Affine Systems of Equations and Counting Infinitary Logic

Albert Atserias Andrei Bulatov Anuj Dawar

January 29, 2007

Abstract

We consider the definability of constraint satisfaction problems (CSP) in various fixed-point and infinitary logics. We show that testing the solvability of systems of equations over a finite Abelian group, a tractable CSP that was previously known not to be definable in Datalog, is not definable in an infinitary logic with counting. This implies that it is not definable in least fixed point logic or its extension with counting. We relate definability of CSPs to their classification obtained from tame congruence theory of the varieties generated by the algebra of polymorphisms of the template structure. In particular, we show that if this variety admits either the unary or affine type, the corresponding CSP is not definable in the infinitary logic with counting. We also study the complexity of determining whether a CSP omits both unary and affine types.

1 Introduction

The classification of constraint satisfaction problems (CSP) according to their tractability has been a major research goal since Feder and Vardi first formulated their dichotomy conjecture [12]. This classification has been closely linked to logic, with definability in Datalog providing one important uniform explanation for tractability. However, it has long been known that there are tractable CSPs, such as the satisfiability of systems of linear equations over finite fields, which are not definable in Datalog. Bulatov [5] (see also [4]) provides a uniform explanation for the tractability of these by showing that any constraint language that has a Mal'tsev polymorphism is solvable in polynomial time. It has remained an open question, however, whether there is an explanation for the tractability of these CSPs in terms of a natural logic whose data complexity is in polynomial time and which can define these problems.

The general form of the *constraint satisfaction problem* takes as instance two finite relational structures $\mathbf{A}$ and $\mathbf{B}$ and asks if there is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. We think of the elements of $\mathbf{A}$ as the variables of the problem and the universe of $\mathbf{B}$ as the domain of values which these variables may take. The individual tuples in the relations of $\mathbf{A}$ act as constraints on the values that must be matched to the relations holding in $\mathbf{B}$. The general form of the problem is NP-complete [26, 27]. In this paper we are mainly concerned with the non-uniform version of the problem which gives rise, for each fixed finite structure $\mathbf{B}$ to a different decision problem that we denote $\mathrm{CSP}(\mathbf{B})$, namely the problem of deciding whether a given $\mathbf{A}$ maps homomorphically to $\mathbf{B}$. For many fixed $\mathbf{B}$, this problem is solvable in polynomial time, while for others it remains NP-complete. A classification of structures for which the problem is tractable remains a major goal of research in the area.

In the present paper we are concerned with classifying constraint satisfaction problems according to their definability in a suitable logic. This is an approach that has proved useful in studying the tractability of constraint satisfaction problems [12, 8, 22]. In particular, it is known that many natural constraint satisfaction problems that are tractable are definable (or, to be precise, their complements are definable) in Datalog, the language of function-free Horn clauses. Any class of structures that is definable in Datalog is necessarily decidable in polynomial time, but there are known constraint satisfaction problems that are tractable but are not definable in Datalog. A classical example is the solvability of systems of linear equations over the two-element field [12], which we denote $\mathrm{CSP}(\mathbb{Z}_2)$. Furthermore, there are NP-complete constraint satisfaction problems, such as 3-colourability of graphs, which we can show are not Datalog-definable, without requiring the assumption that P is different from NP. Indeed, the class of constraint satisfaction problems whose complements are definable in Datalog appears to be a robust, natural class of problems with many independent and equivalent characterisations [9, 21].

A natural question arising from such considerations is whether we can offer any explanation based on logical definability for the tractability of problems such as the satisfiability of systems of linear equations over a finite field. Is there a natural logic such that all problems definable in this logic are polynomial-time decidable and that can express $\mathrm{CSP}(\mathbb{Z}_2)$? In particular, is this problem definable in LFP—the logic extending first-order logic with least fixed points or LFP + C—the extension of LFP with counting? These are both logics that have been extensively studied in the context of descriptive complexity as characterising natural fragments of polynomial time. Interestingly, Blass, Gurevich and Shelah [1] proved that LFP + C is able to define the class of non-singular square matrices over any fixed finite field, so it would not be very surprising if this logic were able to express $\mathrm{CSP}(\mathbb{Z}_2)$. Despite this, it is a consequence of our results that neither of these logics is able to express the solvability of systems of linear equations over any finite field. Indeed, we show that these problems are not definable in $\mathrm{C}^{\omega}_{\infty\omega}$, the infinitary logic with bounded number of variables and counting, a logic much more expressive than LFP + C. Combined with the result of Blass, Gurevich and Shelah about non-singular matrices, our result exhibits a fine-grained distinction between the problem of computing the rank of a square matrix and the problem of computing its determinant.

Another important means of classifying constraint satisfaction problems is on the basis of the algebra of the template structure $\mathbf{B}$. A polymorphism of a structure is an operation of its universe that preserves all its relations (see Section 2 for precise definitions). It is known that whether or not $\mathrm{CSP}(\mathbf{B})$ is tractable depends only on the algebra $\mathcal{B}$ obtained from the universe of $\mathbf{B}$ endowed with its polymorphisms. Indeed, it depends only on the variety generated by this algebra. This is established in [3] by showing that if the algebra $\mathcal{B}'$ of structure $\mathbf{B}'$ is obtained from $\mathcal{B}$ as a power, subalgebra or homomorphic image, then $\mathrm{CSP}(\mathbf{B}')$ is polynomial-time reducible to $\mathrm{CSP}(\mathbf{B})$. We show in the present paper that this can be improved to Datalog-definable reductions. These are weak reductions that, in particular, preserve definability in LFP and $\mathrm{C}^{\omega}_{\infty\omega}$. This allows us to establish that definability of a CSP in these logics is also determined by $\mathrm{var}(\mathcal{B})$, the variety generated by the algebra of $\mathbf{B}$.

Using the tool of Datalog-reductions, which we expect to be useful for other applications in the area, we relate definability of constraint satisfaction problems in $\mathrm{C}^{\omega}_{\infty\omega}$ to the classification of varieties of finite algebras from tame congruence theory [16]. It is known [3] that $\mathrm{CSP}(\mathbf{B})$ is NP-complete if $\mathrm{var}(\mathcal{B})$ admits the unary type (also known as type **1**), and it is conjectured that $\mathrm{CSP}(\mathbf{B})$ is in P otherwise. Similarly, Larose and Zadori showed [24] that $\mathrm{CSP}(\mathbf{B})$ is not definable in Datalog if $\mathrm{var}(\mathcal{B})$ admits the unary or affine types (types **1** and **2**), and conjectured the converse. It is a consequence of our results that we can strengthen the assertion by replacing Datalog with $\mathrm{C}^{\omega}_{\infty\omega}$. This implies that, if the Larose-Zadori conjecture is true, we obtain a dichotomy of definability whereby, for every $\mathbf{B}$, either $\mathrm{CSP}(\mathbf{B})$ is definable in Datalog or it is not definable in $\mathrm{C}^{\omega}_{\infty\omega}$.

Finally, we consider the meta-problems of deciding, given a structure $\mathbf{B}$ or an algebra $\mathcal{B}$ whether or not $\mathrm{var}(\mathcal{B})$ omits the unary and affine types. For algebras, the problem was shown decidable in polynomial time in [23], while for structures we show it is NP-complete.

The rest of the paper is structured as follows. In Section 2 we present some background definitions. Section 3 gives a proof that solvability of linear equations is not definable in $\mathrm{C}^{\omega}_{\infty\omega}$. Section 4 establishes that the definability of $\mathrm{CSP}(\mathbf{B})$ is determined by the variety generated by the algebra of $\mathbf{B}$. Section 5 shows that if the variety admits the unary or affine type, then it contains an algebra with the operations of a module.

These results are tied together in Section 6 to obtain the main conclusion relating definability in $\mathrm{C}^{\omega}_{\infty\omega}$ to the omitting of types from tame congruence theory. Section 7 gives the complexity results for the meta-problem. Proofs that are omitted or abbreviated in the main text can be found in the appendix.

2 Preliminaries

Structures and graphs A vocabulary σ is a finite collection of relation symbols, each with an associated arity. A σ-structure $\mathbf{A}$ consists of a finite set A with a relation $R^{\mathbf{A}} \subseteq A^r$ for each r-ary relation symbol R in σ. A graph is a structure with a binary relation that is symmetric and irreflexive. A *homomorphism* from a σ-structure $\mathbf{A}$ to a σ-structure $\mathbf{B}$ is a map $h : A \to B$ such that for each R in σ and each $\mathbf{a} \in A^r$, if $\mathbf{a} \in R^{\mathbf{A}}$ then $h(\mathbf{a}) \in R^{\mathbf{B}}$. We write $\mathbf{A} \to \mathbf{B}$ to denote that there exists a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. We write $\mathrm{CSP}(\mathbf{B})$ for the class of finite structures $\mathbf{A}$ such that $\mathbf{A} \to \mathbf{B}$ and also for the decision problem of determining membership in this class.

For the standard definition of the treewidth of a graph, we refer the reader to [10]. In our proofs we will use the following alternative characterization in terms of the *cops and robber game* [29]. The game is played by two players, one of whom controls the set of k cops attempting to catch a robber controlled by the other player. The cop player can move any set of cops to any vertices of the graph, while the robber can move along any path in the graph as long as there is no cop currently on the path. It is known [29] that the cop player has a winning strategy on a graph using $k+1$ cops if and only if the graph has treewidth at most k. The treewidth of a graph G is denoted $tw(G)$.

Logic A formula of first-order logic is said to be *positive primitive* if it is formed from the atomic formulas using only conjunctions and existential quantification. A formula is *existential positive* if it is formed from the atomic formulas using conjunctions, disjunctions and existential quantification. Datalog can be seen as the extension of existential positive formulas with a recursion mechanism. Similarly, LFP is the extension of first-order logic with an operator for forming the least fixed points of positive formulas. Finally, $\mathrm{LFP}+\mathrm{C}$ is the extension of LFP with a counting mechanism. For formal definitions, which we will not need in this paper, we refer the reader to [25]. It is known that every class of structures definable in $\mathrm{LFP}+\mathrm{C}$ is decidable in polynomial time.

The formulas of the logic $C_{\infty\omega}$ are obtained from the atomic formulas using negation, infinitary conjunction and disjunction, and counting quantifiers ($\exists^i x\phi$ for any integer $i \geq 0$). The fragment $\mathrm{C}^k_{\infty\omega}$ consists of those formulas of $C_{\infty\omega}$ in which only k distinct variables appear and $\mathrm{C}^{\omega}_{\infty\omega} = \bigcup_{k\in\omega} \mathrm{C}^k_{\infty\omega}$. The significance of $\mathrm{C}^{\omega}_{\infty\omega}$ is that fixed-point logics can be translated into it. That is, any formula of Datalog or LFP, and indeed of $\mathrm{LFP}+\mathrm{C}$ is equivalent to one of $\mathrm{C}^{\omega}_{\infty\omega}$. Moreover, these translations into infinitary logics have provided some of the most effective tools for proving inexpressibility results for the fixed-point logics. See [11, 17] for a discussion of this and the role of these logics in descriptive complexity.

The expressive power of $\mathrm{C}^{\omega}_{\infty\omega}$ is characterised by a game known as the *bijective game* [14]. This is played by two players, Spoiler and Duplicator, on a pair of structures $\mathbf{A}$ and $\mathbf{B}$, with k pairs of pebbles (x_i, y_i) for $1 \leq i \leq k$. For each move, Spoiler chooses a pair of pebbles (x_i, y_i), Duplicator chooses a bijection $f : A \to B$ such that $f(x_j) = y_j$ for $i \neq j$, and Spoiler chooses $a \in A$ and places x_i on a and y_i on $f(a)$. If, after some move, the map $\mathbf{x} \mapsto \mathbf{y}$ is not a partial isomorphism, Spoiler wins; Duplicator wins infinite plays. By a result of Hella [14], Duplicator has a winning strategy if, and only if, $\mathbf{A}$ and $\mathbf{B}$ cannot be distinguished by any formula of $\mathrm{C}^k_{\infty\omega}$, a fact denoted by $\mathbf{A} \equiv^{\mathcal{C}^k} \mathbf{B}$.

Universal algebra An n-ary operation f on a set A is a *polymorphism* of a relation $R \subseteq A^r$ if, for any tuples $\mathbf{a}_1, \dots \mathbf{a}_n \in R$, the tuple $f(\mathbf{a}_1, \dots \mathbf{a}_n)$ obtained by applying f component-wise also belongs to R.

We say that R is invariant under f. The following theorem links polymorphisms and definability of relations by positive primitive formulas (pp-formulas).

Theorem 1 ([13, 2, 19]). *Let* $\mathbf{A}$ *be a finite structure, and let* $R \subseteq A^r$ *be a non-empty relation that is preserved by all polymorphisms of* $\mathbf{A}$. *Then* R *is definable in* $\mathbf{A}$ *by a pp-formula.*

In [18, 20], Jeavons et al. proved that the set of polymorphisms of $\mathbf{B}$ is included in the set of polymorphisms of $\mathbf{A}$, then $\mathrm{CSP}(\mathbf{A})$ is log-space reducible to $\mathrm{CSP}(\mathbf{B})$. Therefore the complexity of non-uniform CSPs is completely determined by the set of polymorphisms of the corresponding relational structures.

A set with a collection of operations on it is called an *algebra*. Every structure $\mathbf{A}$ can be naturally associated with an algebra $\mathrm{Al}(\mathbf{A})$, called the *algebra* of $\mathbf{A}$, whose base set is the universe of $\mathbf{A}$, and whose operations are the polymorphisms of $\mathbf{A}$. A *variety* is a class of algebras which, if it contains $\mathcal{A}$ also contains every subalgebra of $\mathcal{A}$, every homomorphic image of $\mathcal{A}$, and every direct power of $\mathcal{A}$. The smallest variety containing $\mathcal{A}$ is called the *variety generated* by $\mathcal{A}$ and denoted by $\mathrm{var}(\mathcal{A})$. For further background on universal algebra, see [7].

3 Definability of Equations

In this section we show that the problem of determining the solvability of linear equations over the two-element field, which we mentioned above as a canonical example of a tractable CSP whose complement is not definable in Datalog, is also not definable in $\mathrm{C}^{\omega}_{\infty\omega}$. Indeed, we prove a more general result by showing that the solvability of equations over a finite Abelian group $\mathcal{G}$ with at least two elements is not definable in $\mathrm{C}^{\omega}_{\infty\omega}$. In the following we will write $+$ for the group operation in $\mathcal{G}$ and 0 for the identity.

Consider the following formulation of the problem.

Definition 2. *Let* $\mathcal{G}$ *be a finite Abelian group over a set* G *and* r *be a positive integer. We define the structure* $\mathbf{E}_{\mathcal{G},r}$ *to have universe* G *and, for each* $a \in G$ *and* $1 \leq j \leq r$, *it has a relation* R^j_a *of arity* j *that consists of the set of tuples* $(x_1, \ldots, x_j) \in G^j$ *that satisfy the equation* $x_1 + \cdots + x_j = a$.

