ELSEVIER

Information Processing Letters 58 (1996) 23-29

Information Processing Letters

On approximation algorithms for the minimum satisfiability problem

M.V. Marathe, [a,*,1], S.S. Ravi [b,2]

[a] Los Alamos National Laboratory, P.O. Box 1663, MS M986, Los Alamos, NM 87545, USA

[b] Department of Computer Science, University at Albany - SUNY, Albany, NY 12222, USA

Received 30 July 1995; revised 10 January 1996

Communicated by L.A. Hemaspaandra

Abstract

We consider the following minimum satisfiability (MINSAT) problem: Given a CNF formula, find a truth assignment to the variables that *minimizes* the number of satisfied clauses. This problem was shown to be NP-complete by R. Kohli, R. Krishnamurti and P. Mirchandani (1994). They also presented an approximation algorithm whose performance guarantee is equal to the maximum number of literals in a clause. We present an approximation-preserving reduction from MINSAT to the minimum vertex cover (MINVC) problem. This reduction, in conjunction with known heuristics for the MINVC problem, yields a heuristic with a performance guarantee of 2 for MINSAT. Further, we show that if there is an approximation algorithm with a performance guarantee ρ for MINSAT, then there is an approximation algorithm with the same performance guarantee ρ for the minimum vertex cover problem. This result points out the difficulty of devising an approximation algorithm with a performance guarantee better than 2 for MINSAT. We also observe that MINSAT remains NP-complete even when restricted to planar instances.

Keywords: Analysis of algorithms; Computational complexity

1. Introduction

The problem of finding a satisfying assignment to a given formula in conjunctive normal form (CNF), commonly referred to as SAT, is a fundamental problem in theoretical computer science [8,20]. An optimization version (MAXSAT) of this problem is to find a truth assignment that *maximizes* the number of satisfied clauses. In [9] MAXSAT was shown to be NP-complete even when each clause contains only two literals. A number of researchers have investigated approximation algorithms (heuristics) for MAXSAT (see for example [17,21,24,10,11,6]). In [18], Kohli et al. introduced the *minimum satisfiability problem* (MINSAT) where the goal is to find a truth assignment that *minimizes* the number of satisfied clauses. They showed that MINSAT is NP-complete even when each clause contains only two literals. They also showed that MINSAT remains NP-complete even for Horn formulas (i.e., CNF formulas in which each clause has at most one complemented variable). Further, they analyzed a deterministic and a probabilistic greedy heuristic for MINSAT. For the deterministic version, they proved that the performance guarantee provided by

* Corresponding author.

[1] Supported by U.S. Department of Energy Contract W-7405-ENG-36. E-mail: madhar@c3.lanl.gov.

[2] E-mail: ravi@cs.albany.edu.

PII S0020-0190(96)00031-2

the heuristic is equal to the maximum number of literals in any clause. For the probabilistic version, they showed that the expected number of clauses satisfied by any assignment produced by the heuristic is at most twice the number of clauses satisfied by an optimal assignment.

In this paper, our focus is on *deterministic* approximation algorithms for the MINSAT problem. From now on, we will use the word "heuristic" to mean a deterministic approximation algorithm which runs in polynomial time. Note that when the clauses are of size $\Theta(n)$, where n is the number of variables, the heuristic analyzed in [18] provides only a weak performance guarantee of $\Theta(n)$. We present a simple approximation-preserving reduction from MINSAT to the minimum vertex cover (MINVC) problem. This reduction, in conjunction with known heuristics for the MINVC problem (see for example, [8,22]), yields a heuristic with a performance guarantee of 2 for MINSAT, thus improving the result of Kohli et al. [18]. We also show that MINSAT is as hard to approximate as MINVC; that is, if there is a heuristic with a performance guarantee ρ for MINSAT, then there is a heuristic with the same performance guarantee ρ for MINVC. Moreover, we show that this result holds even for MINSAT instances defined by Horn formulas. It has been conjectured in [12] that no polynomial approximation algorithm can provide a performance guarantee of $2 - \varepsilon$ for any fixed $\varepsilon > 0$ for MINVC unless P = NP. Thus, our result provides an indication of the difficulty involved in devising a heuristic with a performance guarantee better than 2 for MINSAT.

