Electronic Colloquium on Computational Complexity, Report No. 119 (2005)

Preferred representations of Boolean relations

Nadia CREIGNOU*
LIF, UMR CNRS 6166
Université de la Méditerranée
163, avenue de Luminy
13 288 Marseille, France

Phokion Kolaitis †
IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120-6099, U.S.A.

Bruno ZANUTTINI‡
GREYC, UMR 6072
Université de Caen
Boulevard du Maréchal Juin
14032 Caen, France

September 15, 2005

Abstract

We introduce the notion of a *plain basis* for a co-clone in Post's lattice. Such a basis is a set of relations B such that every constraint C over a relation in the co-clone is logically equivalent to a conjunction of equalities and constraints over B and the same variables as C; this differs from the usual notion of a basis in that existential quantification of auxiliary variables is not allowed. We give such a basis for every co-clone and in particular for those in the infinite part of the lattice; it turns out that most of these bases correspond to sets of propositional clauses, thus providing a strong link between classes of formulas defined for CSP and CNF representations. We then show that a so-called *preferred representation* of a relation over one of its bases can be computed efficiently, as well as the minimal co-clone including a given relation, which solves some open structure identification problem as well as the open expressibility problem from database theory.

*E-mail: creignou@lif.univ-mrs.fr
†E-mail: kolaitis@cs.ucsc.edu
‡E-mail: zanutti@info.unicaen.fr

ISSN 1433-8092

1 Introduction

Constraints Satisfaction Problems (CSP) constitute a common and natural framework for formalizing computational problems in various areas of computer science: in particular Operational Research, Scheduling, Artificial Intelligence (AI) and Database theory (DB). Informally, given a set of variables each one with a set of allowed values, a CSP is a set of *constraints*, each one restricting the combinations of values for a subset of the variables to those in some *relation*. A solution for the CSP is an assignment of values to all the variables that respects all the constraints. Unfortunately, deciding the existence of a solution for a CSP is an NP-complete problem, and this motivated the study of restrictions on the allowed constraints in order to gain tractability.

We are interested here in restrictions on the *nature* of the constraints, i.e., on the relations that constraints are built upon. We also restrict ourselves to the Boolean case, where the domain of each variable is $\{0, 1\}$. A series of works beginning with Post's [Pos41] show that all the (finitary) Boolean relations can be classified according to their polymorphisms, i.e., the Boolean functions that close them in some precise sense, and that this classification is very helpful to the study of the computational complexity of many problems (see [CKS01] for a survey). According to this classification, Boolean relations are organized into *co-clones* which are themselves organized into the so-called Post's lattice.

Consequently, a deep understanding of Post's lattice is of fundamental importance from both the complexity-theoretic and the application points of view. In particular, in order to apply complexity results derived from Post's theory to many domains, it is very important to study the links between relations that satisfy some properties in Post's lattice and other representations of CSPs. To this aim, independently and simultaneously to our work, the notion of a *basis* of a co-clone has been studied by Böhler *et al.* [BRSV05], who exhibited bases of minimal arity for every co-clone. Informally, such a basis is a minimal set of relations from which every relation in the co-clone can be built using cartesian product, identification of variables and existential quantification of variables (equivalently, projection).

We study here a stronger notion of a basis. We call *plain basis* of a co-clone a minimal set of relations from which every relation in the co-clone can be built using only cartesian product and identification of variables. Thus every relation has what we call a *preferred* representation, as a conjunction of constraints over the basis and equality relations between variables. What is important here is that since existential quantification of variables is not allowed any more, the preferred representation of a CSP is itself a CSP. Whereas existentially quantified variables do not change the existence of a solution for a CSP, this is not the case for other computational problems, e.g., counting of solutions (see [CH96] and [BCCHV04]).

Our contribution is manifold. We exhibit a plain basis for every co-clone in Post's lattice; this is mainly done through the study of prime CNF representations of relations, and it turns out that most plain basis are sets of propositional clauses. Most of these bases were already known, but we complete them by bases for the co-clones in the infinite part of Post's lattice. We then give an efficient algorithm for computing the preferred representation of a relation in a given co-clone. More importantly, we give an efficient algorithm that decides the minimal co-clone of a relation, which is a new result and which does not seem to be allowed by the classical notion of a basis. Finally, we give an application of our results to a fundamental problem in database theory, namely that of deciding whether a given relation can be expressed by a given set of relations.

Our results can be seen from different points of view. From the complexity-theoretic one, they allow to study the complexity of problems according to properties of CNFs or of relations equivalently (in particular, this remark applies to complexity classification results in which the class of IHSB formulas plays a role as in [CZ04], [CKS01, Theorem 6.5] and [KST97]). From the AI point of view, they answer the structure identification [DP92], or inverse-SAT [KS98] problem for IHSB formulas. Finally, from the DB point of view, they answer, as previously evoked, the fundamental problem of expressivity of relations.

2 Preliminaries

2.1 Post's lattice

A *Boolean function* is an application $f : \{0,1\}^n \mapsto \{0,1\}$. The integer n is called the *arity* of f. If f is n-ary and $g_1, \dots, g_n$ are all m-ary Boolean functions, then the composition $f(g_1, \dots, g_n)$ has arity m as well and its value on $(a_1, \dots, a_m)$ is $f(g_1(a_1, \dots, a_m), \dots, g_n(a_1, \dots, a_m))$. For $n \geq m \geq 1$ the *projection function* $\pi_{n,m}$ is defined by $\pi_{n,m}(x_1, \dots, x_n) = x_m$. A (Boolean) *clone* is a set of Boolean functions closed under composition and containing all coordinate projections of all arities. For classical references we refer the reader to [Pip97] and [PK79].

The clones form a lattice under set inclusion, which is now referred to as Post's lattice [Pos41]. The description of this lattice is facilitated by an important property called *duality.* Informally, it reflects the fact that the entire theory remains unchanged if we interchange the roles of 0 and 1. This process assigns to each Boolean function f another function, called the *dual* of f and defined by $dual(f)(a_1, \dots a_n) = \overline{f(\overline{a_1}, \dots, \overline{a_n})}$. The now well-known Post's lattice is presented in Figure 1 [1]. Note that we use the notation of clones and co-clones developed in [BCRV03, BCRV04].

It has been proved that a Galois correspondence can be established between clones in Post's lattice and maximal classes of relations that are closed under the functions in the clone. More precisely, let R be an m-ary relation, that is a set $R \subseteq \{0,1\}^m$, and let f be a Boolean function of arity n. Then f is said to be a *closure property* of R if when f is applied coordinate-wise to m vectors in R (not necessarily all different), the resulting vector is again in R. In this case, we say that R is *closed under* f or that f is a *polymorphism* of R. We denote the set of all polymorphisms of R by $\mathrm{Pol}(R)$, and for a set S of Boolean relations

[1] The authors are grateful to Steffen Reith who provided them with the figure.

