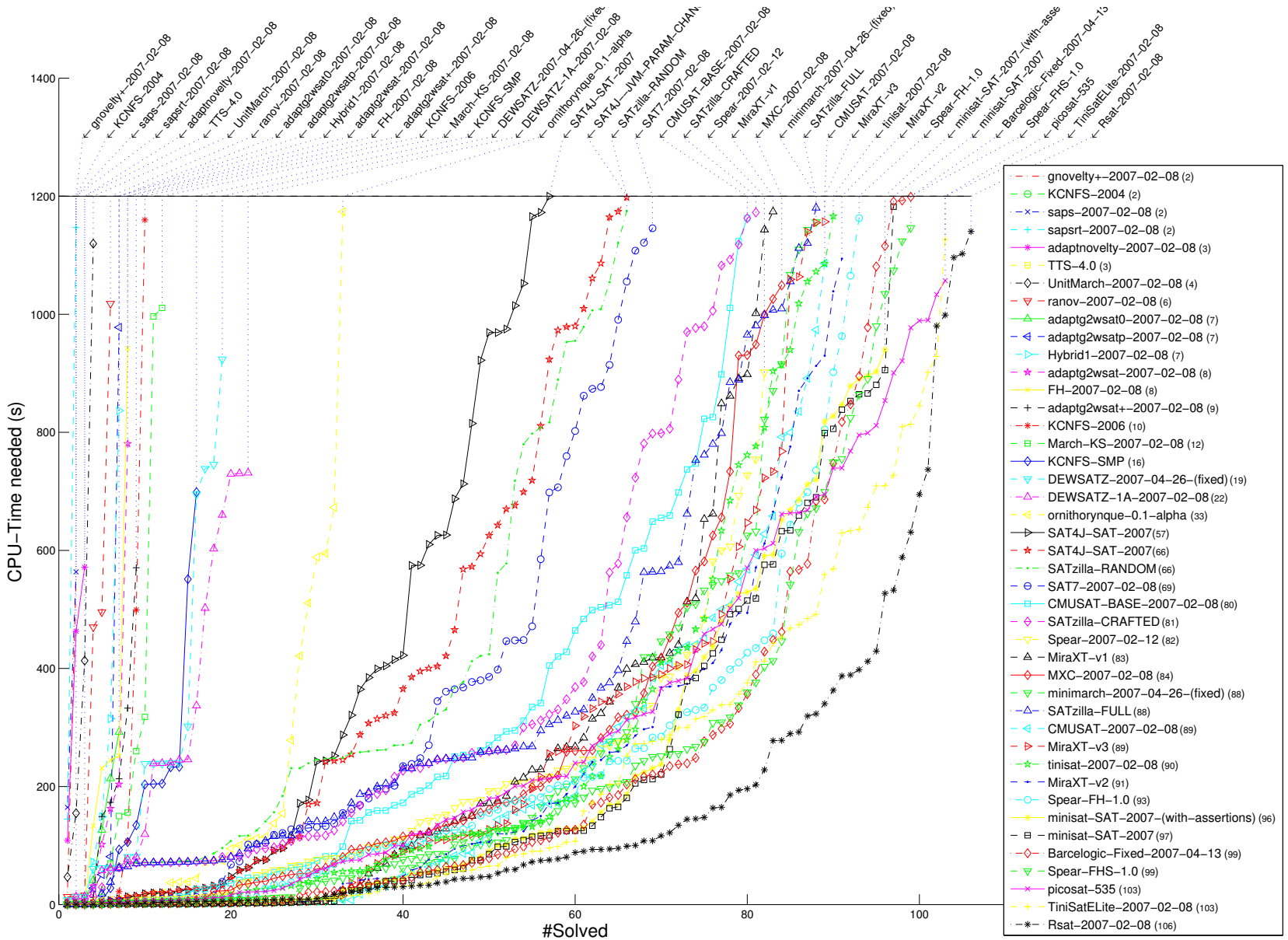
Scores of solvers on Random benchmarks

Solver	#Solv.	Sc. All	Sc. SAT	Sc. UNS
gnovelty+ 2007-02-08	203 (203 0)	69639	65013	0
adaptg2wsat0 2007-02-08	231 (231 0)	68244	68285	0
adaptg2wsat+ 2007-02-08	227 (227 0)	63732	63773	0
Hybrid1 2007-02-08	226 (226 0)	61927	61968	0
adaptnovelty 2007-02-08	218 (218 0)	59476	57517	0
adaptg2wsatp 2007-02-08	222 (222 0)	57506	56880	0
FH 2007-02-08	220 (220 0)	57102	56477	0
March-KS 2007-02-08	208 (121 87)	48845	14386	41953
ranov 2007-02-08	216 (216 0)	45428	45469	0
KCNFS 2004	191 (106 85)	40314	12362	34447
KCNFS 2006	189 (105 84)	38672	11618	33549
sapsrt 2007-02-08	167 (167 0)	37641	37682	0
SATzilla RANDOM	215 (137 78)	37616	16678	27432
KCNFS SMP	165 (80 85)	37031	7862	34827
saps 2007-02-08	169 (169 0)	32931	32972	0
SATzilla CRAFTED	189 (143 46)	23670	17606	8358
MXC 2007-02-08	105 (63 42)	10480	5924	6671
minisat SAT-2007	101 (59 42)	9727	5097	6745

