



Finitely Related Algebras in Congruence Distributive Varieties Have Near Unanimity Terms

Libor Barto

Abstract. We show that every finite, finitely related algebra in a congruence distributive variety has a near unanimity term operation. As a consequence we solve the near unanimity problem for relational structures: it is decidable whether a given finite set of relations on a finite set admits a compatible near unanimity operation. This consequence also implies that it is decidable whether a given finite constraint language defines a constraint satisfaction problem of bounded strict width.

1 Introduction

Since the beginning of the systematic study of universal algebras in the 1930's it has been recognized that an important class of invariants of algebras and classes of algebras are their congruence lattices. Particularly widely studied objects are congruence distributive varieties, *i.e.*, equationally definable classes of algebras whose congruence lattices are distributive (see Section 2 for definitions).

We call an algebra in a congruence distributive variety a CD algebra. Examples of CD algebras include lattices, and, more generally, algebras that have a near unanimity term operation. These operations have also attracted a great deal of attention, not only in universal algebra, but also in graph theory and, recently, in computer science in connection with the constraint satisfaction problem (CSP), where, for instance, near unanimity operations characterize CSPs of bounded strict width [11].

Every finite algebra is, in some sense, determined by a set of relations. We call an algebra finitely related if this set of relations can be chosen to be finite. A useful corollary of a classical result of Baker and Pixley [2] is that every algebra with a near unanimity term operation is finitely related. Our main result provides a partial converse.

Theorem 1.1 *Every finite, finitely related CD algebra has a near unanimity term operation.*

A special case of this theorem for algebras determined by posets was conjectured in [10, 23]. An affirmative answer was given in [24] for bounded posets and in [19] in

Received by the editors May 3, 2011; revised October 14, 2011.

Published electronically December 24, 2011.

Research supported by the Grant Agency of the Czech Republic under the grant No. 201/09/P223 and by the Ministry of Education of the Czech Republic under the grant No. MSM 0021620839 and No. MEB 040915.

AMS subject classification: 08B05, 08B10.

Keywords: congruence distributive variety, Jónsson operations, near unanimity operation, finitely related algebra, constraint satisfaction problem.

full generality. Another special case of the theorem, namely, for algebras determined by a reflexive undirected graph, was proved in [18]. The general version is commonly referred to as the Zádori conjecture, although it has been never stated in a journal paper, perhaps because of scant evidence.

What made this result possible is the connection between the constraint satisfaction problem and universal algebra discovered in [8, 15]. The interaction between these areas is very fruitful in both directions. On one hand, universal algebra has brought a deeper understanding and strong results about the CSP. On the other hand, the CSP has motivated much of the recent work in universal algebra and opened new research directions. This is nicely illustrated by the main result of [3] (Theorem 5.7 in this paper). This theorem contributed to the study of local consistency methods for the CSP (and was an important step toward the full characterization of applicability of local consistency methods given in [4]), and it is also one of the two main ingredients of the proof of our main, purely algebraic result.

We remark that none of the assumptions of Theorem 1.1 is superfluous. In [24], Zádori provides an example of an infinite, bounded poset that determines a CD algebra with no near unanimity term operation. A simple example of a finite CD algebra with no near unanimity operation is the two element set $\{0, 1\}$ together with the implication regarded as a binary operation. Finally, the algebra determined by the complete loopless graph with three vertices does not have any near unanimity operation (it actually has no idempotent operations other than projections).

Of independent interest is a corollary of the main theorem (Corollary 7.1), which gives an affirmative answer to the near unanimity problem for relational structures. It is decidable whether a given set of relations on a finite set admits a compatible near unanimity operation. This consequence is discussed in more detail in Section 7.

1.1 Organization of the Paper

In Section 2 we recall basic notions and results about algebras and relational structures. In Section 3 we show that it is enough to deal with algebras determined by at most binary relations. In Section 4 we associate with such an algebra an instance of the CSP whose solutions are term operations of that algebra. The definitions and results about CSP instances that we require are stated in Section 5, where we also prove the main theorem. The main new tool is only stated in this section; its proof covers Section 6. Finally, in Section 7 we discuss consequences and open problems.

2 Preliminaries

In this section we recall universal algebraic notions and results that will be needed throughout the paper. This material, except for the notion of a Jónsson ideal, is covered in any standard reference on universal algebra, for example, [9].

2.1 Algebras and Varieties

An n -ary operation on a set A is a mapping $f: A^n \rightarrow A$. In this paper we assume that all operations are finitary, *i.e.*, n is a natural number. An operation is *idempotent* if it

satisfies the identity $f(a, a, \dots, a) = a$, *i.e.*, this equation holds for every $a \in A$. An operation of arity at least three is called a *near unanimity operation*, if it satisfies the identity

$$f(a, a, \dots, a, b, a, a, \dots, a) = a$$

for every position of b in the tuple.

An *algebra* is a pair $\mathbf{A} = (A, \mathcal{F})$, where A is a set, called the *universe* of \mathbf{A} , and \mathcal{F} is a set (possibly indexed) of operations on A . We use a boldface letter to denote an algebra and the same letter in plain type to denote its universe. An algebra is *idempotent* if all of its operations are idempotent. Two algebras are *similar* if their operations are indexed by the same set and corresponding operations have the same arities.

A *term operation* of \mathbf{A} is an operation that can be obtained from operations in \mathbf{A} using composition and the projection operations. The set of all term operations of \mathbf{A} is denoted by $\text{Clo}(\mathbf{A})$. Most structural properties of an algebra (such as subalgebras, congruences, automorphisms, *etc.*) depend only on the set of term operations rather than on a particular choice of the basic operations.

There are three fundamental operations on algebras of a fixed similarity type: forming subalgebras, factor algebras, and products.

A subset B of the universe of an algebra \mathbf{A} is called a *subuniverse* if it is closed under all operations (equivalently term operations) of \mathbf{A} . Given a subuniverse B of \mathbf{A} we can form the algebra \mathbf{B} by restricting all the operations of \mathbf{A} to the set B . In this situation we say that \mathbf{B} is a *subalgebra* of \mathbf{A} and we write $B \leq \mathbf{A}$ or $\mathbf{B} \leq \mathbf{A}$.

The *product* of algebras $\mathbf{A}_1, \dots, \mathbf{A}_n$ is the algebra with the universe equal to $A_1 \times \dots \times A_n$ and with operations computed coordinatewise. The product of n copies of an algebra \mathbf{A} is denoted by \mathbf{A}^n . A subalgebra (or a subuniverse) of a product of \mathbf{A} is called a *subpower* of \mathbf{A} .

An equivalence relation \sim on the universe of an algebra \mathbf{A} is a *congruence* if it is a subalgebra of \mathbf{A}^2 . The corresponding *factor algebra* \mathbf{A}/\sim has, as its universe, the set of \sim -blocks and operations that are defined using arbitrarily chosen representatives. The set of congruences of \mathbf{A} forms a lattice, called the *congruence lattice* of \mathbf{A} .

A *variety* is a class of similar algebras closed under forming subalgebras, products (possibly infinite), factor algebras, and isomorphic copies. A fundamental theorem of universal algebra, due to G. Birkhoff, states that a class of similar algebras is a variety if and only if this class can be defined via a set of identities.

2.2 Relational Structures

An *n-ary relation* on a set A is a subset of A^n (again, n is always finite in this article). A relational structure is a pair $\mathbb{A} = (A, \mathcal{R})$, where A is the *universe* of \mathbb{A} and \mathcal{R} is a set of relations on A . We use blackboard bold letters to denote relational structures.

We say that an operation $f: A^n \rightarrow A$ is *compatible* with a relation $R \subseteq A^m$ (or, R is *preserved* by f) if the tuple

$$(f(a_1^1, a_1^2, \dots, a_1^n), f(a_2^1, a_2^2, \dots, a_2^n), \dots, f(a_m^1, a_m^2, \dots, a_m^n))$$

belongs to R whenever $(a_1^i, a_2^i, \dots, a_m^i) \in R$ for all $i \leq n$. In other words, f is compatible with R , if R is a subpower of the algebra $(A, \{f\})$.