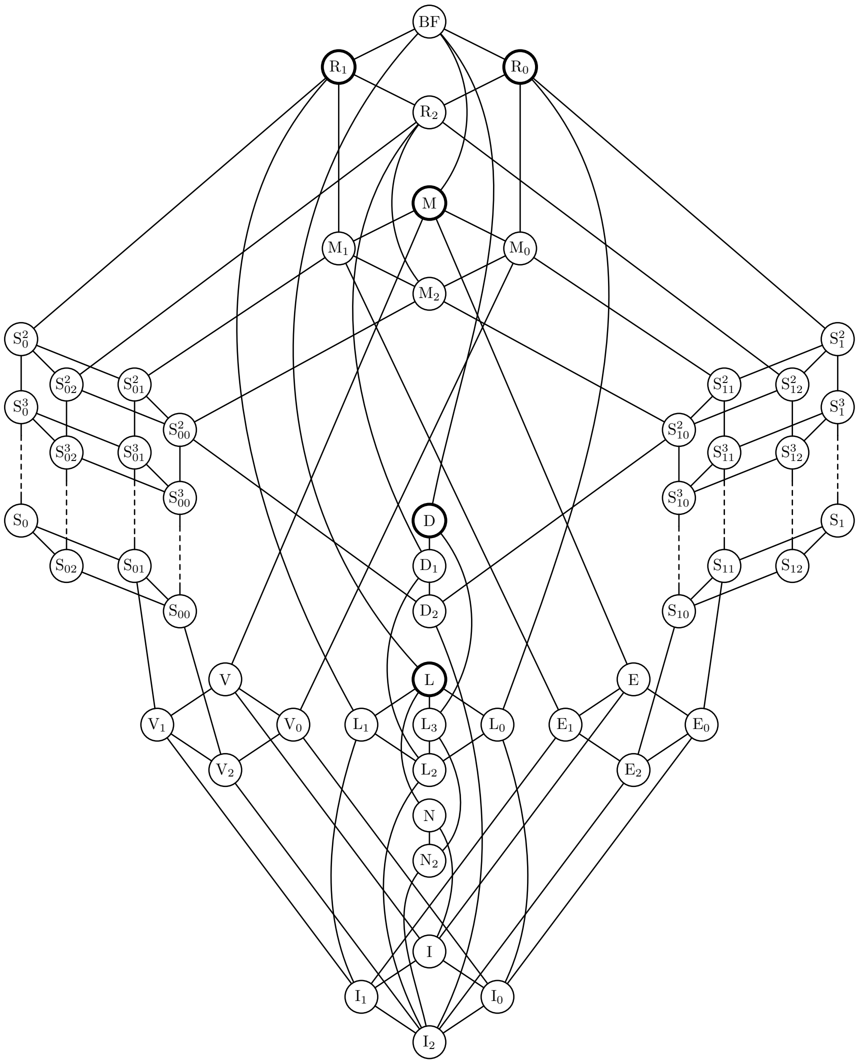

Figure 1: Lattice of all Boolean clones

we define $\mathrm{Pol}(S)$ to be $\bigcap_{R\in S}\mathrm{Pol}(R)$. Importantly, as is easily seen $\mathrm{Pol}(S)$ is a clone.

Conversely, if T is a set of Boolean functions, then $\mathrm{Inv}(T)$ is defined to be the set of all relations which are preserved by every function from T. It turns out that sets $\mathrm{Inv}(T)$ have particular common properties: they contain the equivalence relation, and each such set is closed under Cartesian product, projection and identification of variables; thus they are *co-clones*. The functions Pol and Inv induce an anti-isomorphism between the lattice of clones and the lattice of co-clones ([Gei68, PK79, Pip97], see also [BCRV04] for a survey). The co-clone corresponding to clone Cl is denoted by ICl (e.g., IE_2 for the co-clone corresponding to clone E_2).

2.2 Bases and plain bases

A *constraint* over an n-ary relation R is the application of R to some sequence V of n variables (maybe with repetitions), written $R(V)$. A *solution* for $R(V)$ is an assignment m of values to every variable in V such that when seen as a Boolean vector, m is in R. These notions are straightforwardly extended to conjunctions of constraints, and two conjunctions of constraints are said to be *logically equivalent* if their sets of solutions are the same.

The classical definition of a *basis* for a co-clone is the following. Given a co-clone ICl, a set of relations $B \subseteq ICl$ is called a *basis* for ICl if every constraint $R(x_1,\ldots,x_n)$ with $R \in ICl$ is logically equivalent to $\exists y_1 \ldots y_m \mathcal{C}$ for some set of variables $\{y_1,\ldots,y_m\}$ disjoint from $\{x_1,\ldots,x_n\}$ and some conjunction of constraints $\mathcal{C}$ over $\{x_1,\ldots,x_n\} \cup \{y_1,\ldots,y_m\}$ using only relations in B. Note that since the equality relation is in every co-clone of Post's lattice, this formulation is equivalent to the usual one (R can be obtained from relations in B using cartesian product, identification of variables and existential quantification of variables).

We introduce a new, stronger notion of a basis.

Definition 1 (plain basis) *Let ICl be a co-clone in Post's lattice. A set of relations $B \subseteq ICl$ is called a* plain basis *for ICl if every constraint $R(x_1,\ldots,x_n)$ with $R \in ICl$ is logically equivalent to $\mathcal{C}$ for some conjunction of constraints $\mathcal{C}$ using only relations in B and applied only to variables $x_1,\ldots,x_n$.*

2.3 Conjunctive Normal Form

A *literal* is either a variable x (*positive* literal) or the negation $\neg x$ of one (*negative* literal). A *clause* is a finite disjunction of literals of the form $C = (\ell_1 \vee \cdots \vee \ell_k)$ ($k \geq 0$). A propositional formula is said to be in *Conjunctive Normal Form* (CNF) if it is written as a conjunction of clauses.

If V is a set of variables and φ is a CNF over a subset of V, a *model* of φ over V is an assignment to the variables in V which satisfies φ with the usual rules for the connectives. A formula is said to be *satisfiable* if it has at least one model. If φ_1 and φ_2 are two propositional formulas over sets of variables

V_1 and V_2, respectively, φ_1 is said to *(logically) entail* φ_2 if every model of φ_1 over $V_1 \cup V_2$ is a model of φ_2, and φ_1, φ_2 are said to be *(logically) equivalent*, written $\varphi \equiv \varphi'$, if their sets of models over $V_1 \cup V_2$ are equal.

We will establish the correspondence between relations and formulas in CNF in the following manner. Let R be a n-ary Boolean relation. We see R as a set of assignments to variables $x_1, x_2, \dots, x_n$, i.e., we see any vector $m = (m_1, \dots m_n) \in R$ as the assignment of value m_i to variable x_i for all $i \in \{1, \dots, n\}$. Then a propositional formula φ over variables $x_1, \dots, x_n$ is said to *represent* R if R is its sets of models.

