Detecting Embedded Horn Structure in Propositional Logic *

V. Chandru
School of Industrial Engineering
Purdue University, West Lafayette, IN 47907 USA

J. N. Hooker
Graduate School of Industrial Administration
Carnegie Mellon University, Pittsburgh, PA 15213 USA

April 1991
Revised January 1992

Abstract

We show that the problem of finding a maximum renamable Horn problem within a propositional satisfiability problem is NP-hard but can be formulated as a set packing and therefore a maximum clique problem, for which numerous algorithms and heuristics have been developed.

1 Introduction

Horn clauses are widely used because, for them, satisfiability and inference problems are soluble in linear time. "Renamable Horn" problems (which are Horn up to a rescaling of variables) are also soluble in linear time. We address the problem of obtaining a renamable Horn problem by removing as few variables as possible from a given non-Horn satisfiability problem. One can then solve the original problem by enumerating truth assignments to the removed variables and solving a renamable Horn problem for each assignment.

We show that finding a maximal renamable Horn subproblem can be formulated as a maximum clique problem, for which numerous algorithms and heuristics have been developed and tested [2, 3, 6, 7, 8, 9]. We also observe that finding such a subproblem is NP-hard.

*The first author is partially supported by ONR grant N00014-86-K-0689 and NSF grant DMC 88-07550, and the second by AFOSR grant 91-0287.

2 Embedded Renamable Horn Sets

A set of clauses is *Horn* when each clause in it contains at most one positive literal. To *scale* a set of clauses is to replace every occurrence of x_j with $\neg x_j$ and every occurrence of $\neg x_j$ with x_j, for zero or more variables x_j. A *renamable Horn set* (RHS) is a clause set that can be scaled to obtain a Horn set.

The time required to check the satisfiability of a Horn set is linear in the number of literals [5]. There are also linear-time algorithms that determine whether a given clause set is renamable Horn and find an appropriate scaling when one exists [1, 4]. It is therefore possible to solve the satisfiability problem for an RHS in linear time.

Given a clause set S, let X be a subset of the variables occurring in S, and let $\overline{X}$ contain the variables not in X. Define $S(X)$ to be the result of removing from S all occurrences of variables in $\overline{X}$. If $S(X)$ is renamable Horn, we say that it is an *embedded RHS* of S. An embedded RHS of maximum size is a *maximum* embedded RHS.

Let $v : \overline{X} \to \{T, F\}$ be a mapping that assigns truth values to variables in $\overline{X}$. Then $S(X, v)$ is the clause set that results when each x_j occurring in $\overline{X}$ is fixed to the value $v(x_j)$. That is, $S(X, v)$ is the result of removing from S every clause containing a negated variable x_j for which $v(x_j) = F$, every clause containing a posited x_j for which $v(x_j) = T$, every negated occurrence of a variable x_j for which $v(x_j) = T$, and every posited occurrence of a variable x_j for which $v(x_j) = F$.

Since $S(X, v) \subset S(X)$ for any assignment v, $S(X, v)$ is renamable Horn if $S(X)$ is. Also S is satisfiable if and only if $S(X, v)$ is satisfiable for some v. Thus if $S(X)$ is renamable Horn, we can check S for satisfiability in $O(2^{|\overline{X}|}L)$ time, where L is the number of literals in $S(X)$, by enumerating the $2^{|\overline{X}|}$ assignments v. We naturally prefer $S(X)$ to be a maximum embedded RHS of S, so that $|\overline{X}|$ is as small as possible.

3 Finding a Maximum Embedded RHS

The problem of finding a maximum embedded RHS of a set S of m clauses containing n variables can be formulated as the following set packing problem.

$$\begin{aligned} \max \quad & \textstyle\sum_j y_j + \overline{y}_j \\ \text{s.t.} \quad & Ay + B\overline{y} \le e \\ & y_j + \overline{y}_j \le 1, \quad \text{all } j \\ & y_j, \overline{y}_j \in \{0, 1\}, \quad \text{all } j, \end{aligned} \tag{1}$$

Here e is a vector of 1's and A and B are 0-1 $m \times n$ matrices. We define A by letting $a_{i,j} = 1$ precisely when the literal x_j occurs in clause i, and B by letting $b_{ij} = 1$ precisely when $\neg x_j$ occurs in clause i.

