

Exploiting Resolution-based Representations for MaxSAT Solving

Miguel Neves¹, Ruben Martins^{2,3}, Mikoláš Janota¹,
Inês Lynce¹, and Vasco Manquinho¹

¹ INESC-ID / Instituto Superior Técnico, Universidade de Lisboa, Portugal
`{neves,mikolas,ines,vmm}@sat.inesc-id.pt`

² University of Texas at Austin, USA

³ University of Oxford, Department of Computer Science, United Kingdom
`ruben.martins@cs.ox.ac.uk`

Abstract. Most recent MaxSAT algorithms rely on a succession of calls to a SAT solver in order to find an optimal solution. In particular, several algorithms take advantage of the ability of SAT solvers to identify unsatisfiable subformulas. Usually, these MaxSAT algorithms perform better when small unsatisfiable subformulas are found in early iterations of the algorithm. However, this is not the case in many problem instances, since the whole formula is given to the SAT solver in each call.

In this paper, we propose to partition the MaxSAT formula using a resolution-based graph representation. Partitions are then iteratively joined by using a proximity measure extracted from the graph representation of the formula. The algorithm ends when only one partition remains and the optimal solution is found. Experimental results show that this new approach further enhances a state of the art MaxSAT solver to optimally solve a larger set of industrial problem instances.

1 Introduction

The improvements of Maximum Satisfiability (MaxSAT) technology in recent years lead to a number of applications of MaxSAT. Many real-world problems in different areas such as fault localization in C programs, design debugging, upgradability of software systems, among others, can now be solved using MaxSAT [2,10,12,15,24]. In the last decade, several new techniques and algorithms have been proposed that improved on previous MaxSAT solvers by several orders of magnitude. Moreover, the developments in the underlying SAT technology, namely identification of unsatisfiable subformulas and incrementality have also been a factor in the improvements of MaxSAT solving.

MaxSAT solvers for industrial instances are usually based on iterative calls to a SAT solver. Moreover, most of these MaxSAT algorithms take advantage of the ability of SAT solvers to identify unsatisfiable subformulas. However, in most cases, algorithms deal with the whole formula at each call of the SAT solver. As a result, unnecessarily large unsatisfiable subformulas can be found at each SAT call, resulting in a slow down of the MaxSAT algorithm. In this work,

we try to avoid this behavior by partitioning the formula and taking advantage of structural information obtained from a formula’s graph representation.

In this paper, we improve on the current state of the art MaxSAT solving by proposing a new unsatisfiability-based algorithm for MaxSAT. The new algorithm integrates several new features, namely: (1) usage of resolution-based graphs to represent the MaxSAT formula, (2) partition of soft clauses in the MaxSAT formula using the referred representation, (3) usage of structural information obtained from the graph representation to drive the merge of partitions and, (4) integration of these features into a new fully incremental algorithm that improves on one of the best non-portfolio solvers from the last MaxSAT Solver Evaluation on several partial MaxSAT industrial benchmark sets.

The paper is organized as follows. Section 2 formally defines MaxSAT and briefly reviews the MaxSAT algorithms more closely related to the proposed approach. In section 3, graph representations of CNF formulas are described. Moreover, the adaptation of resolution-based graphs is proposed. The new MaxSAT algorithm is proposed in section 4. Besides a detailed description, we show how to extract structural information from the graph representations and integrate it in the new algorithm. Section 5 presents the experimental results of the new MaxSAT solver on a large set of industrial benchmark sets used at MaxSAT evaluations. Finally, the paper concludes in section 6.

2 Preliminaries

A propositional formula in Conjunctive Normal Form (CNF), using n Boolean variables x_1, x_2, \dots, x_n , is defined as a conjunction of clauses, where a clause is a disjunction of literals. A literal is either a variable x_i or its complement \bar{x}_i . The Propositional Satisfiability (SAT) problem consists of deciding whether there exists a total assignment to the variables such that the formula is satisfied.

The Maximum Satisfiability (MaxSAT) can be seen as an optimization version of the SAT problem. In MaxSAT, the objective is to find a total assignment to the variables of a CNF formula that minimizes the number of unsatisfied clauses. Notice that minimizing the number of unsatisfied clauses is equivalent to maximizing the number of satisfied clauses.

In a partial MaxSAT formula $\varphi = \varphi_h \cup \varphi_s$, some clauses are considered as hard (φ_h), while others are declared as soft (φ_s). The goal in partial MaxSAT is to find a total assignment to the formula variables such that all hard clauses in φ_h are satisfied, while minimizing the number of unsatisfied soft clauses in φ_s . There are also weighted variants of MaxSAT where soft clauses are associated with weights greater than or equal to 1. In this case, the objective is to satisfy all hard clauses and minimize the total weight of unsatisfied soft clauses. In this paper, we focus solely on partial MaxSAT, but the proposed approach can be generalized to its weighted variants. Furthermore, in all algorithms we assume that the set of hard clauses φ_h is satisfiable. Otherwise, the MaxSAT formula does not have a solution. This can easily be checked through a SAT call on φ_h .

