Trichotomies in the Complexity of Minimal Inference*

Arnaud Durand
ELM, Institut de Mathématiques de Jussieu (CNRS UMR 7586)
Université Denis-Diderot Paris 7
case 7012, site Chevaleret
75205 Paris cedex 13, France
durand@logique.jussieu.fr

Miki Hermann
LIX (CNRS UMR 7161)
École Polytechnique
91128 Palaiseau cedex, France
hermann@lix.polytechnique.fr

Gustav Nordh†
Dept. of Computer and Information Sciences
Linköpings Universitet
581 83 Linköping, Sweden
gusno@ida.liu.se

Abstract

We study the complexity of the propositional minimal inference problem. Although the complexity of this problem has been already extensively studied before because of its fundamental importance in nonmonotonic logics and commonsense reasoning, no complete classification of its complexity was found. We classify the complexity of four different and well-studied formalizations of the problem in the version with unbounded queries, proving that the complexity of the minimal inference problem for each of them has a trichotomy (between P, coNP-complete, and Π_2P-complete). One of these results finally settles with a positive answer the trichotomy conjecture of Kirousis and Kolaitis [A dichotomy in the complexity of propositional circumscription, LICS'01]. In the process we also strengthen and give a much simplified proof of the main result from [Durand and Hermann, The inference problem for propositional circumscription of affine formulas is coNP-complete, STACS'03].

1 Introduction and Summary of Results

Reasoning with minimal models of a theory is a general idea widely used in artificial intelligence, especially for capturing various aspects of common sense and nonmonotonic reasoning. In particular, it is the main idea behind circumscription [22, 23], diagnosis [10], and logic programming under stable model semantics [15]. The key idea behind minimal models reasoning, and in particular circumscription, is that the minimal models have as few "exceptions" as possible, and embody therefore common sense. Minimal inference has been shown by Gelfond *et al.* [16] to coincide with reasoning under the extended closed world assumption, which is one of the main formalisms for

*This research was partially supported by the grant ANR-07-BLAN-0327. A preliminary version of the results in Section 5 has been published in [24] and of those in Section 9 has been published in [12].

†Partially supported by the Swedish Research Council (VR) under grant 2008-4675 and the *Swedish-French Foundation*. Part of this work has been completed during a post-doctoral stay of this author at LIX, Ecole Polytechnique.

reasoning with incomplete information. We focus in this paper on the important basic case where the theory is propositional, which is also the most investigated case from a complexity point of view. In this case, the minimality of models is defined with respect to the *pointwise partial order*, extending the order $0 < 1$ on truth values.

The complexity of several basic algorithmic problems have been studied in connection with minimal models of propositional formulas: among them are the *model selection* [7,28], *model checking* [5,20], and *inference* [6,11,13,21,37] problems. Given a propositional formula φ, the minimal model selection problem requires to compute a minimal model of φ. Similarly, given a propositional formula φ and a truth assignment m, the minimal model checking problem asks whether m is a minimal model of φ. Given two propositional formulas φ and ψ, the minimal inference problem asks whether ψ is true in every minimal model of φ. In propositional theories these problems are identical to the corresponding problems for propositional circumscription. Hence, they are often called the model selection, model checking, and inference problems for propositional circumscription, respectively. In this paper we focus exclusively on the minimal inference problem (or equivalently the inference problem for propositional circumscription). We remark that there is also a parallel line of research that studies the complexity of circumscription for a different type of propositional formulas (i.e., so called Post formulas [37,38]).

McCarthy [23] and Gelfond, Przymusinska, and Przymusinski [16] extended the notion of circumscription to a more general setting. They altered the notion of minimal models of a formula φ by dividing the set of variables V of φ into three disjoint sets, namely of the variables P that are minimized, the variables Z that may vary, and the variables Q that maintain a fixed value. Given two models m and m' of φ, the relation $m \leq m'$ holds if $m(x) \leq m'(x)$ is satisfied for all $x \in P$ and the equality $m(y) = m'(y)$ holds for all $y \in Q$. A model m' of φ is defined to be a minimal model if there is no model m of φ satisfying $m \leq m'$ and $m(x) < m'(x)$ for a variable $x \in P$.

The generalized minimal inference problem with the extended notion of minimal models is called (P, Z, Q)-minimal inference. It should be clear that the original minimal inference problem is the special cases of the corresponding generalized problem satisfying the condition $Z = Q = \emptyset$. Cadoli and Lenzerini analyzed in [6] the complexity of (P, Z, Q)-minimal inference, as well as the variants with no fixed variables (i.e., $(P, Z, \emptyset)$-minimal inference), with no free variables (i.e., $(P, \emptyset, Q)$-minimal inference), and where all variables must be minimized (i.e., $(P, \emptyset, \emptyset)$-minimal inference). In fact, they analyzed the complexity of reasoning under several variants of closed world assumptions and their results propagate to minimal inference through the equivalence presented by Gelfond *et al.* in [16].

The topic of this paper is to classify the complexity of the minimal inference problem for the four restrictions on the variable sets Q and Z mentioned above, as well as all restrictions on the types of clauses allowed in the theory φ, presented by a formula in conjunctive normal form. Note that the corresponding classification problem for satisfiability was solved in a seminal paper by Schaefer [35].

The most studied variant of the minimal inference problem and also the most difficult variant to analyse is the classical one where we require all variables to be minimized, i.e., with $Z = \emptyset$ and $Q = \emptyset$. This inference problem was proved Π_2P-complete by Eiter and Gottlob in [13] if no restrictions are imposed on the propositional theory φ. Cadoli and Lenzerini proved in [6] that the inference problem becomes coNP-complete if φ is a bijunctive or a dual Horn formula. Durand and Hermann proved in [11] that also the inference problem for affine formulas φ is coNP-complete. Finally, Kirousis and Kolaitis [21] proved a dichotomy theorem separating the Π_2P-complete cases from the cases in coNP. Moreover, they conjectured that the cases in coNP could be separated into coNP-complete cases and tractable (i.e., polynomial-time decidable) cases.

The proof of the dichotomy theorem for the minimal inference problem by Kirousis and Kolaitis [21] turned out to be difficult for several reasons. One of them was the impossibility to apply the well-known approach through the theory of clones and Post's lattice [3, 17, 32] to obtain a complexity classification. The culprit is the existential quantification which does not combine well with minimality. This implies that the co-clones, which are the sets of relations closed under variable identification, variable permutation, conjunction, and existential quantification, are not invariant with respect to the complexity of minimal inference. This situation precludes the use of the Galois correspondence, which allows first to prove a completeness result for a subset of relations and subsequently to extend it to the whole co-clone.

However, if the pointwise partial order on models is relaxed, which is the consequence of studying the generalized minimal inference problems where some variables are allowed to vary, then it has been observed that the Galois correspondence can be restored [25] and the approach via clone theory works. With the help of these powerful tools it is relatively easy to classify the complexity of these generalized forms of minimal inference for every restriction on the types of allowed clauses in the theory (what we do in Sections 5 and 6). We note in passing that the recent results by Schnoor and Schnoor [36] offer an alternative way of making the approach via clone theory applicable to these problems.

As already mentioned, following the results of Kirousis and Kolaitis [21], it is clear that we cannot use the powerful algebraic approach for attacking the complexity of the minimal inference problem. Our approach is instead based on refinements of Schaefer's approach (via a particular type of *implementations*) to classify the complexity of the satisfiability problem in propositional logic (see [8, 35]). There is one significant difference though between Schaefer's implementations and ours. We can only use conjunction, variable identification, and variable permutation (*without* existential quantifiers) in our implementations. This is due to the fact, as already explained, that existential quantification does not preserve the minimal models, nor the complexity of the problem. Nevertheless, we manage to separate the coNP-hard cases from the tractable ones by using this approach through a more fine-grained analysis, thus, finishing the complexity classification for all four variants of minimal inference with all possible restrictions on the types of allowed clauses, establishing trichotomy theorems for all considered variants of the minimal inference problem in Schaefer's framework. Note that we address the minimal inference problem for unbounded queries. Some remarks on the bounded case can be found in the conclusion.

Our paper finally settles with a positive answer the trichotomy conjecture of Kirousis and Kolaitis [21] for the basic variant of minimal inference (with unbounded queries), which was also listed as one of the open problems (Question 4.1) during the International Workshop on Mathematics of Constraint Satisfaction held in Oxford in March 2006 [27]. In the process we also strengthen and give a much simplified proof of the coNP-completeness result by Durand and Hermann [11]. We also believe that the implementations we present might be interesting in their own right, since they could be useful during the study of other problems, which are not invariant under existential quantification.

2 Preliminaries

Throughout the paper we use the standard correspondence between constraints and relations. We use the same symbol for a constraint and its corresponding relation, since the meaning will always be clear from the context, and we say that the constraint *generates* the relation. Mathematically speaking, a constraint or formula φ *generates* a relation R if the solution set of φ is equal to R.

An n-ary *logical relation* R is a Boolean relation of arity n. Each element of a logical relation R is an n-ary Boolean vector (also called a tuple) $m = (m_1, \ldots, m_n) \in \{0,1\}^n$. To save space, we will often write the vector $(m_1, \ldots, m_n)$ in the form $m_1 \cdots m_n$. We also write $m[i]$ for m_i, denoting this way the i-th coordinate of the Boolean vector m. Let V be a set of variables. A *constraint* is an application of R to an n-tuple of variables from V, i.e., $R(x_1, \ldots, x_n)$.

Consider a relation R represented in the form of a Boolean matrix, i.e., the vectors of R constitute the rows of the matrix. We say that a relation R is *irredundant* if it does not contain two identical columns and it cannot be transformed by column permutation to a relation of the form $Q \times \{0,1\}^k$ for a $k \geq 1$, where Q is another relation. If R is redundant then the corresponding *irredundant reduction* R° is formed by removing the $\{0,1\}^k$ part from $R = Q \times \{0,1\}^k$ and identifying identical columns. A set of relations S is irredundant if every relation in S is irredundant. Given a set of relations S, we form the irredundant reduction S° by replacing all redundant relations R in S by their corresponding irredundant reductions R°.

Throughout the text we refer to different types of Boolean constraint relations following Schaefer's terminology [35] (see also the monograph [8]). We say that a Boolean relation R is

- 1-*valid* if $1 \cdots 1 \in R$ and it is 0-*valid* if $0 \cdots 0 \in R$,
- *Horn* (*dual Horn*) if R can be generated by a conjunctive normal form (CNF) formula having at most one unnegated (negated) variable in each clause,
- *bijunctive* if it can be generated by a CNF formula having at most two variables in each clause,
- *affine* if it can be generated by an affine system of equations $S\colon Ax = b$ over $\mathbb{Z}_2$,
- *bijunctive affine*, sometimes also called width-2 affine, if it is both affine and bijunctive,
- *complementive* if for each $(m_1, \ldots, m_n) \in R$ also $(\neg m_1, \ldots, \neg m_n) \in R$.

A set S of Boolean relations is called 0-valid (1-valid, Horn, dual Horn, affine, bijunctive, complementive) if *every* relation in S is 0-valid (1-valid, Horn, dual Horn, affine, bijunctive, complementive). A relation R (a set of relations S) is called *Schaefer* if it belongs to one of the classes Horn, dual Horn, bijunctive, or affine.

Let $R \subseteq \{0,1\}^k$ be a k-ary Boolean relation. We say that a relation R' is a *direct* 0-*section* of R if there exists an index $i \in \{1, \ldots, k\}$, such that

$$R' \quad = \quad \{(m[1], \ldots, m[i-1], m[i+1], \ldots, m[k]) \mid m \in R \wedge m[i] = 0\}.$$

In other words, R' is a direct 0-section of R if R' can be obtained from R by selecting all tuples in R having a 0 in position i (for some $1 \leq i \leq k$) and projecting onto the other $k-1$ columns/positions. We say that a relation R'' is a 0-*section* of R if there exists a finite sequence of Boolean relations R_0, R_1, $\ldots$, R_n, such that $R_0 = R$, $R'' = R_n$, and R_{j+1} is a direct 0-section of R_j for each $j = 0, \ldots, n-1$. Let S be a finite set of Boolean relations. We say that S^* is a 1-valid *restriction* of S if it contains all relations R^* which are both 1-valid and a 0-section of a relation R from S. Note that starting from an arbitrary relation $R \in S$, we always arrive at a 1-valid 0-section R^* by iterating the 0-section operation long enough, unless $R = \{0 \cdots 0\}$. The aforementioned technique is a well-known concept from coding theory and it was already used for circumscription [20] and minimal inference [21]. Note that if S is Schaefer, then so is S^*. This is because the Schaefer classes are stable under constant substitution. Also note that S is 0-valid if and only if S^* is 0-valid.

An *assignment* is a mapping $m\colon V \to \{0,1\}$ assigning a Boolean value $m(x)$ to each variable $x \in V$. If we arrange the variables in some arbitrary but fixed order, say as a vector $(x_1,\dots,x_n)$, then the assignments can be identified with vectors from $\{0,1\}^n$. The i-th component of a vector m, denoted by $m[i]$, corresponds to the value of the i-th variable, i.e., $m(x_i) = m[i]$. An assignment m satisfies the constraint $R(x_1,\dots,x_n)$ if $(m(x_1),\dots,m(x_n)) \in R$ holds. An assignment m satisfying a constraint R is called a *model* of R.

Let S be a non-empty finite set of Boolean relations, also called a *constraint language*. An *S-formula* is a finite conjunction of clauses $\varphi = c_1 \wedge \cdots \wedge c_k$, where each clause c_i is a constraint application of some logical relation $R \in S$. An assignment m satisfies the formula φ if it satisfies all clauses c_i. Hence the notion of models naturally extends from constraints to formulas. We denote by $[\varphi]$ the set of models of a formula φ. It is clear that each $[\varphi]$ denotes a Boolean relation. In this case we also say that the formula (constraint) φ *generates* the relation $[\varphi]$.

Let $m = (m[1],\dots,m[n])$ and $m' = (m'[1],\dots,m'[n])$ be two Boolean vectors from $\{0,1\}^n$. We write $m \leq m'$ to denote that $m[i] \leq m'[i]$ holds for every $i \leq n$, as well as $m < m'$ for $m \leq m'$ and $m \neq m'$. The relation $\leq$ is called the *pointwise partial order* on models. Suppose $\varphi(x_1,\dots,x_n)$ is a Boolean formula having $x_1,\dots,x_n$ as its variables and let $m \in \{0,1\}^n$ be a truth assignment. We say that m is a *minimal model* of φ if m is a model of φ and there is no other model m' of φ that satisfies the relation $m' < m$. Two vectors $a, b \in R$ satisfying the relation $a < b$ are called *comparable*. We say that a relation R is *incomparable* if it does not contain comparable vectors. We say that a set of relations S is incomparable if each relation $R \in S$ is incomparable.

Let us partition the variables V into three disjoint sets P, Z, and Q, as it is usual in minimality problems in artificial intelligence (see [6,16]), where P is the set of variables to be minimized, Z is the set of variables allowed to vary (also called free variables), and Q is the set of variables that must maintain a constant value (also called constants). We write $m \prec m'$ to denote that the Boolean vectors m and m' satisfy the following conditions: (1) $m(x) < m'(x)$ for all $x \in P$ and (2) $m(x) = m'(x)$ for all variables $x \in Q$. We write $m \preceq m'$ if the first condition is $m \leq m'$. Note that there is no condition on the values of variables in Z. We say that m is a (P,Z,Q)-*minimal* or *generalized minimal model* of φ if m is a satisfying truth assignment of φ and there is no other satisfying truth assignment m' of φ that satisfies the relation $m' \prec m$. If the variable sets Z and Q are empty, then we identify the $(P,\emptyset,\emptyset)$-minimal models with the aforementioned minimal models of φ.

Let φ and ψ be two propositional formulas in conjunctive normal form. We say that ψ follows from φ in propositional minimal inference, denoted by $\varphi \models_{\min} \psi$, if ψ is true in every minimal model of φ. Since ψ is a conjunction $c_1 \wedge \cdots \wedge c_k$ of clauses c_i, the minimal inference $\varphi \models_{\min} \psi$ holds if and only if the minimal inference $\varphi \models_{\min} c_i$ holds for each i. Hence we can restrict ourselves to consider only a single clause instead of a formula ψ at the right-hand side of the propositional inference problem $\varphi \models_{\min} c$. Similarly, we say that ψ follows from φ in (P,Z,Q)-minimal inference, also denoted by $\varphi \models_{\min}^{Z,Q} \psi$, if ψ is true in every (P,Z,Q)-minimal model of φ. The superscripts Z and Q are usually dropped from $\varphi \models_{\min}^{Z,Q} \psi$ when their presence is clear from the context. Notice that we do not impose any length bounds on the right-hand side clause c, also called the *query*, in a minimal inference problem $\varphi \models_{\min} C$, hence we consider the minimal inference problem with unbounded queries.

3 Clones, Co-Clones, and Post's Lattice

The theory of clones and co-clones has previously shown to be very useful for characterizing the complexity of problems with restrictions on the types of allowed constraints (see for instance the survey [2,3]). The underlying theory, a part of universal algebra, can be found in the monograph [31] or in Chapter 1 of the book [29]. Also helpful is the survey paper [30].

The theory of Boolean clones studies the closure properties of Boolean relations with respect to Boolean functions. The notion of closure property of a Boolean relation has been defined in a general way. Let $f\colon \{0,1\}^k \to \{0,1\}$ be a Boolean function of arity k. We say that R is *closed* under f, or that f is a *polymorphism* of R, if for any choice of k vectors $m_1,\ldots,m_k \in R$, we have that

$$\Big(f\big(m_1[1],\ldots,m_k[1]\big),\ f\big(m_1[2],\ldots,m_k[2]\big),\ \ldots,\ f\big(m_1[n],\ldots,m_k[n]\big)\Big) \in R,$$

i.e., that the new vector constructed coordinate-wise from $m_1, \ldots, m_k$ by means of f belongs to R.

We denote by $\mathrm{Pol}(R)$ the set of all polymorphisms of R and by $\mathrm{Pol}(S)$ the set of Boolean functions that are polymorphisms of every relation in S. It turns out that $\mathrm{Pol}(S)$ is a *closed set of Boolean functions*, also called a clone, for every set of relations S. A *clone* is a set of functions containing all projections and closed under composition. In particular, the set of all function $[F]$ constructed from a given set of functions F by means of composition and projection, forms a clone.

If a relation R is closed under a function f, we say that R is *invariant* under f. We denote by $\mathrm{Inv}(F)$ the set of relations *invariant* under the functions F. In other words, for each relation $R \in \mathrm{Inv}(F)$ and each Boolean function $f \in F$ we have that f is a polymorphism of R. It turns out that $\mathrm{Inv}(F)$ is a *closed set of Boolean relations*, also called a *co-clone*, for every set of functions F. Given a set of functions F we denote the corresponding co-clone $\mathrm{Inv}(F)$ by iF.

We denote by $\langle S\rangle$ the *co-clone* generated by the set of relations S, also called the *co-clone closure* of S, i.e., the smallest set containing S and the binary equality relation eq^2, and that are closed under Cartesian product, permutation, restriction, and projection. The co-clone $\langle S\rangle$ can also be understood as the set of all models of formulas built from constraints over S, and closed under conjunction, variable permutation, variable identification, and existential quantification. Keeping in mind that $[\varphi]$ is the set of model of φ, this means that $\langle S\rangle$ can be defined inductively as follows:

1. $S \subseteq \langle S\rangle$ and $eq^2 \in \langle S\rangle$,
2. if $R_1, R_2 \in \langle S\rangle$ then $R_1 \times R_2 = [R_1(\vec{x}) \wedge R_2(\vec{y})] \in \langle S\rangle$, (Cartesian product or conjunction)
3. if $[R(\vec{x}, u, \vec{y}, v, \vec{y})] \in \langle S\rangle$ then $[R(\vec{x}, v, \vec{y}, u, \vec{y})] \in \langle S\rangle$, (permutation)
4. if $R(x_1, x_2, \vec{y}) \in \langle S\rangle$ then $[R(x, x, \vec{y})] \in \langle S\rangle$, where x is a fresh variable, (restriction)
5. if $[R(x, \vec{y})] \in \langle S\rangle$ then $[\exists x R(x, \vec{y})] \in \langle S\rangle$. (projection or existential quantification)

We will also require a weaker closure than $\langle S\rangle$, called *weak co-clone closure* and denoted by $\langle S\rangle_{\nexists}$, where the set is closed under the first four operations, but *not* under projection (existential quantification) depicted in Condition 5. Note that this new closure notion has been studied before in [14,26,34,36]. The *idempotent closure* S^{id} of a set of relations S corresponds to $S \cup \{\ \{0\},\ \{1\}\ \}$, which can be also written as $S \cup \{[x],\ [\neg x]\}$.

