On the Difference of Horn Theories

(Extended Abstract)

Thomas Eiter[1], Toshihide Ibaraki[2], and Kazuhisa Makino[3]

[1] Institut und Ludwig Wittgenstein Labor für Informationssysteme, Technische Universität Wien, Treitlstraße 3, A-1040 Wien, Austria. (eiter@kr.tuwien.ac.at)
[2] Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606, Japan. (ibaraki@kuamp.kyoto-u.ac.jp)
[3] Department of Systems and Human Science, Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka, 560, Japan. (makino@sys.es.osaka-u.ac.jp)

Abstract. In this paper, we consider computing the difference between two Horn theories. This problem may arise, for example, if we take care of a theory change in a knowledge base. In general, the difference of Horn theories is not Horn. Therefore, we consider Horn approximations of the difference in terms of Horn cores (i.e., weakest Horn theories included in the difference), which have been proposed and analyzed extensively in the literature. We study the problem under the familiar representation of Horn theories by Horn CNFs, as well as under the recently proposed model-based representation in terms of the characteristic models. For all problems and representations, polynomial time algorithms or proofs of intractability for the propositional case are provided.

Keywords: computational issues in AI, knowledge compilation, Horn approximation

1 Introduction

Among the basic operations for combining logical theories are the Boolean operations, i.e., conjunction $\wedge$, disjunction $\vee$, and complement $\neg$. We consider another basic operation, the difference $\setminus$, which can represent the complement $\neg$, as for a theory (i.e., set of models) $\Sigma \subseteq \{0,1\}^n$, the complement $\neg\Sigma$ is given by the difference of $\{0,1\}^n$ and Σ (i.e., $\{0,1\}^n \setminus \Sigma$). In principle, these operations are easily accomplished, since the language of propositional logic provides the boolean connectives $\wedge$, $\vee$, and $\neg$, which can be applied to combine formulas representing theories. However, if they are from a particular fragment, it is in general desired to know whether their combination also belongs to this fragment, and to find a suitable representation for it.

In this paper, we study the problem of computing the Boolean difference between two Horn theories Σ_1 and Σ_2, i.e., $\Sigma = \Sigma_1 \setminus \Sigma_2$. In general, the resulting theory Σ is not Horn. Therefore, we consider approximating Σ by a Horn theory, in order to maintain the desired closure property. Different such approximations are possible, among which Horn cores [11, 12] are quite natural.

A Horn theory $\Pi \subseteq \{0,1\}^n$ is a *Horn core* (or a *Horn greatest lower bound*) of a theory $\Sigma \subseteq \{0,1\}^n$, if $\Pi \subseteq \Sigma$, i.e., Π logically implies Σ, and there is no weaker Π of this property, i.e., no Horn Π' exists such that $\Pi \subset \Pi' \subseteq \Sigma$. Observe that, in general, a theory Σ has more than one Horn core; e.g., $\Sigma = \{(110), (101)\}$ has two Horn

Table 1. Complexity of problems on the Horn difference $\Sigma_1 \setminus \Sigma_2$ (Σ_1, Σ_2 and Π are Horn)

	representation	
Problem	*formula-based (Horn CNF)*	*model-based (characteristic set)*
Is $\Sigma_1 \setminus \Sigma_2$ Horn?	P	P
Compute $\Sigma_1 \setminus \Sigma_2$, if it is Horn	P	P
Is Π a Horn core of $\Sigma_1 \setminus \Sigma_2$?	P	P
Compute one Horn core of $\Sigma_1 \setminus \Sigma_2$	P	P
Compute all Horn cores of $\Sigma_1 \setminus \Sigma_2$	polynomial delay	not polynomial total time unless $\mathrm{P} = \mathrm{NP}$
Is Π the Horn envelope of $\Sigma_1 \setminus \Sigma_2$?	co-NP-complete	P
Compute the Horn envelope of $\Sigma_1 \setminus \Sigma_2$	not polynomial total time unless $\mathrm{P} = \mathrm{NP}$	P

cores $\Sigma_1 = \{(110)\}$ and $\Sigma_2 = \{(101)\}$, respectively. Approximating a propositional logic theory by Horn theories (or Horn cores) is used for *knowledge compilation* in [11]. Because of its theoretical and practical importance, semantical and computational issues on Horn cores have been studied extensively, and a number of results have been obtained, cf. [11, 2, 12, 7, 1].

The main contributions of the present paper can be summarized as follows.

• We present characterizations of the Horn cores of a Horn difference $\Sigma_1 \setminus \Sigma_2$, which will form a basis of the algorithms discussed in this paper.
• We either present a polynomial time algorithm or prove intractability (unless P=NP) for each of the problems mentioned above. (For a concise overview, see Table 1.)
• Besides the familiar representation in terms of Horn CNFs, we also consider the model-based representation of Horn theories through their sets of characteristic models [10]. This alternative has also been studied repeatedly, since it offers advantages to formula-based representation in certain cases; see [13, 10, 4] for more details.

