The Power of Primitive Positive Definitions with Polynomially Many Variables

Victor Lagerkvist[1] and Magnus Wahlström[2]

[1] Department of Computer and Information Science, Linköping University, Sweden
victor.lagerkvist@liu.se (Corresponding author)
[2] Department of Computer Science, Royal Holloway, University of London, Great Britain
Magnus.Wahlstrom@rhul.ac.uk

Abstract. Two well-studied closure operators for relations are based on existentially quantified conjunctive formulas, *primitive positive* (p.p.) definitions, and primitive positive formulas without existential quantification, *quantifier-free primitive positive definitions* (q.f.p.p.) definitions. Sets of relations closed under p.p. definitions are known as *co-clones* and sets of relations closed under q.f.p.p. definitions as *weak partial co-clones*. The latter do however have limited expressivity, and the corresponding lattice of strong partial clones is of uncountably infinite cardinality even for the Boolean domain. Hence, it is reasonable to consider the expresiveness of p.p. definitions where only a small number of existentially quantified variables are allowed. In this paper we consider p.p. definitions allowing only polynomially many existentially quantified variables, and say that a co-clone closed under such definitions is *polynomially closed*, and otherwise *superpolynomially closed*. We investigate properties of polynomially closed co-clones and prove that if the corresponding clone contains a k-ary near-unanimity operation for $k \geq 3$ then the co-clone is polynomially closed, and if the clone does not contain a k-edge operation for any $k \geq 2$, then the co-clone is superpolynomially closed. For the Boolean domain we strengthen these results and prove a complete dichotomy theorem separating polynomially closed co-clones from superpolynomially closed co-clones. Using these results, we then proceed to investigate properties of strong partial clones corresponding to superpolynomially closed co-clones. We prove that if Γ is a finite set of relations over an arbitrary finite domain such that the clone corresponding to Γ is essentially unary, then the strong partial clone corresponding to Γ is of infinite order and cannot be generated by a finite set of partial functions.

Keywords: Clone theory, partial clone theory, universal algebra, primitive positive definitions, constraint satisfaction problems

1 Introduction

A finite or infinite set of relations Γ over a finite domain is known as a *constraint language*. Given a constraint language Γ, a natural question to ask is which other relations R can be expressed by first order formulas over Γ, or, equivalently, what is the smallest set of relations that contains Γ and is closed under such definitions. In practice one often considers restricted first order formulas, and two common restrictions are *primitive positive definitions* (p.p. definitions), where one is allowed to use existential quantification, conjunction and equality constraints, and *quantifier-free primitive positive definitions* (q.f.p.p. definitions) where only conjunction and equality constraints are allowed. A *relational clone*, or a *co-clone*, is a set of relations closed under p.p. definitions. Any set of relations which generates a given co-clone using p.p. definitions is called a *base* of the co-clone. Similarly, a set of relations closed under q.f.p.p. definitions is referred to as a *weak partial co-clone*, or a *weak system*. Both co-clones and weak partial co-clones have interesting applications in theoretical computer science, and in particular, for the study of the computational complexity of problems parameterized by constraint languages. One noteworthy example is the *constraint satisfaction problem* over a constraint language Γ (CSP(Γ)), which is the problem of determining whether a conjunctive formula over Γ has a model. The use of algebraic techniques to study the complexity of CSP(Γ) is usually referred to as the *algebraic approach* and was first pioneered by Jeavons [15]. The success of the algebraic approach can be mainly attributed to the fact that for

every constraint language Γ one can associate a particular function algebra, known as the *clone* of Γ, or the *polymorphisms* of Γ. The exact details of this relationship between a constraint language and its polymorphisms will be covered later, but for the moment we merely state that this relationship provides a powerful method for analyzing the complexity of constraint satisfaction and related problems. Using this relationship Jeavons proved that the complexity of $\mathrm{CSP}(\Gamma)$ up to polynomial-time reductions is determined by the polymorphisms of Γ [15]. Since then, this result has been extended and used in numerous applications, cf. the excellent survey by Creignou et al. [10] for a broad introduction to this topic.

There is also a similar relationship between weak partial co-clones and sets of partial functions closed under composition, containing all total and partial projection functions, *strong partial clones*. Again, we omit the exact definitions of these concepts for the moment, and just state that for every weak partial co-clone there exists a set of partial functions, *partial polymorphisms*, which completely characterizes this set. With this relationship Jonsson et al. proved that the partial polymorphisms of a constraint language Γ determines the complexity of $\mathrm{CSP}(\Gamma)$ up to $O(c^n)$ complexity [18], where n denotes the number of variables in a given $\mathrm{CSP}(\Gamma)$ instance. This result was used to give lower bounds for all NP-complete Boolean $\mathrm{CSP}(\Gamma)$ problems. Similar results were given in Jonsson et al. [19] but in the context of Boolean optimization problems. Hence, strong partial clones and weak partial co-clones lead to interesting applications when comparing and relating computational problems vis-à-vis $O(c^n)$ time complexity.

Unfortunately, the seemingly subtle steps from p.p. definitions to q.f.p.p. definitions, and from total to partial functions, makes reasoning much more complex. One of the reasons is that, unlike Post's lattice of Boolean clones [24], the lattice of strong partial clones is of uncountably infinite cardinality even for the Boolean domain [1]. Given this fact it is reasonable to consider the expressive power of closure operators which lie between q.f.p.p. definitions and p.p. definitions. To find logical formulas of such intermediate complexity we in this article restrict the number of existentially quantified variables occurring in formulas, and are therefore interested in which n-ary relations that can be p.p. defined with $1, 2, \ldots, p(n)$ existentially quantified variables, for some reasonably slowly growing function p. In the sequel we assume that p is a polynomial function. If $p(n)$ variables are sufficient to define every n-ary relation R in a co-clone then we say that the co-clone is *polynomially closed*. We remark that if if $p(n) \leq 2$ then the resulting set of definable relations over some language Γ closely corresponds to the closure operator considered in Nordh and Zanuttini [23].

The first contribution of this article is a complete classification of the polynomially closed Boolean co-clones (in Section 3). Our proofs are based on comparing the least expressive base of the co-clone with the most expressive base of the co-clone, in order to obtain an upper bound of p. These languages are known as the weak base and plain base, respectively, and were introduced by Schnoor and Schnoor [28], and Creignou et al. [9]. We first give a general result and provide a sufficient condition for a co-clone over any finite domain to be polynomially closed: a co-clone X is polynomially closed if the clone corresponding to X contains a *k-ary near-unanimity* function for some $k \geq 3$. We then complete this classification for the Boolean domain and in addition prove that a Boolean co-clone X is polynomially closed if the polymorphisms of X can be represented by affine functions, or if X is of infinite order (i.e., that X does not have a finite base). To handle the last case we extend Schnoor and Schnoor's result [28] for constructing weak bases and give a condition for the existence of weak bases for co-clones of infinite order. In Section 5 we then proceed with the problem of determining whether a co-clone is superpolynomially closed. We first prove that if the number of n-ary relations in a co-clone is sufficiently large, then, for any finite base of the co-clone, there exists relations which cannot be p.p. defined using a polynomial number of existentially quantified variables. By a result of Berman et al. [3] we then obtain a sufficient condition for verifying whether a co-clone over any finite domain is superpolynomially closed. We remark that for the Boolean domain, a co-clone of finite order is polynomially closed if and only if the corresponding clone contains a *k-edge function* for some $k \geq 2$, or, equivalently, if the clone has *few subpowers* [3]. Interestingly, this does not hold for co-clones of infinite order, which suggests a quantitative difference between our notion and that of Berman et al.

The second contribution of this article (in Section 6) is an investigation of the structure of the partial polymorphisms of finite constraint languages corresponding to superpolynomially closed co-clones. Before we can present this result, we need a few additional preliminaries. Given a constraint language Γ, say that the

set of partial polymorphisms of Γ is of *finite order* if there exists a finite set of partial functions F which can generate this set, using the standard notion of functional composition, and of *infinite order* otherwise. The set F is in this case called a *base* of the set of partial polymorphisms of Γ. Assume e.g. that $R_{1/3} = \{(0,0,1),(0,1,0),(1,0,0)\}$, and observe that $\mathrm{CSP}(\{R_{1/3}\})$ is an alternative formulation of the well-known NP-complete problem 1-IN-3-SAT. It is easy to verify that the co-clone of $\{R_{1/3}\}$ is the set of all Boolean relations, and from the results in Section 5 we know that this set is superpolynomially closed. Given the fact that the partial polymorphisms of a constraint language has a close relationship with the worst-case time complexity of the corresponding CSP problem [18], obtaining a finite base of the set of partial polymorphisms of $R_{1/3}$ would likely increase our understanding of the time complexity of 1-IN-3-SAT. We prove that such a finite base *cannot exist* (irregardless of any complexity theoretical assumptions). In fact, we prove something stronger: let Γ be a finite constraint language over an arbitrary finite domain. If the co-clone of Γ is superpolynomially closed, and if the polymorphisms of Γ are essentially unary, then the set of partial polymorphisms of Γ is of infinite order. This result can be seen as a continuation of the research by Haddad and Börner [7] who gave a condition for checking whether a strong partial clone is infinitely generated, but our result also has many practical consequences for the applicability of partial clone theory to the study of the computational complexity of NP-hard CSP problems. Also, it is worth noting that the fact that even though a given strong partial clone is of infinite order, it might still be possible to give a reasonably simple characterization of its functions. This problem was investigated in Lagerkvist et al. [22] by considering stronger notions of closure than functional composition.

2 Preliminaries and Notation

If Γ is a constraint language we let $\Gamma^{(n)}$ be defined as $\{R \mid R \in \Gamma, \mathrm{ar}(R) \leq n\}$, where $\mathrm{ar}(R)$ is the arity of the relation R. Given a finite domain D, we let Rel_D be the set of all finitary relations over D, OP_D be the set of all functions over D, and we let Eq_D denote the equality relation $\{(x,x) \mid x \in D\}$ over D. For a tuple $t = (x_1, \ldots, x_i, \ldots, x_n) \in D^n$ we let $t[i] = x_i$. An *n-ary projection function* over D is a function π_i^n which for some $i \in \{1, \ldots, n\}$ satisfies $\pi_i^n(x_1, \ldots, x_i, \ldots, x_n) = x_i$ for all $(x_1, \ldots, x_i, \ldots, x_n) \in D^n$. We typically represent relations and constraint languages by their defining logical formulas, and write $R(x_1, \ldots, x_n) \equiv \phi$, where ϕ is a logical first-order formula, to denote the n-ary relation $R = \{(f(x_1), \ldots, f(x_n)) \mid f \text{ is a model of } \phi\}$. As a convenience we often write $\bar{x}$ instead of $\neg x$.

2.1 Clones, Co-Clones and Galois Connection

Let Γ be a constraint language over a finite domain D. If f is a function over D it is said to be a *polymorphism* of Γ, or that Γ is *invariant* under f, if, for every relation $R \in \Gamma$, $f(t_1, \ldots, t_n) \in R$ for all $t_1, \ldots, t_n \in R$. Here $f(t_1, \ldots, t_n)$ denotes the $\mathrm{ar}(R)$-ary tuple obtained by the component-wise application of f to $t_1, \ldots, t_n$, i.e., $f(t_1, \ldots, t_n) = (f(t_1[1], \ldots, t_n[1]), \ldots, f(t_1[\mathrm{ar}(R)], \ldots, t_n[\mathrm{ar}(R)]))$. If F is a set of functions over D and Γ a set of relations we let $\mathrm{Pol}_D(\Gamma)$ denote the set of all polymorphisms over D of Γ, and $\mathrm{Inv}_D(F)$ denote the set of all relations over D that are invariant under F. If the domain is clear from the context we simply write $\mathrm{Inv}(F)$ and $\mathrm{Pol}(\Gamma)$, respectively.

Sets of the form $\mathrm{Pol}(\Gamma)$ are usually referred to as *clones*, and, as can be verified, are composition-closed sets of functions containing all projection functions. That is, if $f \in \mathrm{Pol}(\Gamma)$ is an n-ary function and $g_1, \ldots, g_n \in \mathrm{Pol}(\Gamma)$ are m-ary functions, then $\mathrm{Pol}(\Gamma)$ also contains the m-ary function

$$(f \circ (g_1, \ldots, g_n))(x_1, \ldots, x_m) = f(g_1(x_1, \ldots, x_m), \ldots, g_n(x_1, \ldots, x_m)).$$