An operation compatible with all relations of a relational structure \mathbb{A} is a *polymorphism* of \mathbb{A} . The set of n -ary polymorphisms of \mathbb{A} is denoted by $\text{Pol}_n(\mathbb{A})$, and the set of all polymorphisms of \mathbb{A} is denoted by $\text{Pol}(\mathbb{A})$. This set of operations is closed under composition and contains the projection operations. On the other hand, every set of operations on a finite set closed under projections and composition can be obtained in this way.

Theorem 2.1 ([6, 13]) *For every finite algebra \mathbf{A} there exists a relational structure \mathbb{A} (with the same universe) such that $\text{Pol}(\mathbb{A}) = \text{Clo}(\mathbf{A})$.*

An algebra is called *finitely related* if finitely many relations suffice to determine $\text{Clo}(\mathbf{A})$.

Definition 2.2 An algebra \mathbf{A} is said to be *finitely related*, if there exists a relational structure \mathbb{A} with finitely many relations such that $\text{Pol}(\mathbb{A}) = \text{Clo}(\mathbf{A})$.

By a classic result of Baker and Pixley [2], every algebra with a near unanimity term operation is finitely related. More generally, every algebra with few subpowers is finitely related [5] (see Subsection 7.3).

2.3 CD Algebras

Definition 2.3 A variety is called *congruence distributive*, if all algebras in it have distributive congruence lattices. A *CD algebra* is an algebra in a congruence distributive variety.

A theorem of Jónsson [16] characterizes CD algebras using operations satisfying certain identities.

Definition 2.4 A sequence p_0, p_1, \dots, p_s of ternary operations on a set A is called a *Jónsson chain*, if the following identities are satisfied:

$$\begin{aligned} p_0(a, b, c) &= a, \\ p_s(a, b, c) &= c, \\ p_i(a, b, a) &= a && \text{for all } i \leq s, \\ p_i(a, a, b) &= p_{i+1}(a, a, b) && \text{for all even } i < s, \\ p_i(a, b, b) &= p_{i+1}(a, b, b) && \text{for all odd } i < s. \end{aligned}$$

Theorem 2.5 ([16]) *An algebra \mathbf{A} has a Jónsson chain of term operations if and only if \mathbf{A} is a CD algebra.*

Example Every algebra with a near unanimity term operation t is a CD algebra.

This can be shown, for instance, by constructing a Jónsson chain:

$$\begin{aligned}
p_0(a, b, c) &= a \\
p_1(a, b, c) &= t(a, a, \dots, a, b, c) \\
p_2(a, b, c) &= t(a, a, \dots, a, c, c) \\
p_3(a, b, c) &= t(a, a, \dots, a, b, c, c) \\
p_4(a, b, c) &= t(a, a, \dots, a, c, c, c) \\
&\dots
\end{aligned}$$

A useful notion for studying CD algebras is a Jónsson ideal.

Definition 2.6 Let \mathbf{A} be a CD algebra with Jónsson chain of term operations p_0, p_1, \dots, p_s . A subuniverse B of \mathbf{A} is a *Jónsson ideal*, if $p_i(b_1, a, b_2) \in B$ for every $a \in A, b_1, b_2 \in B$ and every $i \leq n$.

Every one element subuniverse of a CD algebra is its Jónsson ideal. Therefore, if \mathbf{A} is an idempotent CD algebra, then every singleton is a Jónsson ideal of \mathbf{A} .

3 Reduction to Binary Structures

In this section we show that to prove the main result it is enough to consider algebras determined by *binary relational structures*, i.e., relational structures with at most binary relations. This will make the presentation technically easier.

Proposition 3.1 Let \mathbb{A} be a relational structure whose relations all have arity at most k . Then there exists a binary relational structure $\bar{\mathbb{A}}$ with universe $\bar{A} = A^k$ such that

$$\text{Pol}(\bar{\mathbb{A}}) = \{\bar{f} : f \in \text{Pol}(\mathbb{A})\},$$

where \bar{f} is defined (if f is n -ary) by

$$\begin{aligned}
\bar{f}((a_1^1, a_2^1, \dots, a_k^1), (a_1^2, \dots, a_k^2), \dots, (a_1^n, \dots, a_k^n)) = \\
(f(a_1^1, a_1^2, \dots, a_1^n), f(a_2^1, \dots, a_2^n), \dots, f(a_k^1, \dots, a_k^n)).
\end{aligned}$$

Proof First we replace every relation R in \mathbb{A} with arity $l < k$ by the k -ary relation $R \times A^{k-l}$. This clearly does not change the set of polymorphisms, therefore we may assume that every relation in \mathbb{A} has arity precisely k .

Next we introduce the relations in $\bar{\mathbb{A}}$. For every k -ary relation R (on A) in \mathbb{A} we include in $\bar{\mathbb{A}}$ the unary relation R (on $\bar{A} = A^k$), and for every pair $1 \leq i, j \leq k$ we add a binary relation σ_{ij} defined by

$$((a_1, \dots, a_k), (b_1, \dots, b_k)) \in \sigma_{ij} \quad \text{if and only if} \quad a_i = b_j.$$

It is straightforward to check that $\bar{f} \in \text{Pol}(\bar{\mathbb{A}})$ for every polymorphism f of \mathbb{A} .

To prove the converse inclusion, let $h: (A^k)^n \rightarrow A^k$ be a polymorphism of $\bar{\mathbb{A}}$ and let h_1, \dots, h_k be its components, that is,

$$h((a_1^1, \dots, a_k^1), (a_1^2, \dots, a_k^2), \dots, (a_1^n, \dots, a_k^n)) = \\ \left(h_1((a_1^1, \dots, a_k^1), (a_1^2, \dots, a_k^2), \dots, (a_1^n, \dots, a_k^n)), \dots, \right. \\ \left. h_k((a_1^1, \dots, a_k^1), (a_1^2, \dots, a_k^2), \dots, (a_1^n, \dots, a_k^n)) \right).$$

For any i , $1 \leq i \leq k$, the relation σ_{ii} is compatible with h . Therefore for any elements $a_1^1, a_2^1, \dots, a_1^2, \dots, a_1^n, b_1^1, \dots, b_k^n \in A$ such that, for all $1 \leq l \leq n$, $a_i^l = b_l^i$ we have

$$h_i((a_1^1, \dots, a_k^1), (a_1^2, \dots, a_k^2), \dots, (a_1^n, \dots, a_k^n)) = \\ h_i((b_1^1, \dots, b_k^1), (b_1^2, \dots, b_k^2), \dots, (b_1^n, \dots, b_k^n)).$$

In other words, $h_i((a_1^1, \dots), \dots)$ depends only on $a_i^1, a_i^2, \dots, a_i^n$, and thus there exists an n -ary operation f_i on A such that

$$h_i((a_1^1, \dots, a_k^1), (a_1^2, \dots, a_k^2), \dots, (a_1^n, \dots, a_k^n)) = f_i(a_i^1, a_i^2, \dots, a_i^n).$$

Now we use the relations σ_{ij} for $i \neq j$. For any a_1^1, \dots, a_k^n we choose arbitrarily b_1^1, \dots, b_k^n so that $a_i^l = b_l^i$ (for all $1 \leq l \leq n$). As σ_{ij} is compatible with h , it follows that $f_i(a_i^1, \dots, a_i^n) = f_j(b_j^1, \dots, b_j^n) = f_j(a_i^1, \dots, a_i^n)$. Therefore, $f_1 = f_2 = \dots = f_k$.

We have shown that $h = \bar{f}$ for certain n -ary operation f on A . As each relation R of \mathbb{A} is compatible with h it follows that f is a polymorphism of \mathbb{A} . ■

The proposition implies that Theorem 1.1 follows from the following theorem, to be proved later.