Algorithm 1: Linear Search Unsat-Sat Algorithm

Input: $\varphi = \varphi_h \cup \varphi_s$
Output: satisfying assignment to φ

```
1  $(\varphi_W, V_R, \lambda) \leftarrow (\varphi_h, \emptyset, 0)$ 
2 foreach  $c_i \in \varphi_s$  do
3    $V_R \leftarrow V_R \cup \{r_i\}$  //  $r_i$  is a new relaxation variable
4    $c_R \leftarrow c_i \cup \{r_i\}$ 
5    $\varphi_W \leftarrow \varphi_W \cup \{c_R\}$ 
6 while true do
7    $(st, \nu, \varphi_C) \leftarrow \text{SAT}(\varphi_W \cup \{\text{CNF}(\sum_{r_i \in V_R} r_i \leq \lambda)\})$ 
8   if  $st = \text{SAT}$  then
9     return  $\nu$  // satisfying assignment to  $\varphi$ 
10   $\lambda \leftarrow \lambda + 1$ 
```

The most recent state of the art MaxSAT solvers are based on iterative calls to a SAT solver. One of the most classic approaches is the linear Sat-Unsat algorithm that performs a linear search on the number of unsatisfied clauses. In this case, a new relaxation variable is initially added to each soft clause and the resulting formula is given to a SAT solver. Whenever a solution is found, a new cardinality constraint on the number of relaxation variables is added, such that solutions where a higher or equal number of relaxation variables assigned the value 1 are excluded. The cardinality constraint is encoded into a set of propositional clauses, which are added to the working formula [3,13,17]. The algorithm stops when the SAT call is unsatisfiable. As a result, the last solution found is an optimal solution of the MaxSAT formula.

A converse approach is the linear search Unsat-Sat presented in Algorithm 1. Here, a lower bound λ on the number of unsatisfied soft clauses is maintained between iterations of the algorithm. Initially, λ is assigned value 0. In each iteration, while the working formula given to the SAT solver (line 7) is unsatisfiable, λ is incremented (line 10). Otherwise, an optimal solution to the MaxSAT formula has been found (line 9).

Observe that a SAT solver call on a CNF formula φ_W returns a triple (st, ν, φ_C) , where st denotes the status of the solver: satisfiable (SAT) or unsatisfiable (UNSAT). If φ_W is satisfiable, then ν stores the total assignment found for φ_W . Otherwise, φ_C contains an unsatisfiable subformula that explains a reason for the unsatisfiability of φ_W .

Even though the linear search Unsat-Sat algorithm does not take advantage of current SAT solvers being able to identify unsatisfiable subformulas, there are several more effective algorithms for MaxSAT that use this information to delay the relaxation of soft clauses. An example is the MSU3 algorithm [16] presented in Algorithm 2. Observe that this algorithm also performs an Unsat-Sat linear search, but soft clauses are only relaxed when they appear in an unsatisfiable subformula. The MSU3 algorithm takes as input a MaxSAT formula φ , a set of relaxation variables V_R , and a given lower bound λ . If no structural information

Algorithm 2: MSU3 Algorithm

Input: (φ, V_R, λ)
Output: satisfying assignment to φ

```
1  $\varphi_W \leftarrow \varphi$ 
2 while true do
3    $(st, \nu, \varphi_C) \leftarrow \text{SAT}(\varphi_W \cup \{\text{CNF}(\sum_{r_i \in V_R} r_i \leq \lambda)\})$ 
4   if  $st = \text{SAT}$  then
5     return  $\nu$  // satisfying assignment to  $\varphi$ 
6   foreach  $c_i \in (\varphi_C \cap \varphi_s)$  do
7      $V_R \leftarrow V_R \cup \{r_i\}$  //  $r_i$  is a new variable
8      $c_R \leftarrow c_i \cup \{r_i\}$  //  $c_i$  was not previously relaxed
9      $\varphi_W \leftarrow (\varphi_W \setminus \{c_i\}) \cup \{c_R\}$ 
10   $\lambda \leftarrow \lambda + 1$ 
```

is known about φ , then MSU3 is called with the default values $\varphi = \varphi_h \cup \varphi_s$, $V_R = \emptyset$ and $\lambda = 0$.

Although more sophisticated MaxSAT algorithms exist [20], an implementation of MSU3 algorithm on the Open-WBO framework was one of the best performing non-portfolio algorithms for industrial partial MaxSAT at the MaxSAT Solver Evaluation of 2014¹. One of the crucial features for its success relies on the fact that only one SAT solver instance needs to be created [17]. Therefore, a proper implementation of MSU3 should take advantage of incrementality in SAT solver technology. In this paper, the MSU3 algorithm is further improved with structural information of the problem instance to solve.

3 Graph Representations

In order to extract structural properties of CNF formulas, different graph-based models have been previously proposed. For instance, graph representations have been used to characterize industrial SAT instances [1] and to improve on the performance of MaxSAT algorithms [19]. In this section, we briefly review the Clause-Variable Incidence Graph (CVIG) and adapt the use of Resolution-based Graphs (RES) [26] to model relations in CNF formulas. Although other models exist [1, 25, 19], in the context of our algorithm for MaxSAT solving, these were found to be the best suited.

In the CVIG model, a weighted undirected graph G is built such that a vertex is added for each variable x_j and for each clause c_i occurring in the CNF formula φ . Moreover, for each variable x_j occurring in clause c_i (either as literal x_j or \bar{x}_j), an edge (c_i, x_j) is added to graph G . The edge weight $w(c_i, x_j)$ is defined as:

$$w(c_i, x_j) = \frac{I(x_j)}{|c_i|} \quad (1)$$

¹ Results available at <http://www.maxsat.udl.cat/>

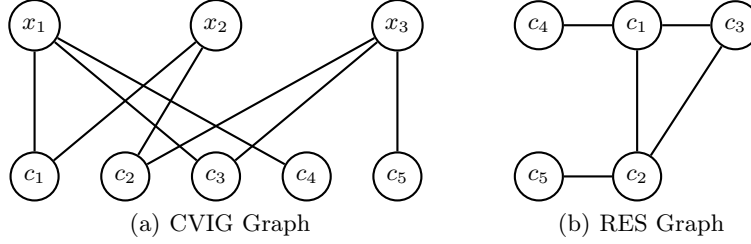


Fig. 1. Example of Graph Models

where $|c_i|$ denotes the number of literals in clause c_i and $I(x_j)$ is defined as the incidence function of x_j in soft clauses as:

$$I(x_j) = 1 + \sum_{x_j \in c \wedge c \in \varphi_s} \frac{1}{|c|} \quad (2)$$

As described in section 2, several MaxSAT solvers rely on the identification of unsatisfiable subformulas. In order to capture sets of clauses more closely related that would result in an unsatisfiable subformula, we propose to adapt Resolution Graphs (RES) to MaxSAT.