Dually, sets of the form $\mathrm{Inv}(F)$ are referred to as *co-clones*, and are sets of relations closed under *primitive positive definitions* (p.p. definitions), i.e, whenever $\Gamma \subseteq \mathrm{Inv}(F)$ then $\mathrm{Inv}(F)$ also contains all n-ary relations R of the form $R(x_1, \ldots, x_n) \equiv \exists y_1, \ldots, y_{n'} . R_1(\mathbf{x_1}) \wedge \ldots \wedge R_m(\mathbf{x_m})$, where each $R_i \in \Gamma \cup \{\mathrm{Eq}\}$ and each $\mathbf{x_i}$ is an $\mathrm{ar}(R_i)$-ary tuple of variables over $x_1, \ldots, x_n, y_1, \ldots, y_{n'}$. Let $[F] = \mathrm{Pol}(\mathrm{Inv}(F))$ and $\langle \Gamma \rangle = \mathrm{Inv}(\mathrm{Pol}(\Gamma))$, and note that $[F]$ is then the smallest clone containing F, while $\langle \Gamma \rangle$ is the smallest co-clone containing Γ. The sets
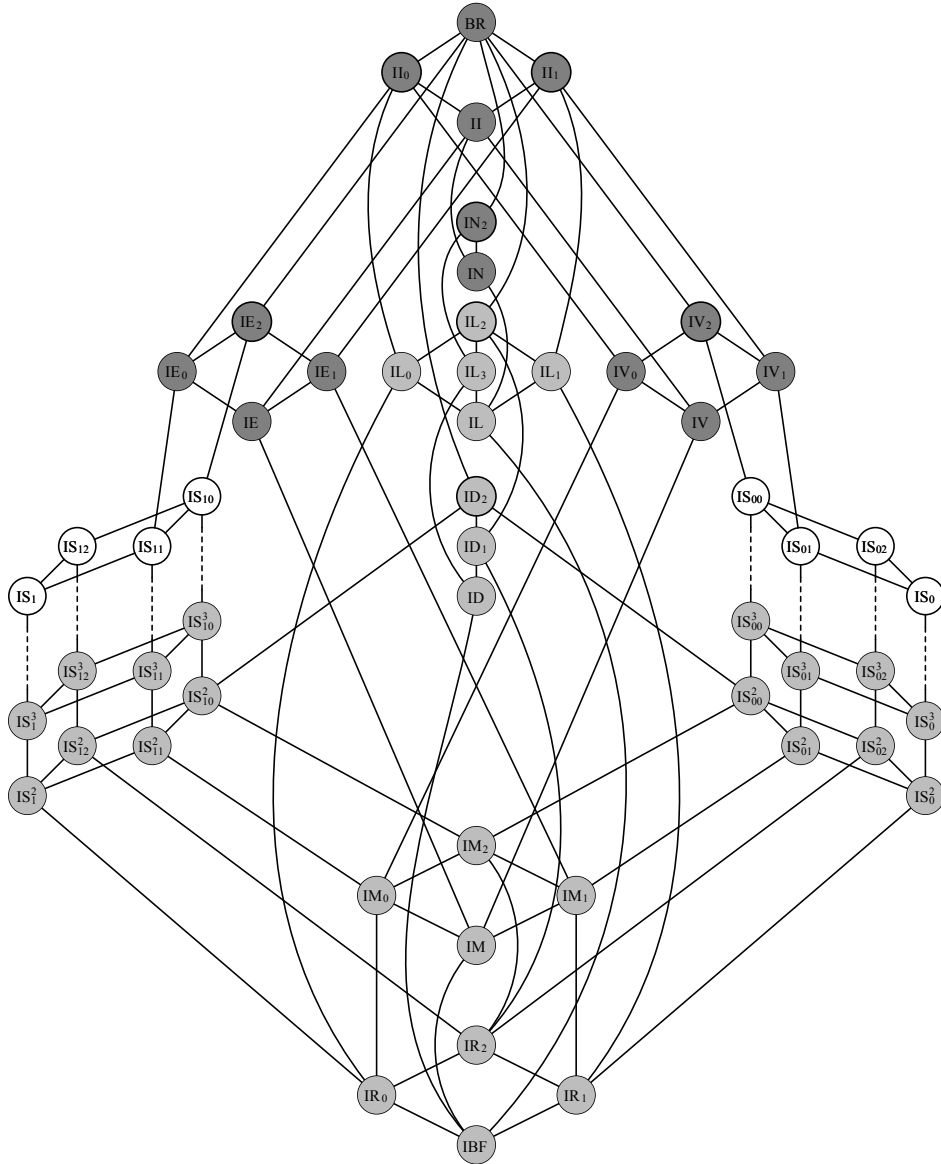
4

Fig. 1. The lattice of Boolean co-clones. The co-clones of finite order which are polynomially closed are coloured in grey. The co-clones of infinite order that are polynomially closed are coloured in white. The superpolynomially closed co-clones are coloured in dark grey.

F and Γ are said to be *bases* of $[F]$ and $\langle \Gamma \rangle$, respectively, and a clone or a co-clone is said to be of *finite order* if it has a finite base, and is said to be of *infinite order* otherwise. We have the following *Galois connection* between $\text{Inv}(\cdot)$ and $\text{Pol}(\cdot)$.

Theorem 1 ([4, 5, 12]). *Let Γ and Δ be two constraint languages. Then $\Gamma \subseteq \langle \Delta \rangle$ if and only if $\text{Pol}(\Delta) \subseteq \text{Pol}(\Gamma)$.*

If D is a finite domain it is well known that the set of all clones over D form a lattice structure when ordered by set inclusion, where the meet-operator $\sqcap$ is defined as $X \sqcap Y = X \cap Y$ and the join-operator $\sqcup$ as $X \sqcup Y = [X \cup Y]$. Similarly, the set of all co-clones over D also form a lattice structure under set inclusion, where $X \sqcap Y = X \cap Y$ and $X \sqcup Y = \langle X \cup Y \rangle$. For the Boolean domain, all clones have been completely determined, and the lattice of Boolean clones is typically referred to as *Post's lattice* due to Post's original classification [24]. See Table 1 for a complete list of Boolean clones and their bases [6]. As a shorthand we let BF denote the set of all Boolean functions and BR the set of all Boolean relations. Due to the Galois connection in Theorem 1, each clone C in Table 1 uniquely determines a co-clone $\text{Inv}(C)$, and it is not difficult to see that the lattice of Boolean co-clones is dually isomorphic to the lattice of Boolean clones. See Figure 1 for a visualization of the Boolean co-clone lattice. In this figure, each node IC is an abbreviation of $\text{Inv}(C)$, where C is a clone from Table 1.

In this article we in addition need more restricted closure operators. Say that an n-ary relation R has a *quantifier-free primitive positive* (q.f.p.p.) definition in a constraint language Γ if $R(x_1, \ldots, x_n) \equiv R_1(\mathbf{x_1}) \wedge \ldots \wedge R_k(\mathbf{x_k})$, where each $R_i \in \Gamma \cup \{\text{Eq}\}$ and each $\mathbf{x_i}$ is an $\text{ar}(R_i)$-ary tuple of variables over $x_1, \ldots, x_n$. Hence, q.f.p.p. definitions are more restricted than p.p. definitions since we do not allow existential quantification. We also need an alternative notion of polymorphisms. An *n-ary partial function* over a finite domain is a map $f : X \mapsto D$, where $X \subseteq D^n$. In other words X is the set of arguments for which the function is defined. We let $\text{domain}(f) = X$, and say that f is *undefined* for all $(x_1, \ldots, x_n) \in D^n \setminus \text{domain}(f)$. Composition of partial functions is defined in an analogous manner to the case of total functions, but the resulting function is only defined for those arguments where all involved functions are defined. Say that an n-ary partial function f is a *partial polymorphism* of a constraint language Γ, or that Γ is *invariant* under f, if, for every $R \in \Gamma$, $f(t_1, \ldots, t_n) \in R$ for all $t_1, \ldots, t_n \in R$ such that $f(t_1, \ldots, t_n)$ is defined. Let $\text{pPol}_D(\Gamma)$ denote the set of all partial polymorphisms over D of a constraint language Γ and $\text{Inv}_D(F)$ the set of all relations over D invariant under the set of partial functions F. Sets of the form $\text{pPol}(\Gamma)$ are known as *strong partial clones*, and are composition-closed sets of partial functions, containing all projection functions, and closed under subfunctions. The latter means that whenever $f \in \text{pPol}(\Gamma)$ then $\text{pPol}(\Gamma)$ also contains all functions g such that $\text{domain}(g) \subseteq \text{domain}(f)$ and such that $g(x_1, \ldots, x_n) = f(x_1, \ldots, x_n)$ for all $(x_1, \ldots, x_n) \in \text{domain}(g)$. Sets of the form $\text{Inv}(\text{pPol}(\Gamma))$ are known as *weak partial co-clones*, or, *weak systems*, and are sets of relations closed under q.f.p.p. definability. Given a set of partial functions F and a constraint language Γ let $[F]_s = \text{pPol}(\text{Inv}(F))$ and $\langle \Gamma \rangle_{\not\exists} = \text{Inv}(\text{pPol}(\Gamma))$. As is easily verified $[F]_s$ is the smallest strong partial clone containing F and $\langle \Gamma \rangle_{\not\exists}$ the smallest weak partial co-clone containing Γ. Observe that this implies that $\langle \Gamma \rangle_{\not\exists}$ is the smallest set of relations which is closed under q.f.p.p. definitions over Γ. Similar to the definitions for clones and co-clones the sets F and Γ are said to be *bases* of $[F]_s$ and $\langle \Gamma \rangle_{\not\exists}$, and we say that a strong partial clone or a weak partial co-clone is of *finite order* if it has a finite base, and is of *infinite order* otherwise. The relationship between strong partial clones and weak partial co-clones is given by the following Galois connection.

Theorem 2 ([12, 26]). *Let Γ and Δ be two constraint languages. Then $\Gamma \subseteq \langle \Delta \rangle_{\not\exists}$ if and only if $\text{pPol}(\Delta) \subseteq \text{pPol}(\Gamma)$.*

2.2 Weak and Plain Bases of Co-Clones

The structure of the lattice of strong partial clones is largely undetermined, since it is of uncountably infinite cardinality for every non-trivial finite domain [1]. Due to the Galois connection in Theorem 2, this also implies that the dually isomorphic lattice of weak partial co-clones is of uncountably infinite cardinality. Despite this, it is possible to describe parts of this lattice by considering a particular kind of sublattice.

Table 1. List of all Boolean clones with definitions and bases, where $\text{id}(x) = x$ and $h_n(x_1,\ldots,x_{n+1}) = \bigvee_{i=1}^{n+1} x_1 \cdots x_{i-1} x_{i+1} \cdots x_{n+1}$, $\text{dual}(f)(a_1,\ldots,a_n) = 1 - f(\overline{a_1},\ldots,\overline{a_n})$.

Clone	Definition	Base
BF	All Boolean functions	$\{x \wedge y, \neg x\}$
R_0	$\{f \mid f$ is 0-reproducing$\}$	$\{x \wedge y, x \oplus y\}$
R_1	$\{f \mid f$ is 1-reproducing$\}$	$\{x \vee y, x \oplus y \oplus 1\}$
R_2	$R_0 \cap R_1$	$\{x \vee y, x \wedge (y \oplus z \oplus 1)\}$
M	$\{f \mid f$ is monotonic$\}$	$\{x \vee y, x \wedge y, 0, 1\}$
M_1	$M \cap R_1$	$\{x \vee y, x \wedge y, 1\}$
M_0	$M \cap R_0$	$\{x \vee y, x \wedge y, 0\}$
M_2	$M \cap R_2$	$\{x \vee y, x \wedge y\}$
S_0^n	$\{f \mid f$ is 0-separating of degree $n\}$	$\{x \to y, \text{dual}(h_n)\}$
S_0	$\{f \mid f$ is 0-separating$\}$	$\{x \to y\}$
S_1^n	$\{f \mid f$ is 1-separating of degree $n\}$	$\{x \wedge \neg y, h_n\}$
S_1	$\{f \mid f$ is 1-separating$\}$	$\{x \wedge \neg y\}$
S_{02}^n	$S_0^n \cap R_2$	$\{x \vee (y \wedge \neg z), \text{dual}(h_n)\}$
S_{02}	$S_0 \cap R_2$	$\{x \vee (y \wedge \neg z)\}$
S_{01}^n	$S_0^n \cap M$	$\{\text{dual}(h_n), 1\}$
S_{01}	$S_0 \cap M$	$\{x \vee (y \wedge z), 1\}$
S_{00}^n	$S_0^n \cap R_2 \cap M$	$\{x \vee (y \wedge z), \text{dual}(h_n)\}$
S_{00}	$S_0 \cap R_2 \cap M$	$\{x \vee (y \wedge z)\}$
S_{12}^n	$S_1^n \cap R_2$	$\{x \wedge (y \vee \neg z), h_n\}$
S_{12}	$S_1 \cap R_2$	$\{x \wedge (y \vee \neg z)\}$
S_{11}^n	$S_1^n \cap M$	$\{h_n, 0\}$
S_{11}	$S_1 \cap M$	$\{x \wedge (y \vee z), 0\}$
S_{10}^n	$S_1^n \cap R_2 \cap M$	$\{x \wedge (y \vee z), h_n\}$
S_{10}	$S_1 \cap R_2 \cap M$	$\{x \wedge (y \vee z)\}$
D	$\{f \mid f$ is self-dual$\}$	$\{(x \wedge \neg y) \vee (x \wedge \neg z) \vee (\neg y \wedge \neg z)\}$
D_1	$D \cap R_2$	$\{(x \wedge y) \vee (x \wedge \neg z) \vee (y \wedge \neg z)\}$
D_2	$D \cap M$	$\{h_2\}$
L	$\{f \mid f$ is affine$\}$	$\{x \oplus y, 1\}$
L_0	$L \cap R_0$	$\{x \oplus y\}$
L_1	$L \cap R_1$	$\{x \oplus y \oplus 1\}$
L_2	$L \cap R_2$	$\{x \oplus y \oplus z\}$
L_3	$L \cap D$	$\{x \oplus y \oplus z \oplus 1\}$
V	$\{f \mid f$ is a disjunction or constant$\}$	$\{x \vee y, 0, 1\}$
V_0	$V \cap R_0$	$\{x \vee y, 0\}$
V_1	$V \cap R_1$	$\{x \vee y, 1\}$
V_2	$V \cap R_2$	$\{x \vee y\}$
E	$\{f \mid f$ is a conjunction or constant$\}$	$\{x \wedge y, 0, 1\}$
E_0	$E \cap R_0$	$\{x \wedge y, 0\}$
E_1	$E \cap R_1$	$\{x \wedge y, 1\}$
E_2	$E \cap R_2$	$\{x \wedge y\}$
N	$\{f \mid f$ depends on at most one variable$\}$	$\{\neg x, 0, 1\}$
N_2	$N \cap R_2$	$\{\neg x\}$
I	$\{f \mid f$ is a projection or a constant$\}$	$\{\text{id}, 0, 1\}$
I_0	$I \cap R_0$	$\{\text{id}, 0\}$
I_1	$I \cap R_1$	$\{\text{id}, 1\}$
I_2	$I \cap R_2$	$\{\text{id}\}$

Definition 3. *Let C be a clone over a finite domain D. The* interval *of C, $\mathcal{I}(C)$, is the set $\mathcal{I}(C) = \{\mathrm{pPol}(\Delta) \mid \Delta \subseteq \mathrm{Rel}_D, \mathrm{Pol}(\Delta) = C)\}$.*

Hence, the interval $\mathcal{I}(C)$ of a clone C is simply the set of all strong partial clones where the total component equals C. Even though $\mathcal{I}(C)$ can still be of uncountably infinite cardinality [29], it is known that there always exists a largest and smallest element [28]. A constraint language Γ_w such that $\mathrm{pPol}(\Gamma_w) \in \mathcal{I}(C)$ satisfying $\mathrm{pPol}(\Gamma_w) \supseteq \mathrm{pPol}(\Delta)$ for any $\mathrm{pPol}(\Delta) \in \mathcal{I}(C)$ is called a *weak base* of $\mathrm{Inv}(C)$ [28]. Note that if a co-clone is of finite order then a weak base can always be given as a single relation. As one can verify with the Galois connection, from the functional point of view, a weak base Γ_w results in the largest element $\mathrm{pPol}(\Gamma_w)$ in $\mathcal{I}(C)$, but from the relational point of view, the weak base has the least expressive power with respect to q.f.p.p. definability. Hence, we have the following theorem.

