CliSAT: a SAT-based exact algorithm for hard maximum clique problems

Pablo San Segundo[1]

Universidad Politécnica de Madrid (UPM), Centre for Automation and Robotics (CAR), Madrid, Spain, pablo.sansegundo@upm.es

Fabio Furini

Department of Computer, Control and Management Engineering Antonio Ruberti, Sapienza University of Rome, Rome, Italy, fabio.furini@uniroma1.it

David Alvarez

Universidad Politécnica de Madrid (UPM), Madrid, Spain, david.asanchez@upm.es

Panos Pardalos

Center for Applied Optimization, University of Florida (UF), Florida, USA, pardalos@ise.ufl.edu

Abstract

Given a graph, the maximum clique problem (MCP) asks for determining a complete subgraph with the largest possible number of vertices. We propose a new exact algorithm, called `CliSAT`, to solve the MCP to proven optimality. This problem is of fundamental importance in graph theory and combinatorial optimization due to its practical relevance for a wide range of applications. The newly developed exact approach is a combinatorial branch-and-bound algorithm that exploits the state-of-the-art branching scheme enhanced by two new bounding techniques with the goal of reducing the branching tree. The first one is based on graph colouring procedures and partial maximum satisfiability problems arising in the branching scheme. The second one is a filtering phase based on constraint programming and domain propagation techniques. `CliSAT` is designed for structured MCP instances which are computationally difficult to solve since they are dense and contain many interconnected large cliques. Extensive experiments on hard benchmark instances, as well as new hard instances arising from different applications, show that `CliSAT` outperforms the state-of-the-art MCP algorithms by several orders of magnitude.

Keywords: Combinatorial Optimization, Exact Algorithm, Branch-and-Bound Algorithm, Maximum Clique Problem.

[1]Corresponding author

Appendix

Extended comparison with state-of-the-art algorithms

This section extends the performance comparison between the algorithms `CliSAT` and `MoMC` reported in the computational section of the article, providing details for individual instances from the `DIMACS` (Table 1) and `BHOSHLIB` (Table 2) datasets, as well as the `evil` family from `MISCLIB` and the `D` family from `CSPLIB` (Table 3). All the tables report the number of steps (number of recursive calls), and the time in seconds spent by both algorithms to prove optimality. Instances with an entry of 0 steps indicate that the corresponding algorithm was able to prove optimality during its initialization phase.

Table 1: Extended performance comparison between the algorithms `CliSAT` and `MoMC` over a subset of 38 `DIMACS` representative instances.

