Counting quantifiers, subset surjective functions, and counting CSPs

Andrei A. Bulatov Amir Hedayaty

***Abstract*—We introduce a new type of closure operator on the set of relations, max-implementation, and its weaker analog max-quantification. Then we show that approximation reductions between counting constraint satisfaction problems (CSPs) are preserved by these two types of closure operators. Together with some previous results this means that the approximation complexity of counting CSPs is determined by partial clones of relations that additionally closed under these new types of closure operators. Galois correspondence of various kind have proved to been quite helpful in the study of the complexity of the CSP. While we were unable to identify a Galois correspondence for partial clones closed under max-implementation and max-quantification, we obtain such results for slightly different type of closure operators, k-existential quantification. This type of quantifiers are known as counting quantifiers in model theory, and often used to enhance first order logic languages. We characterize partial clones of relations closed under k-existential quantification as sets of relations invariant under a set of partial functions that satisfy the condition of k-subset surjectivity.**

I. Introduction

Clones of functions and clones of relations in their various incarnations have proved to be an immensely powerful tool in the study of the complexity of different versions of the Constraint Satisfaction Problem (CSP, for short). In a CSP the aim is to find an assignment of values to a given set of variables, subject to constraints on the values that can be assigned simultaneously to certain specified subsets of variables. A CSP can also be expressed as the problem of deciding whether a given conjunctive formula has a model. In the counting version of the CSP the goal is to find the number of satisfying assignments, and in the quantified version we need to verify if a first order sentence, whose quantifier-free part is conjunctive, is true in a given model.

The general CSP is NP-complete [17]. However, many practical and theoretical problems can be expressed in terms of CSP using constraints of a certain restricted form. One of the most widely used way to restrict a constraint satisfaction problem is to specify the set of allowed constraints, which is usually a collection of relations on a finite set. The key result is that this set of relations can usually be assumed to be a co-clone of a certain kind. More precisely, a generic statement asserts that if a relation R belongs to the co-clone generated by a set Γ of relations then the CSP over $\Gamma \cup \{R\}$ is polynomial time reducible to the CSP over Γ. Then we can use the appropriate Galois connection to transfer the question about sets of relations to a question about certain class of functions.

For the classical decision CSP such a result was obtained by Jeavons et al. [16], who proved that intersection of relations (that is, conjunction of the corresponding predicates) and projections (that is, existential quantification) give rise to polynomial time reducibility of CSPs. Therefore in the study of the complexity of the CSP it suffices to focus on co-clones. Using the result of Geiger [13] or of Bodnarchuk et al. [2] one can instead consider regular clones of functions. A similar result is true for the counting CSP as shown by Bulatov and Dalmau [8]. In the case of quantified CSP, Börner et al. proved [3] that conjunction, existential quantification, and also universal quantification give rise to a polynomial time reduction between quantified problems. The appropriate class of functions is then the class of surjective functions. Along with the usual counting CSP, a version, in which one is required to approximate the number of solutions, has also been considered. The standard polynomial time reduction between problems is not suitable for approximation complexity. In this case, therefore, another type of reductions, approximation preserving, or, AP-reductions, are used. The first author proved in [7] that conjunction of predicates gives rise to an AP-reduction between approximation counting CSPs. By the Galois connection established by Fleischner and Rosenberg [12], the approximation complexity of a counting CSP is a property of a clone of partial functions.

In most cases establishing the connection between clones of functions and reductions between CSPs has led to a major success in the study of the CSP. For the decision problem, a number of very strong results have been proved using methods of universal algebra [9], [4], [5], [1], [15]. For the exact counting CSP a complete complexity classification of such problems has been obtained [6]. Substantial progress has been also made in the quantified CSP [10].

Compared to the results cited above the progress made in the approximation counting CSP is modest. Perhaps, one reason for this is that clones of partial functions are much less studied, and much more diverse than clones of total functions. In this paper we attempt to overcome to some extent the difficulties arising from the weakness of partial clones.

In the first part of the paper we introduce new types of quantification and show that such quantifications, we call them max-implementation and max-quantification, give rise to AP-reductions between approximation counting CSPs. Intuitively, applying the max-existential quantifier to a relation $R(x_1,\dots,x_n,y)$ results in the relation $\exists_{\max} y R(x_1,\dots,x_n,y)$ that contains those tuples $(a_1,\dots,a_n)$ that have a maximal number of extensions $(a_1,\dots,a_n,b)$ such that $R(a_1,\dots,a_n,b)$ is satisfied. Thus we strengthen the closure operator on sets of relation hoping that the sets of functions corresponding to the new type of Galois connection are more tractable. We were unable, however, to describe a Galois connection for sets closed under max-implementation and max-quantification. Instead, we consider a somewhat close type of quantifiers, k-existential quantifiers. This type of quantifiers are known as counting quantifiers in model theory, and often used to enhance first order logic languages (see, e.g. [11]). Counting quantifiers are similar to max-existential quantifiers, although do not capture them completely. We call sets of relations closed under conjunctions and k-existential quantification k-existential co-clones. On the functional side, an n-ary (partial) function on a set D is said to be k-subset surjective if it is surjective on any collection of k-element subsets. More precisely, for any k-element subsets $A_1,\dots,A_n \subseteq D$ the set $f(A_1,\dots,A_n)$ is either empty or contains at least k elements. The second result of the paper asserts that k-existential co-clones are exactly the sets of relation invariant with respect to a set of k-subset surjective (partial) functions.

