

Constraint Satisfaction Problems with Infinite Templates

Manuel Bodirsky¹

Department of Algorithms and Complexity, Humboldt University, Berlin,
`bodirsky@informatik.hu-berlin.de`

Abstract. Allowing templates with infinite domains greatly expands the range of problems that can be formulated as a non-uniform constraint satisfaction problem. It turns out that many CSPs over infinite templates can be formulated with templates that are ω -categorical. We survey examples of such problems in temporal and spatial reasoning, infinite-dimensional algebra, acyclic colorings in graph theory, artificial intelligence, phylogenetic reconstruction in computational biology, and tree descriptions in computational linguistics.

We then give an introduction to the universal-algebraic approach to infinite-domain constraint satisfaction, and discuss how cores, polymorphism clones, and pseudo-varieties can be used to study the computational complexity of CSPs with ω -categorical templates. The theoretical results will be illustrated by examples from the mentioned application areas. We close with a series of open problems and promising directions of future research.

1 Introduction

Some of the oldest results in constraint satisfaction concern constraint satisfaction problems over *infinite* domains. Examples are the constraint satisfaction problem (CSP) for Allen's interval algebra and its fragments, and the CSPs for many other temporal reasoning formalisms. Also in spatial reasoning, some of the earliest computational problems can be formulated as CSPs, e.g., the CSP for the region connection calculus.

In the 1990's, several systematic results about the computational complexity of CSPs in spatial and temporal reasoning were obtained, mostly for binary constraint languages. An important step was Nebel and Bürckert's discovery of Ord-Horn, a tractable fragment of Allen's interval algebra. In a computer-assisted proof, they showed that Ord-Horn is a *largest tractable* fragment of Allen's interval algebra, i.e., adding any (binary) relation from Allen's interval algebra to Ord-Horn results in an NP-complete constraint language. Another important step in complexity classification was the complete classification of all tractable fragments of the region connection calculus RCC-5 in spatial reasoning by Jonsson and Drakengren [44]. The corresponding result for Allen's interval algebra was obtained later by Krokhin, Jeavons, and Jonsson [49]. Similar classifications for the cardinal direction calculus, for constraint languages for branching time, and constraint languages for partially ordered time were mostly obtained by a case-by-case study and ad-hoc methods. There are some common themes [29] and tools (e.g., local consistency techniques), but so far there is no general theory of tractability and hardness for such problems.

A typical property of constraint satisfaction problems is that if we add constraints to an unsatisfiable instance of the CSP, the instance stays unsatisfiable. Moreover, if we

form the disjoint union of two satisfiable instances of the CSP, the resulting instance is again satisfiable. In fact, all computational problems (i.e., all sets of relational structures) that share these two properties can be formulated as a homomorphism problem for a fixed infinite relational structure Γ (in other words, as a non-uniform CSP with an infinite template). Many examples of computational problems that have been studied in the literature have a natural formulation as a homomorphism problem for a fixed infinite structure Γ , and we will mention examples from graph theory, artificial intelligence, phylogenetic reconstruction, finite model theory, and tree description logics in computational linguistics.

This survey is about techniques to study the computational complexity of CSPs with infinite templates. We focus mostly on the border between CSPs that can be solved by a polynomial-time algorithm, and CSPs that are known to be NP-hard. This border has been of central interest in most of the mentioned work on temporal and spatial reasoning. Our main theme will be the question which of the powerful universal-algebraic techniques that are employed to study the complexity of the finite domain CSP (see e.g. [22]) can be generalized to infinite domain templates. It turns out that the universal-algebraic approach applies if the template satisfies an important and central property in model theory, called ω -categoricity.

The article is divided into two parts: in the first part we demonstrate that the class of computational problems that can be formulated with an ω -categorical template is very large (actually, it contains all the infinite-domain CSPs we have mentioned so far). In the second part, we set out to develop the universal-algebraic theory for ω -categorical templates. We recall the definitions of polymorphisms, algebras, and (pseudo-) varieties, and show that these concepts are useful not only for finite but also for infinite ω -categorical structures. Some of the tractability results presented here do not have a finite domain counterpart. We only give proofs of the statements in this text if they are instructive, previously unpublished, or more conceptual than the existing proofs in the literature, and give references to the literature in all other cases.

2 Infinite Templates

The notation used in this text mostly follows Hodges' text book [42]. A *signature* τ is a set of relation and function symbols, each equipped with an arity. A τ -*structure* \mathbf{A} is a set A (the *domain* of \mathbf{A}) together with a relation $R^{\mathbf{A}} \subseteq A^k$ for each k -ary relation symbol in τ and a function $f^{\mathbf{A}} : A^k \rightarrow A$ for each k -ary function symbol in τ . When there is no danger of confusion, we use the same symbol for a function and its function symbol, and for a relation and its relation symbol. By convention, A, B, C, \dots denotes the domain of the structures $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$, respectively. We sometimes write $(A, R_1^{\mathbf{A}}, R_2^{\mathbf{A}}, \dots, f_1^{\mathbf{A}}, f_2^{\mathbf{A}}, \dots)$ for the relational structure \mathbf{A} with relations $R_1^{\mathbf{A}}, R_2^{\mathbf{A}}, \dots$ and functions $f_1^{\mathbf{A}}, f_2^{\mathbf{A}}, \dots$. We say that a structure is infinite if its domain is infinite. The most important special cases of structures that appear in this paper are *relational structures*, that is, structures with a purely relational signature, and *algebras*, that is, structures with a purely functional signature. Algebras with domain A, B, C, \dots are denoted by $\mathbb{A}, \mathbb{B}, \mathbb{C}, \dots$.

Homomorphisms. A *homomorphism* h from a structure \mathbf{A} to a structure \mathbf{B} with the same signature τ is a mapping from A to B that *preserves* each function and each relation for the symbols in τ ; that is,

- if (a_1, \dots, a_k) is in R^A , then $(h(a_1), \dots, h(a_k))$ must be in R^B ;
- if $f^A(a_1, \dots, a_k) = a_0$, then $f^B(h(a_1), \dots, h(a_k)) = h(a_0)$.

In this article, a (*non-uniform*) *constraint satisfaction problem (CSP)* is a computational problem that is specified by a single structure with a finite relational signature, called the *template* of the CSP. Relational structures that denote templates for CSPs will be denoted by capital greek letters Γ, Δ (and their domain by $D(\Gamma), D(\Delta)$, respectively).

