Symmetry Breaking for Maximum Satisfiability*

Joao Marques-Silva[1], Inês Lynce[2], and Vasco Manquinho[2]

[1] School of Electronics and Computer Science, University of Southampton, UK
jpms@ecs.soton.ac.uk
[2] IST/INESC-ID, Technical University of Lisbon, Portugal
{ines,vmm}@sat.inesc-id.pt

Abstract. Symmetries are intrinsic to many combinatorial problems including Boolean Satisfiability (SAT) and Constraint Programming (CP). In SAT, the identification of symmetry breaking predicates (SBPs) is a well-known, often effective, technique for solving hard problems. The identification of SBPs in SAT has been the subject of significant improvements in recent years, resulting in more compact SBPs and more effective algorithms. The identification of SBPs has also been applied to pseudo-Boolean (PB) constraints, showing that symmetry breaking can also be an effective technique for PB constraints. This paper extends further the application of SBPs, and shows that SBPs can be identified and used in Maximum Satisfiability (MaxSAT), as well as in its most well-known variants, including partial MaxSAT, weighted MaxSAT and weighted partial MaxSAT. As with SAT and PB, symmetry breaking predicates for MaxSAT and variants are shown to be effective for a representative number of problem domains, allowing solving problem instances that current state of the art MaxSAT solvers could not otherwise solve.

1 Introduction

Symmetry breaking is a widely used technique for solving combinatorial problems. Symmetries have been extensively studied in Boolean Satisfiability (SAT) [15, 4, 7, 1], and are regarded as an essential technique for solving specific classes of problem instances. Symmetries have also been widely used for solving constraint satisfaction problems (CSPs) [11]. More recent work has shown how to apply symmetry breaking in pseudo-Boolean (PB) constraints [2] and also in soft constraints [24]. It should be noted that symmetry breaking is viewed as an effective problem solving technique, either for SAT, PB or CP, that is often used as an optional technique, to be used when default algorithms are unable to solve a given problem instance.

In recent years there has been a growing interest in algorithms for MaxSAT and variants [16, 17, 26, 13, 14, 18, 21, 20], in part because of the wide range of potential applications. MaxSAT and variants represent a more general framework than either SAT or PB, and so can naturally be used in many practical applications. The interest in MaxSAT and variants motivated the development of a new generation of MaxSAT algorithms, remarkably more efficient than early MaxSAT algorithms [25, 5]. Despite the observed improvements, there are many problems still too complex for MaxSAT

* This paper extends a preliminary technical report [19] on the same subject.

algorithms to solve [3]. Natural lines of research for improving MaxSAT algorithms include studying techniques known to be effective for either SAT, PB or CP. One concrete example is symmetry breaking. Despite its success in SAT, PB and CP, the usefulness of symmetry breaking for MaxSAT and variants has not been thoroughly investigated.

This paper addresses the problem of using symmetry breaking in MaxSAT and in its most well-known variants, partial MaxSAT, weighted MaxSAT and weighted partial MaxSAT. The work extends past recent work on computing symmetries for SAT [1] and PB constraints [2] by computing automorphisms on colored graphs obtained from CNF or PB formulas, and by showing how symmetry breaking predicates [7, 1] can be exploited. The experimental results show that symmetry breaking is an effective technique for MaxSAT and variants, allowing solving problem instances that state of the art MaxSAT solvers could not otherwise solve.

The paper is organized as follows. The next section introduces the notation used throughout the paper, provides a brief overview of MaxSAT and variants, and also summarizes the work on symmetry breaking for SAT and PB constraints. Afterwards, the paper describes how to apply symmetry breaking in MaxSAT and variants. Experimental results, obtained on representative problem instances from the MaxSAT evaluation [3] and also from practical applications [1], demonstrate that symmetry breaking allows solving problem instances that could not be solved by *any* of the available state of the art MaxSAT solvers. The paper concludes by summarizing related work, by overviewing the main contributions, and by outlining directions for future work.

2 Preliminaries

This section introduces the notation used through the paper. Moreover, this section summarizes relevant results in symmetry identification and symmetry breaking, and develops extensions to existing results, which will serve for applying symmetry breaking in MaxSAT. Finally, this section also summarizes the MaxSAT problem and its variants.

2.1 Propositional Satisfiability

The usual definitions of propositional logic are assumed. Let $X = \{x_1, x_2, \ldots, x_n\}$ denote a set of propositional variables. A propositional formula φ in conjunctive normal form (CNF) is a conjunction of clauses. A clause ω is a disjunctions of literals. A literal is either a variable ($x \in X$) or its complement ($\bar{x}$, with $x \in X$). Where appropriate, clauses are viewed as sets of literals, defined on X, and CNF formulas are viewed as set of clauses.