Our results also point out a close relationship between MINSAT and MINVC problems. As a byproduct of this relationship, we also obtain that MINSAT remains NP-complete even when restricted to *planar* instances. As another corollary, we obtain a polynomial time approximation scheme for a restricted version of the planar minimum satisfiability problem.

Recently, another heuristic which also provides a performance guarantee of 2 for MINSAT has been obtained using randomized rounding [23,5]. Also, Bertsimas, Teo and Vohra [4] have obtained a heuristic with an improved performance guarantee for the MIN-2SAT problem (i.e., instances of MINSAT where each clause contains at most two literals). This heuristic, also based on randomized rounding, provides a performance guarantee of $(\sqrt{5}+1)/2 \approx 1.61$.

When we were preparing the final version of the paper, another approach for obtaining a 2-approximation algorithm for MINSAT was brought to our attention by Professor Dorit Hochbaum [13]. This approach is based on expressing a MINSAT instance as a $\{0,1\}$ integer linear program with two variables per inequality and then using the results in [16,15]. Further, it has been shown [14] that $\{0,1\}$ integer linear programs with binary variables and with two variables per inequality are equivalent to MINVC. Thus, this approach also shows that the approximation problem for MINSAT is equivalent to that for MINVC.

The remainder of this paper is organized as follows. We present the necessary definitions in Section 2. Section 3 presents our results that relate the approximabilities of MINSAT and MINVC problems.

2. Definitions and preliminaries

We begin with the formal definition of the MINSAT problem.

Minimum satisfiability (MINSAT).
Instance: A set $C = \{c_1, c_2, \ldots, c_m\}$ of m clauses made up of uncomplemented and complemented occurrences of variables from the set $X = \{x_1, x_2, \ldots, x_n\}$.
Required: Find a truth assignment to the variables that satisfies the *minimum* number of clauses.

We say that a heuristic for a minimization problem Π provides a *performance guarantee* ρ if the value of the solution produced by the heuristic is at most ρ times the value of an optimal solution for all instances of the problem. An *approximation scheme* for a problem Π is a family of heuristics such that given an instance I of Π and an $\varepsilon > 0$, there is a member of the family that returns a solution which is within a factor $(1+\varepsilon)$ of the optimal value for I.

As pointed out in [18], given a CNF formula, it is possible to determine efficiently whether there is an assignment that satisfies *zero* clauses. The necessary and sufficient condition for satisfying zero clauses is stated in the following observation.

Observation 2.1. *For any instance of* MINSAT, *there is a truth assignment that satisfies zero clauses if and* *only if each variable appears only in complemented form or only in uncomplemented form in the CNF formula.*

Note that for a MINSAT instance satisfying the condition of Observation 2.1, a truth assignment that satisfies zero clauses can be constructed by setting each variable that appears in complemented (uncomplemented) form to true (false).

Since we are seeking ratio performance guarantees, we assume for the remainder of this paper that for any given instance of MINSAT, every truth assignment will satisfy at least one clause. We also assume that no clause contains a variable as well the complement of that variable. (Such clauses are satisfied regardless of the truth assignment used.)

The main results of this paper relate the approximabilities of MINSAT and MINVC. For the sake of completeness, we provide a formal definition of MINVC below.

Minimum vertex cover (MINVC).

Instance: An undirected graph $G(V,E)$.

Required: Find a minimum cardinality subset V' of V such that for each edge $\{u,v\}$ in E, at least one of u and v is in V'.

An *independent set* in a graph is a set of nodes which are pairwise non-adjacent. The following is a well-known (and easy to verify) fact that relates independent sets and vertex covers of a graph.

Fact 2.2. *For any graph $G(V,E)$, $V' \subseteq V$ is a vertex cover if and only if $V - V'$ is an independent set.*

3. Approximability of MINSAT

3.1. A heuristic for MINSAT

Our heuristic for MINSAT uses the following definition.