Definition 1. Let Γ be a (possible infinite) relational structure with a finite relational signature τ . Then $\text{CSP}(\Gamma)$ is the computational problem to decide whether a given finite τ -structure homomorphically maps to Γ .

We sometimes also write $\text{CSP}(D, R_1, \dots, R_l)$ instead of $\text{CSP}((D, R_1, \dots, R_l))$. Note that due to the assumption that the signature τ is finite, we can fix any representation of the relation symbols in τ to represent the input structure. $\text{CSP}(\Gamma)$ can also be considered to be a set – the set of all finite structures that homomorphically map to Γ .

Logic. As for finite domain CSPs, there is another way of looking at the constraint satisfaction problem for Γ that uses terminology from logic. A first-order τ -formula is called *primitive positive* if it is of the form

$$\exists x_1, \dots, x_m. \psi_1 \wedge \dots \wedge \psi_l$$

where ψ_1, \dots, ψ_l are atomic τ -formulas.

It is straightforward to verify (in the same way as for finite domain CSPs) that $\text{CSP}(\Gamma)$ is polynomial-time equivalent to the following computational problem (in fact, the two problems can be considered to be the same computational problem, up to formalization). The input consists of a primitive positive τ -sentence Φ (i.e., a primitive positive τ -formula without free variables), and the question is whether Φ holds true in Γ . The conjuncts in the primitive positive sentence are then called the *constraints* of Φ . From this formulation of the constraint satisfaction problem it is obvious that $\text{CSP}(\Gamma)$ is fully determined by the *first-order theory of Γ* (i.e., the set of first-order sentences that are valid in Γ).

We would like to remark that the logic perspective on $\text{CSP}(\Gamma)$ is closely related to the evaluation problem for conjunctive queries studied in database theory. We freely switch between the relational homomorphism and the logic perspective whenever this is convenient. We also say that \mathbf{A} is *satisfiable* (with respect to $\text{CSP}(\Gamma)$) if \mathbf{A} homomorphically maps to Γ , and the homomorphism from \mathbf{A} to Γ is called a *solution* for \mathbf{A} .

Example 1. Consider the problem $\text{CSP}(\mathbb{N}, =, \neq)$. An instance of this problem consists of a set of variables, some linked by equality, some by disequality constraints. Such an instance is unsatisfiable if and only if there is a path x_1, \dots, x_n from a variable x_1 to a variable x_n that uses only equality edges, i.e., ‘ $x_i = x_{i+1}$ ’ is a constraint in the instance for each $1 \leq i \leq n - 1$, and additionally ‘ $x_1 \neq x_n$ ’ is a constraint in the instance. Clearly, it can be tested in linear time in the size of the input instance whether the instance contains such a path.

Example 2. Next, consider the problem $\text{CSP}(\mathbb{Q}, <)$. Here, $<$ denotes the strict linear order on the relational numbers \mathbb{Q} (i.e., $<$ is a binary relation). An instance of this problem can be viewed as a directed graph (potentially with loops), where there is an arc between the

vertices x and y if there is the constraint ‘ $x < y$ ’ in the instance. It is easy to see that an instance homomorphically maps to $(\mathbb{Q}, <)$ if and only if there is no directed cycle in the graph. Again, this can be tested in linear time, e.g., by depth-first search.

Example 3. The so-called *betweenness problem* is $\text{CSP}(\mathbb{Q}, \text{Betw})$ where *Betw* is the ternary relation $\{(x, y, z) \in \mathbb{Q}^3 \mid x < y < z \vee z < y < x\}$. This problem is an NP-complete problem from the famous book of Garey and Johnson [37].

We have defined the constraint satisfaction problem only for relational structures, but the generalization to structures that also include function symbols is straightforward, and has been studied for finite domains [47]. If we allow function symbols *and* infinite domains, we can for instance formulate the famous unification problem (see, e.g., [56]) as $\text{CSP}(\Gamma)$ for an appropriate infinite structure Γ . In this article we focus on relational templates only.

Having structures with function symbols will be convenient in Section 8 where we use algebras to study the computational complexity of CSPs. Moreover, several templates that we use to illustrate the theory have a convenient definition by means of structures that contain functions, e.g., in Subsections 5.3 and 5.2.

The next lemma is a convenient tool to determine whether a computational problem can be formulated as $\text{CSP}(\Gamma)$ for an infinite relational structure Γ . The *disjoint union* of a set of τ -structures \mathcal{C} is a τ -structure \mathbf{B} whose domain is the disjoint union of the domains of the structures in \mathcal{C} . The relations in \mathbf{B} are the union of the corresponding relations in the structures in \mathcal{C} .

Definition 2. *We say that a set \mathcal{C} of relational structures is closed under (finite) disjoint unions iff whenever $\mathbf{A}, \mathbf{B} \in \mathcal{C}$ then the disjoint union of \mathbf{A} and \mathbf{B} is also in \mathcal{C} . We say that \mathcal{C} is closed under inverse homomorphisms iff whenever $\mathbf{B} \in \mathcal{C}$ and \mathbf{A} homomorphically maps to \mathbf{B} then $\mathbf{A} \in \mathcal{C}$.*

The following is a fundamental lemma for constraint satisfaction problems.

Lemma 1. *Let \mathcal{C} be a set of finite τ -structures, for a finite relational signature τ . Then $\mathcal{C} = \text{CSP}(\Gamma)$ for some relational structure Γ over an infinite domain if and only if \mathcal{C} is closed under disjoint unions and inverse homomorphisms.*

Proof. Clearly, the disjoint union of two instances of $\text{CSP}(\Gamma)$ that homomorphically map to Γ also homomorphically maps to Γ . Moreover, if \mathbf{A} does not homomorphically map to Γ , and there is a homomorphism from \mathbf{A} to \mathbf{B} , then \mathbf{B} does not homomorphically map to Γ either.

