

First Max-SAT Evaluation (SAT 2006)

Josep Argelich, Chu Min Li, Felip Manyà and Jordi Planes



Objectives

- Evaluate the performance of existing solvers
- Identify challenging benchmarks
- Identify successful solving techniques
- Stimulate researchers to improve their solvers and do more research on Max-SAT
- Gain new insights for future evaluations

SOLVERS

Max-SAT

- ChaffBS & ChaffLS
Fu & Malik
- Lazy
Alsinet, Manyà & Planes
- MaxSatz
Li, Manyà & Planes
- Sat4Jmaxsat
Le Berre
- Toolbar
de Givry, Heras, Larrosa & Schiex

Weighted Max-SAT

- Lazy
Alsinet, Manyà & Planes
- Sat4Jmaxsat
Le Berre
- Toolbar
de Givry, Heras, Larrosa & Schiex

Benchmarks

- Max-SAT instances
 - 700 instances with a timeout of 30 minutes
 - Random Max-2-SAT and Max-3-SAT without timeout
- Weighted Max-SAT instances
 - 450 instances with a timeout of 30 minutes
 - Random weighed Max-2-SAT and Max-3-SAT without timeout

Experimental Results I

Max-SAT

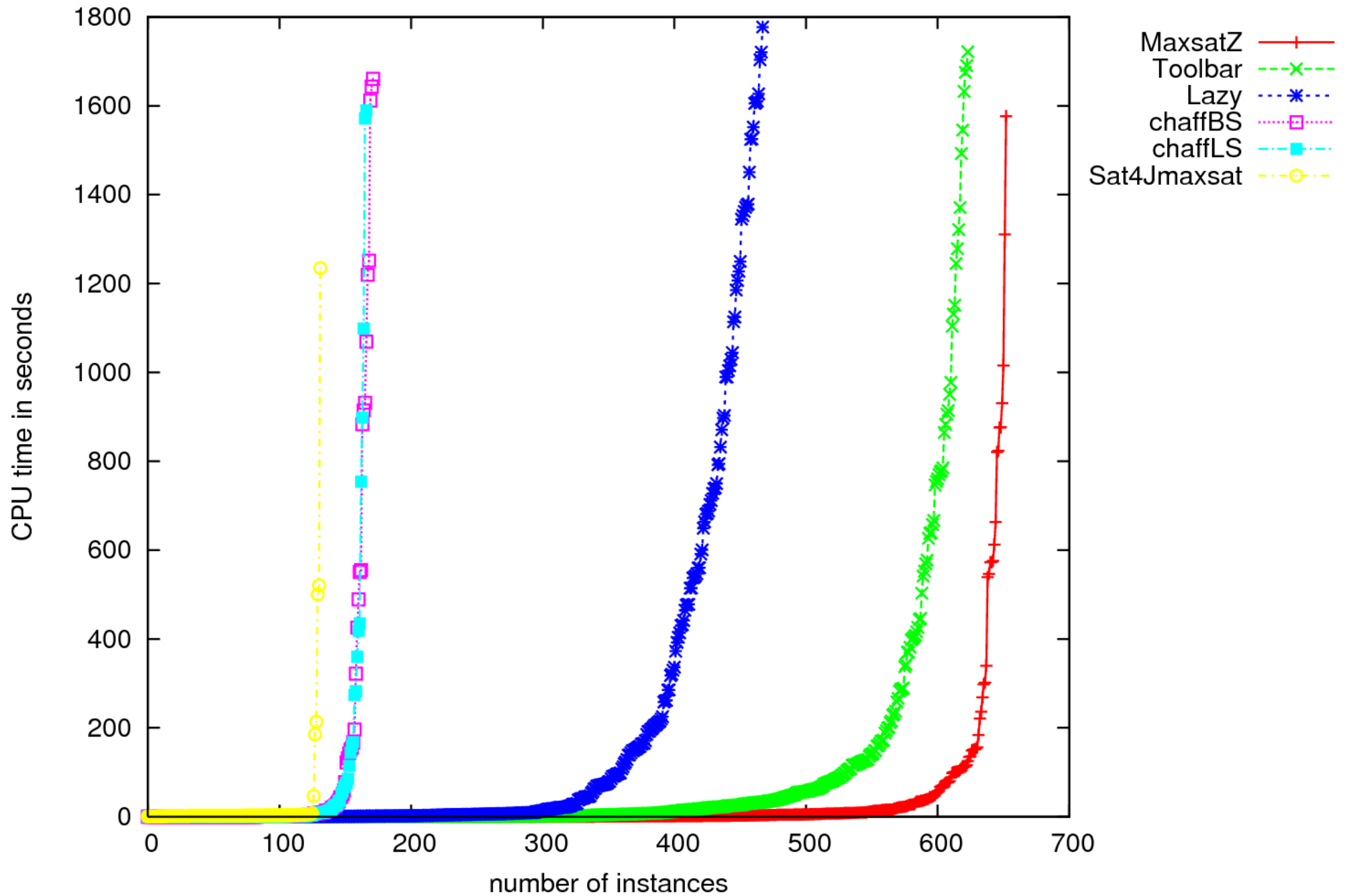
Experiments performed on a Linux cluster with 2GHz AMD Opteron processors with 1Gb of RAM provided by the Universitat de Lleida

