PB24
Pseudo-Boolean Competition 2024

organized by Olivier Roussel
kindly presented by **Jeremias Berg**,
steering committee:
Carlos Ansótegui, Johannes Fichte, Matti Järvisalo,
Jakob Nordström, Olivier Roussel

27th International Conference
on Theory and Applications of Satisfiability Testing,
SAT'27

August 24th, 2024

A new competition 8 years after the last one...

Why the competition stopped after 2016:

- lack of interest (very few solvers submitted in 2016)
- no one decided to organize a new competition
- ...

Why it restarted this year:

- because Jakob Nordström has insisted for two years to have a new competition!
- because the organizer had forgotten how much work was required...
- ...

All details are at the usual URL: **https://www.cril.univ-artois.fr/PB24/**

New this year

- A steering committee for deciding the rules: Carlos Ansótegui, Johannes Fichte, Matti Järvisalo, Jakob Nordström, Olivier Roussel
- Removal of the clearly outdated small/big integers distinction (decided in 2005, when most solvers still used 32 bits integers). Now, a solver decides by its own if it cannot handle an instance and answers 's UNSUPPORTED'. Generation of two separate rankings: ranking of solvers that handle all instances and ranking on the set of instances supported by all solvers.
- Tracks for solvers that generate proofs of unsatisfiability/optimality, checked by VeriPB and VeriPB+CakePB.
- Clarification of the PB format:
  - restricted format for the competition (no significant change)
  - general format with as few restrictions as possible (work still in progress)

Some numbers

- 10 teams,
- 13 submitters,
- 15 solvers (same base name),
- 35 solver variants (different algorithms/arguments),
- 46 different solver versions (bug fixes,...).
- only 10 instances not supported by one or more solvers (because of too large integers). Not useful to generate a separate ranking, those instances were removed from the selection at the very end.
- approximately 7 years of CPU time!

List of solvers

- CASHWMaxSAT-* (Shiwei Pan; Yujiao Zhao; Yiyuan Wang; Minghao Yin)
- DLS-PBO (Zhihan Chen, Peng Lin, Yi Chu, and Shaowei Cai)
- Exact* (Jo Devriendt)
- FiberSCIP (Gioni Mexi, Shanwen Pu, Yuji Shinano, Thorsten Koch)
- Hybrid-* (Yiyuan Wang; Shiwei Pan; Yujiao Zhao; Minghao Yin)
- IPBHS-* (Hannes Ihalainen, Jeremias Berg, Matti Järvisalo)
- LS-Mab* (Yujiao Zhao; Shiwei Pan; Yiyuan Wang; Minghao Yin)
- mixed-bag (Christoph Jabs, Jeremias Berg, Matti Järvisalo)
- ...

List of solvers (continued)

- ...
- NaPS* (Masahiko Sakai and Hidetomo Nabeshima)
- ParLS-PBO (Zhihan Chen, Peng Lin, Yi Chu, and Shaowei Cai)
- pb-oll-rs (Christoph Jabs, Jeremias Berg, Matti Järvisalo)
- Picat (Neng-Fa Zhou)
- RoundingSat* (Jan Elffers, Jo Devriendt, Stephan Gocht, Jakob Nordström, Andy Oertel, Marc Vinyals)
- Sat4j* (Daniel Le Berre and Romain Wallon)
- SCIP (Gioni Mexi, Shanwen Pu, Julian Manns, Marc Pfetsch, Thorsten Koch, Christopher Hojny, Alexander Hoen, Dominik Kamp, Matthias Walter, Ksenia Bestuzheva)