II. Preliminaries

By $[n]$ we denote the set $\{1,\dots,n\}$. For a set D by D^n we denote the set of all *n-tuples* of elements of D. An n-ary relation is any set $R \subseteq D^n$. The number n is called the *arity* of R and denoted $\mathrm{ar}(R)$. Tuples will be denoted in boldface, say, $\mathbf{a}$, and their entries denoted by $\mathbf{a}[1],\dots,\mathbf{a}[n]$. For $I = (i_1,\dots,i_k) \subseteq [n]$ by $\mathrm{pr}_I\mathbf{a}$ we denote the tuple $(\mathbf{a}[i_1],\dots,\mathbf{a}[i_k])$, and we use $\mathrm{pr}_I R$ to denote $\{\mathrm{pr}_I\mathbf{a} \mid \mathbf{a} \in R\}$. We will also need predicates corresponding to relations. To simplify the notation we use the same symbol for a relation and the corresponding predicate, for instance, for an n-ary relation R the corresponding predicate $R(x_1,\dots,x_n)$ is given by $R(\mathbf{a}[1],\dots,\mathbf{a}[n]) = 1$ if and only if $\mathbf{a} \in R$.

For a set of relations Γ over a set D, the set $\langle\langle\Gamma\rangle\rangle$ includes all relations that can be expressed (as a predicate) using (a) relations from Γ, together with the binary equality relation $=_D$ on D, (b) conjunctions, and (c) existential quantification. This set is called the *co-clone generated by* Γ.

Weak co-clone generated by Γ is obtained in a similar way by disallowing existential quantification. $\langle\Gamma\rangle$ includes all relations that can be expressed using (a) relations from Γ, together with $=_D$, and (b) conjunctions,

If $\Gamma = \langle\Gamma\rangle$ or $\Gamma = \langle\langle\Gamma\rangle\rangle$ then the set Γ is said to be a *weak co-clone*, and a *co-clone*, respectively.

Co-clones and weak co-clones can often be conveniently and concisely represented trough functions and partial functions, respectively.

Let R be a (k-ary) relation on a set D, and $f\colon D^n \to D$ an n-ary function on the same set. Function f *preserves* R, or is a *polymorphism* of R, if for any n tuples $\mathbf{a}_1,\dots,\mathbf{a}_n \in R$ the tuple $f(\mathbf{a}_1,\dots,\mathbf{a}_n)$ obtained by component-wise application of f also belongs to R. Relation R in this case is said to be *invariant* with respect to f. The set of all functions that preserve every relation from a constraint language Γ is denoted by $\mathsf{Pol}(\Gamma)$, the set of all relations invariant with respect to a set of functions C is denoted by $\mathsf{Inv}(C)$.

Operators Inv and Pol form a Galois connection between sets of functions and sets of relations. Sets of the form $\mathsf{Inv}(C)$ are precisely co-clones; on the operational side there is another type of closed sets.

A set of functions is said to be a *clone* of functions if it is closed under superpositions and contain all the *projection* functions, that is functions of the form $f(x_1,\dots,x_n) = x_i$. Sets of functions of the form $\mathsf{Pol}(\Gamma)$ are exactly clones of functions [18] .

The study of the #CSP also makes use of another Galois connection, a connection between weak co-clones and sets of *partial functions*. An n-ary partial function f on a set D is just a partial mapping $f\colon D^n \to D$. As in the case of total functions, a partial function f *preserves* relation R, if for any n tuples $\mathbf{a}_1,\dots,\mathbf{a}_n \in R$ the tuple $f(\mathbf{a}_1,\dots,\mathbf{a}_n)$ obtained by component-wise application of f is either undefined or belongs to R. The set of all partial functions that preserve every relation from a set of relations Γ is denoted by $\mathsf{pPol}(\Gamma)$.

The set of all tuples from D^n on which f is defined is called the *domain* of f and denoted by $\mathsf{dom}(f)$. A set of functions is said to be *down-closed* if along with a function f it contains any function f' such that

$\mathsf{dom}(f') \subseteq \mathsf{dom}(f)$ and $f'(a_1,\ldots,a_n) = f(a_1,\ldots,a_n)$ for every tuple $(a_1,\ldots,a_n) \in \mathsf{dom}(f')$. A down-closed set of functions, containing all projections and closed under superpositions is called a *partial clone*. Fleischner and Rosenberg [12] proved that Partial clones are exactly the sets of the form $\mathsf{pPol}(\Gamma)$ for a certain Γ, and the weak co-clones are precisely the sets $\mathsf{Inv}(C)$ for collections C of partial functions.