Definition 3.1. Let I be an instance of MINSATconsisting of the clause set C_I and variable set X_I. The *auxiliary graph* $G_I(V_I,E_I)$ corresponding to I is constructed as follows. The node set V_I is in one-to-one correspondence with the clause set C_I. For any two nodes v_i and v_j in V_I, the edge $\{v_i,v_j\}$ is in E_I if and only if the corresponding clauses c_i and c_j are such that there is a variable $x \in X_I$ that appears in uncomplemented form in c_i and in complemented form in c_j or vice versa.

The intuition behind the auxiliary graph is the following. Whenever there is an edge in the auxiliary graph between the nodes corresponding to clauses c_i and c_j, no truth assignment to the variables can make both c_i and c_j to be false; that is, any truth assignment that forces c_i to be false must necessarily make c_j to be true and vice versa. This property of the auxiliary graph allows us to prove our next lemma. From now on, we do not distinguish between a node of the auxiliary graph and the corresponding clause of the CNF formula.

Lemma 3.2. *Let I be an instance of* MINSAT *with clause set C_I and let G_I be the corresponding auxiliary graph.*

(1) *Given any truth assignment for which the number of satisfied clauses of the* MINSAT *instance I is equal to k, we can find a vertex cover of size k for G_I.*

(2) *Given any vertex cover C' of size k for G_I, we can find a truth assignment that satisfies at most k clauses of the* MINSAT *instance I.*

Proof. (1) Let C', with $|C'| = k$, be the set of all clauses satisfied by the given truth assignment. We claim that the set C'' defined by $C'' = C_I - C'$, is an independent set in G_I. To see this, note that the given truth assignment sets all the clauses in C'' to be false. If there is an edge between some pair of nodes c and c' in C'', then by the definition of the auxiliary graph, no truth assignment can cause both c and c' to false. Thus, C'' is an independent set, and from Fact 2.2, $C' = C_I - C''$ is a vertex cover of size k for G_I.

(2) Let C' be a vertex cover of size k for G_I. Thus, from Fact 2.2, $C'' = C_I - C'$ is an independent set in G_I. Observe that the clauses in C'' are such that each variable appears only in complemented form or in uncomplemented form. (If there are two clauses c, c' in C'' such that some variable x appears in complemented form in c and in uncomplemented form in c' or vice versa, then the edge $\{c,c'\}$ will be in the auxiliary graph, and so C'' cannot be an independent set.)

Therefore, by Observation 2.1, we can find a truth assignment to the variables appearing in the clauses of C'' such that each clause in C'' is set to false. For other variables (if any) appearing in the clauses of C', we can arbitrarily assign true or false values. Since the truth assignment causes at least $|C''| = |C_I - C'|$ clauses to be false, the number of clauses satisfied is at most $|C'| = k$. □

Corollary 3.3. *Let I be an instance of* MINSAT *and let G_I be the corresponding auxiliary graph. Further, let $OPT(I)$ denote the number of clauses satisfied by an optimal assignment to the variables of I and let V^* denote a minimum vertex cover for G_I. Then, $OPT(I) = |V^*|$.*

Our heuristic for MINSAT is shown in Fig. 1. As mentioned earlier, several heuristics that provide a performance guarantee of 2 are available for MINVC [8,22]. Any of these heuristics can be used in Step 2 of our heuristic. Thus, it is clear that Heuristic-MINSAT runs in polynomial time. The performance guarantee provided by the heuristic is an immediate consequence of Lemma 3.2 and Corollary 3.3.

Theorem 3.4. *Let I be an instance of* MINSAT. *Let $OPT(I)$ and $HEU(I)$ denote the number of clauses satisfied by an optimal assignment and that produced by* Heuristic-MINSAT *respectively. Then $HEU(I) \leqslant 2OPT(I)$.*

3.2. Tightness of approximation

We now show that MINSAT is at least as hard to approximate as MINVC by presenting an approximation-preserving reduction from MINVC to MINSAT. Some general techniques for proving non-approximability results using interactive proof systems are presented in [3,7].