An important notion that we will use many times in the paper is that of a *prime implicate*. Let φ be a propositional formula. A clause $C = (\ell_1 \vee \cdots \vee \ell_k)$ is said to be a *prime implicate* of φ if φ entails C but entails no proper subclause of C, i.e., if φ entails C but for no $i \in \{1, \dots, k\}$, φ entails the clause $(\ell_1 \vee \cdots \vee \ell_{k-1} \vee \ell_{k+1} \vee \cdots \vee \ell_k)$. A CNF φ is said to be *prime* if all its clauses are prime implicates of it. We will often use the following remark:

Remark 2 *$C = (\ell_1 \vee \cdots \vee \ell_k)$ is a prime implicate of φ if and only if there is no model of φ that assigns 0 to every literal ℓ_i but for every $i_0 \in \{1, \dots, n\}$ there is a model of φ that assigns 0 to every literal ℓ_i except for ℓ_{i_0}.*

Roughly speaking, in the terms introduced here the aim of this paper is to give a complete characterization of Boolean co-clones by syntactic properties of prime CNF formulas representing the relations.

3 Infinite part of Post's lattice

As evoked in the introduction, plain bases for most co-clones were already exhibited, mainly when studying the *structure identification* problem [DP92, KS98, ZH02]. Nevertheless, there was no plain basis exhibited yet for an important class of co-clones, namely the co-clones $IS_{10}^n, IS_{11}^n \dots$ and the dual ones. These clones constitute the (countably) infinite part of Post's lattice.

These co-clones are presented in Table 1 together with bases of the corresponding clones, i.e., sets of functions from which every function in the clone can be obtained by composition. Throughout the paper, for $n \geq 1$, h_n denotes the $n+1$-ary function defined by $h_n(x_1, \dots, x_{n+1}) = \bigvee_{i=1}^{n+1} x_1 \wedge \cdots \wedge x_{i-1} \wedge x_{i+1} \wedge \cdots \wedge x_{n+1}$, c_0 (resp. c_1) denotes the 0-ary constant function 0 (resp. 1) and imp denotes binary implication, defined by $\mathrm{imp}(x, y) = x \rightarrow y$.

What we show in this section is that these co-clones correspond to classes of *IHSB* CNF formulas, which were introduced as a special class of Horn formulas for which satisfiability is in the parallel complexity class NC (see [GHR95] and [CKS01, Theorem 6.5]) and that have rather good approximability properties (see [KST97]). Thus these characterizations yield plain bases for the corresponding co-clones since every clause can be seen as a relation. We rephrase the results of this section into terms of plain bases in Section 4, but we let them stated in terms of CNF formulas here because they answer the structure identification problem for classes of IHSB formulas.

Co-clone	Base of clone	Co-clone	Base of clone
IS_0^n	$\{\text{imp}, \text{dual}(h_n)\}$	IS_1^n	$\{x \wedge \overline{y}, h_n\}$
IS_0	$\{\text{imp}\}$	IS_1	$\{x \wedge \overline{y}\}$
IS_{02}^n	$\{x \vee (y \wedge \overline{z}), \text{dual}(h_n)\}$	IS_{12}^n	$\{x \wedge (y \vee \overline{z}), h_n\}$
IS_{02}	$\{x \vee (y \wedge \overline{z})\}$	IS_{12}	$\{x \wedge (y \vee \overline{z}\}$
IS_{01}^n	$\{\text{dual}(h_n), c_1\}$	IS_{11}^n	$\{h_n, \ c_0\}$
IS_{01}	$\{x \vee (y \wedge z), c_1\}$	IS_{11}	$\{x \wedge (y \vee z), c_0\}$
IS_{00}^n	$\{x \vee (y \wedge z), \text{dual}(h_n)\}$	IS_{10}^n	$\{x \wedge (y \vee z), h_n\}$
IS_{00}	$\{x \vee (y \wedge z)\}$	IS_{10}	$\{x \wedge (y \vee z)\}$

Table 1: Co-clones in the infinite part of Post's lattice and bases of corresponding clones (where, e.g., $x \vee (y \wedge z)$ denotes the function $(x, y, z) \mapsto x \vee (y \wedge z)$)

Taking advantage of duality in Post's lattice, we focus our attention to the right-side of the lattice. Consequently, we define IHSB$-$ formulas (IHSB+ formulas are defined dually):

Definition 3 (IHSB$-$, IHSB$-^n$) *A clause is said to be* IHSB$-$ *(for* Implicative Hitting Set-Bounded$-$*) if it is of one of the following types:* (x_i)*,* $(\neg x_{i_1} \vee x_{i_2})$*, or* $(\neg x_{i_1} \vee \cdots \vee \neg x_{i_k})$ *for some* $k \geq 0$*. For* $n \geq 2$*, an IHSB$-$ clause is said to be* of width n *(written IHSB$-^n$) if it involves at most* n *literals. A formula in CNF is said to be IHSB$-$ (resp. IHSB$-^n$) if all its clauses are IHSB$-$ (resp. IHSB$-^n$).*

We can now establish the desired correspondences. The following proposition is the main one, since the other ones will use it and its proof.

Proposition 4 (IHSB$-^n$ vs. IS_{10}^n**)** *A relation is in* IS_{10}^n *if and only if every prime CNF representing it is IHSB$-^n$.*

Proof

[$IS_{10}^n \Rightarrow IHSB-^n$] Let R be a relation in IS_{10}^n and let φ be a prime CNF representing it. Since in Post's lattice $E_2 \subset S_{10}^n$ holds we know that φ is Horn, thus each one of its clauses contains either zero or one positive literal (resp. called *negative* and *definite Horn* clauses). We show that all these clauses are IHSB$-^n$.

[Negative clauses] First assume for sake of contradiction that φ contains a negative clause that is "too wide", i.e., a clause C of the form $(\neg x_{i_1} \vee \cdots \vee \neg x_{i_m})$ with $m > n$. Since C is a prime implicate of φ and φ represents R, by Remark 2 there are $n+1$ vectors $m_1, \ldots, m_{n+1} \in R$ whose projections onto $\{x_{i_1}, \ldots, x_{i_m}\}$

are:

$$\begin{array}{lcccccccccccc}
 & & x_{i_1} & x_{i_2} & x_{i_3} & \dots & x_{i_{n-1}} & x_{i_n} & x_{i_{n+1}} & x_{i_{n+2}} & \dots & x_{i_m} \\
m_1 & = & 0 & 1 & 1 & \dots & 1 & 1 & 1 & 1 & \dots & 1 \\
m_2 & = & 1 & 0 & 1 & \dots & 1 & 1 & 1 & 1 & \dots & 1 \\
 & & & & & & \dots & & & & & \\
m_n & = & 1 & 1 & 1 & \dots & 1 & 0 & 1 & 1 & \dots & 1 \\
m_{n+1} & = & 1 & 1 & 1 & \dots & 1 & 1 & 0 & 1 & \dots & 1
\end{array}$$

Then it is easily seen that h_n applied coordinate-wise to these $n+1$ vectors yields the vector $1\dots1$, which is not in R because it falsifies C; thus R is not closed under h_n, which contradicts the hypothesis.

