Configurable SAT Solver Challenge 2014

Frank Hutter[a], Marius Lindauer[a], Sam Bayless[b], Holger Hoos[b], Kevin
Leyton-Brown[b]

[a]*University of Freiburg, Germany*
[b]*University of British Columbia, Vancouver, Canada*

1. Setup

1.1. Participants

- *clasp-3.0.4-p8* [4]

- *cryptominisat* [9]

- *CSCCSat2014* [5, 6]

- *DCCASat+march-rw* [7]

- *lingeling* [3]

- *minisat-HACK-999ED-CSSC*

- *probSAT* [2]

- *Riss-4.27* [8]

- *SparrowToRiss* [1]

- *YalSAT* [3]

1.2. Benchmarks

- *Industrial SAT+UNSAT*:

    - *Bounded Model Checking* : derived by unrolling the HWMCC 2008
      circuits into CNF

    - *Circuit Fuzz*: `http://fmv.jku.at/cnfuzzdd/`

    - *Hardware Verification (IBM)*: BM Formal Verification Benchmark
      Library (offline)

- *crafted SAT+UNSAT*

  - *Graph Isomorphism* : see 2013 SAT Challenge Proceedings
  - *Low Autocorrelation Binary Sequence* see 2013 SAT Challenge Proceedings
  - *Queens*: submitted by Norbert Manthey and Peter Steinke

- *Random SAT+UNSAT*

  - *3cnf*: generated by Sam Bayless
  - *K3*: N/A
  - *unif-k5*: see 2012 SAT Challenge Proceedings, `http://sourceforge.net/projects/ksatgenerator/`

- *Random SAT*

  - *3sat1k*: Captain Jack Paper
  - *5sat500*: Captain Jack Paper
  - *7sat90*: Captain Jack Paper

2. Results

A performance is shown in boldface if it is significant better (according to a permutation test with 100 000 permutations and significance level $\alpha = 0.05$).

2.1. Industrial SAT+UNSAT

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
lingeling	81	81	76.85	76.41	69	69	80.41	**77.16**	paramils-1
minisat-HACK-999ED-CSSC	82	**81**	68.87	**68.01**	70	70	72.91	**72.16**	smac-3
clasp-3.0.4-p8	85	**84**	74.11	**71.11**	71	71	75.59	75.24	smac-0
Riss-4.27	82	82	**72.28**	72.49	72	72	77.32	**76.83**	smac-disc-0
cryptominisat	81	**80**	74.55	75.26	70	**69**	77.48	78.58	smac-1
SparrowToRiss	83	**82**	152.12	**75.93**	72	72	154.35	**80.03**	smac-1

Table 1: Track: *Industrial SAT+UNSAT*; Benchmarks: *Hardware Verification (IBM)*

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
lingeling	23	**11**	40.43	**26.38**	30	**18**	47.79	**31.96**	paramils-1
minisat-HACK-999ED-CSSC	13	**10**	29.01	**24.97**	21	19	38.45	**34.29**	smac-2
clasp-3.0.4-p8	11	**7**	24.87	**21.32**	18	**12**	32.60	**27.63**	smac-disc-1
Riss-4.27	16	**12**	33.88	**29.16**	20	22	39.31	37.21	smac-disc-1
cryptominisat	25	**14**	39.13	**30.18**	31	**20**	51.26	**36.09**	smac-1
SparrowToRiss	26	**13**	167.24	**32.40**	29	**21**	173.79	**39.67**	smac-disc-2

Table 2: Track: *Industrial SAT+UNSAT*; Benchmarks: *Circuit Fuzz*

| Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
lingeling	115	**109**	75.40	**73.62**	20	20	43.65	**42.56**	paramils-1
minisat-HACK-999ED-CSSC	108	**100**	62.12	**60.73**	22	22	36.47	**34.87**	gga-disc-1
clasp-3.0.4-p8	151	**134**	79.57	**75.16**	44	**30**	57.75	**47.29**	smac-3
Riss-4.27	172	**91**	107.65	**70.43**	39	**26**	72.79	**52.62**	smac-disc-2
cryptominisat	147	143	**96.32**	103.54	40	37	70.88	74.47	smac-1
SparrowToRiss	217	**145**	204.66	**87.96**	62	**36**	190.12	**60.85**	smac-2

Table 3: Track: *Industrial SAT+UNSAT*; Benchmarks: *Bounded Model Checking*

Solver	Training performance				Test performance			
	#Timeouts		PAR1		#Timeouts		PAR1	
	default	config.	default	config.	default	config.	default	config.
lingeling	219	201	64.23	58.80	119	107	57.29	50.56
minisat-HACK-999ED-CSSC	203	191	53.33	51.24	113	111	49.28	47.11
clasp-3.0.4-p8	247	225	59.51	55.86	133	113	55.31	50.05
Riss-4.27	270	185	71.27	57.36	131	120	63.14	55.56
cryptominisat	253	237	70.00	69.66	141	126	66.54	63.05
SparrowToRiss	326	240	174.68	65.43	163	129	172.75	60.18

Table 4: Overall results for track: *Industrial SAT+UNSAT*; ranked by number of timeouts of configured solvers on test sets (tie-breaker: PAR1)
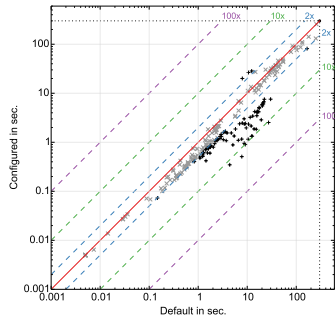
(a) *lingeling*; TOs: 69 → 69
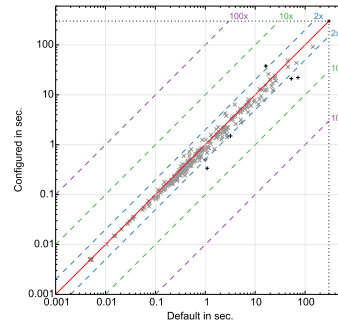
(b) *minisat-HACK-999ED-CSSC*; TOs: 70 → 70
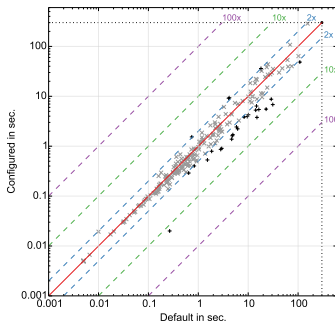
(c) *clasp-3.0.4-p8*; TOs: 71 → 71
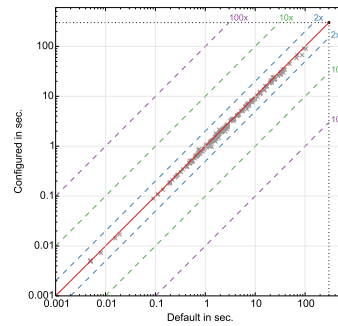
(d) *Riss-4.27*; TOs: 72 → 72
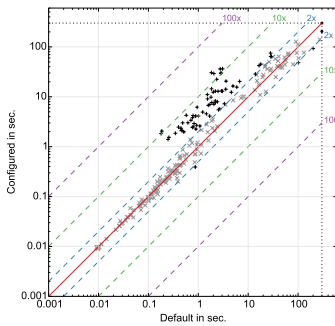
(e) *cryptominisat*; TOs: 70 → 69
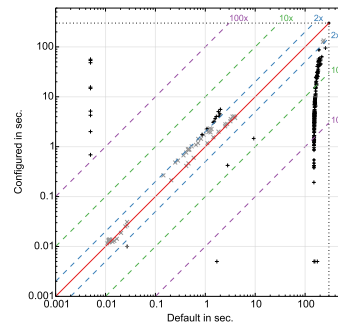
(f) *SparrowToRiss*; TOs: 72 → 72