Max-Cut, Max-One, Ramsey and Random Max-k-SAT instances

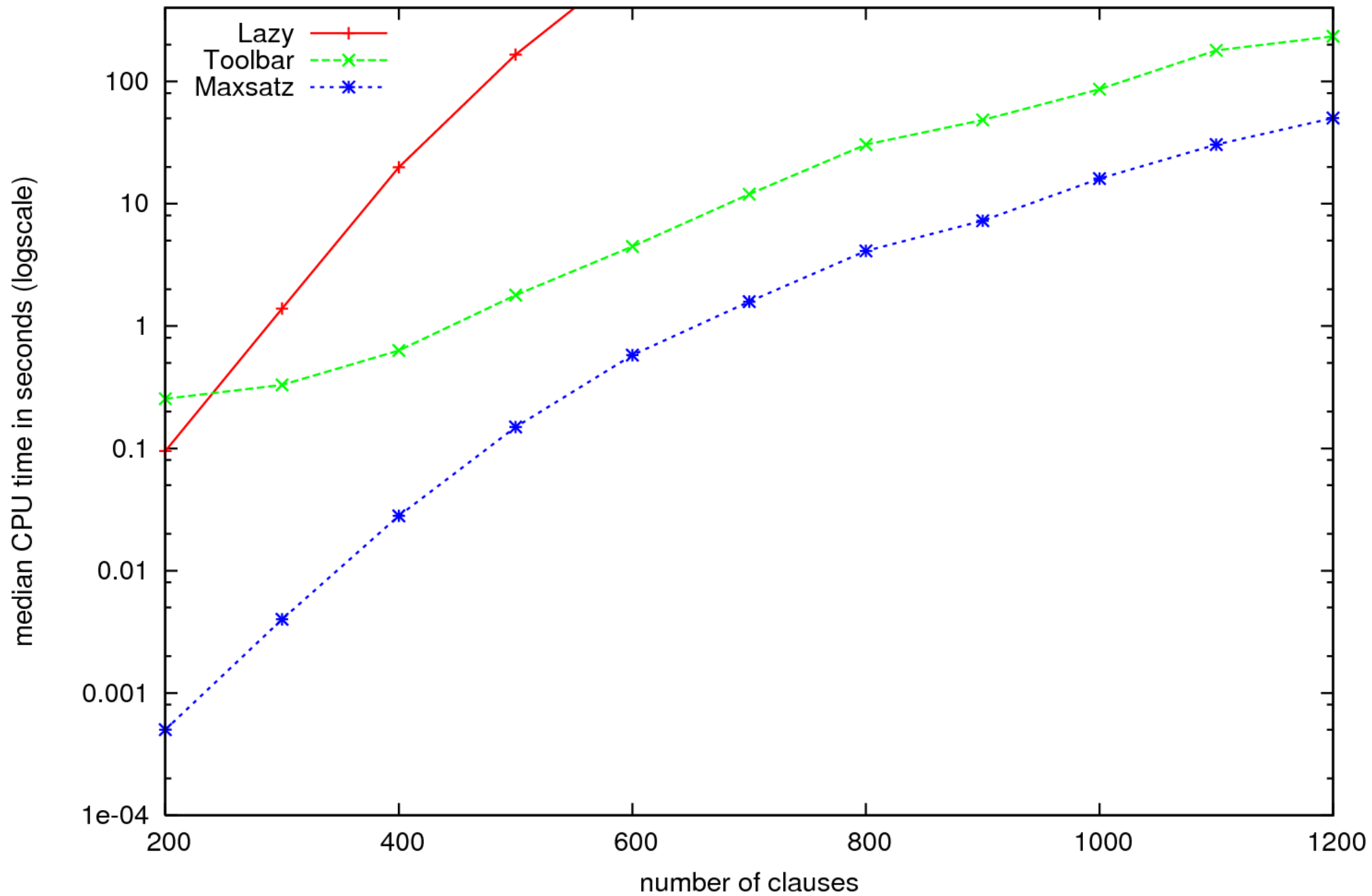
Set Name	#Instances	MaxSatz	Toolbar	Lazy	ChaffBS	ChaffLS	SAT4Jmaxsat
Max-Cut (brock)	12	13.35(12)	57.50(12)	178.48(12)	0.00(0)	0.00(0)	0.00(0)
Max-Cut (c-fat)	7	0.07(5)	21.05(5)	151.13(5)	0.01(2)	0.01(2)	0.85(2)
Max-Cut (hamming)	6	180.12(3)	575.52(3)	42.06(2)	0.00(0)	0.00(0)	0.00(0)
Max-Cut (johnson)	4	45.39(3)	134.68(3)	2.45(2)	0.00(0)	0.00(0)	0.00(0)
Max-Cut (keller)	2	6.12(2)	17.25(2)	69.86(2)	0.00(0)	0.00(0)	0.00(0)
Max-Cut (p_hat)	12	15.84(12)	61.86(12)	192.05(12)	0.00(0)	0.00(0)	0.00(0)
Max-Cut (san)	11	275.05(11)	65.02(7)	249.83(7)	0.00(0)	0.00(0)	0.00(0)
Max-Cut (sanr)	4	71.98(4)	266.86(4)	80.78(3)	0.00(0)	0.00(0)	0.00(0)