Theorem 3.2 *If \mathbb{A} is a finite binary relational structure such that $(A, \text{Pol}(\mathbb{A}))$ is a CD algebra, then \mathbb{A} has a near unanimity polymorphism.*

Proof of Theorem 1.1 assuming Theorem 3.2 Let \mathbf{A} be a finite, finitely related CD algebra and let \mathbb{A} be a relational structure with finitely many relations (say all of them have arity at most k) such that $\text{Pol}(\mathbb{A}) = \text{Clo}(\mathbf{A})$. Let $\bar{\mathbb{A}}$ be the relational structure from the previous proposition. Then $\bar{\mathbf{A}} = (A, \text{Pol}(\bar{\mathbb{A}}))$ is a CD algebra, since $\bar{p}_0, \dots, \bar{p}_s$ is a Jónsson chain of $\bar{\mathbf{A}}$ whenever p_0, \dots, p_s is a Jónsson chain of \mathbf{A} . By Theorem 3.2, $\bar{\mathbb{A}}$ has a near unanimity polymorphism h . Using Proposition 3.1 again we have $h = \bar{f}$ for some polymorphism f of \mathbb{A} , and f is clearly a near unanimity operation. ■

4 CSP Instance Associated with a Binary Relational Structure

Definition 4.1 *An instance of the constraint satisfaction problem (CSP) is a triple $P = (V, A, \mathcal{C})$ with*

- V a nonempty, finite set of *variables*,
- A a nonempty, finite *domain*,
- \mathcal{C} a finite set of *constraints*, where each constraint is a pair $C = (\mathbf{x}, R)$ with
 - \mathbf{x} a tuple of variables of length n , called the *scope* of C , and
 - R an n -ary relation on A , called the *constraint relation* of C .

Let \mathbf{A} be a finite idempotent algebra. An *instance of the CSP over \mathbf{A}* , denoted by $\text{CSP}(\mathbf{A})$, is an instance such that all constraint relations are subpowers of \mathbf{A} .

A *solution* to an instance P is a function $f: V \rightarrow A$ such that, for each constraint $C = (\mathbf{x}, R) \in \mathcal{C}$, the tuple $f(\mathbf{x})$ belongs to R .

Remark 4.2 The CSP is often parametrized by relational structures: an instance whose constraint relations are in a relational structure \mathbb{A} is called an instance of $\text{CSP}(\mathbb{A})$. It was proved in [15] that the computational complexity of deciding whether an instance of $\text{CSP}(\mathbb{A})$ has a solution is fully determined, at least when \mathbb{A} has finitely many relations, by the algebra $\mathbf{A} = (A, \text{Pol}(\mathbb{A}))$. Moreover, Bulatov, Jeavons, and Krokhin proved in [8] that the complexity depends only on the variety generated by \mathbf{A} (*i.e.*, the smallest variety containing \mathbf{A}). These results are at the heart of the connection between universal algebra and the CSP mentioned in the introduction.

For simplicity we will formulate our definitions and results for a special type of CSP instance with a single binary constraint for each pair of variables, although most of the material can be generalized.

Definition 4.3 An instance $P = (V, A, \mathcal{C})$ of the CSP is called a *simple binary instance* if

- $\mathcal{C} = \{((x_1, x_2), R_{x_1, x_2}^P) : x_1, x_2 \in V\}$,
- $R_{x_2, x_1}^P = R_{x_1, x_2}^{P^{-1}} (= \{(b, a) : (a, b) \in R_{x_1, x_2}^P\})$ for every $x_1, x_2 \in V$, and
- $R_{x, x}^P \subseteq \{(a, a) : a \in A\}$ for every $x \in V$.

We omit the superscript P if the instance is clear from the context.

A simple binary instance can be drawn as a $|V|$ -partite graph in the following way. Each part is a copy of A , one for each variable $x \in V$ (the parts are now commonly referred to as potatoes), and elements of R_{x_1, x_2} are edges between the corresponding copies of A . Solutions then correspond to cliques with V vertices (with one vertex in each part).

To every binary relational structure \mathbb{A} and natural number n we can associate, in a natural way, a simple binary instance $P(\mathbb{A}, n)$ of $\text{CSP}((A, \text{Pol}(\mathbb{A})))$ whose solutions are precisely the n -ary polymorphisms of \mathbb{A} .

Definition 4.4 Let \mathbb{A} be a binary relational structure and let $n \geq 2$ be a natural number. The instance $P(\mathbb{A}, n) = (V, A, \mathcal{C})$ is defined by

- $V = A^n$
- $R_{(a_1, \dots, a_n), (b_1, \dots, b_n)}^P = \{(t(a_1, \dots, a_n), t(b_1, \dots, b_n)) : t \in \text{Pol}_n(\mathbb{A})\}$

Note that $P(\mathbb{A}, n)$ is indeed an instance of $\text{CSP}((A, \text{Pol}(\mathbb{A})))$.

Proposition 4.5 *For every binary relational structure \mathbb{A} , the set of solutions of $P(\mathbb{A}, n)$ is equal to $\text{Pol}_n(\mathbb{A})$.*

Proof It is clear that every n -ary polymorphism of \mathbb{A} is a solution of $P(\mathbb{A}, n)$.

Let f be a solution to $P(\mathbb{A}, n)$. We have to show that every relation R of \mathbb{A} is preserved by f , but this is easy. If R is binary and n -tuples $(a_1, \dots, a_n), (b_1, \dots, b_n) \in A^n$ are such that $(a_i, b_i) \in R$ for each $1 \leq i \leq n$, then $(f(a_1, \dots, a_n), f(b_1, \dots, b_n)) \in R_{(a_1, \dots, a_n), (b_1, \dots, b_n)}$ (as f is a solution), therefore

$$(f(a_1, \dots, a_n), f(b_1, \dots, b_n)) = (t(a_1, \dots, a_n), t(b_1, \dots, b_n))$$

for some $t \in \text{Pol}_n(\mathbb{A})$. Since t is a polymorphism, the right hand side is an element of R . The proof for a unary relation R can be done, for instance, by using this reasoning for the relation $R \times A$. ■

We are interested in near unanimity polymorphisms, solutions f of $P(\mathbb{A}, n)$ satisfying the additional conditions $f(a, a, \dots, a, b, a, a, \dots, a) = a$ (for any $a, b \in A$ and any position of b in the tuple). Therefore the following notion of a restriction of an instance will be useful.

Definition 4.6 Let $P = (V, A, \mathcal{C})$ be a simple binary instance of CSP and let $\mathcal{J} = \{J_x : x \in V\}$ be a family of subsets of A . By the *restriction of P to \mathcal{J}* we mean the simple binary instance $P|_{\mathcal{J}} = (V, A, \mathcal{C}')$ with

$$R_{x_1, x_2}^{P|_{\mathcal{J}}} = R_{x_1, x_2}^P \cap (J_{x_1} \times J_{x_2})$$

for every $x_1, x_2 \in V$.

To find an n -ary polymorphism of a binary relational structure \mathbb{A} , we will consider the instance $P = P(\mathbb{A}, n)$ and its restriction to the family $\mathcal{J} = \{J_x : x \in V\}$, where $J_{(a, a, \dots, a, b, a, a, \dots, a)} = \{a\}$ (for every $a, b \in A$ and every position of b in the tuple), and $J_{(a_1, \dots, a_n)} = A$ otherwise. With this choice, the set of solutions of $P|_{\mathcal{J}}$ coincides with the set of n -ary near unanimity polymorphisms of \mathbb{A} . We show that this set is nonempty in two steps. First we prove that $P|_{\mathcal{J}}$ contains a subinstance that is “consistent enough”, and then we apply a result from [3] saying that such instances always have a solution.

In the next section we introduce the required consistency notions.