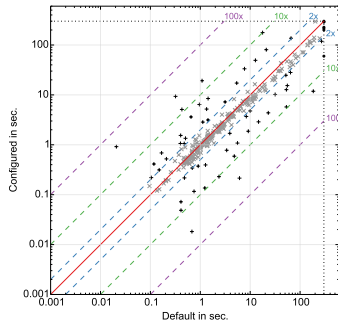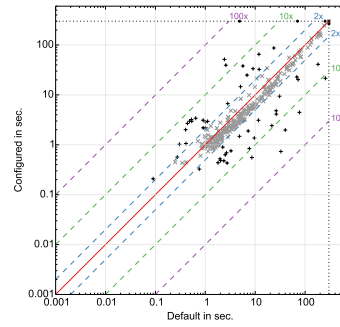Figure 1: Track: *Industrial SAT+UNSAT*, Benchmarks: *Hardware Verification (IBM)*
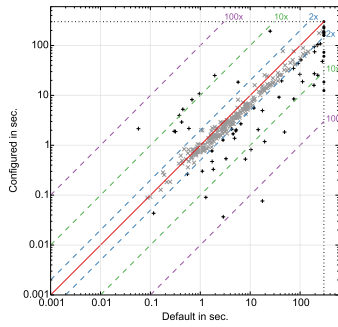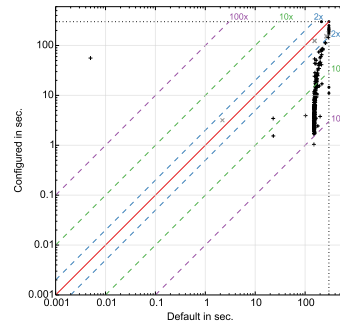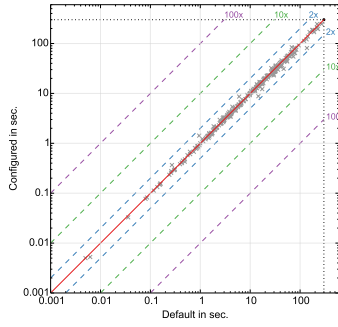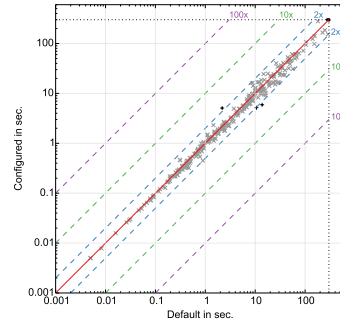
5
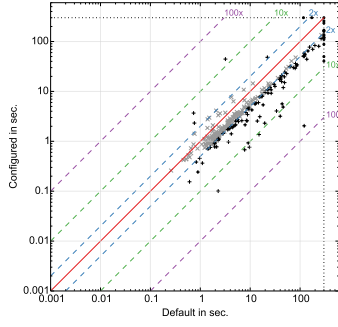
(a) *lingeling*; TOs: $30 \rightarrow 18$
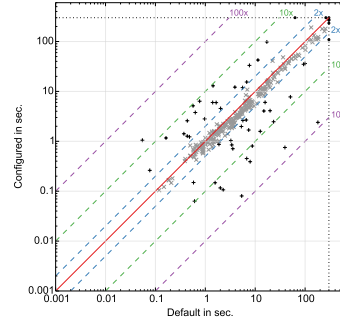
(b) *minisat-HACK-999ED-CSSC*; TOs: $21 \rightarrow 19$

(c) *clasp-3.0.4-p8*; TOs: $18 \rightarrow 12$

(d) *Riss-4.27*; TOs: $20 \rightarrow 22$

(e) *cryptominisat*; TOs: $31 \rightarrow 20$

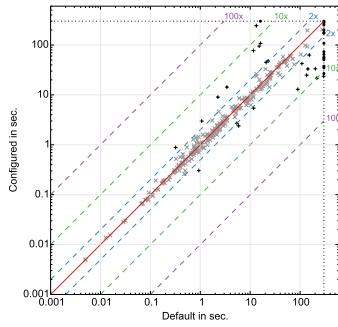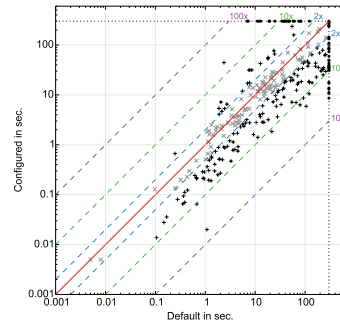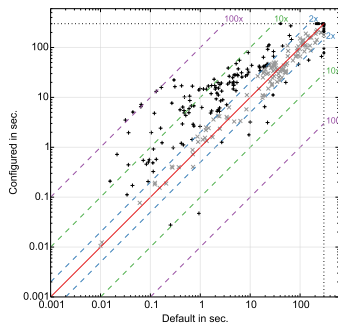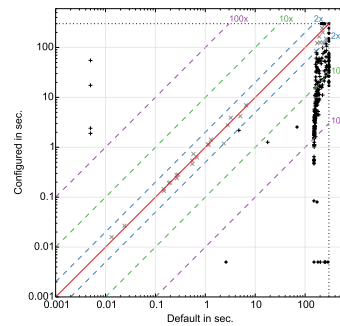(f) *SparrowToRiss*; TOs: $29 \rightarrow 21$