Theorem 3.5. *If there is a heuristic with a performance guarantee of ρ for* MINSAT, *then there is a heuristic with the same performance guarantee ρ for* MINVC.

Proof. Let $\mathcal{A}$ be a heuristic with a performance guarantee of ρ for MINSAT. Consider an arbitrary instance of MINVC given by graph $G(V,E)$. Let V^* denote a minimum vertex cover for G. Construct an instance I of MINSAT as follows. For each node v_i of V, construct a clause c_i. For each edge (v_i, v_j) of G, create a new variable x_{ij}; let the clause c_i contain the uncomplemented occurrence of x_{ij} and let c_j contain the complemented occurrence of x_{ij}. (Thus, the size of clause c_i is simply the degree of node v_i.) Let $OPT(I)$ denote the number of clauses satisfied by an optimal truth assignment to the variables in I.

We note that the resulting MINSAT instance I has the property that the graph G is itself the auxiliary graph corresponding to I. Therefore, by Corollary 3.3, $|V^*| = OPT(I)$.

Suppose we execute $\mathcal{A}$ on I. Let $\mathcal{A}(I)$ be the number of clauses satisfied by the truth assignment produced by $\mathcal{A}$. Since $\mathcal{A}$ provides a performance guarantee of ρ, we have $\mathcal{A}(I) \leqslant \rho OPT(I)$. By Lemma 3.2(1), the truth assignment produced by $\mathcal{A}$ can be converted into a vertex cover of size $\mathcal{A}(I)$ for G. Thus, the resulting vertex cover satisfies the condition $\mathcal{A}(I) \leqslant \rho OPT(I) = \rho|V^*|$. Thus $\mathcal{A}$ can be used to construct a heuristic with a performance guarantee of ρ for MINVC. □

It is known that MINVC is MAX SNP-hard [21,5], and hence unless P = NP, does not have a polynomial time approximation scheme [1]. From the reduction presented in the proof of Theorem 3.5, it can be seen that MINSAT is also MAX SNP-hard. Therefore, unless P = NP, MINSAT does not have a polynomial time approximation scheme.

3.3. The complexity of planar MINSAT

The planar version of SAT was defined by Lichtenstein [19]. We adopt the same definition for planar version of MINSAT (denoted by Pl-MINSAT) as indicated below.

Definition 3.6. Given an instance I of MINSAT, consider the following bipartite graph $\mathcal{B}_I(\mathcal{C}_I, \mathcal{V}_I, \mathcal{E}_I)$. The node set $\mathcal{C}_I$ ($\mathcal{V}_I$) is in one-to-one correspondence with the set of clauses (variables) in I. For any $c \in \mathcal{C}_I$ and $v \in \mathcal{V}_I$, the edge $\{c, v\}$ is in $\mathcal{E}_I$ if the clause corresponding to c contains the variable corresponding to v (in complemented or uncomplemented form). An instance I of MINSAT is planar if the corresponding bipartite graph $\mathcal{B}_I$ is planar.

Heuristic-MINSAT

Step 1. Given the instance I of MINSAT, construct the corresponding auxiliary graph $G_I(V_I, E_I)$.

Step 2. Construct an approximate vertex cover V' for G_I such that $|V'|$ is at most twice that of a minimum vertex cover for G_I.

Step 3. Construct a truth assignment that causes all the clauses in $V_I - V'$ to be false. (See Lemma 3.2 and Observation 2.1.)

Step 4. Output the truth assignment found in Step 3.

Fig. 1. Details of the heuristic for MINSAT.

It is known that MINVC restricted to planar graphs (denoted by Pl-MINVC) is also NP-complete [9]. If we start with an instance of Pl-MINVC, and carry out the construction of the MINSAT instance as described in the proof of Theorem 3.5, it can be verified that we would obtain an instance of Pl-MINSAT. (The node corresponding to the variable x_{ij}, which appears only in the two clauses corresponding to the nodes v_i and v_j, can be placed on the edge $\{v_i, v_j\}$ without affecting planarity.) Since Pl-MINSAT is obviously in NP, we conclude the following.