Pseudo-Boolean Constraints

- **Linear** (LIN) pseudo-Boolean (PB) constraint = sum of integer $\times$ a literal $\{\geq, \leq, =\}$ constant
  Example: $3x_1 - 3x_2 + 2\bar{x}_3 + \bar{x}_4 + x_5 \geq 5$
- **Non-linear** (NLC) pseudo-Boolean (PB) constraint = sum of integer $\times$ a product of literals $\{\geq, \leq, =\}$ constant
  Example: $3x_1\bar{x}_2 - 3x_2x_4 + 2\bar{x}_3 + \bar{x}_4 + x_5x_6x_7 \geq 5$
- As an example, PB allows compact encodings of:
  - cardinalities: $x_1 + x_2 + x_3 \geq 2$
  - adder (C=A+B): $2c_1 + c_0 = 2a_1 + a_0 + 2b_1 + b_0$
  - knapsack:
    $max : 5x_1 + 10x_2 + 2x_3;$
    $5x_1 + 8x_2 + x_3 \leq 10$
  - integer factorization (X*Y=3): $x_0y_0 + 2x_0y_1 + 2x_1y_0 + 4x_1y_1 = 3$
- Cutting-planes proof system stronger than resolution: PHP easily solved in polynomial time

Benchmark categories (1)

Classification based on the linearity of constraints

- **LIN** All constraints are linear
- **NLC** At least one constraint is non linear (contains products of literals)

Classification based on the objective function

- **DEC** No objective function to optimize (decision problem). The solver must simply find a solution.
- **OPT** An objective function is present. The solver must find a solution with the best possible value of the objective function.

Classification based on the existence of hard/soft clauses (generalization of MaxSAT with a top cost as in WCSP)

- **SOFT** All constraints may be violated if needed (no hard constraint). Minimize the weight of unsatisfied constraints.
- **PARTIAL** At least one hard constraint.

Selection of instances

- A domain is a set of similar instances. Basically, one directory in an archive equals one domain, but some directories must be merged into a single domain.
- Past competition instances classified by Johannes Fichte (thanks!) based on data in Pavel Smirnov's thesis.
- New instances classified according to the name of the directory.
- 10 instances randomly selected per domain (15 instances per domain in the DEC-LIN track).
- If a domain doesn't contain enough instances, they are all selected.
- At most 5 % of the whole instance set selected per submitter.
- Duplicates removed.

New instances

- Jo Devriendt: 165 instances
- Hannes Ihalainen: 110 instances
- Jakob Nordström: 15650 instances (3 domains were considered to be well known instances, many instances were merged into preexisting domains, the rest was subject to the limit of 5% of the whole instance set)
- William Pettersson: 1490 instances

Tracks

Listed by decreasing number of registered solvers:

- **OPT-LIN: 37 solvers, 478 instances selected**
- **DEC-LIN: 33 solvers, 397 instances selected**
- **PARTIAL-LIN: 11 solvers, 208 instances selected (no new instances)**
- SOFT-LIN: 11 solvers, 60 instances selected (no new instances)
  not enough instances, ranking not presented
- OPT-NLC: 9 solvers, 54 instances selected (no new instances)
  not enough instances, ranking not presented
- DEC-NLC: 9 solvers, 10 instances selected (no new instances)
  not enough instances, ranking not presented
- **DEC-LIN-CERT and OPT-LIN-CERT: 4 solvers, same selection as in DEC-LIN/OPT-LIN, solvers also ranked in DEC-LIN/OPT-LIN**

Evaluation environment

kindly provided by CRIL, University of Artois, France

- Each solver was given a time limit of 1 hour CPU time (20 hours CPU time for the 2 parallel solvers, which ran on 20 cores)
- The size limit for proofs of unsatisfiability/optimality was set initially to 100 GB, and later extended to 400 GB.
- VeriPB and VeriPB+CakePB were used to verify the proofs and were allowed to run for 5 hours (CPU time).
- Cluster of bi-CPU Intel Xeon E5-2637 v4 3.5Ghz 4 cores, 128 GB RAM *4 jobs per host, 2 cores/job, 31 GB RAM per job* This configuration was tested on a sample of benchmarks and instances and induces a slowdown of at most 10% compared to a solver running alone on the host. *It was the only configuration that allowed to complete the experiments in time.*
- For the 2 parallel solvers: nodes quad-CPU Intel Xeon Gold 6248 2.5Ghz 20 cores, 768 GB RAM, *4 jobs per host, 20 cores/job, up to 195GB/job*

