Discrete Applied Mathematics, 59:23–31, 1995.

Recognizing Renamable Generalized Propositional Horn Formulas is NP-Complete

Thomas Eiter*, Pekka Kilpeläinen, Heikki Mannila

Department of Computer Science, University of Helsinki†

Abstract

Yamasaki and Doshita have defined an extension of the class of propositional Horn formulas; later, Gallo and Scutellà generalized this class to a hierarchy $\Gamma_0 \subseteq \Gamma_1 \subseteq \ldots \subseteq \Gamma_k \subseteq \ldots$, where Γ_0 is the set of Horn formulas and Γ_1 is the class of Yamasaki and Doshita. For any fixed k, the propositional formulas in Γ_k can be recognized in polynomial time, and the satisfiability problem for Γ_k formulas can be solved in polynomial time. A possible way of extending these tractable subclasses of the satisfiability problem is to consider renamings: a renaming of a formula is obtained by replacing for some variables all their positive occurrences by negative occurrences and vice versa. The class of renamings of Horn formulas can be recognized in linear time. Chandru et al. have posed the problem of deciding whether the renamings of Γ_1 formulas can be recognized efficiently. We show that this is probably not the case by proving the NP-completeness of recognizing the renamings of Γ_k formulas for any $k \geq 1$.

Key words: generalized Horn clauses, renamable Horn clauses, satisfiability problem, NP-completeness

*On leave from the Christian Doppler Laboratory for Expert Systems, Technical University of Vienna, Austria.

†Mailing address: Heikki Mannila, Department of Computer Science, P.O.Box 26, (Teollisuuskatu 23), SF-00014 University of Helsinki, Finland.
Internet e-mail: `eiter@vexpert.dbai.tuwien.ac.at`, `kilpelai@cs.Helsinki.FI`, `mannila@cs.Helsinki.FI`

1 Introduction

Checking satisfiability of a set of propositional clauses is a well-known NP-complete problem (SAT). While the general problem of determining satisfiability is hard, it is useful to look for large subclasses for which there exists a polynomial time algorithm for testing satisfiability. Among such subclasses of SAT the class of Horn formulas [5, 11] is possibly the most important. Recently several extensions of the class of Horn formulas with polynomial time satisfiability tests have been presented [12, 6, 3].

A *renaming* of a propositional formula C is obtained by choosing a subset of variables and replacing each positive occurrence of such a variable by the corresponding negative literal and vice versa. The usefulness of this operation stems from the obvious fact that the renamed formula is satisfiable if and only if the original formula is, and that a satisfying truth assignment for the renamed formula gives a satisfying truth assignment for the original formula. Thus if we can find a renaming that maps a formula C to an instance in a subclass which is solvable in polynomial time, we can solve the satisfiability problem for C in polynomial time.

Whether a formula is a renamed Horn formula can be tested in linear time [1, 9, 2].

Yamasaki and Doshita have defined an extension of the class of propositional Horn formulas; later, Gallo and Scutellà generalized this class to a hierarchy $\Gamma_0 \subseteq \Gamma_1 \subseteq \ldots \subseteq \Gamma_k \subseteq \ldots$, where Γ_0 is the set of Horn formulas and Γ_1 is the class of Yamasaki and Doshita. This hierarchy has several nice properties: for any fixed k, the propositional formulas in Γ_k can be recognized in polynomial time, and the satisfiability problem for Γ_k formulas can be solved in polynomial time. Furthermore, any instance of SAT occurs on some level of the hierarchy.

If the renamed instances of, say, Γ_1 were recognizable in polynomial time, one would get a fairly large subclass of SAT with a useful polynomial time satisfiability algorithm. The renamed Horn formulas can be recognized efficiently, but it is not clear what is the situation for the next level of the hierarchy. Chandru et al. [2] pose this question for the class of [12] as an open problem.

The results of this paper are a negative answer to this question. We show that recognizing renamed instances of the classes of generalized Horn formulas in [12, 6] is NP-complete, thus settling the problem in [2].

The rest of this paper is organized as follows. In Section 2 we define the classes of Horn formulas and generalized Horn formulas, and introduce the concept of renaming formally. Section 3 gives a reduction from the exact hitting set problem to the problem of recognizing renamed Γ_1 formulas. Section 4 extends this result to the classes Γ_k for $k \geq 2$. Section 5 is a short conclusion.