Figure 2: Track: *Industrial SAT+UNSAT*, Benchmarks: *Circuit Fuzz*

6

(a) *lingeling*; TOs: $20 \rightarrow 20$

(b) *minisat-HACK-999ED-CSSC*; TOs: $22 \rightarrow 22$

(c) *clasp-3.0.4-p8*; TOs: $44 \rightarrow 30$

(d) *Riss-4.27*; TOs: $39 \rightarrow 26$

(e) *cryptominisat*; TOs: $40 \rightarrow 37$

(f) *SparrowToRiss*; TOs: $62 \rightarrow 36$

Figure 3: Track: *Industrial SAT+UNSAT*, Benchmarks: *Bounded Model Checking*

7

2.2. crafted SAT+UNSAT

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
clasp-3.0.4-p8	86	87	80.79	82.49	**87**	93	**85.30**	88.51	paramils-2
lingeling	93	94	90.40	90.88	101	104	96.94	99.05	smac-disc-0
cryptominisat	89	**86**	84.63	85.40	95	**89**	90.04	89.81	smac-disc-1
Riss-4.27	90	**85**	84.02	81.95	91	88	89.03	**85.72**	paramils-1
SparrowToRiss	92	**86**	130.00	**83.70**	98	**94**	132.19	**90.93**	smac-2
minisat-HACK-999ED-CSSC	88	**83**	82.15	81.35	91	91	85.44	84.90	paramils-0
YalSAT	223	**210**	199.87	**188.63**	218	**207**	191.44	**183.74**	smac-0

Table 5: Track: *crafted SAT+UNSAT*; Benchmarks: *Low Autocorrelation Binary Sequence*

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
clasp-3.0.4-p8	132	**34**	42.01	**23.08**	43	**9**	39.73	**20.31**	smac-2
lingeling	43	**24**	28.09	**21.89**	11	**5**	27.49	**19.81**	smac-0
cryptominisat	115	**77**	40.16	**36.37**	43	**24**	38.83	**34.79**	smac-disc-0
Riss-4.27	133	**98**	45.02	**40.51**	43	**30**	41.56	**36.55**	smac-1
SparrowToRiss	163	**123**	125.86	**52.31**	55	**42**	115.02	**48.01**	smac-disc-1
minisat-HACK-999ED-CSSC	143	**142**	44.76	**43.70**	50	50	45.07	**44.47**	smac-disc-0
YalSAT	590	590	173.26	**172.08**	186	186	159.96	**159.44**	smac-disc-2

Table 6: Track: *crafted SAT+UNSAT*; Benchmarks: *Graph Isomorphism*

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
clasp-3.0.4-p8	90	**0**	71.13	**5.19**	81	**0**	81.84	**4.68**	smac-3
lingeling	3	**0**	26.14	**17.57**	3	**0**	26.92	**17.38**	paramils-2
cryptominisat	4	**0**	19.50	**8.83**	2	1	22.55	**9.62**	smac-disc-1
Riss-4.27	3	**0**	16.54	**6.77**	2	**0**	13.68	**7.31**	smac-disc-0
SparrowToRiss	10	**0**	103.80	**7.79**	3	**0**	91.93	**8.52**	smac-disc-0
minisat-HACK-999ED-CSSC	0	0	10.98	**8.18**	0	0	11.49	**8.38**	gga-disc-1
YalSAT	484	484	300.00	300.00	351	351	300.00	300.00	smac-disc-0

Table 7: Track: *crafted SAT+UNSAT*; Benchmarks: *Queens*

Solver	Training performance				Test performance			
	#Timeouts		PAR1		#Timeouts		PAR1	
	default	config.	default	config.	default	config.	default	config.
clasp-3.0.4-p8	308	121	64.64	36.92	211	102	68.96	37.83
lingeling	139	118	48.21	43.44	115	109	50.45	45.41
cryptominisat	208	163	48.10	43.53	140	114	50.47	44.74
Riss-4.27	226	183	48.53	43.08	136	118	48.09	43.20
minisat-HACK-999ED-CSSC	231	225	45.96	44.41	141	141	47.33	45.92
YalSAT	1297	1284	224.37	220.24	755	744	217.13	214.40
SparrowToRiss(disq.)	265	209	119.89	47.93	156	136	113.05	49.15

Table 8: Overall results for track: *crafted SAT+UNSAT*; ranked by number of timeouts of configured solvers on test sets (tie-breaker: PAR1)
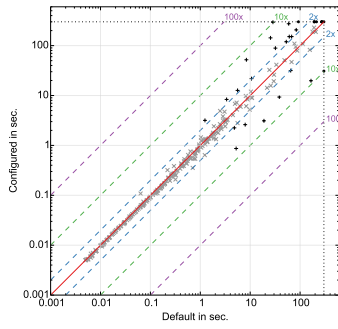
(a) *clasp-3.0.4-p8*; TOs: $87 \rightarrow 93$
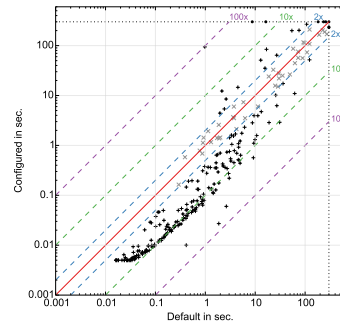
(b) *lingeling*; TOs: $101 \rightarrow 104$
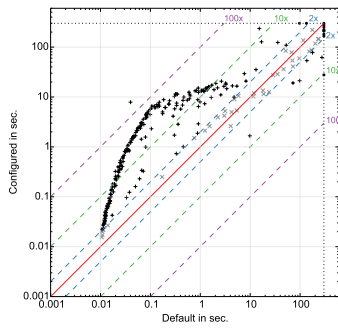
(c) *cryptominisat*; TOs: $95 \rightarrow 89$
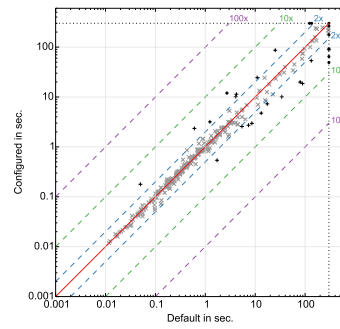
(d) *Riss-4.27*; TOs: $91 \rightarrow 88$
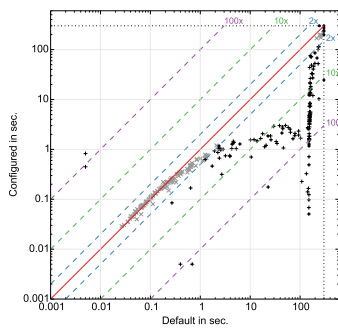
(e) *SparrowToRiss*; TOs: $98 \rightarrow 94$
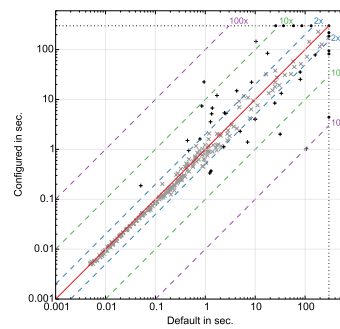
(f) *minisat-HACK-999ED-CSSC*; TOs: $91 \rightarrow 91$

