

Third International CSP Solver Competition CPAI'08

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

CP 2008: 14th International Conference on
Principles and Practice of Constraint Programming

September, 2008

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

Goal of the competition

Help identifying successful techniques in constraint solving

- by comparing solvers in the same environment,
- 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
- comparing one technique with another by implementing them in a single solver is not fair (no guarantee that the implementation has the same efficiency as the original implementation)
- CPU time is the final measure of efficiency for mono-process solvers

Different problems

- CSP: decision problem
find a complete instantiation of variables which satisfies each constraint
- Max-CSP: optimisation problem
find a complete instantiation of variables which maximizes the number of satisfied constraints
- WCSP: optimisation problem
find a complete instantiation of variables with a minimal cost

Representation of instances

Requirements:

- Solvers must be compared on the very same representation of instances
- The representation must be as simple as possible in order to focus on the core features of the solvers: solving constraints

Unfortunately,

- the high expressivity of CSP disallows a simple representation (as in SAT)
- no single, widely used format used in the CSP community

Since 2005, the competition has

- used the simplest possible formats to reach the goal of the competition
- helped contestants by providing them with parsers

This year, the XCSP 2.1 format was used (based on XML).

- based on XML
 - inherits all features of XML
 - easily extendable
 - both human and computer readable (abridged version and pure XML version)
- extensional constraints
- intensional constraints
- global constraints (direct translation from the catalog of global constraints)
- weighted CSP
- quantified CSP

Will evolve to add features currently unused in the competition

XCSP 2.1: example

```
<instance>
  <presentation name="example" format="XCSP 2.1">
    A dummy example
  </presentation>

  <domains nbDomains="1">
    <domain name="D1" nbValues="10"> 1..5 11..15 20 </domain>
  </domains>

  <variables nbVariables="1">
    <variable name="X1" domain="D1"/>
  </variables>

  <relations nbRelations="1">
    <relation name="R1" arity="2" nbTuples="5" semantics="conflicts">
      1 1|1 2|1 3|1 4|1 5
    </relation>
  </relations>
```

XCSP 2.1: example

```
<predicates nbPredicates="1">
  <predicate name="equals">
    <parameters>int A int B</parameters>
    <expression>
      <functional>eq(A,B)</functional>
    </expression>
  </predicate>
</predicates>

<constraints nbConstraints="3">
  <constraint name="C1" arity="2" scope="X1 X2" reference="R1"/>
  <constraint name="C2" arity="2" scope="X1 X2" reference="equals">
    <parameters>X1 X2</parameters>
  </constraint>

  <constraint name="C3" arity="2" scope="X1 X2"
    reference="global:element">
    <parameters>X1 [4 7 22 60] X2</parameters>
  </constraint>
</constraints>
</instance>
```


Classification of solvers

Two categories of solvers:

- complete solvers: can prove unsatisfiability (or optimality)
- incomplete solvers: cannot prove unsatisfiability (nor optimality) (e.g. local search)

Solvers may not have support for all kinds of constraints:

- binary or non-binary constraints
- extension or intension
- global constraints

Solvers are registered in the categories they support.

Instances categories

Five categories were 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

GLOBAL some global constraints

Judges and working group

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

- Pierre FLENER, Upsala University, Sweden
- Richard WALLACE, Cork Constraint Computation Centre, Ireland

A working group was in charge of generating/collecting instances

- Emmanuel HEBRARD, Cork Constraint Computation Centre (4C), Ireland
- Barry O'SULLIVAN, Cork Constraint Computation Centre (4C), Ireland
- Andrea RENDL, School of Computer Science, University of St Andrews, UK
- Sebastien TABARY, Centre de Recherches en Informatique de Lens, Université d'Artois, France

Selection of Instances

- Deepak MEHTA was asked by the judges to select the instances for the competition.
- He was given a list with the different series and number of instances per series.
- He returned the number of instances to select in each series
- Individual instances were selected by a random process in each series

14 submitted solvers (and a few more versions)

Abscon a CSP solver in Java

bpsolver a solver integrated in a Prolog engine

casper C++ library for generic constraint solving

choco a CSP library in Java

CSP4J a CSP library in Java

cpHydra a portfolio solver

galac a BDD based solver, with translation to SAT

...

