FACULTY OF MATHEMATICS AND PHYSICS
Charles University

DOCTORAL THESIS

Miroslav Olšák

Generalizing CSP-related results to infinite algebras

Department of Algebra

Supervisor of the doctoral thesis:	Libor Barto
Study programme:	Mathematics
Study branch:	Algebra, number theory and mathematical logic

Prague 2019

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

The author wishes to express his gratitude to his supervisor, Libor Barto, especially for his help with putting the results into a broader context, and also to the Czech and Polish taxpayers for providing the monetary support in the background.

Title: Generalizing CSP-related results
to infinite algebras

Author: Miroslav Olšák

Department: Department of Algebra

Supervisor: Libor Barto, Department of Algebra

Abstract: The recent research on constraint satisfaction problems (CSPs) on fixed finite templates provided useful tools for computational complexity and universal algebra. However, the research mainly focused on finite relational structures, and consequently, finite algebras. We pursue a generalization of these tools and results into the domain of infinite algebras. In particular, we show that despite the fact that the Maltsev condition $s(r,a,r,e) = s(a,r,e,a)$ does not characterize Taylor algebras (i.e., algebras that satisfy a nontrivial idempotent Maltsev condition) in general, as it does in the finite case, there is another strong Maltsev condition characterizing Taylor algebras, and $s(r,a,r,e) = s(a,r,e,a)$ characterizes another interesting broad class of algebras. We also provide a (weak) Maltsev condition for SD($\wedge$) algebras (i.e., algebras that satisfy an idempotent Maltsev condition not satisfiable in a module). Beyond Maltsev conditions, we study loop lemmata and, in particular, reprove a well known finite loop lemma by two different general (infinite) approaches.

Keywords: Maltsev conditions Constraint satisfaction problems Loop lemmata

Contents

Introduction **2**
1 Background 2
1.1 Constraint satisfaction problem over a fixed finite template 2
1.2 Algebraic approach to CSP 4
1.3 Universal algebra and Maltsev conditions 5
1.4 Loop lemmata 6
2 Results 7
2.1 Maltsev conditions 7
2.2 Loops in digraphs 9
3 Methods and highlights 10
4 Conclusion 11

Bibliography **13**

List of publications **15**

Introduction

The fruitful connection between the constraint satisfaction problem (CSP) and universal algebra, has brought key discoveries to both domains. For example, loop lemmata that originate in the CSP exposed unexpected theorems about finite algebras that can be stated in a purely algebraic form. While the finiteness assumption is often in the core of these proofs, this thesis provides alternative proofs and generalizations that do not involve finiteness. The results are motivated not only by a pure universal algebraic interest but also by potential applications to CSPs over infinite templates.

The thesis consists of the following five articles.

[A] The Weakest Nontrivial Idempotent Equations, Bulletin of the London Mathematical Society, 49(6):1028–1047, 2017,

[B] Maltsev Conditions for General Congruence Meet-Semidistributive Algebras (submitted to Journal of Symbolic Logic)

[C] Loop Conditions (submitted to Algebra Universalis, after second review)

[D] Loop Conditions for Strongly Connected Digraphs (submitted to International Journal of Algebra and Computation, after first review)

[E] Local Loop Lemma, submitted to Algebra Universalis

We begin by summarizing the relevant results from the CSP over finite domains, then we explain the contributions in the attached articles, and, finally, we discuss possible further research directions.

1 Background

1.1 Constraint satisfaction problem over a fixed finite template

A relational structure $\mathbb{R}$ on a domain A is a tuple $(A, R_1, R_2, \ldots)$, where every R_i is a relation on A, that is, a subset of some finite power of A. A constraint over $\mathbb{R}$ is a formal expression of the form "$(x_1, \ldots, x_n) \in R$", where $x_1, \ldots, x_n$ are variables (with possible repetition), and R is an n-ary relation from $\mathbb{R}$. A constraint satisfaction problem over a relational structure (template) $\mathbb{R}$, denoted $\mathrm{CSP}(\mathbb{R})$ is a computational problem, where an algorithm is given a sequence of constraints over $\mathbb{R}$ on the input and it answers whether it is possible to assign values to the variables so that all the constraints are satisfied.

Constraint satisfaction problems over fixed finite templates can model many standard computational problems. We give three examples.

Example 1.1.