For the other direction, suppose that \mathcal{C} is a set of relational structures that is closed under disjoint unions and inverse homomorphisms. Let Γ be the (infinite) disjoint union of all structures in \mathcal{C} . Clearly, every structure in \mathcal{C} homomorphically maps to Γ . Now, let \mathbf{A} be a finite structure with a homomorphism h to Γ . By construction of Γ , the set $h(A)$ is contained in the disjoint union \mathbf{B} of a finite set of structures from \mathcal{C} . Since \mathcal{C} is closed under disjoint unions, \mathbf{B} is in \mathcal{C} . Clearly, \mathbf{A} homomorphically maps to \mathbf{B} , and because \mathcal{C} is closed under inverse homomorphisms, \mathbf{A} is in \mathcal{C} as well. \square

To apply the techniques presented in this paper it will be important that the templates of the CSPs are ω -categorical (ω -categoricity will be defined in Section 3). The structures produced by Lemma 1 are usually not ω -categorical.

Let us conclude this section with remarks concerning the computational complexity of $\text{CSP}(\Gamma)$. So far, we have seen CSPs in Examples 1, 2, and 3 that are in P or NP-complete. However, it is not hard to come up with undecidable CSPs. For example, Hilbert’s 10th problem, the problem to decide whether a given diophantine equation has a solution, is undecidable [54], and is computationally equivalent to $\text{CSP}(\mathbb{Z}, R^+, R^*)$ where $R^+ = \{(x, y, z) \in \mathbb{Z}^3 \mid x + y + z = 1\}$ and $R^* = \{(x, y, z) \in \mathbb{Z}^3 \mid xy = z\}$.

Bauslaugh [6] has constructed for every recursive function f an infinite digraph Γ such that $\text{CSP}(\Gamma)$ is decidable, but has time complexity at least f . In this paper, we focus on CSPs that are in NP.

3 Good Templates: ω -categorical Structures

Many important infinite-domain constraint satisfaction problems can be formulated with templates that are ω -categorical. The concept of ω -categoricity is of central importance in model theory.

Definition 3. *A countably infinite structure Γ is called ω -categorical if all countable models of its first-order theory are isomorphic to Γ .*

One of the first structures that were found to be ω -categorical (by Kantor) is the linear order of the rational numbers $(\mathbb{Q}, <)$. We will see many more examples of ω -categorical structures in this section and in Section 5. One of the standard approaches to verify that a structure is ω -categorical is via a so-called *back-and-forth* argument. We sketch the back-and-forth argument that shows that $(\mathbb{Q}, <)$ is ω -categorical; much more detail about this important concept in model theory can be found in [42, 57]. Let \mathbf{A} be a countable model of the first-order theory of $(\mathbb{Q}, <)$. An isomorphism i between \mathbf{A} and $(\mathbb{Q}, <)$ can be defined inductively as follows. Suppose that we have already defined i on a finite subset S of \mathbb{Q} and that f is an embedding of the structure \mathbf{S} induced by S in $(\mathbb{Q}, <)$ into \mathbf{A} . Since $<^{\mathbf{A}}$ is dense and unbounded, we can extend f to any other element of \mathbb{Q} such that the extension is still an embedding from a substructure of \mathbb{Q} into \mathbf{A} (*going forth*). Symmetrically, for every element v of \mathbf{A} we can find an element $u \in \mathbb{Q}$ such that the extension of f that maps u to v is also an embedding (*going back*). We now alternate between going forth and going back; when going forth, we extend the domain of f by the *next* element of \mathbb{Q} , according to some fixed enumeration of the elements in \mathbb{Q} . When going back, we extend f such that the image of \mathbf{A} contains the *next* element of \mathbf{A} , according to some fixed enumeration of the elements of \mathbf{A} . If we continue in this way, we have defined the value of f on all elements of \mathbb{Q} . Moreover, f will be surjective, and an embedding, and hence an isomorphism between $(\mathbb{Q}, <)$ and \mathbf{A} .

There are many equivalent characterizations of ω -categoricity. Let G be a permutation group acting on a set X . For $n \geq 1$ the *orbit* of $(t_1, \dots, t_n) \in X^n$ under G is the set $\{(\alpha(t_1), \dots, \alpha(t_n)) \mid \alpha \in G\}$. We say that a permutation group is *oligomorphic* if for each $n \geq 1$ there are finitely many orbits of n -tuples. An accessible proof of the following theorem can be found in Hodges’ book (Theorem 6.3.1 in [42]).

Theorem 1 (Engeler, Ryll-Nardzewski, Svenonius). *The following are equivalent:*

1. Γ is ω -categorical;
2. The automorphism group of Γ is oligomorphic;

3. for each $n \geq 1$, there are finitely many inequivalent formulas with n free variables over Γ ;
4. for each $n \geq 1$, the orbits of n -tuples are first-order definable in Γ .

The second condition in Theorem 1 provides another possibility to verify that a structure is ω -categorical. We again illustrate this with the structure $(\mathbb{Q}, <)$. It is not difficult but a good exercise to verify that the orbit of an n -tuple (t_1, \dots, t_n) from \mathbb{Q}^n in the automorphism group of $(\mathbb{Q}, <)$ is determined by the weak linear order that is induced by t_1, \dots, t_n in $(\mathbb{Q}, <)$ (we write *weak linear order*, and not *linear order*, because some of the elements t_1, \dots, t_n might be equal). Hence, there is a finite number of orbits of n -tuples, for all $n \geq 1$.

Lemma 2 below states a useful fact that ω -categorical structures have in common with finite structures. A homomorphism h from \mathbf{A} to \mathbf{B} is called a *strong homomorphism* if it also preserves the complements of the relations from \mathbf{A} . Note that an embedding is an injective strong homomorphism.

Lemma 2. *Let Γ be a finite or infinite ω -categorical structure with relational signature τ , and let \mathbf{A} be a countable relational structure with the same signature τ . If there is no homomorphism (embedding) from \mathbf{A} to Γ , then there is a finite substructure of \mathbf{A} that does not homomorphically map (embed) to Γ .*

The lemma is an easy consequence of Königs tree lemma; a proof for homomorphisms can be found in [11], and the proof for embeddings (and, similarly, for strong homomorphisms and for injective homomorphisms) is analogous.

3.1 Fraïssé Amalgamation – ω -categorical Structures for Everyone

A versatile tool to construct ω -categorical templates is Fraïssé-amalgamation. We present it here for structures that might contain relations *and* functions; this will sometimes be convenient even if we are only interested in relational structures (see Subsections 5.2 and 5.3).