[Definite Horn clauses] Now assume, still for sake of contradiction, that φ contains a clause of the form $C = (x_1 \vee \neg x_2 \vee \dots \vee \neg x_m)$ with $m > 2$. Once again, since C is prime there are at least three vectors in R as below:

$$\begin{array}{lccccccc}
 & & x_{i_1} & x_{i_2} & x_{i_3} & x_{i_4} & \dots & x_{i_m} \\
m_1 & = & 1 & 1 & 1 & 1 & \dots & 1 \\
m_2 & = & 0 & 0 & 1 & 1 & \dots & 1 \\
m_3 & = & 0 & 1 & 0 & 1 & \dots & 1
\end{array}$$

Applying the function $(x, y, z) \mapsto x \wedge (y \vee z)$ to these three vectors coordinate-wise yields the vector $01\dots1$, which does not satisfy C, which contradicts the hypothesis again. Finally, every clause in φ is either negative with at most n literals or definite Horn with at most 2 literals, which shows that φ is IHSB$-^n$.

[$IHSB-^n \Rightarrow IS_{10}^n$] We now show that if R is a relation and φ is an IHSB$-^n$ prime CNF representing R, then R is in IS_{10}^n. This will imply the claim since there is always at least one prime CNF representing a relation. Since co-clones are closed under conjunction and IS_{10}^n is defined to be $\mathrm{Inv}(S_{10}^n)$, we only have to show that any IHSB$-^n$ clause is closed under the functions in a basis of S_{10}^n; we use the base in Table 1, i.e., $\{(x, y, z) \mapsto x \wedge (y \vee z), h_n\}$. This obviously holds for the clauses containing zero or one literal, thus we consider clauses of size at least 2.

[Definite Horn clauses] The only definite Horn clauses to consider are clauses of the form $(\neg x_{i_1} \vee x_{i_2})$. Let C be such a clause; the set of models of C is $\{00, 01, 11\}$. Assume if it not closed under $(x, y, z) \mapsto x \wedge (y \vee z)$; then this operation yields 10 (the only missing vector) on some m_1, m_2, m_3; because of x we thus have $m_1[x_1] = 1$ and thus $m_1 = 11$, but because of y and z we also have $m_2[x_1] = 1$ (without loss of generality). Thus $m_2 = 11$ and $m_1 \wedge (m_2 \vee m_3) = 11$, which contradicts our hypothesis. Similarly, if the set of models of C is not closed under h_n for some n, then it would contain $n+1$ vectors $m_1, \dots, m_{n+1}$ such that $h_n(m_1, \dots, m_{n+1}) = 10$; thus at least n of these vectors would assign 1 to x_{i_1}, i.e., would be 11; thus $h_n(m_1, \dots, m_{n+1})$ would be 10, again contradicting our hypothesis.

[Negative clauses] Finally, let $C = (\neg x_{i_1} \vee \dots \vee \neg x_{i_k})$ with $k \leq n$. The set of models of C is $\{0,1\}^k \backslash 11\dots1$. We first show that $m_1 \wedge (m_2 \vee m_3)$ is always different from $11\dots1$, which will show closure under $(x, y, z) \mapsto x \wedge (y \vee z)$; indeed, as remarked above, $x \wedge (y \vee z) = 1$ implies $x = 1$, thus $m_1 \wedge (m_2 \vee m_3) = 11\dots1$ implies $m_1 = 11\dots1$, which is not a model of C. Finally, let us consider h_n. If the set of models of C is not closed under it then we must have $n+1$ vectors such that $h_n(m_1, \dots, m_{n+1}) = 11\dots1$; thus among them at least n must assign 1 to x_{i_1}, n (maybe different) must assign 1 to x_{i_2} and so on. Since we only have $k \leq n < n+1$ variables, it follows from the pigeonhole principle that at least one m_i is $11\dots1$, which is not a model of C. □

Proposition 5 (IHSB− vs. IS_{10}) *A relation is in IS_{10} if and only if every prime CNF representing it is IHSB−.*

Proof Since a formula is finite, every IHSB− CNF formula is IHSB−^{n} for some n; thus if there is an IHSB− (prime) CNF formula representing a relation R, then R is in IS_{10}^n for some n, and from $S_{10} \subset S_{10}^n$ it follows that R is in IS_{10}. As for the converse, if R is a relation in IS_{10} and φ is a prime CNF representing it, closure of R under $(x, y, z) \mapsto x \wedge (y \vee z)$ is enough to conclude that every definite Horn clause of φ contains at most 2 literals (see the proof of Proposition 4), and thus φ is IHSB−. □

Proposition 6 (IS_{11}^n, IS_{11}) *A relation is in IS_{11}^n (resp. IS_{11}) if and only if every prime CNF representing it is IHSB−^{n} (resp. IHSB−) and does not contain any clause of the form (x_{i_1}).*

Proof From $S_{10}^n \subset S_{11}^n$ it follows that every prime CNF representing a relation in IS_{11}^n is IHSB−^{n}, and because IS_{11}^n has a basis containing c_0 no CNF representing such a relation can contain a unary positive clause. The converse follows from the fact that (i) the set of models of every IHSB−^{n} clause is closed under h_n (Proposition 4), (ii) the only IHSB−^{n} clauses whose set of models is not closed under c_0 are those of the form (x_{i_1}) and finally (iii) $\{h_n, c_0\}$ is a basis of S_{11}^n.

The proof is similar for IS_{11}. □

Proposition 7 (IS_{12}^n, IS_{12}) *A relation is in IS_{12}^n (resp. IS_{12}) if and only if for every prime CNF φ representing it, φ is IHSB−^{n} (resp. IHSB−) and for every two variables x_{i_1}, x_{i_2}, if φ contains the clause $(\neg x_{i_1} \vee x_{i_2})$ then it entails the clause $(x_{i_1} \vee \neg x_{i_2})$.*

Proof Since $S_{10}^n \subset S_{12}^n$ holds, every prime CNF φ representing a relation R in IS_{12}^n is IHSB−^{n}. For sake of contradiction, assume that φ contains the clause $C = (x_1 \vee \neg x_2)$ but does not entail the clause $C' = (\neg x_1 \vee x_2)$. Then since φ is prime, by Remark 2 there are two models m_1, m_2 of φ such that m_1 assigns 1 to x_1 and 1 to x_2, and m_2 assigns 0 to x_1 and 0 to x_2. On the other hand, since φ does not entail C', there is a model m' of φ that does not satisfy C', i.e., that assigns 1 to x_1 and 0 to x_2. Now applying the function $(x, y, z) \mapsto x \wedge (y \vee \overline{z})$ to (m_1, m_2, m') yields a vector which assigns 0 to x_1 and 1 to x_2; thus this vector is not a model of C, which contradicts the fact that R is closed under $(x, y, z) \mapsto x \wedge (y \vee \overline{z})$.