Our results on the complexity of these issues are summarized in Table 1, which gives a complete picture of the tractability/intractability frontier of these problems. The table also shows results on the *Horn envelope* [11, 12] of the difference, i.e., the (unique) least Horn theory E such that $\Sigma_1 \setminus \Sigma_2 \subseteq E$, which we do not discuss here. The interested reader is referred to [6].

The results on formula-based representation complement the results known from [2, 11, 1]. Observe that these works have focussed on theories which are represented by CNFs or restricted classes of CNFs, i.e., the problem input is a (possibly restricted) CNF. In contrast, we deal with non-CNF formulas. The tractability of computing some

Horn core of a Horn difference is a positive result, since this problem is intractable for a general theory. For the case in which a theory is represented by a CNF formula, [2] contains a polynomial time algorithm which consults an NP oracle, and shows that the problem is $\mathrm{P}^{\mathrm{NP}}[O(\log n)]$-hard. Most recently, in [1] another algorithm for computing a Horn core from a given CNF is represented, which is based on the classical Davis-Putnam procedure for the satisfiability problem. Observe, however, that these algorithms are not efficiently applicable to our problem, since they require a CNF formula for input; by rewriting the difference of Horn CNFs φ_1 and φ_2 to a CNF, the size of the formula might exponentially increase. Our result that the Horn cores of a Horn difference $\varphi_1 \setminus \varphi_2$ can be enumerated with polynomial delay parallels the analogous result if all models of a theory Σ are given for input [12].

Our results find an application in the area of theory change, which has grown into an important research area within AI during the last decade. Taking the difference between Horn theories is, for instance, meaningful in the following scenario. Assume that we have a Horn theory Σ, which is a description of all the possible worlds of a state of affairs; i.e., the real "world" amounts to one of the models in Σ, say w, but is unknown to us. Suppose that we now obtain the information that a formula φ is false in w. Since w is a categorical description of the world, either ψ or $\neg\psi$ is true in it (but not both), for any formula ψ. Hence, in this case, we conclude that $\neg\varphi$ is true in w, and thus all models v that satisfy φ can be discarded from Σ. This amounts to updating Σ to $\Sigma_{new} = \Sigma \setminus \Sigma'$, where Σ' is the set of models satisfying φ; if φ is Horn, this gives an instance of our problem. If Σ_{new} is not Horn, the Horn cores and the Horn envelope provide sound and complete approximations of the set Σ_{new} in terms of Horn theories. Our algorithms can be applied for computing Σ_{new} or its Horn approximations. Observe that computing a Horn core and the Horn envelope in the context of theory change was studied in [7], in which the theory $\varphi \circ \psi$, given by two Horn CNFs φ and ψ, is studied for some revision operators "$\circ$" (see [9]). However, this is only weakly related to ours, since taking the difference of φ and ψ was not considered, and moreover, the formula ψ is restricted to be a single Horn clause.

For space reasons, most proofs are omitted. They are given in [6].

2 Preliminaries

We assume a supply of propositional variables $x_1, x_2, \ldots, x_n$, where each x_i evaluates to either 1 (true) or 0 (false). Negated variables are denoted by $\overline{x}_i$. The x_i and $\overline{x}_i$ are called literals. A clause is a disjunction $c = \ell_1 \vee \cdots \vee \ell_k$ of literals, while a term is a conjunction $t = \ell_1 \wedge \cdots \wedge \ell_k$ of literals. By $P(c)$ and $N(c)$ (resp., $P(t)$ and $N(t)$) we denote the sets of variables occurring positively and negatively in c (resp., t); $\perp$ (resp., $\top$) denotes the empty clause (resp., empty term) representing falsity (resp., truth). A *conjunctive normal form* (CNF) (resp., *disjunctive normal form* (DNF)) is a conjunction of clauses $\varphi = \bigwedge_i c_i$ (resp., a disjunction of terms $\varphi = \bigvee_i t_i$).

A *model* is a vector $v \in \{0,1\}^n$, whose i-th component is denoted by v_i. The models $(0, 0, \ldots, 0)$ and $(1, 1, \ldots, 1)$ are denoted by $\mathbf{0}$ and $\mathbf{1}$, respectively. We use $v \leq w$ for the usual bitwise ordering of models, i.e., $v_i \leq w_i$ for all $i = 1, \ldots, n$, where $0 \leq 1$. A *theory* is any set $\Sigma \subseteq \{0,1\}^n$ of models. A model $v \in \Sigma$ is *minimal* in Σ, if no $w \in \Sigma$ exists such that $w < v$.