III. Approximate counting and max-quantifiers

Let D be a set, and let Γ be a finite set of relations over D. An instance of the counting Constraint Satisfaction Problem, $\#\mathrm{CSP}(\Gamma)$ is a pair $\mathcal{P} = (V, \mathcal{C})$ where V is a set of *variables*, and $\mathcal{C}$ is a set of *constraints*. Every constraint is a pair $\langle \mathbf{s}, R\rangle$, in which R is a member of Γ, and $\mathbf{s}$ is a tuple of variables from V of length $\mathrm{ar}(R)$ (possibly with repetitions). A *solution* to $\mathcal{P}$ is a mapping $\varphi : V \to D$ such that $\varphi(\mathbf{s}) \in R$ for every constraint $\langle \mathbf{s}, R\rangle \in \mathcal{C}$. The objective in $\#\mathrm{CSP}(\Gamma)$ is to find the number $\#\mathcal{P}$ of solutions to a given instance $\mathcal{P}$.

We are interested in the complexity of this problem depending on the set Γ. The complexity of the exact counting problem (when we are required to find the exact number of solutions) is settled in [6] by showing that for any finite D and any set Γ of relations over D the problem is polynomial time solvable or is complete in a natural complexity class $\#P$. One of the key steps in that line of research is the following result: For a relation R and a set of relations Γ over D, such that R belongs to the co-clone generated by Γ, then $\#\mathrm{CSP}(\Gamma \cup \{R\})$ is polynomial time reducible to $\#\mathrm{CSP}(\Gamma)$. This results emphasizes the importance of co-clones in the study of constraint problems.

A situation is different if we are concerned about approximating the number of solutions. We will need some notation and terminology. Let A be a counting problem. An algorithm Alg is said to be an *approximation algorithm* for A with relative error ε (which may depend on the size of the input) if it is polynomial time and for any instance $\mathcal{P}$ of A it outputs a certain number $\mathsf{Alg}(\mathcal{P})$ such that

$$\frac{|\#\mathcal{P} - \mathsf{Alg}(\mathcal{P})|}{\#\mathcal{P}} < \varepsilon,$$

where $\#\mathcal{P}$ denotes the exact number of solutions to $\mathcal{P}$.

The following framework is viewed as one of the most realistic models of efficient computations. A *fully polynomial approximation scheme* (FPAS, for short) for a problem A is an algorithm Alg such that: It takes as input an instance $\mathcal{P}$ of A and a real number $\varepsilon > 0$, the relative error of Alg on the input $(\mathcal{P}, \varepsilon)$ is less than ε, and Alg is polynomial in the size of $\mathcal{P}$ and $\log(\frac{1}{\varepsilon})$.

To determine the approximation complexity of problems another type of reductions is used. Suppose A and B are two counting problems whose complexity (of approximation) we want to compare. An *approximation preserving reduction* or *AP-reduction* from A to B is an algorithm Alg, using B as an oracle, that takes as input a pair $(\mathcal{P}, \varepsilon)$ where $\mathcal{P}$ is an instance of A and $0 < \varepsilon < 1$, and satisfies the following three conditions: (i) every oracle call made by Alg is of the form $(\mathcal{P}', \delta)$, where $\mathcal{P}'$ is an instance of B, and $0 < \delta < 1$ is an error bound such that $\frac{1}{\delta}$ is bounded by a polynomial in the size of $\mathcal{P}$ and $\frac{1}{\varepsilon}$; (ii) the algorithm Alg meets the specifications for being approximation scheme for A whenever the oracle meets the specification for being approximation scheme for B; and (iii) the running time of Alg is polynomial in the size of I and $\log(\frac{1}{\varepsilon})$. If an approximation preserving reduction from A to B exists we write $A \le_{\mathrm{AP}} B$, and say that A is *AP-reducible to* B.

Similar to co-clones and polynomial time reductions, weak co-clones can be shown to be preserved by AP-reductions.

Theorem 1 ([7]): Let R be a relation and Γ be a set of relations over a finite set such that R belongs to $\langle\Gamma\rangle$. Then $\#\mathrm{CSP}(\Gamma \cup \{R\})$ is AP-reducible to $\#\mathrm{CSP}(\Gamma)$.

This result however has two significant setbacks. First, weak co-clones are not studied to the same extent as regular co-clones, and, due to greater diversity, are not believed to be ever studied to a comparable level. Second, it does not used the full power of AP-reductions, and therefore leaves significant space for improvements. In the rest of this section we try to improve upon the second issue.