A truth assignment is a function $\mathcal{A} : X \rightarrow \{0, 1\}$. The usual semantics of propositional logic is used for associating values with formulas given truth assignments to the variables. Assignments serve for computing the values of literals, clauses and the complete CNF formula, respectively, $\mathcal{A}(l)$, $\mathcal{A}(\omega)$ and $\mathcal{A}(\varphi)$ [3]. As a result, the truth value of

[3] The use of $\mathcal{A}$ for describing the truth value of clauses and CNF formulas is an often used abuse of notation.

literals, clauses and CNF formulas can be defined as follows:

$$\mathcal{A}(l) = \begin{cases} \mathcal{A}(x_i) & \text{if } l = x_i \\ 1 - \mathcal{A}(x_i) & \text{if } l = \bar{x}_i \end{cases} \tag{1}$$

$$\mathcal{A}(\omega) = \max\{\mathcal{A}(l) \,|\, l \in \omega\} \tag{2}$$

$$\mathcal{A}(\varphi) = \min\{\mathcal{A}(\omega) \,|\, \omega \in \varphi\} \tag{3}$$

A clause is said to be *satisfied* if at least one of its literals assumes value 1. If all literals of a clause assume value 0, then the clause is *unsatisfied*. A formula is satisfied if all clauses are satisfied, otherwise it is unsatisfied. A truth assignment that satisfies φ is referred to as *model*. The set of models of φ is denoted by $\mathcal{M}(\varphi)$. The propositional satisfiability (SAT) problem consists in deciding whether there exists an assignment to the variables such that φ is satisfied.

2.2 Symmetries

A symmetry is an operation that preserves the constraints, and therefore also preserves the solutions of a problem instance [6]. For a set of symmetric objects, it is possible to obtain the whole set of objects from any of the objects. The elimination of symmetries has been extensively studied in CP and SAT [15, 4, 22, 7]. With the goal of developing a solution for breaking symmetries in MaxSAT, this section provides a few necessary definitions related with symmetries in propositional formulas [7].

For a set X of variables, a *permutation* of X is a bijective function $\pi : X \to X$, and the image of x under π is denoted x^π. The set of all permutations of X is denoted by $\mathcal{P}_X$, and this set is a group under the composition operation. Permutations can be extended to literals, clauses and formulas, by replacing each literal by its permuted literal. As a result, $\varphi^\pi = \wedge_i \omega_i^\pi$, and $\omega_i^\pi = \vee_j l_i^\pi$. Moreover, $l_j^\pi = x_j^\pi$ if $l_j = x_j$, and $l_j^\pi = \overline{x_j^\pi}$ if $l_j = \bar{x}_j$.

Permutations also map truth assignments to truth assignments. If $\pi \in \mathcal{P}_X$, then each truth assignment $\mathcal{A}$ is mapped into a new truth assignment ${}^\pi\mathcal{A}$, where ${}^\pi\mathcal{A}(x) = \mathcal{A}(x^\pi)$.

Given a formula φ and $\pi \in \mathcal{P}_X$, π is a *symmetry* (or *automorphism*) iff $\varphi^\pi = \varphi$. Moreover, $\mathcal{S}_\varphi$ represents the set of symmetries of φ. A well-known result in symmetry breaking for SAT is the following [7]:

Proposition 1 (Proposition 2.1 in [7]). *Let φ be a CNF formula over X, $\pi \in \mathcal{S}_\varphi$, and $\mathcal{A}$ a truth assignment of X. Then $\mathcal{A} \in \mathcal{M}(\varphi)$ iff ${}^\pi\mathcal{A} \in \mathcal{M}(\varphi)$.*

Proposition 1 can be extended to account for the number of unsatisfied clauses given an assignment. Essentially, the number of unsatisfied clauses remains *unchanged* in the presence of permutations. A permutation maps each clause to another clause. For each assignment, each unsatisfied clause is mapped to another clause which is also unsatisfied.

Let $\mu(\varphi, \mathcal{A})$ denote the number of unsatisfied clauses of formula φ given assignment $\mathcal{A}$. Clearly $\mathcal{M}(\varphi) = \{\mathcal{A} \,|\, \mu(\varphi, \mathcal{A}) = 0\}$. Then the following holds:

Proposition 2. *Let φ be a CNF formula over X, $\pi \in \mathcal{S}_\varphi$, and $\mathcal{A}$ a truth assignment of X. Then $\mu(\varphi, \mathcal{A}) = \mu(\varphi^\pi, {}^\pi\mathcal{A})$.*

Proof: *The proof follows from the discussion above. Symmetries map clauses into clauses. Unsatisfied clauses will be mapped into unsatisfied clauses, and the mapping is one-to-one.* ■

Proposition 2 is used in the following sections for validating the correctness of symmetry breaking for MaxSAT and extensions.

It is also known that $\mathcal{S}_\varphi$ induces an equivalence relation on the truth assignments of X. Moreover, observe that for each equivalence class, the number of unsatisfied clauses is the same. Symmetry breaking predicates (SBPs) are used for selecting a reduced set of representatives from each equivalence class (ideally one representative from each equivalence class).

2.3 Symmetry Breaking

Given the definition of symmetries, symmetry breaking predicates target the elimination of all but one of the equivalent objects [7, 1]. Symmetry breaking is expected to speed up the search as the search space gets reduced. For specific problems where symmetries may be easily found this reduction may be significant. Nonetheless, the elimination of symmetries necessarily introduces overhead that is expected to be negligible when compared with the benefits it may provide.

The most well-known strategy for eliminating symmetries in SAT consists in adding symmetry breaking predicates (SBPs) to the CNF formula [7]. SBPs are added to the formula before the search starts. The symmetries may be identified for each specific problem, and in that case it is required that the symmetries in the problem are identified when creating the encoding. Alternatively, one may give a formula to a specialized tool for detecting all the symmetries [1]. The resulting SBPs select one representative from each equivalence class. In case all symmetries are broken, only one assignment, instead of n assignments, may satisfy a set of constraints, n being the number of elements in a given equivalence class. The most often used approach for constructing SBPs consists in selecting the least assignment in each equivalence class, e.g. by implementing predicates that compare pairs of truth assignments. Other approaches include remodeling the problem [23] and breaking symmetries during search [12]. Remodeling the problem implies creating a different encoding, e.g. obtained by defining a different set of variables, in order to create a problem with less symmetries or even none at all. Alternatively, the search procedure may be adapted for adding SBPs as the search proceeds to ensure that any assignment symmetric to one assignment already considered will not be explored in the future, or by performing checks that symmetric equivalent assignments have not yet been visited.

