Approximation of relations by propositional formulas: complexity and semantics

Bruno Zanuttini

Approximation of relations by propositional formulas : complexity and semantics

Bruno Zanuttini[1]

GREYC - Université de Caen
bd. Mal Juin, F-14032 Caen Cedex, France
zanutti@info.unicaen.fr

Abstract. Selman and Kautz introduced the notion of approximation of a theory and showed its usefulness for knowledge compilation and on-line reasoning. We study here the complexity of the main computational problems related to the approximation of relations (sets of possible worlds) by propositional formulas, and the semantics of reasoning with these approximations (deduction and abduction). The classes of formulas that we consider are those of (reverse-)Horn, bijunctive and affine formulas, which are the most interesting for storing knowledge.
Concerning the computation of approximations, we survey and complete the results that can be found in the literature, trying to adopt a unified point of view. On the contrary, as far as we know this paper is the first real attempt to study the semantics of abduction with the bounds of a theory.

1 Introduction

Recently there has been a great amount of research about *knowledge approximation* ([DP92,KKS95,Val95,SK96,KR96] for instance). An approximation of a theory Σ representing some knowledge is another theory that approaches it the best (in some precise sense) but allows for efficient reasoning (for instance, deduction) when Σ does not. The idea is to store the approximation with the theory itself, and to use it for speeding up further reasoning. Thus computing approximations can be viewed as a *knowledge compilation* task [SK96] ; practically speaking, this compilation is most of the time intended as an *off-line* process performed in order to speed up further *on-line* reasoning.

We are interested here in *propositional* theories. Many papers have focused on them [DP92,KKS95,SK96,Bou98,CS00], may the original theories be given as relations [DP92,KKS95] or as formulas [Val95,SK96,CS00]. In this paper, we consider the case when the original theory Σ is given as a relation, which corresponds to the case when konwledge is stored as a set of observations [DP92,ZH02a] (or *possible worlds* [KKS95]) ; in this framework, we study the complexity of computing a *propositional formula* approximating Σ. The motivation for approximating by formulas is that most of the time formulas are smaller than their set of models, and since wa aim at storing the approximation and reasoning with it, its size is a very important parameter to take into account.

Following [SK96], we measure closeness of the approximation to the original theory in terms of sets of models ; we then talk about *Upper Bounds* when the approximation is *logically weaker* than the theoy, and of *Lower Bounds* when it is *logically stronger*. Finally, we are interested in approximations into the classes of (reverse-)Horn, bijunctive and affine formulas, for these classes are the most important ones for reasoning and representing knowledge, in many senses. Indeed, they are, among other properties, tractable for the satisfiability problem, stable by conjunction, by conjunction of facts (unit clauses) and by unit resolution, and tractable for deduction of clauses.

Our contribution is the following. We first survey and complete the results concerning the complexity of computing Least Upper Bounds and Greatest Lower Bounds of relations into these classes of formulas, trying to adopt a unified point of view ; moreover, we give explicit examples and counter-examples, or detail the proofs that can be found in the literature when useful. Secondly, we study the semantics of reasoning with such bounds, i.e., the links between the answers obtained when reasoning with the bounds and those that could have been obtained with the original theory. While deduction has been extensively studied in [Val95,SK96], as far as we know the semantics of abduction have not really been studied in the literature.

The paper is structured as follows. Section 2 is devoted to the explanation of our definitions and conventions. In section 3 we discuss the computation of *upper bounds*, in section 4 that of *lower* bounds, and in section 5 we discuss the semantics of reasoning with these bounds. We end with a summary of the results and further directions for research in section 6.

2 Preliminaries

We assume a countable number of propositional variables $x_1, x_2, \ldots$. A *relation* on $\{0,1\}$ is a subset of $\{0,1\}^n$; for instance, $R = \{001, 011, 101\}$ is a relation. We view any vector m of $\{0,1\}^n$ as a $0/1$ assignment to n variables (in the order of their indices), and m_i denotes the value assigned by m to x_i. We write $|R|$ the number of vectors in a relation R.

A *literal* is either a variable x_i (*positive* literal) or its negation $\neg x_i$ (*negative* literal). A *clause* is a finite disjunction of literals, and a formula is in *conjunctive normal form* (CNF for short) if it is written as a finite conjunction of clauses. A $0/1$ assignment m to the variables of a formula ϕ in CNF is a *model* of ϕ (written $m \models \phi$) if it satisfies at least one literal in each clause of ϕ. For instance, the assignment $m = 001$ to the variables x_1, x_2, x_3 is a model of the CNF $\phi = (x_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_2 \vee \neg x_3)$. An *affine* formula [Sch78,KS98] is a conjunction of linear equations modulo 2. A *model* of an affine formula ϕ is an assignment m to its variables that satisfies all the equations (written $m \models \phi$). For instance, $m = 001$ is a model of the affine formula $\phi = (x_1 \oplus x_3 = 1) \wedge (x_1 \oplus x_2 = 0)$. As can be seen, equations play the same role for affine formulas than clauses do for CNFs. If ϕ is a CNF or an affine formula, we denote by $\mathcal{M}(\phi)$ the set of its models.

A formula in conjunctive normal form is *Horn* if it has at most one positive literal per clause ; it is reverse-Horn if it has at most one negative literal per clause. For instance, the formula $\phi_1 = (x_1 \vee \neg x_2 \vee \neg x_3) \wedge (x_2) \wedge (\neg x_1 \vee \neg x_3)$ is Horn, and $\phi_2 = (x_1 \vee \neg x_2) \wedge (x_3)$ is reverse-Horn. Note that if a CNF is reverse-Horn, then replacing each variable with its negation in it yields a Horn formula, and conversely ; the corresponding models are also strongly linked, in the sense that the models of one are the others' where 0 is replaced with 1 and 1 with 0. For that reason, we only study Horn formulas in the paper ; all the results for reverse-Horn formulas are dual. Finally, a CNF is *bijunctive* if it has at most two literals per clause. For instance, ϕ_1 above is not bijunctive while ϕ_2 is. We denote by HORN (resp. REVERSE-HORN, BIJUNCTIVE) the class of all Horn (resp. reverse-Horn, bijunctive) CNFs, and by AFFINE the class of all affine formulas. Unless stated otherwise, when we consider a class of formulas $\mathcal{C}$ it is always one of these. In the paper, the term *formula* groups both CNFs and affine formulas.