Thus, any structure $\mathbf{A}$ in the signature of $\mathbf{E}_{\mathcal{G},r}$ can be seen as a set of equations in which at most r variables occur in each equation. The universe of $\mathbf{A}$ is the set of variables and the occurrence of a tuple $(x_1, \ldots, x_j)$ in a relation R^j_a signifies the equation $x_1 + \cdots + x_j = a$. This set of equations is solvable if, and only if, $\mathbf{A} \to \mathbf{E}_{\mathcal{G},r}$. In the sequel we will say "the equation $x_1 + \cdots + x_j = a$ occurs in $\mathbf{A}$" to mean that the tuple $(x_1, \ldots, x_j)$ is in R^j_a.

Our aim now is to exhibit, for each non-trivial finite Abelian group $\mathcal{G}$ and each positive integer k, a pair of structures $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{A} \equiv^{\mathcal{C}^k} \mathbf{B}$ and such that $\mathbf{A} \in \mathrm{CSP}(\mathbf{E}_{\mathcal{G},3})$ and $\mathbf{B} \notin \mathrm{CSP}(\mathbf{E}_{\mathcal{G},3})$. This will show that $\mathrm{CSP}(\mathbf{E}_{\mathcal{G},3})$ is not definable in $\mathrm{C}^{\omega}_{\infty\omega}$. This, of course, implies the result for all $\mathrm{CSP}(\mathbf{E}_{\mathcal{G},r})$ with $r \geq 3$. The structures we construct are sets of equations derived from 3-regular graphs of large treewidth. From now on, fix a non-trivial finite Abelian group $\mathcal{G}$, a 3-regular graph H, and a distinguished vertex u of H. Let $\{a_1, \ldots, a_m\}$ be the elements of $\mathcal{G}$. We define, for each $a \in \{a_1, \ldots, a_m\}$, a set of equations $\mathbf{E}_a H^u$ as follows (note that $\mathbf{E}_a H^u$ is a structure over the vocabulary of $\mathbf{E}_{\mathcal{G},3}$):

For each vertex $v \in V^H$ and each edge $e \in E^H$ that is incident on v, we have m distinct variables $x^{v,e}_i$ where i ranges over $\{a_1, \ldots, a_m\}$. Since each vertex has three edges incident on it, there are $3m$ variables associated to each vertex. For every vertex v *other than* u, let e_1, e_2, e_3 be the three edges incident on v. We then include the following equation in $\mathbf{E}_a H^u$ for all $i, j, k \in \{a_1, \ldots, a_m\}$:

$$x^{v,e_1}_i + x^{v,e_2}_j + x^{v,e_3}_k = i + j + k. \tag{1}$$

For the distinguished vertex u, instead of the above, we include the following equation, again for all $i, j, k \in \{a_1, \ldots, a_m\}$:

$$x_i^{u,e_1} + x_j^{u,e_2} + x_k^{u,e_3} = i + j + k + a. \tag{2}$$

In addition, for each edge $e \in E^H$ let v_1, v_2 be its endpoints. We include the following equations in $\mathbf{E}_a H^u$ for all $i, j \in \{a_1, \ldots, a_m\}$:

$$x_i^{v_1,e} + x_j^{v_2,e} = i + j. \tag{3}$$

We refer to equations of the form (1) and (2) as *vertex equations* and equations of the form (3) as *edge equations*.

Lemma 3. *$\mathbf{E}_a H^u$ is satisfiable if, and only if, $a = 0$*

Proof. To see that the system of equations $\mathbf{E}_0 H^u$ is satisfiable, just take the assignment that gives the variable $x_i^{v,e}$ the value i. To see that $\mathbf{E}_a H^u$ is unsatisfiable when $a \neq 0$, consider the subsystem S_0 of equations involving only the variables $x_0^{v,e}$ with subscript 0. Note that each such variable occurs exactly twice in S_0, once in a vertex equation and once in an edge equation. Thus, if we add up the left hand sides of all equations in S_0, we get $2\sum x_0^{v,e}$. Note also that each variable $x_0^{v,e}$ has a companion variable $x_0^{v',e}$ where v' is the other endpoint of the edge e and we have the equation $x_0^{v,e} + x_0^{v',e} = 0$. Thus

$$2\sum_{v,e} x_0^{v,e} = 2\sum_{e}(x_0^{v,e} + x_0^{v',e}) = 0.$$

On the other hand, the right-hand side of all equations is 0 except for the one vertex equation for u, which has right-hand side a. Thus summing the right-hand sides of all equations gives the sum a. Since $a \neq 0$, this shows that the subsystem S_0 and hence the system of equations $\mathbf{E}_a H^u$ is unsatisfiable. $\square$

Lemma 4. *If $u, u' \in V^H$ belong to the same connected component of H, then $\mathbf{E}_a H^u \cong \mathbf{E}_a H^{u'}$.*

Lemma 5. *If $\mathrm{tw}(H) > k$ and H is connected, then $\mathbf{E}_0 H^u \equiv^{\mathcal{C}^k} \mathbf{E}_a H^u$ for any $a \in \mathcal{G}$.*

Proof. Our aim is to exhibit a winning strategy for Duplicator in the k-pebble bijective game played on the two structures $\mathbf{A} = \mathbf{E}_0 H^u$ and $\mathbf{B} = \mathbf{E}_a H^u$. Since $\mathrm{tw}(H) > k$, we know that in the k cops and robber game played on H, robber has a winning strategy and Duplicator will make use of this strategy.

For each vertex $v \in V^H$ let X^v denote the set of variables $x_i^{v,e}$ for edges e incident on v. Similarly, for each $e \in E^H$, let X^e denote the set of variables involving e.

We say that a bijection $f : \mathbf{A} \to \mathbf{B}$ is *good* for a vertex $v \in V^H$ if the following conditions hold:

1. for all $w \in V^H$, $fX^w = X^w$;
2. for all $e \in E^H$, $fX^e = X^e$;
3. for all x, y, if $x + y = i$ is an equation in $\mathbf{A}$ then $f(x) + f(y) = i$ is an equation in $\mathbf{B}$; and
4. for all x, y, z, if $x + y + z = i$ is an equation in $\mathbf{A}$, then
   - $f(x) + f(y) + f(z) = i$ is an equation in $\mathbf{B}$ if $x, y, z \notin X^v$; and
   - $f(x) + f(y) + f(z) = i + a$ is an equation in $\mathbf{B}$ if $x, y, z \in X^v$.

Observe that the identity is a bijection that is good for u. Also, observe that a bijection that is good for v preserves all equations except the vertex equations for v.

Claim 6. *Given a bijection $f : \mathbf{A} \to \mathbf{B}$ that is good for v, if there is a path in H from v to w avoiding $u_1, \ldots, u_k$ then there is a bijection $f' : \mathbf{A} \to \mathbf{B}$ that is good for w such that $f|_{(X^{u_1} \cup \cdots \cup X^{u_k})} = f'|_{(X^{u_1} \cup \cdots \cup X^{u_k})}$.*

Proof. : Let the path from v to w avoiding $u_1, \ldots, u_k$ be $v = v_1, \ldots, v_n = w$. For each edge $e = \{v_i, v_{i+1}\}$ along this path, write x_j^{e-} for the variable $x_j^{v_i,e}$ and x_j^{e+} for the variable $x_j^{v_{i+1},e}$. We then define f' by $f'(x_j^{e-}) = f(x_{j-a}^{e-})$ and $f'(x_j^{e+}) = f(x_{j+a}^{e+})$; and f' agrees with f everywhere else. In particular, since the path from v to w avoids $u_1, \ldots, u_k$, f' agrees with f on $X^{u_1} \cup \cdots \cup X^{u_k}$. □

We now describe Duplicator's winning strategy in the bijective k-pebble game. Duplicator responds to Spoiler's first move with the identity bijection. She maintains a board on the side which describes a position in the k cops and robber game played on the graph H. At any point in the game, if Spoiler's pebbles are on the position $x_1, \ldots, x_k$ in $\mathbf{A}$ and $v_1, \ldots, v_k$ are the vertices of H to which these variables correspond, then the current position of the cops and robber game has k cops sitting on the vertices $v_1, \ldots, v_k$. If the robber's position according to its winning strategy is v, then Duplicator will play a bijection that is good for v.

To see that Duplicator can do this forever, suppose Spoiler lifts a pebble from x_i. Duplicator responds with a current bijection f that is good for v. Since the only equations not preserved by f are those associated with the vertex v, Spoiler must place at least three pebbles on variables associated with v to win the game. However, Duplicator responds to Spoiler placing the pebble on a new position x'_i by updating the position of the cops and robber game. Suppose robber's winning strategy dictates that the robber move from v to w. Since robber's move must be along a path avoiding the current cop positions, by Claim 6, Duplicator can update the bijection f to a new f' that is good for w without changing f on any of the currently pebbled positions. It is now clear that Duplicator can play forever. □

Theorem 7. *Let $\mathcal{G}$ be a non-trivial finite Abelian group. Then $\mathrm{CSP}(\mathbf{E}_{\mathcal{G},3})$ is not definable in $\mathrm{C}^{\omega}_{\infty\omega}$*

Proof. Suppose, to the contrary, that there is a k such that $\mathrm{CSP}(\mathbf{E}_{\mathcal{G},3})$ is definable in $\mathrm{C}^{k}_{\infty\omega}$. Let H be any connected, 3-regular graph with $\mathrm{tw}(H) > k$ and u any vertex of H. For instance, H could be a sufficiently large brick graph. Let a be any element of $\mathcal{G}$ distinct from 0. Then, by Lemma 3, $\mathbf{E}_0 H^u \in \mathrm{CSP}(\mathbf{E}_{\mathcal{G},3})$ and $\mathbf{E}_a H^u \notin \mathrm{CSP}(\mathbf{E}_{\mathcal{G},3})$. But, by Lemma 5, $\mathbf{E}_0 H^u \equiv^{\mathcal{C}^k} \mathbf{E}_a H^u$, a contradiction. □

4 Logical Reductions

4.1 Definition

Let σ and $\tau = (R_1, \ldots, R_s)$ be two relational vocabularies. A *k-ary interpretation with p parameters of τ in σ* is an $(s+1)$-tuple $\mathbf{I} = (\varphi_U, \varphi_1, \ldots, \varphi_s)$ of formulas over the vocabulary τ, where $\varphi_U = \varphi_U(\mathbf{x}, \mathbf{y})$ has $k + p$ free variables $\mathbf{x} = (x^1, \ldots, x^k)$ and $\mathbf{y} = (y_1, \ldots, y_p)$, and $\varphi_i = \varphi_i(\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{y})$ has kr free variables where r is the arity of R_i and each $\mathbf{x}_j = (x_j^1, \ldots, x_j^k)$ and $\mathbf{y} = (y_1, \ldots, y_p)$.

Let $\mathbf{A}$ be a σ-structure. A tuple $\mathbf{c} = (a_1, \ldots, a_p)$ of of pairwise different points of $\mathbf{A}$ is called *proper*. The *interpretation of* $\mathbf{A}$ *through* $\mathbf{I}$ *with parameters* $\mathbf{c}$, denoted by $\mathbf{I}(\mathbf{A}, \mathbf{c})$, is the τ-structure whose universe is $\{\mathbf{a} \in A^k : \mathbf{A} \models \varphi_U(\mathbf{a}, \mathbf{c})\}$, and whose interpretation for R_i is the set of tuples $(\mathbf{a}_1, \ldots, \mathbf{a}_r) \in (A^k)^r$ such that $\mathbf{A} \models \varphi_U(\mathbf{a}_1, \mathbf{c}) \wedge \cdots \wedge \varphi_U(\mathbf{a}_r, \mathbf{c}) \wedge \varphi_i(\mathbf{a}_1, \ldots, \mathbf{a}_r, \mathbf{c})\}$. If each formula in $\mathbf{I}$ belongs to a class of formulas Θ, we say that $\mathbf{I}$ is a Θ-interpretation.

Now we are ready to define the notion of logical reduction:

Definition 8. *Let σ and τ be finite relational vocabularies, let $\mathcal{C}$ be a class of σ-structures, let $\mathcal{D}$ be a class of τ-structures that is closed under isomorphisms, and let Θ be a class of formulas. We say that a Θ-interpretation with p parameters $\mathbf{I}$ of τ in σ is a Θ-reduction from $\mathcal{C}$ to $\mathcal{D}$ if, for every σ-structure $\mathbf{A}$ with at least p elements, we have $\mathbf{A} \in \mathcal{C}$ if, and only if, $\mathbf{I}(\mathbf{A}, \mathbf{c}) \in \mathcal{D}$ for some proper $\mathbf{c}$.*

In case there is a reduction as in the definition, we say that $\mathcal{C}$ reduces to $\mathcal{D}$ under Θ-reductions, and write $\mathcal{C} \leq_{\Theta} \mathcal{D}$. We will use the collections of positive quantifier-free formulas, existential positive formulas, and datalog formulas (i.e. datalog programs) and write $\leq_{\mathrm{pqf}}$, $\leq_{\mathrm{ep}}$ and $\leq_{\mathrm{datalog}}$, respectively. Note that these are reductions of increasing power, and that definability in $\mathrm{C}^{\omega}_{\infty\omega}$ is preserved downwards by all three.

4.2 Expansions by reduced definable relations

Let A be a set and let $R \subseteq A^s$ be a relation on A. We define an equivalence relation $\theta(R)$ on $\{1, \ldots, s\}$ by setting $(i, j) \in \theta(R)$ if, and only if, $a_i = a_j$ for every $(a_1, \ldots, a_s) \in R$. We say that R is a *reduced* relation if $\theta(R)$ is the trivial equivalence relation (i.e. equality). Note that the equality relation on A is not reduced.

Lemma 9. *Let $\mathbf{B}$ be a finite structure, and let $\mathbf{D}$ be an expansion of $\mathbf{B}$ by a reduced relation R that is definable in $\mathbf{B}$ by a pp-formula. Then,* $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B})$.

4.3 Reduction to the reduced case

Piece of notation: Let $\mathbf{a} = (a_1, \ldots, a_m)$ be a sequence and let $I = (i_1, \ldots, i_r)$ be a sequence of indices, where $1 \leq i_j \leq m$ for every $j \in \{1, \ldots, r\}$. We write $\mathbf{a}_I$ for the sequence $(a_{i_1}, \ldots, a_{i_r})$. Now let R be a relation of arity s and I a sequence of indices from $\{1, \ldots, s\}$. Then $\mathrm{pr}_I R$ denotes the relation $\{\mathbf{a}_I : \mathbf{a} \in R\}$.

Let R be a relation of arity s and recall the definition of $\theta(R)$, the equivalence relation on $\{1, \ldots, s\}$ defined in the previous section. Let I be a set of representatives of the equivalence-classes of $\theta(R)$, ordered in an arbitrary way, and define $\mathrm{red}(R) = \mathrm{pr}_I R$. Note that $\mathrm{red}(R)$ does not depend on the choice of I. Besides, for every $i \notin I$ there exists some $j \in I$ such that $a_i = a_j$ for every tuple $(a_1, \ldots, a_s) \in R$. We call $\mathrm{red}(R)$ the reduced version of R. A reduced structure is a structure all whose relations are reduced. To every structure $\mathbf{B}$ we can associate a reduced structure, called the *reduced version of* $\mathbf{B}$, whose universe is the universe of $\mathbf{B}$ itself and whose relations are the reduced versions of the relations of $\mathbf{B}$. Note that the vocabularies of a structure and its reduced version may be different. Note that the polymorphisms of $\mathbf{B}$ and its reduced version are the same.