10
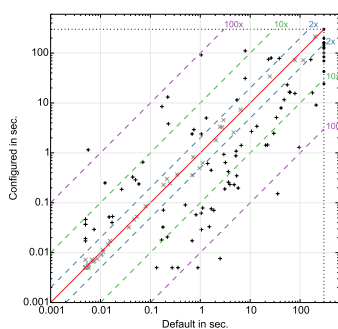
(g) *YalSAT*; TOs: $218 \rightarrow 207$
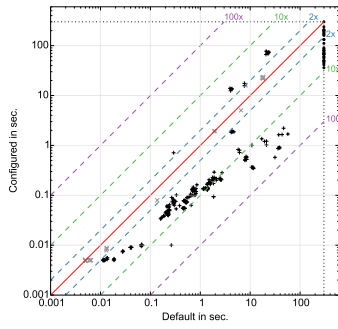
(a) *clasp-3.0.4-p8*; TOs: $43 \rightarrow 9$
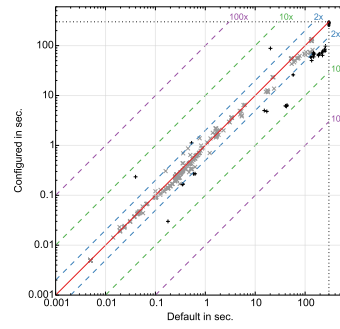
(b) *lingeling*; TOs: $11 \rightarrow 5$
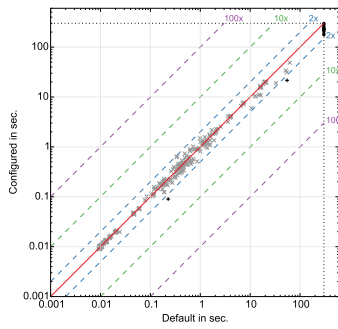
(c) *cryptominisat*; TOs: $43 \rightarrow 24$
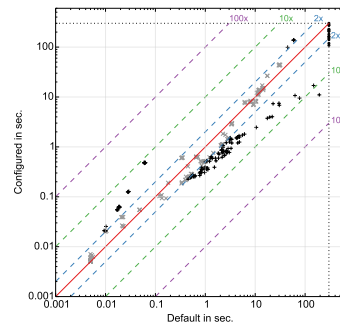
(d) *Riss-4.27*; TOs: $43 \rightarrow 30$
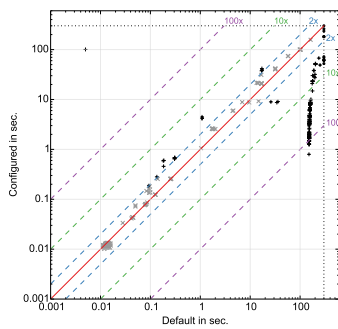
(e) *SparrowToRiss*; TOs: $55 \rightarrow 42$
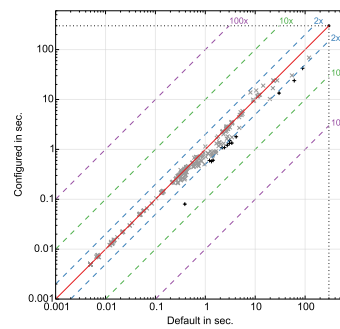
(f) *minisat-HACK-999ED-CSSC*; TOs: $50 \rightarrow 50$
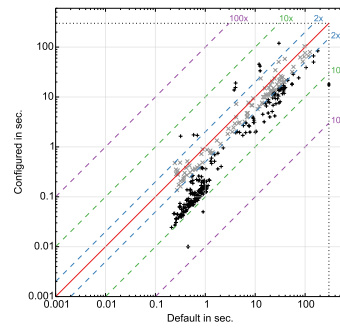
11
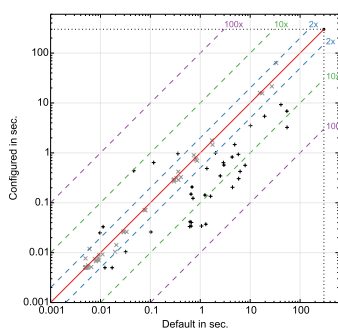
(g) *YalSAT*; TOs: $186 \rightarrow 186$
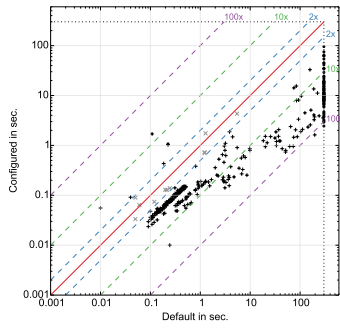
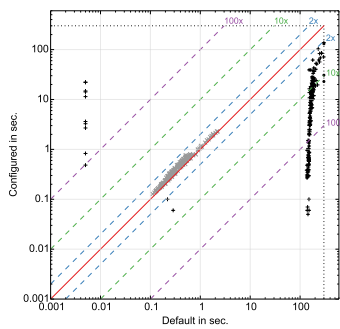(a) *clasp-3.0.4-p8*; TOs: $81 \to 0$
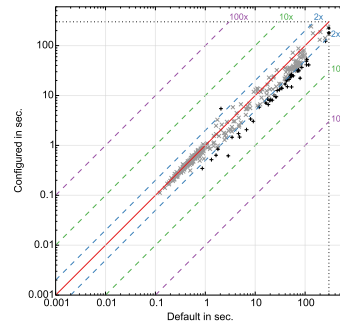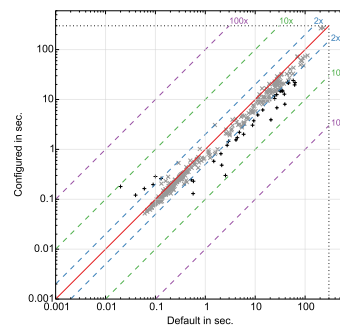