Our motivation for studying these classes of formulas comes from their very good computational properties for reasoning. First of all, they are all polynomial for the satisfiability problem : linear for (reverse-)Horn and bijunctive formulas, and quadratic for affine (more precisely, SAT for affine formulas can be solved in time $O(k^2 n)$ if the formula consists of k equations over n variables, with gaussian elimination, see for instance [Cur84]). Secondly, it follows from their definitions that all these classes are stable by conjunction, which allows for instance to update a Horn knowledge base by adding new Horn rules while preserving the Horn form. From their definitions it also follows that they are stable by conjunction of facts (unit clauses) and by unit resolution, which allows for instance one to specify a value for a variable and propagate it while preserving the class, and which also implies that deduction of clauses, one of the most common reasoning problems, is tractable ; indeed, deciding whether Σ implies $\ell_1 \vee \cdots \vee \ell_p$ is the same as deciding whether $\Sigma \wedge \overline{\ell_1} \wedge \cdots \wedge \overline{\ell_p}$ is unsatisfiable, and thus stability by conjunction of unit clauses and by unit resolution, in addition of tractability for SAT, makes deduction of clauses tractable. Moreover, in some sense these classes are the only ones to possess tractability for SAT and stability by conjunction, which follows from Schaefer's dichotomy theorem [Sch78]. Finally, these classes allow efficient restricted forms of abduction [Val00,Z02b] while this problem, also of major importance for reasoning, is Σ_2^P-complete in general [EG95]. For all these reasons we argue that these four classes are the most interesting ones for reasoning and representing propositional knowledge.

A formula ϕ is said to *describe* a relation R if $R = \mathcal{M}(\phi)$; ϕ is called a *description* of R. For $m, m' \in \{0,1\}^n$, let us write $m \wedge m'$ the vector $m'' \in \{0,1\}^n$ such that $\forall i = 1, \ldots, n, m''_i = m_i \wedge m'_i$, and similarly for $\vee$ and $\oplus$. The following result, proved by Schaefer, is essential for the paper.

Proposition 1 ([Sch78]). *Let R be a relation. Then R has at least one description in*
(i) HORN *if and only if* $\forall m, m' \in R, m \wedge m' \in R$,

(ii) BIJUNCTIVE *iff* $\forall m, m', m'' \in R, (m \vee m') \wedge (m \vee m'') \wedge (m' \vee m'') \in R$,
(iii) AFFINE *iff* $\forall m, m', m'' \in R, m \oplus m' \oplus m'' \in R$.

3 Upper Bounds

We are first interested in the computation of *upper bounds* of relations. The idea is that we weaken the original theory in order to obtain a more interesting representation.

Definition 1 ($\mathcal{C}$-upper bound). *Let $R \subseteq \{0,1\}^n$ be a relation and $\mathcal{C}$ a class of formulas. A formula ϕ on the variables $x_1, \ldots, x_n$ is called a* $\mathcal{C}$-upper bound *($\mathcal{C}$-UB) of R if $\phi \in \mathcal{C}$ and $R \subseteq \mathcal{M}(\phi)$. If this holds for no formula ϕ' with $\mathcal{M}(\phi') \subset \mathcal{M}(\phi)$, ϕ is called a* $\mathcal{C}$-least upper bound *($\mathcal{C}$-LUB) of R.*

Of course, the most interesting upper bounds are the *least* ones, for they approach the original theory the closest and thus yield a smaller loss of information than general upper bounds. Proposition 1 immediately yields the following result for $\mathcal{C}$-LUBs.

Proposition 2. *Assume $\mathcal{C}$ is either* HORN, BIJUNCTIVE *or* AFFINE. *Then the $\mathcal{C}$-LUB of any relation is unique up to logical equivalence.*

Proof. From Proposition 1 it follows that the set of models of the $\mathcal{C}$-LUB of a relation R is the closure of R under the operation corresponding to the class. □

Example 1. Consider the relation $R = \{011, 110, 101\}$. It is easily seen that $011 \wedge 110 = 010$, $011 \wedge 101 = 001$, $110 \wedge 101 = 100$ and $010 \wedge 001 = 011 \wedge 100 = \cdots = 000$, and finally that the set of models of any HORN-LUB of R is $\{0,1\}^3 \backslash \{111\}$.

For sake of simplicity now, we will say a formula ϕ is *the* $\mathcal{C}$-LUB of a relation. Our problem is to compute such a ϕ when $\mathcal{C}$ is one of the classes of interest here.

We first consider HORN-LUBs. Unfortunately, there cannot be an input-polynomial algorithm for computing the HORN-LUB of a relation, since the former can be exponentially bigger than the latter, as shown in [KKS95] (where an explicit example is given).

Proposition 3 ([KKS95, Theorem 6]). *There exist relations on $2n$ variables, consisting of $3n(n-1)$ vectors but whose* HORN-*LUB has 2^n clauses.*

Since there cannot be an input-polynomial algorithm for this problem, it is worth wondering whether there exist an *output-polynomial* one [JYP88]. But it is shown in [KPS93,Kha95] that this problem is harder than the enumeration of all the minimal transversals of an hypergraph, which is open and for which the best known algorithm is subexponential [Kha95].

Now we turn our attention to bijunctive and affine formulas. Contrary to the Horn case, for both classes polynomial algorithms are given in the literature. The idea for bijunctive formulas is that we only need to project the relation

onto every pair of variables and compute the associated clauses, and the proof for affine formulas uses a strong link between the sets of models of affine formulas and vector spaces, which allows to use the notion of *basis* of such a space.