Max-Cut (random)	40	5.58(40)	34.67(40)	752.34(25)	0.00(0)	0.00(0)	0.00(0)
Max-Cut (spinglass)	5	44.92(3)	4.96(2)	48.21(2)	9.97(1)	6.19(1)	0.00(0)
Max-One	45	0.02(45)	5.44(45)	81.34(40)	1.00(45)	0.20(45)	2.31(41)
Ramsey	48	8.99(34)	53.14(33)	81.70(28)	53.39(34)	7.36(33)	2.86(32)
Max-2-SAT (60 vars)	50	0.03(50)	0.62(50)	3.27(50)	13.74(10)	25.69(10)	0.00(0)
Max-2-SAT (100 vars)	50	1.40(50)	17.57(50)	235.83(31)	0.70(10)	1.08(10)	24.37(10)
Max-2-SAT (140 vars)	50	7.02(50)	105.61(49)	204.10(23)	272.77(12)	99.86(11)	47.26(11)
Max-2-SAT (discarded)	180	16.79(180)	99.34(175)	141.39(107)	262.04(18)	172.67(14)	59.87(4)
Max-3-SAT (40 vars)	50	1.50(50)	8.09(50)	6.94(50)	0.31(10)	0.28(10)	50.05(11)
Max-3-SAT (60 vars)	50	23.31(50)	264.98(50)	266.70(43)	84.76(11)	68.55(11)	1.96(10)