A *Galois correspondence* has been exhibited between the sets of Boolean functions $\mathrm{Pol}(S)$ and the sets of Boolean relations S, as well as between the functions F and the relations $\mathrm{Inv}(F)$. The Galois correspondence encompasses the following implications. Let F_1 and F_2 be two sets of Boolean functions, as well as S_1 and S_2 two sets of Boolean relations. The inclusion $[F_1] \subseteq [F_2]$ implies

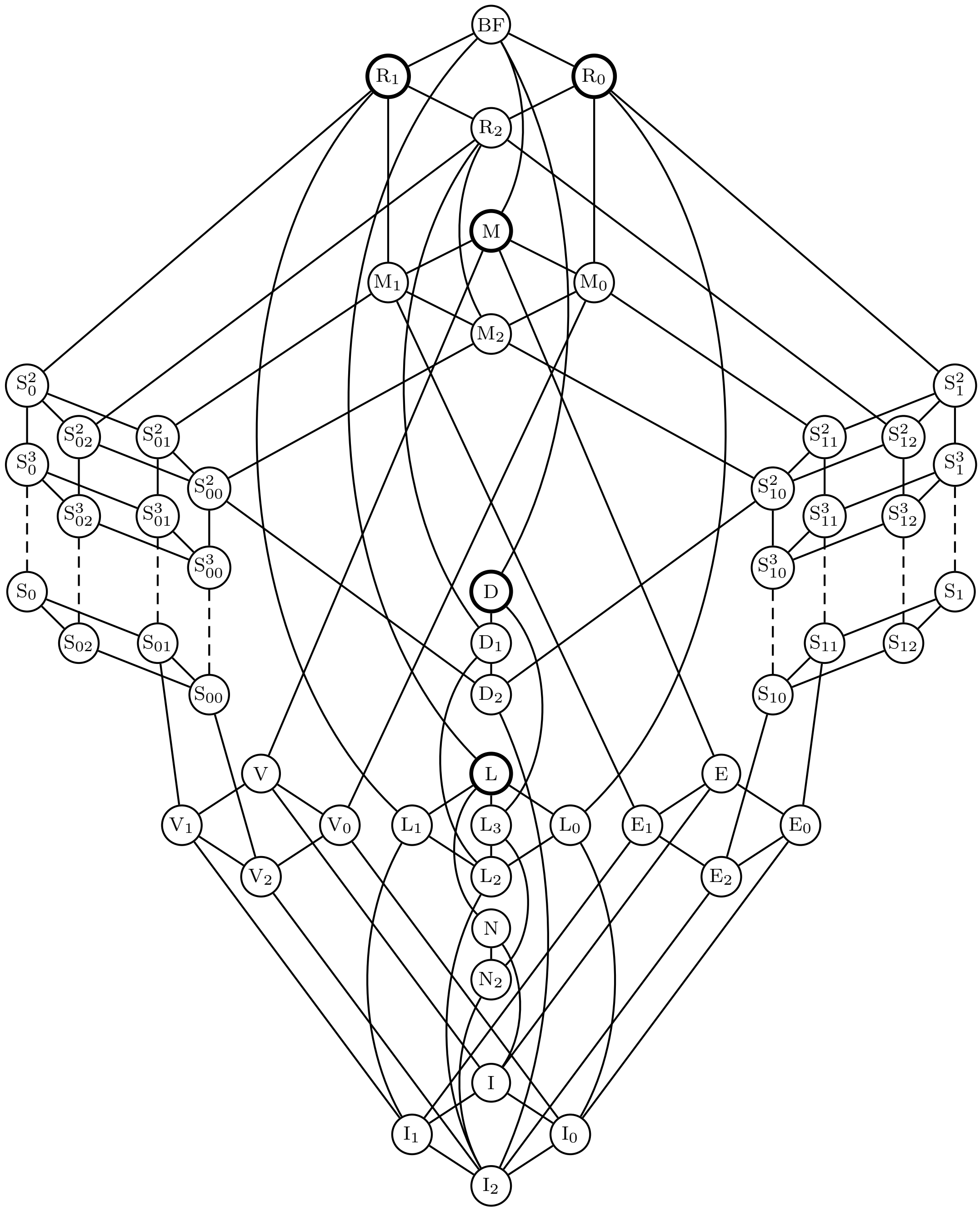

Figure 1: Post's lattice of Boolean clones

Clone	Base(s)	Co-clone	Base(s)
BF	$\{x \wedge y, \neg x\}$	iBF	$\{eq^2\}$
R_0	$\{x \wedge y, x \oplus y\}$	$i R_0$	$\{\neg x\}$
R_1	$\{x \vee y, x \oplus y \oplus 1\}$	$i R_1$	$\{x\}$
R_2	$\{x \vee y, x \wedge (y \oplus z \oplus 1)\}$	$i R_2$	$\{\neg x, x\}$
M	$\{x \wedge y, x \vee y, 0, 1\}$	iM	$\{x \to y\}$
M_1	$\{x \wedge y, x \vee y, 1\}$	$i M_1$	$\{x \to y, x\}$
M_0	$\{x \wedge y, x \vee y, 0\}$	$i M_0$	$\{x \to y, \neg x\}$
M_2	$\{x \wedge y, x \vee y\}$	$i M_2$	$\{x \to y, \neg x, x\}$
S_0^m	$\{x \to y, \text{dual}(h_m)\}$	$i S_0^m$	$\{or^m\}$
S_1^m	$\{x \wedge \neg y, h_m\}$	$i S_1^m$	$\{nand^m\}$
S_0	$\{x \to y\}$	$i S_0$	$\{or^m \mid m \geq 2\}$
S_1	$\{x \wedge \neg y\}$	$i S_1$	$\{nand^m \mid m \geq 2\}$
S_{02}^m	$\{x \vee (y \wedge \neg z)\}$	$i S_{02}^m$	$\{or^m, \neg x, x\}$
S_{02}	$\{x \vee (y \wedge \neg z)\}$	$i S_{02}$	$\{or^m \mid m \geq 2\} \cup \{\neg x, x\}$
S_{01}^m	$\{\text{dual}(h_m), 1\}$	$i S_{01}^m$	$\{or^m, x \to y\}$
S_{01}	$\{x \vee (y \wedge z), 1\}$	$i S_{01}$	$\{or^m \mid m \geq 2\} \cup \{x \to y\}$
S_{00}^m	$\{x \vee (y \wedge z), \text{dual}(h_m)\}$	$i S_{00}^m$	$\{or^m, x \to y, \neg x, x\}$
S_{00}	$\{x \vee (y \wedge z)\}$	$i S_{00}$	$\{or^m \mid m \geq 2\} \cup \{\neg x, x, x \to y\}$
S_{12}^m	$\{x \wedge (y \vee \neg z), \text{dual}(h_m)\}$	$i S_{12}^m$	$\{nand^m, \neg x, x\}$
S_{12}	$\{x \wedge (y \vee \neg z)\}$	$i S_{12}$	$\{nand^m \mid m \geq 2\} \cup \{\neg x, x\}$
S_{11}^m	$\{h_m, 0\}$	$i S_{11}^m$	$\{nand^m, x \to y\}$
S_{11}	$\{x \wedge (y \vee z), 0\}$	$i S_{11}$	$\{nand^m \mid m \geq 2\} \cup \{x \to y\}$
S_{10}^m	$\{x \wedge (y \vee z), h_m\}$	$i S_{10}^m$	$\{nand^m, \neg x, x, x \to y\}$
S_{10}	$\{x \wedge (y \vee z)\}$	$i S_{10}$	$\{nand^m \mid m \geq 2\} \cup \{\neg x, x, x \to y\}$
D	$\{(x \wedge \neg y) \vee (x \wedge \neg z) \vee (\neg y \wedge \neg z)\}$	iD	$\{x \oplus y\}$
D_1	$\{(x \wedge y) \vee (x \wedge \neg z) \vee (y \wedge \neg z)\}$	$i D_1$	$\{x \oplus y, x\}$
D_2	$\{(x \vee y) \wedge (y \vee z) \wedge (z \vee x)\}$	$i D_2$	$\{x \oplus y, x \to y\}$
L	$\{x \oplus y, 1\}$	iL	$\{even^4\}$
L_0	$\{x \oplus y\}$	$i L_0$	$\{even^4, \neg x\}$, $\{even^3\}$
L_1	$\{x \equiv y\}$	$i L_1$	$\{even^4, x\}$, $\{odd^3\}$
L_2	$\{x \oplus y \oplus z\}$	$i L_2$	$\{even^4, \neg x, x\}$
L_3	$\{x \oplus y \oplus z \oplus 1\}$	$i L_3$	$\{even^4, x \oplus y\}$
V	$\{x \vee y, 0, 1\}$	iV	$\{x \vee y \vee \neg z\}$
V_0	$\{x \vee y, 0\}$	$i V_0$	$\{x \vee y \vee \neg z, \neg x\}$
V_1	$\{x \vee y, 1\}$	$i V_1$	$\{x \vee y \vee \neg z, x\}$
V_2	$\{x \vee y\}$	$i V_2$	$\{x \vee y \vee \neg z, \neg x, x\}$
E	$\{x \wedge y, 0, 1\}$	iE	$\{\neg x \vee \neg y \vee z\}$
E_0	$\{x \wedge y, 0\}$	$i E_0$	$\{\neg x \vee \neg y \vee z, \neg x\}$
E_1	$\{x \wedge y, 1\}$	$i E_1$	$\{\neg x \vee \neg y \vee z, x\}$
E_2	$\{x \wedge y\}$	$i E_2$	$\{\neg x \vee \neg y \vee z, \neg x, x\}$
N	$\{\neg x, 0\}$, $\{\neg x, 1\}$	iN	$\{dup^3\}$
N_2	$\{\neg x\}$	$i N_2$	$\{nae^3\}$
I	$\{\text{id}, 0, 1\}$	iI	$\{even^4, x \to y\}$
I_0	$\{\text{id}, 0\}$	$i I_0$	$\{even^4, x \to y, \neg x\}$, $\{dup^3, x \to y\}$
I_1	$\{\text{id}, 1\}$	$i I_1$	$\{even^4, x \to y, x\}$
I_2	$\{\text{id}\}$	BR	{1-in-3}

Table 1: List of Boolean clones and co-clones with bases

Functions	h_m	$=$	$\bigvee_{i=1}^{m+1} x_1 \wedge \cdots \wedge x_{i-1} \wedge x_{x+1} \wedge \cdots \wedge x_{m+1}$
	$\mathrm{dual}(h_m)(x_1, \ldots, x_{m+1})$	$=$	$\neg h_m(\neg x_1, \ldots, \neg x_{m+1})$
	$(x \oplus y)$	$=$	$x + y \mod 2$
	$(x \equiv y)$	$=$	$x + y + 1 \mod 2$
	$\mathrm{id}(x)$	$=$	x
Relations	or^m	$=$	$\{(a_1, \ldots, a_m) \in \{0,1\}^m \mid \sum_{i=1}^m a_i > 0\}$
	$nand^m$	$=$	$\{(a_1, \ldots, a_m) \in \{0,1\}^m \mid \sum_{i=1}^m a_i \neq m\}$
	$even^m$	$=$	$\{(a_1, \ldots, a_m) \in \{0,1\}^m \mid \sum_{i=1}^m a_i$ is even$\}$
	odd^m	$=$	$\{(a_1, \ldots, a_m) \in \{0,1\}^m \mid \sum_{i=1}^m a_i$ is odd$\}$
	dup^3	$=$	$\{0,1\}^3 \smallsetminus \{010, 101\}$
	nae^3	$=$	$\{0,1\}^3 \smallsetminus \{000, 111\}$
	eq^2	$=$	$\{00, 11\}$
	1-in-3	$=$	$\{001, 010, 100\}$

Table 2: List of Boolean functions and relations

R is Horn	$\Leftrightarrow$	$m, m' \in R$ implies $m \wedge m' \in R$
R is dual Horn	$\Leftrightarrow$	$m, m' \in R$ implies $m \vee m' \in R$
R is bijunctive	$\Leftrightarrow$	$m, m', m'' \in R$ implies $\mathrm{maj}(m, m', m'') \in R$
R is affine	$\Leftrightarrow$	$m, m', m'' \in R$ implies $m + m' + m'' \in R$
R is complementive	$\Leftrightarrow$	$m \in R$ implies $\neg m \in R$

Table 3: Correspondence with closure operations

$\mathrm{Inv}(F_1) \supseteq \mathrm{Inv}(F_2)$ and the inclusion $\langle S_1 \rangle \subseteq \langle S_2 \rangle$ implies $\mathrm{Pol}(S_1) \supseteq \mathrm{Pol}(S_2)$. In particular, this Galois correspondence enables the proof of the identities $\mathrm{Pol}(\mathrm{Inv}(F)) = [F]$ and $\mathrm{Inv}(\mathrm{Pol}(S)) = \langle S \rangle$ for every set of functions F and each set of relations S. Moreover, it helps to easily establish complexity results for problems compatible with co-clones.

All clones of Boolean functions were identified by Post in [32]. Post also detected the inclusion structure of these clones, what is now referred to as *Post's lattice* (see Figure 1), as well as the property of each Boolean clone to have a finite *basis*, a smallest finite set of Boolean functions F identifying the clone $[F]$. The list of all clones and co-clones, together with their bases, according to Böhler *et al.* [2–4], can be found in Table 1, with the used functions and relations presented in Table 2. In the last column of Table 1 we often write only the constraints φ instead of the generated relation $[\varphi]$.

In the sequel we need to determine the properties of relations. This will be done by means of closure operations of the corresponding clones. In particular, we study the closure under the Boolean functions called conjunction ($\wedge$), disjunction ($\vee$), majority (maj), addition over the Boolean ring $\mathbb{Z}_2$ called exclusive-or ($+$), and negation ($\neg$), where the majority operation is equivalent to the identity $\mathrm{maj}(m, m', m'') = (m \vee m') \wedge (m' \vee m'') \wedge (m'' \vee m)$. These operations are applied componentwise. The correspondence between the type of relation and the closure operation is specified in Table 3.

4 Generalized Minimal Inference and Extension

We are interested in the complexity of the following problem, which is a generalization of the minimal inference problem studied in [13, 21]. See, e.g., the paper [19] for an overview of existing complexity results for corresponding circumscription and minimal inference problems.

Problem: GMININF(S)
Input: A set of variables V divided into three disjoint sets P, Z, and Q, an S-formula φ over the variables V, and a clause ψ.
Question: Is ψ satisfiable in every (P, Z, Q)-minimal model of φ, i.e., does $\varphi \models_{\min} \psi$ hold?

According to the emptiness of the sets Z and Q, there exist four variants of the generalized minimal inference problem, namely

GMININF	if Z and Q are unrestricted,	VMININF	if $Q = \emptyset$,
CMININF	if $Z = \emptyset$,	MININF	if $Z = Q = \emptyset$.

We will call VMININF the minimal inference with free variables and CMININF the minimal inference with constants. MININF is the classical inference problem where all variables are minimized, as it was considered in [13, 21].

For the non-parametrized case, when we consider the class of formulas as a whole and not produced from a given set of relations S, the following results have been previously proved. Eiter and Gottlob proved in [13] that MININF is Π_2P-complete. Cadoli and Lenzerini [6] showed that the MININF problem for dual Horn or bijunctive formulas φ is coNP-complete. Durand and Hermann [11] showed that the MININF problem for affine formula φ is coNP-complete. For the parametrized case Kirousis and Kolaitis [21] showed that there exists a dichotomy between the Π_2P-complete case of MININF(S) for a whole set of relations S and the special cases of S for which MININF(S) included in coNP. The MININF(S) problem for a set of Horn relations S is trivially known to be polynomial-time decidable, since a Horn formula has at most one minimal model, depending on its satisfiability.

When dealing with the GMININF(S) problem and its complexity we sometimes need to speak about the complexity of infinite sets of relations S. Note that we are interested only in the complexity of GMININF(S) for finite sets of relations S, but we will need these notions when we speak about the problem GMININF($\langle S \rangle$) since $\langle S \rangle$ is in general an infinite set of relations. We say that GMININF(S) is decidable in polynomial time for an infinite set S if for each finite set of relations $S' \subseteq S$ the problem GMININF(S') is decidable in polynomial time. We say that GMININF(S) is coNP-complete if there exists a finite subset $S' \subseteq S$, such that GMININF(S') is coNP-complete and for each finite subset $S'' \subseteq S$ it is the case that GMININF(S'') is in coNP. Finally, we say that GMININF(S) is Π_2P-complete if there exists a finite subset $S' \subseteq S$, such that GMININF(S') is Π_2P-complete and for each finite subset $S'' \subseteq S$ it is the case that GMININF(S'') is in Π_2P.

It should be clear that membership results propagate from more general to more restrictive variants of the minimal inference problem. In the same spirit, hardness results propagate from more restrictive to more general variants of minimal inference. Hence, a polynomial-time decidability of GMININF also holds for VMININF, CMININF, and MININF, whereas a hardness result for MININF also holds for the other three variants.

Nordh and Jonsson [25] observed that the Galois correspondence exists for the GMININF problem, what allowed them to prove the following results. The proof in [25] is done only for GMININF, but the same technique carries over also to VMININF, since the proof is based on a construction concerning variables only.

Proposition 4.1 (Nordh & Jonsson [25]) *If* $\langle S_1\rangle \subseteq \langle S_2\rangle$ *then* GMININF$(S_1) \leq_m^p$ GMININF(S_2) *and* VMININF$(S_1) \leq_m^p$ VMININF(S_2).

Theorem 4.2 (Nordh & Jonsson [25]) *The problems* GMININF(S) *and* GMININF$(\langle S\rangle)$ *are polynomially equivalent for a set of Boolean relations* S*. The same result holds for* VMININF.

As Nordh and Jonsson showed, the proof of this theorem is made possible because the existentially quantified variables $x_{p+1}, \ldots, x_q$ of a constructed relation

$$R(x_1, \ldots, x_p) \quad = \quad \exists x_{p+1} \cdots \exists x_q \; R_1(x_1, \ldots, x_q) \wedge \cdots \wedge R_r(x_1, \ldots, x_q)$$

from a given set of relations $S = \{R_1, \ldots, R_r\}$, can be added to the set of varying variables Z. Hence, the generalized minimal inference problem for a given set of relations S and its co-clone closure $\langle S\rangle$ has the same complexity. This result remains true also for the VMININF problem. However, there is no known way to derive an equivalent to Theorem 4.2 for the CMININF and MININF problems, mainly because existentially quantified variables cannot be handled in the appropriate way and therefore the universal algebra approach cannot be directly applied. We remark that one possible way to deal with this is to work with so called partial algebras instead [26, 36].

Sometimes it is convenient to consider the idempotent closure S^{id} instead of the original set of relations S. However, with respect to the problems GMININF and CMININF there is no difference between S and S^{id} from the complexity point of view, as it was observed by Nordh and Jonsson [25].

Theorem 4.3 (Nordh & Jonsson [25]) *The problems* GMININF(S) *and* GMININF(S^{id})*, as well as* CMININF(S) *and* CMININF(S^{id}) *are polynomially equivalent for any set of Boolean relations* S*.*

As Durand and Hermann showed in [11], it is sometimes more convenient to investigate the dual problem GMINEXT of generalized minimal extension, defined as follows.

