Certified Unsatisfiable Track of the SAT 2007 Competition — Tables of Final Results

Top Number: %RES proof size in MB
Second Number: abbreviated proof size ratio to %RES
Third Number: number of clauses in %RES proof
Lower Number: ratio to previous proof clauses
Legend: **F**: File exceeded 36 GB. **T**: CPU time exceeded 1 hour. **M**: Memory exceeded about 4 GB. **E**: Program Error. **P**: Proof Error.

# *pigeons*:	08	09	10	11	12	13	14	15	16	17	18	19	
THEORY													*THEORY*
	2240	5632	13824	33280	78848	184320	425984	974848	2211840	4980736	11141120	24772608	
(ratio)	2.593	2.514	2.455	2.407	2.369	2.338	2.311	2.288	2.269	2.252	2.237	2.224	(ratio)
booleforce-res	3	22	103	986	5785	**T**							booleforce-res
	0.112	0.103	0.117	0.117	0.108								
	37774	199082	787674	6061520	29150384								
(ratio)		5.270	3.957	7.695	4.809								(ratio)
picosat-res	4	48	458	**T**									picosat-res
	0.096	0.088	0.099										
	49414	409955	3185365										
(ratio)		8.296	7.770										(ratio)
picosat-rup	4	40	330	**T**									picosat-rup
	0.051	0.053	0.085										
	44542	355810	2399402										
(ratio)		7.988	6.743										(ratio)
tts-rpt	0.196	1	2	5	13	35	91	236	603	1515	3763	9248	tts-rpt
	0.200	0.167	0.141	0.120	0.104	0.090	0.079	0.069	0.061	0.055	0.049	0.044	
	2561	6401	15700	37377	88065	204801	471041	1073153	2424833	5439489	12124161	26869761	
(ratio)		2.499	2.453	2.381	2.356	2.326	2.300	2.278	2.260	2.243	2.229	2.216	(ratio)
zchaffSE07-res	1	2	8	29	92	563	**T**						zchaffSE07-res
	8493	25056	73009	215074	600936	3185589							
(ratio)		2.950	2.914	2.946	2.794	5.301							(ratio)
zchaff_0-res	1	2	8	29	92	563	**T**						zchaff_0-res
	8493	25056	73009	215074	600936	3185589							
(ratio)		2.950	2.914	2.946	2.794	5.301							(ratio)

The first section 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.

Top Number: %RES proof size in MB
Lower Number: abbreviated proof size in MB
Legend: **F**: File exceeded 36 GB. **T**: CPU time exceeded 1 hour. **M**: Memory exceeded about 4 GB. **E**: Program Error. **P**: Proof Error.

	itox_ vc965	dspam_ dump_vc973	eq.atree. braun.7	eq.atree. braun.8	eq.atree. braun.9	AProVE 07-21	AProVE 07-02	AProVE 07-22	AProVE 07-20	AProVE 07-15	dated- 5-11-u	dated- 5-15-u	total- 5-11-u	total- 5-13-u	dated- 10-15-u	manol-pipe- c10nidw_s	IBM_FV_2004_ rb30_Sd.k15	
booleforce-res	0.020	E	621	2556	19437	**T**	E	**F**	78826	156472	27785	**F**	2949	6970	3180	98	166699	booleforce-res
	0.032		21	77	361						445		58	152	45	7		
picosat-res	**P**	E	1060	7831	62506	**T**	**T**	E		E	**F**	**F**	7036	**F**	**F**	11	E	picosat-res
			38	247	1655				**F**				63			2		
picosat-rup	0.108	0.051	538	2306	**T**			**F**	**F**	**F**	**T**	**T**	**F**	**T**	**F**	21	88504	picosat-rup
	0.005	0.002	7	43	262	**T**	**T**	65	101	30	41	245	4	14	9	0.192	52	
tts-rpt	**P**	**M**	**T**	**T**	**T**													tts-rpt
zchaffSE07-res	0.112	0.055	310	2166	**T**	**T**	**T**	12589	**T**	21688	**T**	**T**	12863	**F**	36268	**T**	37337	zchaffSE07-res
zchaff_0-res	0.108	0.053	310	2166	**T**	**T**	**T**	12589	**T**	21684	**T**	**T**	12745	**F**	35735	**T**	37336	zchaff_0-res

Bench name	Competition best result	number of variables	number of clauses	max. clause length	#clauses of size 1	#clauses of size 2	#clauses of size 3	#clauses of size 4	#clauses of size 5	#clauses of size > 5	number of literals
itox_vc965	0.24	115769	338946	3	7057	118345	213544	0	0	0	884379
dspam_dump_vc973	3	274379	908191	3	5285	290999	611907	0	0	0	2423004
eq.atree.braun.7.unsat	0.80	505	1696	13	3	719	948	10	10	6	4422
eq.atree.braun.8.unsat	5	684	2300	15	3	978	1284	16	12	7	5992
eq.atree.braun.9.unsat	22	892	3006	17	3	1283	1676	18	16	10	7828
AProVE07-02	686	6196	22741	602	1	12147	9205	597	221	570	59967
AProVE07-15	46	21104	74257	171	1	36142	34380	2745	538	451	192789
AProVE07-20	218	7847	73394	4166	1	62806	5772	389	99	4327	206938
AProVE07-21	412	3189	11039	25	1	5404	5107	389	86	52	28509
AProVE07-22	81	15589	54263	96	1	26004	25780	2100	198	180	140423
dated-5-15-u	864	151952	697321	10	12754	551183	104744	3180	12720	12740	1601192
dated-10-15-u	18	193016	885873	10	16194	700175	133124	4040	16160	16180	2034196
dated-5-11-u	161	108786	482639	10	9230	395289	57400	2300	9200	9220	1096328
total-5-11-u	22	156980	696581	10	13310	570471	82900	3320	13280	13300	1582352
total-5-13-u	75	178708	806681	10	15072	648853	108896	3760	15040	15060	1842626
manol-pipe-c10nidw_s	29	404382	1204273	3	1	802848	401424	0	0	0	2809969
IBM_FV_2004_rb30_Sd.k15	121	29084	119659	12	2072	83935	21996	4194	2369	5093	303691

Experiments, Analysis and Compilation of Results by Allen Van Gelder http://www.cse.ucsc.edu/∼avg/ProofChecker with assistance from Daniel LeBerre