

Hardware Model Checking Competition Report

HWMCC'15

Armin Biere¹ and Keijo Heljanko²

¹Johannes Kepler University, Austria

²Aalto University, Finland

15th International Conference on
Formal Methods in Computer-Aided Design
FMCAD'15

The University of Texas at Austin
Austin, TX, USA

Wednesday, 30th September, 2015

added tip2014bmc data after the competition on 2nd Oct. 2015

AIGER format AVM'06 Ascona	1st HWMCC CAV'07 Berlin	2nd HWMCC CAV'08 Princeton	3rd HWMCC CAV'10 FLOC'10 Edinburgh	4th HWMCC FMCAD'11 Austin	5th HWMCC FMCAD'12 Cambridge UK	6th HWMCC FMCAD'13 Portland USA	7th HWMCC CAV'14 FLOC'14 Vienna Austria	8th HWMCC FMCAD'15 Austin USA
Founding Lunch CAV'06 FLOC'06 Seattle		HWMCC Lunch FMCAD'08 Portland						
2006	2007	2008	2010	2011	2012	2013	2014	2015

- founding lunch at CAV'06, first competition at CAV'07
- HWMCC lunch at FMCAD'08 ⇒ need multiple properties !!!
- affiliated with either CAV (7,8,10,14) or FMCAD (11,12,13,15)
- HWMCC'11: old SINGLE, new LIVEness and new MULTI property track
- HWMCC'12 as HWMCC'11, new DEEP bounds track sponsored by Oski
- in essence no change in HWMCC'12 - HWMCC'15
- HWMCC'15: DEEP, SINGLE, and LIVE, ~~MULTI~~, **1h time limit**, before 15min

■ SINGLE

- single safety property, specified as bad state monitor
 - SAT = bad state reachable dual safety property does not hold
 - UNSAT = bad state not reachable dual safety property does holds
- restricted (pre AIGER 1.9) models only:
 - thus zero initialized latches (actually flip-flops)
 - and no environment constraints
 - new Oski benchmarks went through `aigtosmv | smvtoaig-20071012`

■ DEEP

- focus on hard unsolved benchmarks in SINGLE
- “how deep” can the model checker prove absence of bad state
- for instance in BMC: print bounds for which unrolling UNSAT

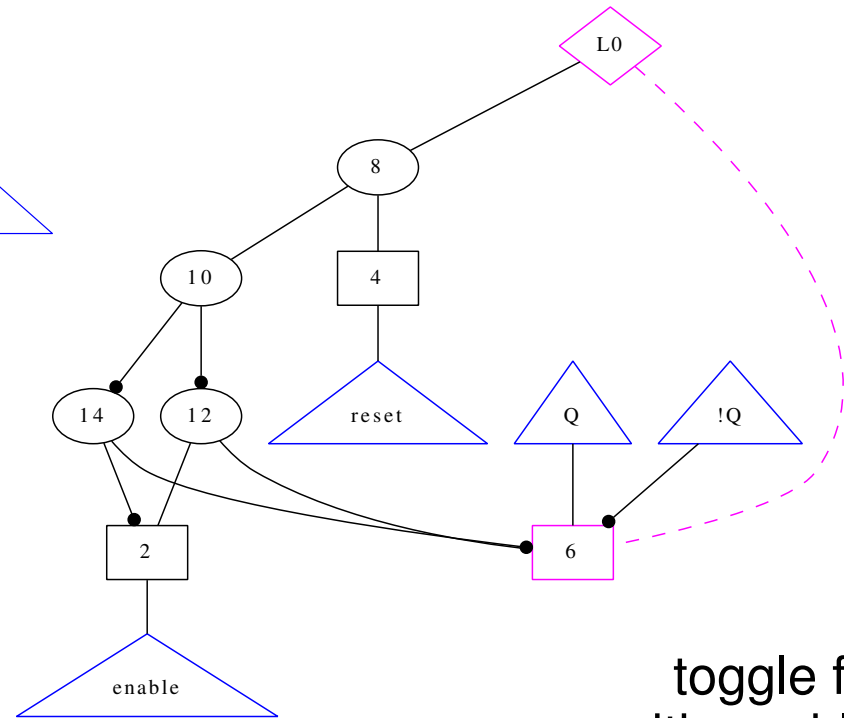
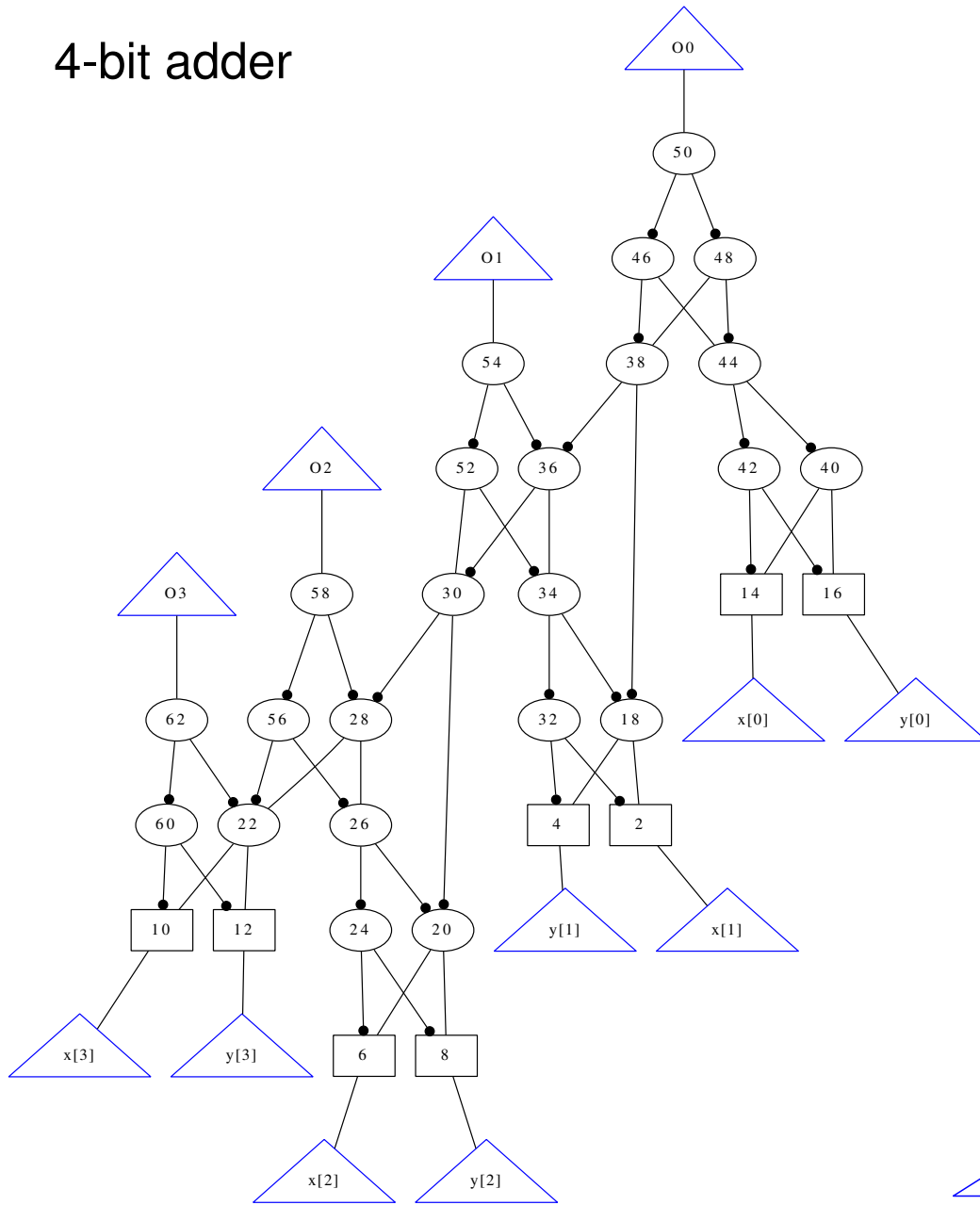
■ LIVE