Problem: GMINEXT(S)
Input: A set of variables V divided into three disjoint sets P, Z, and Q, an S-formula φ over variables V, and a partial assignment m of φ for a subset of variables $V' \subseteq V$.
Question: Can m be extended to a (P, Z, Q)-minimal model $\bar{m}$ of φ?

The problems VMINEXT, CMINEXT, and MINEXT are derived from GMINEXT by the corresponding conditions on the sets Z and Q, as in the case of the GMININF problem. If m can be extended to a (P, Z, Q)-minimal model $\bar{m}$ of φ, but the model $\bar{m}$ is not specified, we also say that m has a (P, Z, Q)-minimal extension. The relationship between GMININF and GMINEXT can be easily established through the following construction. Let S be a finite set of Boolean relations and m an assignment to the variables $x_1, \ldots, x_n$. Let c_x be the largest clause falsified by the assignment m, i.e., $c_m = l_1 \vee \cdots \vee l_n$, where $l_i = x_i$ if $m(x_i) = 0$, and $l_i = \neg x_i$ otherwise. It is clear that the clause c_m is *not* satisfiable in every (P, Z, Q)-minimal model of the formula φ if and only if the assignment m can be extended to a (P, Z, Q)-minimal model $\bar{m}$ of φ. From this follows that for each set of relations S the problem GMININF(S) is Π_2P-complete, coNP-complete, or polynomial-time decidable if and only if GMINEXT(S) is Σ_2P-complete, NP-complete, or polynomial-time decidable, respectively. This also holds for the problems VMININF and VMINEXT, as well as for MININF and MINEXT.

Following Nordh and Jonsson [25], the Galois correspondence holds for GMININF and VMININF. However, no such result is known for CMININF and MININF. In the latter cases we need to consider the weak co-clones $\langle S\rangle_{\nexists}$ instead of the usual clones $\langle S\rangle$. To be able to perform the required complexity analysis, we need reduction theorems between minimal inference problems parametrized by different sets of relations. Since our sets of relations $\langle S\rangle_{\nexists}$ are not closed under existential quantification, we cannot have the usual reduction theorem based on inclusion of polymorphisms.

Proposition 4.4 *Let R be a Boolean relation and S a set of relations. If $R \in \langle S\rangle_{\nexists}$ then there exists a polynomial many-one reduction from* MININF(R) *to* MININF(S)*, as well as from* CMININF(R) *to* CMININF(S)*.*

Proof: Suppose $\varphi \models_{\min} \psi$ is an instance of MININF(R), where φ is a conjunction of constraints built upon the relation R. Since $R \in \langle S\rangle_{\nexists}$ holds, every constraint $R(x_1, \dots, x_k)$ can be written as a conjunction of constraints upon relations from S. Substitute the latter into φ, obtaining the new formula φ'. Now, $\varphi' \models_{\min} \psi$ is an instance of MININF(S), where φ' is only polynomially larger than φ. It is clear that φ and φ' have the same variables and therefore also the same models. Hence, also the minimal models of φ and φ' are the same. The same reasoning also holds for CMININF. □

It is more convenient to work only with irredundant relations. It is intuitively clear that the complexity of GMININF(S) and GMININF(S°) for a set of relations S and its irredundant reduction S° is the same. Moreover, an equivalent result holds for the other three variants of the problem. For GMININF and VMININF it follows from the Galois correspondence. For CMININF and MININF we have the following result.

Proposition 4.5 *If S° is the irredundant reduction of a set of relations S, then* MININF(S) *and* MININF(S°)*, as well as* CMININF(S) *and* CMININF(S°)*, are equivalent under polynomial many-one reductions.*

Proof: Since both sets $\langle S\rangle_{\nexists}$ and $\langle S^\circ\rangle_{\nexists}$ are closed under variable permutation, the order of variables is irrelevant.

We perform the proof only for MININF, since the generalization to CMININF is the same. The reduction MININF$(S^\circ) \leq^p_m$ MININF(S) is straightforward. If $R^\circ \in S^\circ$ is constructed from R by identification of columns i, $k+1$ and φ° contains the constraint $R^\circ(x_1, \dots, x_k)$, then φ is equal to φ° in which we replace $R^\circ(x_1, \dots, x_k)$ by $R(x_1, \dots, x_i, \dots, x_k, x_i)$. It is clear that φ° and φ have the same models. Similarly, if $R^\circ \in S^\circ$ is constructed from $R = R^\circ \times \{0,1\}^p$ for a $p \geq 1$ and φ° contains the constraint $R^\circ(x_1, \dots, x_k)$, then φ is equivalent to φ° in which we replace $R^\circ(x_1, \dots, x_k)$ by $R(x_1, \dots, x_k, x_{k+1}, \dots, x_p)$, where $x_{k+1}, \dots, x_p$ are fresh variables. It is clear that m° is a model of φ° if and only if $m^\circ \cdot m'$ is a model of φ for all $m' \in \{0,1\}^p$. Hence, it is clear that m° is a minimal model of φ° if and only if $m^\circ \cdot 0 \cdots 0$ with p trailing zeros is a minimal model of φ.

Conversely, let φ contain the constraint $R(x_1, \dots, x_i, \dots, x_j, \dots, x_k)$ built upon $R \in S$. If the reduction R° is constructed from R by identification of columns i and j, then we can replace in φ every occurrence of $R(x_1, \dots, x_i, \dots, x_j, \dots, x_k)$ by $R^\circ(x_1, \dots, x_i, \dots, x_{j-1}, x_{j+1}, \dots, x_k)$, followed by the replacement of x_j by x_i in the rest of the formula φ. It is clear that $m = (m[1], \dots, m[k])$ is a minimal model of $R(x_1, \dots, x_k)$ if and only if $m^\circ = (m[1], \dots, m[j-1], m[j+1], \dots, m[k])$ is a minimal model of $R^\circ(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k)$, since the relation R forces the coordinates i and j to be equal. Similarly, if R° is constructed from $R = R^\circ \times \{0,1\}^p$ for a $p \geq 1$ and φ contains the constraint $R(x_1, \dots, x_k, \dots, x_{k+p})$, then we can replace every occurrence of this constraint by $R^\circ(x_1, \dots, x_k)$. It is clear that $m_0 = (m[1], \dots, m[k], 0, \dots, 0)$ is a minimal model of $R(x_1, \dots, x_k, \dots, x_{k+p})$ if and only if $m_0^\circ = (m[1], \dots, m[k])$ is a minimal model of $R^\circ(x_1, \dots, x_k)$. □

As a consequence, we assume throughout the paper that all constraint languages are irredundant.

5 Minimal Inference With Free Variables and Constants (GMININF)

We want to derive in this section a complete characterization of complexity for the GMININF problem. Theorem 4.2 is valid in this context, therefore we can use the algebraic approach based on clones and co-clones, exploiting the Galois correspondence. Moreover, Theorem 4.3 is valid and hence we can further restrict our attention to idempotent co-clones. Only one tractable case for the GMININF problem is known, namely when the constraint language S consists only of relations that are both Horn and dual Horn [6]. We prove two new tractability results for the GMININF problem for the classes of bijunctive affine relations and Horn relations containing only negative literals. These results together with the algebraic approach, combined with well-known complexity results from the literature are sufficient to give a complete complexity classification for the GMININF problem. We start with a relevant result from [25].

Theorem 5.1 (Nordh & Jonsson [25]) *If the constraint language S is Horn, dual Horn, bijunctive, or affine, then* GMININF(S) *is in* coNP. *Otherwise,* GMININF(S) *is* Π_2P*-complete.*

Further we need to introduce some co-clones that will be of particular importance to us later.

Co-clone $i\mathrm{R}_2$: For $a \in \{0,1\}$, a Boolean function f is called a-reproducing if $f(a,\ldots,a) = a$. The clones R_a contain all a-reproducing Boolean functions and the clone R_2 contains all functions that are both 0-reproducing and 1-reproducing. Hence $i\mathrm{R}_2$ is the co-clone consisting of all relations closed under all functions that are both 0-reproducing and 1-reproducing. Note that functions satisfying $f(a,\ldots,a) = a$ for all a in its domain are usually called idempotent.

Co-clone $i\mathrm{D}_1$: The co-clone $i\mathrm{D}_1$ consists of all relations closed under the operations $\mathrm{aff}(x,y,z) = x+y+z$ (affinity) and $\mathrm{maj}(x,y,z) = (x \vee y) \wedge (y \vee z) \wedge (x \vee z)$ (majority). Hence $i\mathrm{D}_1$ is the co-clone of all binary affine relations, which can be generated by systems of linear equations with at most two variables over the Boolean ring $\mathbb{Z}_2$.

Co-clone $i\mathrm{S}_1$: The clone S_1 consists of all 1-separating functions (see [2] for the definition). Dalmau proved in [9, Lemma 39] that $i\mathrm{S}_1$ is the co-clone consisting only of relations of the form $\{0,1\}^n \smallsetminus (1,\ldots,1)$. That is, the co-clone $i\mathrm{S}_1$ consists of all relations corresponding to conjunctions of clauses where all literals are negative, also called negative Horn.

As we have seen in Theorem 4.3, co-clones of the form $\langle S^{\mathrm{id}} \rangle$ are of particular importance to us. Remember that GMININF(S) has the same complexity as GMININF(S^{id}). It is easy to see that a co-clone S is idempotent if and only if it is closed under all a-reproducing functions, i.e., that $i\mathrm{R}_2 \subseteq S$ holds. Hence we can derive the following lemma.

Lemma 5.2 *Let S_1 and S_2 be co-clones such that S_2 is the least upper bound of $i\mathrm{R}_2$ and S_1 in Post's lattice of co-clones. Then* GMININF(S_1) *and* GMININF(S_2) *are polynomially equivalent.*

Proof: Remember that $i\mathrm{R}_2$ is the co-clone of all relations closed under all 0- and 1-reproducing functions. Thus, $\{\{(0)\},\{(1)\}\} \subseteq i\mathrm{R}_2$. If the set F contains a non-reproducing function f, then $\{\{(0)\},\{(1)\}\} \nsubseteq \mathrm{Inv}(F)$. Therefore given a co-clone S. we have $\{\{(0)\},\{(1)\}\} \subseteq S$ if and only if $i\mathrm{R}_2 \subseteq S$. It follows that the least upper bound of $i\mathrm{R}_2$ and S_1 in the lattice of co-clones is $\langle S_1^{\mathrm{id}}\rangle = S_2$. Using Theorem 4.3 we get that GMININF(S_1) and GMININF(S_2) are polynomially equivalent. □

Now we present our new tractable cases for the GMININF problem. The first is the bijunctive affine case.

Proposition 5.3 GMININF($i\mathrm{D}_1$) *is decidable in polynomial time.*

Proof: Instead of GMININF, we will consider the corresponding problem GMINEXT, proving its tractability. Since GMININF is the dual of GMINEXT, its tractability follows.

The bijunctive affine relations are model sets of affine systems containing equations of the following four types:

$$x+y=0, \quad x+y=1, \quad x=0, \quad x=1.$$

Equations of type $x+y=0$ imply $x=y$. Hence we can (1) eliminate the relation generated by this equation, (2) identify the variables x and y in the rest of the problem, and (3) check its compatibility with the partial assignment m.

The equations of the type $x=0$ and $x=1$ fix a value for a given variable. Hence, we can (1) eliminate the relation generated by this equation, (2) substitute the given value for x in the rest of the problem, and (3) check its compatibility with the partial assignment m.

If we have two equations $x+y=1$ and $y+z=1$ with one overlapping variable y, then we can replace the second equation by $x+z=0$. The conjunction $(x+y=1)\wedge(x+z=0)$ is equivalent through Gaussian elimination to the original conjunction $(x+y=1)\wedge(y+z=1)$. We can treat now the equation $x+z=0$ as above.

At the end, we get a conjunction of relations built only from the relation $[x+y=1]$, where each variable occurs in the system exactly once. Check if the partial assignment m, modulo aforementioned variable identifications, is extensible to a satisfying assignment. If YES, the partial assignment m can be extended to a minimal one, OTHERWISE it cannot be extended. □

We consider now the case of relations generated by clauses with only negative literals, also called negative Horn.

Proposition 5.4 GMININF($i\mathrm{S}_1$) *is decidable in polynomial time.*

Proof: Once more we consider the problem GMINEXT instead of GMININF. Recall that the co-clone $i\mathrm{S}_1$ consists of all relations corresponding to negative Horn formulas φ. i.e., relations of the form $\{0,1\}^n \smallsetminus \{1\cdots1\}$. Since φ is negative Horn, it is easy to decide whether a partial assignment m can be extended to a satisfying assignment $\bar{m}$. It should also be clear that such a partial assignment can be extended to a minimal assignment if and only if all variables from P, which are assigned by this partial assignment m, are assigned to the value 0. This is because in any satisfying assignment m' with $m'(x)=1$ there exists another satisfying assignment of φ which coincides with m' on all variables except x and where x is assigned to 0. Hence, GMINEXT($i\mathrm{S}_1$) and therefore also GMININF($i\mathrm{S}_1$) are both polynomial-time decidable. □

The following proposition presents the complexity of GMININF(S) for 8 particular co-clones.

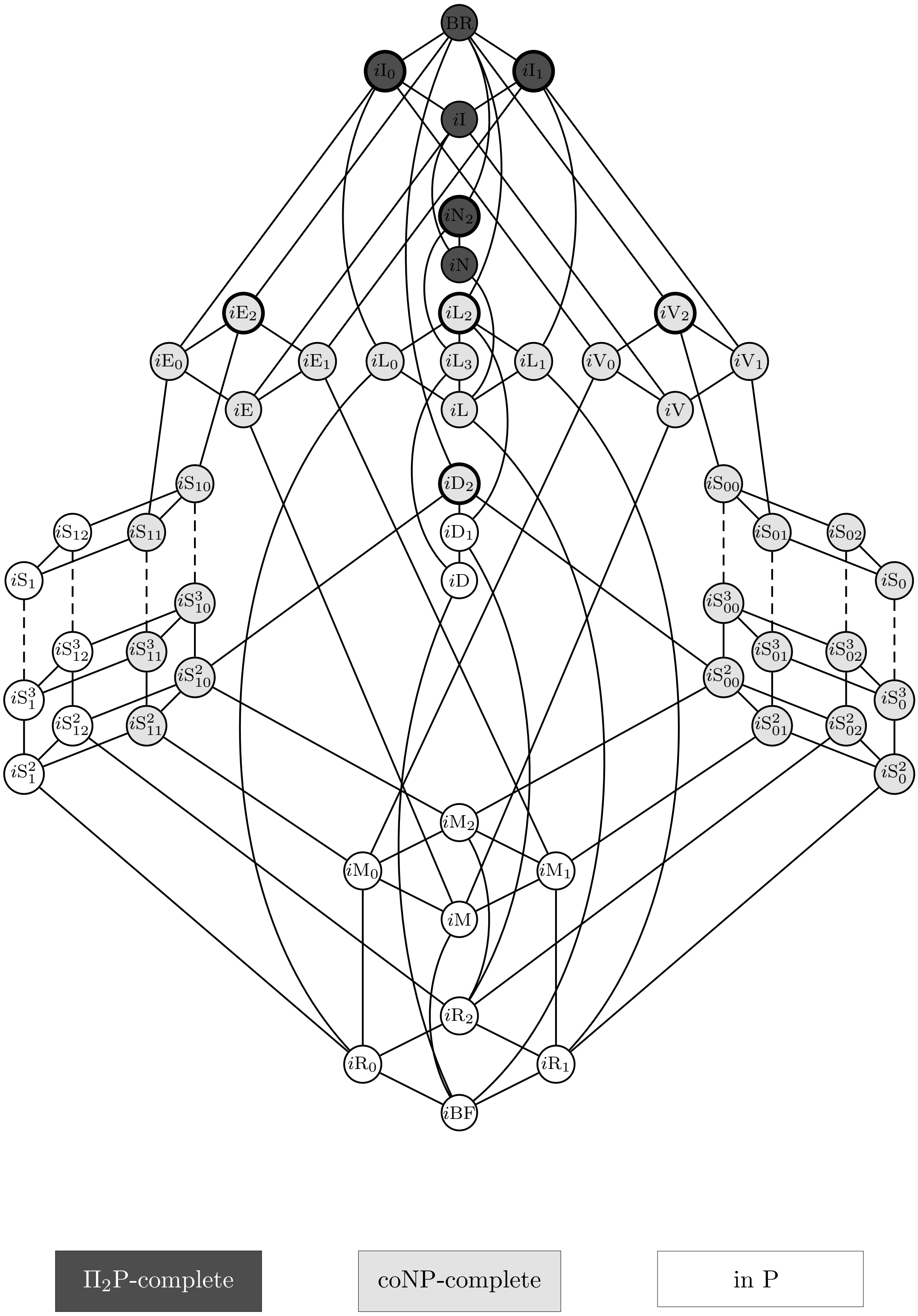

Figure 2: Trichotomy of GMININF

Proposition 5.5 GMININF(S) *is* coNP*-complete when S is one of the following co-clones:* $i\mathrm{S}_{11}^2$*,* $i\mathrm{S}_0^2$*,* $i\mathrm{L}$*,* $i\mathrm{V}$*, or* $i\mathrm{E}$*.* GMININF(S) *is polynomial-time decidable when S is one of the following co-clones:* $i\mathrm{S}_{12}$*,* $i\mathrm{D}_1$*, or* $i\mathrm{M}_2$*.*

Proof: The least upper bound of $i\mathrm{S}_{11}^2$ and $i\mathrm{R}_2$ is $i\mathrm{S}_{10}^2$. The co-clone $i\mathrm{S}_{10}^2$ contains all relations generated by Horn formulas with at most 2 literals per clause. Cadoli and Lenzerini proved in [6] that GMININF($i\mathrm{S}_{10}^2$) is coNP-complete, hence by Lemma 5.2 it follows that GMININF($i\mathrm{S}_{11}^2$) is coNP-complete.

The co-clone $i\mathrm{S}_0^2$ contains the relations generated by formulas with clauses of the form $(x \vee y)$, i.e., containing two positive literals. Cadoli and Lenzerini proved in [6] that GMININF($i\mathrm{S}_0^2$) is coNP-complete.

The least upper bound of the co-clones $i\mathrm{L}$ and $i\mathrm{R}_2$ is $i\mathrm{L}_2$, the set of all relations generated by affine systems. Durand and Hermann proved in [11] that GMININF($i\mathrm{L}_2$) is coNP-complete, hence by Lemma 5.2 it follows that GMININF($i\mathrm{L}$) is coNP-complete.

The least upper bound of the co-clones $i\mathrm{V}$ and $i\mathrm{R}_2$ is $i\mathrm{V}_2$, the set of all relations generated by dual Horn formulas. Since $i\mathrm{S}_0^2 \subseteq i\mathrm{V}_2$ holds and GMININF($i\mathrm{S}_0^2$) is coNP-complete, it follows, using Lemma 5.2, that GMININF($i\mathrm{V}_2$) is coNP-complete.

The least upper bound of the co-clones $i\mathrm{E}$ and $i\mathrm{R}_2$ is $i\mathrm{E}_2$, the set of all relations generated by Horn formulas. Since $i\mathrm{S}_{11}^2 \subseteq i\mathrm{E}_2$ holds and GMININF($i\mathrm{S}_{11}^2$) is coNP-complete, using Lemma 5.2 it follows that GMININF($i\mathrm{E}$) is coNP-complete.

We proved in Proposition 5.4 that GMININF($i\mathrm{S}_1$) is in P. The least upper bound of the co-clones $i\mathrm{S}_1$ and $i\mathrm{R}_2$ is $i\mathrm{S}_{12}$, therefore using Lemma 5.2 it follows that GMININF($i\mathrm{S}_{12}$) is in P.

The membership of GMININF($i\mathrm{D}_1$) in P is proved in Proposition 5.3.