In the RES model, we have one vertex in graph G for each clause $c_i \in \varphi$. Let c_i and c_j denote two clauses such that $x_k \in c_i$ and $\bar{x}_k \in c_j$. Moreover, let c_{ij}^{res} be the resulting clause of applying the resolution operation on these clauses. In this case, if c_{ij}^{res} is not a tautology, then an edge (c_i, c_j) is added to G whose weight is defined as:

$$w(c_i, c_j) = \frac{1}{|c_{ij}^{res}|} \quad (3)$$

Notice that in the RES model, clauses are related if the application of the resolution operation results in a non-trivial resolvent. Moreover, observe that the weight of edges between pairs of clauses is greater when the size of the resolvent is smaller. The goal is to make tighter the relations between clauses that produce smaller clauses when resolution is applied.

Consider the following MaxSAT formula where $c_1 : (x_1 \vee x_2)$, $c_2 : (\bar{x}_2 \vee x_3)$ and $c_3 : (\bar{x}_1 \vee \bar{x}_3)$ are hard clauses and $c_4 : (\bar{x}_1)$, $c_5 : (\bar{x}_3)$ are soft clauses. Figures 1(a) and 1(b) illustrate the structure of the graph representation of this formula when using the CVIG and RES models. The weights of edges are not represented for simplicity but can be obtained via Equations (1) and (3). For example, for the CVIG model $w(c_1, x_1) = \frac{2}{2}$ and for the RES model $w(c_2, c_3) = \frac{1}{2}$. Observe that if the clause $c_6 : (\bar{x}_1 \vee \bar{x}_2)$ was added to the formula, it would not connect to any other clause in the RES graph because the only clause containing x_1 positively is $c_1 = (x_1 \vee x_2)$, but that does not connect to c_6 due to x_2 appearing negatively and positively in c_6 and c_1 , respectively. A similar type of analysis is done in *blocked clause elimination* [14,11] — a technique commonly used in formula preprocessing.

Algorithm 3: Extended MSU3 Algorithm

Input: $\varphi = \varphi_h \cup \varphi_s$
Output: satisfying assignment to φ

```
1  $\gamma \leftarrow \langle \gamma_1, \dots, \gamma_n \rangle \leftarrow \text{partitionSoft}(\varphi_s, \varphi_h)$ 
2 foreach  $\gamma_i \in \gamma$  do
3    $(V_R^i, \lambda_i) \leftarrow (\emptyset, 0)$ 
4    $\nu \leftarrow \text{MSU3}(\varphi_h \cup \gamma_i, V_R^i, \lambda_i)$ 
5 if  $|\gamma| = 1$  then
6   return  $\nu$  // no partitions were identified
7 while true do
8    $(\gamma_i, \gamma_j) \leftarrow \text{selectPartitions}(\gamma)$ 
9    $\gamma \leftarrow \gamma \setminus \{\gamma_i, \gamma_j\}$ 
10   $(\gamma_k, V_R^k, \lambda_k) \leftarrow (\gamma_i \cup \gamma_j, V_R^i \cup V_R^j, \lambda_i + \lambda_j)$ 
11   $\nu \leftarrow \text{MSU3}(\varphi_h \cup \gamma_k, V_R^k, \lambda_k)$ 
12  if  $\gamma = \emptyset$  then
13    return  $\nu$ 
14  else
15     $\gamma \leftarrow \gamma \cup \{\gamma_k\}$ 
```

Although resolution-based graphs are not novel [26] and have been used in other domains [25], in this paper we propose to enhance the resolution-based graph representation by adding weights to edges. Moreover, as far as we know, this representation has never been used for MaxSAT solving.

4 New Partition-based Algorithm for MaxSAT

Despite its very good performance in industrial partial MaxSAT instances, the MSU3 algorithm (see Algorithm 2) may suffer from two issues: (1) identification of unnecessarily large unsatisfiable subformulas and, (2) a potentially large cardinality constraint to be maintained between iterations. In fact these issues are related. If an unsatisfiable subformula with an unnecessarily large number of soft clauses is encountered early, then an unnecessarily large cardinality constraint has to be dealt with through most of the algorithm's iterations.

Our approach to tackle these issues is to split the set of soft clauses. The goal is that, at each iteration, the algorithm should only consider part of the problem, instead of dealing with the whole problem instance in each iteration.

4.1 Algorithm Description

Algorithm 3 presents our enhancement of MSU3 with partitioning the soft clause set. The algorithm starts by partitioning φ_s into n disjoint sets of soft clauses $\gamma_1, \gamma_2 \dots \gamma_n$ (line 1). Observe that several methods can be used to partition φ_s . Details of this procedure are discussed later.

For each set γ_i , we apply the MSU3 algorithm to the formula $\varphi_h \cup \gamma_i$ with starting values $V_R^i = \emptyset$ and $\lambda_i = 0$ (lines 2-4). As a result, we obtain a lower bound value λ_i associated with each set of soft clauses γ_i . If the partitioning procedure creates a single partition, then the algorithm terminates (line 6). Otherwise, it is necessary to build the solution of the MaxSAT instance by merging the different sets of soft clauses.