- *For $i, j \in \{0, 1\}$, let $S_{i,j}$ denote $\{0, 1\}^2 \setminus (i, j)$. Consider the relational structure $\mathbb{R} = (\{0, 1\}, S_{0,0}, S_{0,1}, S_{1,0}, S_{1,1})$. Then* $\mathrm{CSP}(\mathbb{R})$ *is essentially 2-SAT, a well-known computational problem solvable in polynomial time.*

- *For $i, j, k \in \{0,1\}$, let $S_{i,j,k}$ denote $\{0,1\}^3 \setminus (i,j,k)$. Consider the relational structure $\mathbb{R} = (\{0,1\}, S_{0,0,0}, \dots, S_{1,1,1})$. Then* CSP($\mathbb{R}$) *is essentially 3-SAT, a well-known NP-complete problem.*

- *For a prime number p consider a domain $\mathbb{Z}_p = \{0, \dots, p-1\}$. Let $\mathcal{S}$ be the system of all affine subspaces of $\mathbb{Z}_p^3$. Then* CSP($(\mathbb{Z}_p, \mathcal{S})$) *is essentially 3-LIN(p), the problem of solving a system of linear equations in $\mathbb{Z}_p$ (where each equation involves 3 variables) and can be solved by Gaussian elimination in polynomial time.*

Among the major early results in the area were Schaefer's theorem, see Schaefer [1978], which states that every CSP over a two element domain is either solvable in polynomial time or NP-complete, and the Hell-Nešetřil theorem:

Theorem 1.1 (Hell and Nešetřil [1990])**.** *CSP over an undirected loopless graph (considered as a relational structure with a single binary symmetric relation) is solvable in polynomial time if the graph is bipartite, or NP-complete if the graph is not bipartite.*

These results inspired a well-known "dichotomy conjecture" that every CSP over a fixed finite template is either NP-complete or solvable in polynomial time, that was stated in Feder and Vardi [1998] and recently independently proved in Zhuk [2017] and Bulatov [2017].

However, the NP-completeness vs. solvability in polynomial time was not the only important division line in the complexity of CSPs. Some of the CSPs solvable in polynomial time can be solved by a simple class of algorithms, which repeatedly perform such strengthenings of the constraints that are "locally" forced by other constraints. Such an algorithm works for many templates, including 2-SAT. If such an algorithm works, the CSP is said to have *bounded width*. For a detailed definition of bounded width and the appropriate algorithm, we refer the reader to Barto et al. [2017]. It turned out that the only obstacles for a CSP to have bounded width are in some sense the problems 3-LIN(p). More precisely, the following are equivalent for a relational template $\mathbb{D}$.

- CSP($\mathbb{D}$) has bounded width
- $\mathbb{D}$ does not pp-construct the language of 3-LIN(p) for any prime number p.

The term pp-construction (primitive positive construction) stands for a relational construction which generalizes pp-definitions and pp-interpretations.

$$\text{pp-definitions} \subset \text{pp-interpretations} \subset \text{pp-constructions}$$

The significance of pp-constructions in the CSP comes from the fact that CSP($\mathbb{R}_2$) can be reduced to CSP($\mathbb{R}_1$) whenever $\mathbb{R}_1$ pp-constructs $\mathbb{R}_2$. For a definition of these constructions, we refer the reader to Barto et al. [2017].

When studying CSPs, it is convenient to make certain "WLOG" assumptions. Most importantly, we usually focus on idempotent templates. A relational structure $\mathbb{R}$ on A is called *idempotent* if it contains all the constants; more precisely, for every $a \in A$, there is a unary relation $\{a\}$ in $\mathbb{R}$. The following theorem explains why it suffices to study the idempotent templates.

Theorem 1.2 (See Barto et al. [2017])**.** *For every relational structure $\mathbb{R}$ there is an idempotent relational structure $\mathbb{R}_0$ such that $\mathbb{R}$ and $\mathbb{R}_0$ pp-construct each other.*

1.2 Algebraic approach to CSP

Let $\mathbb{R}$ be a relational structure on A. We say that an operation $f\colon A^n \to A$ is compatible with $\mathbb{R}$ (or is a polymorphism of $\mathbb{R}$), if for any k-ary relation R in $\mathbb{R}$ and any k-tuples $\mathbf{r}_1, \dots, \mathbf{r}_n \in R$, also $f(\mathbf{r}_1, \dots, \mathbf{r}_n) \in R$, where we interpret f as an n-ary operation on k-tuples acting element-wise. The algebraic approach is based on the following theorem.

Theorem 1.3 (Jeavons et al. [1997], see also Barto et al. [2017])**.** *The complexity of* CSP($\mathbb{R}$) *is fully determined by the set of polymorphisms of* $\mathbb{R}$*.*

The set of all polymorphisms contains all projections and is closed under composition. In universal algebra, such sets of operations are called *clones*. Using the algebraic terminology, Schaefer's theorem can be phrased as follows.