- single justice properties (with fairness and environment constraints)
- generalized Büchi automata

19 model checkers from 11 teams

- **abcsimple, abcsimplive, abcsuprove** from Berkeley Brayton, Sterin, Mishchenko, ...
- **aigbmc, blimc** from JKU Linz Biere
- **avy** from Technion+SEI+Princeton Vizel, Gurfinkel, Malik
- **iimc** from Boulder Somenzi, Bradley, Hassan
- **iprover, iproverdeep** from Manchester Tsarkov, Korovin
- **nuxmv, nuxmvbmc** from Trento Griggio, Roveri, ...
- **pdtravdeep, pdtravthrd** from Torino Cabodi, Quer, ...
- **ricecnu** from Rice Li, Vardi
- **shiftbmc** from Dresden Manthey
- **tip2014** from Chalmers Sörensson, Claessen
- **v3s** from Taipei Yang, Wu, Huang

4-bit adder

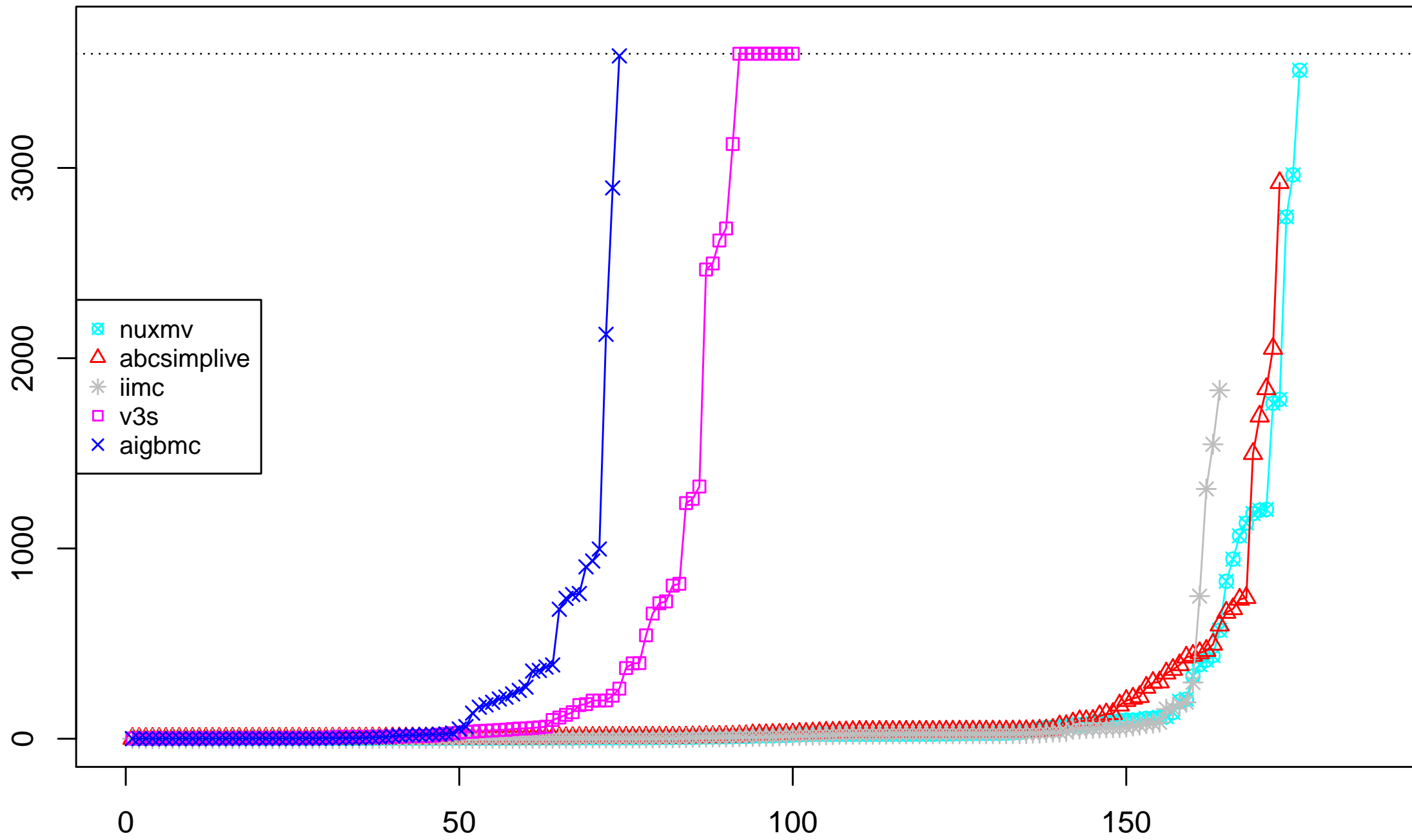


toggle flip-flop
with enable & reset

- LIVE with the identical **223** LIVE benchmarks from HWMCC'14
 - removed “outputs” in `cu*` benchmarks though
- SINGLE
 - included all **79** unsolved instances from HWMCC'14 in SINGLE track:
6s(53), intel(9), oski(7), beem(5), bob(4), nusmv(1)
 - added **156** new benchmarks from Oski submitted by Vigyan Singhal
oski15a01(80), oski15a07(6), oski15a08(18), oski15a10(18), oski15a10(34)
 - classified all 2075 SINGLE benchmarks on my hard disk into 19 classes
 - removed 79 unsolved from HWMCC'14
 - removed 690 solved by 11 out of 16 model checkers in 36 seconds
 - randomly selected further **313** benchmarks out of remaining 1306 as follows:
 - varying fraction of selected benchmarks per class
 - fraction determined to balance origin and importance
 - in total **548** benchmarks in SINGLE