The merge process works as follows. At each iteration, two sets of soft clauses γ_i and γ_j are selected to be merged (line 8) and removed from γ . Let γ_k denote the union of γ_i and γ_j . Since γ_i and γ_j are disjoint, we necessarily have that $\lambda_i + \lambda_j$ is a lower bound for γ_k . Hence, we can safely initialize $\lambda_k = \lambda_i + \lambda_j$ (line 10). Next, the lower bound λ_k is refined by applying the MSU3 algorithm to $\varphi_h \cup \gamma_k$ with starting values $V_R^k = V_R^i \cup V_R^j$ and $\lambda_k = \lambda_i + \lambda_j$ (line 11). When set γ becomes empty, then all soft clauses were merged and the last solution found is an optimal solution (line 13). Otherwise, there are still more sets to be merged and γ_k is added to γ (line 15).

4.2 Partition and Merge of Soft Clauses

Algorithm 3 can be configured differently depending on two procedures: (1) how the set of soft clauses is partitioned (line 1) and (2) how to merge two sets of soft clauses (line 8).

In the partition procedure, our algorithm starts by representing the CNF formula as a graph using one of the models described in section 3. Next, we apply a community-finding algorithm on the graph representation that maximizes a modularity measure [4] in order to obtain a graph partitioning.

Recently, the use of modularity measures has become widespread when analyzing the structure of graphs, in particular for the identification of communities [7,23]. In fact, this has already been used in the analysis of SAT instances [1] and to improve the initial unsatisfiability-based approach proposed by Fu and Malik [6,19]. The purpose of the modularity measure is to evaluate the quality of the partitions, where vertices inside a partition should be densely connected and vertices assigned to different partitions should be loosely connected. However, finding a set of partitions with an optimal modularity value is computationally hard [5]. In our implementation, we use the approximation algorithm proposed by Blondel et al. [4].

At each iteration in Algorithm 3, two partitions are selected to be merged. One can devise several different criteria to select and merge the partitions of soft clauses. In early attempts, the merge process was sequential [19]. Given n partitions $\gamma_1, \gamma_2 \dots \gamma_n$, at iteration i ($i \leq n$) of the algorithm, the first i partitions $\gamma_1, \gamma_2 \dots \gamma_i$ were merged sequentially.

Figure 2(a) illustrates the sequential merging procedure. Observe that the sequential merging process is not balanced. This results in an early growth of the identified subformulas and, as a result, an early growth of the cardinality constraints to be maintained at each iteration of the algorithm.

In this paper, we propose a weighted balanced merge procedure that depends on the strength of the graph connections between partitions. The goal is to delay

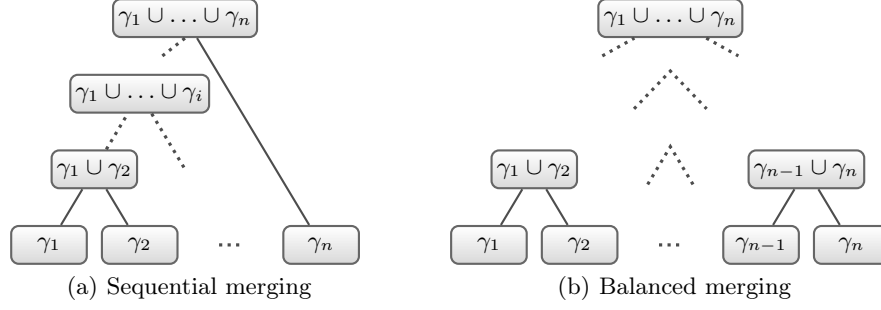


Fig. 2. Examples of merge processes

having to deal with a large number of soft clauses, until the latter iterations of the algorithm. Figure 2(b) illustrates the weighted balanced merging procedure.

Let $G = (V, E)$ denote an undirected weighted graph where V is the set of vertices and E the set of edges. Let $w : E \rightarrow \mathbb{R}$ be a weight function for each edge in the graph. The community-finding algorithm identifies a set of communities $C = \{C_1, C_2, \dots, C_n\}$ where every vertex $u \in V$ is assigned to one and only one community in C . Hence, since in both CVIG and RES model there is a node for each propositional clause, one can build the partitions in a straightforward manner. For each community C_i with vertices representing soft clauses, there is a partition γ_i containing the respective soft clauses.

Based on the graph representation, one can define the strength of the connection between partitions. Let d_{ij} denote the strength between partition γ_i and γ_j . One can define d_{ij} based on the weight between the vertices of their respective communities C_i and C_j in the graph. Hence, d_{ij} can be defined as follows:

$$d_{ij} = \sum_{u \in C_i \wedge v \in C_j} w(u, v) \quad (4)$$

Considering that the graph is undirected, we necessarily have that $d_{ij} = d_{ji}$.

Given an initial set γ of n partitions $\gamma_1, \gamma_2, \dots, \gamma_n$, our algorithm applies a greedy procedure that pairs all partitions γ_i and γ_j from γ to be merged, starting with the pair with largest d_{ij} . After pairing all partitions in the initial set, we perform the same procedure to the next $n/2$ partitions that result from the initial merging iterations. This is iteratively applied until we only have a single partition (see Figure 2(b)).

Observe that if partitions γ_i and γ_j are merged into a new partition γ_k , then the connectivity strength d_{kl} between γ_k to another partition γ_l is given by $d_{kl} = d_{il} + d_{jl}$. This follows from the fact that the communities in the graph are disjoint.