Scores of solvers on Random benchmarks

Solver	#Solv.	Sc. All	Sc. SAT	Sc. UNS
gnovelty+ 2007-02-08	203 (203 0)	69639	65013	0
adaptg2wsat0 2007-02-08	231 (231 0)	68244	68285	0
adaptg2wsat+ 2007-02-08	227 (227 0)	63732	63773	0
Hybrid1 2007-02-08	226 (226 0)	61927	61968	0
adaptnovelty 2007-02-08	218 (218 0)	59476	57517	0
adaptg2wsatp 2007-02-08	222 (222 0)	57506	56880	0
FH 2007-02-08	220 (220 0)	57102	56477	0
March-KS 2007-02-08	208 (121 87)	48845	14386	41953
ranov 2007-02-08	216 (216 0)	45428	45469	0
KCNFS 2004	191 (106 85)	40314	12362	34447
KCNFS 2006	189 (105 84)	38672	11618	33549
sapsrt 2007-02-08	167 (167 0)	37641	37682	0
SATzilla RANDOM	215 (137 78)	37616	16678	27432
KCNFS SMP	165 (80 85)	37031	7862	34827
saps 2007-02-08	169 (169 0)	32931	32972	0
SATzilla CRAFTED	189 (143 46)	23670	17606	8358
MXC 2007-02-08	105 (63 42)	10480	5924	6671
minisat SAT-2007	101 (59 42)	9727	5097	6745

All solvers on Industrial benchmarks



Scores of solvers on Industrial benchmarks

Solver	#Solv.	Sc. All	Sc. SAT	Sc. UNS
Rsat 2007-02-08	106 (47 59)	44626	19422	22801
TiniSatELite 2007-02-08	103 (43 60)	30305	10911	18991
minisat SAT-2007	97 (37 60)	29721	13467	15850
MiraXT v3	89 (35 54)	26595	13292	11359
picosat 535	103 (51 52)	25513	16749	9360
CMUSAT 2007-02-08	89 (28 61)	25339	7275	18186
MiraXT v1	91 (36 55)	23850	12138	11958
tinisat 2007-02-08	90 (40 50)	20241	10264	9274
MXC 2007-02-08	84 (37 47)	19175	10609	9163
MiraXT v2	83 (27 56)	18824	6708	12513
CMUSAT-BASE 2007-02-08	80 (38 42)	18730	10138	8714
SATzilla CRAFTED	81 (35 46)	16727	8006	9317
SATzilla RANDOM	66 (28 38)	14937	6484	8103
SAT7 2007-02-08	69 (30 39)	11454	5883	5898
KCNFS 2006	16 (1 15)	3030	277	2952
KCNFS SMP	10 (0 10)	2877	0	2877
DEWSATZ-1A 2007-02-08	22 (3 19)	2591	249	2342
FH 2007-02-08	8 (5 3)	1871	1542	329

Scores of solvers on Industrial benchmarks

Solver	#Solv.	Sc. All	Sc. SAT	Sc. UNS
Rsat 2007-02-08	106 (47 59)	44626	19422	22801
TiniSatELite 2007-02-08	103 (43 60)	30305	10911	18991
minisat SAT-2007	97 (37 60)	29721	13467	15850
MiraXT v3	89 (35 54)	26595	13292	11359
picosat 535	103 (51 52)	25513	16749	9360
CMUSAT 2007-02-08	89 (28 61)	25339	7275	18186
MiraXT v1	91 (36 55)	23850	12138	11958
tinisat 2007-02-08	90 (40 50)	20241	10264	9274
MXC 2007-02-08	84 (37 47)	19175	10609	9163
MiraXT v2	83 (27 56)	18824	6708	12513
CMUSAT-BASE 2007-02-08	80 (38 42)	18730	10138	8714
SATzilla CRAFTED	81 (35 46)	16727	8006	9317
SATzilla RANDOM	66 (28 38)	14937	6484	8103
SAT7 2007-02-08	69 (30 39)	11454	5883	5898
KCNFS 2006	16 (1 15)	3030	277	2952
KCNFS SMP	10 (0 10)	2877	0	2877
DEWSATZ-1A 2007-02-08	22 (3 19)	2591	249	2342
FH 2007-02-08	8 (5 3)	1871	1542	329

SOTAC solvers

Over all benchmarks, first stage

Solver	#SOTAC
TTS	11
gnovelty+	5
March KS	4
adaptg2wsatp	3
SATzilla	3
adaptnovelty	3
adaptg2wsat+	3
adaptg2wsat0	2
adaptg2wsat	2
Rsat	2

Solver	#SOTAC
Hybrid1	2
MiraXT	2
CMUSAT	2
MiraXT	1
picosat	1
DEWSATZ	1
SAT4J	1
minimarch	1
Spear	1
MXC	1

Solved crafted benchmarks, first stage

Series	SubSeries	#Solved	#Tot
contest02	mix	3	10
contest03	looksrandom	0	10
contest03	others	1	10
contest04	connamacher	2	10
contest04	others	1	10
contest05	counting-clq	0	10
contest05	counting-php	3	10
contest05	jarvisalo	10	10
contest05	others	14	20
contest05	pebbling	10	10
contest05	phnf	2	10
contest05	QG	3	10
contest05	sabharwal	10	10
Difficult	contest-02-03-04	9	9
Difficult	contest05	31	36
spence	hard	2	6
spence	medium	10	10