Definition 2: Let Γ be a set of relations on a set D, and let R be an n-ary relation on D. Let $\mathcal{P}$ be an instance of $\#\mathrm{CSP}(\Gamma)$ over the set of variables consisting of $V = V_x \cup V_y$, where $V_x = \{x_1, x_2, \cdots, x_n\}$ and $V_y = \{y_1, y_2, \cdots, y_q\}$. For any assignment of $\varphi : V_x \to D$, let $\#\varphi$ be the number of assignments $\psi : V_y \to D$ such that $\varphi \cup \psi$ satisfy $\mathcal{P}$. Let M be the maximum value of $\#\varphi$ among all assignments of V_x. The instance $\mathcal{P}$ is said to be a *max-implementation* of R if a tuple φ is in R if and only if $\#\varphi = M$.

Theorem 3: If there is max-implementation of R by Γ, then $\#\mathrm{CSP}(\Gamma \cup \{R\}) \le_{AP} \#\mathrm{CSP}(\Gamma)$.

Proof: For any instance $\mathcal{P}_1 = (V_1, \mathcal{C}_1)$ of $\#\mathrm{CSP}(\Gamma \cup \{R\})$ we construct an instance $\mathcal{P}_2 = (V_2, \mathcal{C}_2)$ of $\#\mathrm{CSP}(\Gamma)$ as follows.

- Choose a sufficiently large integer m (to be determined later).

- Let $C_1, \ldots, C_\ell \in \mathcal{C}_1$ be the constraints from $\mathcal{P}_1$ involving R, $C_i = \langle \mathbf{s}_i, R\rangle$. Set $V_2 = V_1 \cup \bigcup_{i=1}^{\ell}(V_1^i \cup \ldots \cup V_m^i)$, where each V_j^i is a fresh copy of V_y from Definition 2.
- Let $\mathcal{C}$ be the set of constraints of $\mathcal{P}$ (see Definition 2). Set $\mathcal{C}_2 = (\mathcal{C}_1 - \{C_1, \ldots, C_\ell\}) \cup \bigcup_{i=1}^{\ell}(\mathcal{C}_1^i \cup \ldots \cup \mathcal{C}_m^i)$, where each $\mathcal{C}_j^i$ is a copy of $\mathcal{C}$ defined as follows. For each $\langle \mathbf{s}, Q\rangle \in \mathcal{C}$ we include $\langle \mathbf{s}_j^i, Q\rangle$ into $\mathcal{C}_j^i$, where $\mathbf{s}_j^i$ is obtained from $\mathbf{s}$ replacing every variable from V_y with its copy from V_j^i.

Now, as is easily seen, every solution of $\mathcal{P}_1$ can be extended to a solution of $\mathcal{P}_2$ in $M^{\ell m}$ ways. Observe that sometimes the restriction of a solution ψ of $\mathcal{P}_2$ to V_1 is not a solution of $\mathcal{P}_1$. Indeed, it may happen that although ψ satisfies every copy $\mathcal{C}_j^i$ of $\mathcal{P}$, its restriction to $\mathbf{s}_j^i$ does not belong to R, simply because this restriction does not have sufficiently many extensions to solutions of $\mathcal{P}$. However, any assignment to V_1 that is not a solution to $\mathcal{P}_1$ can be extended to a solution of $\mathcal{P}_2$ in at most $(M-1)^m \cdot M^{(\ell-1)m}$ ways. Hence,

$$\begin{aligned} M^{\ell m} \cdot \#\mathcal{P}_1 &\le \#\mathcal{P}_2 \\ &\le M^{\ell m} \cdot \#\mathcal{P}_1 + |V_1|^{|D|} \cdot (M-1)^m \cdot M^{(\ell-1)m} \end{aligned}$$

Then we output $\#\mathcal{P}_2/M^{\ell m}$.

Let $|V_1| = k$ and $|D| = d$. Given a desired relative error ε we have to find m such that $\frac{\#\mathcal{P}_2}{M^{\ell m}} - \#\mathcal{P}_1 < \varepsilon$. A straightforward computation shows that any

$$m > \frac{\log \varepsilon - d \log k}{\log M - \log(M-1)}$$

achieves the goal. ■

Max-implementation can be used as another closure operator on the set of relations. A set of relations Γ over D is said to be a *max-co-clone* if it contains the equality relations, and closed under conjunctions and max-implementations. The smallest max-co-clone containing a set of relations Γ is called the *max-co-clone generated by* Γ and denoted $\langle\Gamma\rangle_{\max}$.

The next natural step would be to find a type of functions and closure operator on the set of functions that give rise to a Galois connection capturing max-co-clones.

Problem 1: Find a class $\mathcal{F}$ of (partial) functions and a closure operator $[\cdot]$ on this class such that for any set of relations Γ and any set $C \subseteq \mathcal{F}$ it holds that $\langle\Gamma\rangle_{\max} = \mathsf{Inv}(\mathcal{F} \cap \mathsf{pPol}(\Gamma))$, and $[C] = \mathcal{F} \cap \mathsf{pPol}\,\mathsf{Inv}(C)$.