If \mathbf{A} and \mathbf{B} are structures, $A \subseteq B$, and the inclusion map from A to B is an embedding, then we say that \mathbf{A} is a *substructure* of \mathbf{B} . Note that for every subset S of B there is a unique smallest substructure \mathbf{A} of \mathbf{B} such that A contains S . We call \mathbf{A} the *substructure of \mathbf{B} generated by S* , also denoted by $\mathbf{B}[S]$. We say that \mathbf{A} is *finitely generated* if S is a finite set of elements.

Let \mathcal{C} be a set of finitely generated structures. We say that \mathcal{C} has the

- HP *Hereditary property* if whenever $\mathbf{A} \in \mathcal{C}$ and \mathbf{B} is a finitely generated substructure of \mathbf{A} then \mathbf{B} is isomorphic to a structure in \mathcal{C} .
- JEP *Joint embedding property* if whenever $\mathbf{A}, \mathbf{B} \in \mathcal{C}$ then there is $\mathbf{C} \in \mathcal{C}$ such that both \mathbf{A} and \mathbf{B} embed into \mathbf{C} .
- AP *Amalgamation property* if whenever $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2 \in \mathcal{C}$ and $e_1 : \mathbf{A} \rightarrow \mathbf{B}_1$ and $e_2 : \mathbf{A} \rightarrow \mathbf{B}_2$ are embeddings there exists $\mathbf{C} \in \mathcal{C}$ and embeddings $f_1 : \mathbf{B}_1 \rightarrow \mathbf{C}$ and $f_2 : \mathbf{B}_2 \rightarrow \mathbf{C}$ such that $f_1 e_1 = f_2 e_2$.

A structure Γ is *homogeneous* (sometimes also called *ultra-homogeneous* [42]) if every isomorphism between finitely generated substructures of Γ can be extended to an automorphism of Γ . It is well-known that a homogeneous structure Γ with a signature that

contains finitely many relation symbols of arity k , for each k , is ω -categorical (this is not explicitly mentioned in Hodges' book [42], but can easily be shown with the background from there). Moreover, Γ admits *quantifier elimination*, i.e., for every first-order τ -formula there exists an equivalent quantifier-free τ -formula [42]. In fact, an ω -categorical structure is homogeneous if and only if it admits quantifier elimination (Statement 2.22 in [25]).

Theorem 2 (Fraïssé [36]; see [42]). *Let τ be a countable signature and let \mathcal{C} be a non-empty finite or countable set of finitely generated τ -structures which has HP, JEP, and AP. Then there is a homogeneous and at most countable τ -structure Γ such that a structure is a finitely generated substructure of Γ if and only if it is isomorphic to a structure in \mathcal{C} . The structure Γ is unique up to isomorphism, and called the Fraïssé-limit of \mathcal{C} .*

With Theorem 2 we have seen a third possibility how to verify the ω -categoricity of our running example $(\mathbb{Q}, <)$: it suffices to verify that the class of all finite weak linear orders is an amalgamation class.

Remark 1. It is sometimes convenient to define an ω -categorical τ -structure Γ by defining an amalgamation class \mathcal{C} with a signature that is larger than τ such that Γ is a reduct¹ of the Fraïssé-limit of \mathcal{C} . It is an easy consequence of Theorem 1 that Γ must then be ω -categorical.

3.2 New ω -categorical structures from old

Many ω -categorical structures can be derived from other ω -categorical structures via first-order interpretations. Our definition of first-order interpretations essentially follows [42].

If $\delta(x_1, \dots, x_k)$ is a first-order formula with the k free variables x_1, \dots, x_k , and \mathbf{A} is a structure, we write $\delta(\mathbf{A}^k)$ for the k -ary relation that is defined by δ on \mathbf{A} .

Definition 4. *A relational σ -structure \mathbf{B} has a (first-order) interpretation in a τ -structure \mathbf{A} if there exists a natural number d , called the dimension of the interpretation, and*

- a τ -formula $\delta(x_1, \dots, x_d)$ – called domain formula,
- for each k -ary relation symbol R in σ a τ -formula $\phi_R(\bar{x}_1, \dots, \bar{x}_k)$ where the \bar{x}_i denote disjoint d -tuples of distinct variables – called the defining formulae,
- a τ -formula $\phi_=(x_1, \dots, x_d, y_1, \dots, y_d)$, and
- a surjective map $h : \delta(\mathbf{A}^d) \rightarrow B$ – called coordinate map,

such that for all relations R in \mathbf{B} and all tuples $\bar{a}_i \in \delta(\mathbf{A}^d)$

$$\begin{aligned} (h(\bar{a}_1), \dots, h(\bar{a}_k)) \in R^{\mathbf{B}} &\Leftrightarrow \mathbf{A} \models \phi_R(\bar{a}_1, \dots, \bar{a}_k), \text{ and} \\ h(\bar{a}_1) = h(\bar{a}_2) &\Leftrightarrow \mathbf{A} \models \phi_=(\bar{a}_1, \bar{a}_2). \end{aligned}$$

If the formulas δ , ϕ_R , and $\phi_ =$ are all primitive positive, we say that \mathbf{B} has a primitive positive interpretation in \mathbf{A} . We say that \mathbf{B} is *interpretable in Γ with finitely many parameters* if there are $c_1, \dots, c_n \in A$ such that \mathbf{B} is interpretable in the expansion of \mathbf{A} by the singleton relations $\{c_i\}$ for all $1 \leq i \leq n$. First-order *definitions* are a special case of interpretations: a structure \mathbf{B} is *(first-order) definable in \mathbf{A}* if \mathbf{B} has an interpretation in \mathbf{B} of dimension one where the domain formula is logically equivalent to true.

¹ A *reduct* of a structure \mathbf{A} is a structure that is obtained from \mathbf{A} by removing relations and/or functions from \mathbf{A} .

Lemma 3 (see e.g. [42]). *If Γ is an ω -categorical structure, then every structure Δ that is first-order interpretable in Γ with finitely many parameters is ω -categorical as well.*

Suppose that \mathbf{A} and \mathbf{B} are finite or ω -categorical structures. In the remainder of the section we show that then the *disjoint union* $\mathbf{A} \uplus \mathbf{B}$ (which was defined in Section 2) and the *composition* $\mathbf{A} \circ \mathbf{B}$ (Definition 5 below) of \mathbf{A} and \mathbf{B} are finite or ω -categorical as well. Again, this will be convenient to specify templates of constraint satisfaction problems.