Verification of results

- The usual verifications are performed:
  - the models given by the solvers are checked
  - the answers given by the different solvers on a given instance are checked for consistency
- [new this year] the unsatifiability/optimality proofs in the *-CERT tracks where checked by both VeriPB and VeriPB+CakePB.
  Two new answers OPTC and UNSC (C for Certified)
- Solvers giving a wrong answer in a category are disqualified in that category. Up to 3 submissions of bug fix were allowed.
- Not all UNSATISFIABLE and OPTIMUM FOUND answers could be checked and therefore some results should be taken with caution.

Ranking of solvers and Virtual Best Solver (VBS)

Main ranking (targeting complete solvers) based on two criteria:

1. the number of solved instances
2. ties are broken by considering the cumulated time on solved instances

A few other rankings targeting incomplete solvers are available on the web site (no perfect solution though!).

The Virtual Best Solver (VBS)

- is the virtual solver obtained by combining the best results of all submitted solvers.
- **can 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 DEC-LIN

OPTC= OPT Certified, UNSC= UNSAT Certified, Ⓘ= incomplete solver

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 397					
	Virtual Best Solver (VBS)	362	147 SAT, 215 UNS	91%	100%
1	Hybrid-CASH...CadSP+Exact	312	126 SAT, 186 UNS	79%	86%
2	Hybrid-CASH...ComS+Exact	309	123 SAT, 186 UNS	78%	85%
3	Exact_veripb2	291	108 SAT, 8 UNS, 175 UNSC	73%	80%
4	RoundingSat *log*	288	99 SAT, 189 UNSC	73%	80%
5	Exact	286	102 SAT, 184 UNS	72%	79%
6	Exact_veripb2_no_SoPlex	284	112 SAT, 7 UNS, 165 UNSC	72%	78%
7	RoundingSat *nolog*	282	98 SAT, 184 UNS	71%	78%
8	Exact_no_SoPlex	275	104 SAT, 171 UNS	69%	76%
9	CASHWMaxSAT-DisjCad-S	273	124 SAT, 149 UNS	69%	75%
10	NaPS	273	122 SAT, 151 UNS	69%	75%
11	Picat	273	123 SAT, 150 UNS	69%	75%
12	FiberSCIP *20 cores*	270	101 SAT, 169 UNS	68%	75%
13	CASHWMaxSAT-DisjCad-SP	268	119 SAT, 149 UNS	68%	74%

Results for DEC-LIN (continued)

Rank	Solver	#solved	Detail	%inst.	%VBS
14	CASHWMaxSAT-DisjCom-S	267	113 SAT, 154 UNS	67%	74%
15	CASHWMaxSAT-DisjCom-SP	267	113 SAT, 154 UNS	67%	74%
16	Sat4j Res. Default	246	104 SAT, 142 UNS	62%	68%
17	Sat4j Res.	244	102 SAT, 1 UNS, 141 UNSC	61%	67%
18	NaPS-GM	241	91 SAT, 150 UNS	61%	67%
19	SCIP	241	80 SAT, 161 UNS	61%	67%
20	NaPS-PB16	236	86 SAT, 150 UNS	59%	65%
21	FiberSCIP *1 core*	235	76 SAT, 159 UNS	59%	65%
22	Sat4j Partial RoundingSAT	217	68 SAT, 149 UNS	55%	60%
23	Sat4j Cutting Planes	199	60 SAT, 139 UNS	50%	55%
24	ParLS-PBO *20 cores* Ⓘ	134	124 SAT, 10 UNS	34%	37%
25	Hybrid-LSMabSP+CASH...Com Ⓘ	130	130 SAT	33%	36%
26	ParLS-PBO *1 core* Ⓘ	122	112 SAT, 10 UNS	31%	34%
27	DLS-PBO Ⓘ	121	111 SAT, 10 UNS	30%	33%
28	Hybrid-LSMabP+Exact Ⓘ	117	117 SAT	29%	32%

