Fourth International CSP Solver Competition CSC'09

C. LECOUTRE, O. ROUSSEL and M.R.C VAN DONGEN
CRIL, University of Artois, Lens, France
University College, Cork, Ireland

CP 2009: 15th International Conference on
Principles and Practice of Constraint Programming

September, 2009

- Competition goal
- Environment
- Solvers
- Results of the CSP competition

Help identifying successful techniques in constraint solving

- by comparing solvers in the same environment, on the very same instances
- on a wide and hopefully representative set of instances
- in the hope of boosting research as other competitions did (e.g. SAT competitions)

Key points:

- solvers are considered as black boxes
- CPU time is the final measure of efficiency

Main points in this year competition

- no new benchmark submitted
- no new global constraint
- not enough submitted solvers to run a Max-CSP or WCSP competition
- fewer solvers than last year (but 2 new solvers submitted)
- same procedure as last year for the selection of instances, but with a different seed.
- better comparison of solvers on global constraints: only compare solvers which have the same capabilities

Classification of solvers

Solvers may not have support for all kinds of constraints:

- binary or non-binary constraints
- extension or intension
- global constraints
  - alldifferent
  - weightedsum
  - element
  - cumulative

Solvers are registered in the categories they support.

Instances categories

4 categories were pre-defined:

2-ARY-EXT   only binary constraints defined in extension

2-ARY-INT   only binary constraints (some of them being defined by a predicate)

N-ARY-EXT   some n-ary constraints (all constraints defined in extension)

N-ARY-INT   some n-ary constraints and some constraints defined by a predicate

3 categories with global constraints corresponding to common solver capabilities

alldiff   no other global constraint than alldifferent

alldiff+elt+wsum   no other global constraint than alldifferent, element, weightedsum

alldiff+elt+wsum+cumulative   no other global constraint than alldifferent, element, weightedsum and cumulative

Judges and working group

This year, three judges were in charge of supervising the competition:

- Helmut SIMONIS, Cork Constraint Computation Centre (4C), Ireland.
- George KATSIRELOS, NICTA, Sydney, Australia.
- Matt STREETER, Google, USA.

They decided to select instances the same way as last year :

- The number of instances to select in each series was kept the same as in 2008 (selection made by Deepak MEHTA).
- Individual instances were selected by a random process in each series with a random seed different from last year.

CSP solvers submitted

9 submitted solvers (and a few more versions)

Abscon	a CSP solver in Java (same version as in 2008)
bpsolver	a solver integrated in a Prolog engine
choco	a CSP library in Java
Concrete	a CSP library in Java
Conquer	a CSP solver in Java
Mistral	a C++ constraint library
pcs	a C++ solver that can generate unsatifiability proofs
SAT4J CSP	based on a translation to SAT
Sugar	a SAT based solver

Evaluation environment

- Cluster of bi-Xeon 3 GHz, 2MB cache, 2GB RAM
  kindly provided by the CRIL, University of Artois, France
- All solvers were run in 32 bits mode
- Each solver was imposed a memory limit of 900 MB (to avoid swapping and to allow two jobs to run concurrently on a node without interference)
- CSP solvers were given a time limit of 30 minutes (1800s).
- 253 days of CPU time used by CSP solvers

- The environment performs the following, efficient checks:
  - for SATISFIABLE answers, solvers must output a complete instantiation and the system checks that it satisfies all constraints
  - for UNSATISFIABLE answers, the system only checks that no other solver proved satisfiability
- UNSATISFIABLE answers cannot be completely checked efficiently and therefore should be taken with caution.
- Solvers giving a wrong answer in a category are disqualified in that category.