				CliSAT		MoMC	
name	$\lvert V\rvert$	$d(G)$	$\omega(G)$	steps	time [sec]	steps	time [sec]
brock200_1	200	0.75	21	6,459	**0.2**	72,181	0.6
brock400_1	400	0.75	27	4,781,237	137.6	10,462,110	**112.8**
brock400_2	400	0.75	29	2,560,166	**76.7**	8,521,285	93.5
brock400_3	400	0.75	31	851,004	**34.0**	8,357,671	67.4
brock400_4	400	0.75	33	303,585	**14.8**	2,384,115	19.0
brock800_1	800	0.65	23	107,277,894	3,652.2	235,564,546	**1,828.2**
brock800_2	800	0.65	24	66,145,240	2,929.1	223,408,604	**1,867.1**
brock800_3	800	0.65	25	49,625,642	1,867.4	108,221,563	**752.3**
brock800_4	800	0.65	26	24,581,677	**1,220.1**	175,510,646	1,266.8
C250.9	250	0.90	44	3,417,260	**128.9**	6,639,713	130.6
C2000.5	2,000	0.50	16	675,567,772	**27,660.5**	3,185,621	34,936.7
dsjc500.5	500	0.50	13	33,179	**0.9**	280,166	1.8
dsjc1000.5	1,000	0.50	15	3,029,020	**86.2**	17,047,147	112.3
gen400_p0.9_55	400	0.90	55	0	**0.1**	4,049	1.1
gen400_p0.9_65	400	0.90	65	1	**0.1**	3,154	0.4
gen400_p0.9_75	400	0.90	75	1	**0.1**	3,360	0.2
hamming10-2	1,024	0.99	512	0	**0.1**	131,840	24.5
keller5	776	0.75	27	222,737	**23.1**	7,821,773	157.7
MANN_a45	1,035	0.996	345	19,130	**5.5**	83,195	8.7
MANN_a81	3,321	0.999	1100	179,444,376	**361,440.5**	406,990,467	970,637.5
p_hat300-3	300	0.74	36	2,555	**0.2**	9,336	0.3
p_hat500-2	500	0.50	36	305	**0.1**	4,075	0.2
p_hat500-3	500	0.75	50	75,644	**6.6**	372,329	9.3
p_hat700-2	700	0.50	44	2,547	**0.3**	16,365	0.8
p_hat700-3	700	0.75	62	650,084	**78.4**	2,678,530	100.5
p_hat1000-1	1,000	0.24	10	5,244	**0.2**	66,268	0.4
p_hat1000-2	1,000	0.49	46	132,244	**13.6**	749,907	17.7
p_hat1000-3	1,000	0.74	68	110,015,398	16,289.6	358,413,660	**14,453.0**
p_hat1500-1	1,500	0.33	12	79,771	**2.0**	498,585	3.1
p_hat1500-2	1,500	0.25	65	3,580,545	629.6	14,955,067	**565.9**
san400_0.7_1	400	0.70	40	1,436	**0.1**	5,185	0.3
san400_0.7_2	400	0.70	30	1,006	**0.1**	882	0.2
san400_0.7_3	400	0.70	22	0	**0.1**	1,153	0.3
san400_0.9_1	400	0.90	100	1	**0.1**	6,591	0.3
san1000	1,000	0.50	15	2,196	**0.2**	17,651	1.3
sanr200_0.9	200	0.90	42	31,120	**1.2**	74,213	1.3
sanr400_0.5	400	0.50	13	8,228	**0.2**	62,483	0.4
sanr400_0.7	400	0.70	21	1,588,204	**39.1**	5,518,655	51.5

Concerning the choice of the sorting procedure during the execution of the initial preprocessing phase of `CliSAT`, we report the following. In the case of the instances reported from the `BHOSHLIB` dataset (Table 2) and the `D` family (Table 3), `CliSAT` always selects the ordering of vertices determined by `COLOR-SORT`. In the case of the `DIMACS` instances (Table 1), the choice is as follows:

(i) COLOR-SORT is selected in 4 instances out of the 9 brock instances reported; (ii) COLOR-SORT is selected for the families gen and keller; (iii) DEG-SORT is selected in the remaining families, i.e. c-fat, dsjc, hamming, MANN, p_hat and san. Finally, in the case of the evil family (Table 3), DEG-SORT is invariably the choice.

Table 2: Extended performance comparison of the algorithms CliSAT and MoMC over the 41 instances of the BHOSHLIB dataset.The time limit (tl) was set to 15 days.