The converse is easily shown as usual by (i) showing closure of clauses of the form (x_1) and $(\neg x_{i_1} \vee \cdots \vee \neg x_{i_k})$ under $(x, y, z) \mapsto x \wedge (y \vee \overline{z})$ and under h_n $(k \leq n)$ (ii) remarking that since the presence of $(\neg x_{i_1} \vee x_{i_2})$ implies that φ entails $(x_{i_1} \vee \neg x_{i_2})$, φ is logically equivalent to a conjunction of clauses of the previous forms and of equality constraints, because $(\neg x_{i_1} \vee x_{i_2}) \wedge (x_{i_1} \vee \neg x_{i_2}) \equiv (x_{i_1} = x_{i_2})$. $\square$

Proposition 8 (IS_1^n, IS_1) *A relation is in IS_1^n (resp. IS_1) if and only if for every prime CNF φ representing it, φ is IHSB$-^n$ (resp. IHSB$-$), φ does not contain any clause of the form (x_{i_1}), and for every two variables x_{i_1}, x_{i_2}, if φ contains the clause $(\neg x_{i_1} \vee x_{i_2})$ then it entails the clause $(x_{i_1} \vee \neg x_{i_2})$.*

Proof If R is in IS_1^n (resp. IS_1), then the result follows from Propositions 6 and 7 and the inclusion $S_{11}^n \bigcap S_{12}^n \subseteq S_1^n$ (resp. $S_{11} \bigcap S_{12} \subseteq S_1$) in Post's lattice. The converse is shown as usual, with the same remark for clauses of the form $(\neg x_{i_1} \vee x_{i_2})$ as in the proof of Proposition 7. $\square$

4 Complete list of plain bases and algorithms

In this section, we summarize our results and those obtained in the literature in order to give a plain basis for every co-clone in Post's lattice, then introduce the notion of a *preferred representation* of a relation with respect to a co-clone, and finally give efficient algorithms for computing this representation and for deciding the minimal co-clone of a relation.

4.1 Plain bases

Table 2 gives a plain basis for every co-clone. In this table, when possible we denote relations by clauses that represent them; e.g., $(\neg x \vee y)$ denotes the binary relation $\{00, 01, 11\}$. The positive clause of width k, $(x_1 \vee \cdots \vee x_k)$, is denoted by P_k, and similarly the negative clause of width k is denoted N_k. We use the same kind of notation for relations that correspond to linear equations, and finally use two special notations: Eq denotes the binary equality relation $\{00, 11\}$ and $Compl_{k,\ell}$ denotes the $(k+\ell)$-ary relation represented by the conjunction of clauses $(x_1 \vee \cdots \vee x_k \vee \neg y_1 \vee \cdots \vee \neg y_\ell) \wedge (\neg x_1 \vee \cdots \vee \neg x_k \vee y_1 \vee \cdots \vee y_\ell)$, i.e., the complementive relation $\{0,1\}^{k+\ell} \setminus \{0 \ldots 01 \ldots 1, 1 \ldots 10 \ldots 0\}$. The last column gives the usual name given to the property satisfied by each clause, equation or relation in the basis.

We summarize the proofs and relevant references in the following proposition. The proofs are given as consequences of the bases of each clone (see [BCRV03, Figure 1]).

Proposition 9 (plain bases) *Each line in Table 2 gives a plain basis for the corresponding co-clone.*