Currently available tools for detecting and breaking symmetries for a given formula are based on group theory. From each formula a group is extracted, where a group is a set of permutations. A permutation is a one-to-one correspondence between a set and itself. Each symmetry defines a permutation on a set of literals. In practice, each permutation is represented by a product of disjoint cycles. Each cycle $(l_1\, l_2\, \ldots\, l_m)$ with size m stands for the permutation that maps l_i on l_{i+1} (with $1 \leq i \leq m-1$) and l_m on l_1. Applying a permutation to a formula will produce exactly the same formula.

Example 1. Consider the following CNF formula:

$$\varphi = (x_1 \vee x_2) \wedge (\bar{x}_1 \vee x_2) \wedge (\bar{x}_2) \wedge (x_3 \vee x_2) \wedge (\bar{x}_3 \vee x_2)$$

The permutations identified for φ are $(x_3\ \bar{x}_3)$ and $(x_1\ x_3)(\bar{x}_1\ \bar{x}_3)$. (The permutation $(x_1\ \bar{x}_1)$ is implicit.) The formula resulting from the permutation $(x_3\ \bar{x}_3)$ is obtained by replacing every occurrence of x_3 by $\bar{x}_3$ and every occurrence of $\bar{x}_3$ by x_3. Clearly, the obtained formula is equal to the original formula. The same happens when applying the permutation $(x_1\ x_3)(\bar{x}_1\ \bar{x}_3)$: replacing x_1 by x_3, x_3 by x_1, $\bar{x}_1$ by $\bar{x}_3$ and $\bar{x}_3$ by $\bar{x}_1$ produces the same formula. From [7, 1], selection of the least assignment in each permutation yields the symmetry breaking predicate $\varphi_{sbp} = (\bar{x}_3) \wedge (\bar{x}_1 \vee x_3)$.

2.4 Maximum Satisfiability

Given a propositional formula φ, the MaxSAT problem is defined as finding an assignment to variables in φ such that the number of satisfied clauses is maximized. (MaxSAT can also be defined as finding an assignment that minimizes the number of unsatisfied clauses.) Well-known variants of MaxSAT include partial MaxSAT, weighted MaxSAT and weighted partial MaxSAT.

For partial MaxSAT, a propositional formula φ is described by the conjunction of two CNF formulas φ_s and φ_h, where φ_s represents the *soft* clauses and φ_h represents the *hard* clauses. The partial MaxSAT problem over a propositional formula $\varphi = \varphi_h \wedge \varphi_s$ consists in finding an assignment to the problem variables such that all hard clauses (φ_h) are satisfied and the number of satisfied soft clauses (φ_s) is maximized.

For *weighted* MaxSAT, each clause in the CNF formula is associated to a non-negative weight. A weighted clause is a pair (ω, c) where ω is a classical clause and c is a natural number corresponding to the cost of unsatisfying ω. Given a weighted CNF formula φ, the *weighted* MaxSAT problem consists in finding an assignment to problem variables such that the total weight of the unsatisfied clauses is minimized, which implies that the total weight of the satisfied clauses is maximized.

For the *weighted partial* MaxSAT problem, the formula is the conjunction of a weighted CNF formula (soft clauses) and a classical CNF formula (hard clauses). The weighted partial MaxSAT problem consists in finding an assignment to the variables such that all hard clauses are satisfied and the total weight of satisfied soft clauses is maximized. Observe that, for both partial MaxSAT and weighted partial MaxSAT, hard clauses can also be represented as weighted clauses. For hard clauses one can consider that the weight is greater than the sum of the weights of the soft clauses. This allows a more uniform treatment of hard and weighted soft clauses.

MaxSAT and variants find a wide range of practical applications, that include scheduling, routing, bioinformatics, and design automation. Moreover, MaxSAT can be used for solving pseudo-Boolean optimization [14]. The practical applications of MaxSAT motivated recent interest in developing more efficient algorithms. The most efficient algorithms for MaxSAT and variants are based on branch and bound search, using dedicated bounding and inference techniques [16, 17, 13, 14]. Lower bounding techniques include, for example, the use of unit propagation for identifying necessarily unsatisfied clauses, whereas inference techniques can be viewed as restricted forms of resolution, with the objective of simplifying the problem instance to solve.

3 Symmetry Breaking for MaxSAT

This section describes how to use symmetry breaking in MaxSAT. First, the construction process for the graph representing a CNF formula is briefly reviewed [7, 1], as it will be modified later in this section. Afterwards, plain MaxSAT is considered. The next step is to address symmetry breaking for partial, weighted and weighted partial MaxSAT.

3.1 From CNF Formulas to Colored Graphs

Symmetry breaking for MaxSAT and variants requires a few modifications to the approach used for SAT [7, 1]. This section summarizes the basic approach, which is then extended in the following sections.