The composition of \mathbf{A} and \mathbf{B} is, roughly speaking, the structure obtained from \mathbf{A} by replacing each element in \mathbf{A} by a copy of \mathbf{B} .

Definition 5. *If \mathbf{A}, \mathbf{B} are τ -structures, then $\mathbf{A} \circ \mathbf{B}$ is the τ -structure with domain $A \times B$, where $((a_1, b_1), \dots, (a_k, b_k)) \in R^{\mathbf{A} \circ \mathbf{B}}$ for a k -ary relation symbol $R \in \tau$ iff $(a_1, \dots, a_k) \in R^{\mathbf{A}}$ or $(a_1 = \dots = a_k$ and $(b_1, \dots, b_k) \in \mathbf{B}$).*

To verify ω -categoricity of disjoint union and composition it is most convenient to use the second condition from Theorem 1 (i.e., we will count the orbits of k -tuples in the resulting structure). For an ω -categorical structure \mathbf{A} , let $f_n(\mathbf{A})$ be the number of orbits of n -tuples in the automorphism group of \mathbf{A} , and let $F(\mathbf{A}, x)$ be the exponential generating function in one formal variable x , defined by $\sum_{n \geq 0} f_n(\mathbf{A}) x^n / n!$. The following reflects well-known facts in enumerative combinatorics (for a presentation in the context of oligomorphic permutation groups, see [25]).

Lemma 4. *Let \mathbf{A}, \mathbf{B} be finite or ω -categorical structures. Then*

$$\begin{aligned} F(\mathbf{A} \uplus \mathbf{B}, x) &= F(\mathbf{A}, x)F(\mathbf{B}, x) \\ F(\mathbf{A} \circ \mathbf{B}, x) &= F(\mathbf{A}, F(\mathbf{B}, x) - 1) \end{aligned}$$

In particular, for all $n \geq 1$ we have that $f_n(\mathbf{A} \uplus \mathbf{B})$ and $f_n(\mathbf{A} \circ \mathbf{B})$ are finite, and therefore $\mathbf{A} \uplus \mathbf{B}$ and $\mathbf{A} \circ \mathbf{B}$ is finite or ω -categorical.

We would like to remark that the automorphism group of $\mathbf{A} \circ \mathbf{B}$ is also known as the *wreath product* of the automorphism group of \mathbf{B} with the automorphism group of \mathbf{A} .

4 Better Templates: Model-complete Cores

It might be that the same CSP can be formulated with different ω -categorical templates. For example, if T_3 is the transitive tournament on three vertices, then $\text{CSP}(\mathbb{Q}, <)$ and $\text{CSP}(T_3 \circ (\mathbb{Q}, <))$ (for the definition of the operation \circ , see Subsection 3.2) are the same computational problem, namely graph acyclicity.

For finite and for ω -categorical templates we have a very elegant characterization of those pairs of templates that have the same CSP. Call two structures Γ and Δ *homomorphically equivalent* if there is a homomorphism from Γ to Δ and a homomorphism from Δ to Γ (this is in fact an equivalence relation). The following is an easy consequence of Lemma 2.

Lemma 5. *Let Γ, Δ be finite or ω -categorical. Then $\text{CSP}(\Gamma)$ equals $\text{CSP}(\Delta)$ if and only if Γ and Δ are homomorphically equivalent.*

Lemma 5 is false for general relational structures. Consider for example the structure $(\mathbb{Z}, \{(x, y) \mid y = x + 1\})$ — the ‘infinite line’, and the structure $(\mathbb{N}, \{(x, y) \mid y = x + 1\})$ — the ‘infinite ray’. Clearly, these two structures give rise to the same CSP, but there is no homomorphism from the line to the ray.

It turns out that every equivalence class of finite or ω -categorical templates (with respect to homomorphic equivalence) has a member with very good properties for the universal-algebraic approach. Moreover, this member is unique up to isomorphism.

Definition 6. *An ω -categorical structure Γ is called model-complete if every embedding of Γ into Γ is elementary, i.e., preserves all first-order definable relations. An ω -categorical structure is called a core if every endomorphism is an embedding.*

The structure $(\mathbb{Q}, <)$, for example, is easily seen to be a model-complete core, in contrast to $T_3 \circ (\mathbb{Q}, <)$, which is not a core.

Note that for every finite structure Γ every embedding of Γ into Γ is elementary (because it must be an automorphism of Γ).

Theorem 3 (of [7]). *Every ω -categorical relational structure Γ is homomorphically equivalent to a model-complete core Δ , which is unique up to isomorphism, and ω -categorical or finite. For all $k \geq 1$, the orbits of k -tuples in Δ are primitive positive definable.*

Hence, if we want to classify the complexity of $\text{CSP}(\Gamma)$ for an ω -categorical structure Γ , we can without loss of generality assume that Γ is a core. We state an important consequence of Theorem 3, which is well-known (and also non-trivial) for finite templates. The consequence will be used in Theorem 10.

Corollary 1 (of [7]). *Let Γ be an ω -categorical model-complete core, and let Δ be the expansion of Γ by finitely many singleton relations, i.e., relations of the form $\{a\}$ for $a \in D(\Gamma)$. Then $\text{CSP}(\Gamma)$ and $\text{CSP}(\Delta)$ are polynomial-time equivalent.*

5 Examples

Most of the examples of CSPs that we present in this Section can be solved in polynomial-time, and we come back to most of these problems in Section 10 when discussing tractability criteria for CSPs. Coming up with NP-hard CSPs is much easier: each CSP in this section becomes NP-hard if we expand the template by the relation $\{(x, y, z) \mid x = y \vee y = z\}$. More on NP-hardness criteria of CSPs follows in later sections, culminating in Section 9.

5.1 Temporal Reasoning

Temporal reasoning is an important sub-discipline in Artificial Intelligence. One of the most basic temporal reasoning problems is the constraint satisfaction problem of the so-called *point algebra*. Here, the variables denote time points, and the constraints are of the form $x = y$, $x < y$, $x \leq y$, and $x \neq y$. In this CSP, it does not matter whether we use as the domain the natural, the rational, or the real numbers. The problem can e.g. be formalized as $\text{CSP}(\mathbb{Q}, =, <, \leq, \neq)$, and can be solved in linear time in the size of the input instances.