Theorem 1.4 (Schaefer [1978], algebraically formulated in Chen [2009])**.** *Consider a relational structure* $\mathbb{R}$ *on* $\{0, 1\}$*. Then* CSP($\mathbb{R}$) *is solvable in polynomial time if one of the following cases happens.*

(1) $\mathbb{R}$ *is compatible with a constant operation.*

(2) $\mathbb{R}$ *is compatible with the binary minimum, or the binary maximum operation.*

(3) $\mathbb{R}$ *is compatible with the majority operation, that is, the ternary operation* m *given by* $m(x, x, y) = m(x, y, x) = m(y, x, x) = x$*.*

(4) $\mathbb{R}$ *is compatible with the minority operation, that is, the ternary operation* m *given by* $m(x, y, y) = m(y, x, y) = m(y, y, x) = x$*.*

If none of the cases applies, every polymorphism of $\mathbb{R}$ *depends on exactly one variable and* CSP($\mathbb{R}$) *is NP-complete. Moreover,* CSP($\mathbb{R}$) *has bounded width if and only if* $\mathbb{R}$ *satisfies one of the conditions (1), (2), (3).*

Note that if $\mathbb{R}$ is idempotent, then all the polymorphisms of $\mathbb{R}$ are idempotent, that is, they satisfy $f(x, \dots, x) = x$. In Schaefer's theorem, the assumption of idempotency disables the trivial case (1).

It turned out that for determining the complexity of CSP of a finite idempotent template, the actual set of polymorphism is not necessary, and the complexity is fully determined by the equations that the polymorphisms satisfy. For example, if there is a polymorphism that satisfies $t(x, x, y) = t(x, y, x) = t(y, x, x) = x$, then the CSP is solvable in polynomial time and has bounded width, no matter what is the size of the domain, or how the polymorphism t behaves for different choices of arguments.

More generally, consider a set of operational symbols Δ together with assigned arities. Let $\mathcal{S}$ be a set of equations using symbols in Δ and symbols of variables. We say that polymorphisms of a relational structure $\mathbb{R}$ satisfy $\mathcal{S}$ if it is possible to assign polymorphisms of $\mathbb{R}$ to the symbols of Δ so that the equations are true for every choice of variables. The complexity of CSP($\mathbb{R}$) is then fully determined by the set of all systems of equations satisfied by polymorphisms of $\mathbb{R}$. A system of equations is said to be *trivial* if it is satisfiable by projections to certain coordinates. The CSP dichotomy theorem can be stated algebraically as follows.

Theorem 1.5 (Zhuk [2017], Bulatov [2017])**.** *Let* $\mathbb{R}$ *be an idempotent template. If polymorphisms of* $\mathbb{R}$ *satisfy a non-trivial system of equations, then* CSP($\mathbb{R}$) *is solvable in polynomial time. Otherwise,* CSP($\mathbb{R}$) *is NP-complete.*

We remark that the NP-completeness part of Theorem 1.5 follows immediately from the theory. The hard part of the theorem was to provide a polynomial time algorithm that is capable of solving CSP($\mathbb{R}$) if polymorphisms of $\mathbb{R}$ satisfy a non-trivial system of equations.

These facts directed the CSP research to the domain of universal algebra, where the study of systems of equations, known under the name Maltsev conditions, had a long history.

1.3 Universal algebra and Maltsev conditions

The central objects in universal algebra are algebras. Algebra $\mathbf{A}$ is a set A (called universe) equipped with so called "basic" operations in a given signature. Algebras inherit some terminology from operations. In particular, an algebra is idempotent if all the basic operations are idempotent, or compatible with a relational structure $\mathbb{R} = (A, \ldots)$ if all the basic operations of $\mathbf{A}$ are compatible with $\mathbb{R}$. A *term* in an algebra is an operation given by an expression using variables and the basic operations of the algebra. The set of all terms in an algebra is another basic example of a clone.

Universal algebra used equational conditions (by which we mean conditions postulating the existence of terms satisfying a fixed set of equations) mostly for classifying the behavior of congruences in algebras. The basic example of such a classification result is the following one (see Bergman [2012]). If an algebra $\mathbf{A}$ has a ternary term that satisfies $m(x, y, y) = m(y, y, x) = x$, then its congruences permute, that is, for any congruences α, β of $\mathbf{A}$, $\alpha \circ \beta = \beta \circ \alpha$. The converse is also true but not in the straightforward sense. Let $\mathbf{F}$ be the $\mathbf{A}$-free algebra generated by three elements. If the congruences of $\mathbf{F}$ permute, then there is a term m such that $m(x, y, y) = m(y, y, x) = x$ in $\mathbf{A}$ and $\mathbf{F}$.

For this reason, universal algebra tests equational conditions in varieties rather than algebras. A *variety* is a class of algebras of a fixed signature closed under products, subalgebras and homomorphic images. For any algebra $\mathbf{A}$ in a variety $\mathcal{V}$, the $\mathbf{A}$-free algebra is in $\mathcal{V}$ as well. Varieties with finite free algebras on finitely many generators are called *locally finite*; in other words, a variety is locally finite if every finitely generated algebra in $\mathcal{V}$ is finite.

Equational conditions in universal algebra appear in the form of *strong Maltsev conditions* that are given by finite sets of equations, and *Maltsev conditions* that are given by infinite countable disjunctions of strong Maltsev conditions. An early result from universal algebra was a description of the weakest non-trivial Maltsev condition for idempotent algebras in terms of the existence of a Taylor term. (An algebra that has a Taylor term is called a *Taylor algebra.*) However, later research on CSP and universal algebra revealed many other equivalent Maltsev conditions for finite algebras.