Lemma 10. *Let $\mathbf{B}$ a finite structure and let $\mathbf{D}$ be the reduced version of $\mathbf{B}$. Then* $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{D})$ *and* $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B})$.

Proof: For space constraints, we only sketch the reduction $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{D})$. Let $\mathbf{A}$ be an instance of $\mathrm{CSP}(\mathbf{B})$. We define an instance $\mathbf{C}$ of $\mathrm{CSP}(\mathbf{D})$. The universe of $\mathbf{C}$ is A itself. For the relations, the basic idea is to project every relation $R^{\mathbf{A}}$ to the coordinates of a set of representatives I of the θ-classes, where $\theta = \theta(R)$. However, before we do that, we need to *close* each $R^{\mathbf{A}}$ under all equalities implied by the equivalences $(i, j) \in \theta$. We do that using Datalog-definable intermediate relations.

So, let E be the binary relation on A defined by the following Datalog program:

$$\begin{aligned} E(x_i, x_j) &\ :-\ R(x_1, \ldots, x_s) \\ E(x, y) &\ :-\ E(y, x) \\ E(x, z) &\ :-\ E(x, y) \wedge E(y, z), \end{aligned}$$

where the first rule is introduced for every symbol R in σ and every $(i, j) \in \theta(R)$. It is obvious that E is an equivalence relation on A; reflexivity follows from the fact that $(i, j) \in \theta(R)$ in the first rule, symmetry is enforced by the second rule, and transitivity is enforced by the third. Next, for every r-ary symbol R in σ, let R' be the relation defined by $R'(\mathbf{x}_I)\ :-\ R(y_1, \ldots, y_s) \wedge E(x_1, y_1) \wedge \cdots \wedge E(x_s, y_s)$, where I is a set of representatives of the $\theta(R)$-classes ordered in an arbitrary way. This defines $\mathbf{C}$, and we defined it by a Datalog program interpreted on $\mathbf{A}$. It remains to argue that this datalog-interpretation is indeed a reduction. The argument may be found in the appendix. $\square$

4.4 Powering, subalgebras, and homomorphic images

In this subsection we show how the basic algebraic constructions of powering, subalgebra and homomorphic images can be handled by Datalog-reductions. We start with homomorphic images.

Let $\mathbf{B}$ be a finite structure and let $\mathcal{B}$ be its corresponding algebra. Suppose $\mathcal{B}'$ is an algebra that has a homomorphic image $\mathcal{A} = h(\mathcal{B}')$ that is a reduct of $\mathcal{B}$. Note that $A = B = h(B')$, i.e. the universes of $\mathcal{A}$ and $\mathcal{B}$ are the same and are the image of the universe of $\mathcal{B}'$ under h. We define a new structure $\mathbf{B}' = \mathrm{pre}(\mathbf{B}, h)$, the *preimage* of $\mathbf{B}$, whose universe is B' and whose relations are the preimages $h^{-1}(R^{\mathbf{B}})$ of the relations $R^{\mathbf{B}}$ of $\mathbf{B}$.

Lemma 11. *Let the algebras $\mathcal{B}$ and $\mathcal{B}'$, and the structures $\mathbf{B}$ and $\mathbf{B}' = \mathrm{pre}(\mathbf{B}, h)$ be as above. Then $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$ and $\mathcal{B}'$ is a reduct of $\mathrm{Al}(\mathbf{B}')$.*

Let $\mathbf{B}$ be a finite structure and let $\mathcal{B}$ be its corresponding algebra. Suppose $\mathcal{B}'$ is an algebra that has a subalgebra $\mathcal{A} \subseteq \mathcal{B}'$ that is a reduct of $\mathcal{B}$. Note that $A = B \subseteq B'$, i.e. the universes of $\mathcal{A}$ and $\mathcal{B}$ are the same and are a subset of the universe of $\mathcal{B}'$. We define a new structure $\mathbf{B}' = \mathrm{ext}(\mathbf{B}, B')$, the *extension* of $\mathbf{B}$, with universe B' and the same relations as $\mathbf{B}$.

Lemma 12. *Let the algebras $\mathcal{B}$ and $\mathcal{B}'$, and the structures $\mathbf{B}$ and $\mathbf{B}' = \mathrm{ext}(\mathbf{B}, B')$ be as above. Then $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$ and $\mathcal{B}'$ is a reduct of $\mathrm{Al}(\mathbf{B}')$.*

Let R be an r-ary relation on the set A^n. Then the *flattening* of R, denoted $\mathrm{fla}(R, n)$, is the rn-ary relation on A that contains all tuples $(x_1, \ldots, x_{rn})$ such that $((x_1, \ldots, x_n), \ldots, (x_{(r-1)n+1}, \ldots, x_{rn})) \in R$. Let $\mathbf{B}$ be a finite structure and let $\mathcal{B}$ be its corresponding algebra. Suppose $\mathcal{B}'$ is an algebra that has a direct power $\mathcal{A} = \mathcal{B}'^n$ that is a reduct of $\mathcal{B}$. Note that $A = B = B'^n$, i.e. the universes of $\mathcal{A}$ and $\mathcal{B}$ are the same and are the n-th power of the universe of $\mathcal{B}'$. We define a new structure $\mathbf{B}' = \mathrm{fla}(\mathbf{B}, n)$, the *flattening* of $\mathbf{B}$, whose universe is B and whose relations are the flattenings of the relations of $\mathbf{B}$.

Lemma 13. *Let the algebras $\mathcal{B}$ and $\mathcal{B}'$, and the structures $\mathbf{B}$ and $\mathbf{B}' = \mathrm{fla}(\mathbf{B}, n)$ be as above. Then $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$ and $\mathcal{B}'$ is a reduct of $\mathrm{Al}(\mathbf{B}')$.*

Finally, we are ready to state and prove the consequence of these Lemmas that we will be using.

Theorem 14. *Let $\mathbf{B}$ and $\mathbf{B}'$ be finite structures and let $\mathcal{B}$ and $\mathcal{B}'$ be their respective algebras. If some algebra of $\mathrm{var}(\mathcal{B}')$ is a reduct of $\mathcal{B}$, then $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{B}')$.*

Proof: Suppose that some algebra $\mathcal{A}$ of $\mathrm{var}(\mathcal{B}')$ is a reduct of $\mathcal{B}$. By the HSP-theorem [7, Theorem 9.5] $\mathcal{A}$ is a homomorphic image of a subalgebra of a direct power of $\mathcal{B}'$. Let $\mathcal{B}_p$, $\mathcal{B}_s$, and $\mathcal{B}_h$ be the direct power, its subalgebra, and the homomorphic image, respectively. We have $\mathcal{A} = \mathcal{B}_h$. Let n be such that $\mathcal{B}_p = \mathcal{B}'^n$, an let h be a homomorphism from $\mathcal{B}_s$ to $\mathcal{B}_h$.

We use three intermediate structures $\mathbf{B}_s = \mathrm{pre}(\mathbf{B}, h)$, $\mathbf{B}_p = \mathrm{ext}(\mathbf{B}_s, B_p)$, and $\mathbf{B}_f = \mathrm{fla}(\mathbf{B}_p, n)$ that, by the definition, have the universes of the algebras $\mathcal{B}_s$, $\mathcal{B}_p$, and $\mathcal{B}'$ respectively. By Lemma 11, $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}_s)$ and $\mathcal{B}_s$ is a reduct of $\mathrm{Al}(\mathbf{B}_s)$. By Lemma 12, $\mathrm{CSP}(\mathbf{B}_s) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}_p)$ and $\mathcal{B}_p$ is a reduct of $\mathrm{Al}(\mathbf{B}_p)$. By Lemma 13, $\mathrm{CSP}(\mathbf{B}_p) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}_f)$, and $\mathcal{B}'$ is a reduct of $\mathrm{Al}(\mathbf{B}_f)$. Now, let $\mathbf{D}$ be the reduced version of $\mathbf{B}_f$. Then $\mathrm{CSP}(\mathbf{B}_f) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{D})$ by Lemma 10. We prove that $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$ and the result will follow by composing.

Let $\mathbf{D}'$ be the expansion of $\mathbf{D}$ obtained by adding all the relations of $\mathbf{B}'$. Since $\mathcal{B}'$ is a reduct of the algebra of $\mathbf{B}_f$ and $\mathbf{D}'$ is the flattening of $\mathbf{B}_f$, it is straightforward that every relation of $\mathbf{D}'$ is invariant with respect to all polymorphisms of $\mathbf{B}'$. Therefore, by Theorem 1, every such relation is pp-definable in $\mathbf{B}'$, and since they are reduced, we have $\mathrm{CSP}(\mathbf{D}') \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$ by Lemma 9. It is also obvious that $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{D}')$ through the mapping that sends a structure to its expansion with empty relations. Composing we get $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$. $\square$

4.5 Reduction from the idempotent case

To every finite structure $\mathbf{B}$ we associate a new structure, the *singleton-expansion* of $\mathbf{B}$, by adding one unary relation $\{b\}$ for every $b \in B$. In other words, if $B = \{b_1, \ldots, b_n\}$, then the structure $(\mathbf{B}, \{b_1\}, \ldots, \{b_n\})$ is the singleton-expansion of $\mathbf{B}$. Note that the polymorphisms of the singleton-expansion of $\mathbf{B}$ are exactly the *idempotent* polymorphisms of $\mathbf{B}$, that is polymorphisms f satisfying the identity $f(x, \ldots, x) = x$. Indeed, every singleton set $\{b\}$ is preserved by any idempotent polymorphism of $\mathbf{B}$, and any polymorphism of $\mathbf{B}$ that preserves every singleton set $\{b\}$ must by idempotent.

Lemma 15. *Let $\mathbf{B}$ be a finite structure, and let $\mathbf{D}$ be the singleton-expansion of $\mathbf{B}$. Then $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{D})$ and if $\mathbf{B}$ is a core with at least two points, then $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{ep}} \mathrm{CSP}(\mathbf{B})$.*

5 Omitting types

Let $\mathcal{A}$ be an algebra. A *congruence* of $\mathcal{A}$ is an equivalence relation α that is invariant with respect to all operations of $\mathcal{A}$. In other words, for any (n-ary) operation f of $\mathcal{A}$ and any $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathcal{A}$ such that $(a_i, b_i) \in \alpha$ we have $(f(a_1, \ldots, a_n), f(b_1, \ldots, b_n)) \in \alpha$. The congruences of $\mathcal{A}$ form its *congruence lattice* $\mathrm{con}(\mathcal{A})$. A *prime quotient* in this lattice is a pair of congruences α, β such that $\alpha \leq \beta$, $\alpha \neq \beta$, and for any γ with $\alpha \leq \gamma \leq \beta$ we have either $\alpha = \gamma$, or $\beta = \gamma$. The fact that α, β is a prime quotient will be denoted by $\alpha \prec \beta$.

Tame congruence theory [16] allows one to assign to each prime quotient of the congruence lattice $\mathrm{con}(\mathcal{A})$ of a finite algebra $\mathcal{A}$ one of five types. The type reflects the local structure of the algebra, which can be one of the following: **1** a finite set with a group action on it (unary type), **2** a finite vector space over a finite field (affine type), **3** a two-element Boolean algebra, **4** a two-element lattice, **5** a two-element semilattice. We use tame congruence as a black box extracting properties we need from existing results, and we do not therefore need a precise definition of the types.

The type of a prime quotient $\alpha \prec \beta$ is denoted by $\mathsf{typ}(\alpha, \beta)$, while $\mathsf{typ}(\mathcal{A})$ denotes the set of types appearing as types of some prime quotient of $\mathcal{A}$. If $\mathfrak{A}$ is a class of algebras, $\mathsf{typ}(\mathfrak{A})$ denotes the set $\bigcup_{\mathcal{A} \in \mathfrak{A}} \mathsf{typ}(\mathcal{A})$. If $\mathbf{i} \notin \mathsf{typ}(\mathfrak{A})$, we say that $\mathfrak{A}$ *omits* type $\mathbf{i}$. Otherwise, we say $\mathfrak{A}$ *admits* type $\mathbf{i}$. We need the following:

Lemma 16. *Let $\mathcal{A}$ be a finite idempotent algebra. If $\mathrm{var}(\mathcal{A})$ admits types **1** or **2** then it contains a finite idempotent reduct of a module.*

Recall from Section 3 the definition of the structure $\mathbf{E}_{\mathcal{G},r}$ for every finite Abelian group $\mathcal{G}$ and every integer $r \geq 1$.

Lemma 17. *Let $\mathcal{M}$ be a finite module, let $\mathcal{G}$ be its underlying Abelian group, and let $\mathcal{A}$ be an idempotent reduct of $\mathcal{M}$. Then $\mathcal{A}$ is a reduct of the algebra of $\mathbf{E}_{\mathcal{G},r}$ for every $r \geq 1$.*

6 Results

We can bring together the results of Section 4 and 5 to establish the following theorem.

Theorem 18. *Let $\mathbf{B}$ be a finite structure and let $\mathcal{B}$ be its algebra. If $\mathrm{var}(\mathcal{B})$ admits the unary or affine types, then there exists a non-trivial finite Abelian group $\mathcal{G}$ such that $\mathrm{CSP}(\mathbf{E}_{\mathcal{G},r}) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{B})$ for every $r \geq 1$.*

Proof. Since $\mathrm{CSP}(\mathbf{B}) = \mathrm{CSP}(\mathrm{core}(\mathbf{B}))$, where $\mathrm{core}(\mathbf{B})$ is the core of $\mathbf{B}$, we may assume that $\mathbf{B}$ is a core. Let $\mathbf{D}$ be the singleton-expansion of $\mathbf{B}$ and let $\mathcal{D}$ be its algebra, which is idempotent. By Lemma 15, we have $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{B})$. Moreover, if $\mathrm{var}(\mathcal{B})$ admits types **1** or **2**, so does $\mathrm{var}(\mathcal{D})$ because $\mathcal{D}$ is a reduct of $\mathcal{B}$ (see [16, Chapter 5]). By Lemma 16, the variety $\mathrm{var}(\mathcal{D})$ contains a finite idempotent reduct $\mathcal{A}$ of a module. Let $\mathcal{G}$ be the Abelian group underlying the module. Then $\mathcal{G}$ is non-trivial and finite. Moreover, $\mathcal{A}$ is a reduct of the algebra of $\mathbf{E}_{\mathcal{G},r}$ for every $r \geq 1$ by Lemma 17. It follows that $\mathrm{CSP}(\mathbf{E}_{\mathcal{G},r}) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{D})$. Composing we get the result. $\square$

We have seen in Section 3 that $\mathrm{CSP}(\mathbf{E}_{\mathcal{G},3})$ is not definable in $\mathrm{C}^{\omega}_{\infty\omega}$ when $\mathcal{G}$ is non-trivial. Since definability in $\mathrm{C}^{\omega}_{\infty\omega}$ is preserved downwards by Datalog-reductions, this yields the following corollary:

Corollary 19. *Let $\mathbf{B}$ be a finite structure and let $\mathcal{B}$ be its algebra. If $\mathrm{CSP}(\mathbf{B})$ is definable in $\mathrm{C}^{\omega}_{\infty\omega}$, then $\mathrm{var}(\mathcal{B})$ omits the unary and affine types.*

Corollary 19 can be seen as a strengthening of the result of Larose and Zadori [24] that if the complement of $\mathrm{CSP}(\mathbf{B})$ is definable in Datalog then $\mathrm{var}(\mathcal{B})$ omits the unary and affine types. Larose and Zadori also conjectured the converse, namely that if $\mathrm{var}(\mathcal{B})$ omits the unary and affine types then the complement of $\mathrm{CSP}(\mathbf{B})$ is definable in Datalog. By Corollary 19 this conjecture would imply that every $\mathrm{CSP}(\mathbf{B})$ is either definable in Datalog or not definable in $\mathrm{C}^{\omega}_{\infty\omega}$, which can be seen as a definability dichotomy.