Theorem 4 ([28]). *Let Γ_w be a weak base of a co-clone $\mathrm{Inv}(\mathsf{C})$. Then $\Gamma_w \subseteq \langle \Gamma \rangle_{\not\exists}$ for any base Γ of $\mathrm{Inv}(\mathsf{C})$.*

Dually, a constraint language Γ_p such that $\Gamma_p \in \mathcal{I}(C)$ and satisfying $\mathrm{pPol}(\Gamma_p) \subseteq \mathrm{pPol}(\Gamma)$ for any $\mathrm{pPol}(\Gamma) \in \mathcal{I}(C)$ is called a *plain base* of $\mathrm{Inv}(C)$ [9]. Again, using the Galois connection, we see that $\mathrm{pPol}(\Gamma_p)$ is the smallest element in $\mathcal{I}(C)$ but that Γ_p is the most expressive language with respect to q.f.p.p. definability. Hence, we have the following theorem.

Theorem 5. *Let Γ_p be a plain base of a co-clone $\mathrm{Inv}(C)$. Then $R \in \langle \Gamma_p \rangle_{\not\exists}$ for any $R \in \mathrm{Inv}(C)$.*

It is not difficult to verify that $\mathrm{Inv}(C)$ is a plain base of itself since $[C]_s = \mathrm{pPol}(\mathrm{Inv}(C))$ and $\langle \mathrm{Inv}(C) \rangle_{\not\exists} = \mathrm{Inv}(C)$. However, Creignou et al. [9] gave a much more systematic and highly regular description of plain bases for Boolean co-clones. These bases can be found in Table 2, and we remark that every such plain base Γ_p in addition fullfils the condition that $R \in \langle \Gamma^{(n)} \rangle_{\not\exists}$ for each n-ary $R \in \langle \Gamma_p \rangle$. Hence, we have the following theorem.

Theorem 6. *Let Γ_p be the plain base from Table 2 for some Boolean co-clone $\mathrm{Inv}(\mathsf{C})$. Then $R \in \langle \Gamma_p^{(n)} \rangle_{\not\exists}$ for any n-ary $R \in \mathrm{Inv}(\mathsf{C})$.*

For weak bases, Schnoor and Schnoor [28] gave a systematic procedure for obtaining weak bases, which was later refined in Lagerkvist [20] in order to get a complete list of weak bases for all Boolean co-clones of finite order. These relations can be found in Table 2. We give a short description of some of the involved relations: for a full description, see Lagerkvist [20, 21] and Creignou et al. [9]. We write F and T for the constant relations $\{(0)\}$ and $\{(1)\}$; $\mathrm{OR}^n(x_1, \ldots, x_n)$ for the disjunction $x_1 \vee \ldots \vee x_n$, $\mathrm{NAND}^n(x_1, \ldots, x_n)$ for the relation $\overline{x_1} \vee \ldots \vee \overline{x_n}$; and define the $(n+m)$-ary relation $\mathrm{Compl}_{m,n}$ as

$$\mathrm{Compl}_{m,n}(x_1, \ldots, x_{m+n}) \equiv (x_1 \vee \ldots \vee x_n \vee \overline{x_{n+1}} \vee \ldots \vee \overline{x_{n+m}}) \wedge (\overline{x_1} \vee \ldots \vee \overline{x_n} \vee x_{n+1} \vee \ldots \vee x_{n+m}).$$

2.3 The Constraint Satisfaction Problem

The *constraint satisfaction problem* over a constraint language Γ $(\mathrm{CSP}(\Gamma))$ is the following computational decision problem.

INSTANCE: A set V of variables and a set C of constraint applications $R(v_1, \ldots, v_k)$ where $R \in \Gamma$, $k = \mathrm{ar}(R)$, and $v_1, \ldots, v_k \in V$.

QUESTION: Is there a function $f : V \to D$ such that $(f(v_1), \ldots, f(v_k)) \in R$ for each $R(v_1, \ldots, v_k)$ in C?

The $\mathrm{CSP}(\Gamma)$ problem is in general NP-complete and can be used to model many classical NP-complete problems such as the k-colorability problem and the k-clique problem [15]. Jeavons et al. proved that the complexity of $\mathrm{CSP}(\Gamma)$, up to polynomial-time reductions, is determined by $\mathrm{Pol}(\Gamma)$ [17]. With this result Schaefer's dichotomy theorem for the Boolean satisfiability problem [27] can be formulated in a particularly simple way: for Boolean constraint languages Γ, $\mathrm{CSP}(\Gamma)$ is NP-complete if and only if $\mathrm{Pol}(\Gamma) \subseteq [\neg x]$. Consulting Table 1 we see that this furthermore holds if and only if $\mathrm{Pol}(\Gamma) \in \{\mathsf{I}_2, \mathsf{N}_2\}$.

Table 2. Weak and plain bases of all Boolean co-clones of finite order.

Co-clone	Weak base	Plain base
$\mathrm{Inv}(\mathsf{BF})$	$\mathrm{Eq}(x_1, x_2)$	$\{\mathrm{Eq}(x_1, x_2)\}$
$\mathrm{Inv}(\mathsf{R_0})$	$\mathrm{F}(c_0)$	$\{\mathrm{F}(c_0)\}$
$\mathrm{Inv}(\mathsf{R_1})$	$\mathrm{T}(c_1)$	$\{\mathrm{T}(c_1)\}$
$\mathrm{Inv}(\mathsf{R_2})$	$\mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{F}(c_0), \mathrm{T}(c_1)\}$
$\mathrm{Inv}(\mathsf{M})$	$(x_1 \to x_2)$	$\{(x_1 \to x_2)\}$
$\mathrm{Inv}(\mathsf{M_0})$	$(x_1 \to x_2) \wedge \mathrm{F}(c_0)$	$\{(x_1 \to x_2), \mathrm{F}(c_0)\}$
$\mathrm{Inv}(\mathsf{M_1})$	$(x_1 \to x_2) \wedge \mathrm{T}(c_1)$	$\{(x_1 \to x_2), \mathrm{T}(c_1)\}$
$\mathrm{Inv}(\mathsf{M_2})$	$(x_1 \to x_2) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{(x_1 \to x_2), \mathrm{F}(c_0), \mathrm{T}(c_1)\}$
$\mathrm{Inv}(\mathsf{S_0^n}), n \geq 2$	$\mathrm{OR}^n(x_1, \ldots, x_n) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n)\}$
$\mathrm{Inv}(\mathsf{S_{02}^n}), n \geq 2$	$\mathrm{OR}^n(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n), \mathrm{F}(c_0)\}$
$\mathrm{Inv}(\mathsf{S_{01}^n}), n \geq 2$	$\mathrm{OR}^n(x_1, \ldots, x_n) \wedge (x \to x_1 \cdots x_n) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n), (x_1 \to x_2)\}$
$\mathrm{Inv}(\mathsf{S_{00}^n}), n \geq 2$	$\mathrm{OR}^n(x_1, \ldots, x_n) \wedge (x \to x_1 \cdots x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n), (x_1 \to x_2), \mathrm{F}(c_0)\}$
$\mathrm{Inv}(\mathsf{S_1^n}), n \geq 2$	$\mathrm{NAND}^n(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n)\}$
$\mathrm{Inv}(\mathsf{S_{12}^n}), n \geq 2$	$\mathrm{NAND}^n(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n), \mathrm{T}(c_1)\}$
$\mathrm{Inv}(\mathsf{S_{11}^n}), n \geq 2$	$\mathrm{NAND}^n(x_1, \ldots, x_n) \wedge \neg(x \to x_1 \cdots x_n) \wedge \mathrm{F}(c_0)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n), (x_1 \to x_2)\}$
$\mathrm{Inv}(\mathsf{S_{10}^n}), n \geq 2$	$\mathrm{NAND}^n(x_1, \ldots, x_n) \wedge \neg(x \to x_1 \cdots x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n), (x_1 \to x_2), \mathrm{T}(c_1)\}$
$\mathrm{Inv}(\mathsf{D})$	$(x_1 \oplus x_2 = 1)$	$\{(x_1 \oplus x_2 = 1)\}$
$\mathrm{Inv}(\mathsf{D_1})$	$(x_1 \oplus x_2 = 1) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{(x_1 \oplus x_2 = 1)\} \cup \{\mathrm{F}(c_0), \mathrm{T}(c_1)\}$
$\mathrm{Inv}(\mathsf{D_2})$	$\mathrm{OR}_{2\neq}^2(x_1, x_2, x_3, x_4) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{F}(c_0), \mathrm{T}(c_1), (x_1 \vee x_2), (\neg x_1 \vee x_2), (\neg x_1 \vee \neg x_2)\}$
$\mathrm{Inv}(\mathsf{L})$	$\mathrm{EVEN}^4(x_1, x_2, x_3, x_4)$	$\{(x_1 \oplus \ldots \oplus x_k = 0) \mid k \, \mathrm{even}\}$
$\mathrm{Inv}(\mathsf{L_0})$	$\mathrm{EVEN}^3(x_1, x_2, x_3) \wedge \mathrm{F}(c_0)$	$\{(x_1 \oplus \ldots \oplus x_k = 0) \mid k \in \mathbb{N}\}$
$\mathrm{Inv}(\mathsf{L_1})$	$\mathrm{ODD}^3(x_1, x_2, x_3) \wedge \mathrm{T}(c_1)$	$\{(x_1 \oplus \ldots \oplus x_k = c) \mid k \in \mathbb{N}, c = k \bmod 2\}$
$\mathrm{Inv}(\mathsf{L_2})$	$\mathrm{EVEN}_{3\neq}^3(x_1, \ldots, x_6) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{(x_1 \oplus \ldots \oplus x_k = c) \mid k \in \mathbb{N}, c \in \{0, 1\}\}$
$\mathrm{Inv}(\mathsf{L_3})$	$\mathrm{EVEN}_{4\neq}^4(x_1, \ldots, x_8)$	$\{(x_1 \oplus \ldots \oplus x_k = c) \mid k \, \mathrm{even}, c \in \{0, 1\}\}$
$\mathrm{Inv}(\mathsf{V})$	$(\overline{x_1} \leftrightarrow \overline{x_2}\,\overline{x_3}) \wedge (\overline{x_2} \vee \overline{x_3} \to \overline{x_4})$	$\{(x_1 \vee \ldots \vee x_k \vee \neg x) \mid k \geq 1\}$
$\mathrm{Inv}(\mathsf{V_0})$	$(\overline{x_1} \leftrightarrow \overline{x_2}\,\overline{x_3}) \wedge \mathrm{F}(c_0)$	$\{(x_1 \vee \ldots \vee x_k \vee \neg x) \mid k \in \mathbb{N}\}$
$\mathrm{Inv}(\mathsf{V_1})$	$(\overline{x_1} \leftrightarrow \overline{x_2}\,\overline{x_3}) \wedge (\overline{x_2} \vee \overline{x_3} \to \overline{x_4}) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n) \mid n \in \mathbb{N}\} \cup \{(x_1 \vee \ldots \vee x_k \vee \neg x) \mid k \geq 1\})$
$\mathrm{Inv}(\mathsf{V_2})$	$(\overline{x_1} \leftrightarrow \overline{x_2}\,\overline{x_3}) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{OR}^n(x_1, \ldots, x_n) \mid n \in \mathbb{N}\} \cup \{(x_1 \vee \ldots \vee x_k \vee \neg x) \mid k \in \mathbb{N}\})$
$\mathrm{Inv}(\mathsf{E})$	$(x_1 \leftrightarrow x_2 x_3) \wedge (x_2 \vee x_3 \to x_4)$	$\{(\neg x_1 \vee \ldots \vee \neg x_k \vee x) \mid k \geq 1\}$
$\mathrm{Inv}(\mathsf{E_0})$	$(x_1 \leftrightarrow x_2 x_3) \wedge (x_2 \vee x_3 \to x_4) \wedge \mathrm{F}(c_0)$	$\{(\neg x_1 \vee \ldots \vee \neg x_k \vee x) \mid k \in \mathbb{N}\}$
$\mathrm{Inv}(\mathsf{E_1})$	$(x_1 \leftrightarrow x_2 x_3) \wedge \mathrm{T}(c_1)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n) \mid n \in \mathbb{N}\} \cup \{(\neg x_1 \vee \ldots \vee \neg x_k \vee x) \mid k \geq 1\})$
$\mathrm{Inv}(\mathsf{E_2})$	$(x_1 \leftrightarrow x_2 x_3) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{\mathrm{NAND}^n(x_1, \ldots, x_n) \mid n \in \mathbb{N}\} \cup \{(\neg x_1 \vee \ldots \vee \neg x_k \vee x) \mid k \in \mathbb{N}\})$
$\mathrm{Inv}(\mathsf{N})$	$\mathrm{EVEN}^4(x_1, x_2, x_3, x_4) \wedge x_1 x_4 \leftrightarrow x_2 x_3$	$\{\mathrm{Compl}_{m,n} \mid m, n \geq 1\}$
$\mathrm{Inv}(\mathsf{N_2})$	$\mathrm{EVEN}_{4\neq}^4(x_1, \ldots, x_8) \wedge x_1 x_4 \leftrightarrow x_2 x_3$	$\{\mathrm{Compl}_{m,n} \mid m, n \in \mathbb{N}\}$
$\mathrm{Inv}(\mathsf{I})$	$(x_1 \leftrightarrow x_2 x_3) \wedge (\overline{x_4} \leftrightarrow \overline{x_2}\,\overline{x_3})$	$\{(x_1 \vee \ldots \vee x_m \vee \neg y_1 \vee \ldots \neg y_n) \mid m, n \geq 1\}$
$\mathrm{Inv}(\mathsf{I_0})$	$(\overline{x_1} \vee \overline{x_2}) \wedge (\overline{x_1}\,\overline{x_2} \leftrightarrow \overline{x_3}) \wedge \mathrm{F}(c_0)$	$\{(x_1 \vee \ldots \vee x_m \vee \neg y_1 \vee \ldots \neg y_n) \mid m \in \mathbb{N}, n \geq 1\}$
$\mathrm{Inv}(\mathsf{I_1})$	$(x_1 \vee x_2) \wedge (x_1 x_2 \leftrightarrow x_3) \wedge \mathrm{T}(c_1)$	$\{(x_1 \vee \ldots \vee x_m \vee \neg y_1 \vee \ldots \neg y_n) \mid m \geq 1, n \in \mathbb{N}\}$
BR	$R_{3\neq}^{1/3}(x_1, \ldots, x_6) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$	$\{(x_1 \vee \ldots \vee x_m \vee \neg y_1 \vee \ldots \neg y_n) \mid m, n \in \mathbb{N}\}$