Proposition 3.7. *The problem* Pl-MINSAT *is* NP-*complete.*

Observe that Proposition 3.7 holds even when restricted to instances in which each variable appears in at most two clauses. Consider a restricted version of Pl-MINSAT in which each variable appears in at most three clauses. Denote this problem by Pl-MIN-3O-SAT. From the above discussion, it follows that Pl-MIN-3O-SAT is NP-complete. By a simple approximation-preserving reduction to Pl-MINVC, we show that the problem Pl-MIN-3O-SAT has a polynomial time approximation scheme. To see this, we observe that the auxiliary graph of any instance of Pl-MIN-3O-SAT is planar. Fig. 2 depicts how to lay out the auxiliary graph so that its planarity becomes evident. As shown in [2], Pl-MINVC has a polynomial time approximation scheme. Thus we have:

Corollary 3.8. *The problem* Pl-MIN-3O-SAT *has a polynomial time approximation scheme.*

It would be of interest to investigate whether Pl-MINSAT has a polynomial time approximation scheme.

3.4. MINSAT with Horn clauses

A Horn formula is a CNF formula in which each clause contains at most one complemented literal. Let Horn-MINSAT denote instances of MINSAT in which the CNF formula is a Horn formula. As mentioned earlier, Horn-MINSAT is also NP-complete [18]. Hence, it is of interest to investigate heuristics for Horn-MINSAT. From Theorem 3.4, we know that Heuristic-MINSAT provides performance guarantee of 2 for Horn-MINSAT. We now show that the restriction of MINSAT to Horn formulas does not make the approximation problem any easier. In other words, Horn-MINSAT is as hard to approximate as MINVC.

Theorem 3.9. *If there is a heuristic with a performance guarantee of ρ for* Horn-MINSAT, *then there is a heuristic with the same performance guarantee ρ for* MINVC.

Proof. Let $\mathcal{A}$ be a heuristic with a performance guarantee of ρ for Horn-MINSAT. Consider an arbitrary instance of MINVC given by graph $G(V, E)$. We show how an instance I of Horn-MINSAT such that the auxiliary graph of I is G can be constructed. For each node v_i of V, construct a variable x_i as well as a clause c_i. The literals in c_i are chosen as follows. Let $v_{i_1}, v_{i_2}, \ldots, v_{i_k}$ be the nodes adjacent to v in G. Then clause c_i is defined by

$$c_i = (\overline{x_i} \vee x_{i_1} \vee x_{i_2} \vee \cdots \vee x_{i_k}).$$

Clearly, each clause contains only one negated literal, and so the reduction produces a Horn formula.

It can be verified that the auxiliary graph for the instance I is G itself. The remainder of the proof is identical to that of Theorem 3.5. □

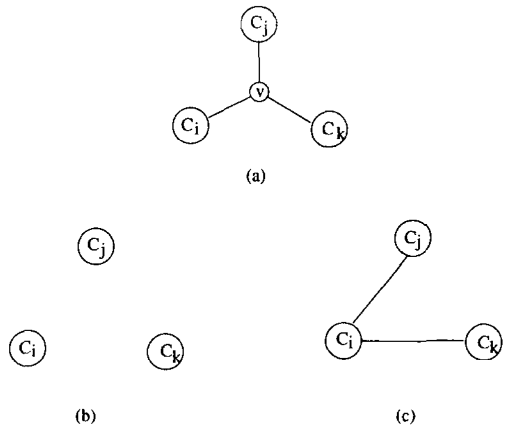

Fig. 2. Schematic diagram showing how to preserve planarity when constructing the auxiliary graph for an instance of PI-MIN-3O-SAT. (a) Variable v and the clauses c_i, c_j and c_k in which v appears. (b) When v appears only in uncomplemented form or only in complemented form in all the clauses. (c) When v appears in uncomplemented form in c_i and in complemented form in c_j and c_k or vice versa. (All other cases are similar to case (b).)