7 Testing omitting types

We consider three decision problems. In ALGEBRA OF TYPE 2 we are given a finite set A and operation tables of idempotent operations $f_1, \ldots, f_n$ on A, and the question is whether $\mathrm{var}(\mathcal{A})$, where $\mathcal{A} = (A; \{f_1, \ldots, f_n\})$, omits types **1** and **2**. In RELATIONAL STRUCTURE OF TYPE 2 we are given a finite relational structure $\mathbf{A}$, and the question is whether $\mathrm{var}(\mathrm{Al}(\mathbf{A}))$ omits types **1** and **2**. In RELATIONAL STRUCTURE OF TYPE $2(k)$ we are given a finite relational structure $\mathbf{A}$, $|A| \leq k$, and the question is whether $\mathrm{var}(\mathrm{Al}(\mathbf{A}))$ omits types **1** and **2**. The problems ALGEBRA OF TYPE 2 and RELATIONAL STRUCTURE OF TYPE $2(k)$ were shown tractable in [23].

Theorem 20. RELATIONAL STRUCTURE OF TYPE 2 *is NP-complete.*

References

[1] Andreas Blass, Yuri Gurevich, and Saharon Shelah. On polynomial time computation over unordered structures. *J. Symbolic Logic*, 67(3):1093–1125, 2002.

[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. [Russian].

[3] A. Bulatov, P. Jeavons, and A. Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM Journal on Computing*, 34:720–742, 2005.

[4] A. A. Bulatov and V. Dalmau. A simple algorithm for mal'tsev constraints. *SIAM J. Comput.*, 36(1):16–27, 2006.

[5] A.A. Bulatov. Mal'tsev constraints are tractable. Technical Report PRG-RR-02-05, Computing Laboratory, University of Oxford, Oxford, UK, 2002.

[6] A.A. Bulatov and P.G. Jeavons. Algebraic structures in combinatorial problems. Technical Report MATH-AL-4-2001, Technische universität Dresden, Dresden, Germany, 2001.

[7] S. Burris and H.P. Sankappanavar. *A course in universal algebra*, volume 78 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1981.

[8] V. Dalmau. Linear datalog and bounded path duality of relational structures. *Logical Methods in Computer Science*, 1(1), 2005.

[9] V. Dalmau, Ph.G. Kolaitis, and M.Y. Vardi. Constraint satisfaction, bounded treewidth, and finite variable logics. In *Proceedings of the 8th International Conference on Principles and Practice of Constraint Programming, CP'02*, Lecture Notes in Computer Science, pages 311–326. Springer-Verlag, 2002.

[10] R. Diestel. *Graph Theory*. Springer, 1997.

[11] H-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer, 2 edition, 1999.

[12] T. Feder and M.Y. Vardi. Computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory. *SIAM Journal of Computing*, 28:57–104, 1998.

[13] D. Geiger. Closed systems of function and predicates. *Pacific Journal of Mathematics*, pages 95–100, 1968.

[14] L. Hella. Logical hierarchies in PTIME. *Information and Computation*, 129:1–19, 1996.

[15] C. Herrman. Affine algebras in congruence-modular varieties. *Acta Sci. Math. (Szeged)*, 41:119–125, 1971.

[16] D. Hobby and R.N. McKenzie. *The Structure of Finite Algebras*, volume 76 of *Contemporary Mathematics*. American Mathematical Society, Providence, R.I., 1988.

[17] N. Immerman. *Descriptive Complexity*. Springer, 1999.

[18] P.G. Jeavons, D.A. Cohen, and M. Gyssens. Closure properties of constraints. *Journal of the ACM*, 44:527–548, 1997.

[19] P.G. Jeavons, D.A. Cohen, and M. Gyssens. How to determine the expressive power of constraints. *Constraints*, 4:113–131, 1999.

[20] P.G. Jeavons, D.A. Cohen, and J.K. Pearson. Constraints and universal algebra. *Annals of Mathematics and Artificial Intelligence*, 24:51–67, 1998.

[21] Ph.G. Kolaitis and M.Y Vardi. A game-theoretic approach to constraint satisfaction. In *Proc. 17th National Conference on Artificial Intelligence, AAAI-2000*, pages 175–181, 2000.

[22] B. Larose, C. Loten, and C. Tardif. A characterisation of first-order constraint satisfaction problems. In *LICS*, pages 201–210, 2006.

[23] B. Larose and M. Valeriote. personal communication, 2006.

[24] B. Larose and L. Zádori. Bounded width problems and algebras. Algebra Universalis, to appear.

[25] L. Libkin. *Elements of Finite Model Theory*. Springer, 2004.

[26] A.K. Mackworth. Consistency in networks of relations. *Artificial Intelligence*, 8:99–118, 1977.

[27] U. Montanari. Networks of constraints: Fundamental properties and applications to picture processing. *Information Sciences*, 7:95–132, 1974.

[28] E.L. Post. *The two-valued iterative systems of mathematical logic*, volume 5 of *Annals Mathematical Studies*. Princeton University Press, 1941.

[29] P. Seymour and R.Thomas. Graph searching and a min-max theorem for treewidth. *Journal of Combinatorial Theory, Series B*, 58:22–33, 1993.

A Proofs omitted from Section 3

Lemma 3. $\mathbf{E}_a H^u$ *is satisfiable if, and only if,* $a = 0$

Proof. To see that the system of equations $\mathbf{E}_0 H^u$ is satisfiable, just take the assignment that gives the variable $x_i^{v,e}$ the value i.

To see that $\mathbf{E}_a H^u$ is unsatisfiable when $a \neq 0$, consider the subsystem S_0 of equations involving only the variables $x_0^{v,e}$ with subscript 0. Note that each such variable occurs exactly twice in S_0, once in a vertex equation and once in an edge equation. Thus, if we add up the left hand sides of all the equations, we get $2\sum x_0^{v,e}$. Note also that each variable $x_0^{v,e}$ has a companion variable $x_0^{v',e}$ where v' is the other endpoint of the edge e and we have the equation $x_0^{v,e} + x_0^{v',e} = 0$. Thus

$$2\sum_{v,e} x_0^{v,e} = 2\sum_{e}(x_0^{v,e} + x_0^{v',e}) = 0.$$

On the other hand, the right-hand side of all equations is 0 except for the one vertex equation for u, which has right-hand side a. Thus summing the right-hand sides of all equations gives the sum a. Since $a \neq 0$, this shows that the subsystem S_0 and hence the system of equations $\mathbf{E}_a H^u$ is unsatisfiable. $\square$

Lemma 4. *For any* $u, u' \in V^H$ *that lie in the same connected component of* H, $\mathbf{E}_a H^u \cong \mathbf{E}_a H^{u'}$.

Proof. The case where $u = u'$ is trivial, so assume that they are distinct.

Let $u = v_1, e_1, \ldots, e_s, v_{s+1} = u'$ be the sequence of vertices and edges along a simple path from u to u'. We now define a map η from $\mathbf{E}_a H^u$ to $\mathbf{E}_a H^{u'}$ as follows:

- for any $v \notin \{v_1, \ldots, v_{s+1}\}$, $\eta(x_j^{v,e}) = x_j^{v,e}$;
- for each $l \in \{1, \ldots, s\}$, $\eta(x_j^{v_l,e_l}) = x_{j+a}^{v_l,e_l}$; and
- for each $l \in \{1, \ldots, s\}$, $\eta(x_j^{v_{l+1},e_l}) = x_{j-a}^{v_{l+1},e_l}$.

To show that η is an isomorphism, we need to argue that it preserves all the equations in $\mathbf{E}_a H^u$. Clearly, all equations corresponding to vertices and edges of H that do not appear on the path are preserved as η is the identity map on the corresponding variables. Consider now the vertex equations corresponding to the vertex u. Note that the edge e_1 (the first edge on the chosen path) is incident on u and let f and g be the two other edges incident on u. Then, the equation

$$x_i^{u,e_1} + x_j^{u,f} + x_k^{u,g} = i + j + k + a$$

is mapped by η to

$$x_{i+a}^{u,e_1} + x_j^{u,f} + x_k^{u,g} = i + j + k + a$$

which is, indeed, an equation of $\mathbf{E}_r H^{u'}$.

Similarly, a vertex equation for u':

$$x_i^{u',e_s} + x_j^{u',f} + x_k^{u',g} = i + j + k$$

is mapped to

$$x_{i-a}^{u',e_s} + x_j^{u,f} + x_k^{u,g} = i + j + k.$$

Now, consider a vertex equaton for an intermediate vertex $v = v_{l+1}$ along the path. In this case, there are two edges e_l, e_{l+1} of the path incident on v. Thus, the equation

$$x_i^{v,e_l} + x_j^{v,e_{l+1}} + x_k^{v,f} = i + j + k$$

is mapped by η to

$$x_{i-a}^{v,e_l} + x_{j+a}^{v,e_{l+1}} + x_k^{v,f} = i + j + k,$$

where f is the third edge incident on v.

Finally, for each edge e_l along the path, the equation

$$x_i^{v_l,e_l} + x_j^{v_{l+1},e_l} = i + j$$

is mapped by η to

$$x_{i+a}^{v_l,e_l} + x_{j-a}^{v_2,e} = i + j.$$

We have thus established that η maps equations to equations. Since η is a bijection, and the number of equations in $\mathbf{E}_a H^u$ and in $\mathbf{E}_a H^{u'}$ is the same, this proves that it is an isomorphism.

□

Lemma 5. *If* $\mathrm{tw}(H) > k$ *and* H *is connected, then* $\mathbf{E}_0 H^u \equiv^{\mathcal{C}^k} \mathbf{E}_a H^u$ *for any* $a \in \mathcal{G}$.

Proof. Our aim is to exhibit a winning strategy for Duplicator in the k-pebble bijective game played on the two structures $\mathbf{A} = \mathbf{E}_0 H^u$ and $\mathbf{B} = \mathbf{E}_a H^u$. Since $\mathrm{tw}(H) > k$, we know that in the k cops and robber game played on H, robber has a winning strategy and Duplicator will make use of this strategy.

For each vertex $v \in V^H$ let X^v denote the set of variables $x_i^{v,e}$ for edges e incident on v. Similarly, for each $e \in E^H$, let X^e denote the set of variables involving e.

We say that a bijection $f : \mathbf{A} \to \mathbf{B}$ is *good* for a vertex $v \in V^H$ if the following conditions hold:

1. for all $w \in V^H$, $fX^w = X^w$;
2. for all $e \in E^H$, $fX^e = X^e$;
3. for all x, y, if $x + y = i$ is an equation in $\mathbf{A}$ then $f(x) + f(y) = i$ is an equation in $\mathbf{B}$; and
4. for all x, y, z, if $x + y + z = i$ is an equation in $\mathbf{A}$, then
   - $f(x) + f(y) + f(z) = i$ is an equation in $\mathbf{B}$ if $x, y, z \notin X^v$; and
   - $f(x) + f(y) + f(z) = i + a$ is an equation in $\mathbf{B}$ if $x, y, z \in X^v$.

Observe that the identity is a bijection that is good for u. Also, observe that a bijection that is good for v preserves all equations except the vertex equations for v.

Claim 21. *6 Given a bijection* $f : \mathbf{A} \to \mathbf{B}$ *that is good for* v*, if there is a path in* H *from* v *to* w *avoiding* $u_1, \ldots, u_k$ *then there is a bijection* $f' : \mathbf{A} \to \mathbf{B}$ *that is good for* w *such that*

$$f|_{(X^{u_1} \cup \cdots \cup X^{u_k})} = f'|_{(X^{u_1} \cup \cdots \cup X^{u_k})}.$$

Proof of Claim: Let the path from v to w avoiding $u_1, \ldots, u_k$ be $v = v_1, \ldots, v_n = w$. For each edge $e = \{v_i, v_{i+1}\}$ along this path, write x_j^{e-} for the variable $x_j^{v_i,e}$ and x_j^{e+} for the variable $x_j^{v_{i+1},e}$. We then define f' by:

- $f'(x_j^{e-}) = f(x_{j-a}^{e-})$;

- $f'(x_j^{e+}) = f(x_{j+a}^{e+})$;

and f' agrees with f everywhere else. In particular, since the path from v to w avoids $u_1, \ldots, u_k$, f' agrees with f on $X^{u_1} \cup \cdots \cup X^{u_k}$. $\square$

We now describe Duplicator's winning strategy in the bijective k-pebble game. Duplicator responds to Spoiler's first pebble with the identity bijection. She maintains a board on the side which describes a position in the k cops and robber game played on the graph H. At any point in the pebble game, if Spoiler's pebbles are on the position $x_1, \ldots, x_k$ in $\mathbf{A}$ and $v_1, \ldots, v_k$ are the vertices of H to which these variables correspond, then the current position of the cops and robber game has k cops sitting on the vertices $v_1, \ldots, v_k$. If the robber's position according to its winning strategy is v, then Duplicator will play a bijection that is good for v.

To see that Duplicator can do this forever, suppose Spoiler lifts a pebble from x_i. Duplicator responds with a current bijection f that is good for v. Since the only equations not preserved by f are those associated with the vertex v, Spoiler must place at least three pebbles on variables associated with v to win the game. However, Duplicator responds to Spoiler placing the pebble on a new position x_i' by updating the position of the cops and robber game. Suppose robber's winning strategy dictates that the robber move from v to w. Since robber's move must be along a path avoiding the current cop positions, by Claim 6, Duplicator can update the bijection f to a new f' that is good for w without changing f on any of the currently pebbled positions. It is now clear that Duplicator can play forever.

$\square$

B Proofs omitted from Section 4

Lemma 9. *Let $\mathbf{B}$ be a finite structure, and let $\mathbf{D}$ be an expansion of $\mathbf{B}$ by a reduced relation R that is definable in $\mathbf{B}$ by a pp-formula. Then,* $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B})$.