	Plain basis	Property
IBF	$\{Eq\}$	only equalities
IR_0	$\{Eq, (\neg x)\}$	neg^1
IR_1	$\{Eq, (x)\}$	pos^1
IR_2	$\{Eq, (\neg x), (x)\}$	unary
IM	$\{(\neg x \vee y)\}$	implicative
IM_0	$\{(\neg x), (\neg x \vee y)\}$	implicative or pos^1
IM_1	$\{(x), (\neg x \vee y)\}$	implicative or neg^1
IM_2	$\{(x), (\neg x), (\neg x \vee y)\}$	implicative or unary
IS_0^n	$\{Eq\} \cup \{P_k \mid k \leq n\}$	pos^n
IS_0	$\{Eq\} \cup \{P_k \mid k \in \mathbb{N}\}$	pos.
IS_1^n	$\{Eq\} \cup \{N_k \mid k \leq n\}$	neg^n
IS_1	$\{Eq\} \cup \{N_k \mid k \in \mathbb{N}\}$	neg.
IS_{02}^n	$\{Eq, (\neg x)\} \cup \{P_k \mid k \leq n\}$	neg^1 or pos^n
IS_{02}	$\{Eq, (\neg x)\} \cup \{P_k \mid k \in \mathbb{N}\}$	neg^1 or positive
IS_{12}^n	$\{Eq, (x)\} \cup \{N_k \mid k \leq n\}$	pos^1 or neg^n
IS_{12}	$\{Eq, (x)\} \cup \{N_k \mid k \in \mathbb{N}\}$	pos^1 or negative
IS_{01}^n	$\{(\neg x \vee y)\} \cup \{P_k \mid k \leq n\}$	implicative or pos^n
IS_{01}	$\{(\neg x \vee y)\} \cup \{P_k \mid k \in \mathbb{N}\}$	implicative or positive
IS_{11}^n	$\{(\neg x \vee y)\} \cup \{N_k \mid k \leq n\}$	implicative or neg^n
IS_{11}	$\{(\neg x \vee y)\} \cup \{N_k \mid k \in \mathbb{N}\}$	implicative or negative
IS_{00}^n	$\{(\neg x), (\neg x \vee y)\} \cup \{P_k \mid k \leq n\}$	IHSB+^n
IS_{00}	$\{(\neg x), (\neg x \vee y)\} \cup \{P_k \mid k \in \mathbb{N}\}$	IHSB+
IS_{10}^n	$\{(x), (\neg x \vee y)\} \cup \{N_k \mid k \leq n\}$	$\text{IHSB}-^n$
IS_{10}	$\{(x), (\neg x \vee y)\} \cup \{N_k \mid k \in \mathbb{N}\}$	IHSB−
ID	$\{(x \oplus y = c) \mid c \in \{0, 1\}\}$	affine of width exactly 2
ID_1	$\{(x = c) \mid c \in \{0, 1\}\} \cup \{(x \oplus y = c) \mid c \in \{0, 1\}\}$	affine of width 2
ID_2	$\{(x), (\neg x), (x \vee y), (\neg x \vee y), (\neg x \vee \neg y)\}$	bijunctive
IL	$\{(x_1 \oplus \cdots \oplus x_k = 0) \mid k \text{ even}\}$	even homogeneous linear equation
IL_0	$\{(x_1 \oplus \cdots \oplus x_k = 0) \mid k \in \mathbb{N}\}$	homogeneous linear equation
IL_1	$\{(x_1 \oplus \cdots \oplus x_k = c) \mid k \in \mathbb{N}, c = k \bmod 2\}$	1-valid linear equation
IL_2	$\{(x_1 \oplus \cdots \oplus x_k = c) \mid k \in \mathbb{N}, c \in \{0, 1\}\}$	linear equation
IL_3	$\{(x_1 \oplus \cdots \oplus x_k = c) \mid k \text{ even}, c \in \{0, 1\}\}$	even linear equation
IV	$\{(x_1 \vee \cdots \vee x_k \vee \neg y) \mid k \geq 1\}$	definite dual Horn and not neg^1
IV_0	$\{(x_1 \vee \cdots \vee x_k \vee \neg y) \mid k \in \mathbb{N}\}$	definite dual Horn
IV_1	$\{P_k \mid k \in \mathbb{N}\} \cup \{(x_1 \vee \cdots \vee x_k \vee \neg y) \mid k \geq 1\}$	dual Horn and not neg^1
IV_2	$\{P_k \mid k \in \mathbb{N}\} \cup \{(x_1 \vee \cdots \vee x_k \vee \neg y) \mid k \in \mathbb{N}\}$	dual Horn
IE	$\{(\neg x_1 \vee \cdots \vee \neg x_k \vee y) \mid k \geq 1\}$	definite Horn and not pos^1
IE_0	$\{N_k \mid k \in \mathbb{N}\} \cup \{(\neg x_1 \vee \cdots \vee \neg x_k \vee y) \mid k \geq 1\}$	Horn and not pos^1
IE_1	$\{(\neg x_1 \vee \cdots \vee \neg x_k \vee y) \mid k \in \mathbb{N}\}$	definite Horn
IE_2	$\{N_k \mid k \in \mathbb{N}\} \cup \{(\neg x_1 \vee \cdots \vee \neg x_k \vee y) \mid k \in \mathbb{N}\}$	Horn
IN	$\{Compl_{k,\ell} \mid k, \ell \geq 1\}$	complementive, 0- and 1-valid
IN_2	$\{Compl_{k,\ell} \mid k, \ell \in \mathbb{N}\}$	complementive
II	$\{(x_1 \vee \cdots \vee x_k \vee \neg y_1 \vee \cdots \vee \neg y_\ell) \mid k, \ell \geq 1\}$	0- and 1-valid
II_0	$\{(x_1 \vee \cdots \vee x_k \vee \neg y_1 \vee \cdots \vee \neg y_\ell) \mid k \in \mathbb{N}, \ell \geq 1\}$	0-valid
II_1	$\{(x_1 \vee \cdots \vee x_k \vee \neg y_1 \vee \cdots \vee \neg y_\ell) \mid k \geq 1, \ell \in \mathbb{N}\}$	1-valid
II_2	$\{(x_1 \vee \cdots \vee x_k \vee \neg y_1 \vee \cdots \vee \neg y_\ell) \mid k, \ell \in \mathbb{N}\}$	any clause

Table 2: CNF properties corresponding to co-clones: "neg^n" means "negative and containing at most n literals", and similarly for "pos^n"

Proof We proceed from the largest co-clone to the smallest one. Since the proofs for co-clones ICl, ICl_0, ICl_1 follow straightforwardly from the proofs for co-clone ICl_2, we only consider the latter.

$[II, II_c]$ Obviously, any relation can be represented by a CNF formula and conversely, any relation is stable under identity.

$[IN, IN_2]$ Obviously again, any complementive relation can be represented by a CNF containing the clause $(\neg x_1 \vee \cdots \vee \neg x_k \vee y_1 \vee \cdots \vee y_\ell)$ as soon as it contains $(x_1 \vee \cdots \vee x_k \vee \neg y_1 \vee \cdots \vee \neg y_\ell)$; grouping these clauses two by two in the CNF yields a conjunction of $Compl_{k,\ell}$ relations, and conversely the set of models of such a conjunction is complementive.

$[IE, IE_c, IV, IV_c, IL, IL_c, ID, ID_c]$ We refer the reader to [Sch78] and to [ZH02] for the proofs for co-clones IE_2, IV_2, IL_2, ID_2.

$[IS_{cd}, IS_c]$ The proofs follow immediately from our results in Section 3; for plain bases containing Eq, the result follows by grouping clauses $(\neg x \vee y)$ and $(x \vee \neg y)$ two by two.

$[IM, IM_c]$ The proof for IM_2 follows from the inclusions $IM_2 \subseteq IS_{00}^2, IS_{10}^2$ in one direction, and from stability of the clauses in the plain basis under *and* and *or* in the other direction.

$[IR_c, IBF]$ The result for IR_2 follows in one direction from the inclusions $IR_2 \subseteq IM_2, ID_1$, because clauses $(\neg x \vee y)$ of the plain basis of IM_2 are not in ID_1 while unary clauses are (equation $(x = 1)$ is equivalent to clause (x), and equation $(x = 0)$ is equivalent to clause $(\neg x)$). The proof in the other direction follows from stability of unary clauses under both *or* and $(x, y, z) \mapsto x \wedge (y \oplus z \oplus 1)$.
$\square$

4.2 Preferred representations

As we pointed out in the introduction, the main advantage of our approach as compared to Böhler *et al.*'s [BRSV05] is that our notion of plain basis allows to derive efficient algorithms for two algorithmical purposes: given a relation, (i) compute a representation of it in a given plain basis, and (ii) find out the minimal co-clone containing it.

Definition 10 (preferred representation) *Let ICl be a co-clone, B a plain basis for it, and let R be a relation in ICl. Then a conjunction of constraints φ is called a* preferred representation *for R with respect to ICl and B if φ is logically equivalent to R and every constraint in φ is built upon a relation in B.*

Consider for instance the ternary relation $R = \{000, 111\}$. It is easily seen that this relation is closed under all polymorphisms, and thus R is in every co-clone. Then a preferred representation for it with respect to co-clone IBF and its plain basis of Table 2 is the conjunction of binary equalities $Eq(x_1, x_2) \wedge Eq(x_2, x_3)$. Note that the conjunction $Eq(x_1, x_2) \wedge Eq(x_2, x_3) \wedge Eq(x_1, x_3)$ is also a preferred representation for R with respect to IBF. Since R is also in

IL_2, it also has a preferred representation with respect to it, e.g., $(x_1 \oplus x_2 = 0) \wedge (x_1 \oplus x_3 = 0)$. As for co-clone IN_2, R has $Compl_{1,1}(x_1, x_2) \wedge Compl_{1,1}(x_1, x_3)$ as a preferred representation, and as for co-clone ID_2 it has $(\neg x_1 \vee x_2) \wedge (\neg x_2 \vee x_3) \wedge (\neg x_3 \vee x_1)$.