CSP solvers submitted (continued)

...

mddc-solv C++, aimed at large arity extensional constraints

MDG derived from Mistral

Minion/Tailor a C++ solver

Mistral a C++ constraint library

SAT4J CSP based on a translation to SAT

spider derived from CPlan

Sugar a SAT based solver

Max-CSP and WCSP solvers submitted

Max-CSP: 4 submitted solvers (and a few more versions)

`AbsconMax` a CSP solver in Java

`CSP4J` a CSP library in Java

`Sugar` a SAT based solver

`toulbar2` a WCSP solver

WCSP: only one solver submitted (Toulbar)

- the competition has been postponed

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)
- CSP solvers were given a time limit of 30 minutes (1800s).
- Max-CSP solvers were given a time limit of 1 hour (3600s).
- 515 days of CPU time used by CSP solvers
204 days of CPU time used by Max-CSP solvers

Verification of results

- 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
 - for OPTIMUM FOUND answers, the system only checks that no other solver found a better solution
- UNSATISFIABLE and OPTIMUM FOUND 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

Part I

Results of the CSP competition

Results for 2-ARY-EXT

622 instances solved out of 635

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	622	100%

Results for 2-ARY-EXT

622 instances solved out of 635

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	622	100%
1	cpHydra <i>k_10</i>	574	92%

Results for 2-ARY-EXT

622 instances solved out of 635

Rank	Solver	#solved	% of VBS
<i>Virtual Best Solver (VBS)</i>		622	100%
1	cpHydra <i>k_10</i>	574	92%
2	cpHydra <i>k_40</i>	567	91%

Results for 2-ARY-EXT

622 instances solved out of 635

Rank	Solver	#solved	% of VBS
<i>Virtual Best Solver (VBS)</i>		622	100%
1	cpHydra <i>k_10</i>	574	92%
2	cpHydra <i>k_40</i>	567	91%
3	MDG-probe	561	90%

Results for 2-ARY-EXT

622 instances solved out of 635

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	622	100%
1	cpHydra <i>k_10</i>	574	92%
2	cpHydra <i>k_40</i>	567	91%
3	MDG-probe	561	90%
4	Mistral-option <i>1.314</i>	559	90%
5	MDG-noprobe	558	90%
6	Mistral-prime <i>1.313</i>	556	89%
7	choco2_dwdeg	554	89%
8	choco2_impwdeg	550	88%
9	Abscon <i>ESAC</i>	546	88%
10	Abscon <i>AC</i>	546	88%

Results for 2-ARY-EXT (continued)

Rank	Solver	#solved	% of VBS
11	Concrete + CSP4J	498	80%
12	bpsolver	495	80%
13	spider	472	76%
14	Sugar <i>v1.13+minisat</i>	470	76%
15	Sugar <i>v1.13+picosat</i>	443	71%
16	casper <i>zito</i>	433	70%
17	SAT4J CSP	426	68%
18	galac <i>E</i>	425	68%
19	galac <i>M</i>	422	68%
20	casper <i>zao</i>	415	67%
21	Minion/Tailor	414	67%
22	Concrete + CSP4J - Tabu	348	56%

Results for 2-ARY-INT

634 instances solved out of 696

Rank	Solver	#solved	% of VBS
<i>Virtual Best Solver (VBS)</i>		634	100%
1	cpHydra <i>k_40</i>	597	94%
2	cpHydra <i>k_10</i>	596	94%
3	choco2_dwdeg	523	82%

Results for 2-ARY-INT

634 instances solved out of 696

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	634	100%
1	cpHydra <i>k_40</i>	597	94%
2	cpHydra <i>k_10</i>	596	94%
3	choco2_dwdeg	523	82%
4	Mistral-prime <i>1.313</i>	522	82%
5	choco2_impwdeg	521	82%
6	Abscon 112v4 <i>ESAC</i>	520	82%
7	MDG-probe	515	81%
8	Abscon 112v4 <i>AC</i>	513	81%
9	MDG-noprobe	499	79%
10	Sugar <i>v1.13+picosat</i>	486	77%

Results for 2-ARY-INT (continued)

Rank	Solver	#solved	% of VBS
11	Sugar <i>v1.13+minisat</i>	484	76%
12	Mistral-option <i>1.314</i>	451	71%
13	Concrete + CSP4J	415	65%
14	bpsolver	375	59%
15	SAT4J CSP	309	49%
16	Concrete + CSP4J - WMC	222	35%
17	Concrete + CSP4J - Tabu	183	29%

Results for N-ARY-EXT

607 instances solved out of 704

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	607	100%
1	cpHydra <i>k_40</i>	587	97%

Results for N-ARY-EXT

607 instances solved out of 704

Rank	Solver	#solved	% of VBS
<i>Virtual Best Solver (VBS)</i>		607	100%
1	cpHydra <i>k_40</i>	587	97%
2	cpHydra <i>k_10</i>	586	97%
3	mddc-solv	576	95%