Proof: Let σ be the vocabulary of $\mathbf{B}$. Let r be the arity of $R^{\mathbf{D}}$ and let $\phi(x_1, \ldots, x_r)$ be the primitive-positive formula that defines $R^{\mathbf{D}}$ in $\mathbf{B}$. The formula has the following form:

$$(\exists x_{r+1}) \cdots (\exists x_m) \left(R_1(\mathbf{x}_{I_1^1}) \wedge \cdots \wedge R_1(\mathbf{x}_{I_{n_1}^1}) \wedge \cdots \wedge R_s(\mathbf{x}_{I_1^s}) \wedge \cdots \wedge R_s(\mathbf{x}_{I_{n_s}^s}) \right),$$

where $R_1, \ldots, R_s$ are all the relation symbols of σ, each I_j^i is a sequence of indices in $\{1, \ldots, m\}$ whose length matches the arity r_i of R_i, and $\mathbf{x}_I$ denotes the projection of the tuple $(x_1, \ldots, x_m)$ to the indices indicated by I. We may assume that all variables $x_{r+1}, \ldots, x_m$ are distinct and disjoint from $x_1, \ldots, x_r$. Moreover, since $R^{\mathbf{D}}$ is reduced, we may also assume that all variables $x_1, \ldots, x_r$ are distinct. Given an instance $\mathbf{C}$ of $\mathrm{CSP}(\mathbf{D})$, we need to define an instance $\mathbf{A}$ of $\mathrm{CSP}(\mathbf{B})$ such that $\mathbf{A} \to \mathbf{B}$ if and only if $\mathbf{C} \to \mathbf{D}$. First we define $\mathbf{A}$ abstractly, and then show how to define it in $\mathbf{C}$ through a positive quantifier-free interpretation with parameters.

The universe of $\mathbf{A}$ is the set

$$C \cup (R^{\mathbf{C}} \times \{x_{r+1}, \ldots, x_m\}),$$

where $x_{r+1}, \ldots, x_m$ are the quantified variables in ϕ, which we assume not to be members of C. Intuitively, we have a new copy of each quantified variable of ϕ for each tuple in $R^{\mathbf{C}}$. The interpretation of the relation R_i in $\mathbf{A}$ consists of $R_i^{\mathbf{C}}$, together with a set of tuples defined next. For every $\mathbf{c} = (c_1, \ldots, c_r)$ in $R^{\mathbf{C}}$ and for every $I_j^i = (i_1, \ldots, i_{r_i})$, add to $R_i^{\mathbf{A}}$ the tuple $(z_1, \ldots, z_{r_i})$ defined by:

1. $z_k = c_{i_k}$ if i_k is the index of a free variable of ϕ, that is, $1 \leq i_k \leq r$,
2. $z_k = (\mathbf{c}, x_{i_k})$ if i_k is the index of a bound variable of ϕ, that is, $r + 1 \leq i_k \leq m$.

This defines the structure $\mathbf{A}$. Let us prove it has the right property:

Claim 22. $\mathbf{A} \to \mathbf{B}$ *if and only if* $\mathbf{C} \to \mathbf{D}$.

Proof: Let h be a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. We claim that the restriction of h to C is a homomorphism from $\mathbf{C}$ to $\mathbf{D}$. For every R_i we have $R_i^{\mathbf{C}} \subseteq R_i^{\mathbf{A}}$ and $R_i^{\mathbf{D}} = R_i^{\mathbf{B}}$. Moreover, h is a homomorphism, so $h(R_i^{\mathbf{A}}) \subseteq R_i^{\mathbf{B}}$. Thus $h(R_i^{\mathbf{C}}) \subseteq R_i^{\mathbf{D}}$. Let us now check that $h(R^{\mathbf{C}}) \subseteq R^{\mathbf{D}}$. Let then $\mathbf{c}$ be any tuple in $R^{\mathbf{C}}$. Let $\mathbf{d} = h(\mathbf{c})$. We want to show that $\mathbf{B} \models \phi(\mathbf{d})$, so $\mathbf{d}$ belongs to $R^{\mathbf{D}}$. By the definition of $\mathbf{A}$, for every $I_j^i = (i_1, \ldots, i_{r_i})$, the tuple $(z_1, \ldots, z_{r_i})$ defined as before belongs to $R_i^{\mathbf{A}}$. Now, if i_k is the index of a bound variable of ϕ, we view $h((\mathbf{c}, x_{i_k}))$ as a witness for x_{i_k} when evaluating $\phi(\mathbf{d})$ in $\mathbf{B}$. On the other hand, if i_k is the index of a free variable of ϕ, we view $d_{i_k} = h(c_{i_k})$ as the interpretation of x_{i_k}. This interpretation is well-defined because, critically, $R^{\mathbf{D}}$ is reduced so all m variables $x_1, \ldots, x_m$ are distinct. Under this interpretation for the free and bound variables, we have $\mathbf{B} \models \phi(\mathbf{d})$ as was to be proved.

Suppose now that h is a homomorphism from $\mathbf{C}$ to $\mathbf{D}$. We need to extend h to map A to B. Fix a tuple $\mathbf{c}$ in $R^{\mathbf{C}}$, and let $\mathbf{d} = h(\mathbf{c})$. Then $\mathbf{d}$ belongs to $R^{\mathbf{D}}$ so $\mathbf{B} \models \phi(\mathbf{d})$. Let $b_{r+1}, \ldots, b_m$ be witnesses to the existentially quantified variables in ϕ. We extend h by defining $h((\mathbf{c}, x_i)) = b_i$ for $r+1 \leq i \leq m$. The claim is that h is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$ and that this follows directly from the definitions. $\square$

We are left with the question of showing that this reduction is indeed a positive quantifier-free interpretation with parameters. This is more or less routine. Fix a pair of distinct variables p_0, p_1 that will play the role of parameters. For concreteness, we can think of p_0 and p_1 as distinct elements of C. Let $q = m - r$ and $t = \lfloor \log_2 q \rfloor + 1$. We can think of the universe of $\mathbf{A}$ as the subset of C^{r+t+2} defined by the following formula with $r + t + 2$ free variables $y_0, y_1, \ldots, y_{r+t+1}$:

$$(y_0 = p_0 \wedge y_1 = \cdots = y_{r+t+1}) \vee (y_0 = p_1 \wedge R(y_1, \ldots, y_r) \wedge \psi(y_{r+1}, \ldots, y_{r+t+1})),$$

where $\psi(y_{r+1}, \ldots, y_{r+t+1})$ is a formula that is satisfied by the set of numbers $k \in \{0, \ldots, q-1\}$ when encoded in binary; the bits are encoded by $y_{r+b} = p_0$ or $y_{r+b} = p_1$. In other words, when q is an exact power of two, which we may as well assume by adding dummy variables, ψ is the following formula:

$$\begin{array}{lllllllll}
 & (y_{r+t+1} = p_0 & \wedge & \cdots & \wedge & y_{r+2} = p_0 & \wedge & y_{r+1} = p_0) & \vee \\
\vee & (y_{r+t+1} = p_0 & \wedge & \cdots & \wedge & y_{r+2} = p_0 & \wedge & y_{r+1} = p_1) & \vee \\
\vdots & & & & & & & & \\
\vee & (y_{r+t+1} = p_1 & \wedge & \cdots & \wedge & y_{r+2} = p_1 & \wedge & y_{r+1} = p_0) & \vee \\
\vee & (y_{r+t+1} = p_1 & \wedge & \cdots & \wedge & y_{r+2} = p_1 & \wedge & y_{r+1} = p_1). &
\end{array}$$

Intuitively, the set of tuples $(y_0, \ldots, y_{r+t+1})$ for which $y_0 = p_0 \wedge y_1 = \cdots = y_{r+t+1}$ holds encodes C, and the set of tuples for which $y_0 = p_1 \wedge R(y_1, \ldots, y_r) \wedge \psi_q(y_{r+1}, \ldots, y_{r+t+1})$ holds encodes $R^{\mathbf{C}} \times \{x_{r+1}, \ldots, x_m\}$. With this universe at hand, the rest of the formal definition is easy to work out. $\square$

Lemma 10. *Let* $\mathbf{B}$ *a finite structure and let* $\mathbf{D}$ *be the reduced version of* $\mathbf{B}$. *Then* $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{D})$ *and* $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B})$.

Proof: We start with the reduction $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B})$. Let σ be the vocabulary of $\mathbf{B}$ and let σ' be the vocabulary of the reduced structure $\mathbf{D}$. Hence, for every symbol R in σ we have a symbol R' in σ' of the arity of $\mathrm{red}(R^{\mathbf{B}})$. Let $\mathbf{C}$ be an instance of $\mathrm{CSP}(\mathbf{D})$. We define an instance $\mathbf{A}$ of $\mathrm{CSP}(\mathbf{B})$. The universe of $\mathbf{A}$ is C itself. The interpretation of the r-ary symbol R in $\mathbf{A}$ is defined as follows: let $\theta = \theta(R^{\mathbf{B}})$ and let I be a set of representatives of the θ-classes, ordered in an arbitrary way. Then, $R^{\mathbf{A}}$ is defined by the formula

$$\psi_R(x_1, \ldots, x_r) = R'(\mathbf{x}_I) \wedge \bigwedge_{(i,j) \in \theta} x_i = x_j.$$

It is clear that $\mathbf{C} \to \mathbf{D}$ if, and only if, $\mathbf{A} \to \mathbf{B}$. Moreover, the reduction is positive quantifier-free.

We proceed now with the reduction $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{datalog}} \mathrm{CSP}(\mathbf{D})$. Let $\mathbf{A}$ be an instance of $\mathrm{CSP}(\mathbf{B})$. We define an instance $\mathbf{C}$ of $\mathrm{CSP}(\mathbf{D})$. The universe of $\mathbf{C}$ is A itself. For the relations, the basic idea is to project every relation $R^{\mathbf{A}}$ to the coordinates of a set of representatives I of the θ-classes, where $\theta = \theta(R)$. However, before we do that, we need to *close* each $R^{\mathbf{A}}$ under all equalities implied by the equivalences $(i,j) \in \theta$. We do that using Datalog-definable intermediate relations.

So, let E be the binary relation on A defined by the following Datalog program:

$$
\begin{aligned}
E(x_i, x_j) &\ :-\ R(x_1, \ldots, x_s) \\
E(x, y) &\ :-\ E(y, x) \\
E(x, z) &\ :-\ E(x, y) \wedge E(y, z),
\end{aligned}
$$

where the first rule is introduced for every symbol R in σ and every $(i,j) \in \theta(R)$. It is obvious that E is an equivalence relation on A; reflexivity follows from the fact that $(i,i) \in \theta(R)$ in the first rule, symmetry is enforced by the second rule, and transitivity is enforced by the third. Next, for every r-ary symbol R in σ, let $\hat{R}$ and R' be the relations defined by

$$
\begin{aligned}
\hat{R}(x_1, \ldots, x_s) &\ :-\ R(y_1, \ldots, y_s) \wedge E(x_1, y_1) \wedge \cdots \wedge E(x_s, y_s) \\
R'(\mathbf{x}_I) &\ :-\ \hat{R}(\mathbf{x}),
\end{aligned}
$$

where I is a set of representatives of the $\theta(R)$-classes ordered in an arbitrary way. This defines $\mathbf{C}$, and we defined it by a Datalog program interpreted on $\mathbf{A}$. Let us now prove that this datalog-interpretation is indeed a reduction. First a technical claim:

Claim 23. *If h is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$ and $(a, a') \in E$, then $h(a) = h(a')$.*

Proof: We proceed by induction on the stage on which (a, a') enters the relation E. If it enters in the first stage, then there exist R in σ, $(i,j) \in \theta(R)$, and $\mathbf{a} \in R^{\mathbf{A}}$ such that $a_i = a$ and $a_j = a'$. Since $h(\mathbf{a}) \in R^{\mathbf{A}}$ and $(i,j) \in \theta(R)$, it follows that $h(a_i) = h(a_j)$. Hence $h(a) = h(a')$. The inductive cases follow trivially from symmetry and transitivity of equality. □

Claim 24. $\mathbf{A} \to \mathbf{B}$ *if and only if* $\mathbf{C} \to \mathbf{D}$.

Proof: Suppose that $\mathbf{A} \to \mathbf{B}$ and let h be a homomorphism. We claim that h itself is also a homomorphism from $\mathbf{C}$ to $\mathbf{D}$. Suppose $\mathbf{c} \in R'^{\mathbf{C}}$. Then there exists $\mathbf{a} \in \hat{R}$ such that $\mathbf{a}_I = \mathbf{c}$, which in turn means that there exists $\mathbf{a}' \in R^{\mathbf{A}}$ such that $(a_i, a'_i) \in E$ for every $i \in \{1, \ldots, s\}$. Now, $h(\mathbf{a}') \in R^{\mathbf{B}}$ because h is a homomorphism. But also $h(\mathbf{a}) = h(\mathbf{a}')$ by the claim above because $(a_i, a'_i) \in E$ for every i. But then

$$
h(\mathbf{c}) = h(\mathbf{a}_I) = h(\mathbf{a})_I = h(\mathbf{a}')_I \in \mathrm{pr}_I(R^{\mathbf{B}}) = \mathrm{red}(R^{\mathbf{B}}) = R'^{\mathbf{D}}.
$$

Thus h is a homomorphism from $\mathbf{C}$ to $\mathbf{D}$.

Suppose now that $\mathbf{C} \to \mathbf{D}$ and let h be a homomorphism. For every $a \in A$, let a^E be a fixed representative of the E-equivalence class of a. Let $g(a) = h(a^E)$ for every a. We claim that g is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. Suppose $\mathbf{a} \in R^{\mathbf{A}}$. Then $\mathbf{a}^E \in \hat{R}$, so $(\mathbf{a}^E)_I \in R'^{\mathbf{C}}$. Then $h((\mathbf{a}^E)_I) \in R'^{\mathbf{D}}$. Note that

$$
g(\mathbf{a})_I = h(\mathbf{a}^E)_I = h((\mathbf{a}^E)_I) \in R'^{\mathbf{D}} = \mathrm{red}(R^{\mathbf{B}}) = \mathrm{pr}_I(R^{\mathbf{B}}).
$$

But then $g(\mathbf{a}) \in R^{\mathbf{B}}$ by the definition of $\theta(R)$ and I. So g is a homomorphism. □

Lemma 11. *Let the algebras $\mathcal{B}$ and $\mathcal{B}'$, and the structures $\mathbf{B}$ and $\mathbf{B}' = h^{-1}(\mathbf{B})$ be as above. Then*

1. $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$.

2. $\mathcal{B}'$ *is a reduct of* $\mathrm{Al}(\mathbf{B}')$.

Proof: 1. We argue that $\mathrm{CSP}(\mathbf{B}) = \mathrm{CSP}(\mathbf{B}')$ by arguing that $\mathbf{B}$ and $\mathbf{B}'$ are homomorphically equivalent. The homomorphism from $\mathbf{B}'$ to $\mathbf{B}$ is just h, and this is easy to check. As a homomorphism from $\mathbf{B}$ to $\mathbf{B}'$ we take any *inverse* of h; that is, any function $f : B \to B'$ such that $f(b)$ belongs to $h^{-1}(b)$ for every $b \in B$. Such a function exists because h is onto B. It is a homomorphism because if $\mathbf{b}$ is a tuple in $R^{\mathbf{B}}$, then $h(f(\mathbf{b})) = \mathbf{b}$, so $f(\mathbf{b}) \in h^{-1}(R^{\mathbf{B}})$.