Instances contributed by de Givry, Heras, Larrosa & Schiex

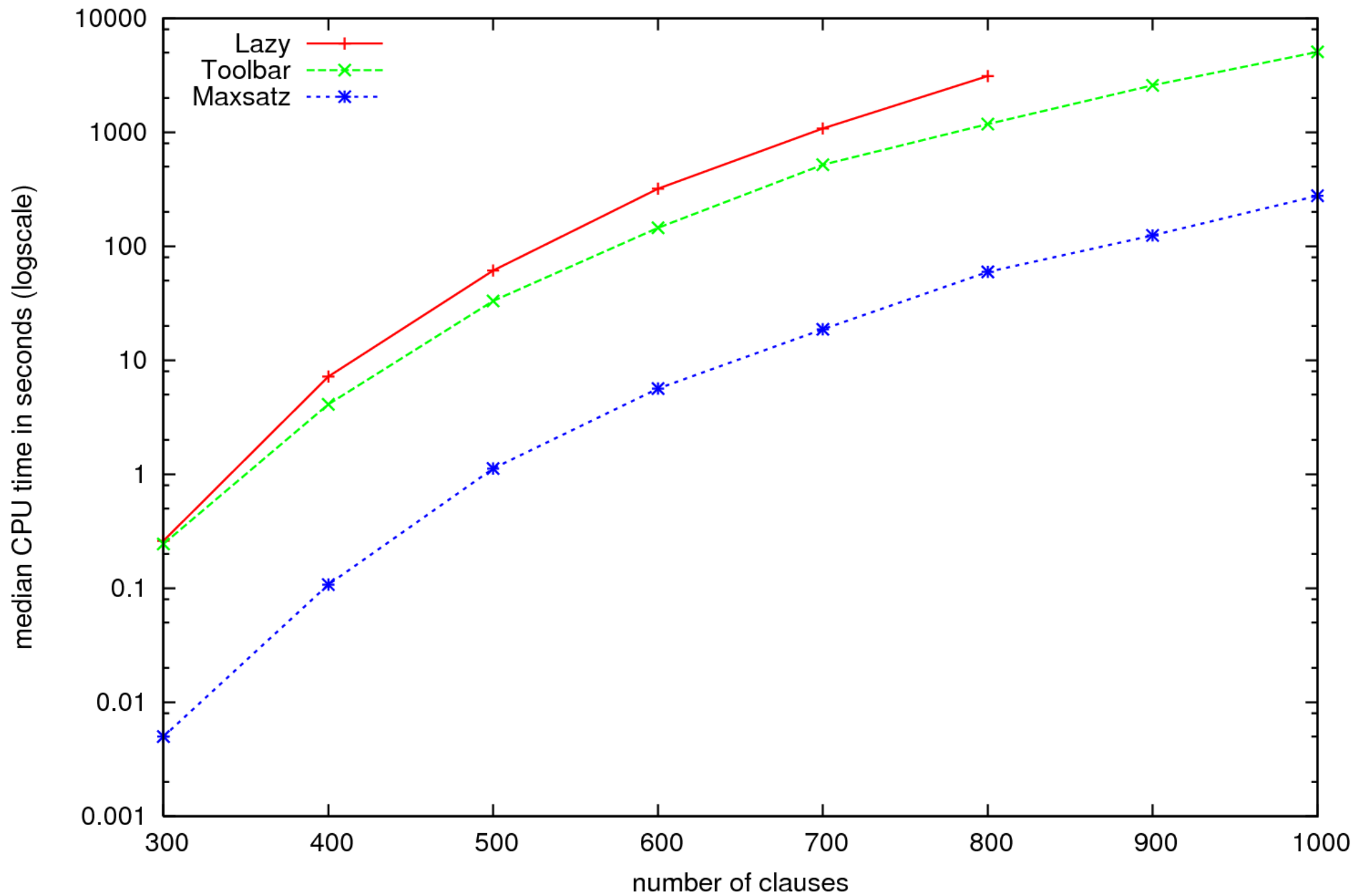
Number x of instances solved in y seconds (unweighted category)