selected	class	remaining	fraction	original
79	6s	237	1/3	314
39	pdt	234	1/6	347
38	bob	76	1/2	147
38	beem	446	1/12	680
26	intel	53	1/2	62
16	bv	98	1/6	129
13	eijk	26	1/2	28
11	nusmv	23	1/2	32
8	vis	16	1/2	26
8	oski	16	1/2	24
8	lmcs	16	1/2	156
6	pj	13	1/2	14
5	ken	10	1/2	16
5	bjrb	10	1/2	44
4	nec	9	1/2	13
4	mentor	9	1/2	15
3	cmu	7	1/2	7
2	texas	5	1/3	19
0	regr	2	1/inf	2
313	total	1306	~1/4.17	2075

HWMCC'15 Cactus LIVE Track SAT+UNSAT



HCMCC'15 Table LIVE SAT+UNSAT

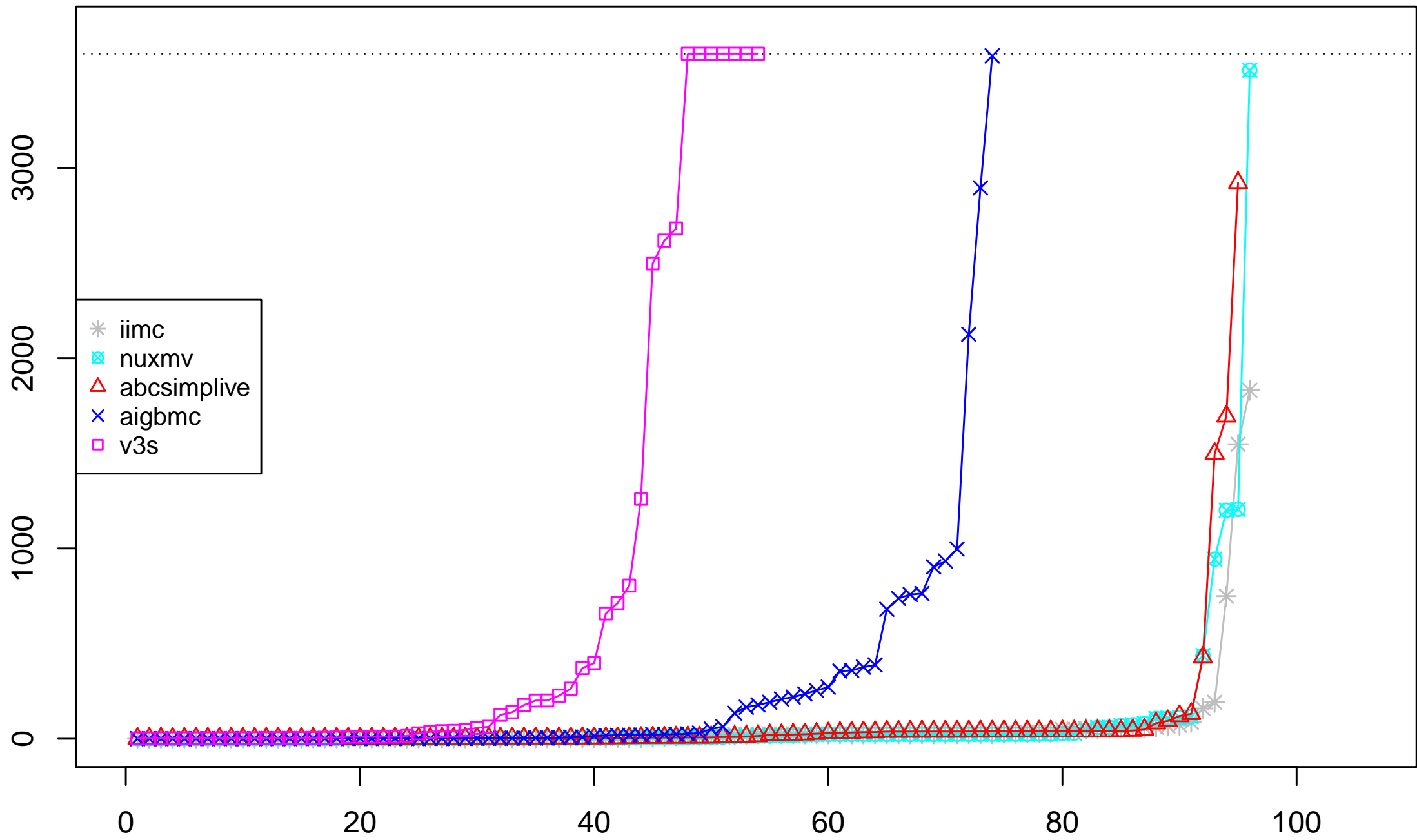
rank		cnt	ok	sat	uns	fld	to	mo	unk	real	time	max	best	uniq
1	nuxmv	223	176	96	80	47	46	1	0	26037	102160	5254	47	1
2	abcsimplive	223	173	95	78	50	19	31	0	21294	74824	6954	41	1
3	iimc	223	164	96	68	59	14	45	0	8709	29772	4354	67	3
	v3s	223	100	54	46	123	68	47	8	57957	117237	7000	3	0
	aigbmc	223	74	74	0	149	66	58	25	18122	18054	4507	26	2

hors concours (not ranked):

aigbmc: organizer model checker

cnt number of benchmarks considered (for ranking)
ok solved benchmarks (sat or unsat)
sat solved satisfiable benchmarks
uns solved unsatisfiable benchmarks
fld unsolved benchmarks (cnt - ok)
to time out (3600 wall clock time)
mo memory out (7GB main memory)
s6 signal 6 caught (abort)
s11 signal 11 caught (segmentation fault)
unk other signal or no result in output
real wall clock time
time process time
max maximum memory usage (MB)
best number of benchmarks solved fastest
uniq number of uniquely solved benchmarks