2 Preliminaries and previous results

Variables are denoted by lower-case letters from the end of the alphabet. A literal is a propositional variable (x) or its negation ($\overline{x}$). For a literal l we use $v(l)$ to denote the variable of l, and $op(l)$ to denote the opposite of literal l (i.e., $op(x) = \overline{x}$, $op(\overline{x}) = x$). A literal l is *positive* if $l = v(l)$ and *negative* otherwise. A *clause* is a set $\{l_1, \ldots, l_n\}$ of distinct literals. A *formula* is a set $\mathcal{C} = \{C_1, \ldots, C_n\}$ of clauses on a finite set of propositional variables. A formula is satisfiable, if there exists an assignment of truth values to the variables such that the formula evaluates to true; a clause $\{l_1, \ldots, l_n\}$ is regarded as the disjunction $l_1 \vee \cdots \vee l_n$, and a formula is regarded as the conjunction of its clauses. A clause C is *Horn* if it contains at most one positive literal, and a set of clauses is Horn if all its clauses are Horn.

A *renaming* of the variables is a mapping r of the variables into the literals where $r(x) \in \{x, \overline{x}\}$; r is identified with $\{x : r(x) = \overline{x}\}$. We say that x is renamed in r if $x \in r$.

For every renaming r and a clause C, the renamed clause Cr is obtained by replacing each l in C by $op(l)$ if $v(l) \in r$; i.e.,

$$Cr = (C - \{l : v(l) \in r\}) \cup \{op(l) : v(l) \in r, l \in C\}.$$

For a formula $\mathcal{C}$, the renamed formula is defined by $\mathcal{C}r = \{Cr : C \in \mathcal{C}\}$.

Gallo and Scutellà [6] introduced a hierarchy of formulas $\Gamma_0, \Gamma_1, \ldots, \Gamma_k, \ldots$ which generalizes Horn formulas. For any clause set $\mathcal{C}$ and clause C define the simplified formulas $\mathcal{C}_C$ and $\mathcal{C}\Theta C$ by

$$\mathcal{C}_C = \{C' \in \mathcal{C} : C \not\subseteq C'\} \text{ and } \mathcal{C}\Theta C = \{C' - C : C' \in \mathcal{C}\}.$$

Definition 2.1 *The class of clause sets Γ_k, $k \geq 0$, is defined by*

$k = 0$: *$\mathcal{C} \in \Gamma_0$ if and only if $\mathcal{C}$ is Horn;*

$k > 0$: *$\mathcal{C} \in \Gamma_k$ if and only if $\mathcal{C}$ is an empty set or a singleton consisting of the empty clause, or there exists a positive literal l such that (i) $\mathcal{C}_{\{l\}} \in \Gamma_{k-1}$, and (ii) $\mathcal{C}\Theta\{l\} \in \Gamma_k$.*

$\square$

Note that $\Gamma_0 \subseteq \Gamma_1 \subseteq \cdots \subseteq \Gamma_k \subseteq \cdots$; it is not hard to see that $\Gamma_k \subset \Gamma_{k+1}$ for all $k \geq 0$, and that $\bigcup_{k=0}^{\infty} \Gamma_k$ contains all propositional formulas. The class Γ_1 is coincides with the class of generalized Horn functions introduced by Yamasaki and Doshita [12], which is defined as follows [6].

Definition 2.2 *A set* $\mathcal{C} = \{C_1, \ldots, C_n\}$ *of clauses is* generalized Horn *if and only if there exist sets of positive literals* $P_1, \ldots, P_n$ *(a "chain") such that*

(i) $P_1 \subseteq P_2 \subseteq \cdots \subseteq P_n$*,*

(ii) $P_i \subseteq C_i$ *for all* $1 \leq i \leq n$*, and*

(iii) $\{C_i - P_i : 1 \leq i \leq n\}$ *is Horn.* $\square$

Deciding whether $\mathcal{C} \in \Gamma_k$ can be done in time $O(n^* n^k)$ where n^* is the input size and n the number of propositional letters [6], and for $k = 1$ also in linear time [2]. Moreover, the satisfiability of $\mathcal{C} \in \Gamma_k$ can be checked in $O(n^* n^k)$ time [6].

Call a set of clauses $\mathcal{C}$ *renamable* Γ_k if there exists a renaming r such that $\mathcal{C}r \in \Gamma_k$.