Random Max-2-SAT (100 variables)



Random Max-3-SAT (70 variables)



Experimental Results II

Weighted Max-SAT

Experiments performed on a Linux cluster with 2GHz AMP Opteron processors with 1 Gb of RAM provided by the Universitat de Lleida

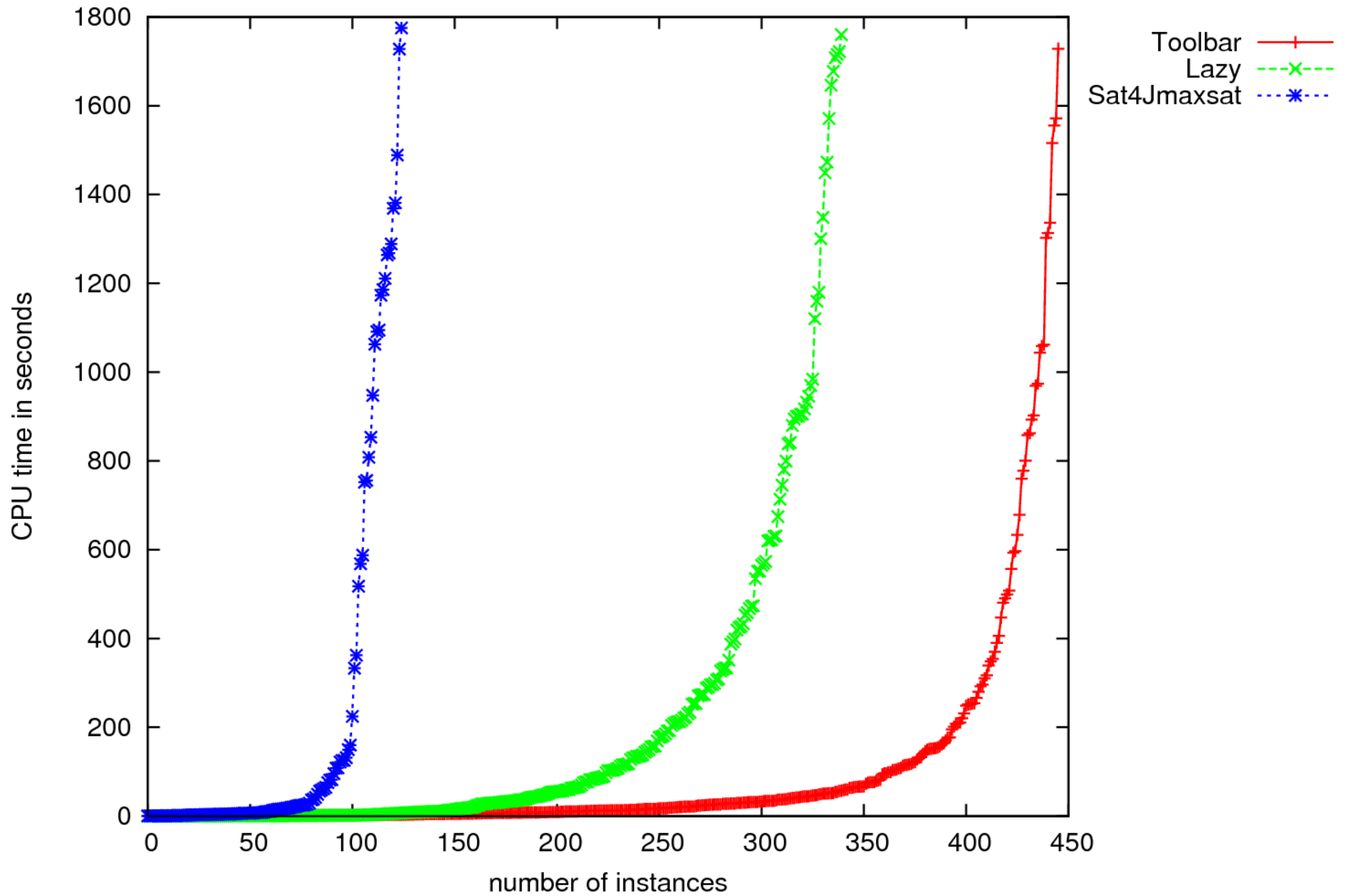
Set Name	#Instances	Toolbar	Lazy	SAT4Jmaxsat
Auction (paths)	30	249.77(26)	81.24(20)	0.00(0)
Auction (regions)	30	8.16 (30)	2.03 (28)	926.99(6)
Auction (scheduling)	30	132.15(30)	63.33(30)	518.41(8)
Max-Clique (brock)	12	96.76(4)	104.69(4)	0.00(0)
Max-Clique (c-fat)	7	25.19(7)	17.36(7)	346.68(4)
Max-Clique (hamming)	6	134.04(5)	195.05(5)	6.32 (2)
Max-Clique (johnson)	4	53.91(3)	38.64(3)	61.73(2)
Max-Clique (keller)	2	34.12(1)	43.38(1)	0.00(0)
Max-Clique (mann_a)	4	45.62(3)	0.31 (1)	726.50(2)
Max-Clique (p_hat)	12	325.70(8)	254.14(6)	0.00(0)
Max-Clique (san)	11	25.01(3)	10.88(1)	0.00(0)
Max-Clique (sanr)	4	821.98(3)	790.55(2)	0.00(0)