Similarly, it has been shown that $\mathrm{pPol}(\Gamma)$ determines the complexity of $\mathrm{CSP}(\Gamma)$ up to $O(c^n)$ time complexity [18], where n denotes the number of variables in a given $\mathrm{CSP}(\Gamma)$ instance. Hence, a better understanding of the partial polymorphisms of a constraint language Γ could lead to a better understanding of the worst-case time complexity of $\mathrm{CSP}(\Gamma)$. However, as we will see in Section 6, obtaining simple characterizations of strong partial clones $\mathrm{pPol}(\Gamma)$ is likely very difficult for many natural choices of constraint languages such that $\mathrm{CSP}(\Gamma)$ is NP-complete.

3 Polynomially Closed Co-Clones of Finite Order

In this section we formally introduce the notion of a polynomially closed co-clone. Intuitively, the notion means that for any base of the co-clone, a polynomial amount of variables is sufficient to p.p. define any relation in the co-clone.

Definition 7. *Let* $\mathrm{Inv}(\mathsf{C})$ *be a co-clone over a finite domain. We say that* $\mathrm{Inv}(\mathsf{C})$ *is* polynomially closed *if there exists a polynomial p such that for all bases Γ of* $\mathrm{Inv}(\mathsf{C})$ *and all n-ary $R \in \mathrm{Inv}(\mathsf{C})$ it holds that R can be p.p. defined in Γ with at most $p(n)$ existentially quantified variables.*

Observe that $\mathrm{Inv}(\mathsf{C})$ in this definition is allowed to be of infinite order. In this section, however, we restrict our focus to co-clones of finite order, while we in Section 4 investigate co-clones of infinite order. If a co-clone is not polynomially closed then we say that it is *superpolynomially closed*. We remark that if a co-clone is polynomially closed then the corresponding CSP problem is globally tractable [14]. As we will see later in Section 5, the opposite is not necessarily true for superpolynomially closed co-clones, however, since there exists superpolynomially closed co-clones where the corresponding CSP problem is trivially tractable.

We now turn to the problem of determining whether a co-clone is polynomially closed. First observe that to prove that a co-clone is polynomially closed it is sufficient to prove that there exists some polynomial p such that the weak base of the co-clone can p.p. define any n-ary relation with $p(n)$ variables.

Say that a plain base Γ_p of a co-clone $\mathrm{Inv}(\mathsf{C})$ is a *polynomial base* if there exists a polynomial p, such that every n-ary $R \in \mathrm{Inv}(\mathsf{C})$ has a q.f.p.p. definition over $\Gamma_p^{(n)}$, with at most $p(n)$ constraints. Polynomial bases and polynomially closed co-clones are related by the following lemma, which states that a polynomial base for a co-clone implies polynomial closure, under some additional conditions.

Lemma 8. *Let* $\mathrm{Inv}(\mathsf{C})$ *be a co-clone with a weak base R_w. If there exists a polynomial, plain base Γ_p of* $\mathrm{Inv}(\mathsf{C})$, *and a polynomial p such that, for each $n \geq 1$, R_w can p.p. define every relation in $\Gamma_p^{(n)}$ with at most $p(n)$ existentially quantified variables, then* $\mathrm{Inv}(\mathsf{C})$ *is polynomially closed.*

Proof. Let $R \in \mathrm{Inv}(\mathsf{C})$ be an n-ary relation. By Theorem 5 and the assumption that Γ_p is a polynomial, plain base it follows that $\Gamma_p^{(n)}$ can q.f.p.p. define R using at most $g(n)$ constraints for some polynomial g. Let ϕ denote the q.f.p.p. definition of R in $\Gamma_p^{(n)}$. For every constraint C_i in ϕ we then replace C_i with its p.p. definition in $\{R_w, \mathrm{Eq}\}$. Let the resulting formula be ϕ'. Since ϕ had $g(n)$ constraints and each constraint in ϕ' introduced at most $p(n)$ new existentially quantified variables, the total number of variables in ϕ' is $g(n) \cdot p(n)$, clearly polynomial with respect to n. Hence, $\mathrm{Inv}(\mathsf{C})$ is polynomially closed. $\qquad\square$

It is not difficult to see that this condition is satisfied whenever a co-clone has a finite plain base.

Lemma 9. *If* $\mathrm{Inv}(\mathsf{C})$ *has a finite plain base then* $\mathrm{Inv}(\mathsf{C})$ *is polynomially closed.*

Proof. Assume that $\mathrm{Inv}(\mathsf{C})$ has a plain base Γ_p of finite cardinality and let R_w denote a weak base of $\mathrm{Inv}(\mathsf{C})$. Observe that Γ_p is trivially a polynomial base. Since Γ_p is finite there exists a polynomial p such that R_w can p.p. define $\Gamma_p^{(n)}$ for every $n \geq 1$ with $p(n)$ variables. To see this, simply take the number of existentially quantified variables of the relation requiring the largest number of quantified variables in the p.p. definition in Γ_p. Such a relation must exist since Γ_p is finite. The result then follows from Lemma 8. $\qquad\square$

For the Boolean domain one can simply consult Table 2 to see which co-clones have finite plain bases. It is however possible to give a general characterization of the co-clones admitting finite plain bases. Let D be an arbitrary finite domain. A k-*ary near-unanimity* (NU) operation on D for $k \geq 3$ is an operation $f : D^k \to D$ such that $f(x, y, \ldots, y) = f(y, x, \ldots, y) = \ldots = f(y, y, \ldots, x) = y$ for all $x, y \in D$ (we may note that this is a special case of a k-*edge operation*, used in Section 5).

Theorem 10. *Let* $\mathrm{Inv}(\mathsf{C})$ *be a co-clone over a finite domain* D *such that* $\mathrm{Pol}(\mathrm{Inv}(\mathsf{C}))$ *contains a* k-*ary NU operation for some* $k \geq 3$. *Then* $\mathrm{Inv}(\mathsf{C})$ *is polynomially closed.*

Proof. We recall some definitions from Jeavons et al. [16]. Let $R \subseteq D^n$ be a relation and $I = \{i_1, \ldots, i_d\}$ a set of indices, $1 \leq i_1 < \ldots < i_d \leq n$. The *projection of* R *onto* I is the relation $\pi_I(R) = \{(t[i_1], \ldots, t[i_d]) \mid t \in R\}$. A relation $R \subseteq D^n$ over D is r-*decomposable* if it is equivalent to the conjunction of all its projections of arity at most r, i.e., for every $t \in (D^n \setminus R)$ there is a set $I = \{i_1, \ldots, i_d\}$ as above such that $(t[i_1], \ldots, t[i_d]) \notin \pi_I(R)$. It is known that any relation preserved by a k-ary, $k \geq 3$, NU operation is $(k-1)$-decomposable [16].

Now let $R \in \mathrm{Inv}(\mathsf{C})$, of arity n. Observe that $\pi_I(R)$ can be defined using existential quantification over R, hence $\pi_I(R) \in \langle \{R\} \rangle$ for every set of indices I. Also note that $\bigwedge_{I : |I| < k} \pi_I(R)$ is a q.f.p.p. definition of R. Hence the set of all relations $R' \in \mathrm{Inv}(\mathsf{C})$ of arity at most $k - 1$ is a plain base of $\mathrm{Inv}(\mathsf{C})$. Clearly, this is a finite set (since $|D|$ is finite). Thus $\mathrm{Inv}(\mathsf{C})$ is polynomially closed by Lemma 9. $\qquad\square$

Observe, however, that Lemma 9 or Theorem 10 are not applicable for $\mathrm{Inv}(\mathsf{L})$, $\mathrm{Inv}(\mathsf{L}_0)$, $\mathrm{Inv}(\mathsf{L}_1)$, $\mathrm{Inv}(\mathsf{L}_3)$ and $\mathrm{Inv}(\mathsf{L}_2)$ since they do not admit finite plain bases. Fortunately, it is rather straightforward to prove that these co-clones admit polynomial bases, since the included relations can be viewed as linear equations over the field $\mathrm{GF}(2)$.

Lemma 11. $\mathrm{Inv}(\mathsf{L}), \mathrm{Inv}(\mathsf{L}_0), \mathrm{Inv}(\mathsf{L}_1), \mathrm{Inv}(\mathsf{L}_3)$ *and* $\mathrm{Inv}(\mathsf{L}_2)$ *have polynomial, plain bases.*

Proof. We only consider $\mathrm{Inv}(\mathsf{L}_2)$ since the other cases follow through similar arguments. Every n-ary relation $R \in \mathrm{Inv}(\mathsf{L}_2)$ can according to Theorem 6 be expressed by a $\Gamma_p^{(n)}$ formula ϕ with m constraints, where Γ_p is the plain base of $\mathrm{Inv}(\mathsf{L}_2)$ in Table 2. Thus every constraint C_i in ϕ is of the form $(x_{i_1} \oplus \ldots \oplus x_{i_n}) = c_i$, where $c_i \in \{0, 1\}$. Create an $m \times (n+1)$-matrix M such that each entry $r_{i,j}$, $1 \leq j \leq n$, is equal to 1 if the variable x_j is included in the constraint C_i, and 0 otherwise. The entry $r_{i,n+1}$ is equal to the constant c_i in C_i. Then it is not hard to verify that if the row r_{i+1} is linearly dependent on $r_1, \ldots, r_i$ then $C_1, \ldots C_i$ entails C_{i+1} in any satisfying assignment. Hence we only need to keep the rows that are linearly independent, which gives the bound $\min(n+1, m)$ on the number of constraints. $\qquad\square$

Lemma 12. $\mathrm{Inv}(\mathsf{L})$, $\mathrm{Inv}(\mathsf{L}_0)$, $\mathrm{Inv}(\mathsf{L}_1)$, $\mathrm{Inv}(\mathsf{L}_3)$ *and* $\mathrm{Inv}(\mathsf{L}_2)$ *are polynomially closed.*

Proof. We only present the proof of $\mathrm{Inv}(\mathsf{L}_2)$ since the other co-clones follow through entirely analogous arguments. Let Γ_p and R_w be the plain and weak base of $\mathrm{Inv}(\mathsf{L}_2)$ from Table 2, respectively. Since $\mathrm{Inv}(\mathsf{L}_2)$ has a polynomial base by Lemma 11 all we need to prove is that there exists a polynomial p such that R_w can p.p. define $\Gamma_p^{(n)}$ using at most $p(n)$ existentially quantified variables. We first and most crucially show that $\Gamma_p^{(n)}$ can p.p. define $\Gamma_p^{(n+1)}$ with only one extra variable, for every $n \geq 3$, with the definition $(x_1 \oplus \ldots \oplus x_{n+1} = c) \equiv \exists x.(x_1 \oplus \ldots \oplus x_{n-1} \oplus x = c) \wedge (x_n \oplus x_{n+1} \oplus x = 0)$. In addition to one quantified variable this requires one extra $\Gamma_p^{(3)}$-constraint. Hence if $3 \leq n \leq n'$ then $\Gamma_p^{(n)}$ can p.p. define every relation in $\Gamma_p^{(n')}$ with $O(n' - n)$ variables and $n' - n$ additional $\Gamma_p^{(3)}$-constraints. By this it first follows that $\Gamma_p^{(3)}$ can p.p. define any relation in $\Gamma_p^{(n)}$ with at most $n - 3$ variables and $n - 2$ constraints. The weak base R_w can then p.p. define $\Gamma_p^{(3)}$ with a fixed number of variables since the arity of each relation is bounded, for example we have that $(x_1 \oplus x_2 \oplus x_3 = 0) \equiv \exists y_1, y_2, y_3, c_0, c_1.R_w(x_1, x_2, x_3, y_1, y_2, y_3, c_0, c_1)$ and $(x_1 \oplus x_2 \oplus x_3 = 1) \equiv \exists y_1, y_2, y_3, c_0, c_1.R_w(y_1, y_2, y_3, x_1, x_2, x_3, c_0, c_1)$. Put together this implies that R_w can p.p. define any $\Gamma_p^{(n)}$ with $O(n)$ existentially quantified variables, and by Lemma 8 that $\mathrm{Inv}(\mathsf{L}_2)$ is polynomially closed. $\qquad\square$

Combining all results so far in this section, we obtain the following characterization of the polynomially closed Boolean co-clones of finite order.