Solved random benchmarks, first stage

Series	SubSeries	#Solved	#Tot
2+p-p0.7	v3500	10	10
2+p-p0.7	v6500	6	10
2+p-p0.8	v1295	10	10
2+p-p0.8	v2405	6	11
2+p-p0.9	v1170	5	10
2+p-p0.9	v990	6	10
LargeSize-3SAT	v7000	10	10
LargeSize-3SAT	v19000	1	10
LargeSize-5SAT	v600	10	10
LargeSize-5SAT	v1100	7	10
LargeSize-7SAT	v140	10	10
LargeSize-7SAT	v240	0	10
OnTreshold-3SAT	v360	10	10
OnTreshold-3SAT	v650	5	10
OnTreshold-5SAT	v70	10	10
OnTreshold-5SAT	v130	5	10
OnTreshold-7SAT	v45	10	10
OnTreshold-7SAT	v75	6	10

Solved industrial benchmarks, first stage

Series	SubSeries	#Solved	#Tot
anbulagan	hard-sat	1	7
anbulagan	hard-unsat	1	24
anbulagan	medium-sat	10	10
anbulagan	medium-unsat	5	6
babic	dspam	10	10
babic	xinetd	10	10
crypto	crypto	10	10
fuhs	hard	1	6
fuhs	medium	10	10
grieu	grieu	10	10
jarvisalo	jarvisalo	5	7
manolios	manolios	10	10
narain	narain	4	5
palacios	hard	2	7
palacios	uts	10	10
velev	vliw-sat-4.0	10	10
velev	vliw-unsat-2.0	4	9
zarpas	IBM-FV-2002-13-rule-1	0	19
zarpas	IBM-FV-2002-31-1-rule-1	2	16
zarpas	IBM-FV-2004-30	10	18

Road Map

1. History, rules and Goals
2. The participants
3. The First Stage
4. The Winners
   - Smallest unsolved benchmarks
   - The crafted winners
   - The random winners
   - The industrial winners
5. Conclusion and next contests
6. Certified UNSAT Special Track
7. And Inverter Graph Special Track

Smallest unsolved crafted benchmarks

Benchmark	size	#clauses	#variables
spence/hard/s117-100	732	244	117
Hard/contest03/looksrand/hgen8-n260	904	399	212
Hard/contest03/looksrand/hgen8-n320-01	1102	486	260
Hard/contest03/looksrand/hgen8-n320-03	1108	489	260
Hard/contest03/looksrand/hgen8-n320-02	1114	492	260

Smallest unsolved random benchmarks

Benchmark	size	#clauses	#va
OnTres/3SAT/v550/unif-k3-r4.26-v550-c2343-07.UNSAT	7029	2343	
OnTres/3SAT/v550/unif-k3-r4.26-v550-c2343-03.UNSAT	7029	2343	
OnTres/3SAT/v550/unif-k3-r4.26-v550-c2343-15.UNSAT	7029	2343	
OnTres/3SAT/v550/unif-k3-r4.26-v550-c2343-01.UNSAT	7029	2343	
OnTres/3SAT/v550/unif-k3-r4.26-v550-c2343-20.UNSAT	7029	2343	

Smallest unsolved industrial benchmarks

Benchmark	size	#clauses	#variables
jarvisalo/eq.atree.braun.12.unsat	14874	5726	1694
jarvisalo/eq.atree.braun.13.unsat	17668	6802	2010
fuhs/hard/AProVE07-01	76290	28770	7502
fuhs/hard/AProVE07-25	83706	31884	8920
fuhs/hard/AProVE07-26	211276	79766	21734

And Now...

The results for the **crafted** category...

Category crafted, SAT+UNS specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, minisat SAT-2007, picosat 535

And the winner is...

Category crafted, Sat+Uns specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, minisat SAT-2007, picosat 535

And the winner is...

SATzilla CRAFTED

Category crafted, Sat+Uns specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, minisat SAT-2007, picosat 535

And the winner is...

SATzilla CRAFTED

minisat SAT-2007

Category crafted, Sat+Uns specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, minisat SAT-2007, picosat 535

And the winner is...

SATzilla CRAFTED

minisat SAT-2007

MXC 2007-02-08

Category crafted, Sat+Uns specialty

The details!

Solver	Score	#SAT	#Uns
SATzilla CRAFTED	74469	27	67
minisat SAT-2007	63371	26	72
MXC 2007-02-08	39848	20	57
MiraXT v3	34236	24	54
CMUSAT 2007-02-08	26461	21	45
Rsat 2007-02-08	19532	15	40
picosat 535	19081	22	38

Category crafted, SAT specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, TTS 4.0, minisat SAT-2007, picosat 535

And the winner is...

Category crafted, SAT specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, TTS 4.0, minisat SAT-2007, picosat 535

And the winner is...

March-KS 2007-02-08

Category crafted, SAT specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, TTS 4.0, minisat SAT-2007, picosat 535

And the winner is...

March-KS 2007-02-08

SATzilla CRAFTED

Category crafted, SAT specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, TTS 4.0, minisat SAT-2007, picosat 535

And the winner is...

March-KS 2007-02-08

SATzilla CRAFTED

minisat SAT-2007

Category crafted, SAT specialty

The details!