Instances contributed by de Givry, Heras, Larrosa & Schiex

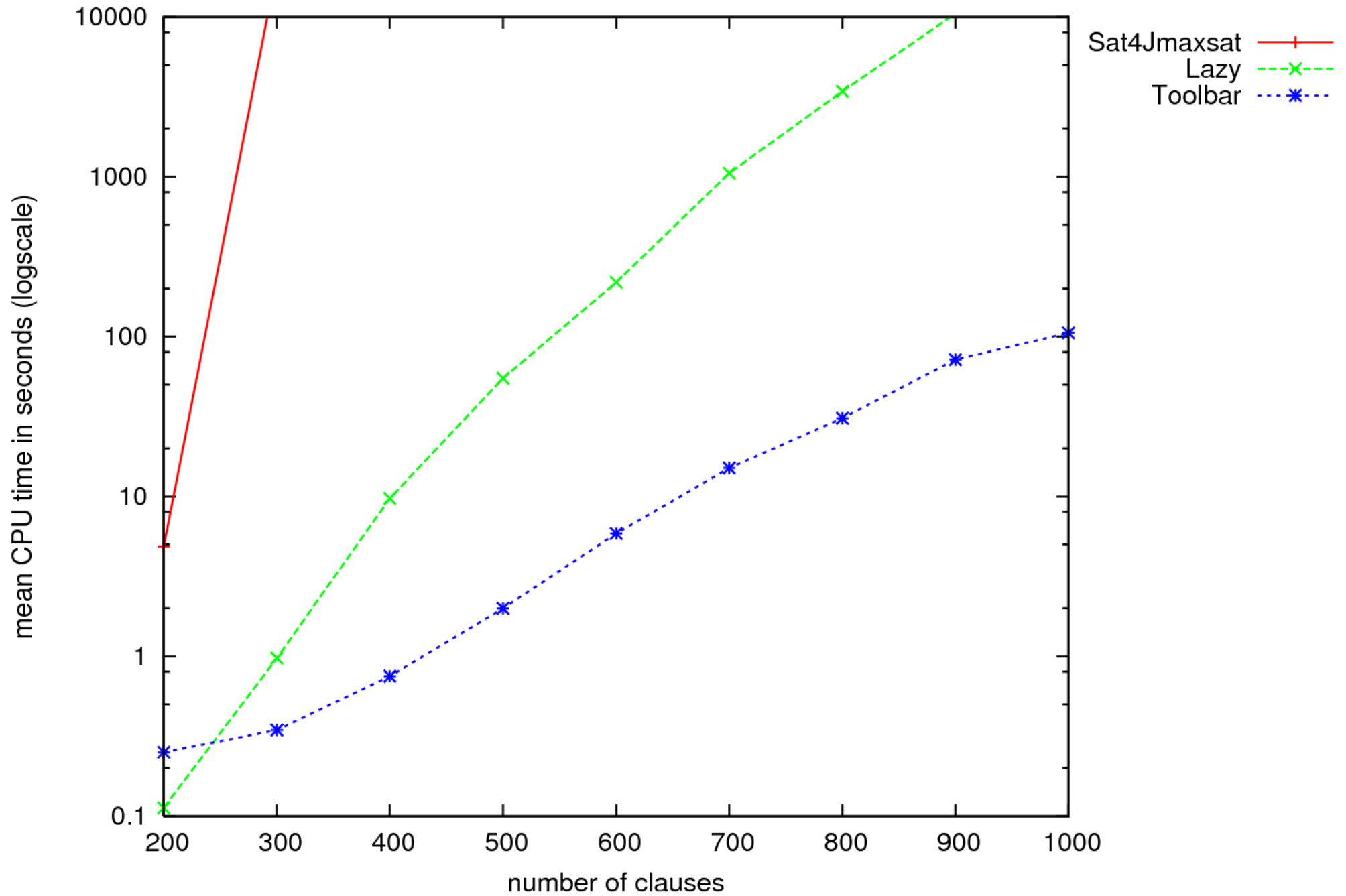
Set Name	#Instances	Toolbar	Lazy	SAT4Jmaxsat
Weighted Max-Cut (brock)	12	12.37(12)	18.01(12)	0.00(0)
Weighted Max-Cut (c-fat)	7	7.80 (7)	25.99(7)	1.07 (2)
Weighted Max-Cut (hamming)	6	105.22(4)	88.93(4)	0.00(0)
Weighted Max-Cut (johnson)	4	71.39(3)	74.29(3)	0.00(0)
Weighted Max-Cut (keller)	2	13.46(2)	17.43(2)	0.00(0)
Weighted Max-Cut (mann_a)	4	1155.02(2)	1015.62(4)	0.00(0)
Weighted Max-Cut (p_hat)	12	11.42(12)	10.92(12)	0.00(0)
Weighted Max-Cut (san)	11	82.66(11)	57.33(11)	0.00(0)
Weighted Max-Cut (sanr)	4	42.26(4)	25.90(4)	0.00(0)
Weighted Max-Cut (random)	40	11.57(40)	246.96(40)	0.00(0)
Weighted Max-Cut (spinglass)	5	40.50(3)	0.26 (2)	0.00(0)
Max-One	45	122.37(45)	343.82(27)	472.66(9)
Quasigroup Completion	25	112.33(10)	94.58(6)	22.04(24)
Ramsey	48	19.04(35)	54.79(29)	32.54(33)
Weighted CSP (dense-loose)	40	236.81(34)	527.02(32)	0.00(0)
Weighted CSP (dense-tight)	60	30.68(30)	0.00(0)	0.00(0)
Weighted CSP (sparse-loose)	40	94.13(32)	392.13(27)	299.68(25)
Weighted CSP (sparse-tight)	40	40.96(20)	0.00(0)	0.00(0)
Weighted CSP (spot)	42	75.77(12)	15.37(5)	35.34(4)