Finally, we would like to reference other solvers that split the set of soft clauses by identifying disjoint unsatisfiable subformulas [8, 21]. However, there are major differences with regard to our proposed approach. First, our solver takes

advantage of an explicit formula representation to split the set of soft clauses, instead of using the unsatisfiable subformulas provided by the SAT solver. Moreover, in our solver, the merge process is also guided by the explicit representation of the formula.

Furthermore, in solvers where disjoint unsatisfiable subformulas are identified [8,21], the split occurs on the cardinality constraints at each iteration. However, each SAT call still has to deal with the whole formula at each iteration. In Algorithm 3, the SAT solver does not have to deal with all soft clauses at each iteration, but only after the final merge step.

4.3 Algorithm Analysis

In this section a proof sketch of the correctness Algorithm 3, as well as an analysis on the number of SAT calls is presented.

Proof (Correctness of Algorithm 3). As mentioned in section 2, we assume the set of hard clauses φ_h is satisfiable. Otherwise, the MaxSAT formula is unsatisfiable. This can be verified by a single SAT call on φ_h before applying Algorithm 3.

For the proof we adopt the following notation. For some set γ_i processed in Algorithm 3, we write $\gamma_i^R \subseteq \varphi_s$ for the set of clauses that were relaxed in the algorithm (but clauses in γ_i^R do not contain the relaxation variables). We will prove by induction the invariant that $\varphi_h \cup \gamma_i^R$ cannot be satisfied unless at least λ_i clauses are removed from γ_i^R . The induction hypothesis is satisfied trivially at the beginning of the algorithm as each λ_i is initialized to 0.

Consider the case where λ_i is augmented by 1 when $\varphi_h \cup \gamma_i \cup \{\sum_{r \in V_R^i} r \leq \lambda_i\}$ is unsatisfiable. Let φ_C be the obtained unsatisfiable subformula from the SAT call, let $\varphi_C^R \subseteq \varphi_s$ be the soft clauses of φ_C that appear as relaxed in γ_i and let $\varphi_C^N = \varphi_s \cap \varphi_C$ be the rest of the soft clauses in the unsatisfiable subformula (not yet relaxed). From induction hypothesis $\varphi_h \cup \varphi_C^R$ cannot be satisfied unless at least λ_i clauses are removed from $\varphi_C^R \subseteq \gamma_i^R$. Since φ_C is an unsatisfiable subformula, it is *impossible* to satisfy $\varphi_h \cup \varphi_C^R \cup \varphi_C^N$ by removing λ_i clauses from φ_C^R . Now we need to also show that it is impossible to satisfy $\varphi_h \cup \gamma_i^R \cup \varphi_C^N$ by removing λ_i clauses from $\gamma_i^R \cup \varphi_C^N$ (this is the new set of relaxed clauses).

Let us assume for contradiction that it is possible to satisfy $\gamma_i^R \cup \varphi_C^N$ by removing some set of clauses ξ s.t. $|\xi| = \lambda_i$. To show the contradiction we consider two cases: (1) $\xi \subseteq \gamma_i$ and (2) $\xi \not\subseteq \gamma_i$. Case (1) yields an immediate contradiction as we would have not obtained unsatisfiability in the SAT call as it would be possible to satisfy $\varphi_h \cup \gamma_i^R$ by removing λ_i clauses from γ_i^R . For case (2) consider that there is a clause $c \in \xi$ s.t. c is not yet relaxed, i.e. $c \notin \gamma_i^R$. This means that $\varphi_h \cup \gamma_i^R$ is satisfiable after removing *less* than λ_i clauses, which is a contradiction with the induction hypothesis.

To show that the invariant is preserved by the merge operation, we observe that any merged γ_i and γ_j are disjoint and therefore so are γ_i^R and γ_j^R . In order to satisfy $\varphi_h \cup (\gamma_i^R \cup \gamma_j^R)$, both $\varphi_h \cup \gamma_i^R$, $\varphi_h \cup \gamma_j^R$ must be satisfied. Consequently, at least $\lambda_i + \lambda_j$ clauses must be removed from $(\gamma_i^R \cup \gamma_j^R)$.

□

Table 1. Experimental evaluation of Open-WBO’s MSU3 algorithm, Eva500a, MSCG and 4 different configurations of the partition-based algorithm.

Instance Group	Total	MSU3	Eva500a	MSCG	S-CVIG	S-RES	W-CVIG	W-RES
aes	7	1	1	1	1	1	1	1
atcoss/mesat	18	11	11	4	11	1	11	11
atcoss/sugar	19	12	11	4	12	3	12	12
bcp/fir	59	59	55	59	56	44	51	51
bcp/hipp-yRa1/simp	17	16	16	16	16	16	16	16
bcp/hipp-yRa1/su	38	35	34	33	34	34	35	33
bcp/msp	64	26	37	29	23	41	27	42
bcp/mtg	40	40	40	40	40	40	40	40
bcp/syn	74	43	48	47	47	48	46	49
circuit-trace-compaction	4	4	4	4	4	3	4	4
close-solutions	50	48	48	46	40	32	40	45
des	50	42	41	41	49	48	50	48
haplotype-assembly	6	5	5	5	5	5	5	5
hs-timetabling	2	1	1	0	1	1	1	1
mbd	46	45	42	43	44	45	45	45
packup-pms	40	40	40	40	40	40	40	40
pbo/mqc/nencdr	84	84	84	84	84	84	84	84
pbo/mqc/nlogencdr	84	84	84	84	84	84	84	84
pbo/routing	15	15	15	15	14	15	15	15
protein_ins	12	12	8	12	12	12	12	12
tpr/Multiple_path	48	48	44	42	48	48	48	48
tpr/One_path	50	50	50	50	50	50	50	50
Total	827	721	719	699	715	695	717	736

Finally, we note that the number of SAT calls performed by Algorithm 3 is larger than the MSU3 algorithm. Observe that the number of unsatisfiable SAT calls is the same for both algorithms. Let λ be the number of unsatisfiable soft clauses at any optimal solution of the MaxSAT instance. In this case, both algorithms perform λ unsatisfiable SAT calls. However, while MSU3 performs only one satisfiable SAT call, Algorithm 3 performs $2n - 1$, where n is the number of identified partitions (line 1).