HWMCC'15 Cactus LIVE Track SAT



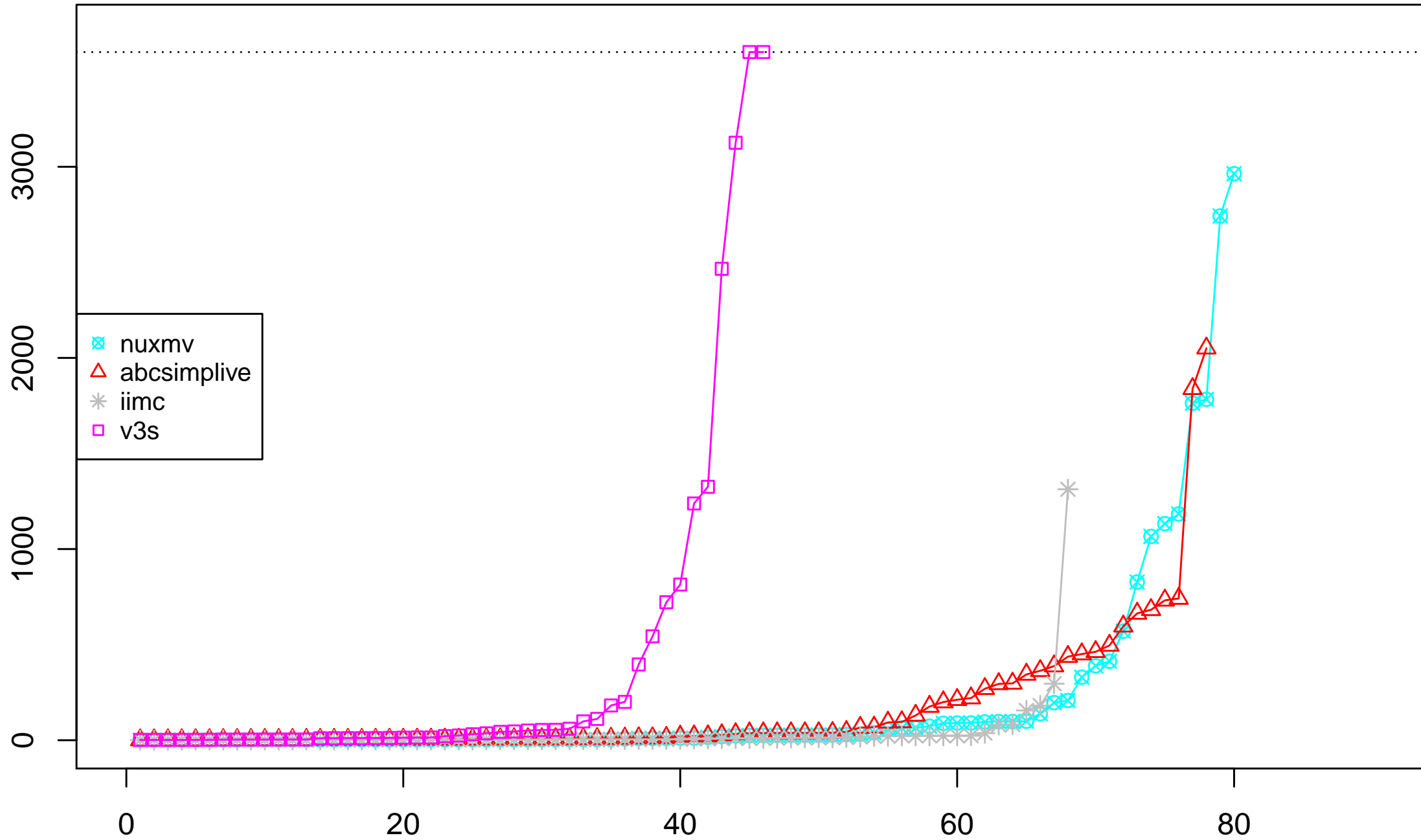
HCMCC'15 Table LIVE SAT

rank		cnt	ok	sat	fld	to	mo	real	time	max	best	uniq
1	iimc	100	96	96	4	3	1	6149	20686	2835	20	2
2	nuxmv	100	96	96	4	4	0	8924	34820	5254	30	1
3	abcsimplive	100	95	95	5	2	3	8313	28860	6954	24	0
	aigbmc	100	74	74	26	4	22	18122	18054	4507	26	2
	v3s	100	54	54	46	23	23	38969	67309	2577	0	0

hors concours (not ranked):