Ranking of solvers and Virtual Best Solver (VBS)

Ranking based on two criteria:

1. the number of solved instances (#solved)
2. and in case of duce, the cumulated time on solved instances

The Virtual Best Solver (VBS)

- is the virtual solver obtained by combining the best results of all submitted solvers.
- could be obtained by running in parallel all submitted solvers
- represents the current state of the art (SOTA)
- is a reference for the evaluation of the other solvers

Results for 2-ARY-EXT

Rank	Solver	#solved	%inst.	%VBS
	Total number of instances: 635			
	Virtual Best Solver (VBS)	609	96%	100%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 635				
Virtual Best Solver (VBS)		609	96%	100%
3	Abscon 112v4 *AC*	551	87%	90%
4	Abscon 112v4 *ESAC*	547	86%	90%

Results for 2-ARY-EXT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 635				
Virtual Best Solver (VBS)		609	96%	100%
2	Choco2.1.1b *2009-07-16*	556	88%	91%
3	Abscon 112v4 *AC*	551	87%	90%
4	Abscon 112v4 *ESAC*	547	86%	90%

Results for 2-ARY-EXT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 635				
Virtual Best Solver (VBS)		609	96%	100%
1	Mistral *1.545*	570	90%	94%
2	Choco2.1.1b *2009-07-16*	556	88%	91%
3	Abscon 112v4 *AC*	551	87%	90%
4	Abscon 112v4 *ESAC*	547	86%	90%

Results for 2-ARY-EXT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 635				
Virtual Best Solver (VBS)		609	96%	100%
1	Mistral *1.545*	570	90%	94%
2	Choco2.1.1b *2009-07-16*	556	88%	91%
3	Abscon 112v4 *AC*	551	87%	90%
4	Abscon 112v4 *ESAC*	547	86%	90%
5	Choco2.1.1 *2009-06-10*	547	86%	90%
6	Concrete DC *2009-07-14*	504	79%	83%
7	Concrete *2009-07-14*	503	79%	83%
8	Sugar *v1.14.6+minisat*	466	73%	77%
9	Sugar *v1.14.6+picosat*	438	69%	72%
10	SAT4J CSP *2.1.1*	421	66%	69%
11	bpsolver *09*	416	66%	68%
12	pcs-restart *0.3.2*	394	62%	65%
13	pcs *0.3.2*	393	62%	65%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 686				
Virtual Best Solver (VBS)		617	90%	100%

Results for 2-ARY-INT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 686				
Virtual Best Solver (VBS)		617	90%	100%
4	Choco2.1.1b *2009-07-16*	510	74%	83%
5	Choco2.1.1 *2009-06-10*	508	74%	82%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 686				
Virtual Best Solver (VBS)		617	90%	100%
3	Mistral *1.545*	511	74%	83%
4	Choco2.1.1b *2009-07-16*	510	74%	83%
5	Choco2.1.1 *2009-06-10*	508	74%	82%

Results for 2-ARY-INT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 686				
Virtual Best Solver (VBS)		617	90%	100%
1	Abscon 112v4 *ESAC*	517	75%	84%
2	Abscon 112v4 *AC*	513	75%	83%
3	Mistral *1.545*	511	74%	83%
4	Choco2.1.1b *2009-07-16*	510	74%	83%
5	Choco2.1.1 *2009-06-10*	508	74%	82%

Results for 2-ARY-INT

Rank	Solver	#solved	%inst.	%VBS
	Total number of instances: 686			
	Virtual Best Solver (VBS)	617	90%	100%
1	Abscon 112v4 *ESAC*	517	75%	84%
2	Abscon 112v4 *AC*	513	75%	83%
3	Mistral *1.545*	511	74%	83%
4	Choco2.1.1b *2009-07-16*	510	74%	83%
5	Choco2.1.1 *2009-06-10*	508	74%	82%
6	Sugar *v1.14.6+picosat*	479	70%	78%
7	Sugar *v1.14.6+minisat*	470	69%	76%
8	Concrete *2009-07-14*	428	62%	69%
9	pcs-restart *0.3.2*	419	61%	68%
10	pcs *0.3.2*	419	61%	68%
11	Concrete DC *2009-07-14*	372	54%	60%
12	bpsolver *09*	349	51%	57%
13	SAT4J CSP *2.1.1*	306	45%	50%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 699				
Virtual Best Solver (VBS)		612	88%	100%