In all the cases studied the projection (or quantification) type operators on relations can be reduced to quantifying away a single variable. It is not clear, however, if this can be done for max-implementations, which seems to inherently involve a number of variables, rather than a single variable. Therefore a meaningful relaxation of max-co-clones restricts the use of max-implementation to one auxiliary variable. Let Φ be a formula with free variables $x_1, \ldots, x_n$ and y over set D and some predicate symbols. Then $a_1, \ldots, a_n$ satisfy

$$\Psi(x_1, \ldots, x_n) = \exists_{\max} y \Phi(x_1, \ldots, x_n, y)$$

if and only if the number of $b \in D$ such that $\Phi(a_1, \ldots, a_n, b)$ is true is maximal among all tuples $(c_1, \ldots, c_n) \in D^n$. The quantifier $\exists_{\max}$ will be called *max-existential quantifier*. A set of relations Γ over D is said to be a *max-existential co-clone* if it contains the equality relation, and closed under conjunctions and max-existential quantification. The smallest max-existential co-clone containing a set of relations Γ is called the *max-existential co-clone generated by* Γ and denoted $\langle\Gamma\rangle^1_{\max}$.

Problem 2: Find a class $\mathcal{F}$ of (partial) functions and a closure operator $[\cdot]$ on this class such that for any set of relations Γ and any set of functions $C \subseteq \mathcal{F}$ it holds that $\langle\Gamma\rangle^1_{\max} = \mathsf{Inv}(\mathcal{F} \cap \mathsf{pPol}(\Gamma))$, and $[C] = \mathcal{F} \cap \mathsf{pPol}\,\mathsf{Inv}(C)$.

In the next section we consider certain constructions approximating max-existential co-clones.

IV. k-Existential and max-existential co-clones

In order to approach max-quantification we consider counting quantifiers that have been used in model theory to increase the power of first order logic.

Let Φ be a formula with free variables $x_1, \ldots, x_n$ and y over set D and some predicate symbols. Then $a_1, \ldots, a_n$ satisfy

$$\Psi(x_1, \ldots, x_n) = \exists_k y \Phi(x_1, \ldots, x_n, y)$$

if and only if $\Phi(a_1, \ldots, a_n, b)$ is true for at least k values $b \in D$. The quantifier $\exists_k$ will be called *k-existential quantifier*.

We now introduce several types of co-clones depending on what kind of k-existential quantifiers are allowed. A set of relations Γ over set D is said to be a *k-existential weak co-clone* if it contains the equality relation $=_D$, and closed under conjunctions and k-existential quantification. It is called *k-existential co-clone* if in addition it is closed under regular existential quantification. The set Γ is said to be a *counting co-clone*[1] if it contains $=_D$, and closed under conjunctions and k-existential quantification for all $k \ge 1$. The smallest k-existential weak co-clone containing a set of relations Γ is called the *k-existential weak co-clone generated by* Γ and

[1] 'Counting' in this term comes from counting quantifiers and has nothing to do with counting constraint satisfaction.

denoted $\langle\Gamma\rangle_k$. Similarly, the smallest k-existential co-clone and the smallest counting co-clone containing Γ are called the *k-existential co-clone* and the *counting clone generated by* Γ and denoted $\langle\langle\Gamma\rangle\rangle_k$ and $\langle\langle\Gamma\rangle\rangle_\infty$, respectively.

The following observation summarizes some relationship between the constructions introduced.

Observation 4: For a set of relations Γ on D, $|D| = m$, the following hold.

- Γ is a 1-existential (weak) co-clone if and only if it is a co-clone.

- Γ is a (weak) m-existential clone if and only if it a (weak) co-clone closed under universal quantification.

- if Γ is a counting co-clone then it is a max-existential co-clone.

- if Γ is a max-existential co-clone then it is a weak m-existential clone.

In all other cases the introduced versions of co-clones are incomparable.

Example 5: Fix a natural number m and let D be a set with $\frac{m(m-1)}{2}$ elements. Consider an equivalence relation R on D with classes $D_1, \ldots, D_m$ such that $|D_i| = i$. Then the co-clone generated by R_m corresponds to one of the Rosenberg's maximal clones [19], and so the structure of relations from this co-clone is well understood. For any n-ary relation $Q \in \langle\langle R_m\rangle\rangle$ there is a partition $I_1, \ldots, I_k$ of $[n]$ such that a tuple $\mathbf{a}$ belongs to Q if and only if for each $j \in [k]$ and every $i, i' \in I_j$ the entries $\mathbf{a}[i], \mathbf{a}[i']$ are R_m-related. This also means that $\langle R_m\rangle = \langle\langle R_m\rangle\rangle$.

Applying k-existential and max-existential quantifiers one can easily find the k-existential, counting, and max-existential clones generated by R:

- $\langle R_m\rangle_k = \langle\langle R_m\rangle\rangle_k$ is the set of relations Q: There is a partition $I_1, \ldots, I_t$ of $[\mathrm{ar}(Q)]$ and $J \subseteq [t]$ such that a tuple $\mathbf{a}$ belongs to Q if and only if for each $j \in [t]$ and every $i, i' \in I_j$ the entries $\mathbf{a}[i], \mathbf{a}[i']$ are R_m-related and $\mathbf{a}[i] \in D_k \cup \ldots \cup D_m$ for $i \in I_j$, $j \in J$.