The co-clone $i\mathrm{M}_2$ consists of all relations closed under the monotone functions, i.e., the relations generated by formulas which are both Horn and dual Horn. Cadoli and Lenzerini proved in [6] that GMININF($i\mathrm{M}_2$) is in P. □

In the previous proof we refer to results of Cadoli and Lenzerini in [6] concerning GMININF. Of course, they did not consider the problem as we defined it in our paper, but proved equivalent results from which the required result concerning GMININF can be obtained by easy syntactic reformulations.

The previous proposition together with the structure of Post's lattice of co-clones and Theorem 5.1 yields the following trichotomy result for the complexity of GMININF(S).

Theorem 5.6 (Trichotomy of GMININF) *Let S be a nonempty finite set of logical relations. If every relation in S is Horn and dual Horn, or bijunctive affine, or negative Horn, then* GMININF(S) *is decidable in polynomial time. Else if every relation in S is dual Horn, bijunctive, or affine, then* GMININF(S) *is* coNP*-complete. Otherwise* GMININF(S) *is* Π_2P*-complete.*

Figure 2 shows the complexity results for GMININF(S) visualized in the lattice of co-clones. The co-clones colored in dark gray correspond to Π_2P-complete problems, those colored in light gray correspond to coNP-complete problems, and finally those colored in white correspond to polynomial-time decidable problems.

6 Minimal Inference With Free Variables But Without Constants (VMININF)

We want to derive in this section a characterization of the complexity for the VMININF problem under restrictions on the types of allowed constraints. Theorem 4.2 is valid in this context, therefore we can use the algebraic approach working with clones and co-clones. Since the polynomial-time decidable cases are known and a dichotomy between Π_2P-complete cases of VMININF and those included in coNP (almost) carries over from [21], it is the case that the most involved proofs concern the intermediate level, especially the case of affine relations. In fact, the coNP-hardness proofs for affine relations that we present here are a significant strengthening of the main result of Durand and Hermann from [11]. Moreover, our proofs are much simpler than the rather involved proof in [11].

6.1 Complementive or 1-valid Relations

Note first that Kirousis and Kolaitis show in [21] the existence of 1-valid relations S (i.e., relations from $i\mathrm{I}_1$) for which the problem MININF(S) is Π_2P-complete. Using Theorem 4.2 we obtain the following result.

Proposition 6.1 VMININF($i\mathrm{I}_1$) *is* Π_2P-*complete.*

We continue with complementive relations. Let $R \subseteq \{0,1\}^n$ be an n-ary Boolean relation, R^0 the closure of R under a 0-prefix, and $R^\neg$ the closure of R under negation, i.e.,

$$\begin{aligned} R^0 &= \{(0, a_1, \ldots, a_n) \mid (a_1, \ldots, a_n) \in R\} \\ R^\neg &= R \cup \{(1-a_1, \ldots, 1-a_n) \mid (a_1, \ldots, a_n) \in R\} \end{aligned}$$

Define $R^{0\neg} = (R^0)^\neg$ to be the $(n+1)$-ary relation produced from R by a 0-prefix and closure under negation.

Lemma 6.2 *If* $\mathrm{Pol}(R) = \mathrm{I}_2$ *then* $\mathrm{Pol}(R^\neg) = \mathrm{N}_2$.

Proof: It is clear that $R^\neg$ is closed under negation, resulting from the closure construction. We investigate the possibility of $\mathrm{Pol}(R^\neg)$ to reside in a larger clone than N_2.

Suppose that $\mathrm{Pol}(R^\neg) \supseteq \mathrm{N}$, hence $R^\neg$ is both 0- and 1-valid. Therefore R must be 0- or 1-valid, what is a contradiction.

Suppose that $\mathrm{Pol}(R^\neg) \supseteq \mathrm{L}_3$. Consider the existence of three vectors a, b, and c from $R^\neg$. From the pigeonhole principle it follows that either (1) $a, b \in R$ and $\bar{c} \in R^\neg \smallsetminus R$, or (2) $a \in R$ and $\bar{b}, \bar{c} \in R^\neg \smallsetminus R$. If the second choice has been made, then there exist the vectors $\neg a \in R^\neg \smallsetminus R$ and $\neg\bar{b}, \neg\bar{c} \in R$, hence we can restrict ourselves to the first case.

Let $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$, and $\bar{c} = (1-c_1, \ldots, 1-c_n)$ be the considered vectors for some $c = (c_1, \ldots, c_n) \in R$. The co-clone $i\mathrm{L}_3$ is closed under the operation $x + y + z + 1 \mod 2$, therefore we must have the vector $d = (d_1, \ldots, d_n) \in R^\neg$, where

$$\begin{aligned} d_i &= a_i + b_i + (1 + c_i) + 1 \mod 2 \\ &= a_i + b_i + c_i \mod 2 \end{aligned}$$

for all $i = 1, \ldots, n$. If the vector d belongs to R, then R must be included in $i\mathrm{L}_2$, what is a contradiction. If $d \in R^\neg \setminus R$, then the complement $\bar{d} = (1 - d_1, \ldots, 1 - d_n)$ must be in R, i.e., $\bar{d}_i = 1 - d_i = d_i + 1 = a_i + b_i + c_i + 1 \mod 2$. But then R must be included in $i\mathrm{L}_3$, what is once more a contradiction. □

Lemma 6.3 *If* $\mathrm{Pol}(R) = \mathrm{I}_2$ *then* $\mathrm{Pol}(R^0) = \mathrm{I}_2$.

Proof: Let R be an n-ary Boolean relation and R^0 its $(n + 1)$-ary extension by the 0-prefix. Suppose that R^0 is closed under a Boolean k-ary function f other than the identity. This means that

$$\big(f(a_1[0], \ldots, a_k[0]),\ f(a_1[1], \ldots, a_k[1]),\ \ldots,\ f(a_1[n], \ldots, a_k[n])\big) \in R^0$$

for any not necessarily distinct k-tuple of vectors $a_1, \ldots, a_k \in R^0$. However, this also implies

$$\big(f(a_1[1], \ldots, a_k[1]),\ \ldots,\ f(a_1[n], \ldots, a_k[n])\big) \in R$$

what is in contradiction with $\mathrm{Pol}(R) = \mathrm{I}_2$. □

Proposition 6.4 VMININF$(i\mathrm{N}_2)$ *is* $\Pi_2\mathrm{P}$*-complete.*

Proof: We know from [13] that there exists a relation R for which MININF(R) is Π_2P-complete, hence MINEXT(R) is Σ_2P-complete. Let

$$\varphi(\vec{x}, \vec{y}) \quad = \quad \bigwedge_{i=1}^{k} R(\vec{x}, \vec{y})$$

be a CNF-formula and $m = (m_1, \ldots, m_n)$ a partial assignment to the variables $\vec{x}$. Let

$$\varphi'(z, \vec{x}, \vec{y}) \quad = \quad \bigwedge_{i=1}^{k} R^{0\neg}(z, \vec{x}, \vec{y})$$

be a CNF-formula and $m' = (0, m_1, \ldots, m_n)$ a partial assignment to the variables $(z, \vec{x})$. It is clear that the vector m can be extended to a minimal assignment of φ if and only if m' can be extended to a minimal assignment of φ'. Indeed, if $\mathrm{Pol}(R) = \mathrm{I}_2$ then $\mathrm{Pol}(R_i^{0\neg}) = \mathrm{N}_2$ for each $i = 1, \ldots, k$, following Lemmas 6.3 and 6.2. Since the formula φ' is a conjunction of complementive relations, using Theorem 4.2 this proves the result. □

6.2 0-valid Relations

We begin by observing that VMININF for 0-valid relations is in coNP.

Proposition 6.5 VMININF$(i\mathrm{I}_0)$ *is in* coNP.

Proof: Any formula $\varphi(x_1, \ldots, x_n)$ built over a set of 0-valid relations has at least one minimal model m which is $m(x_i) = 0$ for all i. Suppose φ is the theory, ψ the clause to be minimally inferred and P, Z the partition of the variables into variables to be minimized and free variables, respectively. If ψ contains a literal $\neg x$ such that $x \in P$, then obviously ψ is true in any minimal model of φ. If ψ consists only of positive literals, then ψ cannot be minimally inferred from φ since the all-zero minimal model is not a model of ψ. Otherwise, let φ_0 (ψ_0) be the formula (clause) resulting from assigning 0 to all the variables in P. Obviously ψ can be minimally inferred from φ if and only if ψ_0 can be (classically) inferred from φ_0. The result follows since classical inference is in coNP. □

Proposition 6.6 VMININF(iN) *is* coNP-*complete.*

Proof: We present a polynomial-time reduction from the NP-complete problem NAE-3-SAT [35], i.e., SAT(nae) where $nae = \{0,1\}^3 \smallsetminus \{000, 111\}$, to the complement of VMININF(iN). Given an instance $\rho(X)$ of SAT(nae), let (φ, ψ, P, Z) be the instance of VMININF(iN) where $Z = X \cup \{x_q\}$, $P = \{x_m\}$, and $\psi = \neg x_q$ (i.e., the clause to be inferred is the single negative literal $\neg x_q$), where x_q, x_m are fresh variables not occurring in X. Let R be the following 5-ary relation

$$R \quad = \quad \{00000, 11111\} \cup \{0,1\} \times nae \times \{0,1\}.$$

which is in iN since it is 0-valid, 1-valid, and complementive. For each constraint $nae(x_i, x_j, x_k)$ in $\rho(X)$ let $R(x_m, x_i, x_j, x_k, x_q) \in \varphi$. Now ψ cannot be minimally inferred from φ if and only if $\varphi \wedge \neg x_m \wedge x_q$ is satisfiable. This is because R is 0-valid, and thus, x_m must have the value 0 in every minimal model of φ. It is easy to realize that $\varphi \wedge \neg x_m \wedge x_q$ is satisfiable if and only if the original instance $\rho(X)$ of SAT(nae) is satisfiable. □

The following result follows directly from co-clone inclusions in Post's lattice, Theorem 4.2 and Propositions 6.5 and 6.6.

Corollary 6.7 *If* $i\mathrm{N} \subseteq \langle S \rangle \subseteq i\mathrm{I}_0$ *holds then* VMININF(S) *is* coNP-*complete.*

6.3 Dual Horn and Bijunctive Relations

For the instances of VMININF included in coNP, it is more convenient to consider the dual instances VMINEXT in NP. We consider the problem VMINEXT($i\mathrm{S}_0^2$). Note that the clone S_0^2 includes the clone D_2, corresponding to bijunctive relations, as well as the clone V_1 corresponding to 1-valid dual Horn relations. We need first a lemma that exhibits an NP-complete problem for a relation belonging to the co-clone $i\mathrm{S}_0^2$, which was already proved by Cadoli and Lenzerini in [6, Theorem 5].

Lemma 6.8 (Cadoli & Lenzerini [6]) MINEXT(R) *is* NP-*complete for* $R = [x \vee y]$*.*

Proposition 6.9 VMININF($i\mathrm{S}_0^2$) *is* coNP-*complete.*

Proof: We have $[x \vee y] \in i\mathrm{S}_0^2$ since $[x \vee y] = \{01, 10, 11\}$ is closed under both the implication function and the majority function. In the same time we have $[x \vee y] \notin \mathrm{R}_1$ since $\mathrm{R}_1 \supseteq \mathrm{E}_2$ holds and $[x \vee y]$ is *not* closed under conjunction, because $01 \wedge 10 = 00 \notin [x \vee y]$. The rest of the proof follows from Lemma 6.8, Theorem 4.2, and the correspondence between VMINEXT and VMININF. □

6.4 Affine Relations

We consider the problems VMININF($i\mathrm{L}_1$) as well as VMININF($i\mathrm{L}_3$). We need first some supporting lemmas to exhibit NP-complete problems built from relations in the corresponding co-clones. The result presented in the supporting lemma is a significant strengthening of the previous result due to Durand and Hermann [11], which states that there exists a set of affine relations S for which MININF(S) is coNP-complete. Moreover, the previous proof is rather involved in comparison to the proofs we present here.

Lemma 6.10 MINEXT(R) *is* NP*-complete for* $R = [x + y + z + w = 1]$.

Proof: The equation $x + y + z + w = 1$ gives rise to $x + y = 1$ when we identify the variables y, z, and w. The membership in NP is clear and we prove NP-hardness by a reduction from 3SAT.

Suppose $\varphi(x_1, \ldots, x_n) = c_1 \wedge \cdots \wedge c_k$ is a 3SAT formula. Associate the variables y_i with the clauses c_i for $i = 1, \ldots, k$. Construct the system φ' in the following way. Add the equations $x_i + x'_i = 1$ for $i = 1, \ldots, n$ to the constructed formula φ', what forces x_i and x'_i to take opposite values. For each clause $c_i = l_i^1 \vee l_i^2 \vee l_i^3$ add the following three equations

$$z_{3i-2} + v_i^1 + y_i + w = 1, \qquad z_{3i-1} + v_i^2 + y_i + w = 1, \qquad z_{3i} + v_i^3 + y_i + w = 1$$

to the formula φ', where

$$v_i^j = \begin{cases} x_p & \text{if } l_i^j = \neg x_p, \\ x'_p & \text{if } l_i^j = x_p. \end{cases}$$

The variable v_i^j is a placeholder for the negation $\neg l_i^j$ of the literal l_i^j. The variable w is the same all over the system.

Claim: The formula φ is satisfiable if and only if the assignment $s(y_i) = 1$ for all $i = 1, \ldots, k$ of φ' has a minimal extension.

Let φ be satisfiable. Then there exists a satisfying assignment m of the formula φ. Since every clause c_i is satisfied, for each i there must be a j such that $m(l_i^j) = 1$. Let $\bar{m}$ be an extension of m that satisfies φ'. Following the construction of φ', there must be a variable v_i^j such that $\bar{m}(v_i^j) = 0$. Then there are three possibilities to get a minimal assignment $\bar{m}$:

1. When we set $\bar{m}(z_{3i+3-j}) = 1$, $\bar{m}(y_i) = 0$, and $\bar{m}(w) = 0$, we will get an assignment which is not interesting for us.

2. When we set $\bar{m}(z_{3i+3-j}) = 0$, $\bar{m}(y_i) = 0$, and $\bar{m}(w) = 1$, we will also get an assignment which is not interesting for us.

3. When we set $\bar{m}(z_{3i+3-j}) = 0$, $\bar{m}(y_i) = 1$, and $\bar{m}(w) = 0$, we get an assignment which is an extension of the assignment $s(y_i) = 1$ for all $i = 1, \ldots, k$.

These three possible assignments are clearly incomparable and no value can be changed from 1 to 0 to get another satisfying assignment of φ'. Therefore the assignment $\bar{m}$ from the third case is minimal.

Let φ be unsatisfiable. Then for each assignment m there exists a falsified clause c_i in φ. This means that we have $m(l_i^j) = 0$ for each $j = 1, 2, 3$. Let $\bar{m}$ be an extension of m that satisfies the formula φ'. Following the construction of φ', we have $\bar{m}(v_i^j) = 1$ for each j. Then we can set $\bar{m}(z_{3i+3-j}) = 0$, $\bar{m}(y_i) = 0$, and $\bar{m}(w) = 0$ for all j. This implies that the assignment $s(y_i) = 1$ for all i cannot be extended to a minimal one. □

Proposition 6.11 VMININF(iL$_3$) *is* coNP-*complete.*

Proof: The relation $[x + y + z + w = 1] = \{0001, 0010, 0100, 1000, 0111, 1011, 1101, 1110\}$ is affine and complementive. It is neither 1-valid nor 0-valid and it is *not* bijunctive since the closure under majority computes $\mathrm{maj}(0001, 0010, 0100) = 0000 \notin [x + y + z + w = 1]$. The proof is finished by means of Lemma 6.10, Theorem 4.2, and the correspondence between VMINEXT and VMININF. □

Lemma 6.12 MINEXT(R) *is* NP-*complete for the relations* $R = [(x + y + z = 1) \wedge (x + w = 1)]$ *and* $R = [(x + y + z = 0) \wedge (x + w = 1)]$.

Proof: The membership in NP is clear, we focus on the NP-hardness proof by means of a reduction from 3SAT. Suppose $\varphi(x_1, \ldots, x_n) = c_1 \wedge \cdots \wedge c_k$ is a 3SAT formula. Associate the variables y_i with the clauses c_i for $i = 1, \ldots, k$. Construct the system φ' in the following way.

For each clause $c_i = l_i^1 \vee l_i^2 \vee l_i^3$ add the following three formulas

$$(z_{3i-2}+v_i^1+y_i = 1)\wedge(v_i^1+u_i^1 = 1),\ (z_{3i-1}+v_i^2+y_i = 1)\wedge(v_i^2+u_i^2 = 1),\ (z_{3i}+v_i^3+y_i = 1)\wedge(v_i^3+u_i^3 = 1)$$

to the formula φ', where

$$v_i^j = \begin{cases} x_p & \text{if } l_i^j = \neg x_p, \\ x'_p & \text{if } l_i^j = x_p. \end{cases} \quad \text{and} \quad u_i^j = \begin{cases} x'_p & \text{if } l_i^j = \neg x_p, \\ x_p & \text{if } l_i^j = x_p. \end{cases}$$

The variable v_i^j is a placeholder for the negation $\neg l_i^j$ of the literal l_i^j and vice versa for u_i^j.

Claim: The formula φ is satisfiable if and only if the assignment $s(y_i) = 1$ for all $i = 1, \ldots, k$ of φ' has a minimal extension.

Let φ be satisfiable. Let m be a satisfying assignment of φ. Then in each clause c_i there must be a literal l_i^j such that $m(l_i^j) = 1$. Hence, for each extension $\bar{m}$ of m that satisfies φ' we must have $\bar{m}(v_i^j) = 0$ following the definition of v_i^j, where the equation $x + x' = 1$ enforces the variables x and x' to take opposite values. The value $\bar{m}(v_i^j) = 0$ implies two different incomparable assignments for the variables y_i and z_{3i+3-j}. The first one is $\bar{m}(z_{3i+3-j}) = 1$ and $\bar{m}(y_i) = 0$, which is not interesting for us. The second is $\bar{m}(z_{3i+3-j}) = 0$ and $\bar{m}(y_i) = 1$, which is the desired minimal assignment. It is clear from the construction of the formula φ' that this assignment is minimal. It is also clear that $\bar{m}$ is an extension of the assignment $s(y_i) = 1$ for all $i = 1, \ldots, k$. Hence, the assignment $s(y_i) = 1$ for all i can be extended to a minimal one.

Let φ be unsatisfiable. Then for all assignments m of the variables V there exists a falsified clause c_i. This implies that $m(l_i^j) = 0$ for all $j = 1, 2, 3$. Let $\bar{m}$ be an extension of m that satisfies the formula φ'. Then the structure of the formula φ' implies that we have $\bar{m}(v_i^1) = \bar{m}(v_i^2) = \bar{m}(v_i^3) = 1$. This implies the existence of a minimal assignment with $y_i = 0$ and $\bar{m}(z_{3i+3-j}) = 0$ for all j. Hence, the assignment $s(y_i) = 1$ for all i cannot be extended to a minimal one.

Finally for the latter relation, by swapping variables we get $[(x + y + z = 0) \wedge (x + w = 1)] = [(w + y + z = 1) \wedge (x + w = 1)]$ and the result follows from the previous relation. □

Lemma 6.13 MINEXT(R) *is* NP*-complete for* $R = [x+y+z = 1]$ *and* $R = [(x+y+z = 1) \wedge (w = 1)]$.

Proof: We proceed by dropping the relation $[x + y = 1]$ from S in Lemma 6.12. This means that we reformulate the proof using only the relation $[x+y+z = 1]$. Write the equation $x_i'' + x_i + x_i' = 1$ instead of $x_i + x_i' = 1$ for each i. Add also the equation $x_i'' + u_i + w_i = 1$ for each i, where u_i and w_i are new variables.