aigbmc: organizer model checker

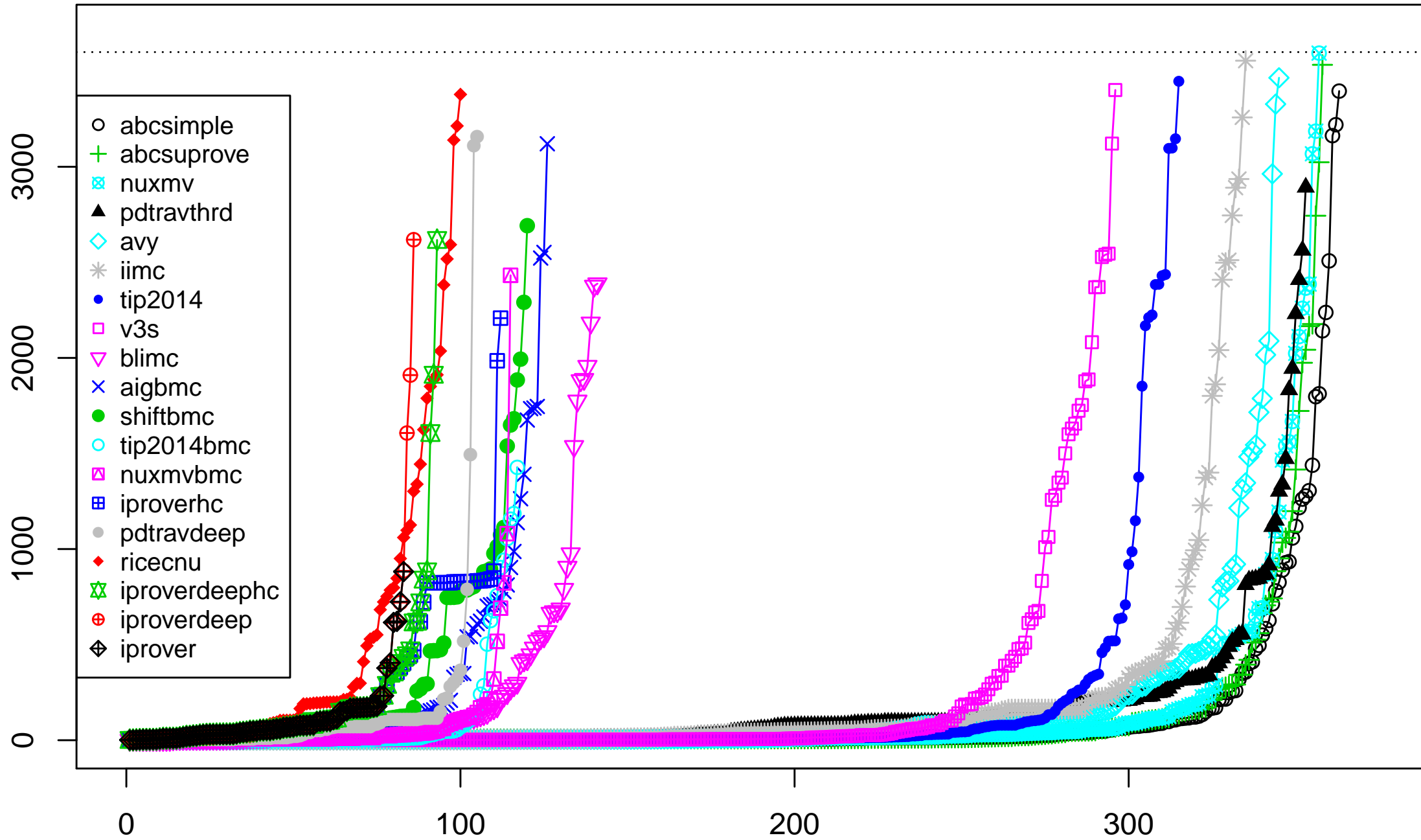
HWMCC'15 Cactus LIVE Track UNSAT



HCMCC'15 Table LIVE UNSAT

rank		cnt	ok	uns	fld	to	mo	unk	real	time	max	best	uniq
1	nuxmv	84	80	80	4	4	0	0	17113	67340	4838	17	0
2	abcsimplive	84	78	78	6	2	4	0	12982	45964	6485	17	1
3	iimc	84	68	68	16	0	16	0	2560	9086	4354	47	1
	v3s	84	46	46	38	31	6	1	18988	49927	7000	3	0

HWMCC'15 Cactus SINGLE Track SAT+UNSAT



HWMCC'15 Table SINGLE SAT+UNSAT

rank		cnt	ok	sat	uns	fld	to	mo	s11	s6	unk	real	time	max	best	uniq
1	abcsimple	548	363	122	241	185	154	31	0	0	0	45438	161407	6785	28	1
	abcsuprove	548	358	120	238	190	156	34	0	0	0	41596	115597	6905	23	1
2	nuxmv	548	357	127	230	191	165	26	0	0	0	47600	186902	6900	30	0
	pdtravthrd	548	353	115	238	195	62	124	3	0	6	52819	158951	6824	18	3
3	avy	548	345	115	230	203	171	30	0	0	2	49724	166878	5924	19	1
	iimc	548	335	105	230	213	163	50	0	0	0	57840	202386	6969	76	7
	tip2014	548	315	98	217	233	233	0	0	0	0	46817	46596	1062	144	0
	v3s	548	296	62	234	252	141	7	102	2	0	54273	152125	6037	28	4
	blimc	548	141	128	13	407	287	20	0	0	100	31687	31584	2241	61	1
	aigbmc	548	126	126	0	422	310	43	0	0	69	34210	34104	3644	47	1
	shiftbmc	548	120	120	0	428	247	12	0	0	169	37819	37697	1810	28	0
	tip2014bmc	548	117	117	0	431	225	50	0	0	156	11054	10911	5496	64	1
	nuxmvbmc	548	115	115	0	433	256	28	0	0	149	8885	8796	2504	45	0
	iproverhc	548	112	64	48	436	156	0	0	0	280	32269	17292	6730	3	0
	pdtravdeep	548	105	46	59	443	345	86	4	0	8	14725	14568	4202	11	0
	ricecnu	548	100	30	70	440	342	88	0	10	0	49153	49054	2204	0	0
	iproverdeephc	548	93	46	47	455	177	0	0	0	278	18116	17482	6728	2	0
	iproverdeep	548	86	46	40	462	149	0	0	0	313	14899	14818	6322	1	0
	iprover	548	83	43	40	465	59	0	0	0	406	8768	8698	6176	3	0

hors concours (not ranked):