(b) *lingeling*; TOs: $3 \to 0$
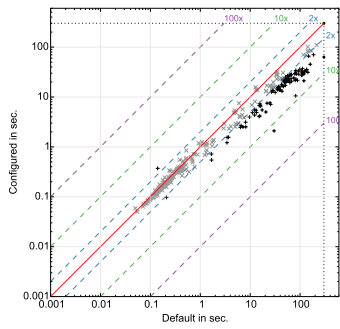
(c) *cryptominisat*; TOs: $2 \to 1$

(d) *Riss-4.27*; TOs: $2 \to 0$

(e) *SparrowToRiss*; TOs: $3 \to 0$

(f) *minisat-HACK-999ED-CSSC*; TOs: $0 \to 0$

12
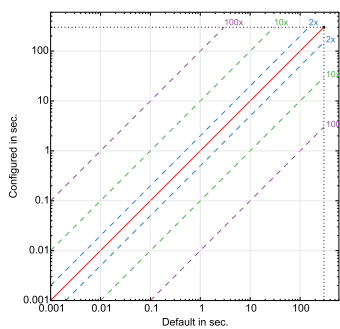
(g) *YalSAT*; TOs: $351 \to 351$

2.3. Random SAT+UNSAT

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
clasp-3.0.4-p8	0	0	11.36	**4.11**	0	0	7.91	**2.66**	smac-3
DCCASat+march-rw	0	0	81.07	**16.44**	0	0	74.75	**15.01**	gga-1
minisat-HACK-999ED-CSSC	7	**0**	35.36	**22.19**	5	**1**	30.77	**14.86**	paramils-1
Riss-4.27	7	**2**	38.34	**26.64**	2	2	27.95	**20.42**	smac-disc-2
SparrowToRiss	24	**3**	104.32	**28.89**	8	**1**	89.67	**20.99**	smac-2

Table 9: Track: *Random SAT+UNSAT*; Benchmarks: *K3*

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
clasp-3.0.4-p8	0	0	0.74	**0.30**	0	0	0.74	**0.30**	paramils-3
DCCASat+march-rw	0	0	149.65	**29.64**	1	**0**	150.73	**29.90**	paramils-2
minisat-HACK-999ED-CSSC	0	0	1.86	**0.84**	0	0	1.83	**0.81**	smac-disc-3
Riss-4.27	0	0	2.55	**1.32**	1	**0**	3.72	**1.31**	paramils-2
SparrowToRiss	0	0	152.06	**1.44**	0	0	149.46	**1.45**	paramils-3

Table 10: Track: *Random SAT+UNSAT*; Benchmarks: *unif-k5*

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
clasp-3.0.4-p8	41	**0**	114.52	**35.83**	18	**0**	115.05	**35.03**	smac-2
DCCASat+march-rw	0	0	84.58	**19.35**	1	**0**	80.53	**18.94**	smac-1
minisat-HACK-999ED-CSSC	301	**171**	236.85	**178.96**	166	**99**	246.70	**190.03**	paramils-3
Riss-4.27	295	**221**	233.84	**207.40**	160	**113**	241.02	**210.81**	smac-3
SparrowToRiss	253	242	**162.00**	213.78	126	126	**156.64**	219.49	paramils-2

Table 11: Track: *Random SAT+UNSAT*; Benchmarks: *3cnf*

13

Solver	Training performance				Test performance			
	#Timeouts		PAR1		#Timeouts		PAR1	
	default	config.	default	config.	default	config.	default	config.
clasp-3.0.4-p8	41	0	42.20	13.42	18	0	41.23	12.66
DCCASat+march-rw	0	0	105.10	21.81	2	0	102.00	21.28
minisat-HACK-999ED-CSSC	308	171	91.36	67.33	171	100	93.10	68.57
Riss-4.27	302	223	91.58	78.45	163	115	90.89	77.51
SparrowToRiss	277	245	139.46	81.37	134	127	131.93	80.64