Define the partial assignment to be $s(y_i) = 1$ and $s(u_i) = 1$ for each i. Let m be a minimal assignment of the formula φ'. Suppose that there exists an i, such that $m(x_i'') = 1$. Then we have also $m(w_i) = 1$. But then we can construct a new assignment m' with $m'(u_i) = 0$, $m'(w_i) = 0$, and $m'(x_i'') = 1$. This is a contradiction to the assumption that the partial assignment s can be extended to a minimal one. Hence, for all i we must have $m(x_i'') = 0$.

The relation $R'(x, y, z, w) = [(x + y + z = 1) \wedge (w = 1)]$ can be constructed from the relation $R(x, y, z) = [x + y + z = 1]$ as $R'(x, y, z, w) = R(x, y, z) \wedge R(w, w, w)$. Since the value of the variable w is determined by the trivial equation $w = 1$, the complexity of MINEXT(R') is equal to that of MINEXT(R). □

Proposition 6.14 VMININF$(i\mathrm{L}_1)$ *is* coNP*-complete.*

Proof: For the relation $[x+y+z = 1]$ from Lemma 6.13 we know that $[x+y+z = 1] \in i\mathrm{L}_1$, since the relation is linear and 1-valid. We also know that $[x + y + z = 1] \notin i\mathrm{L}$, since $[x + y + z = 1]$ is not 0-valid. Finally, $[x + y + z = 1] \notin i\mathrm{R}_1$ since the relation $[x + y + z = 1] = \{001, 010, 100, 111\}$ is not closed under disjunction, i.e., $001 \vee 010 = 011 \notin [x + y + z = 1]$. The proof is finished by means of Lemma 6.13, Theorem 4.2, and the relation between VMINEXT and VMININF. □

6.5 Polynomial-Time Decidable Cases

The polynomial-time decidable cases for GMINEXT(S) propagates to VMINEXT(S). There are three other cases of VMINEXT(S) that are also decidable in polynomial time.

Tractability of VMINEXT(S) when for sets of Horn relations S was already shown by Cadoli and Lenzerini in [6]. This implies the following result.

Proposition 6.15 (Cadoli & Lenzerini [6]) VMININF$(i\mathrm{E}_2)$ *is decidable in polynomial time.*

Proposition 6.16 VMININF$(i\mathrm{L}_0)$ *and* VMININF$(i\mathrm{V}_0)$ *are in* P.

Proof: We only give the proof for $i\mathrm{L}_0$ since the proof for $i\mathrm{V}_0$ uses exactly the same argument.

Given an instance (φ, ψ, P, Z) of VMININF$(i\mathrm{L}_0)$, we know that the all-zero assignment $0 \cdots 0$ is a minimal model of φ since $i\mathrm{L}_0$ is 0-valid. If ψ contains a negative literal $\neg x$ such that $x \in P$, then

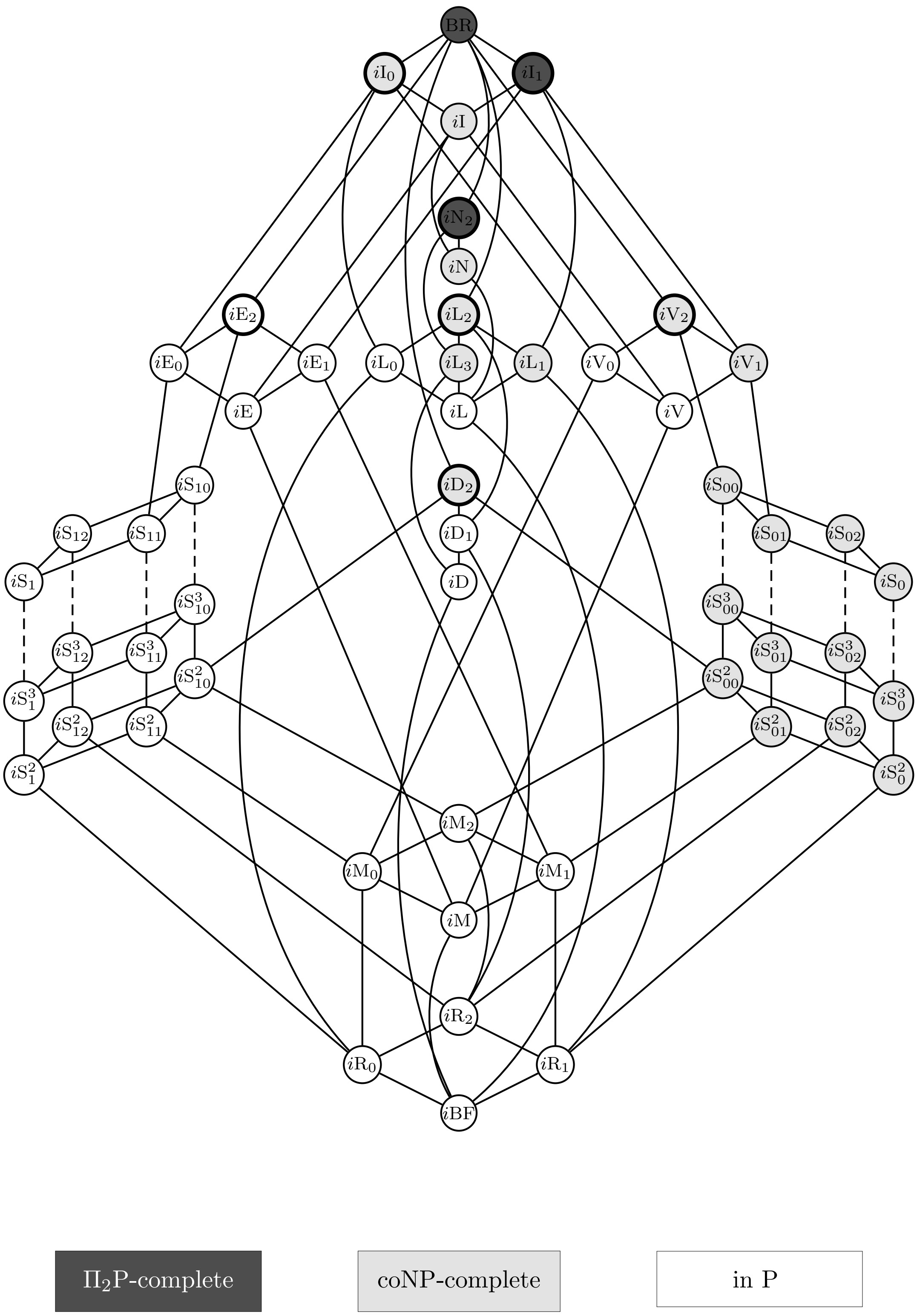

Figure 3: Trichotomy of VMININF

obviously ψ is true in any minimal model of φ. If ψ consists only of positive literals, then ψ cannot be minimally inferred from φ since the all-zero minimal model is not a model of ψ.

Otherwise, we can reduce the problem by assigning 0 to all the variables in P; resulting in the equivalent instance $(\varphi_0, \psi_0, \emptyset, Z)$. The clause ψ can be minimally inferred from φ if and only if ψ_0 can be (classically) inferred from φ_0. Note that φ_0 is still a formula built from relations in $i\mathrm{L}_0$ (i.e., affine and 0-valid), and since inference over affine constraints is polynomial-time decidable it follows that VMININF($i\mathrm{L}_0$) is in P. □

6.6 Main Result

We put all results of this section together and we get the following trichotomy result for minimal inference with free variables and without constants.

Theorem 6.17 (Trichotomy of VMININF) *Let S be a nonempty finite set of logical relations. If every relation in S is Horn, or affine and 0-valid, or dual Horn and 0-valid, or bijunctive affine, then* VMININF(S) *is decidable in polynomial time. Else if every relation in S is 0-valid, or dual Horn, or bijunctive, or affine, then* VMININF(S) *is* coNP*-complete. Otherwise* VMININF(S) *is* Π_2P*-complete.*

Figure 3 shows the complexity results for VMININF(S) visualized in the lattice of co-clones. The co-clones colored in dark gray correspond to Π_2P-complete problems, those colored in light gray correspond to coNP-complete problems, and finally those colored in white correspond to polynomial-time decidable problems. If we compare the lattices in Figures 2 and 3, we see that the absence of constants can sometimes make the minimal inference problem easier.

7 Expressivity Techniques

The problems CMININF and MININF are not compatible with the co-clone closure $\langle S \rangle$ of a set of relations S any more, since the application of existential quantification to a constraint can change the order of the models in the corresponding relation, but only with the weak co-clone closure $\langle S \rangle_{\nexists}$. For this reasons we need to review in this section techniques compatible with the weak co-clone closure $\langle S \rangle_{\nexists}$. The first is due to Schaefer [35], for showing that certain relations belong to $\langle S \rangle_{\nexists}$ based on information about the polymorphisms of S. The second one, namely the 1-restriction, is due to Kirousis and Kolaitis [21] and it was already mentioned in the preliminaries. We begin by stating some basic but useful facts about the set $\langle S \rangle_{\nexists}$.

Lemma 7.1 *If S is not closed under the functions $f_1, \ldots, f_k$ then there exists a single relation $R \in \langle S \rangle_{\nexists}$ which is not closed under the functions $f_1, \ldots, f_k$.*

Proof: Since S is not closed under the functions $f_1, \ldots, f_k$, there exist relations $R_1, \ldots, R_k$ in S such that R_i is not closed under f_i. Let R be the Cartesian product of the relations $R_1, \ldots, R_k$. Obviously R can be expressed as the conjunction $R_1(x_{11}, \ldots, x_{1m}) \wedge \cdots \wedge R_k(x_{k1}, \ldots, x_{kn})$, where $x_{11}, \ldots, x_{kn}$ are all distinct variables. Hence $R \in \langle S \rangle_{\nexists}$ holds and R is not closed under the functions $f_1, \ldots, f_k$. □

If a relation R is not closed under several function, then the witnesses of this non-closure are usually different vectors in R. For instance, if R is neither Horn nor dual Horn, then there exist

$a, b, c, d \in R$, such that $a \wedge b \notin R$ and $c \vee d \notin R$, where the sets $\{a, b\}$ and $\{c, d\}$ are not necessarily equal. The following lemma shows that we can always find a new relation $R' \in \langle R \rangle_{\nexists}$, such that the witnesses of the non-clorure for all functions are the same.

Lemma 7.2 *If a relation R is not closed under the functions $f_1, \ldots, f_k$ of arities $a_1, \ldots, a_k$ where a denotes the maximum arity of the functions, then there is a relation $R' \in \langle R \rangle_{\nexists}$ containing tuples $t_1, \ldots t_a$, such that for any function f_i and any a_i tuples from $t_1, \ldots, t_a$, the tuple resulting from applying f_i to these a_i tuples (in any order), is not in R'.*

Proof: We prove the result by induction on the number of functions k. In the base case we have one function f_1, and we let R' be $R^{a_1!}$. Since we know that R is not closed under f_1 there exist a_1 tuples $t_1, \ldots, t_{a_1}$ such that $f_1(t_1, \ldots, t_{a_1}) \notin R$. Let S_{a_1} denote the group of permutations of $1, \ldots, a_1$, and denote the $a_1!$ permutations by $\pi_1, \ldots, \pi_{a_1!}$. Now construct the a_1 tuples by, for each $1 \leq i \leq a_1$, constructing the tuple $t'_i = t_{\pi_1(i)} \times \cdots \times t_{\pi_{a_1!}(i)}$. Clearly, f_1 applied to the tuples $t'_1, \ldots, t'_{a_1}$ in any order results in a tuple which is not in R'.

Now, assuming the result holds for $k-1$ functions $f_1, \ldots, f_{k-1}$ we show that it must hold also for k functions $f_1, \ldots, f_k$. Assume without loss of generality that f_j, for some $j < k$, is a function of maximum arity among $f_1, \ldots, f_k$. By assumption we know that there is a relation $R_1 \in \langle R \rangle_{\nexists}$ and $a_j = a$ tuples $t'_1, \ldots, t'_{a_j}$ in R_1 such that for any function f_i, for some $i < k$, the function f_i applied to any a_i tuples from $t'_1, \ldots, t'_{a_j}$ (in any order) results in a tuple which is not in R_1. We also know that there is a relation $R_2 \in \langle R \rangle_{\nexists}$ and a_k tuples such that f_k applied to these tuples in any order results in a tuple which is not in R_2. By assumption $a_j \geq a_k$. By taking Cartesian products of R_2 we get a relation R'_2 containing at least a_j tuples $s'_1, \ldots, s'_{a_j}$ such that f_k applied to any a_k of these a_j tuples (in any order), results in a tuple that is not in R'_2. Now take the Cartesian product $R_1 \times R'_2$ and consider the tuples $t'_1 \cdot s'_1, \ldots, t'_{a_j} \cdot s'_{a_j}$ (where $\cdot$ denotes tuple concatenation). By the reasoning above we have that $f_i \in \{f_1, \ldots, f_k\}$ applied (in any order) to any a_i tuples from $t'_1 \cdot s'_1, \ldots, t'_{a_j} \cdot s'_{a_j}$ results in a tuple which is not in $R_1 \times R'_2$, and the result follows. □

The notation $R[x/V]$ means that in the relation R the variables V are replaced with the new variable x. As a shorthand, we also write $R[V]$ instead of $R[x/V]$ when the new variable x is implicitly clear or when V representss both the set of variables to be replaced and the fresh variable. This means that V is a set of variables, but also the identifier of a new variable, by which all variables in V are replaced in $R[V]$. Recall that $[R(x_1, \ldots, x_k)]$ denotes the relation generated by the constraint $R(x_1, \ldots, x_k)$.

The following is the typical situation we are faced with when proving our hardness results. We have a set of relations S for which we want to prove a hardness result. We also know a hardness result for a relation R. The goal is to show that $R \in \langle S \rangle_{\nexists}$. For this purpose we have some information on S, namely we know that S is closed under the operations $f_1, \ldots, f_j$, but it is *not* closed under the operations $g_1, \ldots, g_k$. The first step is to use Lemmas 7.1 and 7.2 to produce a single relation $R' \in \langle S \rangle_{\nexists}$ that has the properties stated in Lemma 7.2 and is not closed under any of the operations $g_1, \ldots, g_k$.

The next (crucial) step is to prove that R can be implemented by R' using variable identification. We first show that R' must contain a number of vectors $a_1, \ldots, a_p$, usually by the argument that R' would otherwise be closed under one of the functions $g_1, \ldots, g_k$. For all Boolean vectors m of length p we construct the variable sets $V_m = \{x \mid a_1(x) = m[1], \ldots, a_p(x) = m[p]\}$. We then construct the constraint $Q(V_0, \ldots, V_{2^p-1}) = R'[V_i \mid i = 0, \ldots, 2^p - 1]$. For instance, when we know the existence of the vectors $a, b, c \in R'$, we create the variable sets $V_{ijk} = \{x \mid a(x) = i, b(x) = j, c(x) = k\}$ and construct the constraint $Q(V_0, \ldots, V_7) = R'[V_{000}, \ldots, V_{111}]$.

Finally, we must show that the constructed constraint $R'[V_i \mid i = 0, \ldots, 2^p - 1]$, possibly after additional variable identification, really generates the relation R. For more information on this technique, see the monograph [8] or Schaefer's original exposition [35]. Note again that our implementations differ from Schaefer's since we are not allowed to use existential quantification to eliminate variables.

8 Minimal Inference Without Free Variables But With Constants (CMININF)

In this section we want to derive a complete complexity characterization of the CMININF problem. Theorem 4.2 is not valid any more in this context since there are for example sets of relations S such that CMININF(S) is in P but CMININF($\langle S \rangle$) is coNP-complete.

The complexity classification for the CMININF problem will be carried out in two phases. First we deal with sets of relations S that are neither Horn, dual Horn, bijunctive, nor affine. We call these types of sets S for non-Schaefer. Then we move on and classify the complexity of CMININF(S) for sets of relations S that are Schaefer (i.e., S which are Horn, dual Horn, bijunctive, or affine).

In the first phase we will use an approach already used by Kirousis and Kolaitis in [21], where we use 1-valid 0-sections, a well-known concept from coding theory and already used for circumscription [20] and minimal inference [21]. The following result is the main theorem in [21].

Proposition 8.1 (Kirousis & Kolaitis [21]) *Let S be a finite set of Boolean relations and S^* the corresponding 1-valid restriction. Then there exists a polynomial many-one reduction from* CMININF(S^*) *to* CMININF(S).

Proposition 8.2 *Let S be non-Schaefer. If S^* is Schaefer then* CMININF(S) *is coNP-complete, otherwise* CMININF(S) *is Π_2P-complete.*

Proof: Kirousis and Kolaitis proved in [21] that MININF(S^{id}) is Π_2P-complete when S^* is not Schaefer. We give a polynomial-time reduction from MININF(S^{id}) to CMININF(S) proving that CMININF(S) is Π_2P-complete when S^* is not Schaefer. Given an instance $\varphi \models_{\min} \psi$ of MININF(S^{id}) consisting of a conjunction of constraints φ over S^{id} and a clause to be inferred ψ. For all constraints of the form $x = 0$ ($x = 1$) we add the literal x ($\neg x$) to ψ, remove the constraint $x = 0$ ($x = 1$), and let x be fixed (i.e., add x to Q). The idea behind the reduction is as follows. If $x = 0$ is a constraint in φ, we remove it and modify ψ to make sure that every minimal model m_* of $\varphi \smallsetminus \{x = 0\}$ such that $m_*(x) = 1$, is a model of ψ, and in the case where $m_*(x) = 0$ we make sure that m_* is a model of the modified ψ if and only if m_* was a model of the original ψ. The case where $x = 1$ is a constraint in φ is handled in the same way. It should be clear that the resulting instance is equivalent to the original instance $\varphi \models_{\min} \psi$.

Kirousis and Kolaitis proved in [20] that the minimal model checking problem for circumscription with constants but no free variables is in P when S^* is Schaefer, and hence by well-known arguments CMININF(S) is in coNP when S^* is Schaefer. It remains to prove that CMININF(S) is coNP-hard when S is non-Schaefer. We give a reduction from UNSAT(S^{id}), the complement of the satisfiability problem, to CMININF(S). Since UNSAT(S^{id}) is coNP-complete when S is non-Schaefer [35], our result will follow. Given a conjunction of constraints φ over S^{id}, let φ' be $\varphi \wedge (x = 1)$ where x is a fresh variable. Then $\neg x$ can be inferred from φ' under propositional circumscription if and only if φ is not satisfiable. This gives a reduction from UNSAT(S^{id}) to MININF(S^{id}). Hence when combined with the reduction from MININF(S^{id}) to CMININF(S), we get the desired result that CMININF(S) is in coNP-hard when S is non-Schaefer. □

The following result appears in [6] in the form that ECWA-reasoning with constants for formulas that are both Horn and bijunctive (HK in their notation) is coNP-complete.

Proposition 8.3 (Cadoli & Lenzerini [6]) *The problem* CMININF(S) *is* coNP*-complete for* $S = \{[\neg x \vee \neg y],\ [\neg x \vee y]\}$.

8.1 Incomparable Relations

In this section we present a new tractable class for CMININF.

Proposition 8.4 *If a set of relations S is Schaefer and incomparable then* CMININF(S) *is in* P.

Proof: Suppose $\varphi \models_{\min} \psi$ is an instance of CMININF with S incomparable. It can be shown that $[\varphi]$ is an incomparable relation. This follows from the fact that incomparable relations are preserved under conjunction and variable identification. Given two formulas φ_1 and φ_2, each having only incomparable models, their conjunction $\varphi_1 \wedge \varphi_2$ has only incomparable models since a model of φ_1 can never be extended to two comparable models of $\varphi_1 \wedge \varphi_2$ (because φ_2 has only incomparable models). Variable identification only reduce the set of models of a formula. If the original set of models was incomparable, then of course the reduced set of models will also be incomparable. Also note that adding some variables to Q (i.e., letting them be fixed) can never make two incomparable models of a formula comparable.

Since φ is expressed by conjunction and variable identification over an incomparable set of relations S it follows that all models of φ are incomparable and hence also minimal. Thus, $\varphi \models_{\min} \psi$ if and only if $\varphi \models \psi$, where the latter can be checked in polynomial time since S is Schaefer. □

The following example shows the impact of incomparable relations on the tractability of CMININF compared to VMININF upon the same relation.