- $\langle\langle R_m\rangle\rangle_\infty$ is the set of relations Q: There is a partition $I_1, \ldots, I_t$ of $[\mathrm{ar}(Q)]$ and a function $\varphi : [t] \to [m]$ such that a tuple $\mathbf{a}$ belongs to Q if and only if for each $j \in [t]$ and every $i, i' \in I_j$ the entries $\mathbf{a}[i], \mathbf{a}[i']$ are R_m-related and $\mathbf{a}[i] \in D_{\varphi(j)} \cup \ldots \cup D_m$ for $i \in I_j$, $j \in J$.

- $\langle R_m\rangle_{\max} = \langle R_m\rangle^1_{\max}$ is the set of relations Q: There is a partition $I_1, \ldots, I_t$ of $[\mathrm{ar}(Q)]$ and $J \subseteq [t]$ such that a tuple $\mathbf{a}$ belongs to Q if and only if for each $j \in [t]$ and every $i, i' \in I_j$ the entries $\mathbf{a}[i], \mathbf{a}[i']$ are R_m-related and $\mathbf{a}[i] \in D_m$ for $i \in I_j$, $j \in J$.

A set Γ such that $\langle\Gamma\rangle_k \neq \langle\langle\Gamma\rangle\rangle_k$ can be easily found among usual weak co-clones. For instance, for any weak co-clone Γ that is not a co-clone we have $\langle\Gamma\rangle_1 \neq \langle\langle\Gamma\rangle\rangle_1$. Such a weak co-clone can be found in, say, [14].

In the example given we have $\langle R_m\rangle^1_{\max} = \langle R_m\rangle_m$. However, since $\langle R_{m-1}\rangle_m = \langle R_{m-1}\rangle$, we have $\langle R_{m-1}\rangle^1_{\max} \neq \langle R_{m-1}\rangle_m$. Distinguishing between $\langle\Gamma\rangle_{\max}$ and $\langle\Gamma\rangle^1_{\max}$ is more involved.

V. Galois correspondence

Let D be a finite set. A (partial) function $f\colon D^n \to D$ is said to be *k-subset surjective* if for any k-element subsets $A_1, \ldots, A_n \subseteq D$ the image $f(A_1, \ldots, A_n)$ is either empty, or has cardinality at least k. A (partial) function that is k-subset surjective for each k, $1 \leq k \leq |D|$ is said to be *subset surjective*. The set of all arity n k-subset surjective partial functions [arity n k-subset surjective functions, subset surjective functions] on D will be denoted by $P_D^{k,(n)}$ [resp., $F_D^{k,(n)}$, $F_D^{(n)}$]; furthermore, $P_D^k = \bigcup_{n\geq 0} P_D^{k,(n)}$, $F_D^k = \bigcup_{n\geq 0} F_D^{k,(n)}$, $F_D = \bigcup_{n\geq 0} F_D^{(n)}$. Any partial function is 1-subset surjective, while $|D|$-subset surjective partial functions are exactly the surjective partial functions, and the function with empty domain. Observe that this definition can be strengthened by allowing the sets A_i, $i \in [n]$, to have at least k elements.

Lemma 6: If an n-ary function f is k-subset surjective, then for any subsets $A_1, \ldots, A_n \subseteq D$ with $|A_i| \geq k$, $i \in [n]$, the image $f(A_1, \ldots, A_n)$ is either empty, or has cardinality at least k.

Proof: Suppose $f(A_1, \ldots, A_n)$ is nonempty. Then there are $B_i \subseteq A_i$, $i \in [n]$, such that $B = f(B_1, \ldots, B_n)$ is nonempty. As f is k-subset surjective, $|B| \geq k$. Finally, $B \subseteq f(A_1, \ldots, A_n)$, and the result follows. ■

The notion of invariance for k-subset surjective functions is the standard one for partial functions and relations. As usual, if C is a set of (k-) subset surjective (partial) functions, $\mathsf{Inv}(C)$ denotes the set of relations invariant with respect to every function from C. For a set Γ of relations, $\mathrm{m}(k)$-$\mathsf{Pol}(\Gamma)$ and $\mathrm{m}(k)$-$\mathsf{pPol}(\Gamma)$ denote the set of all k-subset surjective functions and partial functions, respectively, preserving every relation from Γ. By m-$\mathsf{Pol}(\Gamma)$ we denote the analogous set of subset surjective functions.

The operator Inv on one side and the operators $\mathrm{m}(k)$-$\mathsf{pPol}(\Gamma)$, $\mathrm{m}(k)$-$\mathsf{Pol}(\Gamma)$, m-$\mathsf{Pol}(\Gamma)$ on the other side form Galois correspondences in the standard fashion. We characterize closed sets of relations that give rise from this correspondence.