Solver	Score	#SAT	#Uns
March-KS 2007-02-08	16323	18	0
SATzilla CRAFTED	14275	27	0
minisat SAT-2007	13785	26	0
MiraXT v3	11601	24	0
CMUSAT 2007-02-08	8093	21	0
picosat 535	7153	22	0
MXC 2007-02-08	5136	20	0
Rsat 2007-02-08	4475	15	0
TTS 4.0	1155	3	0

Category crafted, Unsat specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, TTS 4.0, minisat SAT-2007, picosat 535

And the winner is...

Category crafted, Unsat specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, TTS 4.0, minisat SAT-2007, picosat 535

And the winner is...

SATzilla CRAFTED

Category crafted, Unsat specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, TTS 4.0, minisat SAT-2007, picosat 535

And the winner is...

SATzilla CRAFTED

TTS 4.0

Category crafted, Unsat specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, Rsat 2007-02-08, SATzilla CRAFTED, TTS 4.0, minisat SAT-2007, picosat 535

And the winner is...

SATzilla CRAFTED

TTS 4.0

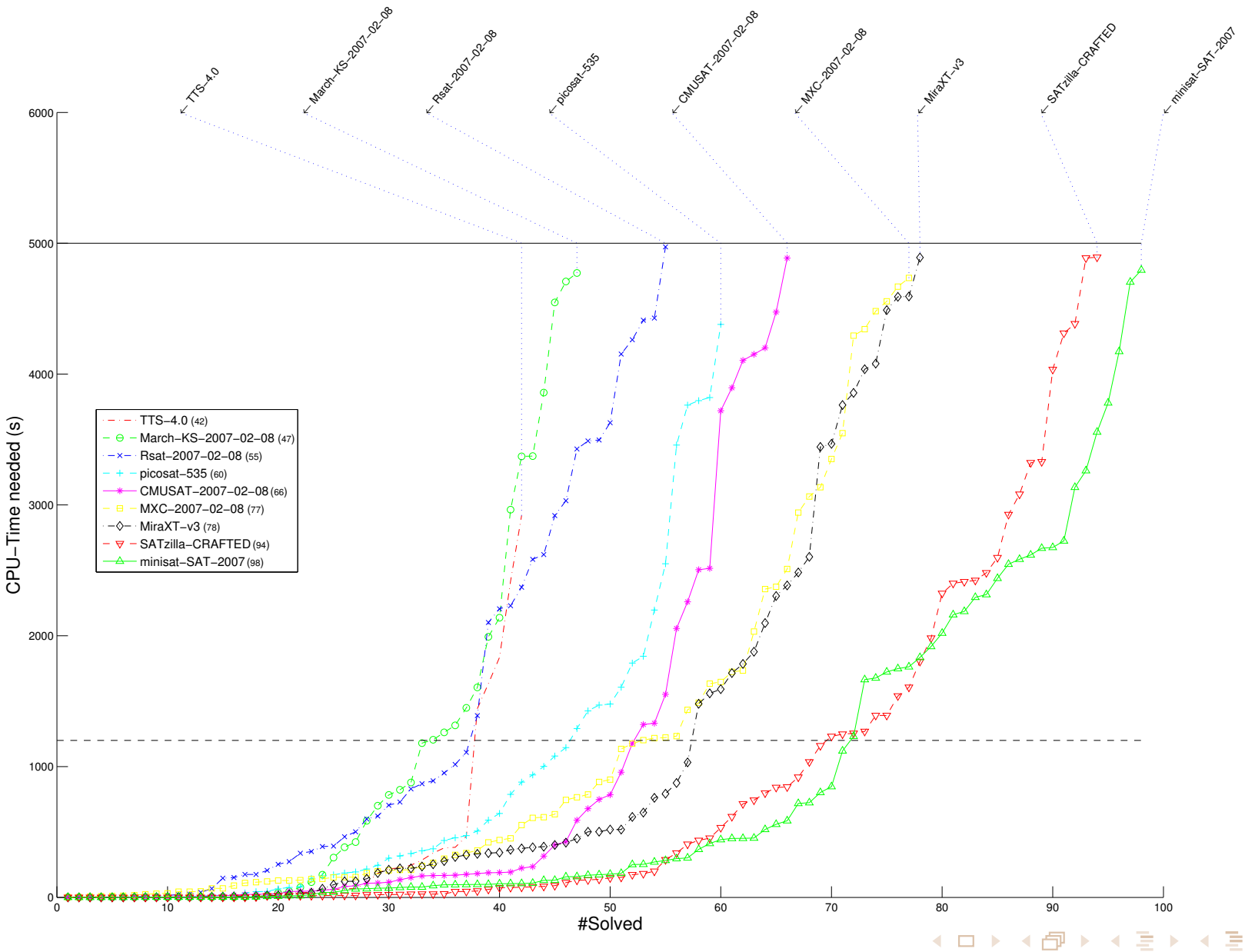
minisat SAT-2007

Category crafted, Unsat specialty

The details!

Solver	Score	#SAT	#Uns
SATzilla CRAFTED	39922	0	67
TTS 4.0	38950	0	39
minisat SAT-2007	38090	0	72
MXC 2007-02-08	26151	0	57
March-KS 2007-02-08	19684	0	29
MiraXT v3	18379	0	54
CMUSAT 2007-02-08	16437	0	45
Rsat 2007-02-08	13373	0	40
picosat 535	11011	0	38

All solvers on Crafted benchmarks



And Now...