Table 12: Overall results for track: *Random SAT+UNSAT*; ranked by number of timeouts of configured solvers on test sets (tie-breaker: PAR1)
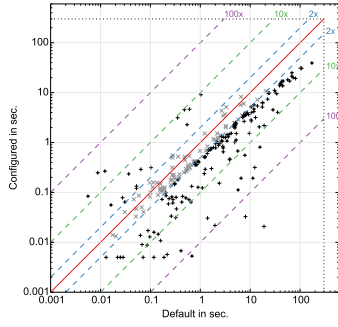
(a) *clasp-3.0.4-p8*; TOs: $0 \to 0$
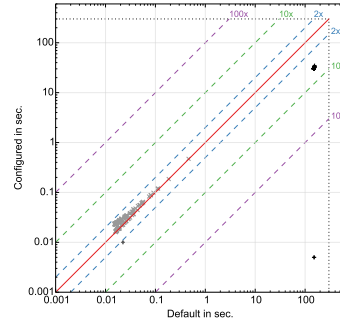
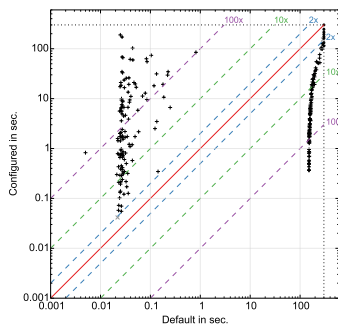(b) *DCCASat+march-rw*; TOs: $0 \to 0$
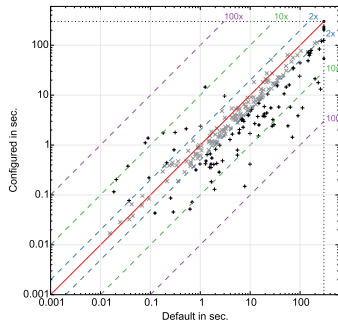
(c) *minisat-HACK-999ED-CSSC*; TOs: $5 \to 1$
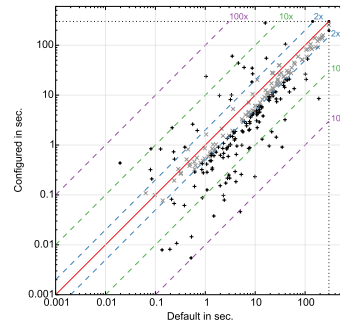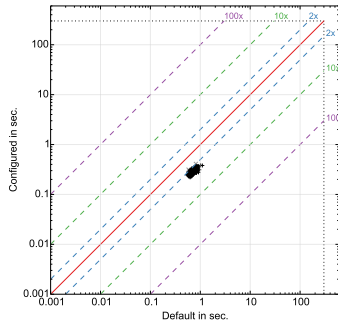
(d) *Riss-4.27*; TOs: $2 \to 2$
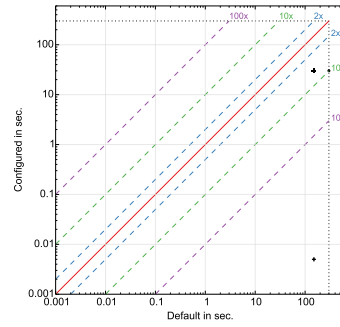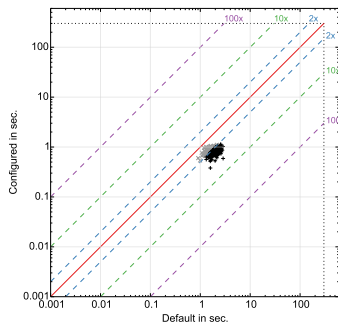
(e) *SparrowToRiss*; TOs: $8 \to 1$
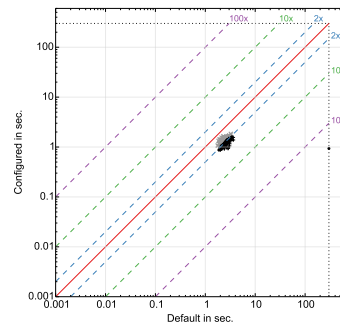
Figure 7: Track: *Random SAT+UNSAT*, Benchmarks: *K3*
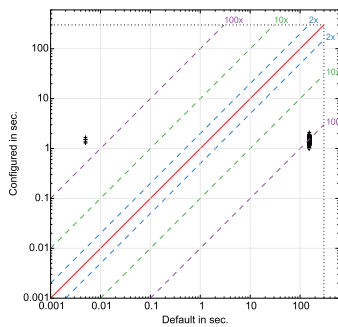
15

(a) *clasp-3.0.4-p8*; TOs: $0 \to 0$

(b) *DCCASat+march-rw*; TOs: $1 \to 0$

(c) *minisat-HACK-999ED-CSSC*; TOs: $0 \to 0$

(d) *Riss-4.27*; TOs: $1 \to 0$

(e) *SparrowToRiss*; TOs: $0 \to 0$

Figure 8: Track: *Random SAT+UNSAT*, Benchmarks: *unif-k5*

16

(a) *clasp-3.0.4-p8*; TOs: $18 \rightarrow 0$

(b) *DCCASat+march-rw*; TOs: $1 \rightarrow 0$
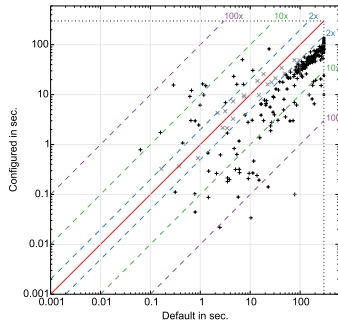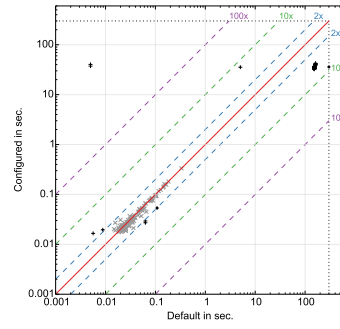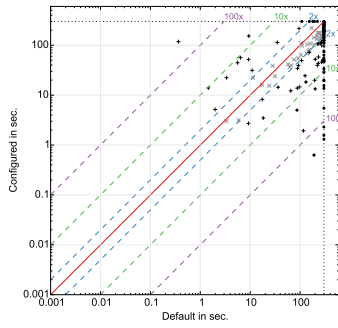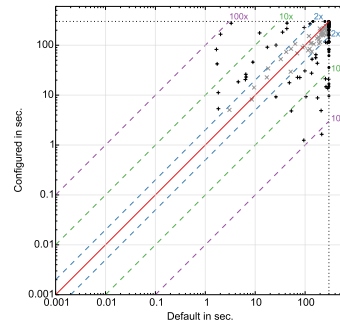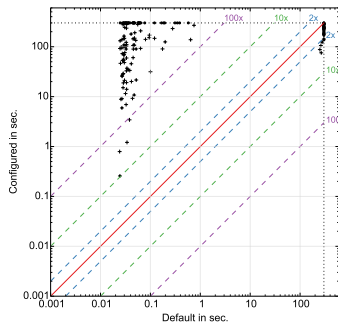
(c) *minisat-HACK-999ED-CSSC*; TOs: $166 \rightarrow 99$

(d) *Riss-4.27*; TOs: $160 \rightarrow 113$

(e) *SparrowToRiss*; TOs: $126 \rightarrow 126$

Figure 9: Track: *Random SAT+UNSAT*, Benchmarks: *3cnf*

2.4. Random SAT

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
probSAT	11	**2**	19.42	**5.94**	10	**4**	24.22	**9.33**	smac-disc-0
SparrowToRiss	11	**3**	18.54	**7.38**	9	**5**	18.61	**9.62**	smac-2
CSCCSat2014	2	**1**	6.47	5.76	2	2	5.64	6.01	smac-2
YalSAT	9	**3**	16.97	**12.72**	**6**	7	13.27	14.22	paramils-2
clasp-3.0.4-p8	250	250	300.00	300.00	250	250	300.00	300.00	smac-disc-0
minisat-HACK-999ED-CSSC	250	250	300.00	300.00	250	250	300.00	300.00	smac-disc-2