For any formula φ, let $T(\varphi) = \{v \in \{0,1\}^n \mid \varphi(v) = 1\}$ be the set of its models. A formula φ represents a theory Σ if $T(\varphi) = \Sigma$. If unambiguous, we do not distinguish a formula from the theory it represents. We write $\varphi \leq \psi$, if $T(\varphi) \subseteq T(\psi)$ holds. A non-tautological clause c (resp., noncontradictory term t) is an *implicate* (resp., *implicant*) of a theory Σ if $c(v) = 1$ for all $v \in \Sigma$, i.e., $\{0,1\}^n \supset T(c) \supseteq \Sigma$ (resp., $t(v) = 0$ for all $v \notin \Sigma$, i.e., $\emptyset \subset T(t) \subseteq \Sigma$); it is *prime*, if no proper subclause (resp., subterm) is an implicate (resp., implicant) of Σ.

A theory Σ is *Horn* if $\Sigma = Cl_\wedge(\Sigma)$ holds, where $Cl_\wedge(S)$ is the closure of $S \subseteq \{0,1\}^n$ under bitwise AND (i.e., intersection) of models v and w, denoted by $v \bigwedge w$. Observe that any Horn theory Σ has the least (unique smallest) model, which is given by $\bigwedge_{v\in\Sigma} v$. For a Horn theory Σ, a model $v \in \Sigma$ is called *characteristic* [10], if $v \notin Cl_\wedge(\Sigma \setminus \{v\})$ holds. The set of all characteristic models of Σ, the *characteristic set of* Σ, is denoted by $C^*(\Sigma)$. Note that every Horn theory Σ has the unique characteristic set $C^*(\Sigma)$. For example, the theory $\Sigma = \{(0101), (1001), (1000), (0001), (0000)\}$ is Horn, and has $C^*(\Sigma) = \{(0101), (1001), (1000)\}$.

A clause c is *Horn* (resp., *negative, positive*) if $|P(c)| \leq 1$ (resp., $|P(c)| = 0$, $|N(c)| = 0$). A CNF is *Horn* (resp., *negative, positive*) if it contains only Horn (resp., negative, positive) clauses. It is well-known that a theory Σ is Horn if and only if it is represented by some Horn CNF, and that all prime implicates of a Horn theory are Horn. A theory is *negative* (resp., *positive*), if it is represented by some negative (resp., positive) CNF. If φ represents a theory Σ, and ψ is Horn CNF representing a Horn core of Σ, then ψ is also called a *Horn core* of φ.

3 Horn Property of the Difference

3.1 Formula-based representation

Let $\varphi_1 = \bigwedge_{j=1}^{m_1} c_{1,j}$ and $\varphi_2 = \bigwedge_{j=1}^{m_2} c_{2,j}$ be Horn CNFs. Then

$$\varphi_1 \setminus \varphi_2 \equiv \bigvee_{j=1}^{m_2} (\varphi_1 \wedge \neg c_{2,j}) \; = \; \bigvee_{j=1}^{m_2} (\varphi_1 \wedge t_{2,j}), \tag{3.1}$$

where $t_{2,j}$ is the term equivalent to $\neg c_{2,j}$; e.g., if $c_{2,j} = (\overline{x}_1 \vee \overline{x}_2 \vee x_3)$, then $t_{2,j} = x_1 x_2 \overline{x}_3$. Note that each $\psi_j = \varphi_1 \wedge t_{2,j}$ is a Horn CNF. Hence $\varphi_1 \setminus \varphi_2$ can be represented by the disjunction of m_2 Horn theories ψ_j. Observe that in general, deciding whether a disjunction of Horn CNFs is Horn is co-NP-complete [5]. However, we sketch a proof that checking the Horn property of the difference of Horn theories is polynomial.

If $\varphi_2 \equiv \perp$, then

$$\varphi_1 \setminus \varphi_2 \equiv \varphi_1 \tag{3.2}$$

clearly holds, and hence the difference is Horn. Otherwise, we compute the least model v in φ_2, and separately consider the two cases: (1) $\varphi_1(v) = 0$ and (2) $\varphi_1(v) = 1$.

Case (1). Since φ_1 is Horn, there is a Horn clause $c^* \in \varphi_1$ such that $c^* \geq \varphi_1$ and $c^*(v) = 0$. For a model u, let $HC(u)$ denote the set of all maximal Horn clauses c such that $c(u) = 0$. E.g., if $u = (111000)$, then $HC(u) = \{(\overline{x}_1 \vee \overline{x}_2 \vee \overline{x}_3 \vee x_4), (\overline{x}_1 \vee \overline{x}_2 \vee \overline{x}_3 \vee x_5), (\overline{x}_1 \vee \overline{x}_2 \vee \overline{x}_3 \vee x_6)\}$. Now, some c in $HC(v)$ satisfies $c \geq c^*(\geq \varphi_1)$.