Definition 2.3 *For* $k \geq 0$*,* $R\Gamma_k$ *denotes the class of all renamable* Γ_k *formulas.* $\square$

Note that $\mathcal{C} \in R\Gamma_0$ if and only if C is renamable Horn, and $\mathcal{C} \in R\Gamma_1$ if and only if $\mathcal{C}$ is renamable generalized Horn.

Finding out how many variables have to be removed from a formula to reach a renamable Horn formula is known to be NP-complete [10, 4]. While it is known that deciding $\mathcal{C} \in R\Gamma_0$ can be done in linear time [1], the complexity of deciding $\mathcal{C} \in R\Gamma_k$ for $k > 0$ was unknown yet; the case $k = 1$ has been posed as an open problem by Chandru et al. [2].

3 The complexity of deciding $\mathcal{C} \in R\Gamma_1$

We describe in this section a polynomial time transformation from the EXACT HITTING SET problem [8] to deciding the membership of a collection of clauses in $R\Gamma_1$, and prove the correctness of this reduction. Since EXACT

HITTING SET is NP-complete [8], the reduction shows the NP-hardness of recognizing the formulas in $R\Gamma_1$.

The EXACT HITTING SET problem is as follows. Given a family $\mathcal{F} = \{F_1, \ldots, F_m\}$ of subsets of a finite set of vertices $V = \{v_1, \ldots, v_n\}$, decide whether there exists a subset H of V such that $|H \cap F_i| = 1$ for all i with $1 \leq i \leq m$. This problem remains NP-complete in the case where each F_i, $1 \leq i \leq m$, consists of two or three vertices. (This follows by an immediate reduction from the ONE-IN-THREE-SAT problem [7]).

The transformation from EXACT HITTING SET to the problem of recognizing formulas of $R\Gamma_1$ is as follows. Let $\mathcal{F} = \{F_1, \ldots, F_m\}$ be a family of subsets on the set of vertices $V = \{v_1, \ldots, v_n\}$, where $|F_i| \in \{2, 3\}$ for all $1 \leq i \leq m$. Let $X_i = \{x_{i,j} : 1 \leq j \leq n\}$, where $1 \leq i \leq m$, and $S = \{s_1, s_2, s_3\}$ be disjoint sets of distinct propositional variables. We will also use some additional variables, which are all assumed to be distinct.

We construct a set of clauses $\mathcal{C}$ which consists of three disjoint sets of clauses: the "Horn part" $\mathcal{H}$, the "separating group" $\mathcal{S}$, and the "chain part" $\mathcal{P}$. The construction is such that the variables $x_{i,j}$ renamed by a renaming r such that $\mathcal{C}r \in \Gamma_1$ will correspond to the vertices v_j of an exact hitting set of $\mathcal{F}$ and vice versa.

The Horn part $\mathcal{H}$ of $\mathcal{C}$ contains the following clauses:

- The clauses $\{x_{i,j}, \overline{x}_{k,j}\}$ and $\{\overline{x}_{i,j}, x_{k,j}\}$ for each i, j and k where $1 \leq i < k \leq m$, $1 \leq j \leq n$, and $v_j \in F_i \cap F_k$.

  (These clauses will serve to assure that renaming $x_{i,j}$ requires also renaming of $x_{k,j}$.)

- For each three-element set $F_i = \{v_{i_1}, v_{i_2}, v_{i_3}\}$ the formula $\mathcal{H}$ contains the clauses

  $$\{\overline{x}_{i,i_1}, \overline{x}_{i,i_2}\}, \{\overline{x}_{i,i_1}, \overline{x}_{i,i_3}\}, \{\overline{x}_{i,i_2}, \overline{x}_{i,i_3}\}.$$

  (These clauses will assure that at most one of the variables x_{i,i_1}, x_{i,i_2} and x_{i,i_3} will be renamable.)

- For each two-element set $F_i = \{v_{i_1}, v_{i_2}\}$ the formula $\mathcal{H}$ contains the clauses

  $$\{\overline{x}_{i,i_1}, \overline{x}_{i,i_2}\}, \{\overline{z}_i, \overline{a}_i\}, \{\overline{z}_i, a_i\},$$

where z_i, a_i are new propositional variables.

(These clauses will assure that at most one of the variables x_{i,i_1}, x_{i,i_2} will be renamable and that z_i can not be renamed).