Given a graph, the *graph automorphism* problem consists in finding isomorphic groups of edges and vertices with a one-to-one correspondence. In case of graphs with colored vertices, the correspondence is made between vertices with the same color. It is well-known that symmetries in SAT can be identified by reduction to a graph automorphism problem [7, 1]. The propositional formula is represented as an undirected graph with colored vertices, such that the automorphism in the graph corresponds to a symmetry in the propositional formula.

Given a propositional formula φ, a colored undirected graph is created as follows:

- For each variable $x_j \in X$ add two vertices to represent x_j and $\bar{x}_j$. All vertices are associated with variables are colored with color 1;
- For each variable $x_j \in X$ add an edge between the vertices representing x_j and $\bar{x}_j$;
- For each binary clause $\omega_i = (l_j \vee l_k) \in \varphi$, add an edge between the vertices representing l_j and l_k;
- For each non-binary clause $\omega_i \in \varphi$ create a vertex colored with color 2;
- For each literal l_j in a non-binary clause ω_i, add an edge between the vertices representing the literal and the clause.

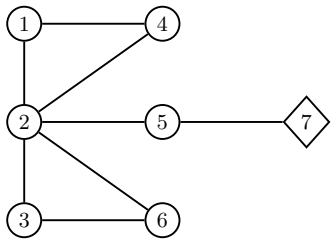
Example 2. Figure 1 shows the colored undirected graph associated with the CNF formula of Example 1. Vertices with shape $\circ$ represent color 1 and vertices with shape $\diamond$ represent color 2. Vertex 1 corresponds to x_1, 2 to x_2, 3 to x_3, 4 to $\bar{x}_1$, 5 to $\bar{x}_2$, 6 to $\bar{x}_3$ and 7 to unit clause $(\bar{x}_2)$. Edges 1-2, 2-3, 2-4 and 2-6 represent binary clauses and edges 1-4, 2-5 and 3-6 link complemented literals. Finally, edge 5-7 associates the correct literal with the unit clause.

Observe that for binary clauses it suffices to connect the vertices of the literals associated with the clause [1].

3.2 Plain Maximum Satisfiability

Let φ represent the CNF formula of a MaxSAT instance. Moreover, let φ_{sbp} be the CNF formula for the symmetry-breaking predicate obtained with a CNF symmetry tool (e.g. Shatter [4] [1] built on top of Saucy [8]). All clauses in φ are effectively *soft* clauses,

[4] Available from http://www.eecs.umich.edu/∼faloul/Tools/shatter/.



(a) Colored graph

Vertex	Mapping
1	literal x_1
2	literal x_2
3	literal x_3
4	literal $\bar{x}_1$
5	literal $\bar{x}_2$
6	literal $\bar{x}_3$
7	clause $(\bar{x}_2)$

(b) Vertex mapping

Fig. 1. Colored graph for Example 2

for which the objective is to maximize the number of satisfied clauses. In contrast, the clauses in φ_{sbp} are *hard* clauses, which must necessarily be satisfied. As a result, the original MaxSAT problem is transformed into a partial MaxSAT problem, where φ denotes the soft clauses and φ_{sbp} denotes the hard clauses. The solution of the partial MaxSAT problem corresponds to the solution of the original MaxSAT problem.

Example 3. As shown earlier, for the CNF formula of Example 1, the generated SBP (e.g. by Shatter) is: $\varphi_{sbp} = (\bar{x}_3) \land (\bar{x}_1 \lor x_3)$. As a result, the resulting instance of partial MaxSAT will be $\varphi' = (\varphi_h \land \varphi_s) = (\varphi_{sbp} \land \varphi)$. The addition of the clauses associated with the SBP imply $x_3 = 0$ and $x_1 = 0$. Observe that if there exists a MaxSAT solution for φ with $x_3 = 1$ or $x_1 = 1$, then not only it cannot have a smaller number of unsatisfied clauses than φ', but also such a solution must be included in an equivalent class for which there is at least one representative in the solutions of φ'.

As the previous example suggests, the hard clauses represented by φ_{sbp} do not change the solution of the original MaxSAT problem. Indeed, the construction of the symmetry breaking predicate guarantees that the maximum number of satisfied soft clauses remains unchanged by the addition of the hard clauses.

Proposition 3. *The maximum number of satisfied clauses for the MaxSAT problem φ and the partial MaxSAT problem $(\varphi \land \varphi_{sbp})$ are the same.*

Proof: *From Proposition 2 it is known that symmetries maintain the number of unsatisfied clauses, and this also holds for the equivalence classes induced by symmetries. Moreover, symmetry breaking predicates allow for at least one truth assignment from each equivalence class. Hence, at least one truth assignment from the equivalence class that maximizes the number of satisfied clauses will satisfy the symmetry breaking predicate, and so the solution of the MaxSAT problem is preserved.* ■

3.3 Partial and Weighted Maximum Satisfiability

For partial MaxSAT, the generation of SBPs needs to be modified. The graph representation of the CNF formula must take into account the existence of hard and soft

clauses, which must be distinguished by a graph automorphism algorithm. Symmetric objects for problem instances with hard and soft clauses establish a correspondence either between hard clauses or between soft clauses. In other words, when applying a permutation hard clauses can only be replaced by other hard clauses, and soft clauses by other soft clauses. In order to address this issue, the colored graph generation needs to be modified. In contrast to the MaxSAT case, binary clauses are *not* handled differently from other clauses, and must be represented as vertices in the colored graph. This is necessary for distinguishing between hard and soft binary clauses, and in general between binary clauses with different weights.