If we do not restrict our attention to binary constraints, it is natural to study expansions of the point-algebra by higher-ary relations with a first-order definition in $(\mathbb{Q}, <)$. For instance, consider the relation

$$R^{\min} = \{(x, y, z) \mid x > y \vee x > z\}.$$

The relation $R^{\min}(x, y, z)$ specifies that x is larger than the minimum of y and z . The constraint satisfaction problem $\text{CSP}(\mathbb{Q}, R^{\min})$ will be discussed in Section 10.

Another famous temporal reasoning problem is the CSP for *Allen's Interval Algebra*. It is easiest to describe the template for this CSP by an interpretation in $(\mathbb{Q}, <)$. The dimension of this interpretation is two, and the domain formula $\delta(x, y)$ is $x < y$. Hence, the variables of the CSP denote non-empty time intervals. The template contains for each inequivalent $\{<\}$ -formula with four variables ϕ a binary relation R such that (a_1, a_2, a_3, a_4) satisfies ϕ if and only if $((a_1, a_2), (a_3, a_4)) \in R$. In particular, we have relations for containment of intervals, disjointness of intervals, and so forth. By Lemma 3, this template is ω -categorical. Its CSP is NP-complete [5]. The computational complexity of the CSP for all fragments of Allen's interval algebra has been determined in [49, 55].

5.2 Spatial Reasoning

One of the most fundamental spatial reasoning formalisms is the RCC-5 calculus (also known as the containment algebra in the theory of relation algebras [31]). In this formalism, the variables denote non-empty regions, and the basic relations in the calculus express containment of regions, disjointness of regions, etc; formal definitions will be provided below. The constraint satisfaction problem for RCC-5 is NP-complete; the computational complexity of its fragments was classified in [44, 58].

To formulate the CSPs for RCC-5 and its fragments with ω -categorical templates, let \mathbb{B} be the countable atomless boolean ring without an identity element. That is, \mathbb{B} is an algebra with an operation $+$ for addition and an operation \cdot for multiplication satisfying the usual axioms for rings. A ring is *boolean* if it satisfies $x \cdot x = x$ for all elements x . A ring is *without (multiplicative) identity element* if there is no element x_1 such that $x_1 \cdot y = y$ for all elements y . A ring without identity element is *atomless* if the partial order \leq defined by $x \leq y$ if $x \cdot y = x$ does not have minimal elements. Every countable atomless boolean ring without an identity element is unique up to isomorphism [1], homogeneous (this can be shown by Theorem 2), and hence ω -categorical. We can interpret the elements of this boolean ring as non-empty sets (regions), where $x + y$ denotes the symmetric difference, and $x \cdot y$ the intersection of x and y .

Now, consider the relational structure that has the same domain B as the boolean ring \mathbb{B} , and that contains all binary relations with a first-order definition in \mathbb{B} . This is exactly the template for the constraint satisfaction problem for RCC-5 (several equivalent definitions can be found in the literature). Ordered by inclusion, there are five minimal non-empty binary relations with a first-order definition in \mathbb{B} , and they are known as the *basic relations* of RCC-5. Traditionally, these relations are denoted by DR, PO, PP, PPI, EQ, and they have the following first-order definitions in \mathbb{B} (with their intuitive meaning in braces).

$\text{DR}(x, y)$ iff $(x + y)x = x$	‘ x and y are disjoint’
$\text{PP}(x, y)$ iff $xy = x \wedge x \neq y$	‘ y properly contains x ’
$\text{PPI}(x, y)$ iff $xy = y \wedge x \neq y$	‘ x properly contains y ’
$\text{EQ}(x, y)$ iff $x = y$	‘ x equals y ’
$\text{PO}(x, y)$ iff $\neg(\text{DR}(x, y) \vee \text{PP}(x, y) \vee \text{PPI}(x, y) \vee \text{EQ}(x, y))$	‘ x and y properly overlap’

It is known that $\text{CSP}(B, \text{DR}, \text{PP}, \text{PPI}, \text{EQ}, \text{PO})$ is in P [58]. A larger tractable fragment of RCC-5 is discussed in Section 10.

5.3 Vector Space CSPs

Next, we discuss a natural algebraic constraint satisfaction problem.

Vector Space CSP

INSTANCE: A set of equalities of the form $x + y = z$ and disequalities of the form $x + y \neq z$ over a set of variables V .

QUESTION: Can we assign d -dimensional boolean vectors to the variables such that all the equalities and disequalities are satisfied, for some d ?

It is easy to observe that if there exists a solution, then there is a solution where $d = |V|$. However, note that $|V|$ grows with the input size, and it is easy to see that this problem cannot be formulated with a finite template.

The vector space CSP has a natural formulation as a CSP with an ω -categorical template. Let \mathbb{V} be the infinite-dimensional vector space over the 2-element field \mathbb{F}_2 . Infinite-dimensional vector spaces over finite fields are known to be homogeneous and ω -categorical [33], and they play an important role in the classification of so-called *strictly minimal sets* in ω -categorical structures [26, 27, 61]. They are also examples of *totally categorical structures*, i.e., structures with a first-order theory that has one model of every infinite cardinality, up to isomorphism.

The signature of \mathbb{V} is operational, with function symbols for $+$ and a unary function for scalar multiplication for each element of the field. Then the template \mathbf{V} of the above CSP is the relational structure with the same domain as \mathbb{V} , and with the ternary relation $\{(x, y, z) \mid x + y = z\}$ and the ternary relation $\{(x, y, z) \mid x + y \neq z\}$. Since \mathbf{V} clearly has a first-order definition in \mathbb{V} , it is ω -categorical as well. It is not hard to see that \mathbf{V} is a model-complete core. An polynomial time algorithm for $\text{CSP}(\mathbf{V})$ will be discussed in Section 10.

5.4 Graph Coloring Problems

The \mathbf{H} -coloring problem in graph theory is the special case of the CSP where the template \mathbf{H} has a single binary relation, and therefore can be viewed as a digraph (directed graph). If we allow infinite digraphs \mathbf{H} , many more graph coloring problems that have been studied in the literature can be formulated as \mathbf{H} -coloring problems.

The complexity of the following class of infinite \mathbf{H} -coloring problems has been classified completely by Feder, Hell, and Mohar [34].