				CliSAT		MoMC	
name	$\|V\|$	$d(G)$	$\omega(G)$	steps	time [sec]	steps	time [sec]
frb30-15-1	450	0.82	30	0	**0.1**	918	0.3
frb30-15-2	450	0.82	30	0	**0.1**	985	0.3
frb30-15-3	450	0.82	30	0	**0.1**	1,095	0.4
frb30-15-4	450	0.82	30	0	**0.1**	1,155	0.4
frb30-15-5	450	0.82	30	126	**0.1**	981	0.4
frb35-17-1	595	0.84	35	0	**0.1**	2,158	1.2
frb35-17-2	595	0.84	35	643	**0.4**	2,928	1.5
frb35-17-3	595	0.84	35	0	**0.1**	1,409	0.7
frb35-17-4	595	0.84	35	102	**0.1**	1,427	0.7
frb35-17-5	595	0.84	35	221	**0.2**	1,727	0.9
frb40-19-1	760	0.86	40	0	**0.1**	2,816	2.1
frb40-19-2	760	0.86	40	121	**0.1**	1,700	1.3
frb40-19-3	760	0.86	40	0	**0.1**	2,660	2.2
frb40-19-4	760	0.86	40	4,140	**2.9**	10,881	5.8
frb40-19-5	760	0.86	40	1,938	**1.8**	8,399	5.0
frb45-21-1	945	0.87	45	2,434	**2.0**	102,403	94.0
frb45-21-2	945	0.87	45	42,513	**28.1**	61,069	60.4
frb45-21-3	945	0.87	45	19,794	**15.1**	29,443	34.1
frb45-21-4	945	0.87	45	13,243	**9.8**	26,477	27.6
frb45-21-5	945	0.87	45	113,072	**100.7**	158,421	168.1
frb50-23-1	1150	0.88	50	333,199	**315.4**	613,547	604.7
frb50-23-2	1150	0.88	50	117,948	**142.0**	214,297	268.2
frb50-23-3	1150	0.88	50	2,149,755	**2,534.4**	4,902,490	5,932.0
frb50-23-4	1150	0.88	50	6,401	**4.3**	12,354	14.8
frb50-23-5	1150	0.88	50	66,901	**68.0**	225,983	183.5
frb53-24-1	1272	0.88	53	1,271,221	**1,557.3**	1,981,396	3,415.1
frb53-24-2	1272	0.88	53	152,126	217.7	126,751	**152.0**
frb53-24-3	1272	0.88	53	747,174	**800.8**	977,570	1,511.1
frb53-24-4	1272	0.88	53	1,320,753	**1,555.4**	1,893,507	2,986.5
frb53-24-5	1272	0.88	53	168,563	**177.1**	428,286	725.4
frb56-25-1	1400	0.89	56	23,724,283	**33,772.0**	41,817,946	73,432.0
frb56-25-2	1400	0.89	56	1,290,189	**1,275.9**	1,355,963	1,982.4
frb56-25-3	1400	0.89	56	3,780,230	**4,218.6**	4,895,951	9,824.0
frb56-25-4	1400	0.89	56	4,668,820	**6,628.6**	13,978,117	19,427.5
frb56-25-5	1400	0.89	56	36,028,666	**53,642.3**	100,513,778	121,379.9
frb59-26-1	1534	0.89	59	62,599,285	**107,239.4**	138,978,307	257,586.4
frb59-26-2	1534	0.89	59	75,062,511	**108,058.4**	64,467,920	178,912.5
frb59-26-3	1534	0.89	59	35,806,103	**56,605.6**	53,434,151	112,232.2
frb59-26-4	1534	0.89	59	44,200,880	**76,077.7**	83,521,902	172,087.0
frb59-26-5	1534	0.89	59	15,343,706	18,326.1	4,451,221	**9,729.0**
frb100-40	4000	0.93	100	-	tl	-	tl

Table 3: Extended performance comparison between the algorithms `CliSAT` and `MoMC` over: *i*) the 20 instances of the `evil` family (`MISCLIB`) with a time limit (*tl*) set to 15 days); *ii*) the 25 instances of the `D` family (`CSPLIB`) with a time limit set to $1,800$ seconds.