2. It suffices to show that every operation of $\mathcal{B}'$ is a polymorphism of $\mathbf{B}'$. Let f' be an m-ary operation of $\mathcal{B}'$, and let f be the corresponding operation of $\mathcal{A}$. Suppose that $\mathbf{a}^1, \ldots, \mathbf{a}^m$ are m tuples that belong to $h^{-1}(R^{\mathbf{B}})$. Then the tuples $h(\mathbf{a}^1), \ldots, h(\mathbf{a}^m)$ all belong to $R^{\mathbf{B}}$. We apply f component-wise and we obtain the tuple

$$(f(h(a_1^1), \ldots, h(a_1^m)), \ldots, f(h(a_r^1), \ldots, h(a_r^m))).$$

Since f is an operation of $\mathcal{A}$, and $\mathcal{A}$ is a reduct of $\mathcal{B}$, it is a polymorphism of $\mathbf{B}$, so this tuple belongs to $R^{\mathbf{B}}$. Now, by the choice of f, this tuple is the same as

$$(h(f'(a_1^1, \ldots, a_1^m)), \ldots, h(f'(a_r^1, \ldots, a_r^m))).$$

We conclude that the tuple

$$(f'(a_1^1, \ldots, a_1^m), \ldots, f'(a_r^1, \ldots, a_r^m))$$

belongs to $h^{-1}(R^{\mathbf{B}})$. This proves that f' preserves every relation of $\mathbf{B}'$. $\square$

Lemma 12. *Let the algebras $\mathcal{B}$ and $\mathcal{B}'$, and the structures* $\mathbf{B}$ *and* $\mathbf{B}' = \mathrm{ext}(\mathbf{B}, B')$ *be as above. Then*

1. $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$.

2. $\mathcal{B}'$ *is a reduct of* $\mathrm{Al}(\mathbf{B}')$.

Proof: 1. The structures $\mathbf{B}$ and $\mathbf{B}'$ are homomorphically equivalent. Indeed the identity mapping on B is a homomorphism of $\mathbf{B}$ to $\mathbf{B}'$, and any mapping $h : B' \to B$ that is the identity on $B \subseteq B'$ and maps elements from $B' \setminus B$ to any element of B is a homomorphism from $\mathbf{B}'$ to $\mathbf{B}$.

2. Let f' be an operation of $\mathcal{B}'$ and let f be the corresponding operation in $\mathcal{A}$. Then f preserves every relation of $\mathbf{B}$ because $\mathcal{A}$ is a reduct of $\mathcal{B}$. But then, trivially, f' also preserves every relation of $\mathbf{B}'$ because the relations in $\mathbf{B}'$ and $\mathbf{B}$ are the same. $\square$

Lemma 13. *Let the algebras $\mathcal{B}$ and $\mathcal{B}'$, and the structures* $\mathbf{B}$ *and* $\mathbf{B}' = \flat(\mathbf{B})$ *be as above. Then*

1. $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{B}')$.

2. $\mathcal{B}'$ *is a reduct of* $\mathrm{Al}(\mathbf{B}')$.

Proof: 1. Given an instance $\mathbf{A}$ of $\mathrm{CSP}(\mathbf{B})$, we need to define an instance $\mathbf{A}'$ of $\mathrm{CSP}(\mathbf{B}')$ such that $\mathbf{A} \to \mathbf{B}$ if, and only if, $\mathbf{A}' \to \mathbf{B}'$. First we define $\mathbf{A}'$ abstractly, and then show how to define it on $\mathbf{A}$ through a positive quantifier-free interpretation with parameters.

The universe of the structure $\mathbf{A}'$ is $A \times \{1, \ldots, n\}$. For every k-ary symbol R in the vocabulary of $\mathbf{B}$, we have a corresponding kn-ary symbol $\overline{R}$ in the vocabulary of $\mathbf{B}'$. The interpretation of $\overline{R}$ in $\mathbf{A}'$ is defined as follows:

$$((x_1, 1), \ldots, (x_1, n), \ldots, (x_k, 1), \ldots, (x_k, n)) \in \overline{R}^{\mathbf{A}'} \iff (x_1, \ldots, x_k) \in R^{\mathbf{A}}.$$

First we prove that this structure has the right property. If $\mathbf{A} \to \mathbf{B}$ and h is a homomorphism, then clearly the mapping $h' : A' \to B'$ defined by the condition $h'((x,i)) = h(x)_i$, where $h(x) = (h(x)_1, \ldots, h(x)_n)$, is a homomorphism. Conversely, if h is a homomorphism from $\mathbf{A}'$ to $\mathbf{B}'$, then the mapping $h'(x) = (h'((x,1)), \ldots, h'((x,n)))$ is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$.

Next we show that this reduction is positive quantifier-free. Fix a pair of distinct variables p_0, p_1 that will play the role of parameters. For concreteness, we can think of p_0 and p_1 as distinct elements of A. Let $t = \lfloor \log_2 n \rfloor + 1$. We can think of the universe of $\mathbf{A}'$ as the subset of A^{t+1} defined by the formula $\psi(y_0, y_1, \ldots, y_t)$ with $t+1$ free variables that is satisfied by the tuples $(y_0, y_1, \ldots, y_t)$ for which $(y_1, \ldots, y_t)$ encodes a number from $\{0, \ldots, n-1\}$ in binary; the bits are encoded by $y_b = p_0$ or $y_b = p_1$ for $1 \leq b \leq t$ (see the proof of Lemma 9). The interpretation of the relational symbol $\overline{R}$ of arity kn is given by the formula

$$\psi_{\overline{R}}(\mathbf{y}^1, \ldots, \mathbf{y}^{kn}) \quad = \quad R(y_0^1, y_0^{n+1}, \ldots, y_0^{(k-1)n+1}) \wedge \bigwedge_{j=1}^{n} \bigwedge_{i=0}^{k-1} (\mathrm{bin}(y_1^{in+j}, \ldots, y_t^{in+j}) = j-1),$$

where $\mathrm{bin}(y_1^{in+j}, \ldots, y_t^{in+j}) = j-1$ abbreviates the expression

$$y_1^{in+j} = b_1 \wedge \cdots \wedge y_t^{in+j} = b_t$$

and $b_1 \ldots b_t$ is the binary representation of $j-1$.

2. Since $\mathcal{A}$ is a reduct of $\mathcal{B}$, every relation of $\mathbf{B}$ is invariant with respect to all operations of $\mathcal{A} = \mathcal{B}'^n$. Now it is straightforward that every relation in the flattening of $\mathbf{B}$ is invariant with respect to every operation of $\mathcal{B}'$. □

Lemma 15. *Let* $\mathbf{B}$ *be a finite structure, and let* $\mathbf{D}$ *be the singleton-expansion of* $\mathbf{B}$*. Then* $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{D})$ *and if* $\mathbf{B}$ *is a core with at least two points, then* $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{ep}} \mathrm{CSP}(\mathbf{B})$*.*

Proof: The reduction $\mathrm{CSP}(\mathbf{B}) \leq_{\mathrm{pqf}} \mathrm{CSP}(\mathbf{D})$ is straightforward: it suffices to map every instance of $\mathrm{CSP}(\mathbf{B})$ to its expansion with n empty unary relations, where n is the cardinality of B. This is clearly a positive quantifier-free reduction without any parameter.

Let us now prove that $\mathrm{CSP}(\mathbf{D}) \leq_{\mathrm{ep}} \mathrm{CSP}(\mathbf{B})$. Given an instance $\mathbf{C}$ of $\mathrm{CSP}(\mathbf{D})$, we need to define an instance $\mathbf{A}$ of $\mathrm{CSP}(\mathbf{B})$ such that $\mathbf{A} \to \mathbf{B}$ if, and only if, $\mathbf{C} \to \mathbf{D}$. First we define $\mathbf{A}$ abstractly, and then show how to define it on $\mathbf{C}$ through an existential-positive interpretation with parameters.

The universe of the structure $\mathbf{A}$ is the disjoint union of C and B. For every relation symbol R of arity r in the vocabulary of $\mathbf{B}$, the interpretation of R in $\mathbf{A}$ is defined by cases: if the sets $P_b^{\mathbf{C}}$ are pairwise disjoint, we let $R^{\mathbf{A}} = A^r$. Otherwise, we let $R^{\mathbf{A}}$ be the set

$$R^{\mathbf{B}} \cup \bigcup_{u \in F} u(R^{\mathbf{C}}),$$

where F is the set of mappings $u : C \to A$ such that the following two conditions are satisfied:

1. $u(y) \in P_b^{\mathbf{C}} \cup \{b\}$ for every $b \in B$ and $y \in P_b^{\mathbf{C}}$,
2. $u(y) = y$ for every $y \in C - \bigcup_{b \in B} P_b^{\mathbf{C}}$.

This defines the structure $\mathbf{A}$. Before we show how to define $\mathbf{A}$ by an existential-positive interpretation, let us show that it has the property we want:

Claim 25. $\mathbf{A} \to \mathbf{B}$ *if, and only if,* $\mathbf{C} \to \mathbf{D}$.

Proof: If the sets $P_b^{\mathbf{C}}$ are not pairwise disjoint, then clearly $\mathbf{C} \not\to \mathbf{D}$. In this case, every relation in $\mathbf{A}$ is the full relation and in particular it is reflexive. But then $\mathbf{A} \not\to \mathbf{B}$ since otherwise $\mathbf{B}$ would also be reflexive and hence not a core with at least two elements.

Suppose in the following that the sets $P_b^{\mathbf{C}}$ are pairwise disjoint. Let h be a homomorphism from $\mathbf{C}$ to $\mathbf{D}$. Note that $h(y) = b$ for every $y \in P_b^{\mathbf{C}}$; this remark will be of use later. Let g be the unique extension of h to $A = B \cup C$ such that $g(b) = b$ for every $b \in B$. We prove that g is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. Let $\mathbf{x} \in R^{\mathbf{A}}$ for some relation symbol R, and we prove $g(\mathbf{x}) \in R^{\mathbf{B}}$. Since $\mathbf{x} \in R^{\mathbf{A}}$, either $\mathbf{x} \in R^{\mathbf{B}}$, or $\mathbf{x} \in u(R^{\mathbf{C}})$ for some $u \in F$. In the first case, $g(\mathbf{x}) = \mathbf{x}$ and hence $g(\mathbf{x}) \in R^{\mathbf{B}}$ as required. In the second case, $\mathbf{x} = u(\mathbf{y})$ for some $\mathbf{y} \in R^{\mathbf{C}}$. Let $\mathbf{y} = (y_1, \ldots, y_r)$ and let us analyze the components y_i dinstinguishing by cases whether they belong to some $P_b^{\mathbf{C}}$ or not. Suppose first $y_i \in P_b^{\mathbf{C}}$ for some b. Then $h(y_i) = b$ by the remark above. Also $u(y_i) \in P_b^{\mathbf{C}} \cup \{b\}$ by the definition of F. Continuing, if $u(y_i) \in P_b^{\mathbf{C}}$ then $g(u(y_i)) = b$ again by the remark above, and if $u(y_i) = b$ then $g(u(y_i)) = g(b) = b$ by the definition of g. Therefore $g(u(y_i)) = h(y_i)$. Suppose next that $y_i \not\in P_b^{\mathbf{C}}$ for all $b \in B$. Then $u(y_i) = y_i$ by the definition of F, and $g(u(y_i)) = h(y_i)$ again. It follows that $g(u(\mathbf{y})) = h(\mathbf{y})$. Since $\mathbf{y} \in R^{\mathbf{C}}$ and h is a homomorphism from $\mathbf{C}$ to $\mathbf{D}$, we have $h(\mathbf{y}) \in R^{\mathbf{D}}$. It follows that $g(\mathbf{x}) \in R^{\mathbf{B}}$ because $g(\mathbf{x}) = g(u(\mathbf{y})) = h(\mathbf{y})$ and $R^{\mathbf{D}} = R^{\mathbf{B}}$. This proves that g is a homomorphism.

Suppose next that f is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. Note that $\mathbf{B}$ is an induced substructure of $\mathbf{A}$, so the restriction of f to B is a homomorphism from $\mathbf{B}$ to itself. Since $\mathbf{B}$ is a core, this restriction must be an automorphism π of $\mathbf{B}$. We may assume then that f is the identity on B; otherwise we start with the homomorphism obtained from f by composing it with π^{-1} on B. Now we define the map $h : C \to B$ as follows: if $y \in P_b^{\mathbf{C}}$ for some $b \in B$, then $h(y) = b$; otherwise, $h(y) = f(y)$. Since we are assuming that the sets $P_b^{\mathbf{C}}$ are pairwise disjoint, the map h is well-defined. We claim that h is a homomorphism from $\mathbf{C}$ to $\mathbf{D}$. First note that if $y \in P_b^{\mathbf{C}}$, then $h(y) \in P_b^{\mathbf{D}}$ by definition. Now, let $\mathbf{y} \in R^{\mathbf{C}}$ for some relation symbol R, and we prove $h(\mathbf{y}) \in R^{\mathbf{D}}$. Define $u : C \to A$ by $u(y) = b$ if $y \in P_b^{\mathbf{C}}$ for some b, and $u(y) = y$ otherwise. Since the sets $P_b^{\mathbf{C}}$ are disjoint, this is well-defined. Note that $u \in F$. Let $\mathbf{y} = (y_1, \ldots, y_r)$ and let us analyze the components y_i distinguishing by cases on whether they belong to some $P_b^{\mathbf{C}}$ or not. Suppose first that $y_i \in P_b^{\mathbf{C}}$ for some b. Then $u(y_i) = b$ by the definition of u, and $f(u(y_i)) = b$ because f is the identity on B. Also $h(y_i) = b$ by the definition of h. Therefore $h(y_i) = f(u(y_i))$. Suppose next that $y_i \not\in P_b^{\mathbf{C}}$ for any b. Then $u(y_i) = y_i$ by the definition of u, and $h(y_i) = f(y_i)$ by the definition of h. Again $h(y_i) = f(u(y_i))$. It follows then that $h(\mathbf{y}) = f(u(\mathbf{y}))$. Now, $u(\mathbf{y}) \in R^{\mathbf{A}}$ because $u \in F$ and $\mathbf{y} \in R^{\mathbf{C}}$. Hence $f(u(\mathbf{y})) \in R^{\mathbf{B}}$ because f is a homomorphism from $\mathbf{A}$ to $\mathbf{B}$. Thus $h(\mathbf{y}) \in R^{\mathbf{D}}$ because $R^{\mathbf{D}} = R^{\mathbf{B}}$. This proves that h is a homomorphism. $\square$