5 Consistency Notions, Proof of Theorem 3.2

5.1 (1, 2)-systems

Definition 5.1 Let $P = (V, A, \mathcal{C})$ be a simple binary instance and let $\{R_x : x \in V\}$ be a family of nonempty subsets of A . We say that P is a *(1, 2)-system with unary projections* $\{R_x : x \in V\}$, if, for any $x_1, x_2 \in V$, the projection of R_{x_1, x_2} to the first coordinate is equal to R_{x_1} . (It follows that the projection to the second coordinate is equal to R_{x_2} .)

If, moreover, \mathbf{A} is an algebra and P is an instance of $\text{CSP}(\mathbf{A})$ we say that P is a *(1, 2)-system over \mathbf{A}* .

Observe that if P is a $(1, 2)$ -system over \mathbf{A} , then each R_x is a subuniverse \mathbf{A} (since the set R_x is equal to the projection of $R_{x,x}$ to the first coordinate). In this case we denote the subalgebra of \mathbf{A} with universe R_x by \mathbf{R}_x .

Also note that the instance $P(\mathbf{A}, n)$ introduced in Section 4 is always a $(1, 2)$ -system with unary projections $\{R_x : x \in V\}$, where

$$R_{(a_1, \dots, a_n)} = \{t(a_1, \dots, a_n) : t \in \text{Pol}_n(\mathbf{A})\}.$$

When a simple binary instance P is drawn as a multipartite graph (see the note after Definition 4.3), then P is a $(1, 2)$ -system if and only if, for every pair x_1, x_2 of variables, every vertex $a \in R_{x_1}$ is adjacent to at least one vertex from R_{x_2} and to no vertex outside R_{x_2} (in particular, vertices outside the sets R_x are isolated).

Whether an instance has a restriction that is a (nonempty) $(1, 2)$ -system can be decided using trees.

Definition 5.2 Let $P = (V, A, \mathcal{C})$ be a simple binary instance. A P -tree T is a tree (i.e., an undirected connected graph without loops or cycles) whose vertices are labeled by variables in V . The vertex set of T is denoted by $\text{vert}(T)$ and the label of a vertex $v \in \text{vert}(T)$ by $\text{lbl}(v)$.

A *realization* of a P -tree T in P is a mapping $r: \text{vert}(T) \rightarrow A$ such that $(r(v_1), r(v_2)) \in R_{\text{lbl}(v_1), \text{lbl}(v_2)}$ whenever v_1, v_2 are adjacent vertices of T . For a vertex v of T we put

$$T[v] = \{r(v) : r \text{ is a realization of } T \text{ in } P\}.$$

If P is a $(1, 2)$ -system with unary projections $\{R_x : x \in V\}$, then every P -tree is clearly realizable. Moreover, for every P -tree T and every vertex v of T , we have $T[v] = R_{\text{lbl}(v)}$. The following proposition provides a converse to this observation.

Proposition 5.3 Let $P = (V, A, \mathcal{C})$ be a simple binary instance over an algebra \mathbf{A} . If every P -tree is realizable in P then, for every $x \in V$, the set

$$R_x = \bigcap_{\substack{T \text{ is a } P\text{-tree} \\ v \in \text{vert}(T) \\ \text{lbl}(v)=x}} T[v]$$

is nonempty and $P_{\{R_x : x \in V\}}$ is a $(1, 2)$ -system over \mathbf{A} .

Proof Since A is a finite set, each R_x can be obtained by intersecting the sets $T[v]$ for only finitely many P -trees T . Moreover, there exists a single tree T_x with vertex v_x labeled by x such that $R_x = T_x[v_x]$. We take the disjoint union of the finite collection of trees and identify the vertices v to a single vertex. It follows that R_x is nonempty for every $x \in V$.

Next we prove that $P_{\{R_x : x \in V\}}$ is a $(1, 2)$ -system. It is enough to show that for every $x_1, x_2 \in V$ and every $a_1 \in R_{x_1}$ there exists $a_2 \in R_{x_2}$ such that $(a_1, a_2) \in R_{x_1, x_2}$. Consider the P -tree T constructed from T_{x_2} by adding a vertex v_1 adjacent to v_{x_2} with label x_1 . This P -tree has a realization r such that $r(v_1) = a_1$ (since $R_{x_1} \subseteq T[v_1]$). Now we can put $a_2 = r(v_{x_2})$, because

$$(r(v_1), r(v_2)) \in R_{x_1, x_2} \quad \text{and} \quad r(v_2) \in T[v_{x_2}] \subseteq T_{x_2}[v_{x_2}] = R_{x_2}.$$

Finally, we have to show that $P_{\{R_x : x \in V\}}$ is an instance of $\text{CSP}(\mathbf{A})$. It is clearly enough to prove that $R_x (= T_x[v_x])$ is a subuniverse of \mathbf{A} for every $x \in V$. But this is a straightforward consequence of the definitions: for any P -tree T , any operation t of \mathbf{A} (say, k -ary) and any k -tuple of realizations r_1, \dots, r_k of T in P , the mapping r defined by $r(v) = t(r_1(v), \dots, r_k(v))$ is a realization of T in P (as R_{x_1, x_2} is a subuniverse of \mathbf{A}^2 for every $x_1, x_2 \in V$). ■

Remark 5.4 The family $\mathcal{R} = \{R_x : x \in V\}$ from the previous proposition is actually the largest family such that $P_{|\mathcal{R}}$ is a $(1, 2)$ -system. Also observe that if some P -tree is not realizable, then no such a family exists.

5.2 $(2, 3)$ -systems

A $(2, 3)$ -system is a $(1, 2)$ -system such that every edge extends to a triangle:

Definition 5.5 A $(1, 2)$ -system $P = (V, A, \mathcal{C})$ is called a $(2, 3)$ -system if for every $x_1, x_2, x_3 \in V$ and every $(a_1, a_2) \in R_{x_1, x_2}$ there exists $a_3 \in A$ such that $(a_1, a_3) \in R_{x_1, x_3}$ and $(a_2, a_3) \in R_{x_2, x_3}$.

Examples of $(2, 3)$ -system include the instances $P(\mathbb{A}, n)$.

The following theorem is the main new ingredient for the proof of the Zádori conjecture. It is proved in Section 6.

Theorem 5.6 Let $P = (V, A, \mathcal{C})$ be a $(2, 3)$ -system with unary projections $\{R_x : x \in V\}$ over a CD algebra \mathbf{A} and let $\mathcal{J} = \{J_x : x \in V\}$ be a family of (nonempty) subsets of A such that each J_x is a Jónsson ideal of \mathbf{R}_x . If all P -trees with at most $4^{8^{|\mathbf{A}|}}$ vertices are realizable in $P_{|\mathcal{J}}$, then all P -trees are realizable in $P_{|\mathcal{J}}$.

The core result of [3] states that every $(2, 3)$ -system over a CD algebra has a solution ([3, Theorem 5.2]). We will need a refinement proved (although not explicitly stated) in the same article.

Theorem 5.7 Let $P = (V, A, \mathcal{C})$ be a $(2, 3)$ -system with unary projections $\{R_x : x \in V\}$ over a CD algebra \mathbf{A} and let $\mathcal{J} = \{J_x : x \in V\}$ be a family of (nonempty) subsets of A such that each J_x is a Jónsson ideal of \mathbf{R}_x . If $P_{|\mathcal{J}}$ is a $(1, 2)$ -system, then $P_{|\mathcal{J}}$ has a solution.

Remark 5.8 The method used to prove [3, Theorem 5.2] was the following. If \mathcal{J} satisfies the assumptions (such families are called absorbing systems in [3]) and some of the sets J_x have more than one element, then it is possible ([3, Lemma 6.9]) to find a family $\mathcal{J}' = \{J'_x : x \in V\}$ such that \mathcal{J}' satisfies the same conditions, $J'_x \subseteq J_x$ and at least one of these inclusions is proper. In this way we eventually get a solution to $\mathcal{P}_{|\mathcal{J}}$.

More recently, the result that every $(2, 3)$ -system over a CD algebra \mathbf{A} has a solution was generalized in two directions. First, a weaker consistency notion than $(2, 3)$ -system is enough to guarantee a solution. It suffices to assume that the instance is a so-called *Prague strategy* (see [4]). A more “modern” proof of Theorem 5.7 would be to show that $P_{|\mathcal{J}}$ is a Prague strategy (which is not hard).