Theorem 1.6 (Barto et al. [2017], Bulatov [2017], Zhuk [2017], A. Kearnes and Kiss [2013], Siggers [2010])**.** *The following are equivalent for a finite idempotent algebra* $\mathbf{A}$*.*

(1) CSP($\mathbb{R}$) *is solvable in polynomial time for every relational structure* $\mathbb{R}$ *compatible with* **A**,

(2) **A** *has a term* s *satisfying* $s(r, a, r, e) = s(a, r, e, a)$ *(called 4-ary Siggers term),*

(3) **A** *has a term* s *satisfying* $s(x, x, y, y, z, z) = s(y, z, z, x, x, y)$ *(called 6-ary Siggers term),*

(4) **A** *has a weak near unanimity term* w *of some arity, that is, a term satisfying*

$$w(y, x, \ldots, x) = w(x, y, x, \ldots, x) = \cdots = w(x, x, \ldots, y),$$

(5) **A** *has a cyclic term* c *of some arity* n*, that is, a term satisfying*

$$c(x_1, \ldots, x_n) = c(x_2, x_3, \ldots, x_n, x_1),$$

(6) **A** *has a Taylor term,*

(7) **A** *satisfies a non-trivial Maltsev condition.*

There are also numerous characterizations for having a bounded width.

Theorem 1.7 (see Barto et al. [2017], A. Kearnes and Kiss [2013] and Kozik et al. [2015])**.** *The following are equivalent for a finite idempotent algebra* **A**.

(1) CSP($\mathbb{R}$) *has bounded width for every relational structure* $\mathbb{R}$ *compatible with* **A**,

(2) **A** *has weak near unanimity terms* w_3, w_4 *of arities 3 and 4 respectively such that* $w_3(y, x, x) = w_4(y, x, x, x)$*,*

(3) the congruence lattice of every algebra in the variety generated by **A** *is meet-semidistributive,*

(4) **A** *satisfies a Maltsev condition that is not satisfiable by affine combinations over a non-trivial field.*

Note that items 6 and 7 in Theorem 1.6 are equivalent even in infinite idempotent algebras, as well as items 3 and 4 in Theorem 1.7

1.4 Loop lemmata

One of the contributions of the CSP research to universal algebra is bringing attention to general binary relations compatible with the algebra, and not primarily equivalences. The Hell and Nešetřil result, Theorem 1.1, was later reproved in Bulatov [2005] by showing that whenever a Taylor algebra is compatible with an undirected non-bipartite graph, the graph is forced to have a loop. Similar statements, that certain relational assumptions (e.g. non-bipartite undirected) and algebraic assumptions (e.g. compatible with a Taylor algebra) forces a digraph (a relational structure with a single binary relation) to have a loop, are referred to as loop lemmata. A theorem often called "the loop lemma" is the following one.

Theorem 1.8 (Barto et al. [2008/09], Barto and Kozik [2012]). *Let $\mathbb{G}$ be a finite digraph compatible with a Taylor algebra $\mathbb{A}$. Suppose that $\mathbb{G}$ is weakly connected, has no sources and no sinks, and that $\mathbb{G}$ cannot be homomorphically mapped to a non-trivial directed cycle. Then $\mathbb{G}$ has a loop.*

Loop lemmata provided important tools for proving various results related to CSPs and Maltsev conditions. Corollaries of Theorem 1.8 include, for example, the positive answer to the Bang-Jensen and Hell conjecture asked in Bang-Jensen and Hell [1990] or the fact that any finite Taylor algebra has a 4-ary Siggers term, see Kearnes et al. [2014].

2 Results

Our results in the attached articles can be primarily categorized into two topics, generalizing the results about Maltsev conditions and generalizing loop lemmata for infinite algebras. The articles discussing loop conditions belong equally to both categories.

2.1 Maltsev conditions

The earliest, and most widely known, contribution is the fact that Taylor algebras (varieties) are characterized by a strong Maltsev condition, proved in [A]:

Theorem 2.1. *Every idempotent algebra that satisfies a non-trivial equational condition has a term t, so called weak 3-cube*[1] *term, satisfying*

$$\begin{aligned} &t(x,y,y,y,x,x)\\ \approx\ &t(y,x,y,x,y,x)\\ \approx\ &t(y,y,x,x,x,y). \end{aligned}$$

Note that the equations in Theorem 2.1 are non-trivial too since there is no constant column of variables.

Attempts to determine whether SD($\wedge$) algebras can be characterized by a strong Maltsev condition was not successful so far. However, we have found at least a neat (but weak) Maltsev condition characterizing these algebras.