The results for the **random** category...

Category random, Sat+Uns specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, SAT7 2007-02-08, SATzilla RANDOM, minisat SAT-2007

And the winner is...

Category random, Sat+Uns specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, SAT7 2007-02-08, SATzilla RANDOM, minisat SAT-2007

And the winner is...

SATzilla RANDOM

Category random, Sat+Uns specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, SAT7 2007-02-08, SATzilla RANDOM, minisat SAT-2007

And the winner is...

SATzilla RANDOM

March-KS 2007-02-08

Category random, Sat+Uns specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, SAT7 2007-02-08, SATzilla RANDOM, minisat SAT-2007

And the winner is...

SATzilla RANDOM

March-KS 2007-02-08

KCNFS 2004

Category random, Sat+Uns specialty

The details!

Solver	Score	#SAT	#Uns
SATzilla RANDOM	189835	147	101
March-KS 2007-02-08	167430	146	111
KCNFS 2004	124280	130	107
minisat SAT-2007	36387	83	57
MXC 2007-02-08	35538	82	53
DEWSATZ-1A 2007-02-08	30754	73	48
SAT7 2007-02-08	30282	74	48
MiraXT v3	26491	69	37

Category random, SAT specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, SAT7 2007-02-08, SATzilla RANDOM, adaptg2wsat+ 2007-02-08, adaptg2wsat0 2007-02-08, adaptnovelty 2007-02-08, gnovelty+ 2007-02-08, minisat SAT-2007, ranov 2007-02-08, sapsrt 2007-02-08

And the winner is...

Category random, SAT specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, SAT7 2007-02-08, SATzilla RANDOM, adaptg2wsat+ 2007-02-08, adaptg2wsat0 2007-02-08, adaptnovelty 2007-02-08, gnovelty+ 2007-02-08, minisat SAT-2007, ranov 2007-02-08, sapsrt 2007-02-08

And the winner is...

gnovelty+ 2007-02-08

Category random, SAT specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, SAT7 2007-02-08, SATzilla RANDOM, adaptg2wsat+ 2007-02-08, adaptg2wsat0 2007-02-08, adaptnovelty 2007-02-08, gnovelty+ 2007-02-08, minisat SAT-2007, ranov 2007-02-08, sapsrt 2007-02-08

And the winner is...

gnovelty+ 2007-02-08

adaptg2wsat0 2007-02-08

Category random, Sat specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, MiraXT v3, SAT7 2007-02-08, SATzilla RANDOM, adaptg2wsat+ 2007-02-08, adaptg2wsat0 2007-02-08, adaptnovelty 2007-02-08, gnovelty+ 2007-02-08, minisat SAT-2007, ranov 2007-02-08, sapsrt 2007-02-08

And the winner is...

gnovelty+ 2007-02-08

adaptg2wsat0 2007-02-08

adaptg2wsat+ 2007-02-08

Category random, Sat specialty

The details!

Solver	Score	#SAT	#Uns
gnovelty+ 2007-02-08	122500	242	0
adaptg2wsat0 2007-02-08	114109	248	0
adaptg2wsat+ 2007-02-08	112877	252	0
adaptnovelty 2007-02-08	96497	240	0
ranov 2007-02-08	86647	242	0
sapsrt 2007-02-08	68218	188	0
SATzilla RANDOM	29895	147	0
March-KS 2007-02-08	26721	146	0
KCNFS 2004	23786	130	0
MXC 2007-02-08	13308	82	0
minisat SAT-2007	12397	83	0
SAT7 2007-02-08	11594	74	0
DEWSATZ-1A 2007-02-08	10832	73	0
MiraXT v3	10614	69	0

Category random, Unsat specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, SAT7 2007-02-08, SATzilla RANDOM, minisat SAT-2007, MiraXT v3

And the winner is...

Category random, Unsat specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, SAT7 2007-02-08, SATzilla RANDOM, minisat SAT-2007, MiraXT v3

And the winner is...

March-KS 2007-02-08

Category random, Unsat specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, SAT7 2007-02-08, SATzilla RANDOM, minisat SAT-2007, MiraXT v3

And the winner is...

March-KS 2007-02-08

KCNFS 2004

Category random, Unsat specialty

The winner!

Were qualified

DEWSATZ-1A 2007-02-08, KCNFS 2004, MXC 2007-02-08, March-KS 2007-02-08, SAT7 2007-02-08, SATzilla RANDOM, minisat SAT-2007, MiraXT v3

And the winner is...