Acyclic H -coloring

INSTANCE: A finite digraph G .

QUESTION: Is there a mapping c from G to H such that for every edge xy in G the digraph H contains the arc $c(x)c(y)$, and such that the pre-image of every vertex is acyclic?

For every digraph H , we can formulate the acyclic H -coloring problem as a constraint satisfaction problem with an ω -categorical template. The acyclic H -coloring problem is $\text{CSP}(H \circ (\mathbb{Q}, <))$.

5.5 Phylogenetic Reconstruction

Modern biology holds the view that the species in the evolution of life on earth developed in a mostly tree-like fashion: at certain time periods, species separated into sub-species. The goal of phylogenetic reconstruction is to determine the evolutionary tree from given partial information about the tree. This motivates the computational problem of *rooted triple consistency*, defined below. In 1981, Aho, Sagiv, Szymanski, and Ullman [4] presented a quadratic time algorithm to this problem, motivated independently from computational biology by questions in database theory.

We fix some terminology concerning rooted trees. Let T be a tree with a distinguished vertex r , the *root* of T . For $u, v \in T$, we say that u *lies below* v if the path from u to r passes through v . We say that u *lies strictly below* v if u lies below v and $u \neq v$. The *youngest common ancestor* (*yca*) of two vertices $u, v \in T$ is the node w such that both u and v lie below w and w has maximal distance from r .

Rooted-Triple-Consistency

INSTANCE: A finite relational structure (V, R^{yca}) , where R^{yca} is a ternary relation.

QUESTION: Is there a rooted tree T with leaves X and a mapping $\alpha : V \rightarrow X$ such that for every triple $(x, y, z) \in R^{yca}$ the yca of $\alpha(x)$ and $\alpha(y)$ lies strictly below the yca of $\alpha(x)$ and $\alpha(z)$ in T ?

The rooted-triple-consistency problem can be formulated as $\text{CSP}(I)$ for an ω -categorical template I . To define I , we first consider the following structure Δ with domain $\mathbb{N} \rightarrow \mathbb{Q}$, i.e., the set of all functions from the natural numbers to the rational numbers (hence, Δ is uncountable). For two elements f, g of Δ , let $k_{f,g}$ be the largest natural number such that $f(i) = g(i)$ for all $i < k_{f,g}$. The ternary relation $f|g|h$ in Δ holds on elements f, g, h of Δ if they are pairwise distinct and either $k_{f,g} > k_{f,h}$ or ($k_{f,g} = k_{f,h}$ and $f(k_{f,g}) < h(k_{f,h})$). It is known that the first-order theory of Δ is ω -categorical [30]. It follows from the theorem of Löwenheim-Skolem (see [42]) that the first-order theory of Δ also has a countable model I , and it is straightforward to verify that $\text{CSP}(I)$ is the rooted triple consistency problem.

5.6 Reasoning over Branching-Time

Another important model in temporal reasoning is *branching time*, where for every time point the past is linearly ordered, but the future is only partially ordered.

Branching-Time-Consistency

INSTANCE: A finite relational structure $I = (V, \leq, \parallel, \neq)$ where \leq , \parallel , and \neq are binary

relations

QUESTION: Is there a rooted tree T and a mapping $\alpha : V \rightarrow T$ such that in T the following is satisfied:

- a) If $x \leq y$, then $\alpha(x)$ lies above $\alpha(y)$;
- b) If $x \parallel y$, then neither $\alpha(x)$ lies strictly above $\alpha(y)$ nor $\alpha(y)$ lies strictly above $\alpha(x)$;
- c) If $x \not\equiv y$, then $\alpha(x) \neq \alpha(y)$.

This problem can be formulated as $\text{CSP}(T)$ for an ω -categorical structure $(D, \leq, \parallel, \neq)$, which has been studied intensively in the theory of infinite permutation groups [25, 30]. The reduct (D, \leq) is a partial order with the property mentioned above: for all $x \in D$, the set $\{y \mid y \leq x\}$ is linearly ordered.

The first polynomial-time algorithm for this problem is due to Hirsch [41], and has a worst-case running time in $O(n^5)$. This was later improved by Broxvall and Jonsson [19], who presented an algorithm running in $O(n^{3.376})$ (this algorithm uses an $O(n^{2.376})$ algorithm for fast integer matrix multiplication). A simpler algorithm which does not use fast matrix multiplication and runs in $O(nm)$ can be found in [15]. It is an easy exercise to efficiently reduce the rooted-triple-consistency problem to the branching-time-consistency problem.

5.7 CSPs without an ω -categorical Template

Of course, there are constraint satisfaction problems that can not be formulated with an ω -categorical template. Since it might not be obvious whether a CSP has this property, we demonstrate with a simple example how one can show that a CSP can *not* be formulated with an ω -categorical template.

Consider the CSP for the relational structure Γ with domain \mathbb{Q} and the three relations $\{(x, y, z) \mid x + y = z\}$, \neq , $\{0\}$, and $\{1\}$. To specify constraints for instances of $\text{CSP}(\Gamma)$ we write $x + y = z$, $x = 0$, and $x = 1$, to keep expressions simple.

Proposition 1. *$\text{CSP}(\Gamma)$ cannot be formulated as $\text{CSP}(\Gamma')$ for an ω -categorical template Γ' .*

Proof. Suppose for contradiction that there is such an ω -categorical template Δ . We construct an infinite sequence $p_0, p_1, p_2 \dots$ of elements from pairwise distinct orbits of the automorphism group of Δ , which will be a contradiction by Theorem 1. Let z_0 and z_1 be elements in Δ such that $z_0 = 0$ and $z_1 = 1$ hold in Δ . Note that if there was an element z'_0 distinct from z_0 such that $z'_0 = 0$ holds in Δ , then $x = 0 \wedge x \neq y \wedge y = 0$ would be a satisfiable instance of $\text{CSP}(\Delta)$, but unsatisfiable in $\text{CSP}(\Gamma)$, which is impossible by assumption. Hence, z_0 is uniquely defined. Let p_0 be z_0 . Assume inductively that p_n is defined uniquely in Γ by a primitive positive formula $\phi(x)$. Consider the following instance of $\text{CSP}(\Gamma)$, which is specified by the primitive positive formula ψ :