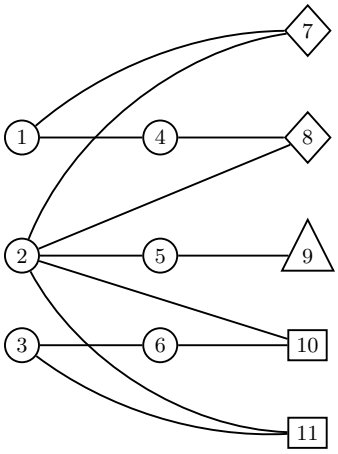
For the partial MaxSAT problem, clauses can now have one of two colors. A vertex with color 2 is associated with each soft clause, and a vertex with color 3 is associated with each hard clause. (As before, a vertex with color 1 corresponds to a literal.) This modification ensures that any identified automorphism guarantees that soft clauses correspond only to soft clauses, and hard clauses correspond only to hard clauses. Moreover, the procedure for the generation of SBPs from the groups found by a graph automorphism tool remains unchanged, and the SBPs can be added to the original instance as *new* hard clauses. The resulting instance is also an instance of partial MaxSAT. Correctness of this approach follows from the correctness of the plain MaxSAT case.

The solution for weighted MaxSAT and for weighted partial MaxSAT is similar to the partial MaxSAT case, but now clauses with different weights are represented by vertices with different colors. This guarantees that the groups found by the graph automorphism tool take into consideration the weight of each clause. Let $\{c_1, c_2, \ldots, c_k\}$ denote the distinct clause weights in the CNF formula. Each clause ω_i of weight c_i, represented as (ω_i, c_i) is associated with a vertex of color $i+1$ in the colored graph. In case there exist hard clauses, an additional color $k+2$ is used, and so each hard clause is represented by a vertex with color $k+2$ in the colored graph. Associating distinct clause weights with distinct colors guarantees that the graph automorphism algorithm can only make the correspondence between clauses with the same weight. Moreover, the identified SBPs result in new *hard* clauses that are added to the original problem. For either weighted MaxSAT or weighted partial MaxSAT, the result is an instance of weighted partial MaxSAT. As before, correctness of this approach follows from the correctness of the plain MaxSAT case.

Example 4. Consider the following weighted partial MaxSAT instance:

$$\varphi = (x_1 \vee x_2, 1) \wedge (\bar{x}_1 \vee x_2, 1) \wedge (\bar{x}_2, 5) \wedge (\bar{x}_3 \vee x_2, 9) \wedge (x_3 \vee x_2, 9)$$

for which the last two clauses are hard. Figure 2 shows the colored undirected graph associated with the formula. Clauses with different weights are represented with different colors (shown in the figure with different vertex shapes). A graph automorphism algorithm can then be used to generate the symmetry breaking predicates $\varphi_{sbp} = (\bar{x}_1) \wedge (\bar{x}_3)$, consisting of two hard clauses. As a result, the assignments $x_1 = 0$ and $x_3 = 0$ become necessary.



(a) Colored graph

Vertex	Mapping
1	literal x_1
2	literal x_2
3	literal x_3
4	literal $\bar{x}_1$
5	literal $\bar{x}_2$
6	literal $\bar{x}_3$
7	clause $(x_1 \vee x_2, 1)$
8	clause $(\bar{x}_1 \vee x_2, 1)$
9	clause $(\bar{x}_2, 5)$
10	clause $(\bar{x}_3 \vee x_2, 9)$
11	clause $(x_3 \vee x_2, 9)$

(b) Vertex mapping

Fig. 2. Colored graph for Example 4

Table 1. Problem transformations due to SBPs

Original	MS	PMS	WMS	WPMS
With Symmetries	PMS	PMS	WPMS	WPMS

Proposition 4. *The maximum number of satisfied clauses for the weighted (partial) MaxSAT problem φ and the resulting weighted partial MaxSAT problem $(\varphi \wedge \varphi_{sbp})$ are the same.*

Proof: *(Sketch) The proof is similar to the proof of Proposition 3, but noting that weights partition the set of clauses into sets of clauses that can be mapped into each other. Since mappings are between clauses with the same weights, the previous results (from Propositions 2 and 3) still hold.* ■

Table 1 summarizes the problem transformations described in this section, where MS represents plain MaxSAT, PMS represents partial MaxSAT, WMS represents weighted MaxSAT, and WPMS represents weighted partial MaxSAT. The use of SBPs introduces a number of hard clauses, and so the resulting problems are either partial MaxSAT or weighted partial MaxSAT.

3.4 Evaluating Alternative Formulations

Even though the proposed approach for breaking symmetries does not seem amenable to further optimizations for the MaxSAT and partial MaxSAT cases, it is interesting to investigate whether it is possible to optimize the approach outlined in the previous section for the weighted variants of MaxSAT, e.g. by reorganizing clause weights. This

section argues that, provided some simple conditions hold, rearranging weights cannot induce stronger symmetry breaking predicates.

Example 5. Consider the weighted MaxSAT formula:

$$\varphi = (x_1 \vee \bar{x}_2, 7) \wedge (\bar{x}_3 \vee x_4, 3) \wedge (\bar{x}_3 \vee x_4, 4) \tag{4}$$

The symmetries for this formula are $(x_1\,\bar{x}_2)$ and $(x_3\,\bar{x}_4)$. Clearly, it is possible to induce more symmetries by considering the following modification:

$$\varphi = (x_1 \vee \bar{x}_2, 3) \wedge (x_1 \vee \bar{x}_2, 4) \wedge (\bar{x}_3 \vee x_4, 3) \wedge (\bar{x}_3 \vee x_4, 4) \tag{5}$$

In addition to the previous symmetries, one now also obtains $(x_1\,x_3)(x_2\,x_4)$.