Results for DEC-LIN (continued)

Rank	Solver	#solved	Detail	%inst.	%VBS
29	Hybrid-LSMabSP+Exact Ⓘ	115	115 SAT	29%	32%
30	LS-MabSP Ⓘ	112	112 SAT	28%	31%
31	LS-MabS Ⓘ	110	110 SAT	28%	30%
32	LS-MabP Ⓘ	108	108 SAT	27%	30%
33	LS-Mab Ⓘ	105	105 SAT	26%	29%

Results for OPT-LIN

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 478					
	Virtual Best Solver (VBS)	339	320 OPT, 19 UNS	71%	100%
1	mixed-bag	279	266 OPT, 13 UNS	58%	82%
2	SCIP	263	248 OPT, 15 UNS	55%	78%
3	Hybrid-CASH...CadSP+Exact	259	245 OPT, 14 UNS	54%	76%
4	IPBHS-GUROBI	257	244 OPT, 13 UNS	54%	76%
5	RoundingSat *nolog*	256	243 OPT, 13 UNS	54%	76%
6	RoundingSat *log*	254	23 OPT, 218 OPTC, 13 UNSC	53%	75%
7	Hybrid-CASH...ComS+Exact	254	241 OPT, 13 UNS	53%	75%
8	Exact	252	238 OPT, 14 UNS	53%	74%
9	FiberSCIP *1 core*	245	230 OPT, 15 UNS	51%	72%
10	IPBHS-SCIP	245	232 OPT, 13 UNS	51%	72%
11	FiberSCIP *20 cores*	245	230 OPT, 15 UNS	51%	72%
12	Exact_veripb2	239	56 OPT, 169 OPTC 1 UNS, 13 UNSC	50%	71%

Results for OPT-LIN (continued)

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 478					
Virtual Best Solver (VBS)		339	320 OPT, 19 UNS	71%	100%
13	CASHWMaxSAT-DisjCad-SP	235	222 OPT, 13 UNS	49%	69%
14	CASHWMaxSAT-DisjCad-S	235	223 OPT, 12 UNS	49%	69%
15	CASHWMaxSAT-DisjCom-S	232	220 OPT, 12 UNS	49%	68%
16	CASHWMaxSAT-DisjCom-SP	231	218 OPT, 13 UNS	48%	68%
17	Exact_no_SoPlex	226	213 OPT, 13 UNS	47%	67%
18	NaPS	208	196 OPT, 12 UNS	44%	61%
19	pb-oll-rs	208	196 OPT, 12 UNS	44%	61%
20	NaPS-GM	204	191 OPT, 13 UNS	43%	60%
21	NaPS-PB16	203	190 OPT, 13 UNS	42%	60%
22	Picat	201	187 OPT, 14 UNS	42%	59%
23	Exact_veripb2_no_SoPlex	195	38 OPT, 144 OPTC 1 UNS, 12 UNSC	41%	58%
24	Sat4j Partial RoundingSAT	146	134 OPT, 12 UNS	31%	43%
25	Sat4j Cutting Planes	133	123 OPT, 10 UNS	28%	39%

Results for OPT-LIN (continued)

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 478					
Virtual Best Solver (VBS)		339	320 OPT, 19 UNS	71%	100%
26	Sat4j Res. Default	128	117 OPT, 11 UNS	27%	38%
27	Sat4j Res.	126	2 OPT, 118 OPTC, 6 UNSC	26%	37%
28	ParLS-PBO *1 core* Ⓘ	8	8 UNS (incomplete)	2%	2%
29	DLS-PBO Ⓘ	8	8 UNS (incomplete)	2%	2%
30	ParLS-PBO *20 cores* Ⓘ	8	8 UNS (incomplete)	2%	2%
31	LS-MabSP Ⓘ	0	(incomplete)	0%	0%
32	LS-MabS Ⓘ	0	(incomplete)	0%	0%
33	LS-MabP Ⓘ	0	(incomplete)	0%	0%
34	LS-Mab Ⓘ	0	(incomplete)	0%	0%
35	Hybrid-LSMabSP+CASH...Com Ⓘ	0	(incomplete)	0%	0%
36	Hybrid-LSMabSP+Exact Ⓘ	0	(incomplete)	0%	0%
37	Hybrid-LSMabP+Exact Ⓘ	0	(incomplete)	0%	0%