March-KS 2007-02-08

KCNFS 2004

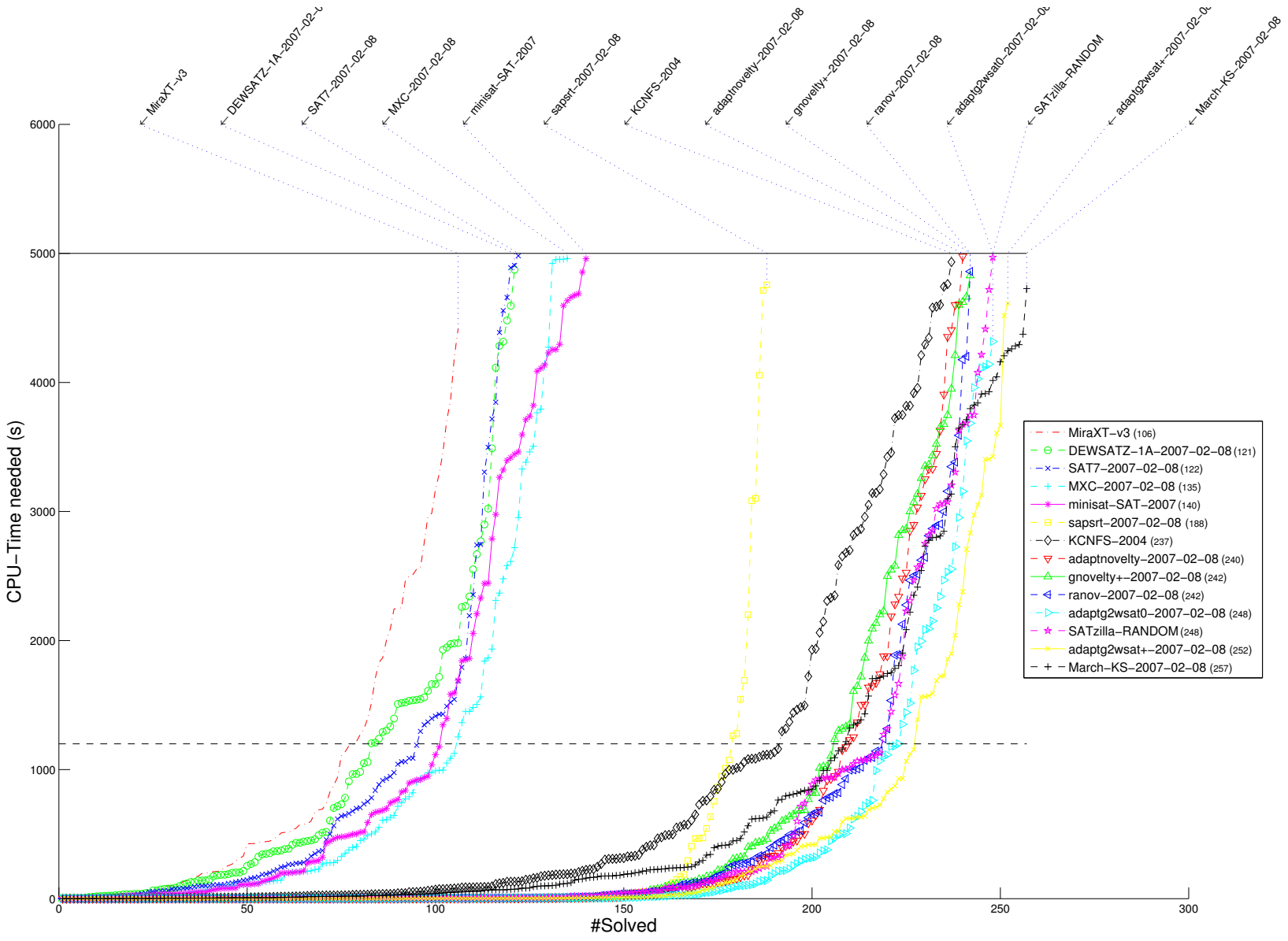
SATzilla RANDOM

Category random, Unsat specialty

The details!

Solver	Score	#SAT	#Uns
March-KS 2007-02-08	88041	0	111
KCNFS 2004	73087	0	107
SATzilla RANDOM	61008	0	101
minisat SAT-2007	16870	0	57
MXC 2007-02-08	14997	0	53
DEWSATZ-1A 2007-02-08	12794	0	48
SAT7 2007-02-08	12628	0	48
MiraXT v3	9573.3	37	37

All solvers on Random benchmarks



And Now...

The results for the **industrial** category...

Category industrial, Sat+Uns specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla CRAFTED, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Category industrial, Sat+Uns specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla CRAFTED, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Rsat 2007-02-08

Category industrial, Sat+Uns specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla crafted, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Rsat 2007-02-08

picosat 535

Category industrial, Sat+Uns specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla crafted, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Rsat 2007-02-08

picosat 535

minisat SAT-2007

Category industrial, Sat+Uns specialty

The details!

Solver	Score	#SAT	#Uns
Rsat 2007-02-08	55452	63	76
picosat 535	51639	72	67
minisat SAT-2007	50398	53	79
TiniSatELite 2007-02-08	49560	55	75
CMUSAT 2007-02-08	41197	46	77
MXC 2007-02-08	38004	47	69
MiraXT v3	37700	52	74
SATzilla CRAFTED	31439	47	67
SAT7 2007-02-08	27607	51	63

Category industrial, SAT specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla CRAFTED, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Category industrial, SAT specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla CRAFTED, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

picosat 535

Category industrial, SAT specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla crafted, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

picosat 535

Rsat 2007-02-08

Category industrial, SAT specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla CRAFTED, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

picosat 535

Rsat 2007-02-08

minisat SAT-2007

Category industrial, SAT specialty

The details!