Example 8.5 Let $R = \{0110, 1001, 1100\}$ then R is a bijunctive relation but not other Schaefer and hence VMININF(R) is coNP-complete. On the other hand CMININF(R) is in P since R is incomparable. Let $R' = \{00111, 01010, 10001, 11100\}$ then R' is affine but not other Schaefer and hence VMININF(R') is coNP-complete. But again CMININF(R') is in P since R' is incomparable.

Proposition 8.6 CMININF(N) *for* $N = \{[\neg x \vee \neg y], [x \not\equiv y]\}$ *is in* P.

Proof: Suppose $\varphi \models_{\min} \psi$ is an instance of CMININF(N) for $N = \{[\neg x \vee \neg y], [x \not\equiv y]\}$. Check first that φ is satisfiable, what can be done in polynomial time since S is bijunctive. If φ is *not* satisfiable, then $\varphi \models_{\min} \psi$ is trivially satisfied.

If the formula φ is produced only from the relation $[\neg x \vee \neg y]$ then it is 0-valid and therefore $\varphi \models_{\min} \psi$ is in P.

The formula φ is a conjunction of binary constraints of the type $(\neg x \vee \neg y)$ and $(x \not\equiv y)$. Let us consider the structure of φ. If a variable v occurs only in negative constraints, i.e., only in constraints of the form $(\neg x \vee \neg y)$, then v has the value 0 in all minimal models of φ. Hence, we can drop all clauses $(\neg v \vee \neg y)$ from φ and remove the variable v from ψ. Repeat this replacement until all variables occur in a $\not\equiv$ constraint.

Let $\varphi' \models_{\min} \psi'$ be the resulting instance. Assume with the aim of reaching a contradiction that there exist two comparable models $m_0 < m_1$ of φ'. Then there exists a variable x, such that $m_0(x) = 0$ and $m_1(x) = 1$. Let y be a variable occurring together with x in a $(x \not\equiv y)$ constraint. Then we have $m_0(y) = 1$ and $m_1(y) = 0$, which implies that $m_0 \not< m_1$, constituting a contradiction. Hence, all models of φ' are incomparable. Therefore $\varphi' \models_{\min} \psi'$ holds if and only if $\varphi' \models \psi'$ which is decidable in polynomial time since φ' is bijunctive. □

Compare the result of Proposition 8.6 with the fact that VMININF(N) for $N = \{[\neg x \vee \neg y], [x \not\equiv y]\}$ is coNP-complete.

8.2 Dual Horn Relations

We apply Schaefer's technique to obtain the following representation lemma.

Lemma 8.7 *Let R be a Boolean relation which is dual Horn, but not Horn. Then we can construct the relation $[x \vee y]$ from R by variable identification and constant substitution.*

Proof: If R is dual Horn, but not Horn, then there exists two vectors $a, b \in R$, such that $a \wedge b \notin R$ and $a \vee b \in R$. Construct the sets of variables $V_{ij} = \{x \mid a(x) = i, b(x) = j\}$. Construct the constraint $Q(x, y) = R[0/V_{00}, x/V_{01}, y/V_{10}, 1/V_{11}]$. It is clear that the constraint $Q(x, y)$ generates the relation $\{01, 10, 11\} = [x \vee y]$. □

Proposition 8.8 CMININF(S) *for all sets S of relations, which are dual Horn but not Horn, is* coNP*-complete.*

Proof: From each set of relation S, which are dual Horn but not Horn, we can produce the relation $[x \vee y]$ according to Lemma 8.7. Cadoli and Lenzerini [6, Theorem 5] proved that MININF$([x \vee y])$ is coNP-complete. Hence, the result follows. □

8.3 Horn But Not Negative Horn Relations

We start with some representation lemmas proved by Schaefer's technique.

Lemma 8.9 *Let R be a Boolean relation which is Horn, but not dual Horn. Then we can construct the relation $[\neg x \vee \neg y]$ from R by variable identification and constant substitution.*

Proof: If R is Horn, but not dual Horn, then there exist two vectors $a, b \in R$, such that $a \vee b \notin R$ and $a \wedge b \in R$. Construct the variable sets $V_{ij} = \{x \mid a(x) = i, b(x) = j\}$. Construct the constraint $Q(x, y) = R[0/V_{00}, x/V_{01}, y/V_{10}, 1/V_{11}]$. It is clear that the constraint $Q(x, y)$ generates the relation $\{00, 01, 10\} = [\neg x \vee \neg y]$. □

Lemma 8.10 *Let R be a Boolean relation which is Horn, but not negative Horn. Then we can construct the relation $[\neg x \vee y]$ from R by variable identification and constant substitution.*

Proof: If R is Horn but not negative Horn, then there exist three vectors $a, b, c \in R$, such that $a \wedge b \in R$, $a \wedge c \in R$, $b \wedge c \in R$, and $a \wedge b \wedge c \in R$, but $a \wedge (b \vee \neg c) \notin R$, what also implies $b \vee \neg c \notin R$. Indeed, because $x \wedge (y \vee \neg z)$ is the basis of the clone S_{12} corresponding to negative Horn relations iS_{12} and Horn relations are closed under conjunction, we must have $a \wedge (b \vee \neg c) \notin R$. Since $a \in R$ holds and R is Horn (i.e., it is closed under conjunction), we must have $(b \vee \neg c) \notin R$, otherwise we would have $a \wedge (b \vee \neg c) \in R$. Construct the variable sets $V_{ijk} = \{x \mid a(x) = i, b(x) = j, c(x) = k\}$. Construct the constraint

$$Q(x, y) \quad = \quad R[0/V_{000}, 0/V_{001}, 0/V_{010}, 0/V_{011}, x/V_{100}, y/V_{101}, x/V_{110}, 1/V_{111}].$$

It can be easily verified that the constraint $Q(x, y)$ generates the relation $\{00, 01, 11\} = [\neg x \vee y]$. □

Proposition 8.11 CMININF(S) *for all sets S which are Horn, but neither dual Horn, nor negative Horn, are* coNP*-complete.*

Proof: Since S is Horn, but not dual Horn, there exists a relation $R_1 \in \langle S \rangle_{\nexists}$ with the same properties according to Lemma 7.1. From R_1 we can produce the relation $[\neg x \vee \neg y]$ according Lemma 8.9. Since S is Horn but not negative Horn, there exists a relation $R_2 \in \langle S \rangle_{\nexists}$ with this property. From R_2 we can produce the relation $[\neg x \vee y]$ according to Lemma 8.10. Hence, we can produce both relations $[\neg x \vee \neg y]$ and $[\neg x \vee y]$ from S. Cadoli and Lenzerini proved in [6] (case HK in their notation) that CMININF(S) for $S = \{[\neg x \vee \neg y], [\neg x \vee y]\}$ is coNP-complete. □

8.4 Affine Relations

We need another representation lemma proved by the usual Schaefer technique.

Lemma 8.12 *If S neither Horn nor dual Horn then we can construct the relation $[x \not\equiv y]$ from S by conjunction and variable identification.*

Proof: Since S is neither Horn nor dual Horn, there exists a relation $R \in \langle S \rangle_{\nexists}$ with the same properties according to Lemma 7.1. Moreover according to Lemma 7.2, R contains two tuples $a, b \in R$ such that $a \wedge b \notin R$ and $a \vee b \notin R$. Construct the variable sets $V_{ij} = \{x \mid a(x) = i, b(x) = j\}$. Construct the constraint $Q(x, y) = R[0/V_{00}, x/V_{01}, y/V_{10}, 1/V_{11}]$. It is clear that the constraint $Q(x, y)$ generates the relation $\{01, 10\} = [x \not\equiv y]$. □

We need to investigate a special property of incomparable relations.

Lemma 8.13 *Given an irredundant affine relation R which is not incomparable, then any two comparable tuples $a < b$ in R must differ in at least two positions. Moreover, there must be a third tuple c which is not constant on the positions where a and b differ.*

Proof: We claim that the existence two comparable tuples $a < b$ differing in just one position implies the relation R to be redundant. Denote by i the coordinate on which a and b differ. For any tuple c in R we construct the tuple $c' = a + b + c$, which is identical to c except that $c'[i] = \neg c[i]$. Since R is affine, the tuple c' must be in the relation. Hence, R is redundant since it is of the form $Q \times \{0, 1\}$. Thus in any irredundant relation any two comparable tuples must differ in at least two positions.

Now let a and b differ on at least two positions, say i and j. If all tuples $c \in R$ are constant on the positions i and j where a and b differ, i.e., $c[i] = c[j]$ for all $c \in R$, then R is again redundant because in particular the columns i and j are identical. □

Proposition 8.14 CMININF(S) *is* coNP*-hard for each set of relations* S*, which is affine but neither incomparable nor other Schaefer.*

Proof: By taking Cartesian products of relations in S, there exists a single relation R in $\langle S \rangle_{\nexists}$ which is affine, but neither incomparable, nor bijunctive, nor Horn, nor dual Horn, according to Lemma 7.1. Since R is not incomparable, there exist two tuples a and b satisfying the condition $a < b$. Without loss of generality we can assume that a and b are closest possible, i.e., there is no tuple t for which $a < t < b$ holds. By Lemma 8.13 we know that a and b differ in at least two positions, and that there is a third tuple c not constant on the coordinates on which a and b differ.

Since R is neither Horn nor dual Horn, we can assume by taking a Cartesian product of R with itself, that c satisfies the conditions $a \wedge c \notin R$, $b \wedge c \notin R$, $a \vee c \notin R$, and $b \vee c \notin R$. Form the Schaefer-style variable identification on R based on the tuples a, b, c. Note that since $a < b$ holds, the variables V_{100} and V_{101} will not appear. Thus we get the relation $R[V_{000}, V_{001}, V_{010}, V_{011}, V_{110}, V_{111}]$. It can be checked that the variables $V_{001}, V_{010}, V_{011}, V_{110}$ must all appear. Observe that if V_{001} does not appear, then $b \vee c = b \in R$ which is a contradiction. Similarly, if V_{110} does not appear, then $a \wedge c = a \in R$ which is a contradiction. Moreover, V_{010} and V_{011} must appear since c is not constant on the coordinates where a and b differ.

Since R is affine, it also contains the tuple $d = a + b + c = 011001$. Moreover, R does not contain the tuple $t = \mathrm{maj}(a, b, c) = 000111$ since this tuple satisfies the condition $a < t < b$. We add the constraint $(V_{110} \not\equiv V_{001}) \wedge (V_{000} = 0) \wedge (V_{111} = 1)$ to $R[V_{000}, V_{001}, V_{010}, V_{011}, V_{110}, V_{111}]$. This addition is allowed following Theorem 4.3 and Lemma 8.12. The resulting constraint contains the tuples 000011, 001111, 010101, 011001, but it does not contain the tuple 000111. There are only three undetermined variables so there can be at most 8 tuples satisfying the affine constraint. Moreover, since the constraint is affine the number of tuples satisfying the constraint is a power of 2. Therefore, since 000111 is not satisfying the constraint, we have that

$$\begin{aligned} & P(V_{000}, V_{001}, V_{010}, V_{011}, V_{110}, V_{111}) \\ &= R[V_{001}, V_{010}, V_{011}, V_{110}] \wedge (V_{110} \not\equiv V_{001}) \wedge (V_{000} = 0) \wedge (V_{111} = 1) \\ &= (V_{001} + V_{010} + V_{011} = 0) \wedge (V_{110} + V_{001} = 1) \wedge (V_{000} = 0) \wedge (V_{111} = 1). \end{aligned}$$

The problem CMININF(P) for the relation P, generated by $P(V_{000}, V_{001}, V_{010}, V_{011}, V_{110}, V_{111})$, is coNP-hard by Lemma 6.12. □

8.5 Bijunctive Relations

Lemma 8.15 MINEXT(R) *for* $R = [(x \vee y) \wedge (x \not\equiv z)]$ *is* NP*-complete.*

Proof: The membership in NP is clear, we focus on the NP-hardness proof by means of a reduction from 3SAT. Suppose $\varphi(x_1, \ldots, x_n) = c_1 \wedge \cdots \wedge c_k$ is a 3SAT formula. Associate the variable y_i with the clause c_i for $i = 1, \ldots, k$. For each clause $c_i = l_i^1 \vee l_i^2 \vee l_i^3$ we add the following three formulas

$$(y_i \vee v_i^1) \wedge (v_i^1 \not\equiv \bar{v}_i^1), \qquad (y_i \vee v_i^2) \wedge (v_i^2 \not\equiv \bar{v}_i^2), \qquad (y_i \vee v_i^3) \wedge (v_i^3 \not\equiv \bar{v}_i^3),$$

to φ', where

$$v_i^j = \begin{cases} x & \text{if } l_i^j = \neg x, \\ x' & \text{if } l_i^j = x. \end{cases} \qquad \bar{v}_i^j = \begin{cases} x' & \text{if } l_i^j = \neg x, \\ x & \text{if } l_i^j = x. \end{cases}$$

Claim: The partial assignment $s(y_i) = 1$ for each $i = 1, \ldots, k$ can be extended to a minimal assignment of φ' if and only if φ is satisfiable.

Let φ be satisfiable. Every clause c_i evaluates to 1. For each clause c_i there exists a j, such that $l_i^j = 1$. Then $v_i^j = 0$ which implies $y_i = 1$. Moreover, since $v_i^j \not\equiv \bar{v}_i^j$ holds, every satisfying assignment to φ' is incomparable, hence minimal.

Let φ be unsatisfiable. Then there exists a falsified clause c_i, i.e., $l_i^1 = l_i^2 = l_i^3 = 0$, which implies $v_i^1 = v_i^2 = v_i^3 = 1$. Therefore there exists a satisfying assignment m of φ' with $m(y_i) = 0$. hence $s(y_i) = 1$ for $i = 1, \ldots, k$ cannot be extended to a minimal solution. □

We need the following result from Jeavons et al. [17], based on a previous algebraic result of Baker and Pixley [1], stating that it is sufficient to consider binary relations when we consider bijunctive constraint languages.

Proposition 8.16 (Jeavons et al. [17]) *Given a n-ary bijunctive constraint $R(x_1, \ldots, x_n)$ then it is equivalent to $\bigwedge_{1 \le i \le j} R_{ij}(x_i, x_j)$ where R_{ij} is the projection of the relation R to the coordinates i and j.*

Proposition 8.18 in the sequel is introduced already in this section concerning the CMININF problem and not in Section 9. The reason is that the complexity result for MININF immediately carries over to CMININF, but we do not want to present the same proof twice. We need first a supporting lemma, which explains how to force variables to take constant values by means of pinning.

Lemma 8.17 *If S is bijunctive but not other Schaefer, then we can construct by conjunction and variable identification the binary relation $R_{01} = \{01\}$.*

Proof: The proof is almost the same as of Lemma 5.24 in [8]. Observe that S is neither 0-valid, nor 1-valid, nor complementive since this (according to Post's lattice) would imply that S is Horn, dual Horn, and affine, respectively.

Consider a non-0-valid relation $R_{n0} \in S$. If R_{n0} is 1-valid, then $R_{n0}(x, \ldots, x) = T(x)$. Similarly if a non-1-valid relation R_{n1} is 0-valid then we can implement $F(x)$ from R_{n1} by variable identification. Hence, we can implement $R_{01}(x, y)$ by $F(x) \wedge T(y)$.

If R_{n0} is neither 0-valid nor complementive, nor 1-valid, then $0 \cdots 0 \notin R_{n0}$, $1 \cdots 1 \notin R_{n0}$, and there exists a tuple $s \in R_{n0}$ such that $\neg s \notin R_{n0}$, where $\neg s$ denotes the complement of s. Define $I_0(s) = \{i \mid s[i] = 0\}$ and $I_1(s) = \{i \mid s[i] = 1\}$. Form the constraint $R_{n0}(x_i, \ldots, x_j)$, where $x_i, \ldots, x_j \in \{x_0, x_1\}$, by placing x_0 at the positions in $I_0(s)$ and x_1 at the positions in $I_1(s)$. Note that $I_0(s)$ and $I_1(s)$ are both non-empty since R_{n0} is neither 0-valid nor 1-valid. Hence, we have $R_{n0}(x_i, \ldots, x_j) = R_{01}(x_0, x_1)$.

Finally, if $R_{n0} \in S$ is complementive but neither 0-valid nor 1-valid, then there must be a relation $R_{nc} \in S$ which is not complementive, since a bijunctive and complementive relations is also affine, therefore also 0- and 1-valid. Hence, there exists a tuple $s \in R_{nc}$ such that $\neg s \notin R_{nc}$. Form the constraint $R_{n10}(x_0, x_1) = R_{nc}(x_i, \ldots, x_j)$ (where $x_i, \ldots, x_j \in \{x_0, x_1\}$) by placing x_0 at the positions in $I_0(s)$ and x_1 at the positions in $I_1(s)$. We have that $01 \in R_{n10}$ and $10 \notin R_{n10}$. Consider again R_{n0} and a tuple $s \in R_{n0}$, then $(x_0 \not\equiv x_1) = R_{n0}(x_i, \ldots, x_j)$ where $x_i, \ldots, x_j \in \{x_0, x_1\}$, x_0 occurs at the positions in $I_0(s)$, and x_1 at the positions in $I_1(s)$. We can again implement the constraint $R_{01}(x_0, x_1) = (x_0 \not\equiv x_1) \wedge R_{n10}(x_0, x_1)$. □

Proposition 8.18 *Let $N = \{[\neg x \vee \neg y], [x \not\equiv y]\}$. For each set of relations S, which is bijunctive, but neither other Schaefer, nor a subset of $\langle N \rangle_{\nexists}$,* MININF$(S)$ *is* coNP*-complete.*

Proof: Let R be the Cartesian product of all relations in S. Obviously, R is bijunctive but neither Horn, nor dual Horn, nor affine, nor a subset of $\langle N \rangle_{\nexists}$. Let φ be the conjunction of binary constraints generating the relation R, produced according to Proposition 8.16.

There must be a clause $(\ell_p \vee \ell) \in \varphi$ with at least one positive literal, say ℓ_p, otherwise we would have $[\varphi] \in \langle N \rangle_{\nexists}$. If ℓ is a negative literal, then ℓ_p must not occur in a $\not\equiv$ constraint, otherwise we would again have $[\varphi] \in \langle N \rangle_{\nexists}$. If both ℓ_p and ℓ are positive, then ℓ_p or ℓ must not occur in a $\not\equiv$ constraint, otherwise we would again have $[\varphi] \in \langle N \rangle_{\nexists}$. In all cases, this is because we have the identity $(\ell_p \vee \ell) \wedge (\ell_p \not\equiv x) = (\neg x \vee \ell) \wedge (\ell_p \not\equiv x)$. Moreover, the literals ℓ_p and ℓ cannot be assigned constant values. Indeed, if $\varphi = \varphi' \wedge (\ell_p \vee \ell) \models \neg \ell_p$ holds for a formula φ' with $[\varphi'] \in \langle N \rangle_{\nexists}$, then φ is equivalent to $\varphi' \wedge (\neg \ell_p \vee \neg \ell_p) \wedge (\ell_p \not\equiv \ell)$, implying $[\varphi] \in \langle N \rangle_{\nexists}$. If $\varphi = \varphi' \wedge (\ell_p \vee \ell) \models \ell_p$ holds then φ is equivalent to φ', implying again $[\varphi] \in \langle N \rangle_{\nexists}$.

Hence, we are in the situation $(\ell_p \vee \ell) \in \varphi$ and ℓ_p is positive and it does not occur in any $\not\equiv$ constraint. Two cases emerge depending on whether ℓ appears in a $\not\equiv$ constraint. For the rest of the proof we focus on the case where ℓ appears in a $\not\equiv$ constraint. The other case can be handled in the same way with only minor and obvious modifications. Assume that $\ell_p = x$, $\mathrm{Var}(\ell) = y$, and y occurs in a $\not\equiv$ constraint together with z, i.e., that $(x \vee y) \wedge (y \not\equiv z) \in \varphi$, or $(x \vee \neg y) \wedge (y \not\equiv z) \in \varphi$. Without loss of generality we assume the former case since $(x \vee \neg y) \wedge (y \not\equiv z) = (x \vee z) \wedge (y \not\equiv z)$.