Results for DEC-LIN-CERT and OPT-LIN-CERT

Excerpt of DEC-LIN:

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 397					
3	Exact_veripb2	291	108 SAT, 8 UNS, 175 UNSC	73%	80%
4	RoundingSat *log*	288	99 SAT, 189 UNSC	73%	80%
6	Exact_veripb2_no_SoPlex	284	112 SAT, 7 UNS, 165 UNSC	72%	78%
17	Sat4j Res.	244	102 SAT, 1 UNS, 141 UNSC	61%	67%

Excerpt of OPT-LIN:

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 478					
6	RoundingSat *log*	254	23 OPT, 218 OPTC, 13 UNSC	53%	75%
12	Exact_veripb2	239	56 OPT, 169 OPTC, 1 UNS, 13 UNSC	50%	71%
23	Exact_veripb2_no_SoPlex	195	38 OPT, 144 OPTC, 1 UNS, 12 UNSC	41%	58%
27	Sat4j Res.	126	2 OPT, 118 OPTC, 6 UNSC	26%	37%

Results for PARTIAL-LIN

Rank	Solver	#solved	Detail	%inst.	%VBS
Total number of instances: 208					
Virtual Best Solver (VBS)		175	175 MOPT	84%	100%
1	FiberSCIP *20 cores, 2024-07-30*	160	160 MOPT	77%	91%
2	SCIP *2024-07-07*	156	156 MOPT	75%	89%
3	Exact_no_SoPlex *2024-07-05*	156	156 MOPT	75%	89%
4	Exact *2024-07-05*	155	155 MOPT	75%	89%
5	FiberSCIP *2024-07-30*	155	155 MOPT	75%	89%
6	NaPS-PB16 *1.02b5*	147	147 MOPT	71%	84%
7	NaPS *1.03a2*	139	139 MOPT	67%	79%
8	NaPS-GM *1.03a2*	138	138 MOPT	66%	79%
9	Sat4j Resolution Default *2024-07-07*	135	135 MOPT	65%	77%
10	Sat4j Partial RoundingSAT 2024 *2024-07-12*	108	108 MOPT	52%	62%
11	Sat4j Cutting Planes 2024 *2024-07-12*	78	78 MOPT	38%	45%

Some contributions of this competition

- evaluate solvers in the same conditions
- help collecting publicly available benchmarks in one place
- help identifying new solvers and new ideas on the market
- help debug solvers and verifiers
- to our knowledge, before June 2024, no PB solver generated a proof for VeriPB v2. The competition has encouraged (maybe forced?) 3 teams to implement proof generation.

More information

- All details are on the web site **https://www.cril.univ-artois.fr/PB24/**
- Questions can be addressed to the organizer: roussel @ cril . fr.
- Thanks to all participants and submitters!
- Will there be another competition? Most probably yes.
  - When? To be determined...
  - Who will organize it? To be determined...
- Keep in mind that a competition cannot exist without both solvers and benchmarks!
- Write solvers, generate instances, be ready to submit!

Links

- PB24: https://www.cril.univ-artois.fr/PB24/
- VeriPB: https://gitlab.com/MIAOresearch/software/VeriPB
- CakePB: https://gitlab.com/MIAOresearch/software/cakepb
- Pavel Smirnov's thesis: https://helda.helsinki.fi/items/569ac68d-759b-4fd3-a145-9913479935ff