Theorem 13. *If* $\mathrm{Inv}(\mathsf{C}) \subseteq \mathrm{Inv}(\mathsf{X})$ *for some* $X \in \{\mathsf{L}_2, \mathsf{D}_2\} \cup \{\mathsf{S}_{00}^n, \mathsf{S}_{10}^n \mid n \geq 2\}$ *then* $\mathrm{Inv}(\mathsf{C})$ *is polynomially closed.*

4 Polynomially closed co-clones of infinite order

So far we have only been concerned with polynomially closed co-clones of finite order. For co-clones of infinite order, we cannot use any of the machinery introduced in Section 3. In particular, Lemma 9 breaks down since there by definition cannot exist a finite plain base of a co-clone of infinite order. In this section we give a general result to obtain weak bases of co-clones of infinite order, and leverage this result to show that the eight Boolean co-clones of infinite order in Figure 1 are polynomially closed.

Theorem 14. *Let* $\mathrm{Inv}(C)$ *be a co-clone of infinite order over a finite domain and let* $\mathrm{Inv}(C_1), \mathrm{Inv}(C_2), \ldots$ *be an infinite chain of co-clones of finite order such that* $C = \bigcup_{i=1}^{\infty} \mathrm{Inv}(C_i)$. *Let* R_{C_i} *denote the weak base of each* $\mathrm{Inv}(C_i)$. *Assume that* $R_{C_i} \in \langle R_{C_{i+1}} \rangle_{\not\exists}$ *for each* $i \geq 1$. *Then the weak base of* $\mathrm{Inv}(C)$ *is the language* $\Gamma_C = \{R_{C_i} \mid i \geq 1\}$.

Proof. First observe that each $\mathrm{Inv}(C_i)$ does indeed have a finite weak base since by assumption they are of finite order. To prove that Γ_C is a weak base of $\mathrm{Inv}(C)$ we must prove that it is a base of $\mathrm{Inv}(C)$ and that $\Gamma_C \subseteq \langle \Gamma \rangle_{\not\exists}$ for each base Γ of $\mathrm{Inv}(C)$. It is easy to see that Γ_C is a base of $\mathrm{Inv}(C)$ since $\mathrm{Inv}(C_i) = \langle \{R_{C_i}\} \rangle \subseteq \langle \Gamma_C \rangle$ for every $i \geq 1$ for some $R_{C_i} \in \Gamma_C$. Let Γ be a constraint language such that $\langle \Gamma \rangle = \mathrm{Inv}(C)$. Observe that Γ must be infinite, and that there for every $R \in \Gamma$ exists some m such that $R \in \mathrm{Inv}(C_m)$, since $\langle \Gamma \rangle = \mathrm{Inv}(C) = \bigcup_{i=1}^{\infty} \mathrm{Inv}(C_i)$.

We must prove that $\Gamma_C \subseteq \langle \Gamma \rangle_{\not\exists}$. Let $R \in \Gamma_C$ be an n-ary relation. Then there exists an m such that R is the weak base of $\mathrm{Inv}(C_m)$. We prove that there exists $\Delta \subseteq \Gamma$ such that $\langle \Delta \rangle = \mathrm{Inv}(C_{m'})$ for some $m' \geq m$, since this implies that $R \in \langle \Delta \rangle_{\not\exists} \subseteq \langle \Gamma \rangle_{\not\exists}$, by the original assumption. Assume for contradiction that no such set Δ exists. But this implies that there exists some $k < m$ such that $\langle \Gamma \rangle = \mathrm{Inv}(C_k)$, which is clearly impossible since C_k is of finite order. Hence, there exists $\Delta \subseteq \Gamma$ such that $R \in \langle \Delta \rangle_{\not\exists}$. Since R was choosen arbitrarily, this in turn implies that $\Gamma_C \subseteq \langle \Gamma \rangle_{\not\exists}$, and that Γ_C is a weak base of $\mathrm{Inv}(C)$. $\square$

Table 3. Weak bases of all Boolean co-clones of infinite order.

Co-clone	Weak base
$\mathrm{Inv}(\mathsf{S}_0)$	$\{\mathrm{OR}(x_1, \ldots, x_n) \wedge \mathrm{T}(c_1) \mid n \geq 2\}$
$\mathrm{Inv}(\mathsf{S}_{02})$	$\{\mathrm{OR}(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1) \mid n \geq 2\}$
$\mathrm{Inv}(\mathsf{S}_{01})$	$\{\mathrm{OR}(x_1, \ldots, x_n) \wedge (x \rightarrow x_1 \cdots x_n) \wedge \mathrm{T}(c_1) \mid n \geq 2\}$
$\mathrm{Inv}(\mathsf{S}_{00})$	$\{\mathrm{OR}(x_1, \ldots, x_n) \wedge (x \rightarrow x_1 \cdots x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1) \mid n \geq 2\}$
$\mathrm{Inv}(\mathsf{S}_1)$	$\{\mathrm{NAND}(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0) \mid n \geq 2\}$
$\mathrm{Inv}(\mathsf{S}_{12})$	$\{\mathrm{NAND}(x_1, \ldots, x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1) \mid n \geq 2\}$
$\mathrm{Inv}(\mathsf{S}_{11})$	$\{\mathrm{NAND}(x_1, \ldots, x_n) \wedge \neg(x \rightarrow x_1 \cdots x_n) \wedge \mathrm{T}(c_1) \mid n \geq 2\}$
$\mathrm{Inv}(\mathsf{S}_{10})$	$\{\mathrm{NAND}(x_1, \ldots, x_n) \wedge \neg(x \rightarrow x_1 \cdots x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1) \mid n \geq 2\}$

We remark that since $R_{C_i} \in \langle R_{C_{i+1}} \rangle_{\not\exists}$ for every $i \geq 1$ we can in fact remove any finite number of relations from the weak base Γ_C and still obtain a weak base of $\mathrm{Inv}(C)$. According to Theorem 14 all that is needed to obtain weak bases for the eight co-clones of infinite order in the Boolean co-clone lattice, is to show that the condition $R_{C_i} \in \langle R_{C_{i+1}} \rangle_{\not\exists}$ is satisfied for every $C_i \in \{\mathsf{S}_0^i, \mathsf{S}_{02}^i, \mathsf{S}_{01}^i, \mathsf{S}_{00}^i, \mathsf{S}_1^i, \mathsf{S}_{12}^i, \mathsf{S}_{11}^i, \mathsf{S}_{10}^i\}$. We only consider the case $\mathrm{Inv}(\mathsf{S}_{00}^n)$ since the remaining proofs are entirely analogous. Hence, we need to show that the weak base of $\mathrm{Inv}(\mathsf{S}_{00}^n)$ can q.f.p.p. define the weak base of $\mathrm{Inv}(\mathsf{S}_{00}^{n-1})$ for each $n \geq 3$. For $n \geq 2$ let $R_{\mathsf{S}_{00}}^n(x_1, \ldots, x_n, x, c_0, c_1) \equiv \mathrm{OR}(x_1, \ldots, x_n) \wedge (x \rightarrow x_1 \cdots x_n) \wedge \mathrm{F}(c_0) \wedge \mathrm{T}(c_1)$. Then we can q.f.p.p. define $R_{\mathsf{S}_{00}}^n$ with $R_{\mathsf{S}_{00}}^{n+1}$ using the definition

$$R_{\mathsf{S}_{00}}^n(x_1, \ldots, x_n, x, c_0, c_1) \equiv R_{\mathsf{S}_{00}}^{n+1}(x_1, \ldots, x_n, x_n, x, c_0, c_1).$$

Hence, we obtain the following corollary, summarized in Table 3.

Corollary 15. *The following statements hold.*

- $\{\text{OR}(x_1, \ldots, x_n) \wedge \text{T}(c_1) \mid n \geq 2\}$ *is a weak base of* $\text{Inv}(\mathsf{S}_0)$,
- $\{\text{OR}(x_1, \ldots, x_n) \wedge \text{F}(c_0) \wedge \text{T}(c_1) \mid n \geq 2\}$ *is a weak base of* $\text{Inv}(\mathsf{S}_{02})$,
- $\{\text{OR}(x_1, \ldots, x_n) \wedge (x \to x_1 \cdots x_n) \wedge \text{T}(c_1) \mid n \geq 2\}$ *is a weak base of* $\text{Inv}(\mathsf{S}_{01})$,
- $\{\text{OR}(x_1, \ldots, x_n) \wedge (x \to x_1 \cdots x_n) \wedge \text{F}(c_0) \wedge \text{T}(c_1) \mid n \geq 2\}$ *is a weak base of* $\text{Inv}(\mathsf{S}_{00})$,
- $\{\text{NAND}(x_1, \ldots, x_n) \wedge \text{F}(c_0) \mid n \geq 2\}$ *is a weak base of* $\text{Inv}(\mathsf{S}_1)$,
- $\{\text{NAND}(x_1, \ldots, x_n) \wedge \text{F}(c_0) \wedge \text{T}(c_1) \mid n \geq 2\}$ *is a weak base of* $\text{Inv}(\mathsf{S}_{12})$,
- $\{\text{NAND}(x_1, \ldots, x_n) \wedge (x \to x_1 \cdots x_n) \wedge \text{T}(c_1) \mid n \geq 2\}$ *is a weak base of* $\text{Inv}(\mathsf{S}_{11})$,
- $\{\text{NAND}(x_1, \ldots, x_n) \wedge (x \to x_1 \cdots x_n) \wedge \text{F}(c_0) \wedge \text{T}(c_1) \mid n \geq 2\}$ *is a weak base of* $\text{Inv}(\mathsf{S}_{10})$.

We are now in a position to prove that all Boolean co-clones of infinite order are polynomially closed. For the proof we use the fact that relations in $\text{Inv}(\mathsf{S}_0)$, $\text{Inv}(\mathsf{S}_1)$, $\text{Inv}(\mathsf{S}_{01})$ and $\text{Inv}(\mathsf{S}_{00})$ can be expressed by a particularly simple form of Boolean formula. Before we can formally state this result we need a few additional preliminaries. If $\varphi = C_1 \wedge \ldots \wedge C_m$ is a Boolean formula with m clauses we say that C_i is a *prime implicate* of φ if φ does not entail any proper subclause of C_i. A formula φ is said to be *prime* if all of its clauses are prime implicates. Obviously any finite Boolean relation is representable by a prime formula. If R is an n-ary Boolean relation we can therefore prove that $R \in \langle \Gamma \rangle_{\not\exists}$ by showing that $R(x_1, \ldots, x_n)$ can be expressed as a conjunction $\varphi_1(\mathbf{y_1}) \wedge \ldots \wedge \varphi_k(\mathbf{y_k})$, where each $\mathbf{y_i}$ is a tuple of variables over $x_1, \ldots, x_n$, and each φ_i is a prime formula representation of a relation in Γ. This is advantageous since relations in $\text{Inv}(\mathsf{S}_0^n)$, $\text{Inv}(\mathsf{S}_{02}^n)$, $\text{Inv}(\mathsf{S}_{01}^n)$, $\text{Inv}(\mathsf{S}_{00}^n)$, $\text{Inv}(\mathsf{S}_1^n)$, $\text{Inv}(\mathsf{S}_{12}^n)$, $\text{Inv}(\mathsf{S}_{11}^n)$ and $\text{Inv}(\mathsf{S}_{10}^n)$ are representable by prime *implicative hitting set-bounded* (IHSB) formulas [9]. We let IHSB_+^n be the set of formulas of the form $(x_1 \vee \ldots \vee x_m), 1 \leq m \leq n, (\neg x_1), (\neg x_1 \vee x_2)$, and dually for IHSB_-^n.

Theorem 16. $\text{Inv}(\mathsf{S}_0)$, $\text{Inv}(\mathsf{S}_{02})$, $\text{Inv}(\mathsf{S}_{01})$, $\text{Inv}(\mathsf{S}_{00})$, $\text{Inv}(\mathsf{S}_1)$, $\text{Inv}(\mathsf{S}_{12})$, $\text{Inv}(\mathsf{S}_{11})$ *and* $\text{Inv}(\mathsf{S}_{10})$ *are polynomially closed.*

Proof. We only consider $\text{Inv}(\mathsf{S}_{00})$ since the other cases follow through similar arguments. Let $\Gamma_{\mathsf{S}_{00}} = \{R_{\mathsf{S}_{00}}^i \mid i \geq 2\}$ denote the weak base of $\text{Inv}(\mathsf{S}_{00})$ from Corollary 15. We must prove that there exists a polynomial p such that $\Gamma_{\mathsf{S}_{00}}$ can p.p. define any n-ary $R \in \text{Inv}(\mathsf{S}_{00})$ using at most $p(n)$ existentially quantified variables. Since $R \in \text{Inv}(\mathsf{S}_{00})$ it is easily seen that there exists some $n' \geq 2$ such that $R \in \text{Inv}(\mathsf{S}_{00}^{n'})$.