The separating group $\mathcal{S}$ consists of the following clauses:

$$\{s_1, s_2\}, \{s_1, s_3\}, \{s_2, s_3\}, \{\overline{s}_1, \overline{s}_2\}, \{\overline{s}_1, \overline{s}_3\}, \{\overline{s}_2, \overline{s}_3\}$$

The aim of these clauses is to assure that in each renaming r that makes the clause set $\mathcal{C}$ generalizable Horn, a suitable chain for $\mathcal{C}r$ can not remove literals from the clauses of the Horn part, while it must remove literals from the clauses of the chain part.

The chain part $\mathcal{P}$ consists of the following sets of clauses $\mathcal{C}_i$, which correspond to the set of nodes F_i. Let $A_0 = \{s_1, s_2, s_3\}$, and define for $1 \leq i \leq m$

$$\mathcal{C}_i = \{D \cup A_{i-1} : D \subseteq B_i, |D| = 2\}, \quad A_i = \bigcup_{C \in \mathcal{C}_i} C,$$

where $B_i = \{x_{i,i_1}, x_{i,i_2}, x_{i,i_3}\}$ if $F_i = \{v_{i_1}, v_{i_2}, v_{i_3}\}$ and $B_i = \{x_{i,i_1}, x_{i,i_2}, z_i\}$ if $F_i = \{v_{i_1}, v_{i_2}\}$. Note that $A_i = A_{i-1} \cup B_i$.

The clauses of $\mathcal{C}_i$ serve for the following purpose: $\mathcal{C}_i = \{C_{i,1}, C_{i,2}, C_{i,3}\} \notin \Gamma_1$ because the intersection of these three clauses with B_i yields sets of positive variables $\{u, v\}, \{u, w\}, \{v, w\}$. If $\mathcal{C}r \in \Gamma_1$ then r renames at least one of u, v, w (say u). The clauses in the Horn part will assure that at most one of these variables can be renamed; hence, exactly one variable (u) must be renamed; v, w can then be removed by a chain.

This finishes the construction.

Consider the following example of the construction. Let $\mathcal{F} = \{F_1, F_2, F_3\}$, $F_1 = \{v_1, v_2, v_3\}$, $F_2 = \{v_1, v_3, v_4\}$ $F_3 = \{v_2, v_5\}$, where $V = \{v_1, \ldots, v_5\}$. The clause set $\mathcal{C}$ is as follows.

$\mathcal{H}$:

$$\begin{aligned}
&\{x_{1,1}, \overline{x}_{2,1}\}, \{\overline{x}_{1,1}, x_{2,1}\}, \\
&\{x_{1,2}, \overline{x}_{3,2}\}, \{\overline{x}_{1,2}, x_{3,2}\}, \\
&\{x_{1,3}, \overline{x}_{2,3}\}, \{\overline{x}_{1,3}, x_{2,3}\}
\end{aligned}$$

$$\begin{aligned}
&\{\overline{x}_{1,1}, \overline{x}_{1,2}\}, \{\overline{x}_{1,1}, \overline{x}_{1,3}\}, \{\overline{x}_{1,2}, \overline{x}_{1,3}\}, \\
&\{\overline{x}_{2,1}, \overline{x}_{2,3}\}, \{\overline{x}_{2,1}, \overline{x}_{2,4}\}, \{\overline{x}_{2,3}, \overline{x}_{2,4}\}, \\
&\{\overline{x}_{3,2}, \overline{x}_{3,5}\}, \{\overline{z}_3, \overline{a}_3\}, \{\overline{z}_3, a_3\}.
\end{aligned}$$

$\mathcal{S}$:

$$\{s_1, s_2\}, \{s_1, s_3\}, \{s_2, s_3\}, \{\overline{s}_1, \overline{s}_2\}, \{\overline{s}_1, \overline{s}_3\}, \{\overline{s}_2, \overline{s}_3\}.$$