5 Experimental Results

In this section we compare different configurations of Algorithm 3 with the top 3 non-portfolio solvers of the MaxSAT 2014 Evaluation’s industrial partial MaxSAT category. The top 3 were Open-WBO’s MSU3 incremental algorithm [18,17], Eva500a [22] and MSCG [9]. The new partition-based algorithm is also implemented using the Open-WBO framework².

The algorithms were evaluated running on the set union of the partial MaxSAT industrial instances of the MaxSAT evaluations of 2012, 2013 and 2014. For each instance, algorithms were executed with a timeout of 1800 seconds and a memory limit of 4 GB. Similar resource limitations were used during the last MaxSAT Evaluation of 2014. These tests were conducted on a machine with 4 AMD Opteron 6376 (2.3 GHz) and 128 GB of RAM, running Debian jessie.

Table 1 presents the number of instances solved by each algorithm, per instance set. Besides MSU3, Eva500a and MSCG, results for the best 4 config-

² Available at <http://sat.inesc-id.pt/open-wbo/>

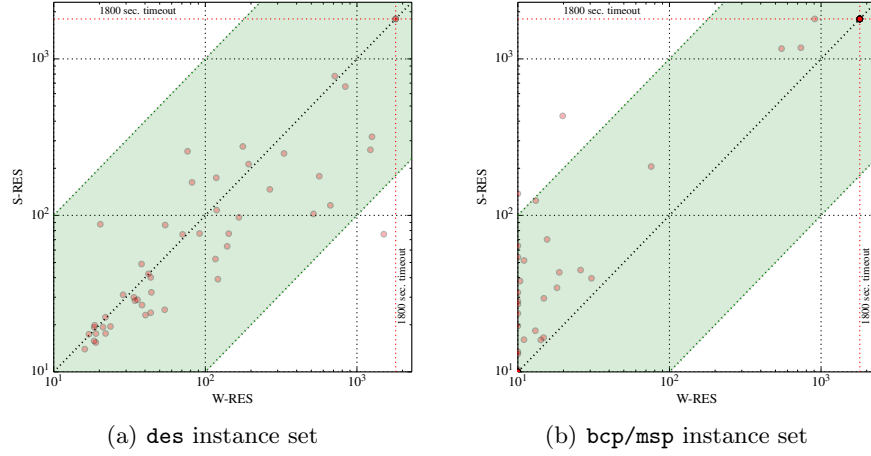


Fig. 3. Comparison between run times of S-RES and W-RES on **des** and **bcp/msp** instance sets

urations of the partition-based enhanced MSU3 algorithm are shown. S-CVIG applies the sequential merging of partitions using the CVIG graph model. S-RES also applies sequential merging, but using the RES graph model. W-CVIG and W-RES apply the weighted balanced merging of partitions, using the CVIG and RES graph models, respectively. Note that all our implementations are fully incremental, i.e. only one instance of the SAT solver is created throughout the execution of the proposed algorithm. As with the MSU3 implementation on Open-WBO, we take advantage of assumptions usage at each SAT call and incremental encoding of cardinality constraints [17].

Results from Table 1 show that all variants of the partition-based algorithm are competitive with the remaining state of the art algorithms. However, overall results clearly show that W-RES outperforms all remaining algorithms, since it is able to solve more instances in total. Moreover, results for the configurations of partition-based algorithm also show that weight-based balanced merging of partitions is preferable to sequential partitioning.

Considering that MSU3 is our base solver, most gains occur in instance sets **bcp/msp**, **bcp/syn** and **des**. While in the **bcp/syn** and **des** instance sets, all partition-based configurations perform better, in **bcp/msp** the resolution-based graph partitioning allowed a significant performance boost.

Figures 3(a) and 3(b) compare the results of S-RES and W-RES on the **des** and **bcp/msp** instance sets. In the **des** instances, the run time of sequential merging is slightly better, despite solving the same number of instances. Nevertheless, in the **bcp/msp** instance set the weight-based balanced merging used in W-RES clearly outperforms the sequential merging approach used in S-RES.

In Figures 4(a) and 4(b) we compare MSU3 and W-RES on the same benchmark sets. It can be observed that W-RES performs much better in these in-

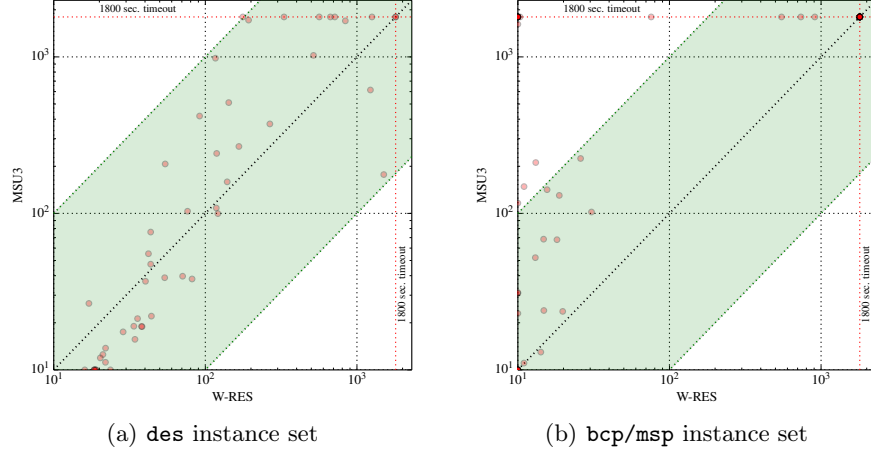


Fig. 4. Comparison between run times of MSU3 and W-RES on **des** and **bcp/msp** instance sets