Lemma 7: Let $R(x_1, \ldots, x_\ell, y)$ be a relation on D, and let $Q(x_1, \ldots, x_\ell) = \exists_k y R(x_1, \ldots, x_\ell, y)$. Then if a k-subset surjective (partial) function f preserves R, it also preserves Q.

Proof: Suppose f is n-ary. Take $\mathbf{a}_1, \ldots, \mathbf{a}_n \in Q$. Since each of them is put into Q by k-existential quantification, it has at least k extensions to a tuple from R. Let $B_1, \ldots, B_n \subseteq D$ be such that $|B_i| \geq k$ and $(\mathbf{a}_i, b) \in R$ for $b \in B_i$ and $i \in [n]$. Let also $\mathbf{b} = f(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. For any $b \in B = f(B_1, \ldots, B_n)$ the tuple $(\mathbf{b}, b)$ belongs to R. As f is k-subset surjective, $|B| \geq k$, hence, $\mathbf{b} \in Q$. ■

Theorem 8: Let Γ be a set of relations on a set D.

(a) $\mathsf{Inv}\, \mathrm{m}(k)\text{-}\mathsf{pPol}(\Gamma) = \langle \Gamma \rangle_k$;
(b) $\mathsf{Inv}\, \mathrm{m}(k)\text{-}\mathsf{Pol}(\Gamma) = \langle\langle \Gamma \rangle\rangle_k$;
(c) $\mathsf{Inv}\, \mathrm{m}\text{-}\mathsf{Pol}(\Gamma) = \langle\langle \Gamma \rangle\rangle_\infty$;

Proof: (a) The equality relation, $=_D$, is invariant with respect to any partial function on D. Let f be a k-subset surjective functions. It is straightforward to verify that the conjunction of any two predicates invariant under f results in a predicate invariant under f. By Lemma 7 applying k-quantification to a predicate invariant under f gives again a predicate invariant under f. Hence, $\langle \Gamma \rangle_k \subseteq \mathsf{Inv}\, \mathrm{m}(k)\text{-}\mathsf{pPol}(\Gamma)$. Moreover, it follows that $\mathsf{Inv}\, \mathrm{m}(k)\text{-}\mathsf{pPol}(\Gamma) = \mathsf{Inv}\, \mathrm{m}(k)\text{-}\mathsf{pPol}(\langle \Gamma \rangle_k)$.

To establish the reverse inclusion, for any ℓ-ary relation $R \in \mathsf{Inv}\, \mathrm{m}(k)\text{-}\mathsf{pPol}(\Gamma)$ we define a relation Q as follows. Let $R = \{\mathbf{a}_1, \ldots, \mathbf{a}_t\}$. We consider sequences $(B_1, \ldots, B_t)$ of k-element subsets of D. Let also $(B_1^1, \ldots, B_t^1), \ldots, (B_1^r, \ldots, B_t^r)$ be a list of all such sequences. Then Q is the union of relations given by

$$\mathbf{a}_j \times \underbrace{B_j^1 \times \ldots \times B_j^1}_{k \text{ times}} \times \ldots \times \underbrace{B_j^r \times \ldots \times B_j^r}_{k \text{ times}},$$

for all $j \in [t]$. We show that there is $S \in \langle \Gamma \rangle_k$ such that $Q \subseteq S$ and $\mathrm{pr}_{[\ell]} S = R$. Then applying k-quantifications to all coordinates of S except for the first ℓ we obtain that $R \in \langle \Gamma \rangle_k$.

Let us consider the relation $S = \bigcap \{Q' \in \langle \Gamma \rangle_k \mid Q \subseteq Q'\}$. Since $\langle \Gamma \rangle_k$ is closed under conjunctions and contains the total relation $D^{\ell + kr}$, we have $S \in \langle \Gamma \rangle_k$ and $Q \subseteq S$.

Now choose any tuple $\mathbf{b} = (b_1, \ldots, b_\ell, d_1, \ldots, d_{kr}) \in S$. There are sets $C_1, \ldots, C_r$ such that $|C_i| = k$, for any $j \in [r]$, $C_{k(j-1)+1} = \ldots = C_{kj}$, $d_i \in C_i$, and for any $d'_i \in C_i$, $i \in [kr]$, the tuple $(b_1, \ldots, b_\ell, d'_1, \ldots, d'_{kr})$, since otherwise we can obtain a smaller relation S' containing Q, by applying a sequence of k-quantifications, followed by a conjunction with the total relation $D^{\ell+kr}$. Therefore we can choose $\mathbf{b}$ such that for any $j \in [r]$ all the values $d_{k(j-1)+1}, \ldots, d_{kj}$ are distinct, and $\{d_{k(j-1)+1}, \ldots, d_{kj}\} = C_{kj}$.