We are left with the question of showing that this reduction is indeed existential-positive. Fix a pair of distinct variables p_0, p_1 that will play the role of parameters. For concreteness, we can think of p_0 and p_1 as distinct elements of C. Let $q = |B|$ and $t = \lfloor \log_2 q \rfloor + 1$. We can think of the universe of $\mathbf{A}$ as the subset of C^{t+1} defined by the following formula with $t+1$ free variables $y_0, y_1, \ldots, y_t$:

$$(y_0 = p_0 \wedge y_1 = \cdots = y_t) \vee (y_1 = p_1 \wedge \psi(y_1, \ldots, y_t)),$$

where $\psi(y_1, \ldots, y_t)$ is a formula that is satisfied by the set of numbers $k \in \{0, \ldots, q-1\}$ when encoded in binary; the bits are encoded by $y_b = p_0$ or $y_b = p_1$. This is the same formula as in the proof of Lemma 9. Intuitively, the set of tuples $(y_0, \ldots, y_t)$ for which $y_0 = p_0 \wedge y_1 = \cdots = y_t$ holds encodes C, and the set of tuples for which $y_0 = p_1 \wedge \psi(y_1, \ldots, y_t)$ encodes B. Now we define the interpretation of the relation symbol R by the following formula:

$$\xi \vee \phi_R \vee \bigvee_{v \in G} \theta_{v,R},$$

where G is the set of mappings $v : \{1, \ldots, r\} \to B \times \{0, 1\}$, and ξ, ϕ_R and $\theta_{v,R}$ are formulas to be described soon. Note the similarity of this formula with the abstract definition of $R^{\mathbf{A}}$ that we gave:

$$R^{\mathbf{B}} \cup \bigcup_{u \in F} u(R^{\mathbf{C}}).$$

The formula $\phi_R(\mathbf{y}^1, \ldots, \mathbf{y}^r, \mathbf{p})$ encodes the set $R^{\mathbf{B}}$ as a finite disjunction of conjunctions of equalities encoding the tuples of $R^{\mathbf{B}}$. This is easy to work out. The formula $\theta_{v,R}(\mathbf{y}^1, \ldots, \mathbf{y}^r, \mathbf{p})$ encodes the set $u(R^{\mathbf{C}})$ as follows:

$$(\exists z_1) \cdots (\exists z_r)(R(z_1, \ldots, z_r) \wedge T_1 \wedge T_2)$$

where

$$\begin{aligned} T_1 &= \bigwedge \{\ y_0^j = p_0 \wedge y_1^j = \ldots = y_t^j \wedge P_b(z_j) \wedge P_b(y_1^j)\ :\ j \in \{1, \ldots, r\}, v(j) = (b, 0)\ \}, \\ T_2 &= \bigwedge \{\ y_0^j = p_1 \wedge y_1^j = p_{b_1} \wedge \ldots \wedge y_t^j = p_{b_t}\ :\ j \in \{1, \ldots, r\}, v(j) = (b, 1)\ \}, \end{aligned}$$

where $b_1, \ldots, b_t$ denote the bits of the binary encoding of b in a fixed numbering of B. Finally, the formula ξ is defined as

$$\bigvee_{b_1 \neq b_2} (\exists z)(P_{b_1}(z) \wedge P_{b_2}(z)),$$

where b_1 and b_2 range over B. This completes the definition of $R^{\mathbf{A}}$. Note that ξ is used to make $R^{\mathbf{A}} = A^r$ whenever the sets $P_b^{\mathbf{C}}$ are not disjoint. $\square$

C Proofs omitted from Section 5

We need some more notation. Let $\mathcal{A}$ be an algebra and α a congruence of $\mathcal{A}$. For any $a \in \mathcal{A}$ we denote by a^α the equivalence class containing a. The set of all equivalence classes is called the *quotient set*, $A/_\alpha$. The fact that α is a congruence allows one to define the action of any operation of $\mathcal{A}$ on the quotient set:

$$f^\alpha(a_1^\alpha, \ldots, a_n^\alpha) = (f(a_1, \ldots, a_n))^\alpha.$$

The quotient set endowed with all the operations f^α is called the *quotient algebra* $\mathcal{A}/_\alpha$ of $\mathcal{A}$. It is not hard to see that $\mathcal{A}/_\alpha$ is the homomorphic image of $\mathcal{A}$ under the homomorphism that takes $a \in \mathcal{A}$ to a^α.

Lemma 16. *Let $\mathcal{A}$ be a finite idempotent algebra. If* $\mathrm{var}(\mathcal{A})$ *does not omit types* **1** *and* **2** *then it contains a finite idempotent reduct of a module.*

Proof. By a result from [6], if $\mathrm{var}(\mathcal{A})$ does not omit type **1** the it contains a finite *set*, that is an algebra all of whose opeartions are projections. So, suppose that $\mathrm{var}(\mathcal{A})$ omits type **1**, but does not omit type **2**.

Since $\mathrm{var}(\mathcal{A})$ does not omit type **2**, there is a finite algebra $\mathcal{B} \in \mathrm{var}(\mathcal{A})$ and a prime quotient $\alpha \prec \beta \in \mathrm{con}(\mathcal{B})$ such that $\mathsf{typ}(\alpha, \beta) = \mathbf{2}$. Note first that taking $\mathcal{B}/_\alpha$ instead of $\mathcal{B}$ we may assume that $\alpha = \underline{0}$, the equality relation, because it follows from tame congruence theory that $\mathsf{typ}(\alpha/_\gamma, \beta/_\gamma) = \mathsf{typ}(\alpha, \beta)$ for any $\gamma \leq \alpha$. Next we notice that $\mathcal{B}$ is an idempotent algebra, every congruence class of β is a subalgebra. Take a non-trivial β-class, and let $\mathcal{C}$ be the corresponding subalgebra. The restriction of β onto C is the total congruence $\underline{1}$.

A congruence θ *centralizes* η *modulo* ϵ if for any term operation $f(x_1, \ldots, x_n, y_1, \ldots, y_n, z_1 \ldots z_k)$, any $c_1, \ldots, c_k \in A$, and any $a_1^1, \ldots, a_n^1$, $a_1^2, \ldots, a_n^2$, $b_1^1, \ldots, b_m^1$, $b_1^2, \ldots, b_m^2 \in A$ such that $(a_i^1, a_i^2) \in \theta$, $(b_i^1, b_i^2) \in \eta$, the following implication holds:

$$\begin{array}{rcl} f(a_1^1, \ldots, a_n^1, b_1^1, \ldots, b_m^1, c_1, \ldots, c_k) & \stackrel{\epsilon}{\equiv} & f(a_1^2, \ldots, a_n^2, b_1^1, \ldots, b_m^1, c_1, \ldots, c_k) \\ & \Downarrow & \\ f(a_1^1, \ldots, a_n^1, b_1^2, \ldots, b_m^2, c_1, \ldots, c_k) & \stackrel{\epsilon}{\equiv} & f(a_1^2, \ldots, a_n^2, b_1^2, \ldots, b_m^2, c_1, \ldots, c_k). \end{array}$$

It is known that $\mathsf{typ}(\eta, \theta) \in \{\mathbf{1}, \mathbf{2}\}$ if and only if θ centralizes itself modulo η (see [16, Theorem 7.2]).

In our situation we have that β centralizes itself modulo $\underline{0}$ in $\mathcal{B}$. Therefore, $\underline{1}$ centralizes itself modulo $\underline{0}$ in $\mathcal{C}$. This implies $\mathsf{typ}(\mathcal{C}) \subseteq \{\mathbf{1}, \mathbf{2}\}$, and, since $\mathrm{var}(\mathcal{A})$ omits type $\mathbf{1}$, we obtain $\mathsf{typ}(\mathcal{C}) = \{\mathbf{2}\}$. By Theorem 9.6 of [16] there is a ternary term operation d that is Mal'tsev on $\mathcal{C}$, that is d satisfies the identities $d(x, y, y) = d(y, y, x) = x$. Therefore $\mathcal{C}$ generates a congruence permutable variety, and by a result of [15] it is an idempotent reduct of a module. □

Lemma 17. *Let $\mathcal{M}$ be a finite module over a ring $\mathcal{R}$ with unit, let $\mathcal{G}$ be its underlying Abelian group, and let $\mathcal{A}$ be an idempotent reduct of $\mathcal{M}$. Then $\mathcal{A}$ is a reduct of the algebra of $\mathbf{E}_{\mathcal{G},r}$ for every $r \geq 1$.*

Proof. Let $\mathbf{E} = \mathbf{E}_{\mathcal{G},r}$. Every m-ary term operation of $\mathcal{A}$ can be represented in the form

$$f(x_1, \ldots, x_m) = r_1 x_1 + \cdots + r_m x_m,$$

and, as f is idempotent, $r_1 + \cdots + r_m = 1$. Take m tuples $\mathbf{a}_1, \ldots, \mathbf{a}_m$ in the relation R_a^j in $\mathbf{E}$, where $\mathbf{a}_i = (a_{i1}, \ldots, a_{ij})$ for $i \in \{1, \ldots, m\}$. Check that the tuple

$$(f(a_{11}, \ldots, a_{m1}), \ldots, f(a_{1j}, \ldots, a_{mj}))$$

also belongs to R_a^j:

$$\begin{array}{rl} & f(a_{11}, \ldots, a_{m1}) + \cdots + f(a_{1j}, \ldots, a_{mj}) \\ = & (r_1 a_{11} + \cdots + r_j a_{m1}) + \cdots + (r_1 a_{1j} + \cdots + r_m a_{mj}) \\ = & r_1(a_{11} + \cdots + a_{1j}) + \cdots + r_m(a_{m1} + \cdots + a_{mj}) \\ = & r_1 a + \cdots + r_m a \\ = & a. \end{array}$$

Therefore, every relation of $\mathbf{E}$ is invariant under every operation of $\mathcal{A}$. That is, $\mathcal{A}$ is a reduct of the algebra of $\mathbf{E}$. □

D Proof omitted from Section 7

Theorem 20. RELATIONAL STRUCTURE OF TYPE 2 *is* NP-*complete.*

Proof. In [6] it is proved that the problem of given a relational structure whether the variety generated by the algebra of this structure omits type $\mathbf{1}$ is NP-complete. Here we actually prove that RELATIONAL STRUCTURE-OF-TYPE-2 is NP-complete even if the input structures give rise to algebras omitting type $\mathbf{1}$.

We reduce NOT-ALL-EQUAL SATISFIABILITY (NAE) to RELATIONAL STRUCTURE-OF-TYPE-2. Let $\mathcal{C} = C_1 \wedge \ldots \wedge C_q$ be an instance of NAE. Let V denote the set of variables occurring in $\mathcal{C}$. We construct a set of relations Γ.

Set $A_0 = \{a, b\}$, $A_v = W_v \cup A_v^0 \cup A_v^1$, $v \in V$, where $W_v = \{a_{0v}, a_{1v}, a_{2v}, a_{3v}\}$, $A_v^0 = \{0_{4v}^0, 0_{5v}^0, 0_{6v}^0, 0_{7v}^0, 0_{4v}^1, 0_{5v}^1, 0_{6v}^1, 0_{7v}^1\}$, $A_v^1 = \{1_{4v}^0, 1_{5v}^0, 1_{6v}^0, 1_{7v}^0, 1_{4v}^1, 1_{5v}^1, 1_{6v}^1, 1_{7v}^1\}$; and $A = A_0 \cup \bigcup_{v \in V} A_v$. For each $1 \leq i \leq q$, let $\{u, v, w\}$ be the variables occurring in C_i. Define a 6-ary relation R_i as

$$R_i = \begin{pmatrix} a & b & b & a \\ b & a & b & a \\ b & b & a & a \\ a_u^0 & a_u^1 & a_u^2 & a_u^3 \\ a_v^0 & a_v^1 & a_v^2 & a_v^3 \\ a_w^0 & a_w^1 & a_w^2 & a_w^3 \end{pmatrix} \cup \left(\begin{pmatrix} a \\ a \\ b \end{pmatrix} \times R_i^4 \right) \cup \left(\begin{pmatrix} a \\ b \\ a \end{pmatrix} \times R_i^5 \right) \cup \left(\begin{pmatrix} b \\ a \\ a \end{pmatrix} \times R_i^6 \right) \cup \left(\begin{pmatrix} b \\ b \\ b \end{pmatrix} \times R_i^7 \right),$$

where

$$R_i^j = \begin{pmatrix} 0_{1u}^0 & 0_{1u}^1 & 0_{1u}^0 & 0_{1u}^1 & 0_{1u}^0 & 0_{1u}^1 & 1_{1u}^0 & 1_{1u}^1 & 1_{1u}^0 & 1_{1u}^1 & 1_{1u}^0 & 1_{1u}^1 \\ 0_{1v}^0 & 0_{1v}^1 & 1_{1v}^0 & 1_{1v}^1 & 1_{1v}^0 & 1_{1v}^1 & 0_{1v}^0 & 0_{1v}^1 & 0_{1v}^0 & 0_{1v}^1 & 1_{1v}^0 & 1_{1v}^1 \\ 1_{1w}^0 & 1_{1w}^1 & 1_{1w}^0 & 1_{1w}^1 & 0_{1w}^0 & 0_{1w}^1 & 1_{1w}^0 & 1_{1w}^1 & 0_{1w}^0 & 0_{1w}^1 & 0_{1w}^0 & 0_{1w}^1 \end{pmatrix}.$$

Finally, set $\Gamma = \{\{(c)\} \mid c \in A\} \cup \{\theta\} \cup R_1 \cup \ldots \cup R_q$ where θ is the equivalence relation whose blocks are $\{a\}$ and $A' = A - \{a\}$. Denote by $\mathcal{A}$ the algebra $(A; \mathrm{Pol}\Gamma)$.

For $c, d \in A$ we shall write $c \equiv d$ if $c = s_{iu}^0, d = t_{iv}^0$, or $c = s_{iu}^1, d = t_{iv}^1$, for some $s, t \in \{0, 1\}$, $u, v \in V$, and $i \in \{4, 5, 6, 7\}$. Furthermore, for $c, d \in A_v^0 \cup A_v^1$ we shall write $c \sim d$ if $c, d \in \{0_{iv}^0, 1_{iv}^0\}$ or $c, d \in \{0_{iv}^1, 1_{iv}^1\}$ some $v \in V$. Denote the tuples $(a, b, b), (b, a, b), (b, b, a), (a, a, a), (a, a, b), (a, b, a), (a, a, b), (b, b, b)$ by $\mathbf{a}^0, \ldots, \mathbf{a}^7$ respectively.

We prove two claims:

CLAIM 1. The class of all homomorphic images of subalgebras of $\mathcal{A}$ omits type **1**. Moreover, it omits type **2** if and only if $\mathcal{A}$ has a term operation satisfying one of the following two conditions

$$f \begin{pmatrix} a & b & a \\ b & a & a \\ b & b & a \end{pmatrix} = \begin{pmatrix} a \\ a \\ b \end{pmatrix} \quad \text{or} \quad f \begin{pmatrix} a & b & a \\ b & a & a \\ b & b & a \end{pmatrix} = \begin{pmatrix} b \\ b \\ a \end{pmatrix}.$$

CLAIM 2. $\mathcal{C}$ has a solution if and only if a term operation f exists satisfying the first condition.