Proposition 4 ([DP92, Lemma 3.18]). *The* BIJUNCTIVE-*LUB of a relation* $R \in \{0,1\}^n$ *can be computed in time* $O(|R|n^2)$.

Proposition 5 ([Z02a, Proposition 6]). *The* AFFINE-*LUB of a relation* $R \in \{0,1\}^n$ *can be computed in time* $O(|R|n^3 + n^4)$.

4 Lower Bounds

We are now interested in *lower bounds* of relations, i.e., in *strengthening* the original theory. The following definition is dual to Definition 1.

Definition 2 ($\mathcal{C}$-lower bound). *Let* $R \subseteq \{0,1\}^n$ *be a relation and* $\mathcal{C}$ *a class of formulas. A formula* ϕ *on the variables* $x_1, \ldots, x_n$ *is called a* $\mathcal{C}$-lower bound *($\mathcal{C}$-LB) of* R *if* $\phi \in \mathcal{C}$ *and* $\mathcal{M}(\phi) \subseteq R$. *If this holds for no formula* ϕ' *with* $\mathcal{M}(\phi) \subset \mathcal{M}(\phi')$, ϕ *is called a* $\mathcal{C}$-greatest lower bound *($\mathcal{C}$-GLB) of* R.

Dually to the case of upper bounds we are mainly interested in *greatest* lower bounds. But contrary to its $\mathcal{C}$-LUB, the $\mathcal{C}$-GLB of a relation is not unique in general. We are thus interested in two problems : find one $\mathcal{C}$-GLB of a given theory, and find one with the maximum number of models. Indeed, two $\mathcal{C}$-GLBs of the same relation have not even the same number of models in the general case, and one of them may even be exponentially bigger than another, as the next example illustrates it.

Example 2. Consider the relation over the variables $x_1, \ldots, x_n$:

$$R = \{m100 | m \in \{0,1\}^{n-3}\} \cup \{0 \ldots 0010, 0 \ldots 0001\}$$

where $m100$ stands for the assignment of m_i to x_i for $i = 1, \ldots, n-3$, 1 to x_{n-2} and 0 to x_{n-1}, x_n, $0 \ldots 0010$ stands for the assignment of 0 to $x_1, \ldots, x_{n-3}$, 0 to x_{n-2}, 1 to x_{n-1} and 0 to x_n, and similarly for $0 \ldots 0001$. It is easily seen with Proposition 1 that $S_{ha} = \{m100 | m \in \{0,1\}^{n-3}\}$ is the set of models of a HORN- and AFFINE-GLB of R, and $S_b = \{m100 | m \in \{0,1\}^{n-3} \cup \{0 \ldots 0010\}$ that of a BIJUNCTIVE-GLB of it. On the other hand, $S'_h = \{0 \ldots 0010\}$ is also the set of models of a HORN-GLB of R, and $S'_{ba} = \{0 \ldots 0010, 0 \ldots 0001\}$ of a BIJUNCTIVE- and AFFINE-GLB. However, for all n it holds that $|R| = 2^{n-3} + 2$, $|S_{ha}| = 2^{n-3}$, $|S_b| = 2^{n-3} + 1$ and $|S'_h| = 1$, $|S'_{ba}| = 2$.

4.1 Computing one GLB

The first natural problem concerning the computation of GLBs is the one of computing *one* GLB. For the classes we are interested in the problem appears to be polynomial, mainly because one can compute in polynomial time a formula $\phi \in \mathcal{C}$ given the set of its models [DP92,KS98,ZH02a]. Thus we can first compute

the set of models S of a $\mathcal{C}$-GLB of R, which can be done by selecting vectors from R as long as the closure of the current S is included in R, and then a description of S. Since the size of S is by definition less than the size of R, the whole process is polynomial.

Proposition 6 (see also [KPS93, Theorem 2]). *A* HORN-*GLB of a relation* $R \subseteq \{0,1\}^n$ *can be computed in time* $O(|R|^2 n^2)$.

Proof. We begin by computing the set of models S of the HORN-GLB. Initialize S to $\emptyset$, and perform the following until there is no more convenient vector m in R : pick $m \in R$ such that $\forall m' \in S, m \wedge m' \in R$ and set $S \leftarrow S \cup \{m\} \cup \{m \wedge m' | m' \in S\}$. Then S is always closed under $\wedge$ and included in R, and when no more convenient m can be chosen in R, then S is obviously maximal for set inclusion. Then compute a formula ϕ describing S.

The running time is established as follows : since at most $|R|$ vectors can be added to S, and at each step one has to check that at most $|R|$ products $m \wedge m'$ are in R (which requires $O(n)$ steps if R is initially sorted into a decision tree, in time $O(|R|n)$), computing S requires time $O(|R|^2 n)$. Then $|S| \leq |R|$ holds, and describing S into a Horn formula requires $O(|S|^2 n^2)$ steps [DP92,ZH02a]. □

Proposition 7. *A* BIJUNCTIVE- *or* AFFINE-*GLB of a relation* $R \in \{0,1\}^n$ *can be computed in time* $O(|R|^3 n + |R|^2 n^2)$.

Proof. The proof is the same as for Proposition 6, except that the test for stability requires three vectors ; thus computing S requires time $O(|R|^3 n)$ instead of $O(|R|^2 n)$, and describing S still requires $O(|S|^2 n^2)$ steps [ZH02a]. □

However, computing one GLB may turn out to be rather uninteresting ; indeed, as shown in Example 2 one can find a GLB ϕ whose set of models is exponentially smaller than that of another GLB ϕ'. In such a case, it is natural to consider that the bounds with exponentially many models are the best ones. For this reason, the problem of computing a GLB *with the maximum number of models over all GLBs* is more interesting.