A weaker condition can also be put on the algebra. It is enough to assume that \mathbf{A} lies in a meet semi-distributive variety (actually, for an idempotent finite algebra \mathbf{A} , the statement “every (2, 3)-system (or Prague strategy) over \mathbf{A} has a solution” is equivalent to “ \mathbf{A} lies in a meet semi-distributive variety” [4]).

5.3 Proof of the Zádori Conjecture

We are ready to prove the main theorem. As discussed in Section 3, it is enough to prove Theorem 3.2.

Theorem 3.2 *If \mathbb{A} is a finite binary relational structure such that $(A, \text{Pol}(\mathbb{A}))$ is a CD algebra, then \mathbb{A} has a near unanimity polymorphism.*

Proof Let p_0, \dots, p_s be a Jónsson chain of operations of the algebra $(A, \text{Pol}(\mathbb{A}))$. Let \mathbf{A} be the algebra with universe A whose operations are *idempotent* polymorphisms of \mathbb{A} . Since p_i 's are idempotent, p_0, \dots, p_s is a Jónsson chain for the algebra \mathbf{A} .

Let n be a natural number greater than $4^{8^{|A|}}$ and let $P = P(\mathbb{A}, n)$. It was observed above that P is a (2, 3)-system over \mathbf{A} with unary projections $\{R_x : x \in V\}$.

Let $\mathcal{J} = \{J_x : x \in V\}$, where $J_{(a,a,\dots,a,b,a,a,\dots,a)} = \{a\}$ (for every $a, b \in A$ and every position of b in the tuple), and $J_{(a_1,\dots,a_n)} = A$ otherwise. Since \mathbf{A} is idempotent, each J_x is a Jónsson ideal of \mathbf{R}_x . As discussed in Section 4, the solutions to the instance $Q = P|_{\mathcal{J}}$ are n -ary near unanimity polymorphisms of \mathbb{A} ; therefore, it is enough to show that Q has a solution.

First we observe that every P -tree T with at most $n - 1$ vertices is realizable in Q . Indeed, the variables are n -tuples and T has less than n vertices; therefore, there exists a natural number i (with $1 \leq i \leq n$) such that b is *not* on the i -th position of any tuple $x = (a, a, \dots, a, b, a, a, \dots, a)$, $a \neq b$ which is a label of a vertex of T . Then the mapping assigning a_i to a vertex of label (a_1, \dots, a_n) is a realization of T in Q .

By Theorem 5.6 every tree is realizable in Q .

Proposition 5.3 (applied to the simple binary instance Q) gives us a system $\mathcal{J}' = \{J'_x : x \in V\}$ such that $Q|_{\mathcal{J}'}$ is a (1, 2)-system with unary projections \mathcal{J}' .

For every $x \in V$, J'_x is a Jónsson ideal of \mathbf{R}_x . Indeed, in the proof of Proposition 5.3 we have shown that $J'_x = T_x^Q[v_x]$ for certain tree T_x and its vertex x . If $a_1, a_2 \in J'_x$ and $b \in R_x$, then there exists a realization r_1 (resp. r_2) of T_x in Q such that $r_1(v_x) = a_1$ (resp. $r_2(v_x) = a_2$), and, since P is a (1, 2)-system, there exists a realization r_3 of T_x in P such that $r_3(v_x) = b$. Now we apply the Jónsson term operation p_i to r_1, r_2, r_3 (in the same way as in the last paragraph of the proof of Proposition 5.3), and we get a realization r of T_x in P such that $r(v_x) = p_i(a_1, b, a_2)$. From the assumption that J'_x is a Jónsson ideal of \mathbf{R}_x (for every $x' \in V$), it follows that r is a realization in Q . Therefore, $p_i(a_1, b, a_2) \in T_x^Q[v_x] = J'_x$.

Finally, P is a (2, 3)-system, \mathcal{J}' is formed by Jónsson ideals of appropriate R_x 's and $P|_{\mathcal{J}'}$ ($= Q|_{\mathcal{J}'}$) is a (1, 2)-system; thus, by Theorem 5.7, $P|_{\mathcal{J}'}$ has a solution, which is of course also a solution to $P|_{\mathcal{J}}$. \blacksquare

6 Proof of Theorem 5.6

Theorem 5.6 *Let $P = (V, A, \mathbb{C})$ be a $(2, 3)$ -system with unary projections $\{R_x : x \in V\}$ over a CD algebra \mathbf{A} and let $\mathcal{J} = \{J_x : x \in V\}$ be a family of subsets of A such that each J_x is a Jónsson ideal of \mathbf{R}_x . If all P -trees with at most $4^{|\mathcal{A}|}$ vertices are realizable in $P_{|\mathcal{J}}$, then all P -trees are realizable in $P_{|\mathcal{J}}$.*

We argue by contradiction. We take a tree that is not realizable in $P_{|\mathcal{J}}$, and we eventually obtain a configuration (a tuple $(\mathbf{B}, L, U, E, F, a, b)$) that will contradict the following auxiliary result. In this lemma we look at the binary relations E, F on B as digraphs.

Lemma 6.1 *Let \mathbf{B} be a finite CD algebra and let $U, L \subseteq B, E, F \leq \mathbf{B}^2, a, b \in B$ be such that*

- E is a Jónsson ideal of F ;
- U is disjoint from L ;
- $a \in U, b \in L, (a, b) \in F$;
- the digraph $E \cap U^2$ has no sources (that is, for all $c \in U$ there exists $d \in U$ such that $(d, c) \in E$);
- the digraph $E \cap L^2$ has no sinks (that is, for all $c \in L$ there exists $d \in L$ such that $(c, d) \in E$).

Then there exist $c \in U$ and $d \in B \setminus U$ such that $(c, d) \in E$.

Proof We take a counterexample to the lemma and fix a Jónsson chain p_0, p_1, \dots, p_s of term operations of \mathbf{B} . We may assume that \mathbf{B} is idempotent, otherwise we can replace \mathbf{B} by the algebra $(B, \{p_0, p_1, \dots, p_s\})$.

Let us quickly sketch the proof on a smallest choice that does not satisfy the conclusion:

$$\begin{aligned} B &= \{1, 2\}, & U &= \{1\}, & L &= \{2\}, & E &= \{(1, 1), (2, 2)\}, \\ F &= \{(1, 1), (2, 2), (1, 2)\}, & a &= 1, & b &= 2. \end{aligned}$$

The first step of the proof is to transform our counterexample into a form closer to this simplest one. Next we prove that E must at least contain the edge $(2, 1)$, and finally we show that in this case we would have a directed path from 1 to 2 in the digraph E .

Since $E \cap U^2$ has no sources, we can find a sequence $a = a_1, a_2, \dots$ of elements in U such that $(a_{i+1}, a_i) \in E$ for all i . As U is finite, there are positive numbers k and l such that $a_k = a_{k+l}$. Similarly, we find a sequence $b = b_1, b_2, \dots$ of elements in L such that $(b_i, b_{i+1}) \in E$ and positive numbers k' and l' such that $b_{k'} = b_{k'+l'}$. Let m be a natural number greater than or equal to $k + k' - 1$ and divisible by l and l' , let

$$E' = \underbrace{E \circ E \circ \dots \circ E}_{m\text{-times}}, \quad F' = \underbrace{F \circ F \circ \dots \circ F}_{m\text{-times}}, \quad U' = U, \quad \mathbf{B}' = \mathbf{B}, \quad a' = a_k,$$

where \circ denotes the composition of relations defined by

$$S \circ S' = \{(s, s'') : \exists s' \in B (s, s') \in S, (s', s'') \in S'\}$$

and let $b' \in \{b_{k'}, b_{k'+1}, \dots, b_{k'+l'}\}$ be an element such that there exists a directed path in the digraph F from a' to b' of length m (we can take the element of appropriate distance from a' on the path $a' = a_k, a_{k-1}, \dots, a_1, b_1, b_2, \dots, b_{k'}, b_{k'+1}, \dots, b_{k'+l'} = b_{k'}, b_{k'+1}, \dots$).