For positive integers n, m, we define the $(n+m)$-terms as the triple of idempotent terms (f, g_1, g_2) where g_1 is n-ary, g_2 is m-ary and f is $(n+m)$-ary, and they satisfy the following identities for any $1 \le i \le n$ and $1 \le j \le m$:

$$\begin{aligned} f(x,x,\dots,x,\underset{i}{y},x,\dots,x) &= g_1(x,x,\dots,x,\underset{i}{y},x,\dots,x),\\ f(x,x,\dots,x,\underset{n+j}{y},x,\dots,x) &= g_2(x,x,\dots,x,\underset{j}{y},x,\dots,x). \end{aligned}$$

Recall that the SD($\wedge$) property is equivalent with satisfying an equational condition that cannot be satisfied by affine combinations over a non-trivial field, and

[1] The fact that the term is sometimes associated with the author's name suggests that the notation "weak 3-cube" was not chosen optimally.

note that $(n+m)$-terms provide an example of such a condition. To see this assume that the $(n+m)$-terms are represented by affine combinations:

$$f(x_1,\dots,x_{n+m}) = \sum_{i=1}^{n+m} \alpha_i x_i,$$
$$g_1(x_1,\dots,x_n) = \sum_{i=1}^{n} \beta_i x_i,$$
$$g_2(x_1,\dots,x_m) = \sum_{i=1}^{m} \gamma_i x_i,$$

The equations force $\alpha_i = \beta_i$ and $\alpha_{n+j} = \gamma_j$ for $1 \leq i \leq n$ and $1 \leq j \leq m$. If the linear combinations were affine, we get a contradiction:

$$1 = \sum_{i=1}^{n+m} \alpha_i = \sum_{i=1}^{n} \beta_i + \sum_{i=1}^{m} \gamma_i = 1 + 1 = 2.$$

In [B], we proved the following result.

Theorem 2.2. *Every congruence* SD($\wedge$) *algebra has* $(3+n)$*-terms for some* n*.*

The third Maltsev condition that was investigated is inspired by the characterization in Theorem 1.6. There are given two strong Maltsev conditions that characterize Taylor varieties among the locally finite varieties – the 6-ary Siggers term and the 4-ary Siggers term. However, it was shown in Kazda [2017] that there is an infinite Taylor algebra that has none of them. We studied the strong Maltsev conditions given by exactly one linear equation, calling them *loop conditions* in articles [C] and [D]. It turned out that the Siggers terms characterizes a quite natural class of general algebras. In particular, we proved the following.

Theorem 2.3. *Let $\mathcal{V}$ be a variety. The following are equivalent.*

(1) $\mathcal{V}$ satisfies a non-trivial loop condition.

(2) $\mathcal{V}$ has a 6-ary Siggers term,

(3) $\mathcal{V}$ has a 4-ary Siggers term,

(4) Let $\mathbb{G}$ be a digraph compatible with an algebra in $\mathcal{V}$. If $\mathbb{G}$ has a strongly connected component that cannot be homomorphically mapped to a directed cycle, then $\mathbb{G}$ has a loop.

Showing that a certain variety satisfies a non-trivial loop condition is often not very difficult (if it is true) since it is sufficient to prove that the operations can generate a loop from a clique which is as large as necessary. This simple idea was applied to show that the existence of a Siggers term is implied by an existence of a Hobby-McKenzie term, equivalently, by satisfying a non-trivial congruence identity.

2.2 Loops in digraphs

Our goal is to find useful infinite generalizations of Theorem 1.8 which was so fruitful in the finite world.

Even the weak 3-cube term (our weakest non-trivial idempotent equational condition) can be used for generating loops from a symmetric graph containing a triangle. Let t be a weak 3-cube term and let A, B, C form a triangle in a compatible graph. Then the nodes

$$\begin{gathered} t(A,B,B,B,A,A) \\ t(C,B,C,B,C,B) \\ t(A,A,C,C,C,A) \\ t(B,C,A,A,B,C) \end{gathered}$$

form a 4-clique in the graph. Using this approach, it is possible to generate a clique of arbitrary size from a triangle. Therefore, if there is a finite number of vertices, a loop must exist.

We introduce two other approaches and, in particular, reprove the following two theorems using each of them.

Theorem 2.4. *Let $\mathbb{G}$ be a finite strongly connected digraph that cannot be homomorphically mapped to a non-trivial directed cycle. If $\mathbb{G}$ is compatible with a Taylor term, then $\mathbb{G}$ has a loop.*

Theorem 2.5. *Let $\mathbb{G}$ be a strongly connected undirected graph that contains a cycle of odd length. If $\mathbb{G}$ is compatible with a near unanimity term, then $\mathbb{G}$ has a loop.*

The first theorem is a slight weakening of Theorem 1.8, the second one first appeared in [A] and can be considered as one of the most basic versions of an infinite loop lemma.

Our first approach was already introduced in Subsection 2.1 – the loop conditions. To prove the above theorems, it suffices to show that both locally finite Taylor varieties and varieties with a near unanimity term satisfy a non-trivial loop condition. This is done in papers [C] and [D]. Moreover, note that Hobby-McKenzie term is known to be weaker than a near unanimity term, so Theorem 2.5 follows from the result mentioned earlier. An advantage of loop conditions is that they do not require any form of idempotency, which makes them suitable for infinite domain CSPs, as explained in Section 4.