It can be shown that if $v \neq \mathbf{1}$, then

$$\varphi_1 \setminus \varphi_2 \equiv \varphi_1 \setminus (\varphi_2 \wedge x_i) \tag{3.3}$$

holds for the x_i such that $P(c) = \{x_i\}$. On the other hand, if $v = \mathbf{1}$, then $T(\varphi_2) = \{\mathbf{1}\}$ and it follows

$$\varphi_1 \setminus \varphi_2 \equiv \varphi_1. \tag{3.4}$$

Case (2). Clearly $(\varphi_1 \setminus \varphi_2)(v) = 0$. Thus, if $\varphi_1 \setminus \varphi_2$ represents a Horn theory, there exists a Horn clause c^* such that $c^* \geq \varphi_1 \setminus \varphi_2$ and $c^*(v) = 0$. Some c in $HC(v)$ then satisfies $c \geq c^*(\geq \varphi_1 \setminus \varphi_2)$. It can be shown that if $v \neq \mathbf{1}$, then

$$\varphi_1 \setminus \varphi_2 \equiv (\varphi_1 \wedge c) \setminus (\varphi_2 \wedge x_i) \tag{3.5}$$

holds, where the x_i appears in c (i.e., $P(c) = \{x_i\}$).

On the other hand, if $v = \mathbf{1}$, then in a manner similar to (3.4), we prove

$$\varphi_1 \setminus \varphi_2 \equiv \varphi_1 \wedge c. \tag{3.6}$$

Now we iterate as long as possible. As soon as $\varphi_2 = \perp$ or $v = \mathbf{1}$ holds, we can conclude that $\varphi_1 \setminus \varphi_2$ is Horn by (3.2), (3.4) or (3.6). If there is no $c \in HC(v)$ such that $c \geq \varphi_1 \setminus \varphi_2$, then $\varphi_1 \setminus \varphi_2$ is not Horn. In the remaining cases, we apply (3.3) or (3.5) (i.e., φ_1 is modified to $\varphi_1 \wedge c$ in case of (3.5), and φ_2 to $\varphi_2 \wedge x_i$ in both cases).

It can be shown that this procedure halts in finitely many steps. Formally, the algorithm can be written as follows.

Algorithm CHECK-HORN
Input: Horn CNFs φ_1 and φ_2.
Output: If $\varphi_1 \setminus \varphi_2$ is a Horn theory Σ, then output a Horn CNF for Σ; otherwise, "No".

Step 0. $\phi_1 := \varphi_1$; $\phi_2 := \varphi_2$;
Step 1. **if** $\phi_2 \equiv \perp$ **then** output ϕ_1 and halt
    else begin compute the least model v of ϕ_2
        if $\phi_1(v) = 0$ **then goto** Step 2
        else (i.e., $\phi_1(v) = 1$) **goto** Step 3
    end;
Step 2. **if** $v = \mathbf{1}$ **then** output ϕ_1 and halt
    else begin
      find a Horn clause c in $HC(v)$ such that $c \geq \phi_1$;
      $\phi_2 := \phi_2 \wedge x_i$, where $P(c) = \{x_i\}$;
      goto Step 1
    end;
Step 3. Find a Horn clause c in $HC(v)$ such that $c \geq \phi_1 \setminus \phi_2$;
    if no such c exists **then** output "No" and halt
    else begin $\phi_1 := \phi_1 \wedge c$;
      if $v = \mathbf{1}$ **then** output ϕ_1 and halt
      else $\phi_2 := \phi_2 \wedge x_i$, where $P(c) = \{x_i\}$;
      goto Step 1
    end. □

Example 1. Let us apply CHECK-HORN to $\varphi_1 = (\overline{x}_1 \vee \overline{x}_2) \wedge (\overline{x}_1 \vee x_3) \wedge (\overline{x}_2 \vee x_3) \wedge (\overline{x}_2 \vee x_4)$ and $\varphi_2 = \overline{x}_4 \wedge (\overline{x}_1 \vee \overline{x}_3) \wedge (\overline{x}_2 \vee \overline{x}_3)$. As we can check, the theory represented by $\varphi_1 \setminus \varphi_2$ is $\Sigma = \{(0111), (1011), (1010), (0011), (0001)\}$. Let $\phi_1 := \varphi_1$ and $\phi_2 := \varphi_2$. Since $\phi_2 \not\equiv \perp$ in Step 1, the least model $v = (0000)$ of ϕ_2 is computed. Since $\phi_1(v) = 1$, we branch to Step 3. Then $HC(v) = \{x_i \mid i = 1, 2, \ldots, 4\}$, and we can check that $\phi_1 \setminus \phi_2 \not\leq x_i$ holds for all $i = 1, 2, \ldots, 4$. Thus, the algorithm outputs "No" in Step 3. Observe that this is correct, since (1010), $(0001) \in \Sigma$ but $(0000) = (1010) \wedge (0001) \notin \Sigma$, which means that Σ is not Horn.