These new sets \mathbf{B}', E', F', U' and elements a', b' have the following properties.

- \mathbf{B}' is a CD algebra, $E' \leq F' \leq \mathbf{B}'^2$, E' is a Jónsson ideal of F' . This is straightforward. (That E', F' are subalgebras follows from a more general fact that any relation positively primitively defined from subpowers is a subpower, but it is easy to check the claims directly.)
- $a' \in U', b' \in B' \setminus U', (a', a') \in E', (b', b') \in E', (a', b') \in F'$. We have chosen b' so that there exists a directed path in F of length m between a' and b' , thus $(a', b') \in F'$. Since a' (resp. b') are in a closed path of length l (resp. l') and this length divides m , it follows that $(a', a'), (b', b') \in E'$.
- *There do not exist $c \in U', d \in B' \setminus U'$ such that $(c, d) \in E'$.* Otherwise there is a path in E from c to d in E , which is impossible as there is no edge from U' to $B' \setminus U'$.

We will show that it is impossible to find $\mathbf{B}', U', E', F', a', b'$ satisfying the conditions above. For contradiction, assume that $\mathbf{B}', U', E', F', a', b'$ satisfy these three conditions and $|B'|$ is the smallest possible.

The minimality assumption has some useful consequences.

- $(c, c) \in E'$ for any $c \in B'$. Otherwise the following choice would form a smaller counterexample: $B'' = \{c : (c, c) \in E'\}, U'' = U' \cap B'', E'' = E' \cap B''^2, F'' = F' \cap B''^2, a'' = a', b'' = b'$. That B'' is a subuniverse of \mathbf{B}' is straightforward to check (and it again follows from the general fact about positive primitive definitions of subpowers).
- $(a', c) \in F'$ for any $c \in B'$. Otherwise we could take $B'' = \{c : (a', c) \in F'\}$ and restrict all the sets to B'' as above, i.e., $U'' = U' \cap B'', E'' = E' \cap B''^2, F'' = F' \cap B''^2, a'' = a', b'' = b'$. Note that we need idempotency to show that B'' is a subuniverse of \mathbf{B}' (B'' is defined using F' and the subuniverse $\{a'\}$ of \mathbf{B}').
- $(c, d) \in F'$ for any $c \in U', d \in B' \setminus U'$. Otherwise we take $B'' = \{e : (e, d) \in F'\}, a'' = a', b'' = d$ and, again, restrict U', E', F' to B'' . From the first item it follows that $(b'', b'') \in E''$, and the second item implies that $(a'', b'') \in F''$.

Now we consider the sequence

$$\begin{aligned} a' &= p_1(a', a', b'), p_1(a', b', b') = p_2(a', b', b'), p_2(a', a', b') \\ &= p_3(a', a', b'), \dots, p_{s'}(a', b', b') = b', \end{aligned}$$

where $s' = s$ if s is odd and $s' = s - 1$ if s is even.

As $(a', a'), (b', b') \in E', (a', b') \in F'$, and E' is a Jónsson ideal of F' , it follows that the first pair of elements of this sequence is in E' . Similarly, the second pair is in E'^{-1} , the third pair in E' , and so on. Thus we have a “fence” in E' from a' to b' , and, since we are assuming that there are no $c \in U', d \in B' \setminus U'$ such that $(c, d) \in E'$, there must exist elements $c \in U'$ and $d \in B' \setminus U'$ such that $(d, c) \in E'$.

We have $(c, c), (d, d), (d, c) \in E'$ and $(c, d) \in F'$. It follows that

$$c = p_1(c, c, d), p_1(c, c, d) = p_2(c, d, d), p_2(c, d, d) = p_3(c, c, d), \dots, d$$

is a sequence where all the pairs are in E' . This contradicts the assumption that there is no element in U' that is E' -related to an element outside U' . ■

For the remainder of this section we fix P , R_x 's, and \mathcal{J} satisfying the hypotheses of Theorem 5.6, and we assume that there exists a tree that is not realizable in $P|_{\mathcal{J}}$.

To obtain a configuration contradicting the previous lemma we will first transform our non-realizable tree to a tree whose every vertex has degree 1 or 3 and that has no realization in P with leaves realized in $P|_{\mathcal{J}}$. We require the following definition.

Definition 6.2 Let T be a P -tree and let S be a subset of vertices of T . A realization r of T in P is called an S -realization if $r(v) \in J_{\text{lbl}(v)}$ for every $v \in S$.

For a vertex v of T we define

$$T_S[v] = \{r(v) : r \text{ is an } S\text{-realization of } T \text{ in } P\}$$

The set S from the definition will often be the set of all leaves of T , which we denote by $\text{leaves}(T)$.

Lemma 6.3 *There exists a P -tree T such that*

- *the degree of any vertex of T is 1 or 3;*
- *T has no $\text{leaves}(T)$ -realization;*
- *T has a S -realization for every proper subset S of $\text{leaves}(T)$.*

Proof We start with a P -tree T that is not realizable in $P|_{\mathcal{J}}$. To every inner vertex (*i.e.*, a vertex of degree greater than one) we add an adjacent vertex with the same label. Since $R_{x,x}$ is a subset of the equality relation, any realization maps the new leaf to the same element of A as the inner vertex. It follows that the new tree is not $\text{leaves}(T)$ -realizable.

In a similar way we can modify the tree so that all the vertices have degree at most 3. If a vertex v has degree at least 4, we can split it into two adjacent vertices v_1, v_2 with the same label in such a way that v_1 is adjacent to 2 of the original neighbors of v and v_2 is adjacent to the remaining neighbors. Clearly, v_1 and v_2 have smaller degree than v ; therefore, we can repeat this splitting procedure until we obtain a tree whose every vertex has degree at most 3 and that is not $\text{leaves}(T)$ -realizable.

Let T be such a tree with minimal number of vertices.

Now we show that T has no vertex of degree 2. Suppose otherwise, that is, there is a vertex v with precisely two neighbors v_1, v_2 . The tree T' obtained by removing the vertex v and adding the edge $v_1 - v_2$ is smaller than T , therefore T' has a $\text{leaves}(T')$ -realization r' . As $(r'(v_1), r'(v_2)) \in R_{\text{lbl}(v_1), \text{lbl}(v_2)}$ and P is a $(2, 3)$ -system, there exists $a \in A$ such that $(r'(v_1), a) \in R_{\text{lbl}(v_1), \text{lbl}(v)}$ and $(r'(v_2), a) \in R_{\text{lbl}(v_2), \text{lbl}(v)}$ (This is the only place in this section where we use the assumption that P is a $(2, 3)$ -system. For the rest it would suffice to assume that P is a $(1, 2)$ -system.) It follows that the extension r of the mapping r' by $r(v) = a$ is a $\text{leaves}(T)$ -realization of T , a contradiction.

It remains to show that T is S -realizable for every proper subset S of $\text{leaves}(T)$, but this is easy. If we remove a leaf outside S , the remaining tree is S -realizable (from the minimality of T), and this realization can be extended to an S -realization of T as P is a $(1, 2)$ -system. ■

For the remainder of the proof we fix a P -tree T with the properties stated in the previous lemma.

Lemma 6.4 T contains a path of length at least $2 \cdot 8^{|A|}$.

Proof It can be easily computed that a tree, which has all vertices of degree at most 3 and which does not contain any path with more than k vertices, has size at most 2^k (this is a crude estimate, one computes that the most accurate estimate is $3 \cdot 2^{k/2} - 2$ for even k and $2^{(k+3)/2} - 2$ for odd $k > 1$).