On the other hand, the idea of the second approach, described in [E], is to rely primarily on idempotency, but then to add just a little local algebraic assumption to avoid projections, or Kazda's counterexample. The local loop lemma can be stated as follows.

Theorem 2.6 (local loop lemma)**.** *Let $\mathbb{G}$ be a digraph compatible with an n-ary idempotent operation t. Suppose that $\mathbb{G}$ is strongly connected and that $\mathbb{G}$ contains directed cycles of all lengths greater than 1. Moreover suppose that for every $i \in \{1, \ldots, n\}$ it is possible to find vertices $\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n} \in \mathbb{G}$ such that there is an edge*

$$\alpha_{i,i} \to t(\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n}).$$

Then $\mathbb{G}$ has a loop.

Note that the assumptions of the local loop lemma are so weak that even its finite version does not follow from the thoroughly developed finite theory. Again, Theorems 2.4, 2.5 follow from Theorem 2.6. We also remark that the local loop lemma provides an even more convenient tool for proving that Taylor algebras (varieties) can be characterized by a strong Maltsev condition. Details can be found in [E].

3 Methods and highlights

In the papers [A, B, C, D], Maltsev conditions are obtained by finding an appropriate pattern (for example, a loop for loop conditions) in a suiably "freely generated" relation. This is a fairly standard approach in universal algebra, however, particular papers differ in the methods of obtaining such a required pattern.

The article [A] takes Barto and Kozik [2012] as an inspiration, and creates a loop in a graph by squashing a cycle of odd length into absorbing subuniverses. The rest of the derivations is done by composing operations.

Worth mentioning is a correspondence discovered in the article [B] between certain terms in the free algebra and elements of certain semirings. In particular, the problem of showing that every SD($\wedge$) algebras has $(n+n)$-terms for some n reduces to proving the following statement about semirings:

Proposition 3.1. *Let* $\mathbf{S}$ *be a semiring and let* $X_1, X_2, X_{n+1}, Y_1, Y_2, \dots, Y_n \subset \mathbf{S}$ *be finite sets of its elements such that* $X_1, \dots, X_{n+1}$ *are pairwise disjoint,* $Y_1, \dots, Y_n$ *are pairwise disjoint, and*

$$\bigcup_{i=1}^{n+1} X_i = \bigcup_{j=1}^{n} Y_j.$$

If $\sum_{x \in X_i} x = \sum_{y \in Y_j} y = 1$ *for every* $1 \le i \le n+1$ *and* $1 \le j \le n$*, then* $1 = 2$ *in* $\mathbf{S}$*.*

Loop conditions are tackled with pp-definitions and pp-interpretations. These tools are extensively used in CSP and were also used for the first proof of Theorem 1.1. In the study of Maltsev conditions, pp-definitions are nowadays sometimes used but the usage of another pp-interpretations is rather atypical. For handling the undirected loop conditions in [C], we use a hand-made gadget while the directed loop conditions in [D] require complex graphs of clique paths, or cycle walks (denoted CLQP(k, l, s) and CCLW(k, l, s)).

For the proof of the local loop lemma in [E], we, again, use a non-standard approach, although a very straightforward one – we explicitly define what to plug into a star power of the given term to obtain the required loop. Although the local approach is entirely different from the approach used for loop conditions, there is a flagrant similarity between the two articles. Let n, k be natural numbers, and $\mathbb{H}_{n,k}$ be a digraph defined as follows. The nodes of $\mathbb{H}_{n,k}$ are all sequences of length k of the letters $\{1, \dots, n\}$. There is an edge $a \to b$ in $\mathbb{H}_{n,k}$ if the suffix of a of length $k-1$ is equal to the prefix of b of length $k-1$. In both articles, we are proving a theorem that, very roughly, states "If k is large enough, there is a homomorphism from $\mathbb{H}_{n,k}$ to a fixed digraph $\mathbb{G}$". In loop conditions for directed graphs (paper [D]) we prove such a statement in Lemma 6, and in the local lemma (paper [E]) as the main part of the proof. Of course, the statement is, as formulated, false. First, $\mathbb{G}$ has to satisfy certain conditions, being strongly connected and not admitting a homomorphism to a non-trivial directed cycle, at least. However, just postulating some requirements on $\mathbb{G}$ is never enough since $\mathbb{H}$ contains loops on the constant tuples. The two articles differ in the way how this issue is handled. In the proof of the local loop lemma, we allow some exceptional places where the mapping $\mathbb{H} \to \mathbb{G}$ does not have to preserve the edges. Unfortunately, handling the exceptions brings significant technical difficulties. In the paper on loop conditions, we focus only on several significant subgraphs of $\mathbb{H}_{n,k}$, denoted $\mathrm{CLQP}(k, 0, n)$ and $\mathrm{CCLW}(k, 0, n)$ in the paper, which do not contain the loops.