On the other hand, let φ_1 as above and $\varphi_2' = \overline{x}_4 \wedge (\overline{x}_2 \vee \overline{x}_3)$. Then, $\varphi_1 \setminus \varphi_2'$ represents $\Sigma' = \{(0111), (1011), (0011), (0001)\}$. Let $\phi_1 := \varphi_1$ and $\phi_2 := \varphi_2'$. Then, in Step 1 the least model is $v = (0000)$, and we again branch to Step 3, where $HC(v)$ is as previous. Now, $\varphi_1 \setminus \varphi_2' \leq x_4$ holds. Hence, we update ϕ_1 to $\phi_1 = \phi_1 \wedge x_4$ and ϕ_2 to $\phi_2 \wedge x_4$. Returning to Step 1, we find that $\phi_2 \equiv \perp$ holds. Hence,

$$\begin{aligned}\phi_1 &= (\overline{x}_1 \vee \overline{x}_2) \wedge (\overline{x}_1 \vee x_3) \wedge (\overline{x}_2 \vee x_3) \wedge (\overline{x}_2 \vee x_4) \wedge x_4 \\ &= (\overline{x}_1 \vee \overline{x}_2) \wedge (\overline{x}_1 \vee x_3) \wedge (\overline{x}_2 \vee x_3) \wedge x_4\end{aligned}$$

is output. Observe that Σ' is represented by this ϕ_1, and thus the output is correct. $\square$

An analysis of the time complexity yields the first result. Note $\varphi \equiv \perp$ for Horn φ is decidable in linear time [3], and that by (3.1), a Horn difference $\varphi_1 \setminus \varphi_2$ is representable as Horn disjunction $\psi_1 \vee \cdots \vee \psi_{m_2}$. Thus, $c \geq \varphi_1 \setminus \varphi_2$ is equivalent to Horn $\psi_i \wedge \neg c \equiv \perp$ for all $i = 1, 2, \ldots, m_2$.

Theorem 1. *Let φ_1 and φ_2 be Horn CNFs. Then, algorithm* CHECK-HORN *checks whether $\varphi_1 \setminus \varphi_2$ is Horn in $O(n^4 + n^3|\varphi_2| + n^2|\varphi_1| + n|\varphi_1||\varphi_2|)$ time.*

3.2 Model-based representation

Let Σ_1 and Σ_2 be Horn theories. Let S be the set of models defined by

$$S = C^*(\Sigma_1) \cup \{v \bigwedge w \mid v, w \in C^*(\Sigma_1)\}. \tag{3.7}$$

That is, S augments $C^*(\Sigma_1)$ by one step of the closure operator $Cl_\wedge(\cdot)$. We split S into S_1 and S_2 as follows:

$$S_1 \ \cap \ \Sigma_2 \ = \ \emptyset \quad \text{and} \quad S_2 \ \subseteq \ \Sigma_2. \tag{3.8}$$

It is easy to see that $\quad S_1 \subseteq \Sigma_1 \setminus \Sigma_2$ and $Cl_\wedge(S_2) \subseteq \Sigma_2$.

Lemma 1. *For a Horn theory Σ, let S as in (3.7), and let S_1 and S_2 be sets of models such that $S_1 \cup S_2 = S$. Then $Cl_\wedge(S_1) \cup Cl_\wedge(S_2) = \Sigma$.*

By this lemma, it holds for the S_1 and S_2 of (3.8) that

$$Cl_\wedge(S_1) \ \cup \ Cl_\wedge(S_2) = \Sigma_1. \tag{3.9}$$

The next lemma leads then to a polynomial time algorithm for our problem.

Lemma 2. *Let Σ_1 and Σ_2 be Horn theories, and let S_1 as in* (3.8)*. Then $Cl_\wedge(S_1) \cap \Sigma_2 = \emptyset$ holds if and only if $\Sigma_1 \setminus \Sigma_2$ is a Horn theory. Furthermore, if $\Sigma_1 \setminus \Sigma_2$ is a Horn theory, then $Cl_\wedge(S_1) = \Sigma_1 \setminus \Sigma_2$ (i.e., $C^*(S_1) = C^*(\Sigma_1 \setminus \Sigma_2)$) holds.*

It is known [4] that given $Q_1, Q_2 \subseteq \{0,1\}^n$, deciding $Cl_\wedge(Q_1) \cap Cl_\wedge(Q_2) = \emptyset$ is possible in $O(n(|Q_1| + |Q_1|))$ time. Since $\Sigma_2 = Cl_\wedge(C^*(\Sigma_2))$, we thus obtain from Lemma 2, turned into a straightforward algorithm, the following result.