4.2 Computing a GLB with the maximum number of models

We now consider the problem of computing a GLB that has the maximum number of models over all GLBs. More formally, for R a relation and $\mathcal{C}$ a class of formulas, let us call a $\mathcal{C}$-*maxGLB of* R any $\mathcal{C}$-GLB ϕ of R such that no other $\mathcal{C}$-GLB ϕ' of R has more models than ϕ ; we emphasize that we do not consider here the relations of inclusion between the sets of models of ϕ and ϕ', but only the number of these models.

As for $\mathcal{C}$-LUBs, the results here are not the same for all the classes we are interested in. In each case computing a $\mathcal{C}$-maxGLB is hard, but it is NP-hard for HORN and BIJUNCTIVE while subexponential, and thus unlikely to be NP-hard [SH90], for AFFINE.

In order to show NP-hardness, we show that it is NP-hard to compute the *set of models* of a maxGLB ; indeed, the models of any Horn or bijunctive formula

can be enumerated in output-polynomial time, and since a GLB of R has by definition less than $|R|$ models, we deduce that up to a polynomial overcost computing the set of models of a maxGLB ϕ is not harder than computing ϕ ; thus if the former computation is NP-hard, then the latter also is.

Proposition 8 (see also [KPS93, Theorem 2]). *If P≠NP, a* HORN-*max-GLB of a relation $R \subseteq \{0,1\}^n$ cannot be computed in time polynomial in $|R|$ and n.*

Proof. We consider the associated decision problem : given R and $k \leq |R|$, is there a HORN-glb of R that consists in at least k vectors ?, and we reduce the problem `INDEPENDENT-SET`, which is known to be NP-complete (see [GJ83]), to this one. Let us recall that an independent set of a graph G is a subset V' of its vertices such that no two vertices in V' are joined by an edge of G. The problem `INDEPENDENT-SET` is the following :

Input : A graph $G = (V, E)$ and a positive integer $k \leq |V|$

Question : Does G contain an independent set of size at least k ?

The reduction is the following : first consider $|E|$ variables $x_1, \ldots, x_{|E|}$, each variable x_i corresponding to the edge e_i of G, and associate to each vertex v_j of G a vector $m^{[j]}$ such that for each variable x_i, $m_i^{[j]}$ is 1 if and only if $v_j \in e_i$. Then introduce $|V|$ new variables $x_{|E|+1}, \ldots, x_{|E|+|V|}$, each variable $x_{|E|+k}$ corresponding to the vertex v_k of G, and complete each vector $m^{[j]}$ with $m^{[j]}_{|E|+k} = 1$ if and only if $j = k$, i.e., if and only if $m^{[j]}$ is the vector associated to v_k (this distinguishes, for example, the two vertices of the connected graph with only one edge). Finally, add the all-0 vector $\mathbf{0}$. Call R this relation ; it clearly contains $|V| + 1$ vectors and $|V| + |E|$ variables, and its construction from G is polynomial.

We show that G has an independent set of size k if and only if R has a HORN-glb of size at least $k + 1$. Assume first that G has an independent set $V' = \{v_1, v_2, \ldots, v_k\}$ (without loss of generality). Consider the set of vectors $S = \{\mathbf{0}, m^{[1]}, \ldots, m^{[k]}\} \subseteq R$. Obviously, for all i, $\mathbf{0} \wedge m^{[i]} = \mathbf{0} \in S$. Now for $m^{[i]}, m^{[j]} \in S, m^{[i]} \neq m^{[j]}$, since V' is an independent set there is no edge joining v_i and v_j, thus $m^{[i]}$ and $m^{[j]}$ have no 1 in common on the first $|E|$ variables ; since $m^{[i]} \neq m^{[j]}$, they have no 1 in common on the last $|V|$ variables neither, and finally $m^{[i]} \wedge m^{[j]} = \mathbf{0} \in S$. We conclude that S is stable by $\wedge$, thus R has a HORN-GLB of size at least $k + 1$.

Conversely, assume that G has no independent set of size k. This means that for any subset $V' \subseteq V, |V'| = k$, there exist $v_i, v_j \in V'$ joined by an edge e. We conclude that for any subset S of R of size at least $k+1$, there exist $m^{[i]}, m^{[j]} \in S$ with one 1 in common on the first $|E|$ variables, and since $m^{[i]} \neq m^{[j]}$ none in common among the last $|V|$ variables. Thus $m^{[i]} \wedge m^{[j]}$ is all-0 except for at least one of the first $|E|$ variables. Since every vector in R is either all-0, or 1 on exactly one of the last $|V|$ variables, $m^{[i]} \wedge m^{[j]} \notin R$. Thus no subset of R is both of size at least $k + 1$ and stable by $\wedge$, thus R has no HORN-GLB of size $k + 1$ or more. □

Proposition 9. *If P≠NP, a* BIJUNCTIVE-*maxGLB of a relation* $R \subseteq \{0,1\}^n$ *cannot be computed in time polynomial in* $|R|$ *and* n.

Proof. As for Proposition 8, we reduce the problem `INDEPENDENT-SET` to the associated decision problem. We encode a given instance of `INDEPENDENT-SET` in the same manner as for Horn, except that we replace 0 with 1 and 1 with 0 for each of the first $|E|$ variables. We let the last $|V|$ variables distinguish all the vectors as before, with exactly one 1 per vector. The vector $\mathbf{0}$ is similarly replaced with $\mathbf{A} = 1 \ldots 10 \ldots 0$.
Assume G has an independent set $\{v_1, v_2, \ldots, v_k\}$, and consider the subset $S = \{\mathbf{A}, m^{[1]}, \ldots, m^{[k]}\}$ of R. We show that for any $m, m', m'' \in S$, $m_0 = (m \vee m') \wedge (m \vee m'') \wedge (m' \vee m'') \in S$. Indeed, m and m' have no 0 in common on the first $|E|$ variables, for either the associated vertices have no edge in common or one of the two vectors is all-1 on these variables. Thus $m \vee m'$ is 1 on each one of the first $|E|$ variables. The same holds for $m \vee m''$ and $m' \vee m''$, thus m_0 assigns 1 to each one of the first $|E|$ variables. It is easily seen that it assigns 0 to the last $|V|$ variables whatever m, m' and m'' may be, thus $m_0 = \mathbf{A} \in S$. Thus R has a BIJUNCTIVE-GLB of size at least $k+1$. The converse is easily seen to be true, for if m and m' (without loss of generality) have one edge in common, then $m \vee m'$ assigns 0 to the corresponding variable, thus also m_0, but m_0 still assigns 0 to all of the last $|V|$ variables ; since such a vector does not exist in R, the associated subset of R cannot be completed into a BIJUNCTIVE-LB of R. □