$\mathcal{P}$:

$$\{x_{1,1}, x_{1,2}, s_1, s_2, s_3\}, \{x_{1,1}, x_{1,3}, s_1, s_2, s_3\}, \{x_{1,2}, x_{1,3}, s_1, s_2, s_3\}$$

$$\begin{array}{l} \{x_{2,1}, x_{2,3}, s_1, s_2, s_3, x_{1,1}, x_{1,2}, x_{1,3}\} \\ \{x_{2,1}, x_{2,4}, s_1, s_2, s_3, x_{1,1}, x_{1,2}, x_{1,3}\} \\ \{x_{2,3}, x_{2,4}, s_1, s_2, s_3, x_{1,1}, x_{1,2}, x_{1,3}\} \end{array}$$

$$\begin{array}{l} \{x_{3,2}, x_{3,5}, s_1, s_2, s_3, x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, x_{2,3}, x_{2,4}\} \\ \{x_{3,2}, z_3, s_1, s_2, s_3, x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, x_{2,3}, x_{2,4}\} \\ \{x_{3,5}, z_3, s_1, s_2, s_3, x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, x_{2,3}, x_{2,4}\}. \end{array}$$

We now prove the correctness of the reduction. We first observe an important property of the "separating group" $\mathcal{S} = \{\{s_1, s_2\}, \{s_1, s_3\}, \{s_2, s_3\}, \{\overline{s}_1, \overline{s}_2\}, \{\overline{s}_1, \overline{s}_3\}, \{\overline{s}_2, \overline{s}_3\}\}$.

Lemma 3.1 $\mathcal{S} \in R\Gamma_1 - R\Gamma_0$.

Proof. It is easily checked that $\mathcal{S}r$ contains a clause with two positive literals for any renaming r. Hence $\mathcal{S} \notin R\Gamma_0$.

Consider then the renaming $r = \{s_1\}$ and denote $\mathcal{S}r = \{C_1, \ldots, C_6\}$. The clauses can be ordered in such a way that $C_1 = \{\overline{s}_1, s_2\}$, $C_2 = \{\overline{s}_1, s_3\}$, $C_3 = \{s_1, \overline{s}_2\}$, $C_4 = \{s_1, \overline{s}_3\}$, $C_5 = \{\overline{s}_2, \overline{s}_3\}$, and $C_6 = \{s_2, s_3\}$. Now choose $P_i = \emptyset$, for $1 \le i \le 5$, and let $P_6 = \{s_2, s_3\}$. Then $\{C_i - P_i : 1 \le i \le 6\}$ is Horn, and hence $\mathcal{S} \in R\Gamma_1$. $\Box$

Lemma 3.2 *Let $\mathcal{F} = \{F_1, \ldots, F_m\}$ be a family of subsets on $V = \{v_1, \ldots, v_n\}$, where $|F_i| \in \{2, 3\}$ for all $1 \le i \le m$. The formula $\mathcal{C}$ constructed from the family $\mathcal{F}$ satisfies $\mathcal{C} \in R\Gamma_1$ if and only if $\mathcal{F}$ has an exact hitting set.*

Proof. *If:* Assume that $H \subseteq \{v_1, \ldots, v_n\}$ is an exact hitting set of $\mathcal{F}$. Define the renaming

$$r = \{x_{i,j} : v_j \in F_i \cap H, 1 \le i \le m\} \cup \{s_1\}.$$

We can then write $\mathcal{C}r = \{R_1, \ldots, R_t, \ldots, R_l\}$, where

$$\begin{array}{rcl} \{R_1, \ldots, R_{t-1}\} & = & [\mathcal{H} \cup (\mathcal{S} - \{\{s_2, s_3\}\})]r, \\ R_t & = & \{s_2, s_3\}r, \text{ and} \\ \{R_{t+3i-2}, R_{t+3i-1}, R_{t+3i}\} & = & \mathcal{C}_i r \text{ for } 1 \le i \le m. \end{array}$$

Moreover, the clauses can be ordered in such a way that

$$R_{t+3i} \cap \{\overline{x} : x \in X_i\} = \emptyset$$

for all $1 \leq i \leq m$.

Define the sets $P_1, \ldots, P_l$ as follows.

$$\begin{aligned} P_i &= \emptyset \text{ for } 1 \leq i < t, \\ P_t &= \{s_2, s_3\}, \\ P_{t+3i-2} &= P_{t+3i-1} = P_{t+3i-3} \text{ for } 1 \leq i \leq m, \text{ and} \\ P_{t+3i} &= P_{t+3i-3} \cup (R_{t+3i} \cap (X_i \cup \{z_i\})) \text{ for } 1 \leq i \leq m. \end{aligned}$$

Since H is an exact hitting set, it is not hard to verify from the definition of r that R_i is Horn for $1 \leq i < t$. Furthermore, it is straightforward to check that $R_j - P_j$ contains no positive literal if $j = t + 3i$ for some $0 \leq i \leq m$ and that it contains exactly one positive literal if $j = t+3i-2$ or $j = t+3i-1$ for some $1 \leq i \leq m$. Hence $\{R_i - P_i : 1 \leq i \leq l\}$ is Horn. Since $P_1 \subseteq \cdots \subseteq P_l$ and $P_i \subseteq R_i$ for $1 \leq i \leq l$, this entails $\mathcal{C} \in R\Gamma_1$.