Table 13: Track: *Random SAT*; Benchmarks: *3sat1k*

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
probSAT	30	**2**	93.53	**24.16**	24	**0**	78.10	**14.72**	smac-disc-2
SparrowToRiss	13	**1**	32.80	**11.33**	3	3	19.75	**11.04**	smac-disc-0
CSCCSat2014	9	7	25.30	23.32	**3**	6	17.69	20.11	gga-disc-1
YalSAT	7	**4**	27.82	28.08	5	5	24.49	21.70	paramils-1
clasp-3.0.4-p8	250	**249**	300.00	**299.13**	250	250	300.00	300.00	smac-disc-0
minisat-HACK-999ED-CSSC	250	250	300.00	300.00	250	250	300.00	300.00	smac-disc-0

Table 14: Track: *Random SAT*; Benchmarks: *7sat90*

Solver	Training performance				Test performance				Configurator
	#Timeouts		PAR1		#Timeouts		PAR1		
	default	config.	default	config.	default	config.	default	config.	
probSAT	250	**0**	300.00	**1.88**	250	**0**	300.00	**1.97**	smac-3
SparrowToRiss	250	**0**	300.00	**6.57**	250	**0**	300.00	**6.23**	smac-disc-2
CSCCSat2014	0	0	7.08	7.07	0	0	**6.77**	6.80	paramils-1
YalSAT	0	0	6.80	**4.44**	0	0	4.89	4.60	smac-disc-0
clasp-3.0.4-p8	250	250	300.00	300.00	250	250	300.00	300.00	smac-disc-0
minisat-HACK-999ED-CSSC	250	250	300.00	300.00	250	250	300.00	300.00	smac-disc-0

Table 15: Track: *Random SAT*; Benchmarks: *5sat500*

Solver	Training performance				Test performance			
	#Timeouts		PAR1		#Timeouts		PAR1	
	default	config.	default	config.	default	config.	default	config.
probSAT	291	4	137.65	10.66	284	4	134.11	8.67
SparrowToRiss	274	4	117.11	8.43	262	8	112.78	8.96
CSCCSat2014	11	8	12.95	12.05	5	8	10.03	10.97
YalSAT	16	7	17.20	15.08	11	12	14.22	13.51
clasp-3.0.4-p8	750	749	300.00	299.71	750	750	300.00	300.00
minisat-HACK-999ED-CSSC	750	750	300.00	300.00	750	750	300.00	300.00

Table 16: Overall results for track: *Random SAT*; ranked by number of timeouts of configured solvers on test sets (tie-breaker: PAR1)
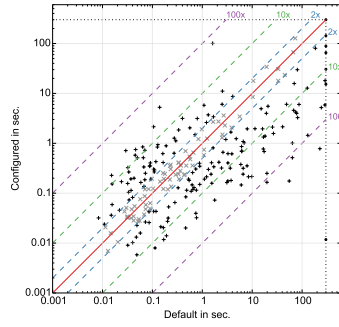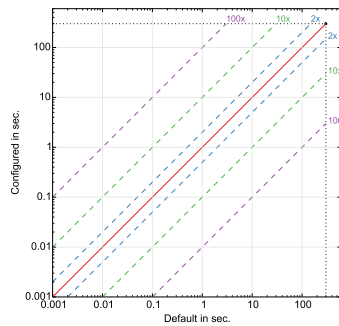
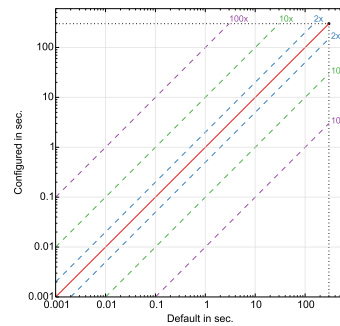(a) *probSAT*; TOs: $10 \rightarrow 4$
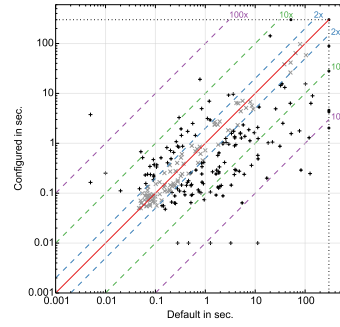
(b) *SparrowToRiss*; TOs: $9 \rightarrow 5$
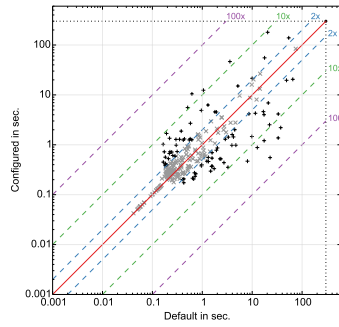
(c) *CSCCSat2014*; TOs: $2 \rightarrow 2$
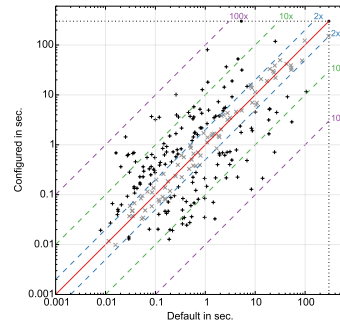
(d) *YalSAT*; TOs: $6 \rightarrow 7$

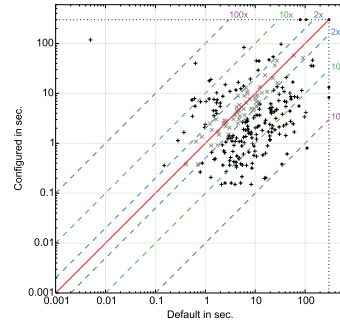(e) *clasp-3.0.4-p8*; TOs: $250 \rightarrow 250$
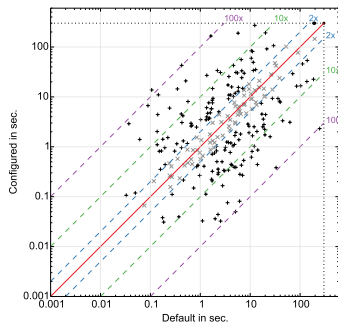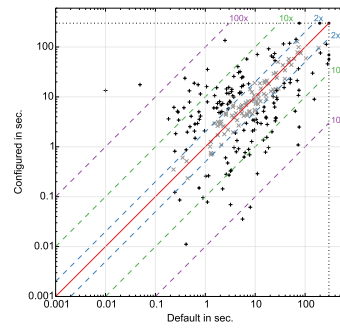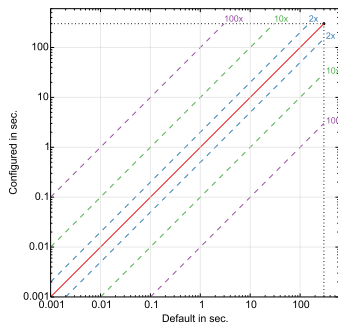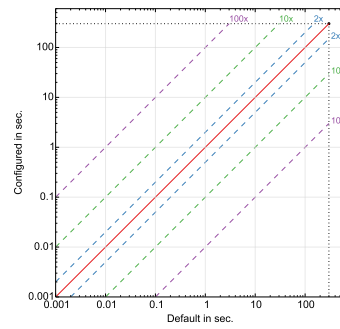