Since T has more than $4^{8^{|A|}}$ vertices by our assumption (smaller P -trees are even realizable in $P_{[g]}$), the claim follows. ■

We fix a subpath v_1, v_2, \dots, v_m of T , where $m \geq 2 \cdot 8^{|A|}$. We define subsets S_i of $\text{leaves}(T)$, $i = 1, 2, \dots, m$, as follows. A leaf of T is in S_i if and only if the shortest path from this leaf to v_i contains neither v_{i-1} nor v_{i+1} . (For v_1 only the vertex v_2 is considered. If v_1 is a leaf, then $S_1 = \{v_1\}$. Similarly for v_m .) In other words, we straighten the line v_1, \dots, v_m and shake the tree. Then S_i is the set of leaves below v_i .

The next lemma will enable us to find the sought after configuration.

Lemma 6.5 *There exist natural numbers k, l such that*

- $1 \leq k, l \leq m, k \leq l + 2$;
- $T_{S_k}[v_k] = T_{S_l}[v_l]$;
- $T_{S_1 \cup S_2 \cup \dots \cup S_k}[v_k] = T_{S_1 \cup S_2 \cup \dots \cup S_l}[v_l] \neq \emptyset$;
- $T_{S_k \cup S_{k+1} \cup \dots \cup S_m}[v_k] = T_{S_l \cup S_{l+1} \cup \dots \cup S_m}[v_l] \neq \emptyset$.

Proof There is at least $m/2 - 1 \geq 8^{|A|} - 1$ even numbers less than m . For each such number i we consider the triple

$$(T_{S_i}[v_i], T_{S_1 \cup \dots \cup S_i}[v_i], T_{S_i \cup \dots \cup S_m}[v_i])$$

of subsets of A (note that these subsets are nonempty by the third item of Lemma 6.3). There are less than $(2^{|A|} - 1)^3 < 8^{|A|} - 2$ possible triples, therefore, by the pigeonhole principle, there exist distinct k, l with the same associated triples and the lemma follows.

Again, the estimates we used are very rough. For instance, the second and third sets in the triple are disjoint subsets of the first subset. This significantly reduces the number of possibilities, etc. ■

Let

$$Q_1 = S_1 \cup S_2 \cup \dots \cup S_k, \quad Q_2 = S_k \cup \dots \cup S_l, \quad Q_3 = S_l \cup \dots \cup S_m.$$

Now we define

$$\begin{aligned}
B &= T_{S_k}[v_k] = T_{S_l}[v_l], \\
L &= T_{Q_2 \cup Q_3}[v_k] = T_{Q_3}[v_l], \\
U &= T_{Q_1}[v_k] = T_{Q_1 \cup Q_2}[v_l], \\
E &= \{(r(v_k), r(v_l)) : r \text{ is a } Q_2\text{-realization of } T\}, \\
F &= \{(r(v_k), r(v_l)) : r \text{ is a } (S_k \cup S_l)\text{-realization of } T\}, \\
(a, b) &= (r(v_k), r(v_l)) \text{ for a chosen } (Q_1 \cup Q_3)\text{-realization } r \text{ of } T.
\end{aligned}$$

Since $k \leq l - 2$ and $S_{k+1} \neq \emptyset$ (by the first item of Lemma 6.3), $Q_1 \cup Q_3$ is a proper subset of $\text{leaves}(T)$; therefore, T has a $(Q_1 \cup Q_3)$ -realization by the third item of Lemma 6.3, and the definition of a and b makes sense. This choice satisfies all the assumptions of Lemma 6.1:

- B is a subuniverse of \mathbf{A} . It follows directly from the definitions (see the last paragraph of the proof of Proposition 5.3). Let \mathbf{B} be the subalgebra of \mathbf{A} with universe B .
- $E, F \leq \mathbf{B}^2$, E is a Jónsson ideal of F . This is also straightforward. That E is a Jónsson ideal of F follows from the assumption that J_x is a Jónsson ideal of \mathbf{R}_x for every $x \in Q_2$.
- U and L are disjoint. Suppose $c \in U \cap L$. Since $U = T_{Q_1}[v_k]$, there exists a Q_1 -realization r_1 of T such that $r_1(v_k) = c$. Similarly, since $L = T_{Q_2 \cup Q_3}[v_k]$, there exists a $(Q_2 \cup Q_3)$ -realization r_2 of T such that $r_2(v_k) = c$. The realizations r_1 and r_2 can be joined in the following way. We put $r(v) = r_1(v)$ for vertices v whose shortest path to v_k does not contain v_{k+1} , and $r(v) = r_2(v)$ for the other vertices. Now r is a $(Q_1 \cup Q_2 \cup Q_3)$ -realization of T . But $Q_1 \cup Q_2 \cup Q_3$ is the set of all leaves of T , a contradiction (see the second item of Lemma 6.3).
- $a \in U, b \in L, (a, b) \in F$. The element a is defined as $r(v_k)$ for a $(Q_1 \cup Q_3)$ -realization r of T . Since $Q_1 \subseteq Q_1 \cup Q_3$, we have $T_{Q_1 \cup Q_3}[v_k] \subseteq T_{Q_1}[v_k] = U$; therefore, $a = r(v_k) \in U$. Similarly, $b \in L$ follows from $b = r(v_l)$, $Q_3 \subseteq Q_1 \cup Q_3$ and $L = T_{Q_3}[v_l]$, and $(a, b) \in F$ follows from $S_k \cup S_l \subseteq Q_1 \cup Q_3$.
- $E \cap U^2$ has no sources, $E \cap L^2$ has no sinks. Let c be an arbitrary element of U . Since $U = T_{Q_1 \cup Q_2}[v_l]$, there exists a $(Q_1 \cup Q_2)$ -realization r of T such that $r(v_l) = c$. But r is also a Q_2 -realization of T , hence $(r(v_k), r(v_l)) \in E$. The element $d = r(v_k)$ lies in $T_{Q_1 \cup Q_2}[v_k] \subseteq T_{Q_1}[v_k] = U$. We can analogically show that $E \cap L^2$ has no sinks: any $c \in L$ is equal to $r(v_k)$ for a $(Q_2 \cup Q_3)$ -realization r of T , and $r(v_l) \in T_{Q_3}[v_l] = L$.
- There do not exist $c \in U$ and $d \in B \setminus U$ such that $(c, d) \in E$. If $c \in U = T_{Q_1}[v_k]$ and $(c, d) \in E$, then there exists a Q_1 -realization r_1 of T and a Q_2 -realization r_2 of T such that $r_1(v_k) = c = r_2(v_k)$ and $r_2(v_l) = d$. When we join r_1 and r_2 in the same way as in the proof that U and L are disjoint, we get a $(Q_1 \cup Q_2)$ -realization r of T such that $r(v_l) = d$. But $U = T_{Q_1 \cup Q_2}[v_l]$, thus $d \in U$.

The last property contradicts Lemma 6.1, and this concludes the proof of Theorem 5.6. \blacksquare

7 Conclusion

7.1 Decidability of Near Unanimity for Relational Structures

As a corollary of the main theorem we obtain an affirmative answer to the near unanimity problem for relations.

Corollary 7.1 *It is decidable whether a finite relational structure with finitely many relations has a near unanimity polymorphism.*

Proof It is enough to decide whether the given relational structure has a Jónsson chain of polymorphisms. This can be decided as follows. We first compute the set P of all ternary idempotent polymorphisms satisfying $p(a, b, a) = a$ and then compute the graph whose vertices are idempotent binary operations having f and g the vertices of an edge if and only if there exist $p_1, p_2 \in P$ such that for all a, b ,

$$f(a, b) = p_1(a, a, b), p_1(a, b, b) = p_2(a, b, b) \quad \text{and} \quad g(a, b) = p_2(a, a, b).$$

A Jónsson chain exists if and only if π_1 is connected to π_2 , where π_i is the idempotent binary operation that is the projection operation on the i -th coordinate. ■

It was shown in [22] that the corresponding decision problem for algebras (that is, does a given finite algebra with finitely many operations have a near unanimity term operation?) is decidable. This was a surprising development after undecidability results about closely related questions [21].