Only if: Assume that r is such a renaming that $\mathcal{C}r = \{R_1, \ldots, R_l\} \in \Gamma_1$, and that $P_1 \subseteq \cdots \subseteq P_l$ is a suitable chain such that $P_i \subseteq R_i$ for $1 \leq i \leq l$ and $\{R_i - P_i : 1 \leq i \leq l\}$ is Horn.

We observe that regardless of r, we must for some i have $R_i \subseteq \{s_1, s_2, s_3\}$ and $|R_i| = 2$ by Lemma 3.1, and hence $P_i \subseteq \{s_1, s_2, s_3\}$. Let then C be any clause from the Horn part $\mathcal{H}$. Since $Cr \cap \{s_1, s_2, s_3\} = \emptyset$, it follows that $Cr = R_j$ and $P_j = \emptyset$ for some j.

We next show that r must rename at least one of the propositional variables $\{x_{i,i_1}, x_{i,i_2}, x_{i,i_3}\}$ corresponding to any set $F_i = \{v_{i_1}, v_{i_2}, v_{i_3}\}$ with three elements, and that r must rename at least one of the variables $\{x_{i,i_1}, x_{i,i_2}\}$ corresponding to any set $F_i = \{v_{i_1}, v_{i_2}\}$ with two elements.

Assume to the contrary that $\{x_{i,i_1}, x_{i,i_2}, x_{i,i_3}\} \cap r = \emptyset$. Since $\mathcal{C}r \in \Gamma_1$, we thus get that $P_{i_j} \cap \{x_{i,i_1}, x_{i,i_2}, x_{i,i_3}\} \neq \emptyset$, where $C_{i,j}r = R_{i_j}$, $1 \leq j \leq 3$. Without loss of generality assume $i_1 < i_2 < i_3$ and $x_{i,i_1} \in P_{i_1}$. Since $P_{i_1} \subseteq P_{i_2} \subseteq P_{i_3}$, this implies $x_{i,i_1} \in P_{i_2} \cap P_{i_3} \subseteq C_{i,2} \cap C_{i,3}$. But this is a contradiction since x_{i,i_1} is not contained in both $C_{i,2}$ and $C_{i,3}$. Consequently, $|\{x_{i,i_1}, x_{i,i_2}, x_{i,i_3}\} \cap r| \geq 1$ holds. The argument for $|\{x_{i,i_1}, x_{i,i_2}\} \cap r| \geq 1$ is analogous.

As a consequence, $H = \{v_j : \exists x_{i,j} \in r\}$ is a hitting set of $\mathcal{F} = \{F_1, \ldots, F_m\}$ by the construction of $\mathcal{P}$.

On the other hand, since $\mathcal{H}r$ is Horn, we can derive from the construction of $\mathcal{H}$ that *(i)* $x_{i,j} \in r$ if and only if $x_{k,j} \in r$, for all $1 \leq i, k \leq m$, $1 \leq j \leq n$, where $v_j \in F_i \cap F_k$, and that *(ii)* $|\{x_{i,i_1}, x_{i,i_2}, x_{i,i_3}\} \cap r| \leq 1$ (resp. $|\{x_{i,i_1}, x_{i,i_2}\} \cap r| \leq 1$), and hence we have that $|\{x_{i,i_1}, x_{i,i_2}, x_{i,i_3}\} \cap r| = 1$ (resp. $|\{x_{i,i_1}, x_{i,i_2}\} \cap r| = 1$) must hold for all $1 \leq i \leq m$. Thus we conclude that $|H \cap F_i| = 1$ holds for $1 \leq i \leq m$, i.e., H is an exact hitting set of $\mathcal{F}$. $\square$

Theorem 3.3 *Let $\mathcal{C}$ be a set of clauses. Deciding whether $\mathcal{C} \in R\Gamma_1$, i.e., if $\mathcal{C}$ is renamable generalizable Horn, is NP-complete.*

Proof. This problem is clearly in NP since a guess for r such that $\mathcal{C}r \in \Gamma_1$ can be verified in polynomial time, e.g., by the linear time algorithm of Chandru et al. [2].