The previous example suggests that by rearranging weights one may be able to increase the number of identified symmetries. As the example also suggests, this can only happen when a clause is associated with *more* than one single weight. For the previous example $(\bar{x}_3 \vee x_4)$ is associated with weights 3 and 4. One simple way to tackle this problem is to require that multiple occurrences of the same clause be aggregated into a single clause, i.e. multiple occurrences of the same clause are represented by a single clause and the multiple weights are added.

Proposition 5. *If each clause has a single occurrence in formula φ, then splitting the weight of a clause induces no additional symmetries.*

Proof: *Suppose that each clause has a single occurrence, and that additional symmetries could be identified by splitting the weight c_i of a single clause ω_i. Without loss of generality assume that weight c_i is split into c_{i_1} and c_{i_2}. If additional symmetries can now be identified, then ω_i is mapped to clause ω_{j_1} due to c_{i_1} and to clause ω_{j_2} due to c_{i_2}. However, since each variable is mapped to some other variable, then ω_{j_1} and ω_{j_2} must be the same clause; but this is a contradiction.* ■

The previous result ensures that the approach outlined in Section 3.3, for computing symmetry breaking predicates for the weighted variations of MaxSAT, cannot be improved upon by rearranging clause weights, provided each clause has a single occurrence in the formula. Clearly, this is not the case with the earlier example.

4 Experimental Results

The approach outlined in the previous sections for generating SBPs for MaxSAT has been implemented in MAXSATSBP [5]. MAXSATSBP interfaces SAUCY [8], and is organized similarly to SHATTER [1] and SHATTERPB [2].

The experimental setup has been organized as follows. First, all the instances from the first and second MaxSAT evaluations (2006 and 2007) [3] were run. A timeout of 1000s of CPU time was considered, and instances requiring more than 1000s of CPU

[5] The MAXSATSBP tool is available on request from the authors.

time are declared as *aborted*. These results allowed selecting relevant benchmark families, for which symmetries occur and which require a non-negligible amount of time for being solved by both approaches (with or without SBPs). Afterwards, the instances for which both approaches aborted were removed from the tables of results. This resulted in selecting the `hamming` and the `MANN` instances for plain MaxSAT, the `ii32` and again the `MANN` instances for partial MaxSAT, the `c-fat500` instances for weighted MaxSAT and the `dir` and `log` instances for weighted partial MaxSAT.

Besides the instances that participated in the MaxSAT competition, we have included additional plain MaxSAT problem instances (`hole`, `Urq` and `chnl`). The `hole` instances refer to the well-known pigeon hole problem, the `Urq` instances represent randomized instances based on expander graphs and the `chnl` instances model the routing of wires in the channels of field-programmable integrated circuits. These instances refer to problems that can be naturally encoded as MaxSAT problems and are known to be highly symmetric [1]. The approach outlined above was also followed for selecting the instances to be included in the results.

We have run different publicly available MaxSAT solvers, namely MINIMAXSAT [6], TOOLBAR [7] and MAXSATZ [8]. (MAXSATZ accepts only plain MaxSAT instances.) Evidence from the MaxSAT evaluation suggests that the behavior of MINIMAXSAT is similar to TOOLBAR and MAXSATZ, albeit being in general more robust. For this reason, the results focus on MINIMAXSAT.

Tables 2 and 3 provide the results obtained. In the tables, TO denotes a timeout, and so the run time is in excess of 1000s. Table 2 refers to plain MaxSAT instances and Table 3 refers to partial MaxSAT (PMS), weighted MaxSAT (WMS) and weighted partial MaxSAT (WPMS) instances. For each instance, the results shown include the number of clauses added as a result of SBPs (#ClsSbp), the time required for solving the original instances (OrigT), i.e. without SBPs, and the time required for breaking the symmetries plus the time required for solving the extended formula afterwards (SbpT). (The best configuration for each instance is outlined in bold.) Moreover, the SbpT column is split into the time to run MAXSATSBP (MXSBP) and the time to run MINIMAXSAT. In practice, the time required for generating SBPs is negligible. The results were obtained on an Intel Xeon 5160 server (3.0GHz, 1333Mhz FSB, 4MB cache) running Red Hat Enterprise Linux WS 4.

The experimental results allow establishing the following conclusions:

– The inclusion of symmetry breaking is *essential* for solving a number of problem instances. We should note that *all* the plain MaxSAT instances in Table 2 for which MINIMAXSAT aborted, are also aborted by TOOLBAR and MAXSATZ. After adding SBPs all these instances become easy to solve by any of the solvers. For the aborted partial, weighted and weighted partial MaxSAT instances in Table 3 this is not always the case, since a few instances aborted by MINIMAXSAT could be solved by TOOLBAR without SBPs. However, the converse is also true, as there are instances that were initially aborted by TOOLBAR (although solved by MINIMAXSAT) that are solved by TOOLBAR after adding SBPs.