Our goal is to produce the constraint $(x \vee y) \wedge (y \not\equiv z) \wedge R_{01}(v, w)$ from φ by conjunction and variable identification. It is clear that both aforementioned constraints belong to $\langle S \rangle_{\nexists}$, since we can produce from S the constraint $R_{01}(x_0, x_1)$ following Lemma 8.17.

Simplify φ by identifying all variables x_1, x_2 such that $(x_1 = x_2) \in \varphi$, then remove all equality constraints from φ. Form the sets of variables $V_0 = \{v \mid \varphi \models \neg v\}$ and $V_1 = \{w \mid \varphi \models w\}$. Remove from φ all clauses containing variables in $V_0 \cup V_1$, then add the constraint $R_{01}(V_0, V_1)$ to φ. Obviously, this transformation of φ does not change the fact that the projection of φ onto $\{x, y\}$ is equivalent to $(x \vee y)$, whereas the projection of φ onto $\{y, z\}$ is equivalent to $(y \not\equiv z)$. Note that there can be no $(x_1 = x_2) \in \varphi$ where $\{x_1, x_2\} \subseteq \{x, y, z\}$. In the same manner, we must have $\{x, y, z\} \cap V_0 = \emptyset$ and $\{x, y, z\} \cap V_1 = \emptyset$.

Continue simplifying the formula φ by eliminating all $(x_1 \not\equiv x_2)$ constraints, except the one $(y \not\equiv z)$, by replacing x_1 by $\neg x_2$, as well as $\neg x_1$ by x_2, throughout the formula. Denote the resulting formula by φ'. Since x does not occur in any $\not\equiv$ constraint and neither y nor z occur in any other $\not\equiv$ constraint except in $(y \not\equiv z)$, otherwise there would have to be some $(x_1 = x_2)$ constraint still left, we have $(x \vee y) \wedge (y \not\equiv z) \in \varphi'$.

Recall that resolution between two clauses $(c \vee v)$ and $(\neg v \vee c')$ produces the new clause $(c \vee c')$ and discards the two previous clauses. Note that resolution on binary clauses produces a binary clause. For any variable $x_i \notin \{x, y, z\}$ occurring both positively and negatively in φ', apply resolution to get a formula, where every variable not in $\{x, y, z\}$ only occurs either positively or negatively, but not both. Denote the resulting formula by φ''. Form the variable sets $V_0'' = \{v \mid v \text{ only occurs negatively in } \varphi''\}$ and $V_1'' = \{w \mid w \text{ only occurs positively in } \varphi''\}$. Discard from φ'' the clauses containing variables from $V_0'' \cup V_1''$ and add the constraint $R_{01}(V_0'', V_1'')$ to the formula.

Hence using only conjunction and variable identification, we can implement the constraint $(x \vee y) \wedge (y \not\equiv z) \wedge R_{01}(v, w)$ from S. Using Lemma 8.15 we obtain the desired coNP-completeness result. □

8.6 Main Result

We put together all results in this section and obtain the following trichotomy result for minimal inference without free variables but with constants.

Theorem 8.19 (Trichotomy of CMININF) *Let S be a finite nonempty set of Boolean relations and S^* the corresponding 1-valid restriction. If every relation in S is both Horn and dual Horn, or negative Horn, or Schaefer and incomparable, or a subset of the weak co-clone $\langle N \rangle_{\nexists}$, where $N = \{[\neg x \vee \neg y], [x \not\equiv y]\}$, then* CMININF$(S)$ *is decidable in polynomial time. Else if S^* is Schaefer, then* CMININF(S) *is* coNP*-complete. Otherwise,* CMININF(S) *is* Π_2P*-complete.*

Proof: The parts concerning Π_2P-completeness and membership in coNP follow from the dichotomy theorem due to Kirousis and Kolaitis [21]. As for tractability, the fact that CMININF(S) is tractable when every relation in S is both Horn and dual Horn, or negative Horn, follows directly from the corresponding tractability results for GMININF proved in [6], and in Proposition 5.4, respectively. Tractability of MININF(S) for S being Schaefer and incomparable, or a subset of $\langle N \rangle_{\nexists}$, is proved in Propositions 8.4, and 8.6, respectively. The coNP-hardness for CMININF(S) when S is not Schaefer, and S^* is Schaefer, is proved in Proposition 8.2.

Hence, what remains to be done is to prove coNP-hardness for all sets of relations S that do not fall into one of the tractable classes when S is dual Horn, Horn, affine, or bijunctive. This is done in Sections 8.2, 8.3, 8.4, and 8.5, respectively, by Propositions 8.8, 8.11, 8.14, and 8.18. □

9 Minimal Inference Without Free Variables Nor Constants (MININF)

We want to derive in this section a complete characterization of complexity for the MININF problem, which is the most challenging among the considered minimal inference problems. It requires *all* variables of φ in a MININF instance $\varphi \models_{\min} c$ to be minimized. The difficulty of the problem originates in the impossibility to use the Galois correspondence or the possibility to substitute logical constants for variables in the relations. In fact, MININF is the most difficult problem to analyse among all four cases of minimal inference. MININF was listed as one of the important open problems (Question 4.1) during the International Workshop on Mathematics of Constraint Satisfaction held in Oxford in March 2006 [27].

Many results for MININF were already presented in previous sections, since they represent necessary stepping stones to prove other complexity results for the three previously considered variants of minimal inference. Recall also that MININF is a special case of GMININF, VMININF, or CMININF with the supplementary condition $Q = \emptyset$ and $Z = \emptyset$. Hence a tractability result for the more general case of minimal inference, namely Proposition 5.3, Proposition 5.4, Proposition 5.5, Proposition 6.15, Proposition 6.16, and Proposition 8.4, propagates.

We use the same two phase approach as in the previous section, considering first non-Schaefer relations S, where we proceed with a sharpening of a result from Kirousis and Kolaitis [21].

Proposition 9.1 *Let S be a non-Schaefer and non-0-valid set of Boolean relations, with S^* being the corresponding 1-valid restriction. If S^* is Schaefer, then* MININF(S) *is* coNP*-complete, otherwise it is* Π_2P*-complete.*

Proof: The Π_2P-completeness part and the fact that MININF(S) is in coNP when S^* is Schaefer is proved in [21, Theorem 3.10]. Hence, what remains to be proved is that MININF(S) is coNP-hard when S^* is Schaefer. First note that the case where S is non-Schaefer and non-0-valid but 1-valid is known to be Π_2P-complete by Theorem 3.10 in [21], hence we can assume that S is non-1-valid. We give a reduction from the coNP-complete unsatisfiability problem UNSAT(S) to MININF(S). Let φ be a conjunction of constraints over S. Let $R(x_1, x_2, \ldots, x_n)$ be a constraint build over a relation R in S where the variables $x_1, \ldots, x_n$ do not occur in φ. Let m be a minimal model of $R(x_1, x_2, \ldots, x_n)$. We construct the clause to be inferred ψ as follows: If $m(x_i) = 1$ ($m(x_i) = 0$) then add $\neg x_i$ (x_i) to ψ for $1 \leq i \leq n$. Then φ is unsatisfiable if and only if ψ can be minimally inferred from $\varphi \wedge R(x_1, x_2, \ldots, x_n)$. □

Consider now S to be Schaefer or 0-valid. Tractability of MININF(S) for S being 0-valid is trivial and the polynomial-time decidability of MININF(S) for S being Horn follows from the fact that a satisfiable Horn formula has a unique minimal model computable in polynomial time. In the rest of the paper we investigate the other cases for which MININF(S) is in coNP, namely when S is dual Horn, bijunctive, or affine.

9.1 Affine Relations

We need first a representation lemma proved by Schaefer's technique.

Lemma 9.2 *Let R be a Boolean relation which is affine and 1-valid, but neither 0-valid, nor other Schaefer. Then we can construct the relation $[(x + y + z = 1) \wedge (w = 1)]$ from R by conjunction and variable identification.*

Proof: Since R is not Horn, there exist two vectors $a, b \in R$, such that $a \wedge b \notin R$. Construct the variable sets $V_{ij} = \{x \in V \mid a(x) = i, b(x) = j\}$. Identify the variables in each set V_{ij} and construct the relation R' generated by the constraint $R[V_{00}, V_{01}, V_{10}, V_{11}]$. It can be easily seen that the vectors 0011, 0101, and 1111 belong to R', whereas $0000 \notin R'$ (not 0-valid) and $0001 \notin R'$ (not Horn). Since R' is affine, we have $0011 + 0101 + 1111 = 1001 \in R'$.

Let $m, m' \in R'$ and $m'' \notin R'$. Then also $m+m'+m'' \notin R'$, since otherwise from the membership of m, m', and $m+m'+m''$ in R' follows that $(m+m'+m'')+m+m' = m'' \in R'$, because R' is affine. Using this result, from $0011, 0101 \in R'$ and $0001 \notin R'$ follows $0110 \notin R'$. From $0101, 1111 \in R'$ and $0001 \notin R'$ follows $1011 \notin R'$. Finally from $0011, 1111 \in R'$ and $0001 \notin R'$ follows $1101 \notin R'$. We can force the variable V_{11} to take the value 1 by the constraint $R'(V_{11}, V_{11}, V_{11}, V_{11})$, since R as well as R' are both 1-valid but not 0-valid. Hence, the constraint $R'(x, y, z, w) \wedge R'(w, w, w, w)$ generates the relation $[(x + y + z = 1) \wedge (w = 1)] = \{0011, 0101, 1001, 1111\}$. □

Proposition 9.3 MININF(S) *is* coNP*-complete for each set of relations S which is affine and* 1*-valid, but neither* 0*-valid, nor other Schaefer.*

Proof: We can construct a relation R which is affine and 1-valid, but not 0-valid, not Horn, not dual Horn, nor bijunctive, by Cartesian product of relations from S. The rest follows from Lemma 9.2, Lemma 6.13, and from the relationship between MINEXT and MININF. □

We need a supporting lemma explaining in the context of affine constraint languages how to force variables to take constant values.

Lemma 9.4 *If S is neither complementive nor 0-valid nor 1-valid, then the relation $R_{01} = \{01\}$ can be constructed from S by conjunction and variable identification.*

Proof: We can assume the existence of a relation R which is neither complementive nor 0-valid nor 1-valid by taking Cartesian products of relations in S according to Lemma 7.1. Take a tuple $a \in R$ such that $\neg a \notin R$ (since R is not complementive) and perform Schaefer's construction on a, i.e., construct the relation Q generated by the constraint $R[V_0, V_1]$. Obviously $01 \in Q$ holds, and neither 00 nor 11 is in Q (because R is neither 0-valid nor 1-valid). Moreover $10 \notin Q$ holds since $\neg a \notin R$. Thus we have constructed the desired relation $Q = \{01\}$. □

Proposition 9.5 MININF(S) *is* coNP*-hard for each set of relations S, which is affine, but neither incomparable, nor other Schaefer, nor 0-valid, nor 1-valid, nor complementive.*

Proof: By Lemma 9.4 we know that we have access to the constraint $R_{01}(x, y)$, based on the relation $R_{01} = \{01\}$, forcing the variables x and y to take constant values. Since S is neither Horn nor dual Horn, by taking Cartesian products we can assume that there is a relation $N \in \langle S \rangle$ and two tuples $a, b \in N$, such that $a \wedge b \notin N$ and $a \vee b \notin N$. Use Schaefer's variable identification on a and b to get the relation $N[V_{00}, V_{01}, V_{10}, V_{11}]$. Now we need to impose the constant 0 on V_{00} and 1 on V_{11}, using the relation R_{01}. We construct the constraint $N(z, x, y, w) \wedge R_{01}(z, w) = (x \not\equiv y) \wedge (z = 0) \wedge (w = 1)$ by conjunction and variable identification. The rest of the proof is totally identical to that of Proposition 8.14. □

We need a representation lemma for complementive relations and a complexity result for a minimal inference problem upon a particular relation R which is both affine and complementive.

Lemma 9.6 *If S is complementive and neither 0-valid nor 1-valid, then we can construct $[x \not\equiv y]$ from S by conjunction and variable identification.*

Proof: We can assume the existence of a relation R by taking Cartesian products according to Lemma 7.1, such that R is neither 0-valid nor 1-valid. Take two tuples $a, b \in R$ such that $\neg a = b$. Perform Schaefer's construction $R[V_{00}, V_{01}, V_{10}, V_{11}]$ on a and b. Obviously the variables V_{00} and V_{11} do not appear since a is the complement of b. Hence we have the binary relation $R[V_{01}, V_{10}]$ containing the tuples 01 and 10, but not the tuples 11 and 00 (since otherwise R would be 1-valid and 0-valid, respectively). Hence the relation generated by the constraint $R[V_{01}, V_{10}]$ is $\{01, 10\} = [x \not\equiv y]$. □

Lemma 9.7 MINEXT(R) *is* NP*-complete for* $R = [(x + y + z + w = 0) \wedge (x + u = 1) \wedge (y + v = 1)]$.

Proof: NP-membership is clear, we focus on the NP-hardness proof by a reduction from NAE3SAT (not-all-equal 3SAT). Suppose $\varphi = c_1 \wedge \cdots \wedge c_k$ is a not-all-equal 3SAT formula with variables $x_1, \ldots, x_n$ and clauses $c_i = l_i^1 \vee l_i^2 \vee l_i^3$. We construct a formula φ' as the conjunction of the following equations for each clause $c_i = l_i^1 \vee l_i^2 \vee l_i^3$:

$$\begin{aligned}
(z_{3i-2} + v_i^1 + v_i^2 + y_i = 0) \ &\wedge\ (v_i^1 + u_i^1 = 1) \ \wedge\ (v_i^2 + u_i^2 = 1),\\
(z_{3i-1} + v_i^2 + v_i^3 + y_i = 0) \ &\wedge\ (v_i^2 + u_i^2 = 1) \ \wedge\ (v_i^3 + u_i^3 = 1),\\
(z_{3i} + v_i^3 + v_i^1 + y_i = 0) \ &\wedge\ (v_i^3 + u_i^3 = 1) \ \wedge\ (v_i^1 + u_i^1 = 1).
\end{aligned}$$

where

$$v_i^j = \begin{cases} x_p & \text{if } l_i^j = x_p, \\ x_p' & \text{if } l_i^j = \neg x_p. \end{cases} \qquad \text{and} \qquad u_i^j = \begin{cases} x_p' & \text{if } v_i^j = x_p, \\ x_p & \text{if } v_i^j = x_p'. \end{cases}$$

The variable v_i^j is a placeholder for the literal l_i^j, whereas u_i^j is the negation of v_i^j.

Claim: The formula φ is satisfiable if and only if the assignment $s(y_i) = 1$ for all $i = 1, \ldots, k$ of φ' has a minimal extension.

Let φ be not-all-equal satisfiable. Then there exists a not-all-equal satisfying assignment m of the formula φ. Since every clause c_i is not-all-equal satisfied, for each i there must be $a, b \in \{1, 2, 3\}$ such that $m(l_i^a) \neq m(l_i^b)$. Let $\bar{m}$ be an extension of m that satisfies φ'. Following the construction of φ', we must have $\bar{m}(v_i^a) + \bar{m}(v_i^b) = 1$. Then there are two possibilities to get a minimal assignment $\bar{m}$:

1. When we set $\bar{m}(z_{3i-p}) = 1$ and $\bar{m}(y_i) = 0$, we get an uninteresting assignment.
2. When we set $\bar{m}(z_{3i-p}) = 0$ and $\bar{m}(y_i) = 1$, we get an assignment which is an extension of the assignment $s(y_i) = 1$ for all $i = 1, \ldots, k$.

These two possible assignments are clearly incomparable and no value can be changed from 1 to 0 to get another satisfying assignment of φ'. Therefore the assignment $\bar{m}$ from the second case is minimal.

Let φ be not-all-equal unsatisfiable. Then for each assignment m there must always be a clause c_i which literals are assigned the same values, i.e., $m(l_i^1) = m(l_i^2) = m(l_i^3) = 0$ or $m(l_i^1) = m(l_i^2) = m(l_i^3) = 1$. Let $\bar{m}$ be an extension of m that satisfies the formula φ'. Following the construction of φ', we have $\bar{m}(v_i^1) + \bar{m}(v_i^2) = \bar{m}(v_i^2) + \bar{m}(v_i^3) = \bar{m}(v_i^3) + \bar{m}(v_i^1) = 0$. Then we can set $\bar{m}(z_{3i-2}) = \bar{m}(z_{3i-1}) = \bar{m}(z_{3i}) = \bar{m}(y_i) = 0$ to produce a minimal assignment. This implies that the assignment $s(y_i) = 1$ for all i cannot be extended to a minimal one. □

Proposition 9.8 MININF(S) *is* coNP*-hard for each set of relations S, which is affine and complementive, but neither incomparable, nor other Schaefer, nor* 0*-valid, nor* 1*-valid.*

Proof: By taking Cartesian products of relations in S there exists a relation R in $\langle S \rangle_{\nexists}$, according to Lemma 7.1, which is affine, but neither incomparable, nor bijunctive, nor Horn, nor dual-Horn. Since R is not incomparable there are two tuples a and b satisfying the condition $a < b$. Without loss of generality we can assume that a and b are closest, i.e., there is no tuple t satisfying $a < t < b$. By Lemma 8.13 we know that a and b differ in at least two positions and that there exists a third tuple c which is not constant on the coordinates where a and b differ.

Since R is neither Horn, nor dual-Horn we can assume, by taking a Cartesian products of R with itself, that c satisfies the conditions $a \wedge c \notin R$, $b \wedge c \notin R$, $a \vee c \notin R$, and $b \vee c \notin R$. Form the Schaefer style implementation on R based on the tuples a, b, c. Note that since $a < b$ holds, the variables V_{100} and V_{101} will not appear. Hence we get the relation $R[V_{000}, V_{001}, V_{010}, V_{011}, V_{110}, V_{111}]$. It can be checked that the variables $V_{001}, V_{010}, V_{011}, V_{110}$ must all appear. If V_{001} does not appear, then $b \vee c = b \in R$ which is a contradiction. Similarly, if V_{110} does not appear, then $a \wedge c = a \in R$ which is a contradiction. Moreover, V_{010} and V_{011} must appear since c is not constant on the coordinates where a and b differ.

Since R is affine, it must also contain the tuple $d = a + b + c = 011001$. Moreover, R does not contain the tuple $t = \mathrm{maj}(a, b, c) = 000111$ since this tuple satisfies $a < t < b$. Furthermore, since R is complementive it also contains the tuples 100110, 111100 101010, 110000. Now, adding the constraints $V_{000} \not\equiv V_{111}$ and $V_{110} \not\equiv V_{001}$, to which we have access according to Lemma 9.6, we get the relation $R[V_{000}, V_{001}, V_{010}, V_{011}, V_{110}, V_{111}] \wedge (V_{000} \not\equiv V_{111}) \wedge (V_{110} \not\equiv V_{001})$. Since the value of V_{110} is determined by V_{001} and vice versa, and the same for V_{000} and V_{111}, there can be at most 16 tuples satisfying the affine constraint. We already have 8 tuples that satisfy the constraint and we know that the tuple 000111 does not satisfy the constraint. Since the constraint is affine, the number of tuples satisfying the constraint must be a power of 2. The constraint $R[V_{000}, V_{001}, V_{010}, V_{011}, V_{110}, V_{111}] \wedge (V_{000} \not\equiv V_{111}) \wedge (V_{110} \not\equiv V_{001})$ generates the relation consisting of the tuples {000011, 001111, 010101, 011001, 100110, 111100, 101010, 110000}. It is equivalent to the constraint $(V_{000} + V_{001} + V_{010} + V_{011} = 0) \wedge (V_{000} + V_{111} = 1) \wedge (V_{001} + V_{110} = 1)$, for which the minimal inference problem is coNP-complete according to Lemma 9.7. □

9.2 Dual Horn Relations

Proposition 9.9 MININF(S) *is* coNP*-hard for each set of relations S, which is dual Horn, but neither Horn, nor* 0*-valid, nor* 1*-valid.*

Proof: Construct a relation R from S which is neither 0-valid nor 1-valid. Such a relation exists by taking Cartesian products. Take a tuple $a \in R$ and construct the relation $R[V_0, V_1]$ on a. Both variables V_0 and V_1 must appears since R is neither 0-valid nor 1-valid. Hence $01 \in R[V_0, V_1]$. Neither 00 nor 11 is in $R[V_0, V_1]$ since R is neither 0-valid nor 1-valid. Moreover, 10 is not in R, since if it were then the fact that R is dual-Horn together with $01 \in R$ implies $11 \in R$ which we have already ruled out. Thus $R[V_0, V_1]$ generates the relation {01}.