aigbmc blimc: organizer model checkers

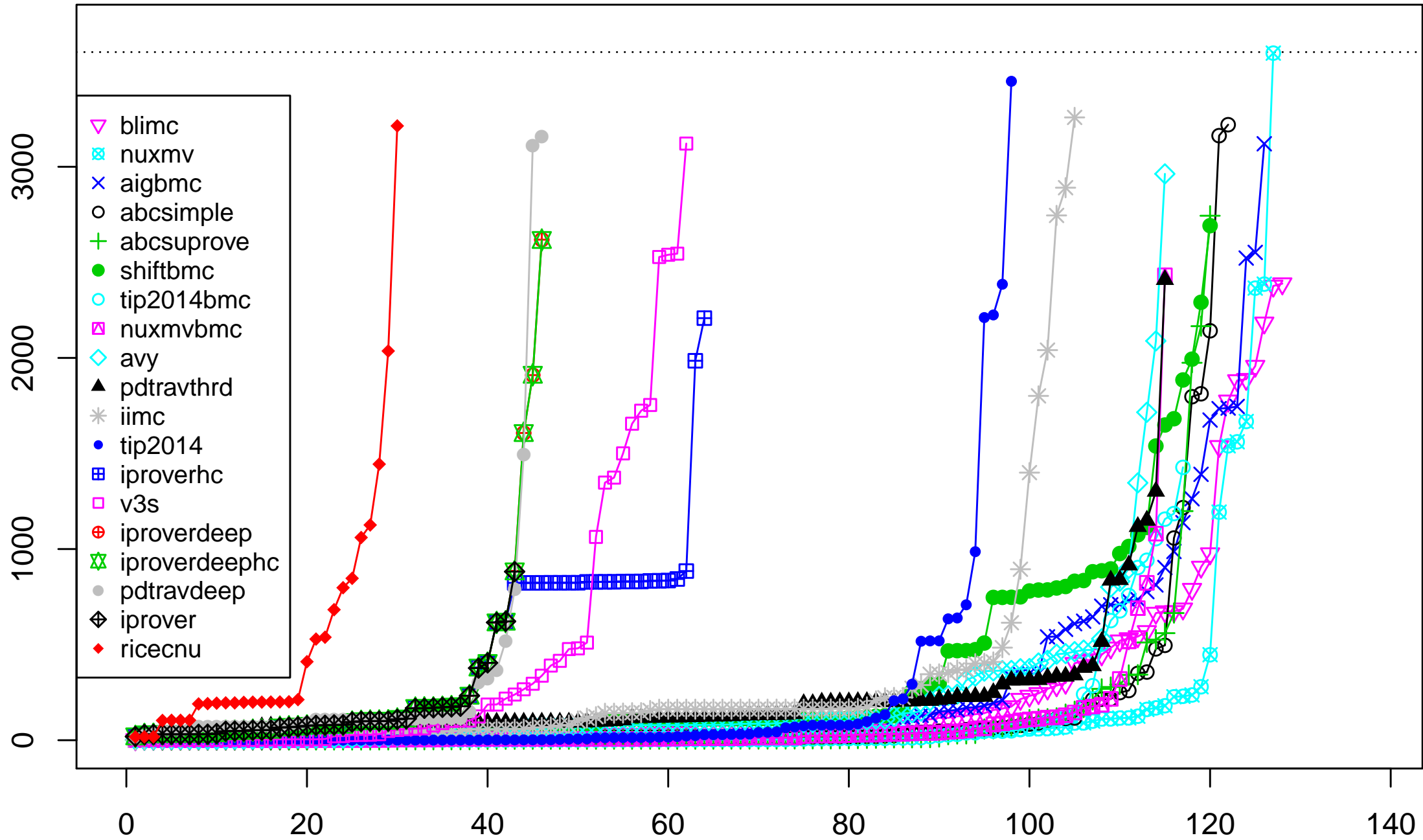
pdtravthrd: issue catching 'FATAL' for 'intel045' (not counted)

ricecnu: reports 8 instances SAT which are UNSAT (not counted)

iprover*hc: last minute (script) fixes after deadline

each team / submitter only ranked once (one medal maximum)

HWMCC'15 Cactus SINGLE Track SAT



HWMCC'15 Table SINGLE SAT

rank		cnt	ok	sat	fld	to	mo	s11	s6	unk	real	time	max	best	uniq
1	blimc	144	128	128	16	14	1	0	0	1	31680	31583	2241	4	1
	nuxmv	144	127	127	17	9	8	0	0	0	18941	74202	6899	4	0
2	aigbmc	144	126	126	18	16	1	0	0	1	34210	34104	3644	7	1
	abcsimple	144	122	122	22	11	11	0	0	0	19559	66584	6267	4	1
3	abcsuprove	144	120	120	24	15	9	0	0	0	13884	30568	5623	16	0
	shiftbmc	144	120	120	24	23	0	0	0	1	37819	37697	1810	2	0
	tip2014bmc	144	117	117	27	23	3	0	0	1	11054	10911	5496	46	1
	nuxmvbmc	144	115	115	29	27	1	0	0	1	8885	8796	2504	5	0
	avy	144	115	115	29	21	6	0	0	2	22939	65806	5924	2	0
	pdtravthrd	144	115	115	29	7	15	2	0	5	23817	57597	6108	3	0
	iimc	144	105	105	39	19	20	0	0	0	27651	87239	6969	7	4
	tip2014	144	98	98	46	46	0	0	0	0	17110	17037	1062	43	0
	iproverhc	144	64	64	80	19	0	0	0	61	26045	11701	6695	0	0
	v3s	144	62	62	82	24	0	58	0	0	26253	75591	6037	0	0
	iproverdeep	144	46	46	98	20	0	0	0	78	12214	12164	6322	0	0
	iproverdeephc	144	46	46	98	20	0	0	0	78	12247	12165	6340	0	0
	pdtravdeep	144	46	46	98	83	4	3	0	8	13354	13255	4202	1	0
	iprover	144	43	43	101	3	0	0	0	98	6079	6042	6176	0	0
	ricecnu	144	30	30	114	75	38	0	1	0	15532	15500	2204	0	0

hors concours (not ranked) :