The naive algorithm described in the proof of Corollary 7.1 runs in exponential time.

Open Problem 7.2 *Determine the computational complexity of deciding whether a finite relational structure with finitely many relations has a near unanimity polymorphism.*

There exist polynomial time algorithms for finite posets [17] and for finite reflexive undirected graphs [18].

The complexity of the same problem for algebras is also unknown. There exists a polynomial time algorithm for deciding whether a finite idempotent algebra (with finitely many operations) is a CD algebra, and the same problem without assuming idempotency is exponential time complete [12].

7.2 Arities

Our proof gives some upper bound on the minimal arity of a near unanimity polymorphism; namely, a binary relational structure \mathbb{A} either has a near unanimity polymorphism of arity $4^{8^{|\mathbb{A}|}} + 1$ or has none. From the reduction presented in Section 3 it follows that for a relational structure whose relations have maximum arity k an upper bound is $4^{8^{|\mathbb{A}|^k}} + 1$. We have used quite rough estimates in a couple of places; however, this proof most likely cannot provide a better upper bound than doubly exponential.

For finite algebras with finitely many operations the upper bound also exists, but is tremendously large and is not even computed in [22].

Therefore we have the following open problem.

Open Problem 7.3 Give a better upper bound for the minimal arity of a near unanimity polymorphism (resp. term operation) for relational structures with finitely many relations (resp. finite algebras).

7.3 Valeriote’s Conjecture

The most important open problem related to this work is the Valeriote conjecture (also known as the Edinburgh conjecture [7]).

Conjecture 7.4 Every finite, finitely related algebra in a congruence modular variety has few subpowers.

Congruence modularity is a widely studied generalization of congruence distributivity. An algebra \mathbf{A} has *few subpowers* if the logarithm of the number of subalgebras of \mathbf{A}^n is bounded by a polynomial in n . This property was defined and its importance in the CSP demonstrated in [5, 14]. Examples of algebras with few subpowers include algebras with a Maltsev operation (e.g., groups, rings, modules) and algebras with a near unanimity operation. It is known [5, 20] that every finite CD algebra with few subpowers has a near unanimity term. Therefore, a positive solution to the Valeriote conjecture would imply the main result of this paper. It would also have deep consequences in the complexity of constraints.

A converse to the conjecture generalizing the Baker–Pixley result [2] was proved recently. E. Aichinger, P. Mayr, and R. McKenzie [1] have shown that every finite algebra with few subpowers is finitely related.

References

- [1] E. Aichinger, P. Mayr, and R. McKenzie, *On the number of finite algebraic structures*. arxiv:1103.2265.
- [2] K. A. Baker and A. F. Pixley, *Polynomial interpolation and the Chinese remainder theorem for algebraic systems*. Math. Z. **143**(1975), no. 2, 165–174. <http://dx.doi.org/10.1007/BF01187059>
- [3] L. Barto and M. Kozik, *Congruence distributivity implies bounded width*. SIAM J. Comput. **39**(2009/10), no. 4, 1531–1542. <http://dx.doi.org/10.1137/080743238>
- [4] ———, *Constraint satisfaction problems of bounded width*. In: 2009 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2009), IEEE Computer Soc., Los Alamitos, CA, 2009, pp. 595–603.
- [5] J. Berman, P. Idziak, P. Marković, R. McKenzie, M. Valeriote, and R. Willard, *Varieties with few subalgebras of powers*. Trans. Amer. Math. Soc. **362**(2010), no. 3, 1445–1473. <http://dx.doi.org/10.1090/S0002-9947-09-04874-0>
- [6] V. G. Bodnarčuk, L. A. Kalužnin, V. N. Kotov, and B. A. Romov, *Galois theory for Post algebras. I, II*. (Russian), Kibernetika (Kiev) **1969**, no. 3, 1–10; *ibid.* **1969**, no. 5, 1–9.
- [7] S. Bova, H. Chen, and M. Valeriote, *Generic expression hardness results for primitive positive formula comparison*. 38th International Colloquium on Automata, Languages and Programming (ICALP), Zürich, Switzerland, 2011.
- [8] A. Bulatov, P. Jeavons, and A. Krokhin, *Classifying the complexity of constraints using finite algebras*. SIAM J. Comput. **34**(2005), no. 3, 720–742. <http://dx.doi.org/10.1137/S0097539700376676>
- [9] S. N. Burris and H. P. Sankappanavar, *A course in universal algebra*. Graduate Texts in Mathematics, 78, Springer-Verlag, New York-Berlin, 1981.
- [10] B. A. Davey, *Monotone clones and congruence modularity*. Order **6**(1990), no. 4, 389–400. <http://dx.doi.org/10.1007/BF00346133>
- [11] T. Feder and M. Y. Vardi, *The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory*. SIAM J. Comput. **28**(1999), no. 1, 57–104. <http://dx.doi.org/10.1137/S0097539794266766>

- [12] R. Freese and M. A. Valeriote, *On the complexity of some Maltsev conditions*. Internat. J. Algebra Comput. **19**(2009), no. 1, 41–77. <http://dx.doi.org/10.1142/S0218196709004956>
- [13] D. Geiger, *Closed systems of functions and predicates*. Pacific J. Math. **27**(1968), 95–100.
- [14] P. Idziak, P. Marković, R. McKenzie, M. Valeriote, and R. Willard, *Tractability and learnability arising from algebras with few subpowers*. In: Proceedings of the Twenty-Second Annual IEEE Symposium on Logic in Computer Science (LICS 2007), IEEE Computer Society Press, 2007, pp. 213–222.
- [15] P. Jeavons, D. Cohen, and M. Gyssens, *Closure properties of constraints*. J. ACM **44**(1997), no. 4, 527–548. <http://dx.doi.org/10.1145/263867.263489>
- [16] B. Jónsson, *Algebras whose congruence lattices are distributive*. Math. Scand. **21**(1968), 110–121.
- [17] G. Kun and C. Szabó, *Order varieties and monotone retractions of finite posets*. Order **18**(2001), no. 1, 79–88. <http://dx.doi.org/10.1023/A:1010681409599>
- [18] B. Larose, C. Loten, and L. Zádori, *A polynomial-time algorithm for near-unanimity graphs*. J. Algorithms **55**(2005), no. 2, 177–191. <http://dx.doi.org/10.1016/j.jalgor.2004.04.011>
- [19] B. Larose and L. Zádori, *Algebraic properties and dismantlability of finite posets*. Discrete Math. **163**(1997), no. 1–3, 89–99. [http://dx.doi.org/10.1016/0012-365X\(95\)00312-K](http://dx.doi.org/10.1016/0012-365X(95)00312-K)
- [20] P. Marković and R. McKenzie, *Few subpowers, congruence distributivity and near-unanimity terms*. Algebra Universalis **58**(2008), no. 2, 119–128. <http://dx.doi.org/10.1007/s00012-008-2049-1>
- [21] M. Maróti, *On the (un)decidability of a near-unanimity term*. Algebra Universalis **57**(2007), no. 2, 215–237. <http://dx.doi.org/10.1007/s00012-007-2037-x>
- [22] M. Maróti, *The existence of a near-unanimity term in a finite algebra is decidable*. J. Symbolic Logic **74**(2009), no. 3, 1001–1014. <http://dx.doi.org/10.2178/jsl/1245158096>
- [23] ———, *Monotone clones, residual smallness and congruence distributivity*. Bull. Austral. Math. Soc. **41**(1990), no. 2, 283–300. <http://dx.doi.org/10.1017/S0004972700018104>
- [24] L. Zádori, *Monotone Jónsson operations and near unanimity functions*. Algebra Universalis **33**(1995), 216–236. <http://dx.doi.org/10.1007/BF01190934>

Department of Mathematics and Statistics, McMaster University, Hamilton, ON

and

Department of Algebra, Charles University, Prague, Czech Republic

e-mail: libor.barto@gmail.com