On the contrary, a subexponential algorithm is given in [Z02a] for the affine case. As for computing an AFFINE-GLB, it uses the correspondence with vector spaces and the notion of basis. Note that subexponential algorithms may be reasonable in practice, while NP-hard problems are unlikely to be tractable. Note also that subexponential algorithms are unlikely to be NP-hard [SH90].

Proposition 10 ([Z02a, Proposition 7]). *An* AFFINE-*maxGLB of a relation* $R \subseteq \{0,1\}^n$ *can be computed in time* $O(|R|n(\log\log|R|)2^{(\log|R|)^2})$.

5 Reasoning with bounds

Finally, we study the semantics of reasoning with the bounds of a theory. The problem is to characterize the results obtained when reasoning with the bounds in function of those obtained when reasoning with the theory itself. Of course, such characterizations are of major importance for being able to use the bounds of the theory for speeding up further answering to queries. Indeed, the idea is to use as often as possible *only* the approximation for answering the queries that are asked to the original theory, for it is meant to have much better computational properties for reasoning ; that is why we must be able to deduce the right answer (the one that the original theory would have given) from the one that we compute with the approximation. Moreover, if it appears that the bounds allow to solve any reasoning task the theory will be possibly sollicited for, we may even store only these bounds and forget the theory itself.

We however emphasize that the size of a bound must be comparable to or lower than the size of the original theory, for otherwise, on one hand there would be no gain in reasoning with it, even if reasoning is more efficient in the class of the approximation than in the general case, and on the other hand its storage could need too much memory.

5.1 Deduction

We talk about deduction when we want to answer the question "does Σ logically implies α ?", where Σ is a theory and α (the *query*) is another theory ; the answer is positive (written $\Sigma \models \alpha$) if and only if every model of Σ is a model of α, i.e., α is true in all the situations where Σ is. Deduction is in general coNP-complete, since Σ implies α if and only if $\Sigma \wedge \overline{\alpha}$ is unsatisfiable, and thus it is worth considering the semantics of deduction with bounds in classes of formulas for which it becomes tractable.

Selman and Kautz [SK96] have studied the semantics of deduction with bounds. The following result is given in [SK96] for Horn bounds, but it can be straightforwardly extended to any class of formulas.

Proposition 11 ([SK96, Section 2.2]). *Let Σ and α be two propositional theories, and let $\mathcal{C}$ be a class of propositional formulas. Let Σ_{LUB} be a $\mathcal{C}$-LUB of Σ, and Σ_{GLB} a $\mathcal{C}$-GLB. Then :*
(i) If Σ_{LUB} implies α, then Σ implies α
(ii) If Σ_{GLB} does not imply α, then Σ does not imply α.

It follows that we can use the bounds for speeding up reasoning with the original theory in the following manner. When asked whether $\Sigma \models \alpha$, we first decide whether $\Sigma_{LUB} \models \alpha$; if the bound Σ_{LUB} is taken in a class of formulas for which deduction is tractable, and if the size of Σ_{LUB} is comparable to or less than the size of Σ, then this test is efficient. If its answer is positive, then we can conclude. Otherwise, we decide whether $\Sigma_{GLB} \models \alpha$; if the answer is negative, once again we can conclude. In case it is positive, then we have no other solution than to decide directly $\Sigma \models \alpha$, but we have only lost a small amount of time. For more details and examples we refer the reader to [SK96] for instance.

When restricting the form of the queries, the $\mathcal{C}$-LUB of a relation can even allow to solve *exactly* deduction tasks without any use of the original theory, as stated in the next proposition (whose proof can be straightforwardly adapted from del Val's).

Proposition 12 (see [Val95, Theorem 1]). *Assume $\mathcal{C}$ is either* HORN, BIJUNCTIVE *or* AFFINE. *Let Σ be a theory and $\alpha \in \mathcal{C}$. Then the $\mathcal{C}$-LUB of Σ implies α if and only if Σ itself implies α.*

5.2 Abduction

We now consider the problem of *abduction*. Contrary to the case of deduction, as far as we know the semantics of this process with bounds have not really been

studied in the literature, except in a quite special case in [KKS95, Section 5]. This section is a first step into this direction.

Abduction consists in searching for *explanations* of observations, knowing a background theory. More formally, a background theory is simply a propositional theory Σ, supposed to be satisfiable, and an observation is another theory α. We explicit what an explanation is in the following definition.

Definition 3 (explanation). *Let Σ and α be two propositional theories, and H a subset of $Var(\Sigma)\backslash Var(\alpha)$ (the* set of hypotheses*). Then an* explanation *of α knowing Σ over H is a conjunction E of literals such that :*
(i) $Var(E) \subseteq H$
(ii) $\Sigma \wedge E$ implies α
(iii) $\Sigma \wedge E$ is satisfiable.
If in addition there is no proper subset of E with the same properties, E is called a minimal *explanation of α knowing Σ over H.*

Example 3. Let Σ be the theory represented by the CNF formula $(x_1 \vee x_3 \vee \neg x_5) \wedge (x_1 \vee \neg x_3) \wedge (\neg x_1 \vee x_3 \vee x_4) \wedge (\neg x_1 \vee x_2 \vee \neg x_3 \vee x_6) \wedge (\neg x_4 \vee x_5)$, let α be the clause $(x_5 \vee x_6)$ and H the set of hypotheses $\{x_1, x_2, x_3\}$. Then $E = \{x_1, \neg x_2\}$ is an explanation of α knowing Σ over H. Moreover, it is a *minimal* explanation, since neither $E' = \{x_1\}$ nor $E'' = \{\neg x_2\}$ are explanations.