[6] http://www.lsi.upc.edu/∼fheras/docs/m.tar.gz
[7] http://carlit.toulouse.inra.fr/cgi-bin/awki.cgi/ToolBarIntro
[8] http://www.laria.u-picardie.fr/∼cli/maxsatz.tar.gz

Table 2. Results for MINIMAXSAT on plain MaxSAT instances

Name	#ClsSbp	OrigT	SbpT	MXSBP	MiniMaxSat
hamming10-2	81	TO	**0.19**	0.009	0.178
hamming10-4	1	886.57	**496.79**	0.01	496.777
hamming6-4	437	0.17	**0.15**	0.013	0.137
hamming8-2	85	TO	**0.21**	0.016	0.189
hamming8-4	253	0.36	**0.11**	0.011	0.102
MANN_a27	85	TO	**0.24**	0.012	0.226
MANN_a45	79	TO	**0.20**	0.011	0.185
MANN_a81	79	TO	**0.19**	0.01	0.184
hole10	758	42.11	**0.24**	0.023	0.213
hole11	922	510.90	**0.47**	0.023	0.442
hole12	1102	TO	**1.78**	0.028	1.752
hole7	362	**0.10**	0.11	0.007	0.103
hole8	478	0.40	**0.13**	0.008	0.122
hole9	610	3.68	**0.17**	0.016	0.15
Urq3_5	29	83.33	**0.27**	0.033	0.236
Urq4_5	43	TO	**50.88**	0.07	50.806
chnl10_11	1954	TO	**41.79**	0.053	41.737
chnl10_12	2142	TO	**328.12**	0.057	328.063
chnl11_12	2370	TO	**420.19**	0.075	420.111

- For several instances, breaking only a few symmetries can make the difference. We have observed that in some cases the symmetries are broken with unit clauses.
- Adding SBPs is beneficial for most cases where symmetries exist. However, for a few examples, SBPs may degrade performance.
- There is no clear relation between the number of SBPs added and the impact on the search time.
- The run time of the symmetry breaking tool is in general negligible.

Overall, the inclusion of SBPs should be considered when a hard problem instance is known to exhibit symmetries. This does not necessarily imply that after breaking symmetries the instance becomes trivial to solve, and there can be cases where the new clauses may degrade performance. However, in a significant number of cases, highly symmetric problems become much easier to solve after adding SBPs. In many of these cases the problem instances become *trivial* to solve.

5 Related Work

Symmetries are a well-known research topic, that serve to tackle complexity in many combinatorial problems. The first ideas on symmetry breaking were developed in the 80s and 90s [15, 4, 22, 7], by relating symmetries with the graph automorphism problem, and by proposing the first approach for generating symmetry breaking predicates. This work was later extended and optimized for propositional satisfiability [1].

Table 3. Results for MINIMAXSAT on partial, weighted and weighted partial MaxSAT instances

Name	MStype	#ClsSbp	OrigT	SbpT	MXSBP	MiniMaxSat
ii32e3	PMS	1756	94.40	**37.63**	0.482	37.15
ii32e4	PMS	2060	175.07	**129.06**	0.787	128.277
c-fat500-10	WMS	2	57.79	**11.62**	0.028	11.591
c-fat500-1	WMS	112	**0.03**	0.06	0.016	0.046
c-fat500-2	WMS	12	0.16	**0.11**	0.011	0.049
c-fat500-5	WMS	4	0.16	**0.11**	0.016	0.091
MANN_a27	WMS	1	TO	**880.58**	0.047	880.533
MANN_a45	WMS	1	TO	**530.86**	0.048	530.807
MANN_a81	WMS	1	TO	**649.13**	0.042	649.084
1502.dir	WPMS	1560	**0.34**	10.67	0.754	9.912
29.dir	WPMS	132	TO	**28.09**	0.031	28.055
54.dir	WPMS	98	4.14	**0.32**	0.029	0.292
8.dir	WPMS	58	**0.03**	0.05	0.008	0.039
1502.log	WPMS	812	0.76	**0.71**	0.32	0.385
29.log	WPMS	54	17.55	**0.82**	0.026	0.792
404.log	WPMS	124	TO	**64.24**	0.094	64.151
54.log	WPMS	48	2.37	**0.16**	0.021	0.139

Symmetries are an active research topic in CP [11]. Approaches for breaking symmetries include not only adding constraints before search [22] but also reformulation [23] and dynamic symmetry breaking methods [12]. Recent work has also shown the application of symmetries to soft CSPs [24].

The approach proposed in this paper for using symmetry breaking for MaxSAT and variants builds on earlier work on symmetry breaking for PB constraints [2]. Similarly to the work for PB constraints, symmetries are identified by constructing a colored graph, from which graph automorphisms are obtained, which are then used to generate the symmetry breaking predicates.

6 Conclusions

This paper shows how symmetry breaking can be used in MaxSAT and in its most well-known variants, including partial MaxSAT, weighted MaxSAT, and weighted partial MaxSAT. Experimental results, obtained on representative instances from the MaxSAT evaluation [3] and practical instances [1], demonstrate that symmetry breaking allows solving problem instances that no state of the art MaxSAT solver could otherwise solve. For all problem instances considered, the computational effort of computing symmetries is negligible. Nevertheless, and as it is the case with symmetry breaking for SAT and PB constraints, symmetry breaking should be considered as an optional problem solving technique, to be used when standard techniques are unable to solve a given problem instance.

The experimental results motivate additional work on computing symmetry breaking predicates for MaxSAT. A new more efficient version of Saucy has recently been

developed [9] and is likely to further reduce the run time for computing symmetries. Moreover, the use of conditional symmetries could be considered [10, 24].