Results for N-ARY-EXT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 699			
	Virtual Best Solver (VBS)	612	88%	100%
3	Concrete *2009-07-14*	544	78%	89%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 699				
Virtual Best Solver (VBS)		612	88%	100%
2	Abscon 112v4 *AC*	545	78%	89%
3	Concrete *2009-07-14*	544	78%	89%

Results for N-ARY-EXT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 699				
Virtual Best Solver (VBS)		612	88%	100%
1	Mistral *1.545*	585	84%	96%
2	Abscon 112v4 *AC*	545	78%	89%
3	Concrete *2009-07-14*	544	78%	89%

Results for N-ARY-EXT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 699				
Virtual Best Solver (VBS)		612	88%	100%
1	Mistral *1.545*	585	84%	96%
2	Abscon 112v4 *AC*	545	78%	89%
3	Concrete *2009-07-14*	544	78%	89%
4	Abscon 112v4 *ESAC*	543	78%	89%
5	Conquer *2009-07-10*	532	76%	87%
6	Choco2.1.1 *2009-06-10*	532	76%	87%
7	Concrete DC *2009-07-14*	532	76%	87%
8	Choco2.1.1b *2009-07-16*	528	76%	86%
9	pcs *0.3.2*	497	71%	81%
10	pcs-restart *0.3.2*	496	71%	81%
11	bpsolver *09*	394	56%	64%
12	Sugar *v1.14.6+minisat*	374	54%	61%
13	Sugar *v1.14.6+picosat*	350	50%	57%
14	SAT4J CSP *2.1.1*	209	30%	34%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 709				
Virtual Best Solver (VBS)		630	89%	100%

Results for N-ARY-INT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 709				
	Virtual Best Solver (VBS)	630	89%	100%
4	pcs-restart *0.3.2*	546	77%	87%
5	pcs *0.3.2*	542	76%	86%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 709				
Virtual Best Solver (VBS)		630	89%	100%
2	Choco2.1.1 *2009-06-10*	560	79%	89%
3	Choco2.1.1b *2009-07-16*	548	77%	87%
4	pcs-restart *0.3.2*	546	77%	87%
5	pcs *0.3.2*	542	76%	86%

Results for N-ARY-INT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 709				
Virtual Best Solver (VBS)		630	89%	100%
1	Mistral *1.545*	572	81%	91%
2	Choco2.1.1 *2009-06-10*	560	79%	89%
3	Choco2.1.1b *2009-07-16*	548	77%	87%
4	pcs-restart *0.3.2*	546	77%	87%
5	pcs *0.3.2*	542	76%	86%

Results for N-ARY-INT

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 709				
Virtual Best Solver (VBS)		630	89%	100%
1	Mistral *1.545*	572	81%	91%
2	Choco2.1.1 *2009-06-10*	560	79%	89%
3	Choco2.1.1b *2009-07-16*	548	77%	87%
4	pcs-restart *0.3.2*	546	77%	87%
5	pcs *0.3.2*	542	76%	86%
6	bpsolver *09*	513	72%	81%
7	Abscon 112v4 *ESAC*	489	69%	78%
8	Abscon 112v4 *AC*	481	68%	76%
9	Sugar *v1.14.6+minisat*	481	68%	76%
10	Sugar *v1.14.6+picosat*	478	67%	76%
11	Concrete *2009-07-14*	439	62%	70%
12	Concrete DC *2009-07-14*	342	48%	54%
13	SAT4J CSP *2.1.1*	171	24%	27%

Results for alldiff

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 118				
Virtual Best Solver (VBS)		107	91%	100%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 118				
Virtual Best Solver (VBS)		107	91%	100%
3	Sugar *v1.14.6+minisat*	88	75%	82%

Results for alldiff

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 118				
Virtual Best Solver (VBS)		107	91%	100%
2	Mistral *1.545*	98	83%	92%
3	Sugar *v1.14.6+minisat*	88	75%	82%

Results for alldiff

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 118				
Virtual Best Solver (VBS)		107	91%	100%
1	Sugar *v1.14.6+picosat*	104	88%	97%
2	Mistral *1.545*	98	83%	92%
3	Sugar *v1.14.6+minisat*	88	75%	82%