We interpret $y_j = 1$ as indicating a positive scaling for x_j and $\overline{y}_j = 1$ as indicating a negative scaling. If $y_j = \overline{y}_j = 0$, we omit variable x_j altogether. Thus problem (1) finds a largest set of variables that, when rescaled in some fashion, yield a Horn set. In other words, it finds a maximum embedded RHS. We have shown the following.

Theorem 1 *If $(y, \overline{y})$ solves (1), then $S(\{x_j \mid y_j + \overline{y}_j = 1\})$ is a maximum embedded RHS of S.*

It is well known that a set packing problem,

$$\begin{array}{rl} \max & \sum_j z_j \\ \text{s.t.} & Qz \leq e \quad z_j \in \{0,1\}, \end{array} \tag{2}$$

can be formulated as a maximum clique problem on a graph. The graph contains a node for each z_j and an arc (z_j, z_k) whenever columns j and k of Q are orthogonal. A clique is a set of nodes in which every pair is connected by an arc. If C is a clique of maximum size, then z given by $z_j = 1$ if node $z_j \in C$, and $z_j = 0$ otherwise, is an optimal solution of the set packing problem. In the present case $z = (y, \overline{y})$ and $Q = \begin{bmatrix} A & B \\ I & I \end{bmatrix}$.

We can not only solve the maximum embedded RHS problem as a set packing problem but can do the reverse as well. Given an $m \times n$ set packing problem (2), consider the clause set,

$$\begin{array}{l} \bigvee_{q_{ij}=1} x_j, \quad i = 1, \ldots, m, \\ \neg x_1 \vee \ldots \vee \neg x_n \vee y_1 \vee \neg y_2, \\ \neg x_1 \vee \ldots \vee \neg x_n \vee \neg y_1 \vee y_2. \end{array} \tag{3}$$

None of the x_j's can be negatively scaled in the maximum embedded RHS of (3). Clearly, at most one can be negatively scaled, and if one is, y_1 and y_2 must be deleted. In this case one could do better by deleting the negatively scaled variable and retaining y_1 and y_2. Thus the maximum number of variables retained in (3) is the maximum number of z_j's equal to 1 in a solution of (2). Since (2) is NP-hard, we have,

Theorem 2 *Finding a maximum embedded RHS is NP-hard.*

References

[1] Aspvall, B., Recognizing disguised NR(1) instance of the satisfiability problem, *Journal of Algorithms* **1** (1980) 97-103.

[2] Balas, E., and C. S. Yu, Finding a maximum clique in an arbitrary graph, *SIAM Journal on Computing* **15** (1986) 1054-1068.

[3] Carraghan, R., and P. Pardalos, An exact algorithm for the maximum clique problem, technical report, Computer Science Dept., Pennsylvania State University, University Park, PA 16802 USA, 1990.

[4] Chandru, V., C. R. Coullard, P. L. Hammer, M. Montañez, and X. Sun, On renamable Horn and generalized Horn functions, *Annals of Mathematics and AI* **1** (1990) 33-48.

[5] Dowling, W. F., and J. H. Gallier, Linear-time algorithms for testing the satisfiability of propositional Horn formulae, *Journal of Logic Programming* **1** (1984) 267-284.

[6] Feo, T., M. Resende and S. Smith, A greedy randomized adaptive search procedure for maximum independent set, technical report, Mechanical Engineering Department, University of Texas, Austin, TX 78712 USA, 1989.

[7] Kopf, R., and G. Ruhe, A computational study of the weighted independent set problem for general graphs, *Foundations of Control Engineering* **12** (1987) 167-180.

[8] Pardalos, P., and G. Rodgers, A branch and bound algorithm for the maximum clique problem, technical report, Computer Science Dept., Pennsylvania State University, University Park, PA 16802 USA, 1990.

[9] Xue, J., Fast algorithms for vertex packing and related problems, Ph.D. thesis, Graduate School of Industrial Administration, Carnegie Mellon University, Pittsburgh, PA 15213 USA, 1991.