Theorem 2. *Let Σ_1, Σ_2 be Horn theories. Given $C^*(\Sigma_1)$ and $C^*(\Sigma_2)$, we can check whether $\Sigma_1 \setminus \Sigma_2$ is Horn in $O(n|C^*(\Sigma_1)|^2|C^*(\Sigma_2)|)$ time. Furthermore, if $\Sigma_1 \setminus \Sigma_2$ is Horn, $C^*(\Sigma_1 \setminus \Sigma_2)$ can be computed in $O(n|C^*(\Sigma_1)|^2(|C^*(\Sigma_1)|^2 + |C^*(\Sigma_2)|))$ time.*

4 Horn Cores

4.1 Formula-based representation

We first consider the problem of computing one Horn core of the difference $\varphi_1 \setminus \varphi_2$, where φ_1 and φ_2 are Horn CNFs. The algorithm is a modification of algorithm CHECK-HORN, which checks whether the difference $\varphi_1 \setminus \varphi_2$ is Horn.

In Step 3 of algorithm CHECK-HORN, if no Horn clause c is in $HC(v)$ such that $c \geq \phi_1 \setminus \phi_2$, we conclude that the difference is not Horn and halt. To compute a Horn core, however, we update ϕ_1 and ϕ_2 as $\phi_1 := \phi_1 \wedge c$ and $\phi_2 := \phi_2 \wedge x_i$, respectively, for an *appropriate* $c \in HC(v)$. This is because no Horn core contains v.

Denote for a formula ψ and a model w, by $\min_{\geq w}(\psi)$ the set of minimal models u such that $\psi(u) = 1$ and $u \geq w$. Then, it holds that any clause $c \in HC(v)$ such that $c(u) = 1$ for some $u \in \min_{\geq v}(\phi_1 \setminus \phi_2)$ is appropriate for our purpose.

For models v, u with $v \leq u$, let $HC(v; u)$ be the set of Horn clauses $c \in HC(v)$ whose positive literal x_j satisfies $u_j = 1$. E.g., if $v = (111000)$ and $u = (111110)$, then $HC(v; u) = \{(\overline{x}_1 \vee \overline{x}_2 \vee \overline{x}_3 \vee x_4), (\overline{x}_1 \vee \overline{x}_2 \vee \overline{x}_3 \vee x_5)\}$. To make the discussion clear and speed up the algorithm, based on a model $u \in \min_{\geq v}(\phi_1 \setminus \phi_2)$, we update ϕ_1 and ϕ_2 respectively by

$$\phi_1 \;:=\; \phi_1 \wedge \bigwedge_{c \in HC(v;u)} c \quad \text{and} \quad \phi_2 \;:=\; \phi_2 \wedge \bigwedge_{u_j = 1 \wedge v_j = 0} x_j. \tag{4.10}$$

Thus, we have the following algorithm.

Algorithm HORN-CORE1
Input: Horn CNFs φ_1 and φ_2.
Output: A Horn core of $\varphi_1 \setminus \varphi_2$.

Steps 0.–2. as in CHECK-HORN
Step 3. **if** there exists a Horn clause c in $HC(v)$ such that $c \geq \phi_1 \setminus \phi_2$
  then begin $\phi_1 := \phi_1 \wedge c$;
    if $v = \mathbf{1}$ **then** output ϕ_1 and halt
    else $\phi_2 := \phi_2 \wedge x_i$, where $P(c) = \{x_i\}$
  end
  else begin find a model $u \in \min_{\geq v}(\phi_1 \setminus \phi_2)$;

$\phi_1 := \phi_1 \wedge \bigwedge_{c \in HC(v;u)} c$;
$\phi_2 := \phi_2 \wedge \bigwedge_{u_j = 1 \wedge v_j = 0} x_j$

end

goto Step 1. □

Let $\phi_1^{(k)}$ and $\phi_2^{(k)}$ respectively denote the ϕ_1 and ϕ_2 in Step 1 of the k-th iteration, and let $v^{(k)}$ be the least model in $\phi_2^{(k)}$. Then, it can be shown that

$$v^{(k)} < v^{(k+1)}, \qquad \text{for } k = 1, 2, \ldots. \tag{4.11}$$

This implies that the number of iterations is at most $n + 1$. The only nonobvious polynomial operation in HORN-CORE1 is finding some $u \in \min_{\geq v}(\phi_1 \setminus \phi_2)$ in Step 3. The models $\{w \geq v \mid (\phi_1 \setminus \phi_2)(w) = 1\}$ can be described by a disjunction of m Horn CNFs $\psi_1, \ldots, \psi_m$, where m is the number of clauses in ϕ_2. Consequently, the models in $\min_{\geq v}(\phi_1 \setminus \phi_2)$ are the minimal models among the least models of $\psi_1, \ldots, \psi_m$. Each ψ_i results by fixing the value of some literals in ϕ_1, and its least model can be computed in $O(|\phi_1| + n)$ time exploiting [3]. Looking at the number of 1-bits in each model, we can thus find some $u \in \min_{\geq v}(\phi_1 \setminus \phi_2)$ in $O(m(|\phi_1| + n) + mn)$ = $O(m(|\phi_1| + n))$ time. An analysis of the time complexity gives us then the following result.