The notion of *preferred representation* is very important for structure identification purposes. Structure identification [DP92] is the problem of deciding, for a fixed class of formulas C, whether a given relation R can be represented by a formula in C, and to compute such a formula if the answer is affirmative. As Dechter and Pearl discuss it, this problem is a fundamental one for knowledge acquisition processes, where the aim is in general to compute a formula-based representation for some knowledge described by a set of *examples*, themselves given as assignments to descriptors (variables). Structure identification has also been studied in [KS98], where the problem is called *Inverse Satisfiability*, and in [ZH02].

The following proposition shows that our approach allows to solve efficiently and in an unified manner the structure identification problem for many classes of formula. Importantly, the preferred representations with respect to the plain bases given in Table 2 have polynomial algorithms for deciding satisfiability as soon as satisfiability is polynomial for the corresponding co-clone; i.e., these representations preserve the algorithmical properties of the initial relations.

Proposition 11 *Given a relation R of arity n and containing m vectors and a co-clone ICl to which R belongs, a preferred representation of R with respect to ICl and its plain basis of Table 2 can be found in time $O(m^2n^2)$.*

Proof Zanuttini and Hébrard [ZH02] show that a prime CNF formula φ representing R can be computed in time $O(m^2n^2)$, and that φ contains $O(mn)$ clauses. Therefore a polynomial-time algorithm follows from Table 2 and the case study below (where φ denotes a prime CNF representing R). Notice indeed that the additional operations performed on φ are always linear in its size.

[$ICl \subseteq IS_{10}$ (or dually $ICl \subseteq IS_{00}$)] According to Propositions 4–8 φ is a preferred representation, up to replacing every implicative clause $(\neg x_1 \vee x_2)$ with the equality constraint $Eq(x_1, x_2)$ when $ICl \subseteq IS_{12}$.

[$IE \subseteq ICl \subseteq IE_2$ (or dually $IV \subseteq ICl \subseteq IV_2$)] In this case φ is a Horn (or dually a dual Horn) formula, and thus a preferred representation [ZH02].

[If $IN \subseteq ICl \subseteq IN_2$] In this case φ is a preferred representation up to replacing every one of its clauses of the form $(\neg x_1 \vee \cdots \vee \neg x_k \vee y_1 \vee \cdots \vee y_\ell)$ (or the dual one) with the constraint $Compl_{k,\ell}(x_1, \ldots, x_k, y_1, \ldots, y_\ell)$.

[If $IL \subseteq ICl \subseteq IL_2$] By using the results in [ZH02] one can get a preferred representation by replacing $\vee$ with $\oplus$ and $\neg x$ with $x \oplus 1$ in each clause of φ, setting each obtained sum equal to 1 and simplifying constants.

[If $ID \subseteq ICl \subseteq ID_2$] In this case φ is in 2CNF [ZH02], and thus is a convenient preferred representation of R. $\square$

4.3 Computing the minimal co-clone of a relation

We now show that our notion of plain basis yields an efficient algorithm for deciding the minimal co-clone which contains a given relation. As far as we know this result is new and, as discussed below, could not be taken for granted.

We first wish to point out that there is a straightforward, efficient algorithm for this problem if we forget the infinite part of Post's lattice, since every other clone has a finite basis in which every function is of arity at most 3. Thus one can test a relation for stability under a basis in cubic time and deduce the minimal co-clone containing the relation.

This approach, however, cannot be applied to the infinite part of Post's lattice. Indeed, though each basis is finite (and, in fact, contains at most two functions –see Table 1), the arity of the functions in these bases is unbounded, which implies that testing a relation for stability under one of those bases cannot be done in polynomial time using the naive approach.

This problem, however, is easily circumvented by our approach, as we show now.

Lemma 12 *Given a relation $R \in IS_{10}$ of arity n and containing m vectors, the minimal co-clone in $\{IS_1^n, IS_{10}^n, IS_{12}^n, IS_{11}^n \mid n \in \mathbb{N}\}$ that contains R can be found in time $O(m^2n^2)$. The dual result holds for $R \in IS_{00}$.*

Proof As already mentionned Zanuttini and Hébrard [ZH02] showed that a prime CNF formula φ representing R can be computed in time $O(m^2n^2)$, and that φ contains $O(mn)$ clauses. By scanning φ once, in time $O(mn^2)$, one can find the maximum size of its clauses and deduce n, and at the same time decide whether φ contains unary positive clauses. Finally, for every clause of the form $(\neg x \vee y)$ in φ one can decide whether φ entails $(x \vee \neg y)$ in time $O(mn)$ by deciding whether every vector in R satisfies it; since φ contains $O(mn)$ clauses, this requires $O(m^2n^2)$ operations. Once these informations collected one can find the minimal co-clone containing R by referring to our Table 2, which concludes the proof. $\square$

We are finally able to show that computing the minimal co-clone that contains a given relation can be done in quadratic time.

Proposition 13 *Given a relation R of arity n and containing m vectors, the minimal co-clone of Post's lattice that contains R can be found in time $O(m^2n^2)$.*

Proof The algorithm is as follows. First compute a prime CNF φ representing R in time $O(m^2n^2)$ with Zanuttini and Hébrard's algorithm; φ contains $O(mn)$ clauses. By reading φ once, by the results of [ZH02] and a reasoning similar to that of Proposition 11 and Lemma 12 one can decide which co-clones contain R among $IBF, IR_c, IM, IM_c, ID_2, IV, IV_c, IE, IE_c, II, II_c$ ($c \in \{0,1,2\}$); this requires $O(mn^2)$ operations.

Now, still using the results in [ZH02] one can handle the affine co-clones by essentially replacing $\vee$ with $\oplus$ in φ and testing whether each vector in R satisfies the resulting affine formula; if the answer is affirmative, then R is affine and thus in ID, ID_1, IL or IL_c ($c \in \{0,1,2,3\}$); the exact co-clones among these ones can then be found in linear time by reading the formula once and counting the variables in each equation. Now membership of R in IN or IN_2 can be decided by testing whether the componentwise complement of each vector of R is still in R, in time $O(mn + n^2)$ by first sorting R into a trie. Finally, concerning the infinite part of the lattice, one can decide by reading φ which co-clones contain R among $\{IS_1, IS_{10}, IS_{12}, IS_{11}\}$ and the dual ones, and then apply our Lemma 12.