stances. In the **des** instance set, there are some instances where W-RES is not as fast, since there is some time spent in finding partitions and additional SAT calls. We note that there is always some time spent in building the graph, applying the community finding algorithm and splitting the set of soft clauses. However, this partitioning step is usually not very time consuming. Nevertheless, W-RES is able to scale better and solve more instances. In the **bcp/msp** instances, the proposed techniques allow W-RES to be much better than MSU3, as well as all other algorithms tested.

Resolution-based graph models performed worst in the **bcp/fir** category. It was observed that the modularity values obtained for the resolution-based graphs were low in this particular instance set. As a result, the partitioning obtained for S-RES and W-RES in **bcp/fir** instances is not as meaningful as for other instance sets. When this occurs, it can deteriorate the algorithm’s performance, since the partition-based algorithm performs more SAT calls than MSU3.

When considering all benchmark sets, W-CVIG and W-RES solve different instances and the Virtual Best Solver³ (VBS) between them solves 747 instances (11 more than W-RES). Furthermore, there are a few instances which are only solved by MSU3 but not by W-CVIG nor W-RES. The VBS between MSU3, W-CVIG, W-RES can solve 752 (5 more than the VBS between W-CVIG and W-RES). Even though W-RES outperforms the remaining algorithms, this suggests that dynamically choosing the partition type could further improve the performance of the solver.

Finally, Figure 5 shows a cactus plot with the run times of all algorithms considered in the experimental evaluation. Here we can observe that S-RES is

³ The Virtual Best Solver between a set of solvers shows the total number of instances that can be solved by at least one of those solvers.

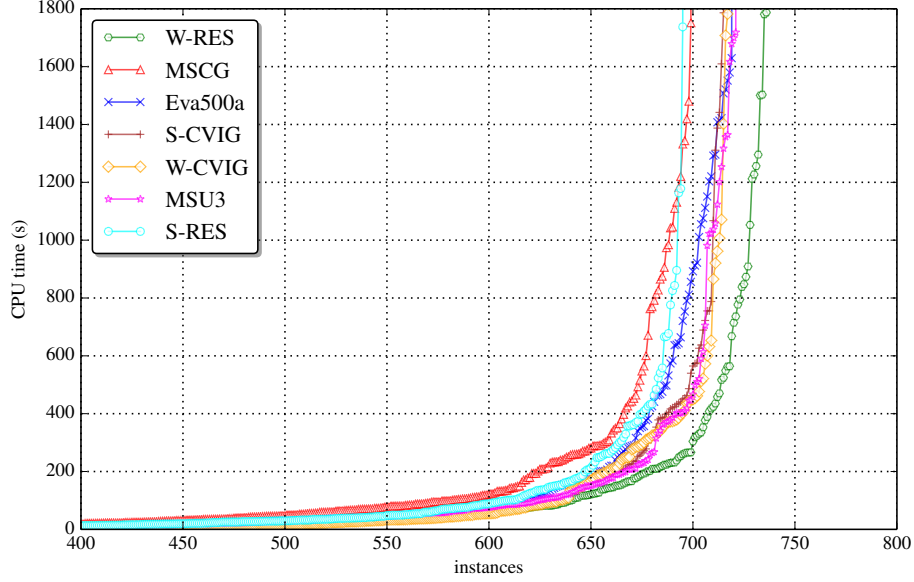


Fig. 5. Cactus plot with the run times for MSU3, Eva500a, MSCG, S-CVIG, S-RES, W-CVIG and W-RES.

much slower than W-RES, clearly showing the effectiveness of the newly proposed weight-based merging. Overall, W-RES clearly outperforms the remaining algorithms, being able to solve 700 instances in 300 seconds or less.

6 Conclusions and Future Work

In this paper we exploit resolution-based graph representations of CNF formulas in order to develop a new state of the art algorithm for MaxSAT. In the proposed approach, soft clauses are initially partitioned in disjoint sets by analyzing the formula structure. The partitioning process is attained by applying a community-finding algorithm on weighted resolution-based graphs. Next, at each iteration of the algorithm, partitions are merged using structural information from the graph representation until an optimal solution is found.

The proposed approach is novel in many aspects. First, the use of a resolution-based graph representation allows to better model the interaction between clauses. Furthermore, instead of applying a sequential merging process, the graph representation is also used in a weight-based balanced merging procedure. Moreover, since the algorithm does not have to deal with the whole formula at each iteration, smaller unsatisfiable cores are identified. As a result from this process, smaller cardinality constraints are encoded into CNF at each iteration, thus improving the algorithm’s performance.

Experimental results obtained in industrial partial MaxSAT instances clearly show the effectiveness of the proposed algorithm. As a result, our solver improves upon one of the best non-portfolio solvers on the industrial partial category from the 2014 MaxSAT solver evaluation.

The source code of the new solver will become available as part of the Open-WBO framework. This will allow the research community to build upon the current work to further improve MaxSAT solving.