Acknowledgement. This work is partially supported by EPSRC grant EP/E012973/1, by EU grants IST/033709 and ICT/217069, and by FCT grants POSC/EIA/61852/2004 and PTDC/EIA/76572/2006.

References

1. F. Aloul, K. A. Sakallah, and I. Markov. Efficient symmetry breaking for boolean satisfiability. In *International Joint Conference on Artificial Intelligence*, pages 271–276, August 2003.
2. F. A. Aloul, A. Ramani, I. L. Markov, and K. A. Sakallah. ShatterPB: symmetry-breaking for pseudo-Boolean formulas. In *Asian and South-Pacific Design Automation Conference*, pages 883–886, 2004.
3. J. Argelich, C. M. Li, F. Manyà, and J. Planes. MaxSAT evaluation. www.maxsat07.udl.es, May 2007.
4. B. Benhamou and L. Sais. Theoretical study of symmetries in propositional calculus and applications. In *Eleventh International Conference on Automated Deduction*, pages 281–294, 1992.
5. B. Borchers and J. Furman. A two-phase exact algorithm for MAX-SAT and weighted MAX-SAT problems. *Journal of Combinatorial Optimization*, 2(4):299–306, 1998.
6. D. A. Cohen, P. Jeavons, C. Jefferson, K. E. Petrie, and B. M. Smith. Symmetry definitions for constraint satisfaction problems. In *International Conference on Principles and Practice of Constraint Programming*, pages 17–31, 2005.
7. J. M. Crawford, M. L. Ginsberg, E. Luks, and A. Roy. Symmetry-breaking predicates for search problems. In *International Conference on Principles of Knowledge Representation and Reasoning*, pages 148–159, 1996.
8. P. T. Darga, M. H. Liffiton, K. A. Sakallah, and I. L. Markov. Exploiting structure in symmetry detection for CNF. In *Design Automation Conference*, pages 530–534, June 2004.
9. P. T. Darga, K. A. Sakallah, and I. L. Markov. Faster symmetry discovery using sparsity of symmetries. In *Design Automation Conference*, pages 149–154, June 2008.
10. I. P. Gent, T. Kelsey, S. Linton, I. McDonald, I. Miguel, and B. M. Smith. Conditional symmetry breaking. In *International Conference on Principles and Practice of Constraint Programming*, pages 256–270, 2005.
11. I. P. Gent, K. E. Petrie, and J.-F. Puget. *Handbook of Constraint Programming*, chapter Symmetry in Constraint Programming, pages 329–376. Elsevier, 2006.
12. I. P. Gent and B. M. Smith. Symmetry breaking in constraint programming. In *European Conference on Artificial Intelligence*, pages 599–603, 2000.
13. F. Heras and J. Larrosa. New inference rules for efficient Max-SAT solving. In *AAAI Conference on Artificial Intelligence*, 2006.
14. F. Heras, J. Larrosa, and A. Oliveras. MiniMaxSat: a new weighted Max-SAT solver. In *International Conference on Theory and Applications of Satisfiability Testing*, pages 41–55, May 2007.
15. B. Krishnamurthy. Short proofs for tricky formulas. *Acta Informatica*, 22(3):253–275, 1985.
16. C. M. Li, F. Manyà, and J. Planes. Exploiting unit propagation to compute lower bounds in branch and bound Max-SAT solvers. In *International Conference on Principles and Practice of Constraint Programming*, pages 403–414, 2005.

17. C. M. Li, F. Manyà, and J. Planes. Detecting disjoint inconsistent subformulas for computing lower bounds for Max-SAT. In *AAAI Conference on Artificial Intelligence*, July 2006.
18. H. Lin and K. Su. Exploiting inference rules to compute lower bounds for MAX-SAT solving. In *International Joint Conference on Artificial Intelligence*, pages 2334–2339, 2007.
19. J. Marques-Silva, I. Lynce, and V. Manquinho. Symmetry breaking for maximum satisfiability. *Computing Research Repository*, abs/0804.0599, April 2008. Available from http://arxiv.org/abs/0804.0599.
20. J. Marques-Silva and V. M. Manquinho. Towards more effective unsatisfiability-based maximum satisfiability algorithms. In *International Conference on Theory and Applications of Satisfiability Testing*, pages 225–230, 2008.
21. J. Marques-Silva and J. Planes. Algorithms for maximum satisfiability using unsatisfiable cores. In *Design, Automation and Testing in Europe Conference*, pages 408–413, March 2008.
22. J.-F. Puget. On the satisfiability of symmetrical constrained satisfaction problems. In *International Symposium on Methodologies for Intelligent Systems*, pages 350–361, 1993.
23. B. M. Smith. Reducing symmetry in a combinatorial design problem. Technical Report 2001.01, School of Computing, University of Leeds, January 2001. Presented at the CP-AI-OR Workshop, April 2001.
24. B. M. Smith, S. Bistarelli, and B. O'Sullivan. Constraint symmetry for the soft CSP. In *International Conference on Principles and Practice of Constraint Programming*, pages 872–879, September 2007.
25. R. Wallace and E. Freuder. Comparative studies of constraint satisfaction and Davis-Putnam algorithms for maximum satisfiability problems. In D. Johnson and M. Trick, editors, *Cliques, Coloring and Satisfiability*, volume 26 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 587–615. American Mathematical Society, 1996.
26. Z. Xing and W. Zhang. MaxSolver: An efficient exact algorithm for (weighted) maximum satisfiability. *Artificial Intelligence*, 164(1-2):47–80, 2005.