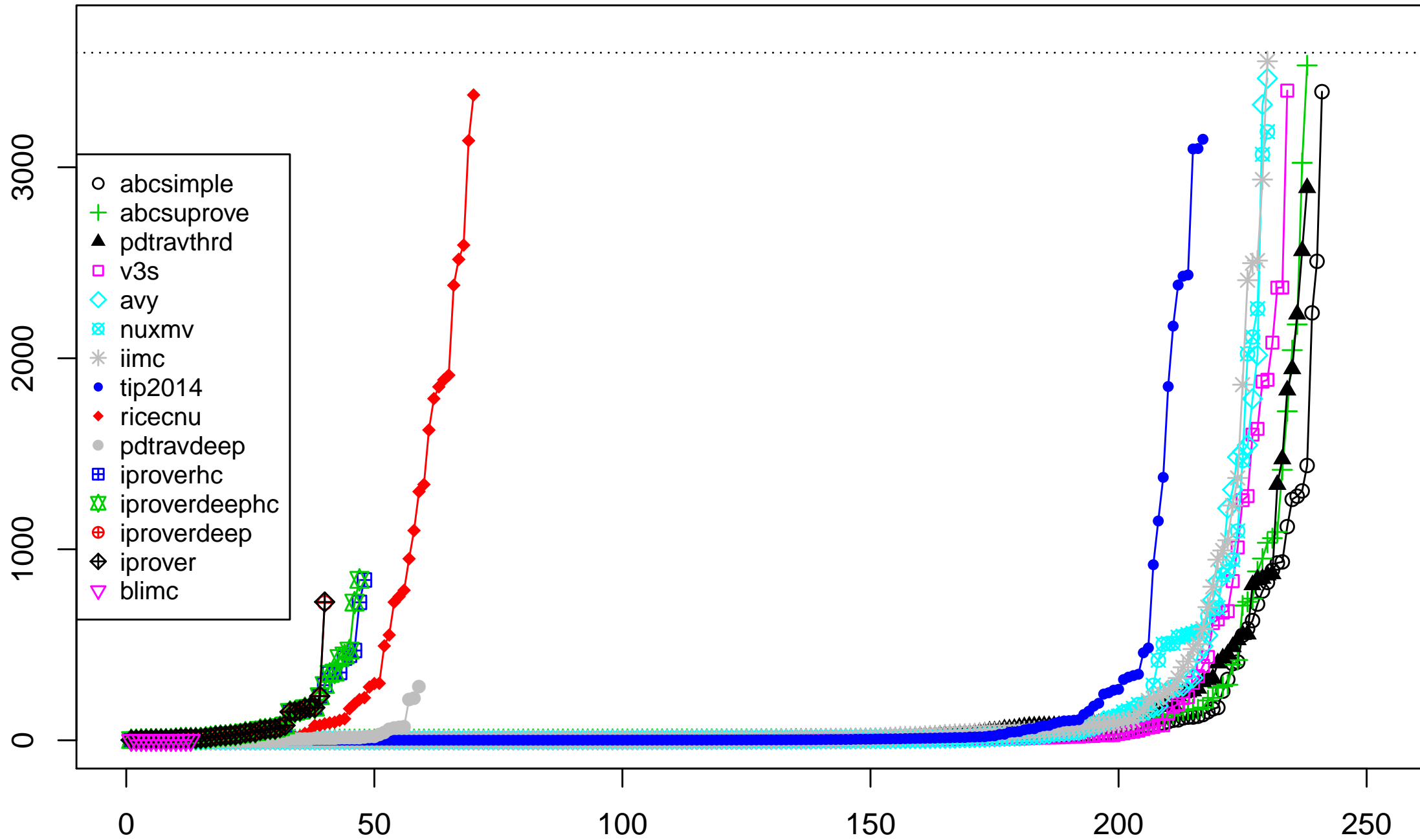
aigbmc blimc: organizer model checkers

pdtravthrd: issue catching 'FATAL' for 'intel045' (not counted)

ricecnu: reports 8 instances SAT which are UNSAT (not counted)

iprover*hc: last minute (script) fixes after deadline

HWMCC'15 Cactus SINGLE Track UNSAT



HWMCC'15 Table SINGLE UNSAT

rank		cnt	ok	uns	dis	fld	to	mo	s11	s6	unk	real	time	max	best	uniq
1	abcsimple	269	241	241	0	28	22	6	0	0	0	25879	94823	6785	24	0
	abcsuprove	269	238	238	0	31	18	13	0	0	0	27711	85028	6905	7	1
	pdtravthrd	269	238	238	0	31	15	16	0	0	0	29002	101354	6824	15	3
2	v3s	269	234	234	0	35	26	3	6	0	0	28020	76534	5074	26	4
3	avy	269	230	230	0	39	30	9	0	0	0	26785	101073	5513	17	1
	nuxmv	269	230	230	0	39	33	6	0	0	0	28660	112700	6900	13	0
	iimc	269	230	230	0	39	31	8	0	0	0	30189	115147	4613	56	3
	tip2014	269	217	217	0	52	52	0	0	0	0	29707	29559	431	101	0
	ricecnu	269	70	70	8	191	167	15	0	9	0	33621	33554	1766	0	0
	pdtravdeep	269	59	59	0	210	152	57	1	0	0	1371	1314	2391	2	0
	iproverhc	269	48	48	0	221	82	0	0	0	139	6225	5592	6730	0	0
	iproverdeephc	269	47	47	0	222	101	0	0	0	121	5869	5317	6728	1	0
	iproverdeep	269	40	40	0	229	80	0	0	0	149	2686	2653	5390	0	0
	iprover	269	40	40	0	229	35	0	0	0	194	2688	2655	5391	0	0
	blimc	269	13	13	0	256	147	12	0	0	97	7	1	2	7	0

hors concours (not ranked):

aigbmc blimc: organizer model checkers

pdtravthrd: issue catching 'FATAL' for 'intel045' (not counted)

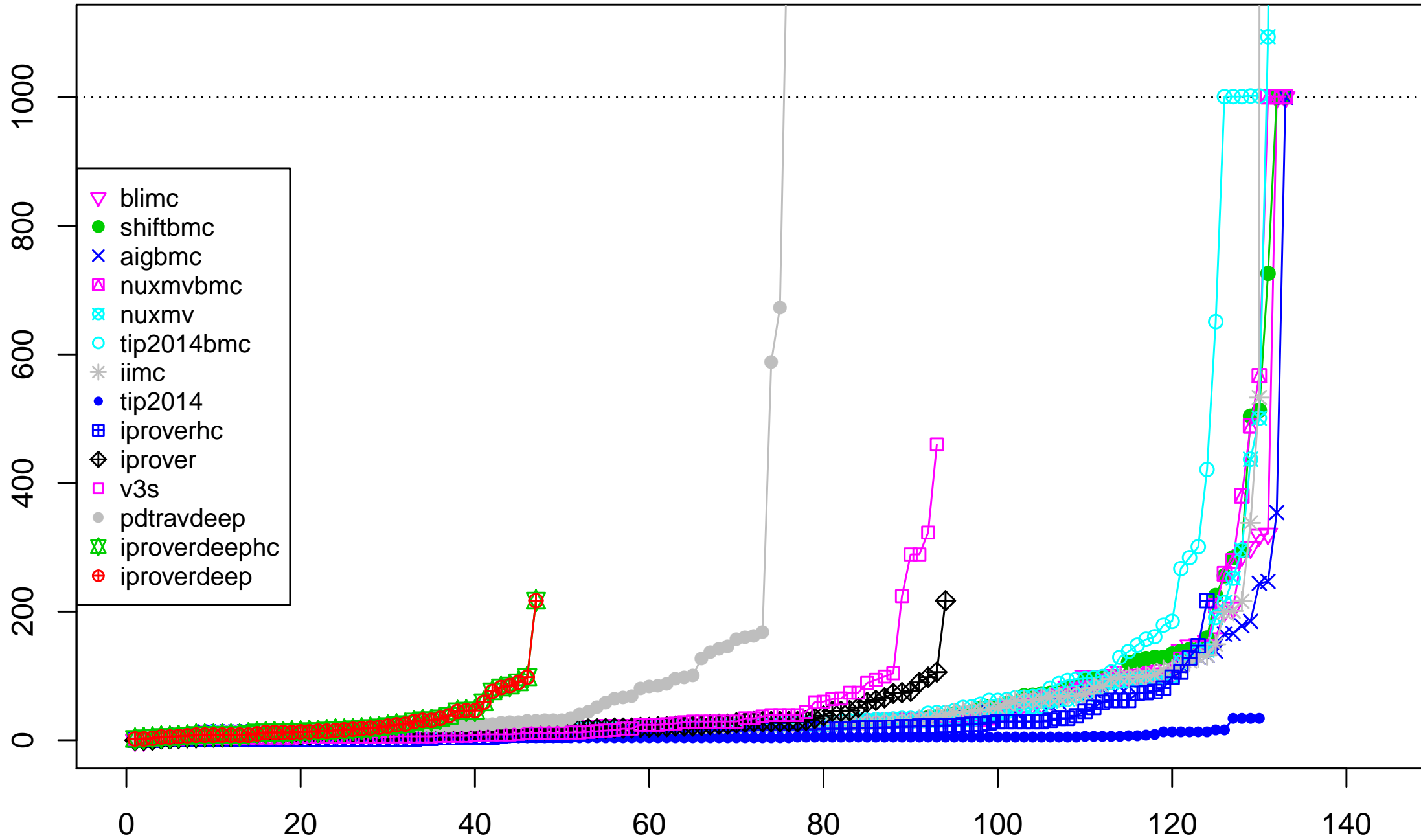
ricecnu: reports 8 instances SAT which are UNSAT (not counted)