Solver	Score	#SAT	#Uns
picosat 535	33812	72	0
Rsat 2007-02-08	24850	63	0
minisat SAT-2007	20639	53	0
TiniSatELite 2007-02-08	19977	55	0
MiraXT v3	18288	52	0
CMUSAT 2007-02-08	17571	46	0
MXC 2007-02-08	17006	47	0
SATzilla CRAFTED	13491	47	0
SAT7 2007-02-08	12363	51	0

Category industrial, Unsat specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla crafted, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Category industrial, Unsat specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla crafted, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Rsat 2007-02-08

Category industrial, Unsat specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla CRAFTED, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Rsat 2007-02-08

minisat SAT-2007

Category industrial, Unsat specialty

The winner!

Were qualified

CMUSAT 2007-02-08, MXC 2007-02-08, MiraXT v3, Rsat 2007-02-08, SAT7 2007-02-08, SATzilla CRAFTED, TiniSatELite 2007-02-08, minisat SAT-2007, picosat 535

And the winner is...

Rsat 2007-02-08

minisat SAT-2007

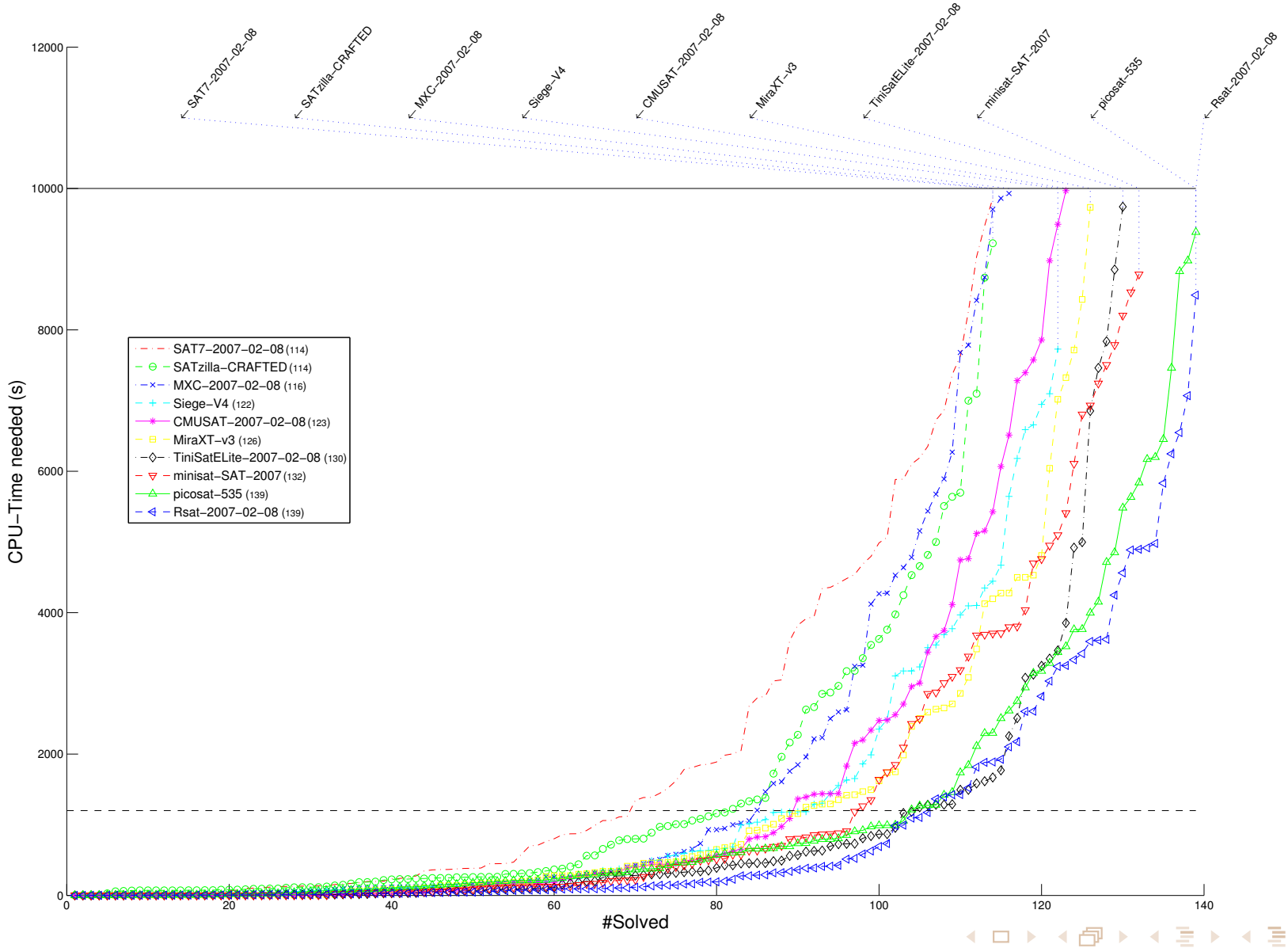
TiniSatELite 2007-02-08

Category industrial, Unsat specialty

The details!