4 Conclusion

We have found a strong Maltsev condition for general Taylor algebras, a simple Maltsev condition for SD($\wedge$) algebras, revealed a structural property that is characterized by the Siggers term, and generalized loop lemmata to infinite algebras in two ways. However, this line of research is far from being finished.

Each of the attached articles formulates several open questions. Since the article [A] is the oldest one, most of its open problems were already answered. The only intact one is Open problem 7.1: does every Taylor algebra have a 4-ary term e satisfying $e(y, y, x, x) = e(y, x, y, x) = e(x, x, x, y)$? Problem 7.4 was answered in the affirmative in [E], Problem 7.2 was answered negatively in Bodirsky et al. [2019]. Problem 7.3 asks, in the terminology of loop conditions, whether a near unanimity term implies the loop condition of any strongly connected digraph which does not admit a homomorphism to a directed cycle and have no sources and no sinks. This was confirmed for strongly connected digraphs in [D] while we do not have a proof even of 4-ary near unanimity term implying any loop condition which is not implied by the existence of a siggers term (Question 7.1 in [D]).

The question whether SD($\wedge$) algebras are characterized by a strong Maltsev condition remains open and is perhaps one of the main open question in universal algebra that concern strong Maltsev conditions. However, we at least provided a reasonable candidate for such a strong Maltsev condition – the $(3 + 3)$-terms.

Although the results concerning loops in graphs are quite self-contained, there is still plenty of room for generalizations to digraphs that are not strongly connected, or even to hypergraphs. One such a generalization to hypergraphs is studied in Gillibert et al. [2018].

Even though our work is not primarily focused on infinite domain CSP, it has the potential for applications. The algebraic approach to infinite domain CSP works particularly well for templates whose algebras of polymorphisms are oligomorphic, see Bodirsky [2008] for a survey. An issue with such algebras is that they are in a sense opposite to idempotent algebras – while idempotent algebras have only trivial unary operations, oligomorphic algebras have in a sense cofinite amount of them. Since the results presented in the thesis often require idempotency, most of our methods are not directly applicable. However, some variants of loop lemmata are already known to be significant for this kind of CSPs too. The paper Barto and Pinsker [2016] studying oligomorphic algebras introduces "pseudo-loop lemma" which guarantees an edge connecting two vertices in the same equivalence class modulo the unary operations instead of a loop. This translates into "pseudo-loop conditions" of the form

$$\alpha(t(\text{variables})) = \beta(t(\text{variables})).$$

The paper Gillibert et al. [2018] studies such conditions, inspired by the results in the thesis.

Thus the presented work offers many ways to continue. It is just the choice of the reader which way they pick if they decide to devote their time and talent to CSP-related algebraic tools.

Bibliography

Keith A. Kearnes and Emil Kiss. *The shape of congruence lattices*, volume 1046. Providence, Rhode Island : American Mathematical Society, 03 2013.

Jørgen Bang-Jensen and Pavol Hell. The effect of two cycles on the complexity of colourings by directed graphs. *Discrete Applied Mathematics*, 26(1):1 – 23, 1990. ISSN 0166-218X. doi: https://doi.org/10.1016/0166-218X(90)90017-7. URL `http://www.sciencedirect.com/science/article/pii/0166218X90900177`.

Libor Barto and Marcin Kozik. Absorbing subalgebras, cyclic terms, and the constraint satisfaction problem. *Logical Methods in Computer Science*, 8, 01 2012. doi: 10.2168/LMCS-8(1:7)2012.

Libor Barto and Michael Pinsker. The algebraic dichotomy conjecture for infinite domain constraint satisfaction problems. In *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '16, pages 615–622, New York, NY, USA, 2016. ACM. ISBN 978-1-4503-4391-6. doi: 10.1145/2933575.2934544. URL `http://doi.acm.org/10.1145/2933575.2934544`.

Libor Barto, Marcin Kozik, and Todd Niven. The CSP dichotomy holds for digraphs with no sources and no sinks (a positive answer to a conjecture of Bang-Jensen and Hell). *SIAM J. Comput.*, 38(5):1782–1802, 2008/09. ISSN 0097-5397. doi: 10.1137/070708093. URL `http://dx.doi.org/10.1137/070708093`.

Libor Barto, Andrei Krokhin, and Ross Willard. Polymorphisms, and How to Use Them. In Andrei Krokhin and Stanislav Zivny, editors, *The Constraint Satisfaction Problem: Complexity and Approximability*, volume 7 of *Dagstuhl Follow-Ups*, pages 1–44. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2017. ISBN 978-3-95977-003-3. doi: 10.4230/DFU.Vol7.15301.1. URL `http://drops.dagstuhl.de/opus/volltexte/2017/6959`.

Clifford Bergman. *Universal algebra*, volume 301 of *Pure and Applied Mathematics (Boca Raton)*. CRC Press, Boca Raton, FL, 2012. ISBN 978-1-4398-5129-6. Fundamentals and selected topics.