Theorem 3. *Let φ_1 and φ_2 be Horn CNFs. Then algorithm* HORN-CORE1 *outputs a Horn core of $\varphi_1 \setminus \varphi_2$ in $O(n^4 + n^3|\varphi_2| + n^2|\varphi_1| + n|\varphi_1||\varphi_2|)$ time.*

As regards computing all Horn cores of $\varphi_1 \setminus \varphi_2$, we note that this problem is provably not solvable in polynomial time, since in general, an exponential number of Horn cores might exist [6]. However, we have a *polynomial total time* algorithm (called *output-polynomial* in [8]), i.e., it runs in polynomial time in the combined size of the input and the output. In fact, it enumerates all Horn cores with polynomial delay [8].

Note that in HORN-CORE1, the choice of $u \in \min_{\geq v}(\phi_1 \setminus \phi_2)$ in the else-statement of Step 3 results in a Horn core ψ of $\varphi_1 \setminus \varphi_2$ such that $\psi(u) = 1$ and $\psi(u') = 0$ for all other models $u' \in \min_{\geq v}(\phi_1 \setminus \phi_2)$. We can also show that every Horn core ψ satisfies $\psi(u) = 1$ for some u. This means that every Horn core can be constructed by algorithm HORN-CORE1, if it properly chooses a model u in the else-statement of Step 3.

Thus, all Horn cores can be generated as follows.

Algorithm ALL-HORN-CORES(φ_1, φ_2)
Input: Horn CNFs φ_1 and φ_2.
Output: All Horn cores of $\varphi_1 \setminus \varphi_2$.

Steps 0.–2. as in HORN-CORE1 and CHECK-HORN.
Step 3. **if** there exists a Horn clause c in $HC(v)$ such that $c \geq \phi_1 \setminus \phi_2$
then begin $\phi_1 := \phi_1 \wedge c$;
if $v = \mathbf{1}$ **then** output ϕ_1 and exit;
else begin $\phi_2 := \phi_2 \wedge x_i$, where $P(c) = \{x_i\}$; **goto** Step 1 **end**
end
else for each a model $u \in \min_{\geq v}(\phi_1 \setminus \phi_2)$ **do**

begin $\phi_1 := \phi_1 \wedge \bigwedge_{c \in HC(v;u)} c$;
$\phi_2 := \phi_2 \wedge \bigwedge_{u_j=1 \wedge v_j=0} x_j$;
call ALL-HORN-CORES(ϕ_1,ϕ_2)
end{for}. □

An analysis of the time complexity leads us to the following result.

Theorem 4. *Algorithm* ALL-HORN-CORES *correctly generates all Horn cores of* $\varphi_1 \setminus \varphi_2$ *with polynomial delay, where the delay is bounded by the time of computing one Horn core, i.e.,* $O(n^4 + n^3|\varphi_2| + n^2|\varphi_1| + n|\varphi_1||\varphi_2|)$.

4.2 Model-based representation

In Subsection 3.2, we have seen that S_1 of (3.8) characterizes the difference $\Sigma_1 \setminus \Sigma_2$ of two Horn theories Σ_1 and Σ_2. In this subsection, we show that any maximal set $Q \subseteq S_1$ such that $Cl_\wedge(Q) \cap \Sigma_2 = \emptyset$ gives a Horn core of $\Sigma_1 \setminus \Sigma_2$, and conversely any Horn core of $\Sigma_1 \setminus \Sigma_2$ can be generated from such a maximal set $Q \subseteq S_1$.

Note that, by the definition of S_1, any maximal set $Q \subseteq S$ with $Cl_\wedge(Q) \cap \Sigma_2 = \emptyset$ is contained in S_1. Thus we prove the above statements by using a set S instead of S_1.

Theorem 5. *Let* Σ_1, Σ_2 *be Horn theories and let* S *as in* (3.7)*. Then,* Π *is a Horn core of* $\Sigma_1 \setminus \Sigma_2$ *iff* $\Pi = Cl_\wedge(Q)$ *for some maximal* $Q \subseteq S$ *such that* $Cl_\wedge(Q) \cap \Sigma_2 = \emptyset$. □

Based on Theorem 5, we obtain the following polynomial time algorithm.

Algorithm HORN-CORE2
Input: Characteristic sets $C^*(\Sigma_1)$ and $C^*(\Sigma_2)$ of Horn theories Σ_1 and Σ_2.
Output: The characteristic set $C^*(\Pi)$ of a Horn core Π of $\Sigma_1 \setminus \Sigma_2$.