The NP-hardness follows from Lemma 3.2, since $\mathcal{C} = \mathcal{H} \cup \mathcal{S} \cup \mathcal{P}$ is clearly constructible in polynomial time. $\square$

4 NP-completeness of $\mathcal{C} \in R\Gamma_k$ for $k > 0$

In this section, we show that deciding whether $\mathcal{C} \in R\Gamma_k$ or not is NP-hard for every fixed $k \geq 1$. For this purpose, we need some intermediate results.

Lemma 4.1 *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be sets of clauses on disjoint sets of variables such that $\mathcal{C}_i \in \Gamma_{k_i} - \Gamma_{k_i-1}$ for $i = 1, 2$ and $k_1, k_2 \geq 0$ (we denote $\Gamma_{-1} = \emptyset$). Then $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2 \in \Gamma_{k_1+k_2} - \Gamma_{k_1+k_2-1}$.*

Proof. The proof is by induction on a pairing $\phi(k_1, k_2)$. The claim obviously holds if $k_1, k_2 = 0$. Assume it holds for all $\phi(k_1', k_2') < \phi(k_1, k_2)$ and consider the case $\phi(k_1, k_2)$. We show by induction on the number q of positive literals in $\mathcal{C}$ that the claim holds. If $q = 0$, then $k_1, k_2 = 0$ and the claim holds. If $q > 0$, then without loss of generality assume $k_1 > 0$. Hence there exists a positive literal l such that $\mathcal{C}_{1\{l\}} \in \Gamma_{k_1-1}$ and $\mathcal{C}_1\Theta\{l\} \in \Gamma_{k_1}$. Note that $\mathcal{C}_{1\{l\}} \in \Gamma_a - \Gamma_{a-1}$ for some $a < k_1$ and that $\mathcal{C}_1\Theta\{l\} \in \Gamma_b - \Gamma_{b-1}$, for some $b \leq k_1$. By the induction hypothesis on $\phi(k_1', k_2')$, we have that $\mathcal{C}_{\{l\}} = \mathcal{C}_{1\{l\}} \cup \mathcal{C}_2 \in \Gamma_{a+k_2} - \Gamma_{a+k_2-1}$, and together with the induction hypothesis on $q' < q$ that $\mathcal{C}\Theta\{l\} = (\mathcal{C}_1\Theta\{l\}) \cup \mathcal{C}_2 \in \Gamma_{b+k_2} - \Gamma_{b+k_2-1}$. It follows $\mathcal{C} \in \Gamma_{k_1+k_2}$. Assume $\mathcal{C} \in \Gamma_{k_1+k_2-1}$. This implies that $k_1 + k_2 > 1$, since $\mathcal{C}_1 \subseteq \mathcal{C}$ and $k_1 > 0$. Without loss of generality, for a positive literal l from $\mathcal{C}_1$, we get $\mathcal{C}_{\{l\}} = \mathcal{C}_{1\{l\}} \cup \mathcal{C}_2 \in \Gamma_{k_1+k_2-2}$ and $\mathcal{C}\Theta\{l\} = \mathcal{C}_1\Theta\{l\} \cup \mathcal{C}_2 \in \Gamma_{k_1+k_2-1}$. Using the hypothesis on $q' < q$, it can be readily shown that $\mathcal{C}_{1\{l\}} \in \Gamma_a - \Gamma_{a-1}$ for some

$a < k_1$. By the hypothesis on $\phi(k_1', k_2')$, we get $a + k_2 \leq k_1 + k_2 - 2$, hence $a \leq k_1 - 2$. Similarly, we get $\mathcal{C}_1\Theta\{l\} \in \Gamma_b - \Gamma_{b-1}$ for some $b \leq k_1$, and by the hypothesis on $q' < q$, we get $b+k_2 \leq k_1+k_2-1$ and hence $b \leq k_1-1$. But this entails by the definition of the hierarchy the contradiction that $\mathcal{C}_1 \in \Gamma_{k_1-1}$. Thus $\mathcal{C} \in \Gamma_{k_1+k_2} - \Gamma_{k_1+k_2-1}$; hence the hypothesis on q holds, which means that the hypothesis on $\phi(k_1, k_2)$ also holds. Thus the lemma follows. $\square$

Corollary 4.2 *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be sets of clauses on disjoint sets of variables such that $\mathcal{C}_i \in R\Gamma_{k_i} - R\Gamma_{k_i-1}$ for $i = 1, 2$, where $k_1, k_2 \geq 0$. Then $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2 \in R\Gamma_{k_1+k_2} - R\Gamma_{k_1+k_2-1}$.* $\square$

We obtain the following result.