name	$\|V\|$	$d(G)$	$\omega(G)$	CliSAT steps	CliSAT time [sec]	MoMC steps	MoMC time [sec]
evil-N120-p98-chv12x10	120	0.92	20	2,782	**0.1**	392,883	1.2
evil-N120-p98-myc5x24	120	0.97	48	47	0.05	1,505	**0.02**
evil-N121-p98-myc11x11	121	0.93	22	4,675	**0.1**	572,074	1.3
evil-N125-p98-s3m25x5	125	0.89	20	13,728	**0.2**	606,103	1.4
evil-N138-p98-myc23x6	138	0.87	12	82,781	**0.6**	983,869	2.6
evil-N150-p98-myc5x30	150	0.97	60	112	0.05	2,148	**0.02**
evil-N150-p98-s3m25x6	150	0.90	24	121,391	**1.2**	14,702,870	36.7
evil-N154-p98-myc11x14	154	0.94	28	109,508	**1.2**	17,449,596	54.2
evil-N180-p98-chv12x15	180	0.94	30	2,156,387	**21.9**	1,582,875,479	4,863.0
evil-N180-p98-myc5x36	180	0.97	72	115	0.06	3,259	**0.04**
evil-N184-p98-myc23x8	184	0.90	16	2,195,219	**17.0**	54,927,834	183.5
evil-N187-p98-myc11x17	187	0.95	34	1,838,267	**19.5**	1,371,470,102	5,115.4
evil-N200-p98-s3m25x8	200	0.92	32	8,425,011	**101.3**	976,980,140	3,491.2
evil-N210-p98-myc5x42	210	0.98	84	236	**0.1**	4,219	0.1
evil-N220-p98-myc11x20	220	0.95	40	78,774,365	**889.3**	-	tl
evil-N230-p98-myc23x10	230	0.91	20	145,397,825	**1,237.0**	1,756,363,669	53,718.4
evil-N240-p98-chv12x20	240	0.95	40	1,160,983,608	**13,353.0**	-	tl
evil-N240-p98-myc5x48	240	0.97	96	138	**0.1**	5,248	**0.1**
evil-N250-p98-s3m25x10	250	0.93	40	893,359,445	**13,057.7**	1,311,951,648	165,792.0
evil-N253-p98-myc11x23	253	0.95	46	4,643,934,432	**54,828.4**	-	tl
rand-2-40-8-753-010-04	320	0.88	39	501	**0.6**	3,162	0.7
rand-2-40-8-753-010-32	320	0.89	40	0	**0.1**	2,144	0.4
rand-2-40-8-753-010-60	320	0.88	39	265	**0.4**	3,028	0.6
rand-2-40-8-753-010-88	320	0.88	39	615	**0.8**	4,528	1.1
rand-2-40-11-414-020-00	440	0.87	39	956	**1.0**	4,409	1.5
rand-2-40-11-414-020-28	440	0.87	39	828	**0.8**	5,363	1.5
rand-2-40-11-414-020-56	440	0.87	40	459	**0.4**	2,947	0.9
rand-2-40-11-414-020-84	440	0.87	40	246	**0.2**	1,902	0.6
rand-2-40-16-250-035-12	640	0.87	40	107	**0.1**	2,247	1.1
rand-2-40-16-250-035-40	640	0.87	39	3,979	**2.3**	7,657	4.6
rand-2-40-16-250-035-68	640	0.87	39	775	**0.6**	3,124	1.5
rand-2-40-16-250-035-96	640	0.87	39	2,634	**1.6**	5,873	4.1
rand-2-40-180-84-090-24	7200	0.88	40	311	**13.7**	-	tl
rand-2-40-180-84-090-52	7200	0.88	40	147,375	**858.4**	-	tl
rand-2-40-180-84-090-80	7200	0.88	39	91,718	**191.7**	-	tl
rand-2-40-25-180-050-08	1000	0.86	40	638	**0.5**	2,904	2.9
rand-2-40-25-180-050-36	1000	0.86	40	3,578	**1.5**	6,005	5.0
rand-2-40-25-180-050-64	1000	0.86	39	872	**0.5**	2,603	3.2
rand-2-40-25-180-050-92	1000	0.86	40	474	**0.4**	3,235	2.8
rand-2-40-40-135-065-20	1600	0.87	40	179	**0.8**	2,497	7.0
rand-2-40-40-135-065-48	1600	0.87	40	2,599	**2.5**	4,079	9.3
rand-2-40-40-135-065-76	1600	0.87	39	6,188	**4.9**	13,064	18.5
rand-2-40-80-103-080-16	3200	0.87	39	47,252	**137.8**	38,205	191.5
rand-2-40-80-103-080-44	3200	0.87	40	74,117	**61.8**	9,158	93.2
rand-2-40-80-103-080-72	3200	0.87	40	739	**9.9**	5067	52.0