Hence, R can be written as a prime $\text{IHSB}_+^{n'}$ formula ϕ over $x_1, \ldots, x_n$ [9], and we need to show that it is possible to p.p. define this formula without requiring more than a polynomial number of existentially quantified variables. There are a few different cases to consider depending on the clauses of ϕ. Let c_0 and c_1 be two fresh variables distinct from $x_1, \ldots, x_n$. First, we implement every clause in ϕ of the form $(x_{i_1} \vee \ldots \vee x_{i_j})$ for some $j \leq n'$ with the constraint $R_{\mathsf{S}_{00}}^j(x_{i_1}, \ldots, x_{i_j}, c_0, c_0, c_1)$. Second, we implement every clause of the form $(\neg x_i)$ as $R_{\mathsf{S}_{00}}^2(c_1, c_1, c_1, x_i, c_1)$. Third, we implement every clause of the form $(\neg x_{i_1} \vee x_{i_2})$ as $R_{\mathsf{S}_{00}}^2(x_{i_2}, c_1, x_{i_1}, c_0, c_1)$. Let ϕ' be the $\Gamma_{\mathsf{S}_{00}}$-formula resulting from replacing every clause in ϕ in the above manner. We see that $R(x_1, \ldots, x_n) \equiv \exists c_0 \exists c_1 . \phi'$, and since we in total only require 2 existentially quantified variables, it follows that $\text{Inv}(\mathsf{S}_{00})$ is polynomially closed. $\qquad\square$

5 Superpolynomially closed co-clones

From Section 3 and Section 4 we now have straightforward, necessary conditions for verifying whether a given co-clone is polynomially closed. We now turn to the problem of determining whether a co-clone is not polynomially closed, i.e., superpolynomially closed. We show that this question is related to counting the number of n-ary relations in a co-clone — a problem that has attracted significant attention in universal algebra and conceptual learning problems [3, 14]. Before we can present this result, we for every finite domain D, introduce a particular constraint language Γ_D, which will turn out to be a plain base of Rel_D. The language Γ_D is defined as

$$\Gamma_D = \{R \mid n \geq 1, t \in D^n, R = D^n \setminus \{t\}\}.$$

In other words each n-ary relation in Γ_D contains all n-ary tuples over D except one. Observe that $\Gamma_{\{0,1\}}$ is equivalent to the plain base of BR in Table 2.

Lemma 17. *For any finite domain D the language Γ_D is a plain base of Rel_D.*

Proof. We must prove that $\langle \Gamma_D \rangle_{\nexists} = \mathrm{Rel}_D$, i.e. that we can q.f.p.p. define all relations over D. Hence, let $R \in \mathrm{Rel}_D$ be an n-ary relation. For every $t \in D^n \setminus R$ we let $R_t \in \Gamma_D$ denote the unique relation satisfying $R_t = D^n \setminus \{t\}$. Hence, a constraint of the form $R_t(x_1, \ldots, x_n)$ implies that $x_1, \ldots, x_n$ can take any value except for $t[1], \ldots, t[n]$. With this observation it is then easy to see that we can define R with the q.f.p.p. definition

$$R(x_1, \ldots, x_n) \equiv R_{t_1}(x_1, \ldots, x_n) \wedge \ldots \wedge R_{t_m}(x_1, \ldots, x_n),$$

where $\{t_1, \ldots, t_m\} = D^n \setminus R$. □

Also observe that $\langle \Gamma_D^{(n)} \rangle_{\nexists} \subseteq \langle \Gamma_D^{(n+1)} \rangle_{\nexists}$ for each $n \geq 1$. To see this, simply note that there for every n-ary relation R exists a $(n+1)$-ary relation R' defined as $R'(x_1, \ldots, x_n, x_{n+1}) \equiv R(x_1, \ldots, x_n) \wedge \mathrm{Eq}(x_{n+1}, x_{n+1})$, which is equivalent with respect to q.f.p.p. definitions. We will now prove that if a co-clone $\mathrm{Inv}(\mathsf{C})$ contains a sufficiently large number of n-ary relations, then for every polynomial p there will exist some n-ary relation in $\mathrm{Inv}(\mathsf{C})$ that Γ_D cannot p.p. define using only $p(n)$ existentially quantified variables. To make this counting argument more precise we, given a constraint language Γ, first let $\Gamma^{=n} = \{R \mid R \in \Gamma, \mathrm{ar}(R) = n\}$, and then define the function s_Γ as

$$s_\Gamma(n) = \log_2(|\{R \mid R \in \Gamma, \mathrm{ar}(R) = n\}|).$$

With this notation we see that $s_{\mathrm{Inv}(\mathsf{C})}(n)$ denotes the exponent of the number of n-ary relations in the co-clone $\mathrm{Inv}(\mathsf{C})$, and obtain the following lemma.

Lemma 18. *Let $\mathrm{Inv}(\mathsf{C})$ be a co-clone of finite order over a finite domain D. If $\mathrm{Inv}(\mathsf{C})$ is polynomially closed, then $s_{\mathrm{Inv}(\mathsf{C})}(n) \leq g(n)$ for some polynomial g.*

Proof. Let Γ be a finite base of $\mathrm{Inv}(\mathsf{C})$ and let R be the relation with the highest arity k in Γ. We make a few observations before the proof: first, $\langle \Gamma \rangle_{\nexists} \subseteq \langle \Gamma_D^{(=k)} \rangle_{\nexists}$; second, if some $R' \notin \langle \Gamma_D^{(=k)} \rangle_{\nexists}$ then $R' \notin \Gamma$. This also implies that if Γ can p.p. define some n-ary relation R with $p(n)$ existentially quantified then the same is true for $\Gamma_D^{(=k)}$. By contraposition this also implies that if $\Gamma_D^{(=k)}$ cannot p.p. define some n-ary relation R with $p(n)$ variables then neither can Γ. It is not too difficult to see that the number of q.f.p.p. definitions with $\Gamma_D^{(=k)}$ over n variables is bounded by $2^{|D|^k n^k}$, since (1) $\Gamma_D^{(=k)}$ contains $|D|^k$ relations and (2) for each relation in $\Gamma_D^{(=k)}$ one can form at most n^k distinct constraints. Since $\mathrm{Inv}(\mathsf{C})$ is polynomially closed, we are allowed to introduce at most $p(n)$ existentially quantified variables to define any n-ary relation, hence, the number of definable relations is at most $2^{|D|^k (p(n)+n)^k}$, which implies that $s_{\mathrm{Inv}(\mathsf{C})}(n) \leq |D|^k (p(n) + n)^k$ and that there exists a polynomial g such that $s_{\mathrm{Inv}(\mathsf{C})}(n) \leq g(n)$. □

Since the number of n-ary relations over a finite domain D is $2^{|D|^n}$ it immediately follows that Rel_D is superpolynomially closed. To handle the other cases where it is not so obvious how to count the number of n-ary relations we utilize a result from Berman et al. [3]. Before we can present their result, we need a few additional preliminaries. If Γ is a constraint language over D the algebra $\mathbf{A}_\Gamma = (D, \mathrm{Pol}(\Gamma))$ is said to have *few subpowers* if $s_{\langle \Gamma \rangle}(n) \in O(n^k)$ for some polynomial $k \geq 1$, and to have *many subpowers* if $c^n \in O(s_{\langle \Gamma \rangle}(n))$ for some real number $c > 1$. A *k-edge* operation over D, $k \geq 2$, is a $(k+1)$-ary operation f satisfying the k identities

- $f(x, x, y, y, y, \ldots, y, y) = y,$
- $f(x, y, x, y, y, \ldots, y, y) = y,$
- $f(y, y, y, x, y, \ldots, y, y) = y,$
- $f(y, y, y, y, x, \ldots, y, y) = y,$
  ⋮
- $f(y, y, y, y, y, \ldots, x, y) = y,$
- $f(y, y, y, y, y, \ldots, y, x) = y.$

We then have the following useful theorem from Berman et al. [3], restated in terms of our terminology of clones and co-clones

Theorem 19. *[3] Let Γ be a constraint language over a finite domain D. If $\mathrm{Pol}(\Gamma)$ does not contain a k-edge operation for any $k \geq 2$ then (1) the algebra $(D, \mathrm{Pol}(\Gamma))$ has many subpowers and (2) $s_{\langle \Gamma \rangle}(n) \notin O(n^l)$ for any $l \geq 0$.*

Hence, if $(D, \mathrm{Pol}(\Gamma))$ has many subpowers, then, intuitively, $\langle \Gamma \rangle$ contains too many relations for it to be polynomially closed. Combining Lemma 18 and Theorem 19 we obtain the following classification of the super-polynomially closed co-clones.

Theorem 20. *Let $\mathrm{Inv}(C)$ be a co-clone of finite order over a finite domain D. If C does not contain a k-edge operation for any $k \geq 2$ then $\mathrm{Inv}(C)$ is superpolynomially closed.*

With the help of Table 1 one can verify that any Boolean co-clone of finite order above or equal to $\mathrm{Inv}(\mathsf{V})$, $\mathrm{Inv}(\mathsf{E})$, or $\mathrm{Inv}(\mathsf{N})$ in Figure 1, fullfil this property.

Theorem 21. *If $\mathrm{Inv}(\mathsf{C}) \supseteq \mathrm{Inv}(\mathsf{X})$ for some $\mathrm{Inv}(\mathsf{X}) \in \{\mathrm{Inv}(\mathsf{V}), \mathrm{Inv}(\mathsf{E}), \mathrm{Inv}(\mathsf{N})\}$ then $\mathrm{Inv}(\mathsf{C})$ is superpolynomially closed.*

Due to the close relationship between a polynomially closed co-clone and the existence of a polynomial, plain base, one might suspect that superpolynomially closed co-clones are unlikely to admit such polynomial bases. This can in fact be proven by a straightforward counting argument, using the bounds from Theorem 19 on the number of n-ary relations in these co-clones.

Theorem 22. *Let $\mathrm{Inv}(\mathsf{C})$ be a superpolynomially closed co-clone over a finite domain D such that there exists a plain base Γ_p of $\mathrm{Inv}(\mathsf{C})$ satisfying $\Gamma_p^{(n)} \leq 2^{p(n)}$ for some polynomial p. Then Γ_p is not a polynomial, plain base of $\mathrm{Inv}(\mathsf{C})$.*

Proof. Assume that $\mathrm{Inv}(\mathsf{C})$ has a polynomial base with respect to a polynomial c. We show the theorem with a counting argument, using the results of Section 5. First, recall from Lemma 18 that $s_{\mathrm{Inv}(\mathsf{C})}(n)$ cannot be bounded by a polynomial function since $\mathrm{Inv}(\mathsf{C})$ is superpolynomially closed. In other words it cannot hold that $|\{R \in \mathrm{Inv}(\mathcal{C}) \mid \mathrm{ar}(R) = n\}| \leq 2^{p(n)}$ for some polynomial p.

Now observe that for each $R \in \Gamma_p^{(n)}$, there are at most n^n different possible constraints one can form with R; thus the number of different possible constraints overall is bounded by $|\Gamma_p^{(n)}| \cdot n^n$. The number of possible formulas with at most $c(n)$ constraints is then bounded by $(|\Gamma_p^{(n)}| \cdot n^n)^{c(n)} \leq (2^{p(n)} \cdot n^n)^{c(n)} \leq 2^{q(n)}$ for a polynomial $q(n)$, which implies that $s_{\mathrm{Inv}(\mathsf{C})}(n) \leq q(n)$, contradicting the original assumption. $\square$

Using Table 2 we see that each Boolean plain base Γ_p contains at most polynomially many n-ary relations. Hence, we obtain the following theorem for Boolean co-clones.

Theorem 23. *If $\mathrm{Inv}(\mathsf{C}) \supseteq \mathrm{Inv}(\mathsf{X})$ for some $\mathrm{Inv}(\mathsf{X}) \in \{\mathrm{Inv}(\mathsf{V}), \mathrm{Inv}(\mathsf{E}), \mathrm{Inv}(\mathsf{N})\}$ then the plain base of $\mathrm{Inv}(\mathsf{C})$ in Table 2 is not a polynomial, plain base.*

Thus a Boolean co-clone of finite order has a polynomial, plain base in Table 2 if and only if it is polynomially closed. In conjunction, the results of Section 3 and Section 5 therefore imply the following corollary.

Corollary 24. *Let $\langle \Gamma \rangle$ be a Boolean co-clone of finite order. Then the following statements are equivalent.*

- *$\langle \Gamma \rangle$ is polynomially closed.*
- *$\langle \Gamma \rangle$ has a polynomial, plain base in Table 2.*
- *The algebra $(\{0, 1\}, \mathrm{Pol}(\Gamma))$ has few subpowers.*
- *There exists a polynomial p such that the number of n-ary relations in $\langle \Gamma \rangle$ is not larger than $2^{p(n)}$.*

For arbitrary finite domains our result do not form a sharp dichotomy. Combining Theorem 10 and Theorem 20, we however obtain the following corollary.

Corollary 25. *Let $\langle \Gamma \rangle$ be a co-clone of finite order over a finite domain. Then the following statements hold.*

- *If $\mathrm{Pol}(\Gamma)$ does not contain a k-edge operation for any $k \geq 2$ then $\langle \Gamma \rangle$ is superpolynomially closed.*
- *If $\mathrm{Pol}(\Gamma)$ contains a k-ary near-unanimity operation for some $k \geq 3$ then $\langle \Gamma \rangle$ is polynomially closed.*

For co-clones of infinite order this situation differs drastically, as evident in Section 4, since even in the Boolean domain it can be the case that a co-clone of infinite order is polynomially closed even if the corresponding algebra has many subpowers.

6 Strong Partial Clones of Finite and Infinite Order

So far we have been interested in obtaining conditions for separating polynomially closed co-clones from superpolynomially closed co-clones, and obtained a complete dichotomy theorem for the Boolean domain. Since we for polynomially closed co-clones can define all relations in the co-clone with a comparably few number of existentially quantified variables, one might conjecture that a strong partial clone $\mathrm{pPol}(\Gamma)$ has a more complex structure if $\langle \Gamma \rangle$ is superpolynomially closed. To make this intuition a bit more precise, given a co-clone $\mathrm{Inv}(C)$ and a base Γ of $\mathrm{Inv}(C)$, we are interested in determining when $\mathrm{pPol}(\Gamma)$ is of infinite order and when it is of finite order. Hence, we make the following definition (recall from Section 2.2 that $\mathcal{I}(C)$ denotes the interval of all strong partial clones where the total component equals C).