With this definition, the task of abduction consists in finding, given Σ, α and H, a minimal explanation E of α knowing Σ over H. It is well-known that in general abduction is Σ_2^P-complete [EG95], and NP-hard if we restrict Σ to be a Horn formula [EG95]. On the other hand, for bijunctive and affine formulas it becomes polynomial if in addition we impose some restrictions to the form of α (see for instance [Val00,Z02b]). Therefore, in all cases it is interesting to study the semantics of abduction with bounds.

We first give general remarks that help understanding these semantics in the general case, independently of the class of the approximation and of the restrictions of the general problem.

Proposition 13. *Let Σ and α be two propositional theories, and H a subset of $Var(\Sigma)\backslash Var(\alpha)$. Let $\mathcal{C}$ be a class of formulas, and Σ_{LUB} (resp. Σ_{GLB}) be a $\mathcal{C}$-LUB (resp. a $\mathcal{C}$-GLB) of Σ. Let E be a conjunction of literals formed upon the variables in H. Then :*
(i) If $\Sigma_{GLB} \wedge E$ is satisfiable, then $\Sigma \wedge E$ is satisfiable
(ii) If $\Sigma_{GLB} \wedge E$ does not imply α, then $\Sigma \wedge E$ does not imply α
(iii) If $\Sigma_{LUB} \wedge E$ is not satisfiable, then $\Sigma \wedge E$ is not satisfiable
(iv) If $\Sigma_{LUB} \wedge E$ implies α, then $\Sigma \wedge E$ implies α.

Proof. Points (ii) and (iv) are stated in Proposition 11, and points (i) and (iii) directly follow from the definition of the bounds. □

Now we study more precise classes of formulas. When restricting the form of the observations and explanations, it sometimes becomes possible to perform abduction with the $\mathcal{C}$-LUB of a theory. In these cases, it thus becomes possible

to use the approximations of a theory for performing abduction in the same manner as we can use them for performing deduction.

We first consider the case of HORN-LUBs. We show that if the query is a conjunction of positive literals, the *positive* explanations that can be found with Σ are exactly the same as those that can be found with its HORN -LUB. The point is that when the background knowledge is Horn, a positive observation can be explained only by a conjunction of positive literals, as shown in the next lemma.

Lemma 1 (generalization of [RK87, Corollary 4]). *Let Σ be a Horn theory, α a positive formula and E a minimal explanation of α knowing Σ. Then E contains only positive literals.*

Proof. Assume, for sake of contradiction, that E contains $\neg x_i$ for a certain $x_i \in Var(\Sigma)\backslash Var(\alpha)$. Then there is a model m of Σ with $m \models \alpha$ and $m \models E$, thus in particular $m_i = 0$. We show that $\Sigma \wedge E\backslash\{\neg x_i\}$ implies α, and thus that E is not minimal ; for that purpose, we show that there cannot be a model m' of Σ with $m' \not\models \alpha$ and $m' \models E\backslash\{\neg x_i\}$. Indeed, if there is such an m', then $m'' = m \wedge m'$ must satisfy Σ by Proposition 1. But by construction, m'' satisfies E, since m and m' are equal over the variables in $E\backslash\{\neg x_i\}$ and $m_i = 0$, thus $m''_i = 0$; but m'' does not satisfy α, since $m' \not\models \alpha$, m'' assigns 0 to all the variables of α to which m' assigns 0, and α is positive. Thus m'' contradicts the fact that $\Sigma \wedge E$ implies α. □

Using Lemma 1, we can then show that performing abduction with the HORN-LUB of Σ is the same as with Σ itself when considering only positive observations and conjunctions of positive literals as explanations.

Proposition 14. *Let Σ be a theory, α a conjunction of positive literals, and H a subset of $Var(\Sigma)\backslash Var(\alpha)$. Let Σ_{LUB} be the* HORN-*LUB of Σ. If E is a minimal explanation of α knowing Σ_{LUB} over H, then E is a minimal explanation of α knowing Σ over H. Conversely, if E is a minimal* positive *explanation of α knowing Σ over H, then E is a minimal explanation of α knowing Σ_{LUB} over H.*

Proof. Let E be a minimal explanation of α knowing Σ_{LUB}. We know by Proposition 13 that $\Sigma \wedge E$ implies α. Since $\Sigma_{LUB} \wedge E$ is satisfiable, there is a model m of Σ_{LUB} with $m \models E$, thus there is a model m' of Σ with m' bitwise greater than or equal to m (by the construction of Proposition 2) ; since E is positive (Lemma 1), m' also satisfies E, thus $\Sigma \wedge E$ is satisfiable. There is only left to show that E is minimal knowing Σ. If it were not the case, then there would be a $E' \subset E$ with $\Sigma \wedge E'$ satisfiable, thus $\Sigma_{LUB} \wedge E'$ satisfiable by Proposition 13, and such that $\Sigma \wedge E'$ would imply α but $\Sigma_{LUB} \wedge E'$ would not (for otherwise E would not be minimal knowing Σ_{LUB}). Thus there would be a model m of Σ_{LUB} satisfying E' but not α, thus a model m' of Σ satisfying E' but not α (again by the construction of Proposition 2) ; thus $\Sigma \wedge E'$ would not imply α, contradiction.