iprover*hc: last minute (script) fixes after deadline

135 unsolved benchmarks in SINGLE within one hour

6s100 6s105 6s119 6s148 6s149 6s160 6s161 6s171 6s179 6s185 6s186 6s187
6s188 6s190 6s195 6s22 6s23 6s24 6s266rb2 6s267rb3 6s268r 6s274r 6s279r
6s280r 6s29 6s329rb19 6s329rb20 6s33 6s340rb27 6s340rb63 6s341r 6s342rb122
6s351rb02 6s36 6s365r 6s366r 6s376r 6s377r 6s382r 6s39 6s392r 6s398b09
6s398b16 6s399b02 6s399b03 6s402rb0342 6s44 6s45 6s514r 6s516r 6s517rb0
beembkry8b1 beemcmbrdg7f2 beemfwt4b1 beemfwt5b1 beemfwt5b3 beemhanoi4b1
beemkrebs4b1 beemldfilt7b1 beemlifts3b1 beemloyd3b1 beempgmprot1f2
beemskbn2b1 bob12s02 bobpcihm bobsmcodic bobsmminiuart intel012 intel013
intel016 intel027 intel028 intel032 intel065 intel066 intel067 nusmvdme216
oski15a01b00s oski15a01b02s oski15a01b08s oski15a01b12s oski15a01b14s
oski15a01b18s oski15a01b26s oski15a01b28s oski15a01b30s oski15a01b32s
oski15a01b34s oski15a01b36s oski15a01b38s oski15a01b40s oski15a01b42s
oski15a01b44s oski15a01b48s oski15a01b50s oski15a01b52s oski15a01b54s
oski15a01b58s oski15a01b60s oski15a01b62s oski15a01b64s oski15a01b66s
oski15a01b68s oski15a01b70s oski15a01b72s oski15a01b74s oski15a01b76s
oski15a01b78s oski15a07b0s oski15a07b2s oski15a07b4s oski15a08b00s
oski15a08b02s oski15a08b04s oski15a08b10s oski15a08b12s oski15a10b02s
oski15a10b04s oski15a10b06s oski15a10b08s oski15a10b10s oski15a10b16s
oski15a14b02s oski15a14b04s oski15a14b06s oski15a14b20s oski15a14b26s
oski15a14b28s oski15a14b30s oski1rub00i oski1rub01i oski1rub02i
oski1rub08i oski1rub09i oski1rub10i

HWMCC'15 Cactus DEEP Bound Distribution



DEEP Bound Track award of \$500 sponsored by Oski Technology

135 unsolved benchmarks in SINGLE (out of 548)

reached bounds capped at 1000 (last year 100) \Rightarrow bound_{*i*}

$$\text{deep} = \frac{1}{135} \cdot \sum_{i=1}^{135} (1 - 1/(2 + \text{bound}_i))$$

bound _{<i>i</i>}	=	-1	contributes	0%
bound _{<i>i</i>}	=	0	contributes	50%
bound _{<i>i</i>}	=	1	contributes	75%
			⋮	
bound _{<i>i</i>}	=	8	contributes	90%
			⋮	
bound _{<i>i</i>}	=	98	contributes	99%
			⋮	

HWMCC'15 Table DEEP

rank		cnt	to	mo	s11	s6	unk	best	deep
	blimc	135	126	7	0	0	2	50	93952
1	shiftbmc	135	127	6	0	0	2	26	93733
	aigbmc	135	128	6	0	0	1	40	93728
2	nuxmvbmc	135	124	8	0	0	3	40	93423
	nuxmv	135	123	12	0	0	0	13	93058
3	tip2014bmc	135	100	30	0	0	5	18	92031
	iimc	135	113	22	0	0	0	13	91396
	tip2014	135	135	0	0	0	0	0	82167
	iproverhc	135	55	0	0	0	80	3	79281
	iprover	135	21	0	0	0	114	3	64248
	v3s	135	91	4	38	2	0	2	61874
	pdtravdeep	135	110	25	0	0	0	8	55305
	iproverdeephc	135	56	0	0	0	79	1	32492
	iproverdeep	135	49	0	0	0	86	1	32398
									*1e-5

'tip2014bmc' was added after competition and solved uniquely '6s299b685' which lead to 'nuxmvbmc' and 'shiftbmc' swapping 1st and 2nd place

hors concours (not ranked):

aigbmc blimc: organizer model checkers

iprover*hc: last minute (script) fixes after deadline

others not participating and / or producing incorrect bounds

- changes this year
 - smaller machines
 - larger time limit of 1 hour
 - two model checkers per team maximum
- changes for next year
 - use AIGER version 1.9 format in SINGLE too
 - reencode benchmarks with environment constraints and uninitialized latches
 - strict policy for discrepancies after final deadline
 - more powerful HW?
 - MULTI?
- plans
 - regression, fuzzing and delta-debugging suite
 - continue with word-level format
- please help us getting **more benchmarks**