As future work, we propose to extend the proposed approach for weighted MaxSAT solving. Moreover, different model representations of CNF formulas are to be tested, as well as new techniques for building and merging partitions of soft clauses in MaxSAT formulas. Furthermore, the proposed techniques are not exclusive to MSU3 and can also be integrated into other MaxSAT algorithms. Additionally, these techniques can also be applied to other extensions of SAT.

Acknowledgments

This work was partially supported by DARPA MUSE award #FA8750-14-2-0270, FCT grant POLARIS (PTDC/EIA-CCO/123051/2010), FCT grant AMOS (CMUP-EPB/TIC/0049/2013), and national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013. The views, opinions, and/or findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

References

1. Ansótegui, C., Giráldez-Cru, J., Levy, J.: The Community Structure of SAT Formulas. In: International Conference on Theory and Applications of Satisfiability Testing. LNCS, vol. 7317, pp. 410–423. Springer (2012)
2. Asín, R., Nieuwenhuis, R.: Curriculum-based course timetabling with SAT and MaxSAT. *Annals of Operations Research* 218(1), 71–91 (2014)
3. Bailleux, O., Boufkhad, Y.: Efficient CNF Encoding of Boolean Cardinality Constraints. In: Principles and Practice of Constraint Programming. LNCS, vol. 2833, pp. 108–122. Springer (2003)
4. Blondel, V., Guillaume, J., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. *Journal of Statistical Mechanics* 2008(10), P10008 (2008)
5. Brandes, U., Delling, D., Gaertler, M., Goerke, R., Hoefer, M., Nikoloski, Z., Wagner, D.: Maximizing modularity is hard. arXiv: physics, 0608255 (2006)
6. Fu, Z., Malik, S.: On Solving the Partial MAX-SAT Problem. In: International Conference on Theory and Applications of Satisfiability Testing. LNCS, vol. 4121, pp. 252–265. Springer (2006)
7. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99(12), 7821–7826 (2002)
8. Heras, F., Morgado, A., Marques-Silva, J.: Core-Guided Binary Search Algorithms for Maximum Satisfiability. In: AAAI Conference on Artificial Intelligence. AAAI Press (2011)

9. Ignatiev, A., Morgado, A., Manquinho, V., Lynce, I., Marques-Silva, J.: Progression in Maximum Satisfiability. In: European Conference on Artificial Intelligence. pp. 453–458. IOS Press (2014)
10. Janota, M., Lynce, I., Manquinho, V., Marques-Silva, J.: PackUp: Tools for Package Upgradability Solving. *Journal on Satisfiability, Boolean Modeling and Computation* 8(1/2), 89–94 (2012)
11. Järvisalo, M., Biere, A., Heule, M.: Blocked clause elimination. In: Tools and Algorithms for the Construction and Analysis of Systems. LNCS, vol. 6015, pp. 129–144. Springer (2010)
12. Jose, M., Majumdar, R.: Cause clue clauses: error localization using maximum satisfiability. In: Programming Language Design and Implementation. pp. 437–446. ACM (2011)
13. Koshimura, M., Zhang, T., Fujita, H., Hasegawa, R.: QMaxSAT: A Partial MaxSAT Solver. *Journal on Satisfiability, Boolean Modeling and Computation* 8(1/2), 95–100 (2012)
14. Kullmann, O.: On a generalization of extended resolution. *Discrete Applied Mathematics* 96–97, 149–176 (1999)
15. Le Berre, D., Rapicault, P.: Dependency management for the eclipse ecosystem: An update. In: International Workshop on Logic and Search
16. Marques-Silva, J., Planes, J.: On Using Unsatisfiability for Solving Maximum Satisfiability. *CoRR* (2007)
17. Martins, R., Joshi, S., Manquinho, V., Lynce, I.: Incremental Cardinality Constraints for MaxSAT. In: Principles and Practice of Constraint Programming. LNCS, vol. 8656, pp. 531–548. Springer (2014)
18. Martins, R., Manquinho, V., Lynce, I.: Open-WBO: a Modular MaxSAT Solver. In: International Conference on Theory and Applications of Satisfiability Testing. LNCS, vol. 8561, pp. 438–445. Springer (2014)
19. Martins, R., Manquinho, V.M., Lynce, I.: Community-based partitioning for maxsat solving. In: International Conference on Theory and Applications of Satisfiability Testing. LNCS, vol. 7962, pp. 182–191. Springer (2013)
20. Morgado, A., Heras, F., Liffiton, M., Planes, J., Marques-Silva, J.: Iterative and core-guided MaxSAT solving: A survey and assessment. *Constraints* 18(4), 478–534 (2013)
21. Morgado, A., Heras, F., Marques-Silva, J.: Improvements to Core-Guided Binary Search for MaxSAT. In: International Conference on Theory and Applications of Satisfiability Testing. LNCS, vol. 7317, pp. 284–297. Springer (2012)
22. Narodytska, N., Bacchus, F.: Maximum Satisfiability Using Core-Guided MaxSAT Resolution. In: AAAI Conference on Artificial Intelligence. pp. 2717–2723. AAAI Press (2014)
23. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. *Physical Review E* 69(026113) (2004)
24. Safarpour, S., Mangassarian, H., Veneris, A.G., Liffiton, M.H., Sakallah, K.A.: Improved design debugging using maximum satisfiability. In: Formal Methods in Computer-Aided Design. pp. 13–19. IEEE Computer Society (2007)
25. Van Gelder, A.: Variable independence and resolution paths for quantified boolean formulas. In: Principles and Practice of Constraint Programming. LNCS, vol. 6876, pp. 789–803. Springer (2011)
26. Yates, R.A., Raphael, B., Hart, T.P.: Resolution graphs. *Artificial Intelligence* 1(4), 257–289 (1970)