Theorem 4.3 *Deciding whether a set of clauses $\mathcal{C}$ is in the class $R\Gamma_k$, for a constant $k > 0$, is NP-complete.*

Proof. Membership of the problem in NP holds since a guess for a renaming r such that $\mathcal{C}r \in \Gamma_k$ can be verified with an algorithm of Gallo and Scutellà in $O(n^*n^k)$ time [6].

The NP-hardness part is shown as follows. Take the clause set $\mathcal{C}$ constructed in the proof of NP-hardness for the case $k = 1$ from above, and add $k-1$ copies $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{k-1}$ of the separating group $\mathcal{S}$ on completely different variable sets to $\mathcal{C}$, and let $\mathcal{C}'$ denote the resulting clause set. Clearly, $\mathcal{C}' \notin R\Gamma_0$. Hence by Lemma 3.1 and repeated application of Corollary 4.2, $\mathcal{C}' \in R\Gamma_k$ if and only if $\mathcal{C} \in R\Gamma_1$, from which the result follows. $\square$

5 Conclusions

We have shown that recognizing the formulas in the class $R\Gamma_k$ is NP-complete for all k with $k \geq 1$. This can be contrasted with the linearity of recognizing renamed Horn formulas. Hence the interesting hierarchy of Γ_k formulas cannot be extended by renamings and still retaining the polynomial time identifiability of the resulting class.

It seems that the techniques of Section 4 can be extended to show that given a formula $\mathcal{C}$, it is difficult even to approximate the smallest k such that $\mathcal{C} \in R\Gamma_k$.

An interesting open problem is characterizing the classes of formulas whose renamed versions can be recognized in polynomial time.

References

[1] B. Aspvall. Recognizing Disguised NR(1) Instance of the Satisfiability Problem. *Journal of Algorithms*, 1:97–103, 1980.

[2] V. Chandru, C. Coullard, P. Hammer, M. Montañez, and X. Sun. On Renamable Horn and Generalized Horn Functions. *Annals of Mathematics and Artificial Intelligence*, 1:33–48, 1990.

[3] V. Chandru and J. Hooker. Extended Horn Sets in Propositional Logic. *Journal of the ACM*, 38:205–221, 1991.

[4] V. Chandru and J. Hooker. Detecting Embedded Horn Structure in Propositional Logic. *Information Processing Letters*, 42:109–111, 1992.

[5] W. Dowling and J. H. Gallier. Linear-Time Algorithms for Testing the Satisfiability of Propositional Horn Theories. *Journal of Logic Programming*, 3:267–284, 1984.

[6] G. Gallo and M. Scutellà. Polynomially Solvable Satisfiability Problems. *Information Processing Letters*, 29:221–227, 1988.

[7] M. Garey and D. S. Johnson. *Computers and Intractability – A Guide to the Theory of NP-Completeness*. W. H. Freeman, New York, 1979.

[8] R. Karp. Reducibility Among Combinatorial Problems. In R. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103, New York, 1972. Plenum Press.

[9] H. Lewis. Renaming a Set of Clauses as a Horn Set. *Journal of the ACM*, 25:134–135, 1978.

[10] H. Mannila. *Instance Complexity for Sorting and NP-Complete Problems*. PhD thesis, University of Helsinki, Helsinki, Finland, March 1985.

[11] M. Minoux. LTUR: A Simplified Linear Time Unit Resolution for Horn Formulae and Computer Implementation. *Information Processing Letters*, 29:1–12, 1988.

[12] S. Yamasaki and S. Doshita. The Satisfiability Problem for a Class Consisting of Horn Sentences and Some Non-Horn Sentences in Propositional Logic. *Information and Control*, 59(1–3):1–12, 1983.