Results for N-ARY-EXT

607 instances solved out of 704

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	607	100%
1	cpHydra <i>k_40</i>	587	97%
2	cpHydra <i>k_10</i>	586	97%
3	mddc-solv	576	95%
4	MDG-noprobe	570	94%
5	Mistral-prime <i>1.313</i>	569	94%
6	Mistral-option <i>1.314</i>	565	93%
7	Abscon 112v4 <i>AC</i>	545	90%
8	Concrete + CSP4J	544	90%
9	Abscon 112v4 <i>ESAC</i>	541	89%
10	choco2_dwdeg	477	79%

Results for N-ARY-EXT (continued)

Rank	Solver	#solved	% of VBS
11	choco2_impwdeg	441	73%
12	Minion/Tailor	406	67%
13	casper <i>zito</i>	391	64%
14	casper <i>zao</i>	386	64%
15	Sugar <i>v1.13+minisat</i>	370	61%
16	Sugar <i>v1.13+picosat</i>	347	57%
17	galac <i>M</i>	326	54%
18	galac <i>E</i>	323	53%
19	bpsolver	322	53%
20	SAT4J CSP	218	36%
21	Concrete + CSP4J - Tabu	210	35%
22	Concrete + CSP4J - WMC	135	22%

Results for N-ARY-INT

660 instances solved out of 716

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	660	100%
1	cpHydra <i>k_10</i>	569	86%
2	cpHydra <i>k_40</i>	569	86%

Results for N-ARY-INT

660 instances solved out of 716

Rank	Solver	#solved	% of VBS
<i>Virtual Best Solver (VBS)</i>		660	100%
1	cpHydra <i>k_10</i>	569	86%
2	cpHydra <i>k_40</i>	569	86%
3	casper <i>zao</i>	562	85%

Results for N-ARY-INT

660 instances solved out of 716

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	660	100%
1	cpHydra <i>k_10</i>	569	86%
2	cpHydra <i>k_40</i>	569	86%
3	casper <i>zao</i>	562	85%
4	Mistral-prime <i>1.313</i>	560	85%
5	MDG-probe	560	85%
6	casper <i>zito</i>	555	84%
7	Mistral-option <i>1.314</i>	554	84%
8	MDG-noprobe	530	80%
9	choco2_impwdeg	518	78%
10	Abscon 112v4 <i>ESAC</i>	488	74%

Results for N-ARY-INT (continued)

Rank	Solver	#solved	% of VBS
11	Abscon 112v4 AC	486	74%
12	Sugar v1.13+minisat	486	74%
13	Sugar v1.13+picosat	481	73%
14	bpsolver	459	70%
15	Concrete + CSP4J	456	69%
16	Concrete + CSP4J - Tabu	199	30%
17	SAT4J CSP 2008-06-13	174	26%
18	Concrete + CSP4J - WMC	174	26%

Results for GLOBAL

501 instances solved out of 556

Warning: some solver may not have support for some global constraints!

Rank	Solver	#solved	% of VBS
<i>Virtual Best Solver (VBS)</i>		501	100%
1	Sugar v1.13+picosat	424	85%
2	cpHydra k_40	420	84%
3	cpHydra k_10	419	84%

Results for GLOBAL

501 instances solved out of 556

Warning: some solver may not have support for some global constraints!

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	501	100%
1	Sugar v1.13+picosat	424	85%
2	cpHydra k_40	420	84%
3	cpHydra k_10	419	84%
4	Sugar v1.13+minisat	405	81%
5	Mistral-prime 1.313	403	80%
6	casper zito	397	79%
7	casper zao	390	78%
8	Mistral-option 1.314	383	76%
9	choco2_dwdeg	358	71%
10	MDG-noprobe	353	70%

Results for GLOBAL (continued)

Rank	Solver	#solved	% of VBS
11	choco2_impwdeg	347	69%
12	bpsolver	347	69%
13	MDG-probe	337	67%
14	Minion/Tailor	216	43%
15	Abscon 112v4 <i>AC</i>	184	37%
16	Abscon 112v4 <i>ESAC</i>	173	35%
17	SAT4J CSP	63	13%

A few points on cpHydra

cpHydra

- is a portfolio solver
- was trained on instances available before the start of the competition
- is the great winner of this competition

But the number of instances which were kept hidden was small in this competition.

Two experiments show that cpHydra is really a strong solver

- when only hidden instances are considered, cpHydra still has a good ranking
- when tested on modified instances (for each variable X , a clone X' is added and $X=X'$ is enforced), cpHydra has very similar results.