Now construct from S a relation Q which is dual Horn, but not Horn. There must be two tuples $a, b \in Q$, such that $a \wedge b \notin Q$ and $a \vee b \in Q$. Construct the relation $Q[V_{00}, V_{01}, V_{10}, V_{11}]$ on a and b. Form the conjunction $Q[V_{00}, V_{01}, V_{10}, V_{11}] \wedge R[V_{00}, V_{11}]$ forcing the variables V_{00} and V_{11} to take the values 0 and 1, respectively. The constraint $Q'(V_{00}, V_{01}, V_{10}, V_{11}) = Q[V_{00}, V_{01}, V_{10}, V_{11}] \wedge R[V_{00}, V_{11}]$ is equal to $(V_{01} \vee V_{10}) \wedge (V_{00} = 0) \wedge (V_{11} = 1)$. Cadoli and Lenzerini [6, Theorem 5] proved that MININF($[x \vee y]$) is coNP-complete, therefore also MININF(Q') where $Q' = [(x \vee y) \wedge (z = 0) \wedge (w = 1)]$ is coNP-complete. □

Lemma 9.10 MININF(R) *is* coNP*-complete for* $R = [(x \vee y) \wedge (\neg z \vee x) \wedge (\neg z \vee y)]$ *and* $R = [(x \vee y) \wedge (\neg z \vee x)]$.

Proof: The coNP membership is clear, we focus on the hardness proof done by a simple reduction from MININF($[x \vee y]$). Every constraint $(x \vee y)$ in the MININF($[x \vee y]$) instance is replaced by a constraint $(x \vee y) \wedge (\neg z_i \vee x) \wedge (\neg z_i \vee y)$ or $(x \vee y) \wedge (\neg z_i \vee x)$, respectively, where the z_i's are fresh variables (one for each new constraint). Note that all the z_i's must be 0 in all minimal models of the new instance. Hence, there is an obvious one-to-one correspondence between the minimal models of the old and new MININF instances. □

Proposition 9.11 MININF(S) *is* coNP*-hard for each set of relations S, which is dual-Horn and* 1*-valid, but neither* 0*-valid nor Horn.*

Proof: We know that there is a relation R in $\langle S \rangle_{\nexists}$ which is dual-Horn and 1-valid but not 0-valid, constructed by taking Cartesian products. Now, do Schaefer's construction on two tuples $a, b \in R$ such that $a \wedge b \notin R$ and $a \vee b \in R$. Such tuples exist, since R is dual Horn, but not Horn.

This construction gives us the relation Q generated by the constraint $R[V_{00}, V_{01}, V_{10}, V_{11}]$. Since Q is 1-valid but not 0-valid, we have access to the constraint $(x = 1)$. Construct the conjunction $Q'(V_{00}, V_{01}, V_{10}, V_{11}) = Q(V_{00}, V_{01}, V_{10}, V_{11}) \wedge Q(V_{11}, V_{11}, V_{11}, V_{11})$. Now, V_{01} and V_{10} both exists, otherwise we would have $a \wedge b \in R$. Moreover, $0001 \notin Q'$ since $a \wedge b = 0001$. If V_{00} does not exist then $Q' = [(x \vee y) \wedge (z = 1)]$. Cadoli and Lenzerini proved in [6] that MININF($[x \vee y]$) is coNP-complete, hence also MININF(Q') is coNP-complete.

Assume now that V_{00} exists. We know that Q' contains the tuples 1111, 0011, 0101, and 0111. We now have to consider 5 cases according to the presence of the tuples 1001, 1011, and 1101:

Case 1: $Q' = \{1111, 0011, 0101, 0111\}$ which is equal to $[(y \vee z) \wedge (\neg x \vee y) \wedge (\neg x \vee z) \wedge (w = 1)]$ for which MININF is coNP-complete according to Lemma 9.10.

Case 2: Q' contains the tuple 1001. Because Q' is dual Horn, it must also contain the tuples 1011 and 1101. Since 0001 is not in the relation, we have constructed $[(x \vee y \vee z) \wedge (w = 1)]$. By variable identification we can get the relation $[(x \vee y) \wedge (w = 1)]$ for which we already know MININF to be coNP-complete. Hence we can assume further on that Q' does not contain the tuple 1001.

Case 3: Q' contains 1011. Then it is equal to $[(y \vee z) \wedge (\neg x \vee z) \wedge (w = 1)]$ for which MININF is coNP-complete according to Lemma 9.10.

Case 4: Q' contains 1101. Then it is equal to $[(y \vee z) \wedge (\neg x \vee y) \wedge (w = 1)]$ for which MININF is coNP-complete according to Lemma 9.10.

Case 5: Q' contains both tuples 1101 and 1011. Then we can construct the new constraint $Q''(x, y, z, w) = Q'(x, y, z, w) \wedge Q(x, x, x, x)$ which is equivalent to $(y \vee z) \wedge (x = 1) \wedge (w = 1)$. Since we know that MININF($[x \vee y]$) is coNP-complete, we also know that MININF(Q'') is coNP-complete.

There are no more cases to consider, what terminates the proof. □

9.3 Main Result

Putting together all previous propositions, we obtain the following trichotomy result for the minimal inference problem.

Theorem 9.12 (Trichotomy of MININF) *Let S be a finite nonempty set of Boolean relations and S^* the corresponding 1-valid restriction. If every relation in S is Horn, or 0-valid, or both Schaefer and incomparable, or a subset of the weak co-clone $\langle N \rangle_{\nexists}$ (where $N = \{[\neg x \vee \neg y], [x \not\equiv y]\}$), then MININF($S$) is decidable in polynomial time. Else if S^* is Schaefer, then MININF(S) is coNP-complete. Otherwise, MININF(S) is Π_2P-complete.*

Proof: The parts concerning Π_2P-completeness and membership in coNP follow from the dichotomy theorem due to Kirousis and Kolaitis [21]. As for tractability, this is, as already explained, trivial for S being Horn or 0-valid. Tractability of MININF(S) for S being Schaefer and incomparable, or a subset of $\langle N \rangle_{\nexists}$, is proved in Propositions 8.4, and 8.6, respectively. The coNP-hardness for MININF(S) when S is neither Schaefer, nor 0-valid, and S^* is Schaefer, is proved in Proposition 9.1.

Hence, what remains to be done is to prove coNP-hardness for all sets of relations S that do not fall into one of the tractable classes when S is affine, dual Horn, or bijunctive. This is done in Sections 9.1, 9.2, and 8.5, respectively. For S being affine, the analysis is divided into three cases, depending on whether S is 1-valid (Proposition 9.3), complementive (Proposition 9.8), or neither 1-valid nor complementive (Proposition 9.5). Similarly, the analysis for the dual Horn case is divided into two parts depending on whether S is 1-valid (Proposition 9.11) or not (Proposition 9.9). Finally, the case where S is bijunctive is treated in Proposition 8.18. □

Checking whether a relation R is Schaefer can be done in polynomial time by testing for closure under corresponding operations (see Table 3). Checking whether a relation is 0-valid or incomparable can be done in polynomial time by inspection. Checking whether R is a subset of $\langle N \rangle_{\nexists}$ is more subtle, since we cannot use the Galois correspondence with polymorphisms, because $\langle N \rangle_{\nexists}$ is *not* closed under existential quantification. Recall that $\langle N \rangle_{\nexists}$ is the set of all relations that can be expressed using relations from N, conjunction, and variable identification. However, we can apply Proposition 8.16, since the inclusion $R \subseteq \langle N \rangle_{\nexists}$ implies that R must be bijunctive. It is then sufficient to check that R is closed under majority and that all projections R_{ij} of R are equal to one of the relations $[\neg x \vee \neg y]$ or $[x \not\equiv y]$. This test can be performed in polynomial time.

10 Concluding Remarks

We have proved trichotomy theorems for the complexity of four variants of the generalized minimal inference problem. This also covers the original minimal inference problem, where all variables of the formula φ are being minimized. A careful reader certainly discovered that we rely on many previous results, previously published by Eiter and Gottlob [13], Cadoli and Lenzerini [6], Kirousis and Kolaitis [21], and finally by Nordh and Jonsson [25].

Our trichotomy result for MININF, presented in Theorem 9.12, finally settles with a positive answer the trichotomy conjecture of Kirousis and Kolaitis formulated in [21] and also appearing as one of the open problems (Question 4.1) during the International Workshop on Mathematics of Constraint Satisfaction held in Oxford in March 2006 [27]. In the process we also strengthen and give a much simplified proof of the main result from [11], which concerns the coNP-hardness of MININF over affine relations.

Our proof techniques for the two variants with free variables (i.e., GMININF and VMININF) are based on the well-known and powerful algebraic approach via Post's lattice of co-clones. For the two variants without free variables (i.e., CMININF and MININF) this approach provably does not work since the complexity borderlines cuts through the co-clones. For these problems we instead refine Schaefer's approach via a particular type of *implementations*. There is one significant difference between Schaefer's implementations and ours, namely we only use conjunction, variable identification, and variable permutation (*no* existential quantification) in our implementations. The reason for not allowing existential quantification in our implementations is that the structure of minimal models of a formula is destroyed by existentially quantifying some variables. Moreover as manifested by both the results in this paper and the results due to Kirousis and Kolaitis [21], the complexity of CMININF and MININF is not preserved by implementations using existential quantification.

Our implementations could be interesting in their own right since they might be useful in classifying the complexity of other problems where existential quantification does not preserve the complexity. Note that Schaefer's original implementations have indeed been reused many times and for many different problems [8].

Another observation that we make is that our coNP-hardness proofs for CMININF and MININF seems to require that the clause (query) to be minimally inferred can be arbitrary long (i.e., its length is allowed to depend on the size of the knowledge base). This is in sharp contrast with the Π_2P-hardness proofs for CMININF and MININF proved by Kirousis and Kolaitis [21], where they observe that a Π_2P-hardness holds even when the clause to be inferred consists of two literals.

We have proved a trichotomy for the complexity of the minimal inference problem with unbounded queries. It is natural to ask if such a result can be obtained in the case of bounded queries. Gottlob and Eiter [13] proved the Π_2P-completeness of the minimal inference problem in general already for queries with a single literal which was used by Kirousis and Kolaitis [21] to obtain their dichotomy theorem. However, Cadoli and Lenzerini [6] showed for the bijunctive case that the minimal inference problem is coNP-complete for unbounded queries, but it becomes polynomial-time decidable when the query is restricted to a single literal. The same effect was observed by Durand and Hermann in [11] for the affine case. Our coNP-hardness proofs for the bijunctive, affine, and dual Horn cases are not valid for bounded queries. In fact, the bijunctive case is easily seen to be tractable for bounded queries. The result of [6], showing that there are coNP-complete dual Horn minimal inference problems even for single literal queries, together with the fact that there exist tractable bounded query dual Horn minimal inference problems, indicate that the classification for bounded queries in the dual Horn case is more intricate.

The affine case might be more complicated to solve. It is easy to see that it corresponds to the following problem over representable matroids: given a set S of t elements, find a circuit passing through S. When $t \geq 2$ is bounded, the complexity of this problem is widely open [18]. Notice that the (simpler) problem restricted to graphs can be reduced to the t-disjoint path problem, showed to be polynomial-time decidable after considerable effort [33]. Therefore the final answer to complexity classification of the minimal inference problem with bounded queries still remains a challenging open question.

Acknowledgment: We thank Phokion Kolaitis for valuable contributions in the initial stages of this work. We also thank Steffen Reith for kindly providing us with the visualization of Post's lattice, presented in Figures 1, 2, and 3.

References

[1] K. A. Baker and A. F. Pixley. Polynomial interpolation and the Chinese Remainder Theorem for algebraic systems. *Mathematische Zeitschrift*, 143(2):165–174, 1975.

[2] E. Böhler, N. Creignou, S. Reith, and H. Vollmer. Playing with Boolean blocks, part I: Post's lattice with applications to complexity theory. *SIGACT News, Complexity Theory Column 42*, 34(4):38–52, 2003.

[3] E. Böhler, N. Creignou, S. Reith, and H. Vollmer. Playing with Boolean blocks, part II: Constraint satisfaction problems. *SIGACT News, Complexity Theory Column 43*, 35(1):22–35, 2004.

[4] E. Böhler, S. Reith, H. Schnoor, and H. Vollmer. Bases for Boolean co-clones. *Information Processing Letters*, 96(2):59–66, 2005.

[5] M. Cadoli. The complexity of model checking for circumscriptive formulae. *Information Processing Letters*, 44(3):113–118, 1992.

[6] M. Cadoli and M. Lenzerini. The complexity of propositional closed world reasoning and circumscription. *Journal of Computer and System Sciences*, 48(2):255–310, 1994.

[7] Z.-Z. Chen and S. Toda. The complexity of selecting maximal solutions. *Information and Computation*, 119(2):231–239, 1995.

[8] N. Creignou, S. Khanna, and M. Sudan. *Complexity Classifications of Boolean Constraint Satisfaction Problems*, volume 7 of *SIAM Monographs on Discrete Mathematics and Applications*. SIAM, Philadelphia (PA), 2001.

[9] V. Dalmau. *Computational complexity of problems over generalized formulas*. PhD thesis, Department de Llenguatges i Sistemes Informàtica, Universitat Politécnica de Catalunya, 2000.

[10] J. de Kleer, A. K. Mackworth, and R. Reiter. Characterizing diagnoses and systems. *Artificial Intelligence*, 56(2-3):197–222, 1992.

[11] A. Durand and M. Hermann. The inference problem for propositional circumscription of affine formulas is coNP-complete. In H. Alt and M. Habib, editors, *Proceedings 20th Symposium on Theoretical Aspects of Computer Science (STACS 2003), Berlin (Germany)*, volume 2607 of *Lecture Notes in Computer Science*, pages 451–462. Springer-Verlag, February 2003.

[12] A. Durand, M. Hermann, and G. Nordh. Trichotomy in the complexity of minimal inference. In *Proceedings 24th Annual IEEE Symposium on Logic in Computer Science (LICS 2009), Los Angeles (California, USA)*, pages 387–396, August 2009.

[13] T. Eiter and G. Gottlob. Propositional circumscription and extended closed-world reasoning are Π_2^{P}-complete. *Theoretical Computer Science*, 114(2):231–245, 1993.

[14] D. Geiger. Closed systems of functions and predicates. *Pacific Journal of Mathematics*, 27(1):95–100, 1968.

[15] M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In R. A. Kowalski and K. A. Bowen, editors, *Proceeding 5th International Conference and Symposium on Logic Programming ICLP/SLP, Seatle (Washington)*, pages 1070–1080. MIT Press, 1988.

[16] M. Gelfond, H. Przymusinska, and T. C. Przymusinski. On the relationship between circumscription and negation as failure. *Artificial Intelligence*, 38(1):75–94, 1989.

[17] P. Jeavons, D. Cohen, and M. Gyssens. Closure properties of constraints. *Journal of the Association for Computing Machinery*, 44(4):527–548, 1997.

[18] L. Khachiyan, E. Boros, K. Elbassioni, V. Gurvich, and K. Makino. On the complexity of some enumeration problems for matroids. *SIAM Journal on Discrete Mathematics*, 19(4):966–984, 2005.

[19] L. M. Kirousis and P. G. Kolaitis. On the complexity of model checking and inference in minimal models. In T. Eiter, W. Faber, and M. Truszczyński, editors, *Proceedings 6th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNRM 2001), Vienna (Austria)*, volume 2173 of *Lecture Notes in Computer Science*, pages 42–53. Springer-Verlag, September 2001.

[20] L. M. Kirousis and P. G. Kolaitis. The complexity of minimal satisfiability problems. *Information and Computation*, 187(1):20–39, 2003.

[21] L. M. Kirousis and P. G. Kolaitis. A dichotomy in the complexity of propositional circumscription. *Theory of Computing Systems*, 37(6):695–715, 2004.

[22] J. McCarthy. Circumscription — A form of non-monotonic reasoning. *Artificial Intelligence*, 13(1-2):27–39, 1980.

[23] J. McCarthy. Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28(1):89–116, 1986.

[24] G. Nordh. A trichotomy in the complexity of propositional circumscription. In F. Baader and A. Voronkov, editors, *Proceedings 11th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR 2004), Montevideo (Uruguay)*, volume 3452 of *Lecture Notes in Computer Science*, pages 257–269. Springer-Verlag, March 2005.

[25] G. Nordh and P. Jonsson. An algebraic approach to the complexity of propositional circumscription. In *Proceedings 19th IEEE Symposium on Logic in Computer Science (LICS 2004), Turku (Finland)*, pages 367–376, July 2004.

[26] G. Nordh and B. Zanuttini. Frozen Boolean partial co-clones. In *Proceedings 39th International Symposium on Multiple-Valued Logic (ISMVL 2009), Naha (Okinawa, Japan)*, pages 120–125. IEEE Computer Society, May 2009.

[27] International Workshop on Mathematics of Constraint Satisfaction. Open problems list. URL = `http://www.cs.rhul.ac.uk/home/green/mathscsp/slides/problems/OpenProblemsList.pdf`, March 2006.

[28] C. H. Papadimitriou. On selecting a satisfying truth assignment (extended abstract). In *Proceedings 32nd IEEE Symposium on Foundations of Computer Science (FOCS'91), San Juan (Puerto Rico)*, pages 163–169. IEEE Computer Society, October 1991.

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

[30] R. Pöschel. Galois connection for operations and relations. In K. Denecke, M. Erné, and S. L. Wismath, editors, *Galois connections and applications*, pages 231–258. Kluwer Academic Publishers, 2004.

[31] R. Pöschel and L. A. Kalužnin. *Funktionen- und Relationenalgebren.* Deutscher Verlag der Wissenschaften, Berlin, 1979.

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

[33] N. Robertson and P. D. Seymour. Graph minors XIII: The disjoint paths problem. *Journal of Combinatorial Theory, Series B*, 63(1):65–110, 1995.

[34] B. A. Romov. The algebras of partial functions and their invariants. *Cybernetics and Systems Analysis*, 17(2):157–167, 1981.

[35] T. J. Schaefer. The complexity of satisfiability problems. In *Proceedings 10th Symposium on Theory of Computing (STOC'78), San Diego (California, USA)*, pages 216–226, 1978.

[36] H. Schnoor and I. Schnoor. Partial polymorphisms and constraint satisfaction problems. In N. Creignou, P.G. Kolaitis, and H. Vollmer, editors, *Complexity of Constraints*, volume 5250 of *Lecture Notes in Computer Science*, pages 229–254. Springer-Verlag, 2008.

[37] M. Thomas. The complexity of circumscriptive inference in Post's lattice. In E. Erdem, F. Lin, and T. Schaub, editors, *Proceedings 10th International Conference om Logic Programming and Nonmonotonic Reasoning (LPNMR 2009), Potsdam (Germany)*, volume 5753 of *Lecture Notes in Computer Science*, pages 290–302. Springer-Verlag, 2009.

[38] M. Thomas and H. Vollmer. Complexity of non-monotonic logics. *Bulletin of the EATCS*, 102:53–82, 2010.