Solver	Score	#SAT	#Uns
Rsat 2007-02-08	30193	0	76
minisat SAT-2007	29683	0	79
TiniSatELite 2007-02-08	29459	0	75
CMUSAT 2007-02-08	22146	0	77
MXC 2007-02-08	20922	0	69
MiraXT v3	19086	0	74
SATzilla CRAFTED	17873	0	67
picosat 535	17203	0	67
SAT7 2007-02-08	15430	0	63

All solvers on Industrial benchmarks



Detailed results available on the web

Go to
`http://www.cril.univ-artois.fr/SAT07`

Road Map

1. History, rules and Goals
2. The participants
3. The First Stage
4. The Winners
5. Conclusion and next contests
   - Conclusion
6. Certified UNSAT Special Track
7. And Inverter Graph Special Track

Let's give authors a warm applause

They deserve it.

Subject to discussion

Scoring mechanism

Some solvers were ranked before other solvers that solved more instances. However, it guarantees that benchmarks have different hardness.
Solving a hard benchmark gives more credit than an easy one.
One solution may be not to compute the score of benchmarks during the contest, but before the contest, based on previous solvers (like in CASC).
Any suggestion welcome!

Subject to discussion

Scoring mechanism

Some solvers were ranked before other solvers that solved more instances. However, it guarantees that benchmarks have different hardness.
Solving a hard benchmark gives more credit than an easy one.
One solution may be not to compute the score of benchmarks during the contest, but before the contest, based on previous solvers (like in CASC).
Any suggestion welcome!

Other hot topic:

- Taking multi-core architecture into account for time measures. Real time vs CPU time?

Next Competition?

Next year, **there will be a competitive event** (race or competition).
New organizers are (almost) ready :)

Next Competition?

Next year, **there will be a competitive event** (race or competition).
New organizers are (almost) ready :)

Road Map

1. History, rules and Goals
2. The participants
3. The First Stage
4. The Winners
5. Conclusion and next contests
6. Certified UNSAT Special Track
7. And Inverter Graph Special Track

The Rules

The first part consisted of the well-studied Pigeon-Hole formulas because there is a known proof scheme that produces what are believed to be the shortest possible proofs (see "Theory" in the table), due to S. Cook in 1971, and rediscovered by A. Hakin about 1985, and again by A. Van Gelder about 2003. The formula for P pigeons has $P(P-1)$ variables and $P(P^2-2P+3)/2$ clauses. The "Theory" derivation has $(P-1)(P+2)2^{P-3}$ clauses.
Note that the tts solver produces proofs that grow close to the optimum ratio!
The second section consisted of 17 benchmarks from the regular track, industrial category, of the competition.

Resources: 1 CPU hour at 2.6 GHz, 8 GB real memory, about 6 GB swap, and file size limit 36 GB (most cases) on a local disk.

The Contestants

Program	Proof Format	Authors
booleforce	%RES	A. Biere
picosat	%RES	A. Biere
picosat	%RUP	A. Biere
tts	%RPT	I. Spence
zchaffSE07	%RES	Z. Fu, L. Zhang, (others?)
zchaff_0	%RES	A. Van Gelder, Z. Fu, L. Zhang, (others?)

Who are the winners?

It was a mixed bag. tts was the clear winner on pigeon-hole formulas, but did very poorly on industrial formulas. booleforce edged out zchaff on industrial formulas, but had a few programming problems, and did poorly on pigeon-hole formulas. zchaff had the best balance between the two sections. picosat with the RUP format produced the most proofs, but many RUP proofs could not be verified within the allocated time and disk space.

Road Map

1. History, rules and Goals
2. The participants
3. The First Stage
4. The Winners
5. Conclusion and next contests
6. Certified UNSAT Special Track
7. And Inverter Graph Special Track

And Inverter Graphs special track

The idea

- In 2005, the special ”non clausal” track using the eDimacs format provided by Fahiem Bacchus and Toby Walsh received no submission.
- In 2006, Armin Biere proposed the And Inverter Graph format for ”structural/non clausal” input. Many translators from and to AIG are made available.
- In 2007, the first HardWare Model Checking Competition is organized in conjonction with CAV’07 using the AIGER format.

The idea is to experiment the AIG format for benchmarks in the industrial category that tend to be huge, with a structure lost during CNF translation.

And Inverter Graphs special track

The competitors

Only 3 submitters for that special track, for a total of 5 solvers.

- **AIG-CMUSAT** from Himanshu Jain. Uses aigtocnf translator available in AIGER distribution to feed CMUSAT SAT solver.
- **AIGSAT-SP** from Himanshu Jain, C. Bartzis and E. Clarke. AIG to NNF translation.
- **AIGSAT-WC** from Himanshu Jain, C. Bartzis and E. Clarke. AIG to NNF translation. BCP different from previous one.
- **MiniCirc** from Niklas Eèn. Uses a three stage preprocessing from AIG to CNF to feed MiniSAT 2.0 SAT solver.
- **PICOAIGERSAT** from Armin Biere. Uses a basic AIG to CNF translation to feed picosat SAT solver.

The benchmarks were provided by Armin Biere (available from AIGER web site) or translated from CNF benchmarks in the industrial category (IBM and Velev benchmarks too big to be translated with current translator developed by Armin Biere).

And Inverter Graphs special track

The results

- Available in details from http://www.cril.univ-artois.fr/SAT07/
- Need to be taken with care: no certificates/witnesses provided for SAT answers.
- Solvers received late: minor I/O and stability problems no fixed for some solvers (e.g. MiniCirc) on current results.

Conclusion:

- Check by yourself the results.
- Check the competition website for additional information in the next few weeks.
- Embrace the AIGER format.