Manuel Bodirsky. Constraint satisfaction problems with infinite templates. In Heribert Vollmer, editor, *Complexity of Constraints (a collection of survey articles)*, volume 5250 of *Lecture Notes in Computer Science*, pages 196–228. Springer, 2008.

Manuel Bodirsky, Antoine Mottet, Miroslav Olšák, Jakub Opršal, Michael Pinsker, and Ross Willard. Topology is relevant (in the infinite-domain dichotomy conjecture for constraint satisfaction problems). Preprint arXiv:1901.04237, 2019.

A. A. Bulatov. A dichotomy theorem for nonuniform CSPs. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, volume 00, pages 319–330, Oct. 2017. doi: 10.1109/FOCS.2017.37. URL doi.ieeecomputersociety.org/10.1109/FOCS.2017.37.

Andrei A. Bulatov. H-coloring dichotomy revisited. *Theoret. Comput. Sci.*, 349 (1):31–39, 2005. ISSN 0304-3975. doi: 10.1016/j.tcs.2005.09.028. URL http://dx.doi.org/10.1016/j.tcs.2005.09.028.

Hubie Chen. A rendezvous of logic, complexity, and algebra. *ACM Comput. Surv.*, 42, 12 2009. doi: 10.1145/1592451.1592453.

Tomás Feder and Moshe Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28:57–104, 01 1998. doi: 10.1137/S0097539794266766.

Pierre Gillibert, Julius Jonušas, and Michael Pinsker. Pseudo-loop conditions. Preprint arXiv:1812.00396, 2018.

Pavol Hell and Jaroslav Nešetřil. On the complexity of H-coloring. *J. Combin. Theory Ser. B*, 48(1):92–110, 1990. ISSN 0095-8956. doi: 10.1016/0095-8956(90)90132-J. URL http://dx.doi.org/10.1016/0095-8956(90)90132-J.

Peter Jeavons, David Cohen, and Marc Gyssens. Closure properties of constraints. *Journal of the ACM*, 44, 07 1997. doi: 10.1145/263867.263489.

Alexandr Kazda. Taylor term does not imply any nontrivial linear one-equality Maltsev condition. Preprint arXiv:1409.4601, 2017.

Keith Kearnes, Petar Marković, and Ralph McKenzie. Optimal strong Mal'cev conditions for omitting type 1 in locally finite varieties. *Algebra Universalis*, 72(1):91–100, 2014. ISSN 0002-5240. doi: 10.1007/s00012-014-0289-9. URL http://dx.doi.org/10.1007/s00012-014-0289-9.

Marcin Kozik, Andrei Krokhin, Matt Valeriote, and Ross Willard. Characterizations of several maltsev conditions. *Algebra Universalis*, 73(3):205–224, Jun 2015. ISSN 1420-8911. doi: 10.1007/s00012-015-0327-2. URL https://doi.org/10.1007/s00012-015-0327-2.

Thomas J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing*, STOC '78, pages 216–226, New York, NY, USA, 1978. ACM. doi: 10.1145/800133.804350. URL http://doi.acm.org/10.1145/800133.804350.

Mark H. Siggers. A strong Mal'cev condition for locally finite varieties omitting the unary type. *Algebra Universalis*, 64(1-2):15–20, 2010. ISSN 0002-5240. doi: 10.1007/s00012-010-0082-3. URL http://dx.doi.org/10.1007/s00012-010-0082-3.

Dmitriy Zhuk. A proof of CSP dichotomy conjecture. *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 331–342, 2017.

List of publications

List of publications

[1] The Weakest Nontrivial Idempotent Equations, Bulletin of the London Mathematical Society, 49(6):1028–1047, 2017,

[2] The equivalence of two dichotomy conjectures for infinite domain constraint satisfaction problems (with L. Barto, M. Kompatscher, M. Pinsker, and T. van Pham), Proceedings of the Symposium on Logic in Computer Science (LICS) 2017, 1-12.

[3] Reinforcement Learning of Theorem Proving (with C. Kaliszyk, J. Urban, and H. Michalewski), Advances in Neural Information Processing Systems 31 (NIPS 2018),

[4] Equations in oligomorphic clones and the Constraint Satisfaction Problem for omega-categorical structures (with L. Barto, M. Kompatscher, M. Pinsker, and T. van Pham), to appear in Journal of Mathematical Logic.

[5] Topology is relevant (with M. Bodirsky, A. Mottet, M. Pinker, J. Opršal, and R. Willard), submitted to LICS 2019

[6] Dichotomy for symmetric Boolean PCSPs (with M. Ficak, M. Kozik, and S. Stankiewicz), submitted to ICALP 2019

[7] Maltsev Conditions for General Congruence Meet-Semidistributive Algebras, submitted to Journal of Symbolic Logic

[8] Loop Conditions, submitted to Algebra Universalis

[9] Loop Conditions for Strongly Connected Digraphs, submitted to International Journal of Algebra and Computation

[10] Local Loop Lemma, submitted to Algebra Universalis