Step 0. $Q := \emptyset$;
compute the set of models S given by (3.7);
Step 1. **for** each $w \in S$ **do begin**
if $Cl_\wedge(Q \cup \{w\}) \cap \Sigma_2 = \emptyset$ **then** $Q := Q \cup \{w\}$
end{for};
Step 2. output $C^*(Q)$ and halt. □

Theorem 6. *Let* Σ_1 *and* Σ_2 *be Horn theories. Then,* HORN-CORE2 *outputs the characteristic set of a Horn core of* $\Sigma_1 \setminus \Sigma_2$ *in* $O(n|C^*(\Sigma_1)|^2(|C^*(\Sigma_1)|^2 + |\Sigma_2|))$ *time.*

Is also possible to decide whether a given theory Π is a Horn core of the difference $\Sigma_1 \setminus \Sigma_2$ in polynomial time. A suitable algorithm is straightforward by proper modification of algorithm HORN-CORE2. We omit the details (see [6]).

As for the efficient computation of all Horn cores, we have a negative result. There is no polynomial total time algorithm for generating all Horn cores, unless P = NP.

Theorem 7. *There is no algorithm which, given* $C^*(\Sigma_1)$ *and* $C^*(\Sigma_2)$ *of Horn theories* Σ_1 *and* Σ_2*, computes the characteristic sets* $C^*(\Pi)$ *of all Horn cores* $\Pi \in \mathcal{HC}(\Sigma_1 \setminus \Sigma_2)$ *in polynomial total time* (*i.e., polynomial in the combined size of the input and the output*)*, unless* P = NP.

5 Conclusion and Further Work

The results of the present and companion papers [4, 5] establish operations towards a Boolean "calculus" on Horn theories, in which Horn theories are combined using the operations of conjunction ($\wedge$), disjunction ($\vee$), and difference ($\setminus$). These operations are useful, for example, in the context of changing theories that are possible world representations of a state of affairs.

Several issues remain for our ongoing work. One is giving a precise account to computing Horn cores and the Horn envelope of the complement $\overline{\Sigma} = \{0, 1\}^n \setminus \Sigma$ of a Horn theory. Our results imply polynomial time algorithms in some cases, but a complete picture remains to be drawn. Another issue is a more accurate account of the complexity of the polynomial cases in Table 1. Under formula-based representation, all these problems are complete for P under logspace reductions; this is an easy consequence of the fact that deciding the satisfiability of a Horn CNF is complete for P under logspace reductions. Therefore, parallelization of these problems is most likely not possible. Another interesting issue is a study of the effect of Horn renamings (cf. [1]).

Acknowledgments. The authors appreciate the comments of the anonymous reviewers.

References

1. Y. Boufkhad. Algorithms for Propositional KB Approximation. In *Proc. National Conference on AI (AAAI '98)*, Madison, Wisconsin, pp. 280–285, July 26–30 1998.
2. M. Cadoli. Semantical and Computational Aspects of Horn Approximations. In *Proc. IJCAI-93*, pp. 39–44, 1993.
3. W. Dowling and J. H. Gallier. Linear-time Algorithms for Testing the Satisfiability of Propositional Horn Theories. *J. Logic Programming*, 3:267–284, 1984.
4. T. Eiter, T. Ibaraki, and K. Makino. Computing Intersections of Horn Theories for Reasoning with Models. In: *Proc. AAAI '98*, pp. 292–297, 1998.
5. T. Eiter, T. Ibaraki, and K. Makino. Disjunctions of Horn Theories and their Cores. In *Proc. 9th Int'l Symp. on Algorithms and Computation (ISAAC '98)*, Dec. 1998.
6. T. Eiter, T. Ibaraki, and K. Makino. On the Difference of Horn Theories. RUTCOR Research Report RRR 25-98, Rutgers University 1998.
7. G. Gogic, C. Papadimitriou, and M. Sideri. Incremental Recompilation of Knowledge. *J. Artificial Intelligence Research*, 8:23–37, 1998.
8. D. S. Johnson, M. Yannakakis, and C. H. Papadimitriou. On Generating All Maximal Independent Sets. *Information Processing Letters*, 27:119–123, 1988.
9. H. Katsuno and A. O. Mendelzon. Propositional Knowledge Base Revision and Minimal Change. *Artificial Intelligence*, 52:253–294, 1991.
10. H. Kautz, M. Kearns, and B. Selman. Reasoning With Characteristic Models. In *Proc. AAAI '93*, 1993.
11. H. Kautz and B. Selman. Knowledge Compilation and Theory Approximation. *JACM*, 43(2):193–224, 1996.
12. D. Kavvadias, C. Papadimitriou, and M. Sideri. On Horn Envelopes and Hypergraph Transversals. In *Proc. ISAAC '93*, LNCS 762, pp. 399–405, 1993.
13. R. Khardon and D. Roth. Reasoning with Models. *Art. Intelligence*, 87(1/2):187–213, 1996.