Conversely, let E be a minimal positive explanation of α knowing Σ. We then know by Proposition 13 that $\Sigma_{LUB} \wedge E$ is satisfiable. We also know that $\Sigma_{LUB} \wedge E$ implies α by the same proof as with E' in the previous paragraph. Finally, by minimality of E for Σ the first part of the Proposition shows that there cannot be a $E' \subset E$ with E' an explanation knowing Σ_{LUB}, thus that E is minimal knowing Σ_{LUB}. □

Now we consider the case of BIJUNCTIVE-LUBs. Parallel to the restriction to conjunctions of positive literals for HORN-LUBs, we must now restrict to observations consisting of one clause of one or two literals. Now parallel to Lemma 1, we show in the next lemma that a minimal explanation of such an observation knowing a bijunctive theory must contain at most one literal.

Lemma 2. *Let Σ be a bijunctive theory, α a clause of one or two literals and E a minimal explanation of α knowing Σ. Then E contains zero or one literal.*

Proof. Since E is an explanation of α knowing Σ, it holds that Σ implies $C = (E \rightarrow \alpha)$; writing $E = \{\ell_{i_1}, \ldots, \ell_{i_p}\}$ (maybe empty) and $\alpha = \ell_{j_1} \vee \ell_{j_2}$ (the case $\alpha = \ell_{j_1}$ is similar), we get $C = (\ell_{i_1} \wedge \ldots \wedge \ell_{i_p} \rightarrow \ell_{j_1} \vee \ell_{j_2})$, i.e., $C = (\overline{\ell_{i_1}} \vee \cdots \vee \overline{\ell_{i_p}} \vee \ell_{j_1} \vee \ell_{j_2})$. Since Σ is bijunctive, there is a subclause C' of C consisting of two literals that is implied by Σ ; thus either (i) C' is a $\overline{\ell_{i_a}} \vee \overline{\ell_{i_b}}$, but this implies $\Sigma \rightarrow \overline{E}$, which contradicts the fact that $\Sigma \wedge E$ is satisfiable, or (ii) C' is $\ell_{j_1} \vee \ell_{j_2}$, but this means $\Sigma \rightarrow \alpha$ and thus E, being minimal, is empty, or finally (iii) C' is a $\overline{\ell_{i_a}} \vee \ell_{j_b}$, and thus Σ implies $\ell_{i_a} \rightarrow \ell_{j_b}$, which in turn implies $\ell_{i_a} \rightarrow \alpha$, and thus $E = \emptyset$ or $E = \{\ell_{i_a}\}$ is a minimal explanation. □

Using Lemma 2, we can now show that, parallel to Proposition 14, performing abduction with the BIJUNCTIVE-LUB of Σ is the same as with Σ itself when considering only clauses of one or two literals as observations and explanations containing zero or one literal.

Proposition 15. *Let Σ be a theory, α a clause of one or two literals, and H a subset of $Var(\Sigma) \backslash Var(\alpha)$. Let Σ_{LUB} be the* BIJUNCTIVE-*LUB of Σ. If E is a minimal explanation of α knowing Σ_{LUB} over H, then E is a minimal explanation of α knowing Σ over H. Conversely, if E is a minimal explanation containing one or zero literal of α knowing Σ over H, then E is a minimal explanation of α knowing Σ_{LUB} over H.*

Proof. Let E be a minimal explanation of α knowing Σ_{LUB} over H. Proposition 13 tells us that $\Sigma \wedge E$ implies α. Now by Lemma 2 we know that E contains at most one literal ; since Σ is satisfiable, if $E = \emptyset$ then $\Sigma \wedge E$ is satisfiable ; now if $E = \{\ell_{i_a}\}$, then the clause $(\overline{\ell_{i_a}})$ is not implied by Σ_{LUB} (since $\Sigma_{LUB} \wedge E$ is satisfiable) ; but Σ_{LUB} is logically equivalent to the conjunction of all the clauses of length 2 or less that are implied by Σ, thus Σ does not imply $(\overline{\ell_{i_a}})$ and we deduce that $\Sigma \wedge (\ell_{i_a})$ is satisfiable. Finally, if E is not minimal for Σ then the only possibility is that E contains one literal and $\emptyset$ is an explanation of α knowing Σ ; but this means that Σ implies α, and since α contains at most two literals we deduce that Σ_{LUB} implies α, i.e., that E is not minimal for Σ_{LUB}.

The converse is established as for Proposition 14. □

Finally, we consider abduction with AFFINE-LUBs. Unfortunately, in the framework for abduction defined here, no link other than those exhibited in Proposition 13 seems to exist between the explanations found with this LUB and those that can be found with the theory itself, even if we restrict ourselves to observations consisting of one variable and to explanations containing at most two literals. The next example illustrates that fact.

Example 4. Consider $\Sigma = \{001, 010, 100\}$ over the variables x_1, x_2, x_3. It is easily seen that the AFFINE-LUB of Σ is $(x_1 \oplus x_2 \oplus x_3 = 1)$. Now consider the observation $\alpha = x_3$ and the set of hypotheses $H = \{x_1, x_2\}$. Then $E = x_1 \wedge x_2$ is a minimal explanation of α knowing the AFFINE-LUB of Σ over H ; but $\Sigma \wedge E$ is not satisfiable.

The intuition behind this fact is that the relations between variables that are preserved between a theory and its AFFINE-LUB are the relations that can be expressed by a *linear equation* ; but explanations as defined here are meant to be *conjunctions* of literals, and this is intuitively why they do not behave well. We could define explanations in a more general form (for instance as any propositional formula over H), but this is out of scope here.