Since $\langle \Gamma \rangle_k$ is closed under conjunctions, by the Fleischer and Rosenberg result [12] it satisfies $\langle \Gamma \rangle_k = \mathsf{Inv}\, \mathsf{pPol}(\langle \Gamma \rangle_k)$. Moreover, by the proof of Theorem 2 of [12] S is the set of all tuples of the form $f(\mathbf{c}_1, \ldots, \mathbf{c}_n)$ for $n \geq 1$, $\mathbf{c}_1, \ldots, \mathbf{c}_n \in Q$, and $f \in \mathsf{pPol}(\langle \Gamma \rangle_k)$. Therefore there exist $n \geq 1$, $\mathbf{c}_1, \ldots, \mathbf{c}_n \in Q$ and $f \in \mathsf{pPol}(\langle \Gamma \rangle_k)$ such that $\mathbf{c} = f(\mathbf{c}_1, \ldots, \mathbf{c}_n)$. Let $\mathrm{pr}_{[\ell]} \mathbf{c}_q = \mathbf{a}_{i_q}$. For any selection $E_1, \ldots, E_n$ of k-element subsets of D there is $j \in [r]$ such that $E_q = B_{i_q}^j$ for $q \in [n]$. By the choice of $\mathbf{c}$ the range of f on $E_1 \times \ldots \times E_n = B_{i_1}^j \times \ldots \times B_{i_n}^j$ contains C_{kj}. Hence f is k-subset surjective, and so $f \in \mathsf{Inv}\, \mathrm{m}(k)\text{-}\mathsf{pPol}(\Gamma)$, as it is equal to $\mathsf{Inv}\, \mathrm{m}(k)\text{-}\mathsf{pPol}(\langle \Gamma \rangle_k)$. Therefore R is invariant under f, and so $(b_1, \ldots, b_\ell) \in R$. Relation S satisfies the required conditions, which completes the proof.

Proofs in parts (b), and (c) are quite similar. ■

REFERENCES

[1] L. Barto and M. Kozik. Constraint satisfaction problems of bounded width. In *FOCS*, 2010.
[2] V.G. Bodnarchuk, L.A. Kaluzhnin, V.N. Kotov, and B.A. Romov. Galois theory for post algebras. i. *Kibernetika*, 3:1–10, 1969.
[3] F. Börner, A.A. Bulatov, H. Chen, P. Jeavons, and Andrei A. Krokhin. The complexity of constraint satisfaction games and qcsp. *Inf. Comput.*, 207(9):923–944, 2009.
[4] A.A. Bulatov. Tractable conservative constraint satisfaction problems. In *LICS* 2003, pages 321–330 .
[5] A.A. Bulatov. A dichotomy theorem for constraint satisfaction problems on a 3-element set. *J. ACM*, 53(1):66–120, 2006.
[6] A.A. Bulatov. The complexity of the counting constraint satisfaction problem. In *ICALP (1)*, pages 646–661, 2008.
[7] A.A. Bulatov. Counting problems and clones of functions. In *ISMVL*, pages 1–6, 2009.
[8] A.A. Bulatov and V. Dalmau. Towards a dichotomy theorem for the counting constraint satisfaction problem. *Inf. Comput.*, 205(5):651–678, 2007.
[9] A.A. Bulatov, P. Jeavons, and A.A. Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM J. Comput.*, 34(3):720–742, 2005.
[10] H. Chen. The complexity of quantified constraint satisfaction: Collapsibility, sink algebras, and the three-element case. *SIAM J. Comput.*, 37(5):1674–1701, 2008.
[11] A. Dawar and D. Richerby. The power of counting logics on restricted classes of finite structures. In *CSL*, pages 84–98, 2007.
[12] I. Fleischner and I.G. Rosenberg. The Galois connection between partial operations and relations. *Pacific J. Math.*, 79:93–97, 1978.
[13] D. Geiger. Closed systems of function and predicates. *Pacific Journal of Mathematics*, pages 95–100, 1968.
[14] L. Haddad and I.G. Rosenberg. Partial clones containing all permutations. *Bull. Austral. Math. Soc.*, 52:263–278, 1995.
[15] P.M. Idziak, P. Markovic, R. McKenzie, M. Valeriote, and R. Willard. Tractability and learnability arising from algebras with few subpowers. *SIAM J. Comput.*, 39(7):3023–3037, 2010.
[16] P.G. Jeavons, D.A. Cohen, and M. Gyssens. Closure properties of constraints. *Journal of the ACM*, 44:527–548, 1997.
[17] U. Montanari. Networks of constraints: Fundamental properties and applications to picture processing. *Information Sciences*, 7:95–132, 1974.
[18] R. Pöschel and L.A. Kalužnin. *Funktionen- und Relationenalgebren.* DVW, Berlin, 1979.
[19] I.G. Rosenberg. Über die funktionale Vollständikeit in dem mehrwertigen Logiken. *Rozpravy Čs. Akademie Věd. Ser. Math. Nat. Sci.*, 80:3–93, 1970.