Definition 26. *Let C be a clone over a finite domain. We say that $\mathcal{I}(C)$ is* finitely generated *if every $\mathrm{pPol}(\Delta) \in \mathcal{I}(C)$ is of finite order and that $\mathcal{I}(C)$ is* infinitely generated *if $\mathrm{pPol}(\Delta)$ is of infinite order for every $\mathrm{pPol}(\Delta) \in \mathcal{I}(C)$.*

A few basic observations are in place. First, determining whether a partial clone is of finite or infinite order is a problem that has attracted significant attention in the literature, see e.g. [7, 8, 13]. However, observe that the authors in this case study partial clones that are not necessarily strong, and that a partial clone P might be of infinite order even though the smallest strong partial clone containing P is of finite order. Second, if C is a clone of finite order then $\mathrm{pPol}(\mathrm{Inv}(C))$ is of finite order. This implies that as long as C is of finite order, $\mathcal{I}(C)$ will contain at least one strong partial clone of finite order. Hence, in general, intervals of the form $\mathcal{I}(C)$ may contain strong partial clones of both finite and infinite order. If we restrict ourself to strong partial clones $\mathrm{pPol}(\Gamma)$ where Γ is finite, this phenomenon is not as likely to occur, however. We thus make the following definition as well.

Definition 27. *Let C be a clone over a finite domain D. The* finite interval *of C, $\mathcal{I}_{\mathrm{fin}}(C)$, is the set $\mathcal{I}_{\mathrm{fin}}(C) = \{\mathrm{pPol}(\Delta) \mid \Delta \subseteq \mathrm{Rel}_D,\ \Delta$ is finite, $C = \mathrm{Pol}(\Delta)\}$.*

In Section 6.1 we prove that the existence of finitely generated intervals is related to the question whether the cardinality of the interval is finite or uncountably infinite, and give examples of polynomially closed co-clones over arbitrary finite domains resulting in finitely generated intervals. Since not much is currently known about the lattice of strong partial clones over arbitrary finite domains, these results are necessarily inconclusive, and we cannot yet hope to provide a complete classification of finitely generated intervals. In Section 6.2 we study the opposite question: given a superpolynomially closed co-clone $\mathrm{Inv}(C)$, is $\mathcal{I}_{\mathrm{fin}}(C)$ infinitely generated? We give a general result and prove that $\mathcal{I}_{\mathrm{fin}}(C)$ is infinitely generated whenever C consists of essentially unary functions, i.e., if $C = [\{e_1, \ldots, e_k\}]$ for some unary functions $e_1, \ldots, e_k$. The results are summarized in Figure 2.
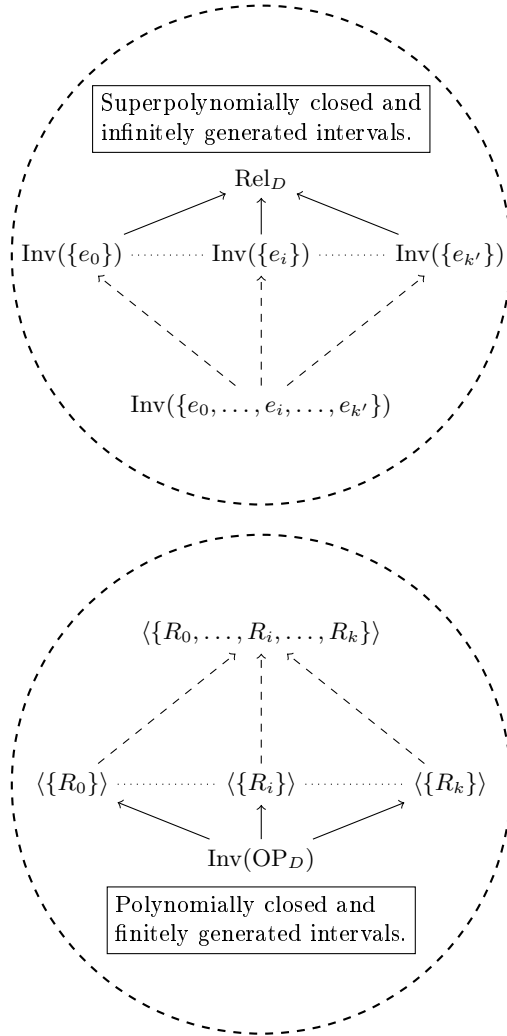
Fig. 2. An illustration of some fragments of the the structure of $\mathcal{I}(\Gamma)$ for Γ over an arbitrary finite domain $D = \{0, \ldots, i, \ldots, k\}$. For $a \in D$ let R_a denote the relation $\{(a)\}$. Let $e_0, \ldots, e_{k'}$ be an enumeration of the unary functions over D which are not projections. A directed arrow from node A to B means that $A \subset B$. A dashed arrow from node A to B means that there exists A' such that $A \subset A' \subset B$. Some inclusions have been omitted.

6.1 Strong Partial Clones of Finite Order

We first remark that if $\mathcal{I}(C)$ is finitely generated then the cardinality of $\mathcal{I}(C)$ is at most countably infinite. Hence, we get the following proposition.

Theorem 28. *Let C be clone such that $\mathcal{I}(C)$ is of uncountably infinite cardinality. Then $\mathcal{I}(C)$ is not finitely generated.*

On the other hand, if $\mathcal{I}(C)$ is finite and C is of finite order, it is not to difficult to see that $\mathcal{I}(C)$ must be finitely generated.

Lemma 29. *Let C be a clone of finite order over D such that $\mathcal{I}(C)$ is finite. Then $\mathcal{I}(C)$ is finitely generated.*

Proof. Let F denote an arbitrary finite base of C. Then $[F]_s$ is the least element in $\mathcal{I}(C)$. Assume, for contradiction, that there exists a strong partial clone $C' \in \mathcal{I}(C)$ of infinite order. Obviously $C' \supset [F]_s$. Let $f \in C' \setminus [F]_s$. Then it is easy to see that $C' \supset [F \cup \{f\}]_s \supset [F]_s$ since by assumption C' is of infinite order and cannot be generated by $F \cup \{f\}$. This procedure can be repeated arbitrarily many times, which contradicts the assumption that $\mathcal{I}(C)$ was finite. $\square$

Hence, the question of whether an interval is finitely generated or not is tightly connected to whether the interval is finite. In the Boolean domain it has been proven that $\mathcal{I}(\mathrm{Pol}(\Gamma))$ is finite if $\langle \Gamma \rangle$ is a subset of $\mathrm{Inv}(\mathsf{M}_2)$ or $\mathrm{Inv}(\mathsf{D}_1)$, and is of uncountably infinite cardinality otherwise [29]. Hence, we obtain the following proposition.

Proposition 30. *Let Γ be a Boolean constraint language. Then $\mathcal{I}(\mathrm{Pol}(\Gamma))$ is finitely generated if and only if $\langle \Gamma \rangle \subseteq \mathrm{Inv}(\mathsf{X})$ for $\mathsf{X} \in \{\mathsf{M}_2, \mathsf{D}_1\}$.*

In Schölzel [29] it is conjectured that intervals of the form $\mathcal{I}(\mathrm{Pol}(\Gamma))$ are either finite or uncountably infinite for arbitrary finite domains. Such a dichotomy theorem would therefore also answer the question which intervals are finitely generated and which are not. We remark that such a dichotomy theorem is likely very difficult to obtain, since not much is known of the structure of the lattice of strong partial clones over arbitrary finite domains. We give an examplary case of a simple kind of constraint language where the intervals of strong partial clones is always finite.

Given a finite domain $D = \{0, \ldots, k\}$ let R_i, $i \in D$, denote the unary, constant relation $\{(i)\}$. Say that a co-clone $\mathrm{Inv}(\mathsf{C})$ over D is *essentially constant* if there exists a set $\Gamma \subseteq \{R_0, \ldots, R_k\}$ such that $\langle \Gamma \rangle = \mathrm{Inv}(\mathsf{C})$. In other words $\mathrm{Inv}(\mathsf{C})$ can be generated from a finite set of constant relations.

Theorem 31. *Let $\mathrm{Inv}(C)$ be an essentially constant co-clone over some finite domain D. Then $\mathcal{I}(C)$ is finitely generated.*

Proof. Since $\mathrm{Inv}(C)$ is essentially constant there exists $\Gamma \subseteq \{R_0, \ldots, R_k\}$ such that $\langle \Gamma \rangle = \mathrm{Inv}(C)$. It is known both that $\mathrm{Pol}(\{R_0, \ldots, R_k\})$, the clone consisting of all idempotent functions over D, is finitely generated [25] and that there exists a finite number of (strong) partial clones containing $\mathrm{Pol}(\{R_0, \ldots, R_k\})$ [11]. From this it easily follows that $\mathrm{Pol}(\Gamma)$ is of finite order and that $\mathcal{I}(C)$ is finite. By applying Lemma 29 it follows that $\mathcal{I}(C)$ is finitely generated. $\square$

The reader might well ask why we do not attempt to prove a more general result than Theorem 31. The reason is that, currently, not much is known about the structure of finitely generated intervals of strong partial clones over arbitrary finite domains. For instance, it is not even known whether $\mathrm{pPol}(\{u_1, \ldots, u_n\})$, where each $u_i \subseteq D$, is of finite order. Moreover, it is known that the intersection of two strong partial clones of finite order can be of infinite order [8], which suggests that this problem is more difficult than one might believe at a first glance.

6.2 Strong Partial Clones of Infinite Order

We now turn to the problem of determining whether an interval $\mathcal{I}_{\mathrm{fin}}(C)$ is infinitely generated. We show that $\mathcal{I}_{\mathrm{fin}}(C)$ is always infinitely generated if C is an essentially unary clone over an arbitrary finite domain D. For finite Boolean constraint languages Γ this implies that if $\langle \Gamma \rangle \supseteq \mathrm{Inv}(\mathsf{N}_2)$, i.e. $\mathrm{CSP}(\Gamma)$ is NP-hard assuming P $\neq$ NP, then $\mathrm{pPol}(\Gamma)$ is of infinite order. For the proofs, we first need the following construction of a universal hash family, due to Alon et al.[2]. Given a natural number k we let $[k] = \{1, \ldots, k\}$.

Theorem 32 (Section 4 of [2]). *For any k and n, there is a family H of $2^{O(k)} \log n$ functions $h_i : [n] \mapsto [k]$ such that for every $S \subset [n]$ of size k there is a function in H that is injective on S.*

Note that the bound $O(k)$ has no hidden dependency on n. Hence, if k is a constant, then $2^{O(k)} \log n \in O(\log(n))$. The purpose of a universal hash family in this paper is to, given an n-ary relation R, create an n'-ary relation R' using the universal hash family such that $\mathrm{pPol}(R) \subseteq \mathrm{pPol}(R')$, and such that $n' = O(n)$. In the following definition we exploit the fact that any n-ary relation R can be viewed as an $|R| \times n$ matrix where each row corresponds to a tuple in R.

Definition 33. *Let R be a relation over D, $|R| = m$, let $r \geq 1$ and let H be the universal hash family from $[m]$ to $[r]$. The* closure of R under H, $H(R)$, *is the relation defined as follows.*

1. *let M be the matrix corresponding to R,*
2. *let $g_1, \ldots, g_{|D|^r}$ be an enumeration of all functions $g : [r] \mapsto D$,*
3. *for every $h_i \in H$ and every g_j add the column $y_{i,j}$ to M which in row $x \in [m]$ takes the value $g_j(h_i(x))$,*
4. *let $H(R)$ be the relation corresponding to M.*

Say that a relation R over D is *n-saturated* if for every $t_1, \ldots, t_{n'} \in R$, $n' \leq n$, for every $(x_1, \ldots, x_{n'}) \in D^{n'}$ there exists an i such that $(t_1[i], \ldots, t_{n'}[i]) = (x_1, \ldots, x_{n'})$.

Lemma 34. *Let R be a relation with m tuples and let $r \geq 1$. Let H be the universal hash family from $[m]$ to $[r]$. Then $H(R)$ is r-saturated.*

Proof. Let $t_1, \ldots, t_q \in H(R)$, $q \leq r$, let M be the matrix corresponding to $H(R)$. For every $(x_1, \ldots, x_q) \in D^q$ we must prove that there exists some j such that $(x_1, \ldots, x_q) = (t_1[j], \ldots, t_q[j])$. Let $P = (p_1, \ldots, p_q) \in [m]^q$ be the row indices of $t_1, \ldots, t_q$, i.e., $t_i = M[p_i, \cdot]$ for each $i \in [q]$. Since H is a universal hash family, there is some $h \in H$ which is injective on P. Let $g : D^q \mapsto D$ be the function satisfying $(g(h(p_1)), \ldots g(h(p_q))) = (x_1, \ldots, x_q)$. Due to the construction of $H(R)$ this implies that the column in M corresponding to h and g will enumerate $(x_1, \ldots, x_q)$. Hence, there is a j such that $(x_1, \ldots, x_q) = (t_1[j], \ldots, t_q[j])$. $\qquad\square$

If R is a relation and Γ a constraint language we let $\mathrm{Pol}(\Gamma)(R)$ denote the closure of R under $\mathrm{Pol}(\Gamma)$. Formally this relation can be defined as $\mathrm{Pol}(\Gamma)(R) = \bigcap_{R' \in \langle \Gamma \rangle, R \subseteq R'} R'$.

Lemma 35. *Let $\mathrm{Pol}(\Gamma)$ be an essentially unary clone. If $\mathrm{pPol}(\Gamma)$ is of finite order, then Γ can p.p. define all n-ary relations $R \in \langle \Gamma \rangle$ with at most $O(n)$ existentially quantified variables.*

Proof. Let R be an n-ary relation in $\langle \Gamma \rangle$, and let $m \leq |D|^n$ be the number of tuples in R. Let S be a finite base of $\mathrm{pPol}(\Gamma)$, let r be the largest arity of any function in S, and let H be the r-universal hash family from $[m]$ to $[r]$ of Theorem 32. Let $R' = H(R)$. By the construction of $H(R)$ in Definition 33 it follows that $\mathrm{ar}(R') = n + |D|^r |H| = n + |D|^r 2^{O(r)} \log m = |D|^r 2^{O(r)} O(n)$. To see that the last equality holds simply note that $\log(m) \leq \log(|D|^n) = O(n)$. Moreover, since r is a constant, it also holds that $\mathrm{ar}(R') = O(n)$. Let $p = \mathrm{ar}(R')$, and let $R'' = \mathrm{Pol}(\Gamma)(R')$, i.e. R' closed under all polymorphisms of Γ. Note that $\mathrm{Pol}(\Gamma) \subseteq \mathrm{Pol}(R'')$. Note that

$$R(x_1, \ldots, x_n) \equiv \exists x_{n+1}, \ldots, x_p R''(x_1, \ldots, x_n, x_{n+1}, \ldots, x_p),$$

or, put in other words, R'' can p.p. define R with at most $O(n)$ existentially quantified variables. To see that this holds, simply note that $\{(x_1, \ldots, x_n) \mid (x_1, \ldots, x_n, x_{n+1}, \ldots, x_p) \in R''\} = R$.