6 Conclusion

Computing...	(REVERSE-)HORN	BIJUNCTIVE	AFFINE
the $\mathcal{C}$-LUB	exp output	$\|R\|n^2$	$\|R\|n^3 + n^4$
one $\mathcal{C}$-GLB	$\|R\|^2 n^2$	$\|R\|^3 n + \|R\|^2 n^2$	$\|R\|^3 n + \|R\|^2 n^2$
a $\mathcal{C}$-maxGLB	NP-hard	NP-hard	subexponential

Table 1. Summary of results for computing approximations

We have settled the complexity of the main problems concerning the computation of Least Upper Bounds and Greatest Lower Bounds of relations into the main classes of formulas that allow efficient storage of knowledge and reasoning, namely the classes of Horn, reverse-Horn, bijunctive and affine formulas. The results are summarized in Table 1. It appears that computing one GLB is polynomial, while computing the LUB is polynomial only for the classes of bijunctive and affine formulas, and computing a GLB with the maximum number of models is NP-hard for Horn and bijunctive formulas while subexponential for affine.

Beside this, we have studied the semantics of reasoning with the bounds of a theory. As far as we know, the semantics of abduction with LUBs had not really been studied before. We have shown that assuming restrictions about the form of the observation and of the explanations, we can perform abduction with the HORN- or BIJUNCTIVE-LUB of a theory and get the same explanations as if

we had reasoned with the theory itself. These results show that assuming these restrictions, abduction can be performed with the bounds of a theory in the same manner as deduction.

This study of the semantics of abduction is of course incomplete, and this problem would certainly benefit further exploration. Also, an interesting problem is the one of computing Horn-renamable approximations of relations and studying their semantics ; Horn-renamable formulas are those formulas that we can transform into a Horn one by replacing some variables with their negation everywhere in the formula. This class includes Horn and reverse-Horn formulas as well as satisfiable bijunctive formulas, and is tractable for satisfiability and deduction as well. Thus although it is not stable by conjunction, which makes it less interesting for storing knowledge as a set of rules, it is certainly worth studying the complexity of computing approximations of relations into this class. As far as we know, the only study about Horn-renamable approximations can be found in [Bou98].

References

[Bou98] Boufkhad, Y., Algorithms for propositional KB approximation, in : *Proc. 15th National Conference on Artificial Intelligence (AAAI'98)* (Madison, USA), Menlo Park : AAAI Press / MIT Press (1998) 280-285

[CS00] Cadoli, M. and Scarcello, F., Semantical and computational aspects of Horn approximations, *Artificial Intelligence* **119 (1-2)** (2000) 1-17

[Cur84] Curtis, C.W., *Linear algebra. An introductory approach*, Springer-Verlag (1984)

[DP92] Dechter, R. and Pearl, J., Structure identification in relational data, *Artificial Intelligence* **58** (1992) 237-270

[EG95] Eiter, T. and Gottlob, G., The complexity of logic-based abduction, *Journal of the ACM* **42 (1)** (1995) 3–42

[GJ83] Garey, M.R. and Johnson, D.S., *Computers and intractability : a guide to the theory of NP-completeness* , San Francisco : W.H. Freeman and Company (1983)

[JYP88] Johnson, D.S., Yannakakis, M. and Papadimitriou, C.H., On generating all maximal independent sets, *Inform. Process. Lett.* **27 (3)** (1988) 119–123

[KKS95] Kautz, H., Kearns, M. and Selman, B., Horn approximations of empirical data, *Artificial Intelligence* **74** (1995) 129-145

[KPS93] Kavvadias, D., Papadimitriou, C.H. and Sideri, M., On Horn envelopes and hypergraph transversals (extended abstract), in : *Proc. 4th International Symposium on Algorithms And Computation (ISAAC'93)* (Honk Kong), Berlin : Springer Lecture Notes in Computer Science **762** (1993) 399-405

[KS98] Kavvadias, D. and Sideri, M., The inverse satisfiability problem, *SIAM J. Comput.* **28 (1)** (1998) 152–163

[KSS00] Kavvadias, D.J., Sideri, M. and Stavropoulos, E.C., Generating all maximal models of a Boolean expression, *Inform. Process. Lett.* **74** (2000) 157–162

[Kha95] Khardon, R., Translating between Horn representations and their characteristic models, *Journal of Artificial Intelligence Research* **3** (1995) 349–372

[KR96] Khardon, R. and Roth, D., Reasoning with models, *Artificial Intelligence* **87** (1996) 187-213

[RK87] Reiter, R. and de Kleer, J., Foundations of assumption-based truth maintenance systems : preliminary report, in : *Proc. 6th National Conefrence on Artificial Intelligence (AAAI'87)* (Seattle, USA), Menlo Park : AAAI Press / MIT Press (1987) 183–188

[Sch78] Schaefer, T.J., The complexity of satisfiability problems, in : *Proc. 10th Annual ACM Symposium on Theory Of Computing (STOC'78)* (San Diego, USA), New York : ACM Press (1978) 216-226

[SK96] Selman, B. and Kautz, H., Knowledge compilation and theory approximation, *Journal of the ACM* **43 (2)** (1996) 193-224

[SH90] Stearns, R.E. and Hunt III, H.B., Power indices and easier hard problems, *Math. Systems Theory* **23** (1990) 209–225

[Val95] del Val, A., An analysis of approximate knowledge compilation, in : *Proc. 14th International Joint Conference on Artificial Intelligence (IJCAI'95)* (Montréal, Canada), San Francisco : Morgan Kaufmann (1995) 830-836

[Val00] del Val, A., The complexity of restricted consequence finding and abduction, in : *Proc. 17th National Conference on Artificial Intelligence (AAAI'00)* (Austin, Texas), AAAI Press / MIT Press (2000) 337–342

[ZH02a] Zanuttini, B. and Hébrard, J.-J., A unified framework for structure identification, *Information Processing Letters* **81** (2002) 335-339

[Z02a] Zanuttini, B., Approximating propositional knowledge with affine formulas, to appear in : *Proc. 15th European Conference on Artificial Intelligence (ECAI'02)* (Lyon, France), (2002)

[Z02b] Zanuttini, B., Des classes polynomiales pour l'abduction en logique propositionnelle (in French), *Proc. 8èmes Journées nationales sur la résolution de problèmes NP-Complets (JNPC'02)* (Nice, France), available from the author (2002)