Proof first Claim 1. As is easily seen $\{a, b\}$ is a subalgebra. Post's description of clones on a 2-element set [28] implies that if this subalgebra is not of type **1** or **2**, then it has either a semilattice, or majority term, and therefore, an operation satisfying one of the two conditions exists. Thus, if there is no such operation then $\{a, b\}$ is a subalgebra of type **1** or **2**.

If h' is an affine operation on $\{a, b\}$ then its action on the set $\{\mathbf{a}_0, \ldots, \mathbf{a}_7\}$ defines an affine operation

h'' on the set $\{0,\dots,7\}$. Let also $+$ denote addition modulo 2. It is not hard to see that the operation

$$g(x,y,z)=\begin{cases} h'(x,y,z), & \text{if } x,y,z\in\{a,b\},\\ a_{h''(i,j,k)v}, & \text{if } x=a_{iv}, y=a_{jv}, z=a_{kv}, \text{ or } x=a_{iv},\\ & y\in A^0_{jv}\cup A^1_{jv}, z\in A^0_{kv}\cup A^1_{v}, \text{ or}\\ & x\in A^0_{iv}\cup A^1_{iv}, y=a_{jv}, z\in A^0_{kv}\cup A^1_{kv},\\ & \text{or } x\in A^0_{iv}\cup A^1_{iv}, y\in A^0_{jv}\cup A^1_{jv}, z=a_{kv},\\ (s_1+s_2+s_3)^{(t_1+t_2+t_3)}_{h''(i,j,k)}, & \text{if } x=(s_1)^{t_1}_{iv}, y=(s_2)^{t_1}_{jv}, z=(s_3)^{t_3}_{kv},\\ s^t_{h''(i,j,k)v}, & \text{if } x=a_{iv}, y=a_{jv}, z=s^t_{kv}, \text{ or } x=a_{iv},\\ & y=s^t_{jv}, z=a_{kv}, \text{ or } x=s^t_{iv}, y=a_{jv}, z=a_{kv},\\ x, & \text{if } \{x,y,z\}\not\subseteq A_v \text{ and } \{x,y,z\}\not\subseteq\{a,b\}, \text{ and } (x,y)\in\theta,\\ y, & \text{if } \{x,y,z\}\not\subseteq A_v \text{ and } \{x,y,z\}\not\subseteq\{a,b\}, \text{ and } (x,z)\in\theta,\\ z, & \text{otherwise.}\end{cases}$$

is a polymorphism of Γ. This operation is an affine operation on $\{a,b\}$, and so it witnesses that this subalgebra has at least type **2**. Now we show that all other divisors of $\mathcal{A}$ have types **3**,**4**,**5**, which will follow from Claims 3 and 4.

CLAIM 3. If a subset of A is a subalgebra of $\mathcal{A}$, then it is either a singleton, or is one of $A, A',\ \{a,b\}$, or $C\cup D$ where $C\subseteq W_v$, $v\in V$, and $D\subseteq\bigcup_{v\in V}A^0_v\cup A^1_v$ such that if $|D\cap(A^0_{iv}\cup A^1_{iv})|\ge 2$ for some i then $A^0_{iv}\cup A^1_{iv}\subseteq D$.

Obviously, $A',\{a,b\},W_v\cup A^0_v\cup A^1_v$ are subalgebras, because they are projections of certain relations from Γ, or a class of θ. Take a subset $B\subseteq A$. Suppose first $B\ne A'$, $B\not\subseteq\{a,b\}$, $B\not\subseteq W_v\cup A^0_v\cup A^1_v$, for any $v\in V$, say $B=\{a_1,\dots,a_k\}$ ($a_1\ne a$). Then let k-ary operation $g_B(x_1,\dots,x_k)$ be such that

$$g_B(a_1,\dots,a_k)=c\in A'-B,$$
$$g_B(x_1,\dots,x_k)=x_1, \text{ otherwise.}$$

We have to prove that g_B preserves relations from Γ. By (a), (b), (c) we mark the parts of the proof corresponding to the parts of Γ: unary relations, equivalence relation θ, and the R_i.

(a) g_B obviously preserves unary one-element relations, but destroys B.

(b) Take $(x_1,y_1),\dots,(x_k,y_k)\in\theta$. Then $g_B\left(\begin{pmatrix}x_1\\y_1\end{pmatrix}\dots\begin{pmatrix}x_k\\y_k\end{pmatrix}\right)=\begin{pmatrix}x_1\\y_1\end{pmatrix}$ whenever neither $(x_1,\dots,x_k)$ nor $(y_1,\dots,y_k)$ equals $(a_1,\dots,a_k)$. If $(x_1,\dots,x_k)=(a_1,\dots,a_k)$ then $x_1\in A'$, and hence, $y_1\in A'$. Thus $g_B(x_1,\dots,x_k),g_B(y_1,\dots,y_k)\in A'$.

(c) For any $\mathbf{x}_1,\dots,\mathbf{x}_k\in R_i$ where $\mathbf{x}_j=(x_{1j},\dots,x_{6j})$, none of $(x_{l1},\dots,x_{lk})$ equals $(a_1,\dots,a_k)$. Hence, $g_B(\mathbf{x}_1,\dots,\mathbf{x}_k)=\mathbf{x}_1$ which means g_B preserves R_i.

Further, suppose that $|(A^0_{iv}\cup A^1_{iv})\cap B|\ge 2$, but $A^0_{iv}\cup A^1_{iv}\not\subseteq B$ for some $v\in V$ and $i\in\{4,5,6,7\}$. Let $c,d\in(A^0_{iv}\cup A^1_{iv})\cap B$ and $e\in(A^0_{iv}\cup A^1_{iv})-B$. There are 12 possibilities of what c,d,e are. We consider one of them, namely, when $c=0^0_{iv}, d=0^1_{iv}, e=1^0_{iv}$. The other 11 cases are quite similar. Define the operation $g_{c,d,e}(x,y)$ as follows

$$g_{c,d,e}(x,y)=\begin{cases} e, & \text{if } x=c, y=d,\\ 0^0_{iu}, & \text{if } x\in A^1_{iu}, y\in A^0_{iu}\cup A^1_{iu}, u\in V, x\not\sim y;\\ 1^0_{iu}, & \text{if } x\in A^0_{iu}, y\in A^0_{iu}\cup A^1_{iu}, u\in V, x\not\sim y;\\ x & \text{otherwise.}\end{cases}$$

a) is obvious.

b) Let $\begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = g_{c,d,e}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right)$ where $(x_1, x_2), (y_1, y_2) \in \theta$. If $x_1, x_2 \neq a$ then $(g_1, g_2) \in A'^2 \subseteq \theta$. If $x_1 = a$ then $x_2 = a$ and $(g_1, g_2) = (a, a) \in \theta$.

c) Let $\mathbf{g} = g_{c,d,e}(\mathbf{x}, \mathbf{y})$, $\mathbf{x}, \mathbf{y} \in R_i$. If (x_1, x_2, x_3) or (y_1, y_2, y_3) does not equal (a, a, b) if $i = 4$, (a, b, a) if $i = 5$, (b, a, a) if $i = 6$, and (b, b, b) if $i = 7$, then $\mathbf{g} = \mathbf{x}$. If (x_1, x_2, x_3) and (y_1, y_2, y_3) are equal to the corresponding triple, then $x_4 \equiv x_5 \equiv x_6$, $y_4 \equiv y_5 \equiv y_6$. Therefore (g_1, g_2, g_3) equals the same triple as $(x_1, x_2, x_3), (y_1, y_2, y_3)$; $g_4 \equiv g_5 \equiv g_6$; $(g_4, g_5, g_6) = (x_4, x_5, x_6)$ when $x_j \sim y_j$, $j \in \{4, 5, 6\}$, and $g_j = 0^0_{1v}$ if $x_j \in A^1_{1v}$, $g_j = 1^0_{1v}$ if $x_j \in A^0_{1v}$ when $x_j \not\sim y_j$, $j \in \{4, 5, 6\}$. In both cases $\mathbf{g} \in R_i$.

CLAIM 4. For each subalgebra $\mathcal{B}$ of $\mathcal{A}$, and a congruence η of $\mathcal{B}$, $\mathcal{B}/_\eta$ is not a set, and if $\{a, b\}$ has one of the operations f_1, f_2 then every such simple subalgebra also omits type **2**.

Note that if subalgebra $\{a, b\}^3$ has either a semilattice, or majority, or minority operation, then the subalgebra $W_v \cup A^0_v \cup A^1_v$ being factorised modulo η_v, whose blocks are $\{a_{0v}\}, \{a_{1v}\}, \{a_{2v}\}, \{a_{3v}\}, A^0_{4v} \cup A^1_{4v}, A^0_{5v} \cup A^1_{5v}, A^0_{6v} \cup A^1_{6v}, A^0_{7v} \cup A^1_{7v}$, also has such an operation. Moreover, this is also true for any subalgebra $\mathcal{B}$ of $W_v \cup A^0_v \cup A^1_v$ factorized modulo the restriction of η_v onto $\mathcal{B}$. Since $\mathcal{B}$ is idempotent, any its congruence class is a subalgebra. Claim 3 implies that every congruence of $\mathcal{B}$ is either the restriction of η_v, or is isomorphic to a certain divisor of $\{a, b\}^3$, or is non-trivial on one of the sets $A^0_{iv} \cup A^1_{iv}$. In the first and second cases operation g witnesses that such a divisor is not a set, and if an operation f satisfying one of the two conditions stated in Claim 1 is present then the divisor omits type **2**. So, the only subalgebras to check are $A, A', A^0_{iv} \cup A^1_{iv}$.

1) $A^0_{iv} \cup A^1_{iv}$ has no nontrivial subalgebras. Since it is idempotent, this implies simplicity. The following operation witnesses that this subalgebra omits types **1** and **2**:

$$h(x, y) = \begin{cases} s^0_{iv}, & \text{if } x = s^0_{iv}, y \in A^0_{iv} \cup A^1_{iv}, \\ & \text{or } x \in A^s_{iv} \text{ and } y \in \{0^0_{iv}, 1^0_{iv}\}, \\ x, & \text{otherwise.} \end{cases}$$

2) $\mathcal{A}/_\theta$ is isomorphic to $\{a, b\}$, and therefore, is not a set and if a term operation satisfying the conditions of Claim 1 is present it omits type **2**. Furthermore, any congruence of $\mathbb{A}$ which differs from θ and the total relation is not total on A'.

3) $\mathcal{A}' = \mathcal{A}|_{A'}$ is simple.

Since each congruence block is a subalgebra, any nontrivial congruence η is a subset of the equivalence relation η' whose blocks are $\{b\}$, $W_v \cup A^0_v \cup A^1_v$, $v \in V$. For each $c, d \in W_v \cup A^0_v \cup A^1_v$ define an operation $h_{c,d}$ as follows

$$h_{c,d}(x, y) = \begin{cases} c, & \text{if } x = c, y = b, \\ b, & \text{if } x = d, y = b, \\ x, & \text{otherwise.} \end{cases}$$

(a) $h_{c,d}$ is idempotent, consequently, it preserves all unary relations from Γ.

(b) $\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = h_{c,d}\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ whenever $a \in \{x_1, x_2, y_1, y_2\}$, and $h_1, h_2 \in A'$ otherwise.

(c) Set $\mathbf{h} = h_{c,d}(\mathbf{x}, \mathbf{y})$ where $\mathbf{x}, \mathbf{y} \in R_i$. Since $h_{c,d}(x, y) = x$ if $x, y \in \{a, b\}$, we have $(h_1, h_2, h_3) = (x_1, x_2, x_3)$. Further, since $y_4, y_5, y_6 \neq b$, $(h_4, h_5, h_6) = (x_4, x_5, x_6)$.

This means that if $(c, d) \in \eta$, then $(c, b) \in \eta$, but $(c, b) \notin \eta'$, a contradiction with $\eta \subseteq \eta'$. Operation h constructed above guarantees that $\mathcal{A}'$ omits types **1,2**. Claim 4 is proved. This also completes the proof of Claim 1.

Let us prove Claim 2. Note that if f_1 or f_2 are present, then for each C_i with $V_i = \{u, v, w\}$ we have

$$\begin{aligned} f(a_{0u}, a_{1u}, a_{2u}, a_{3u}) &\in A_u^a \\ f(a_{0v}, a_{1v}, a_{2v}, a_{3v}) &\in A_v^b \\ f(a_{0w}, a_{1w}, a_{2w}, a_{3w}) &\in A_w^c \end{aligned}$$

where (a, b, c) is a solution for C_i. Therefore, $\mathcal{C}$ has a solution.

Conversely, suppose $\mathcal{C}$ has a solution $\varphi\colon V \to \{0, 1\}$. Let $\cdot$ denote the semilattice operation on $\{a, b\}$ ($a \cdot b = b \cdot a = a$), its extention to the set $\{\mathbf{a}_0, \ldots, \mathbf{a}_7\}$, and the corresponding operation on $\{0, \ldots, 7\}$. Set

$$f_1(x, y, z, t) = g(x, y) = \begin{cases} x, & \text{if } x = y, \\ a, & \text{if } x = a \text{ or } y = a, \\ \varphi(v)^0_{(i\cdot j)v}, & \text{if } x \in \{a_{iv}, 0^0_{iv}, 0^1_{iv}, 1^0_{iv}, 1^1_{iv}\}, \\ & y \in \{a_{jv}, 0^0_{jv}, 0^1_{jv}, 1^0_{jv}, 1^1_{jv}\}, \text{ and } i \cdot j \in \{4, 5, 6, 7\}, \\ a_{(i\cdot j)v}, & \text{if } x \in \{a_{iv}, 0^0_{iv}, 0^1_{iv}, 1^0_{iv}, 1^1_{iv}\}, \\ & y \in \{a_{jv}, 0^0_{jv}, 0^1_{jv}, 1^0_{jv}, 1^1_{jv}\}, \text{ and } i \cdot j \in \{0, 1, 2, 3\}, \\ x, & \text{otherwise}. \end{cases}$$

(a) Since f is idempotent, it preserves all unary relations from Γ.

(b) As is easily seen, f preserves A'. Take $(x_1, x_2), (y_1, y_2) \in \theta$. If $a \notin \{x_1, x_2, y_1, y_2\}$ then $g\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \in A'^2 \subseteq \theta$. Otherwise, if $x_1 = a$ ($y_1 = a$) then $x_2 = a$ ($y_2 = a$); therefore $g\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = \begin{pmatrix} a \\ a \end{pmatrix} \in \theta$.

(c) Denote the set of tuples $\{\mathbf{b} \mid \begin{pmatrix} \mathbf{c}^i \\ \mathbf{b} \end{pmatrix} \in R_t\}$ by D_{it}. We just have to prove that $g(\mathbf{c}^i, \mathbf{c}^j) \in \{\mathbf{c}^0, \ldots, \mathbf{c}^7\}$, and if $g(\mathbf{c}^i, \mathbf{c}^j) = \mathbf{c}^k$ then $g(D_{it}, D_{jt}) \subseteq D_{kt}$. However, this follows straightforwardly from the definition of g. $\square$