Finally, membership in each co-clone has been decided in time at most quadratic, which concludes since obviously the minimal co-clone among the ones found can then be computed in constant time with Post's lattice. $\square$

5 The expressibility problem

In this section, we apply our characterization to a fundamental problem in database theory, namely that of deciding whether a given relation R can be expressed by a given set S of relations. We first define what we mean by "expressibility".

Definition 14 (expressible) *Let S be a set of relations and R a relation, and write n for the arity of R. Then R is said to be* expressible *by S if there are a set of variables $\{y_1, \ldots, y_m\}$ and a conjunction of constraints $\mathcal{C}$ over $\{x_1, \ldots, x_n\} \cup \{y_1, \ldots, y_m\}$ such that the constraint $R(x_1, \ldots, x_n)$ is logically equivalent to the formula $\exists y_1 \ldots \exists y_m \mathcal{C}$. In other words R is* expressible *by S if R is in the minimal co-clone containing every relation in S.*

We are thus interested in the following decision problem:

Definition 15 (Expressibility**)**
Input: *A finite set of relations S and a relation R*
Output: *Is R expressible by S? i.e. does R belongs to the minimal co-clone containing every relation in S?*

Note that by definition, this problem could be reformulated into that of deciding whether S is a basis for R (in Böhler *et al.*'s sense). However, it seems that this notion of a basis is of no help in solving it, whereas our stronger notion gives an efficient and simple algorithmic solution to it.

Indeed, since R is expressible by S if and only if the minimal co-clone (wrt set inclusion) containing R is included in the minimal co-clone containing every relation in S, our Proposition 13 gives an efficient algorithm.

Proposition 16 *Problem* Expressibility *is polynomial-time solvable.*

Proof Given S and R the algorithm is as follows. For every relation $R' \in S$ find the minimal co-clone containing R' by using our Proposition 13. Then compute the union C_S of the co-clones found for all relations, which can be done in linear time with Post's lattice. Finally, compute in the same manner the minimal co-clone C_R that contains R, and conclude that R is expressible by S if and only if $C_R \subseteq C_S$, which again can be decided in constant time with Post's lattice. □

6 Conclusion

We have introduced the notion of a *plain basis* for co-clones in Post's lattice, which is stronger than the classical notion of a basis. We have then given a plain basis for every co-clone in Post's lattice, mainly using the prime CNF representation of relations. In particular, we have shown that relations in the infinite part of Post's lattice are those that can be represented by IHSB CNF formulas, which is a new result, though a similar one has been obtained independently and simultaneously for classical bases by Böhler *et al.* [BRSV05].

Based on this notion of a plain basis, we have defined *preferred representations* of relations and shown that these representations can be computed in quadratic time given a relation with respect to our plain bases. This problem is of great importance for the structure identification problem of AI, since our preferred representations preserve the algorithmic properties of relations for the satisfiability problem.

We have also shown that our approach yields a quadratic algorithm for deciding the minimal co-clone of Post's lattice that contains a given relation. This closes an important fundamental open question. We have finally shown that in particular, this allows to derive an efficient solution to the expressibility problem of Database theory, which was also an important fundamental open problem.

Our approach also exhibits strong links between co-clones and CNF representations. Such links are very useful when considering complexity questions and especially, complexity classifications. For instance, two of the authors studied the complexity of the abduction problem for conjunctions of constraints over fixed sets of relations S, and obtained a trichotomic classification involving the IHSB classes [CZ04]. While they observed that the complexity of the abduction problem for some S is also determined by the expressive power of S, thus suggesting that the algebraic approach can also be applied for this problem, their proof was Schaefer's like. The result presented here sheds some light on this classification, since IHSB classes now correspond to identified co-clones and thus the trichotomy result can now also be visualized on Post's lattice. We also hope that this complete characterization of co-clones by means of plain bases will be of help in identifying computational goals for which the complexity for constraints can be studied through the algebraic approach, which constitutes an intriguing fundamental problem.

References

[BCCHV04] M. Bauland, P. Chapdelaine, N. Creignou, M. Hermann and H. Vollmer. An algebraic approach to the complexity of generalized conjunctive queries. In *Proc. 7th International Conference on Theory and Applications of Satisfiability Testing* , SAT'2004, Vancouver (British Columbia, Canada), pages 181-190, May 2004.

[BCRV03] E. Böhler, N. Creignou, S. Reith and H. Vollmer. Playing with Boolean Blocks, Part I: Post's Lattice with Applications to Complexity Theory. *ACM-SIGACT News*, 34(4), Complexity Theory Column 42, pages 38–52, 2003.

[BCRV04] E. Böhler, N. Creignou, S. Reith and H. Vollmer. Playing with Boolean Blocks, Part II: Constraint satisfaction problems. *ACM-SIGACT News*, 35(1), Complexity Theory Column 43, pages 22–35, 2004.

[BRSV05] E. Böhler, S. Reith, H. Schnoor and H. Vollmer. Bases for Boolean co-clones. *Information Processing Letters* 96:59–66, 2005

[CH96] N. Creignou and M. Hermann. Complexity of generalized satisfiability counting problems. *Information and Computation*, 125(1):1-12, 1996.

[CKS01] N. Creignou, S. Khanna and M. Sudan. *Complexity classifications of Boolean constraint satisfaction problems.* SIAM Monographs on Discrete Mathematics and Applications, 2001.

[CZ04] N. Creignou and B. Zanuttini. *A complete classification of the complexity of propositional abduction.* Submitted for publication, October 2004.

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

[Gei68] D. Geiger. Closed systems of functions and predicates. *Pac. J. Math*, 27(2):228–250, 1968.

[GHR95] R. Greenlaw, H. J. Hoover and W. L. Ruzzo. *Limits to parallel computation: P-completeness theory.* Oxford University Press, 1995.

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

[KST97] S. Khanna, M. Sudan and L. Trevisan. Constraint satisfaction: the approximability of minimization problems. In *Proc. 12th Computational Complexity Conference*,IEEE Computer Society Press pages 282–296, 1997.

[KV00] Ph. G. Kolaitis and M. Y. Vardi. Conjunctive queries containment and constraint satisfaction *Journal of Computer and System Sciences*, 61(2):302–332, 2000.

[Pip97] N. Pippenger. *Theories of Computability.* Cambridge University Press, Cambridge, 1997.

[PK79] R. Pöschel and L.A. Kalužnin. *Funktionen- und Relationenalgebren.* DVW, Berlin, 1979.

[Pos41] E.L. Post. The two-valued iterative systems of mathematical logic. *Annals of Mathematical Studies*, 5:1–122, 1941.

[Sch78] T. Schaefer. The complexity of satisfiability problems. In *Proc. 10th STOC, San Diego (CA, USA)*, pages 216–226. Association for Computing Machinery, 1978.

[ZH02] B. Zanuttini and J.-J. Hébrard. A unified framework for structure identification. *Information Processing Letters*, 81(6):335–339, 2002.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'