We note that while the construction presented in the proof of Theorem 3.5 preserves planarity, the construction presented in the proof of Theorem 3.9 will not, in general, preserve planarity.

Acknowledgements

We thank Professor Dorit Hochbaum (University of California, Berkeley) for bringing to our attention the relationship between MINSAT and integer linear programs with two variables per inequality. We are grateful to the referees for providing several valuable suggestions.

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy, Proof verification and hardness of approximation problems, in: *Proceedings 33rd Annual Symposium on Foundations of Computer Science*, Pittsburgh, PA (1992) 14–23.

[2] B.S. Baker, Approximation algorithms for NP-complete problems on planar graphs, *J. ACM* **41** (1994) 153–180.

[3] M. Bellare, O. Goldreich and M. Sudan, Free bits, PCPs and non-approximability-towards tight results, in: *Proceedings 36rd Annual Symposium on Foundations of Computer Science*, Milwaukee, WI (1995) 422–431.

[4] D. Bertsimas, C. Teo and R. Vohra, New randomized rounding algorithms, Preprint (1995).

[5] P. Crescenzi and V. Kann, A compendium of NP optimization problems, Manuscript (1995).

[6] U. Feige and M.X. Goemans, Approximating the value of two prover proof systems, with applications to MAX 2SAT and MAX DICUT, in: *Proceedings Third Israel Symposium on Theory of Computing and Systems*, Tel Aviv (1995).

[7] U. Feige, S. Goldwasser, L. Lovasz, S. Safra and M. Szegedy, Approximating clique is almost NP-complete, in: *Proceedings 32nd Annual Symposium on Foundations of Computer Science*, San Juan, PR (1991) 2–12.

[8] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, CA, 1979).

[9] M.R. Garey, D.S. Johnson and L. Stockmeyer, Some simplified NP-complete graph problems, *Theoret. Comput. Sci.* **1** (1976) 237–267.

[10] M.X. Goemans and D.P. Williamson, New 3/4 approximation algorithms for MAXSAT, *SIAM J. Discrete Math.* **7** (1994) 656–666.

[11] M.X. Goemans and D.P. Williamson, .878 approximation algorithms for MAX CUT and MAX 2SAT, in: *Proceedings 26th Annual ACM Symposium on Theory of Computing*, Montreal, Que. (1994) 422–431.

[12] D.S. Hochbaum, Approximation algorithms for the weighted set covering and node cover problems, *SIAM J. Comput.* **11** (1982) 555–556.

[13] D.S. Hochbaum, Personal communication, 1996.

[14] D.S. Hochbaum, ed., *Approximation Algorithms for NP-Hard Problems* (PWS Publishing Company, Boston, MA, to appear).

[15] D.S. Hochbaum, N. Meggido, J. Naor and A. Tamir, Tight bounds and 2-approximation algorithms for integer programs with two variables per inequality, *Math. Programming* **62** (1993) 69–83.

[16] D.S. Hochbaum and J. Naor, Simple and fast algorithms for linear and integer programs with two variables per inequality, *SIAM J. Comput.* **23** (1994) 1179–1192.

[17] D.S. Johnson, Approximation algorithms for combinatorial problems, *J. Comput. System Sci.* **9** (1974) 256-278.
[18] R. Kohli, R. Krishnamurti and P. Mirchandani, The minimum satisfiability problem, *SIAM J. Discrete Math.* **7** (1994) 275-283.
[19] D. Lichtenstein, Planar formulae and their uses, *SIAM J. Comput.* **11** (1982) 329-343.
[20] C.H. Papadimitriou, *Computational Complexity* (Addison-Wesley, Reading, MA, 1994).
[21] C.H. Papadimitriou and M. Yannakakis, Optimization, approximation and complexity classes, *J. Comput. System Sci.* **43** (1991) 425-440.
[22] C. Savage, Depth-first search and the vertex cover problem, *Inform. Process. Lett.* **14** (1982) 233-235.
[23] R. Vohra, Personal communication, 1995.
[24] M. Yannakakis, On the approximation of maximum satisfiability, *J. Algorithms* **17** (1994) 475-502.