Instances contributed by de Givry, Heras, Larrosa & Schiex

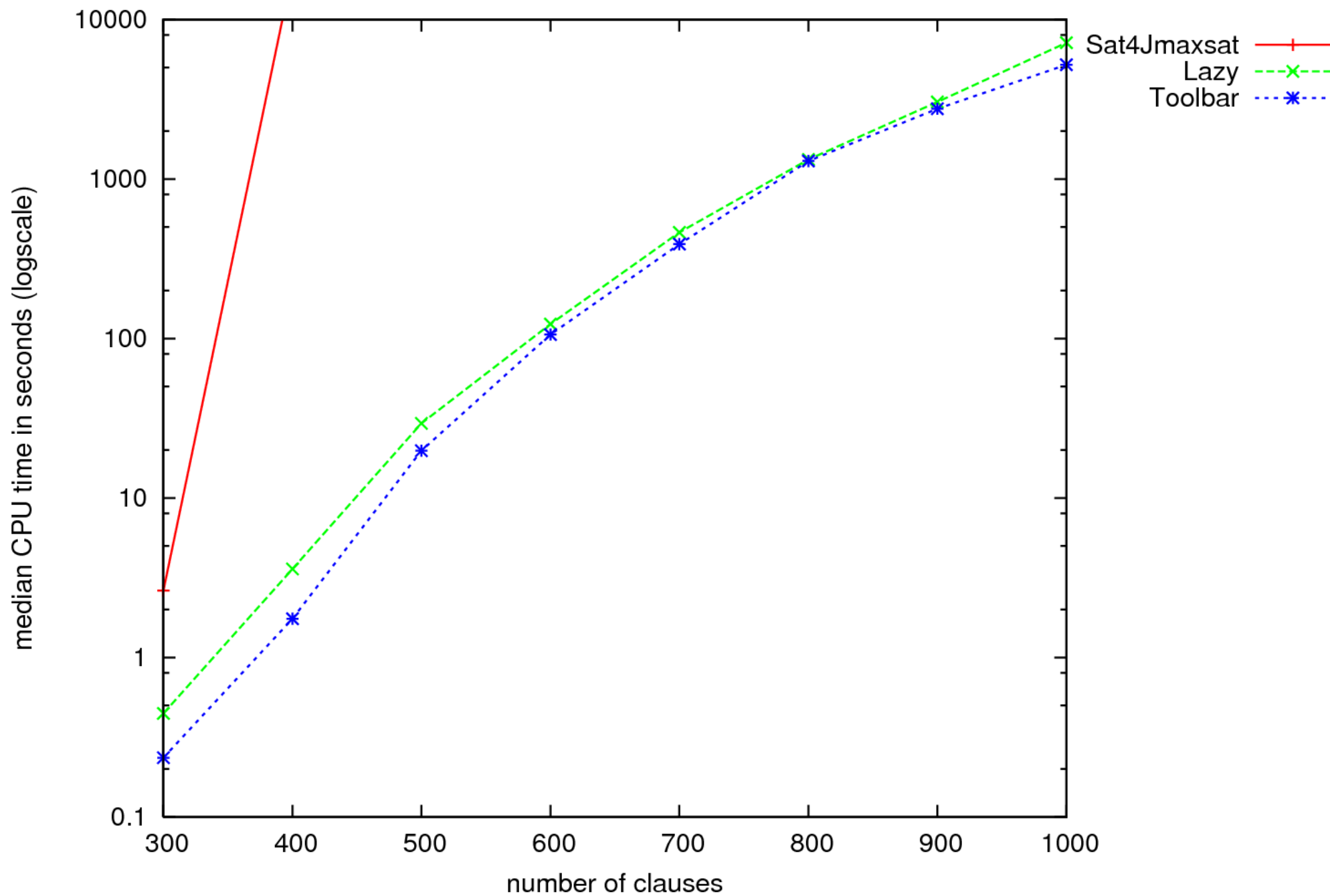
Number x of instances solved in y seconds (weighted category)



Random Weighted Max-2-SAT (100 variables)



Random Weighted Max-3-SAT (70 variables)



Thanks to the people that contributed solvers and benchmarks, and to the Universitat de Lleida for allowing to use its cluster.

We encourage you to participate in the next Max-SAT evaluation.