(f) *minisat-HACK-999ED-CSSC*; TOs: $250 \rightarrow 250$

Figure 10: Track: *Random SAT*, Benchmarks: *3sat1k*

20
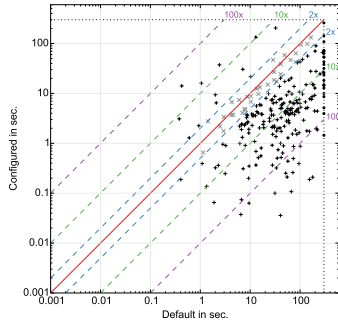
(a) *probSAT*; TOs: $24 \to 0$

(b) *SparrowToRiss*; TOs: $3 \to 3$

(c) *CSCCSat2014*; TOs: $3 \to 6$

(d) *YalSAT*; TOs: $5 \to 5$

(e) *clasp-3.0.4-p8*; TOs: $250 \to 250$

(f) *minisat-HACK-999ED-CSSC*; TOs: $250 \to 250$

Figure 11: Track: *Random SAT*, Benchmarks: *7sat90*

(a) *probSAT*; TOs: $250 \rightarrow 0$

(b) *SparrowToRiss*; TOs: $250 \rightarrow 0$

(c) *CSCCSat2014*; TOs: $0 \rightarrow 0$
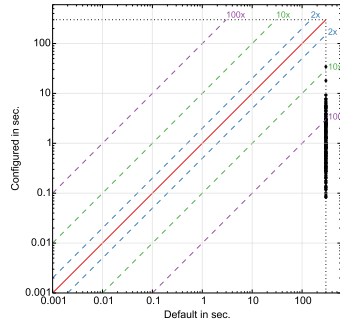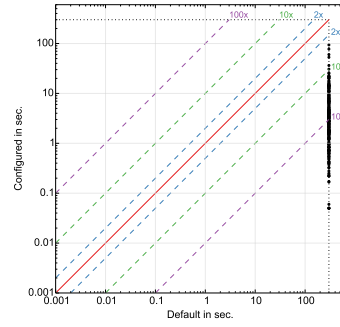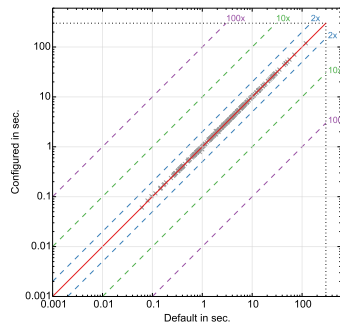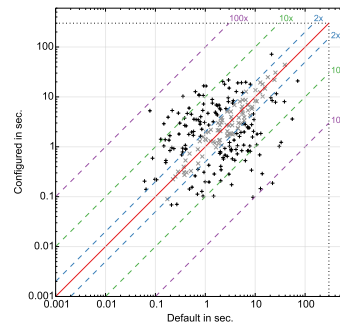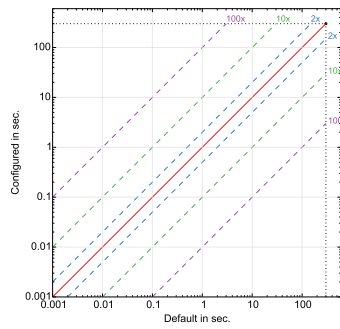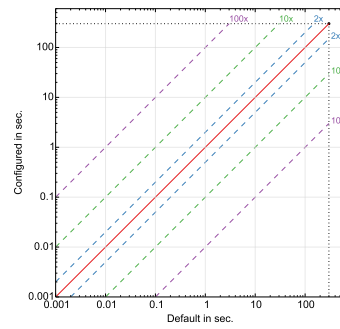
(d) *YalSAT*; TOs: $0 \rightarrow 0$

(e) *clasp-3.0.4-p8*; TOs: $250 \rightarrow 250$

(f) *minisat-HACK-999ED-CSSC*; TOs: $250 \rightarrow 250$

Figure 12: Track: *Random SAT*, Benchmarks: *5sat500*

[1] Balint, A., Manthey, N., 2014. SparrowToRiss. In: Belov, A., Diepold, D., Heule, M., Järvisalo, M. (Eds.), Proceedings of SAT Competition 2014. To appear.

[2] Balint, A., Schöning, U., 2012. Choosing probability distributions for stochastic local search and the role of make versus break. In: Theory and Applications of Satisfiability Testing–SAT 2012. Springer, pp. 16–29.

[3] Biere, A., 2014. Yet another local search solver and lingeling and friends entering the sat competition 2014. In: Belov, A., Diepold, D., Heule, M., Järvisalo, M. (Eds.), Proceedings of SAT Competition 2014. To appear.

[4] Gebser, M., Kaufmann, B., Schaub, T., 2012. Conflict-driven answer set solving: From theory to practice. Artificial Intelligence 187-188, 52–89.

[5] Luo, C., Cai, S., Wu, W., Su, K., 2013. Focused random walk with configuration checking and break minimum for satisfiability. In: Proceedings of the 2013 Conference on Constraint Programming (CP'13). pp. 481–496.

[6] Luo, C., Cai, S., Wu, W., Su, K., 2014. Double configuration checking in stochastic local search for satisfiability. In: Proceeding of the 2014 Conference on Artificial Intelligence (AAAI'14). p. to appear.

[7] Luo, C., Cai, S., Wu, W., Su, K., 2014. Double configuration checking in stochastic local search for satisfiability. In: Proceeding of the 2014 Conference on Artificial Intelligence (AAAI'14). p. to appear.

[8] Manthey, N., 2014. Riss 4.27. In: Belov, A., Diepold, D., Heule, M., Järvisalo, M. (Eds.), Proceedings of SAT Competition 2014. To appear.

[9] Soos, M., 2014. CryptoMiniSat v4. In: Belov, A., Diepold, D., Heule, M., Järvisalo, M. (Eds.), Proceedings of SAT Competition 2014. To appear.