Results for alldiff

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 118				
Virtual Best Solver (VBS)		107	91%	100%
1	Sugar *v1.14.6+picosat*	104	88%	97%
2	Mistral *1.545*	98	83%	92%
3	Sugar *v1.14.6+minisat*	88	75%	82%
4	Abscon 112v4 *ESAC*	78	66%	73%
5	Abscon 112v4 *AC*	77	65%	72%
6	Concrete *2009-07-14*	73	62%	68%
7	Choco2.1.1 *2009-06-10*	72	61%	67%
8	Choco2.1.1b *2009-07-16*	71	60%	66%
9	Concrete DC *2009-07-14*	68	58%	64%
10	SAT4J CSP *2.1.1*	65	55%	61%
11	bpsolver *09*	60	51%	56%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 276				
Virtual Best Solver (VBS)		249	90%	100%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 276				
Virtual Best Solver (VBS)		249	90%	100%
4	Choco2.1.1 *2009-06-10*	194	70%	78%
5	Choco2.1.1b *2009-07-16*	193	70%	78%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 276				
Virtual Best Solver (VBS)		249	90%	100%
3	Mistral *1.545*	217	79%	87%
4	Choco2.1.1 *2009-06-10*	194	70%	78%
5	Choco2.1.1b *2009-07-16*	193	70%	78%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 276				
Virtual Best Solver (VBS)		249	90%	100%
1	Sugar *v1.14.6+picosat*	229	83%	92%
2	Sugar *v1.14.6+minisat*	222	80%	89%
3	Mistral *1.545*	217	79%	87%
4	Choco2.1.1 *2009-06-10*	194	70%	78%
5	Choco2.1.1b *2009-07-16*	193	70%	78%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 276				
Virtual Best Solver (VBS)		249	90%	100%
1	Sugar *v1.14.6+picosat*	229	83%	92%
2	Sugar *v1.14.6+minisat*	222	80%	89%
3	Mistral *1.545*	217	79%	87%
4	Choco2.1.1 *2009-06-10*	194	70%	78%
5	Choco2.1.1b *2009-07-16*	193	70%	78%
6	bpsolver *09*	186	67%	75%
7	Abscon 112v4 *AC*	113	41%	45%
8	Abscon 112v4 *ESAC*	95	34%	38%

Rank	Solver	#solved	%inst.	%VBS
	Total number of instances: 162			
	Virtual Best Solver (VBS)	144	89%	100%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 162				
Virtual Best Solver (VBS)		144	89%	100%
5	Mistral *1.545*	120	74%	83%

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 162				
Virtual Best Solver (VBS)		144	89%	100%
3	Choco2.1.1 *2009-06-10*	125	77%	87%
4	Choco2.1.1b *2009-07-16*	122	75%	85%
5	Mistral *1.545*	120	74%	83%

Results for alldiff/element/weightedsum/cumulative

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 162				
Virtual Best Solver (VBS)		144	89%	100%
1	Sugar *v1.14.6+minisat*	136	84%	94%
2	Sugar *v1.14.6+picosat*	135	83%	94%
3	Choco2.1.1 *2009-06-10*	125	77%	87%
4	Choco2.1.1b *2009-07-16*	122	75%	85%
5	Mistral *1.545*	120	74%	83%

Results for alldiff/element/weightedsum/cumulative

Rank	Solver	#solved	%inst.	%VBS
Total number of instances: 162				
Virtual Best Solver (VBS)		144	89%	100%
1	Sugar *v1.14.6+minisat*	136	84%	94%
2	Sugar *v1.14.6+picosat*	135	83%	94%
3	Choco2.1.1 *2009-06-10*	125	77%	87%
4	Choco2.1.1b *2009-07-16*	122	75%	85%
5	Mistral *1.545*	120	74%	83%
6	bpsolver *09*	111	69%	77%

Website

The competition is designed to be as transparent as possible.

The website http://www.cril.univ-artois.fr/CSC09

- contains all results
- presents the trace of each solver execution
- gives many more details (selection of instances, ...)
- contains the archive of the instances that were used
- contains a description of the solvers