Part II

Results of the Max-CSP competition

Results for 2-ARY-EXT

440 instances were solved out of 534

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	440	100%

Results for 2-ARY-EXT

440 instances were solved out of 534

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	440	100%
1	toulbar2	412	94%

Results for 2-ARY-EXT

440 instances were solved out of 534

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	440	100%
1	toulbar2	412	94%
2	Sugar v1.13+minisat	240	55%

Results for 2-ARY-EXT

440 instances were solved out of 534

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	440	100%
1	toulbar2	412	94%
2	Sugar v1.13+minisat	240	55%
3	Sugar++ v1.13+minisat-inc	229	52%

Results for 2-ARY-EXT

440 instances were solved out of 534

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	440	100%
1	toulbar2	412	94%
2	Sugar v1.13+minisat	240	55%
3	Sugar++ v1.13+minisat-inc	229	52%
4	AbsconMax 112 pc-w	121	28%
5	AbsconMax 112 pc-d	86	20%
6	Concrete + CSP4J - Tabu	32	7%
7	Concrete + CSP4J - MCRW	32	7%

Results for 2-ARY-INT

103 instances were solved out of 276

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	103	100%

Results for 2-ARY-INT

103 instances were solved out of 276

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	103	100%
1	<i>Sugar v1.13+minisat</i>	101	98%

Results for 2-ARY-INT

103 instances were solved out of 276

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	103	100%
1	<i>Sugar v1.13+minisat</i>	101	98%
2	<i>Sugar++ v1.13+minisat-inc</i>	99	96%

Results for 2-ARY-INT

103 instances were solved out of 276

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	103	100%
1	Sugar v1.13+minisat	101	98%
2	Sugar++ v1.13+minisat-inc	99	96%
3	AbsconMax 112 pc-w	32	31%

Results for 2-ARY-INT

103 instances were solved out of 276

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	103	100%
1	Sugar v1.13+minisat	101	98%
2	Sugar++ v1.13+minisat-inc	99	96%
3	AbsconMax 112 pc-w	32	31%
4	AbsconMax 112 pc-d	30	29%
5	Concrete + CSP4J - MCRW	0	0%
6	Concrete + CSP4J - Tabu	0	0%

Results for N-ARY-EXT

171 instances were solved out of 278

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	171	100%

Results for N-ARY-EXT

171 instances were solved out of 278

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	171	100%
1	toulbar2/BTD	153	89%

Results for N-ARY-EXT

171 instances were solved out of 278

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	171	100%
1	toulbar2/BTD	153	89%
2	Sugar v1.13+minisat	118	69%

Results for N-ARY-EXT

171 instances were solved out of 278

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	171	100%
1	toulbar2/BTD	153	89%
2	Sugar v1.13+minisat	118	69%
3	Sugar++ v1.13+minisat-inc	118	69%

Results for N-ARY-EXT

171 instances were solved out of 278

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	171	100%
1	toulbar2/BTD	153	89%
2	Sugar v1.13+minisat	118	69%
3	Sugar++ v1.13+minisat-inc	118	69%
4	AbsconMax 112 pc-w	103	60%
5	AbsconMax 112 pc-d	71	42%
6	Concrete + CSP4J - Tabu	4	2%
7	Concrete + CSP4J - MCRW	4	2%

Results for N-ARY-INT

42 instances were solved out of 109

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	42	100%

Results for N-ARY-INT

42 instances were solved out of 109

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	42	100%
1	<i>Sugar v1.13+minisat</i>	39	93%

Results for N-ARY-INT

42 instances were solved out of 109

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	42	100%
1	Sugar v1.13+minisat	39	93%
2	Sugar++ v1.13+minisat-inc	39	93%

Results for N-ARY-INT

42 instances were solved out of 109

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	42	100%
1	Sugar v1.13+minisat	39	93%
2	Sugar++ v1.13+minisat-inc	39	93%
3	AbsconMax 112 pc-w	14	33%

Results for N-ARY-INT

42 instances were solved out of 109

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	42	100%
1	Sugar v1.13+minisat	39	93%
2	Sugar++ v1.13+minisat-inc	39	93%
3	AbsconMax 112 pc-w	14	33%
4	AbsconMax 112 pc-d	10	24%
5	Concrete + CSP4J - MCRW	0	0%
6	Concrete + CSP4J - Tabu	0	0%

Results for GLOBAL

65 instances were solved out of 98

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	65	100%

Results for GLOBAL

65 instances were solved out of 98

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	65	100%
1	Sugar v1.13+minisat	65	100%

Results for GLOBAL

65 instances were solved out of 98

Rank	Solver	#solved	% of VBS
	<i>Virtual Best Solver (VBS)</i>	65	100%
1	Sugar v1.13+minisat	65	100%
2	Sugar++ v1.13+minisat-inc	50	77%

The competition is designed to be as transparent as possible.

The website <http://www.cril.univ-artois.fr/CPAI08>

- 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