It remains to prove that $\mathrm{pPol}(\Gamma) \subseteq \mathrm{pPol}(R'')$, since this, due to the Galois connection in Theorem 2, implies that $\langle R'' \rangle_{\nexists} \subseteq \langle \Gamma \rangle_{\nexists}$ and that Γ can p.p. define R using at most $O(n)$ existentially quantified variables. Hence, let $f \in S$ be a q-ary, $q \leq r$, function. If $f \notin \mathrm{pPol}(R'')$ then there exists $t_1, \ldots, t_q \in R''$ such that $f(t_1, \ldots, t_q) \notin R''$. We may assume that all $t_1, \ldots, t_q$ are distinct, as otherwise the application of f is equivalent to the application of some q'-ary partial polymorphism f' on distinct rows, where q' is the number of distinct rows represented in $(t_1, \ldots, t_q)$ [22].

Our strategy is now, using Lemma 34, to prove that we can define a total function h using the partial function f such that h does not preserve R''. However, this also implies that $h \notin \mathrm{Pol}(\Gamma)$, which is a contradiction since $f \in \mathrm{pPol}(\Gamma)$. Before this proof we make one observation: for every $t \in R''$ there exists $t' \in R'$ and a unary function $h \in \mathrm{Pol}(\Gamma)$ such that $h(t') = t$. Hence, for the tuples $t_1, \ldots, t_q$ there exists $t'_1, \ldots, t'_q \in R'$ and $h_1, \ldots, h_q \in \mathrm{Pol}(\Gamma)$ such that $h_i(t'_i) = t_i$. We now define the q-ary function h as

$$h(x_1, \ldots, x_q) = f(h_1(\pi_1^q(x_1, \ldots, x_q)), \ldots, h_q(\pi_q^q(x_1, \ldots, x_q))).$$

Obviously, $h \in \mathrm{pPol}(\Gamma)$ since it is a composition of $f, h_1, \ldots, h_q$, and projection functions. This in turn implies that

$$h(t'_1, \ldots, t'_q) = f(h_1(t'_1), \ldots, h_q(t'_q)) = f(t_1, \ldots, t_q) \notin R'',$$

but since $t'_1, \ldots, t'_q \in R'$, R' is r-saturated and $q \leq r$, h must be a total polymorphism, i.e. $h \in \mathrm{Pol}(\Gamma) \subseteq \mathrm{Pol}(R'')$. This is a contradiction since $h(t'_1, \ldots, t'_q) \notin R''$. Hence, $f \in \mathrm{pPol}(R'')$. $\square$

With the help of this Lemma we can now prove that $\mathrm{pPol}(\Gamma)$ is of infinite order whenever Γ is finite and $\mathrm{Pol}(\Gamma)$ is essentially unary.

Theorem 36. *Let C be an essentially unary clone over a finite domain D. Then $\mathcal{I}_{\mathrm{fin}}(C)$ is infinitely generated.*

Proof. Let Γ be a finite constraint language such that $\mathrm{Pol}(\Gamma) = C$. Assume that $\mathrm{pPol}(\Gamma)$ can be finitely generated. By Lemma 35 we then have that Γ can p.p. define all n-ary relations in $\mathrm{Inv}(C)$ with $O(n)$ existentially quantified variables. However, this is a contradiction since $\langle \Gamma \rangle$ is superpolynomially closed by Theorem 20. To see this simply note that C cannot contain a k-edge operation for any $k \geq 2$ since a k-edge operation by definition is not essentially unary. This fact together with Lemma 18 results in a contradiction. Hence, $\mathrm{pPol}(\Gamma)$ cannot be of finite order. $\square$

This theorem has a number of interesting applications. First, recall from Section 2.3 that for Boolean constraint languages Γ, $\mathrm{CSP}(\Gamma)$ is NP-complete if and only if $\mathrm{Pol}(\Gamma) \subseteq [\neg x]$. Hence, assuming $\mathrm{P} \neq \mathrm{NP}$, $\mathrm{pPol}(\Gamma)$ is of infinite order whenever Γ is finite and $\mathrm{CSP}(\Gamma)$ is NP-complete. This implies that describing partial polymorphisms of finite constraint languages resulting in NP-hard CSP problems is a very difficult problem. For some illustrative usages of this theorem, let $R_{1/k} = \{(x_1, \ldots, x_k) \mid x_1, \ldots, x_k \in \{0, 1\}, \Sigma_{i=1}^{k} x_i = 1\}$, and let Γ_{SAT}^k and Γ_{NAE}^k be the restrictions of the plain bases of BR and $\mathrm{Inv}(\mathsf{N}_2)$, respectively, restricted to relations of arity at most k. It is easy to see that $\mathrm{CSP}(\{R_{1/k}\})$, $\mathrm{CSP}(\Gamma_{\mathrm{NAE}}^k)$, and $\mathrm{CSP}(\Gamma_{\mathrm{SAT}}^k)$ can be seen as alternative formulations of the well-known NP-complete problems 1-IN-k-SAT, NOT-ALL-EQUAL-k-SAT, and k-SAT, respectively. Since all these languages are finite we obtain the following corollary to Theorem 36.

Corollary 37. *Let $k \geq 3$. Then $\mathrm{pPol}(\Gamma_{\mathrm{SAT}}^k)$, $\mathrm{pPol}(\Gamma_{\mathrm{NAE}}^k)$, and $\mathrm{pPol}(R_{1/k})$ are of infinite order.*

It is worth noting that a complete dichotomy theorem for $\mathrm{CSP}(\Gamma)$ for constraint languages Γ defined over arbitrary finite domains is not yet known. However, if $\mathrm{Pol}(\Gamma)$ is essentially unary and every $f \in \mathrm{Pol}(\Gamma)$ is injective, then $\mathrm{CSP}(\Gamma)$ is NP-complete [17]. Hence, Theorem 36 also extends to many non-Boolean cases where $\mathrm{CSP}(\Gamma)$ is NP-complete.

7 Concluding Remarks and Open Research Questions

We have studied the question of whether a polynomial amount of variables is sufficient to define any relation in a given co-clone, have provided a complete dichotomy theorem for the Boolean co-clones where this is possible, and have also given several given several general results for arbitrary finite domains. In the process, we have also extended the concept of a weak base from Schnoor and Schnoor [28] and have given weak bases of all Boolean co-clones of infinite order. Using these notions we have then studied the question of whether a given strong partial clone is of finite or infinite order, and proven that the latter holds for a large variety of well-studied constraint languages. We now discuss some possibilities of future research.

Polynomially closed co-clones and few subpowers

From the results in Section 3 and Section 5 we see that the question whether a co-clone of finite order is polynomially closed is related to the question whether the corresponding algebra has few subpowers. For the Boolean domain, these two notions exactly coincide, and it would be interesting to see whether this holds in the more general setting of arbitrary finite domains, possibly using some of the machinery developed in Berman et al. [3].

Partial polymorphisms and superpolynomially closed co-clones

Theorem 36 states that a $\mathrm{pPol}(\Gamma)$ is always of infinite order whenever Γ is finite and $\mathrm{Pol}(\Gamma)$ is essentially unary. It would be interesting to try to extend this theorem to the case when $\langle \Gamma \rangle$ is an arbitrary superpolynomially closed co-clone, and a possible starting point is to investigate the case when $\mathrm{Pol}(\Gamma)$ can be generated from a finite set of binary functions. However, this appears to be far from straightforward, and even in the apparently simple case when $\mathrm{Pol}(\Gamma) = [x_1 \wedge x_2]$, the proof strategy in Lemma 35, based on constructing a universal hash family, breaks down.

Partial Polymorphisms of finite Boolean constraint languages

In the light of Theorem 36, describing the partial polymorphisms of *any* finite Boolean constraint language Γ such that $\mathrm{CSP}(\Gamma)$ is NP-complete is a challenging problem since $\mathrm{pPol}(\Gamma)$ is of infinite order. Nevertheless, recent research shows that this problem can be circumvented by considering stronger closure operators than functional composition [22]. Using this approach it would be interesting to attempt to give a general characterization of the partial polymorphisms of the constraint languages in the "bottom" of BR, e.g., all constraint languages Γ such that $\langle \Gamma \rangle = \mathsf{BR}$ and $\langle \Gamma \rangle_{\overline{\not\exists}} \subseteq \langle R_{1/3} \rangle_{\overline{\not\exists}}$.

Acknowledgements

References

1. V. B. Alekseev and A. A. Voronenko. On some closed classes in partial two-valued logic. *Discrete Mathematics and Applications*, 4(5):401–419, 1994.
2. N. Alon, R. Yuster, and U. Zwick. Color-coding. *Journal of the ACM*, 42(4):844–856, July 1995.
3. J. Berman, P. Marković, R. Mckenzie, M. Valeriote, and R. Willard. Varieties with few subalgebras of powers. *Transactions of the American Mathematical Society*, 362(3):1445 – 1473, 2006.
4. V. G. Bodnarchuk, L. A. Kaluzhnin, V. N. Kotov, and B. A. Romov. Galois theory for Post algebras. I. *Cybernetics*, 5:243–252, 1969.

5. V. G. Bodnarchuk, L. A. Kaluzhnin, V. N. Kotov, and B. A. Romov. Galois theory for Post algebras. II. *Cybernetics*, 5:531–539, 1969.

6. E. Böhler, N. Creignou, S. Reith, and H. Vollmer. Playing with Boolean blocks, part I: Post's lattice with applications to complexity theory. *ACM SIGACT-Newsletter*, 34(4):38–52, 2003.

7. F. Börner and L. Haddad. Maximal partial clones with no finite basis. *Algebra Universalis*, 40(4):453–476, 1998.

8. M. Couceiro and L. Haddad. Finitely generated maximal partial clones and their intersections. In *Proceedings of the 40th IEEE International Symposium on Multiple-Valued Logic, (ISMVL-2010)*, pages 122–127. IEEE Computer Society, 2010.

9. N. Creignou, P. Kolaitis, and B. Zanuttini. Structure identification of Boolean relations and plain bases for co-clones. *Journal of Computer and System Sciences*, 74(7):1103–1115, November 2008.

10. N. Creignou, P. G. Kolaitis, and H. Vollmer, editors. *Complexity of Constraints - An Overview of Current Research Themes*, volume 5250 of *Lecture Notes in Computer Science*. Springer, 2008.

11. J. Fugère and L. Haddad. On partial clones containing all idempotent partial operations. In *Proceedings of the 28th IEEE International Symposium on Multiple-Valued Logic (ISMVL-98)*, pages 369–373, 1998.

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

13. L. Haddad and J. Fugère. Maximal chains of partial clones containing all idempotent partial functions. In *Proceedings of the 29th IEEE International Symposium on Multiple-Valued Logic (ISMVL-99)*, pages 80–84. IEEE Computer Society, 1999.

14. P. Idziak, P. Marković, R. McKenzie, M. Valeriote, and R. Willard. Tractability and learnability arising from algebras with few subpowers. *SIAM Journal on Computing*, 39(7):3023–3037, June 2010.

15. P. Jeavons. On the algebraic structure of combinatorial problems. *Theoretical Computer Science*, 200:185–204, 1998.

16. P. Jeavons, D. Cohen, and M. C. Cooper. Constraints, consistency and closure. *Artificial Intelligence*, 101(1–2):251 – 265, 1998.

17. P. Jeavons, D. Cohen, and M. Gyssens. Closure properties of constraints. *Journal of the ACM*, 44(4):527–548, July 1997.

18. P. Jonsson, V. Lagerkvist, G. Nordh, and B. Zanuttini. Complexity of SAT problems, clone theory and the exponential time hypothesis. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA-2013)*, pages 1264–1277, 2013.

19. P. Jonsson, V. Lagerkvist, J. Schmidt, and H. Uppman. Relating the time complexity of optimization problems in light of the exponential-time hypothesis. In *Proceedings of the 39th International Symposium on Mathematical Foundations of Computer Science (MFCS-14)*, pages 408–419, Berlin, Heidelberg, 2014. Springer-Verlag.

20. V. Lagerkvist. Weak bases of Boolean co-clones. *Information Processing Letters*, 114(9):462–468, 2014.

21. V. Lagerkvist. *Strong Partial Clones and the Complexity of Constraint Satisfaction Problems: Limitations and Applications*. PhD thesis, Linköping University, The Institute of Technology, 2016.

22. V. Lagerkvist, M. Wahlström, and B. Zanuttini. Bounded bases of strong partial clones. In *Proceedings of the 45th International Symposium on Multiple-Valued Logic (ISMVL-2015)*, pages 189–194, 2015.

23. G. Nordh and B. Zanuttini. Frozen Boolean partial co-clones. In *Proceedings of the 39th International Symposium on Multiple-Valued Logic (ISMVL-2009)*, pages 120 –125, 2009.

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

25. R. W. Quackenbush. On the composition of idempotent functions. *Algebra Universalis*, 1(1):7–12, 1971.

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

27. T. Schaefer. The complexity of satisfiability problems. In *Proceedings of the 10th Annual ACM Symposium on Theory Of Computing (STOC-78)*, pages 216–226. ACM Press, 1978.

28. 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 Berlin Heidelberg, 2008.

29. K. Schölzel. Dichotomy on intervals of strong partial boolean clones. *Algebra Universalis*, 73(3-4):347–368, 2015.