$$\phi(s) \wedge (s + x = y) \wedge (y + x = l) \wedge l = 1 .$$

Clearly, if we set s to p_n , x to $(1 - p_n)/2$, and y to $(1 + p_n)/2$, all conjuncts of ψ are satisfied in Γ . By assumption, ψ is also satisfiable over Δ , i.e., there is an assignment $h : \{s, x, y, l\} \rightarrow D(\Delta)$ to the free variables that satisfies all conjuncts in ψ . Let p_{n+1} be $h(y)$. We claim that p_{n+1} is uniquely determined: if there was solution with another value

for y , then $\psi(s, x, y, l) \wedge \psi(s, x', y', l) \wedge x' \neq x$ would be satisfiable over Δ , but not over Γ , a contradiction. We also claim that p_{n+1} lies in an orbit that is distinct from the orbits of p_1, \dots, p_n , respectively: otherwise, if there was an automorphism α of Γ that maps p_i to p_{n+1} , for some $i \leq n$, then p_{n+1} would not have been unique, since $\alpha(h)$ is an assignment that also satisfies all conjuncts but assigns a different value to y . \square

6 CSPs and Existential Second-Order Logic

One of the motivating questions in the landmark paper of Feder and Vardi [35] is

Question 1. Which natural subclasses of NP do exhibit a *complexity dichotomy*, i.e., only contain problems that are either NP-complete or in P?

Whether the class FCSP of finite domain constraint satisfaction problems has such a complexity dichotomy is one of the greatest open problems in constraint satisfaction theory.

Feder and Vardi approached Question 1 from two sides. On the one hand, they conjectured that the class FCSP exhibits a dichotomy and made important contributions to this end. On the other hand, they investigated larger classes of computational problems that provably do not exhibit a complexity dichotomy. These larger classes are defined with syntactic restrictions of existential second order logic (ESO).

By Fagin's famous theorem, a set \mathcal{C} of finite τ -structures is in NP if and only if there is a sentence Φ in existential second-order logic such that Φ is true on a τ -structure \mathbf{A} if and only if \mathbf{A} is in \mathcal{C} . An important syntactic restriction of existential second-order logic is SNP (for *strict NP*): here we require that the sentence is of the form

$$\exists R_1, \dots, R_k. \forall x_1, \dots, x_l. \psi$$

where R_1, \dots, R_k are (existentially quantified) relation symbols, x_1, \dots, x_l are (universally quantified) first-order variables, and ψ is a quantifier-free formula. Furthermore, we say that an SNP-formula is *monadic* if all second-order variables are unary. Let MSNP be the class of all problems that can be described by a monadic SNP sentence.

Theorem 4 (of [35]). *Every problem in NP has a polynomial-time equivalent problem in MSNP.*

Theorem 4 shows that there is no complexity dichotomy for all problems in MSNP, because if $P \neq NP$ there are problems in NP that are neither in P nor NP-complete [52].

However, Feder and Vardi also showed that if we restrict our attention to (infinite domain) constraint satisfaction problems in MSNP, we obtain a class that is 'computationally equivalent' to FCSP. Let CSP be the class of all constraint satisfaction problems (with domains of arbitrary cardinality).

Theorem 5 (of [35], [50]). *Every problem in $CSP \cap MSNP$ is polynomial-time equivalent to an FCSP (and FCSP is contained in $CSP \cap MSNP$).*

Hence, a complexity dichotomy for FCSP is equivalent to a complexity dichotomy for $CSP \cap MSNP$. The dichotomy conjecture for FCSP is wide open, and not undisputed. Instead of disproving the dichotomy conjecture for FCSP, it might be easier to disprove it for CSPs in SNP that can be formulated with an ω -categorical template.

Question 2. Is every problem in NP polynomial-time equivalent to a CSP with an ω -categorical template?

The logic of SNP and MSNP are natural logics in the context of constraint satisfaction with ω -categorical templates, as witnessed by the following theorem.

Theorem 6 (of [11]). *Every CSP in MSNP can be formulated as $\text{CSP}(T)$ for an ω -categorical structure T .*

So far, we have not touched the question how infinite templates T might be finitely represented. For the definition of $\text{CSP}(T)$ we don't have to fix a representation. But for several algorithmic "meta-questions" concerning $\text{CSP}(T)$ such a representation is necessary.

Most natural CSPs with ω -categorical templates are in SNP. Thus, SNP sentences can be used to specify constraint satisfaction problems for ω -categorical templates (up to homomorphic equivalence, see Lemma 5). However, the following questions remain open.

Question 3. Is every CSP in SNP a CSP over an ω -categorical template?

7 Preservation Theorems

Primitive positive definability plays an important role in hardness proofs for the CSP. In this section, we discuss a model-theoretic preservation theorem that characterizes primitive positive definability. We then use this result to establish that the computational complexity of a CSP with a finite or ω -categorical template is fully described by an algebra that can be associated to a template.

Let \mathbf{A} and \mathbf{B} be two structures with the same signature τ . Then the (*direct*) *product* $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ of \mathbf{A} and \mathbf{B} is the τ -structure with domain $A \times B$, and for each k -ary $R \in \tau$ the structure \mathbf{C} has the relation that contains a tuple $((a_1, b_1), \dots, (a_k, b_k))$ if and only if $R(a_1, \dots, a_k)$ holds in \mathbf{A} and $R(b_1, \dots, b_k)$ holds in \mathbf{B} . For each k -ary $f \in \tau$ the structure \mathbf{C} has the operation that maps $((a_1, b_1), \dots, (a_k, b_k))$ to $(f(a_1, \dots, a_k), f(b_1, \dots, b_k))$. We write \mathbf{A}^k for $\mathbf{A} \times \mathbf{A}^{k-1}$ if $k > 1$, and \mathbf{A}^1 for \mathbf{A} .

Definition 7. *Let \mathbf{A} be a structure. Then $f : A^k \rightarrow A$ is called a polymorphism of \mathbf{A} if f is a homomorphism from \mathbf{A}^k to \mathbf{A} .*

If R is a relation over D , and f is a polymorphism of (D, R) , we also say that f *preserves* R . A relation R has a primitive positive definition in a *finite* structure if and only if R is preserved by all polymorphisms of this structure. This was discovered by Geiger [38] and by Bodnarcuk et al. [18], and is of central importance in universal algebra.

The following generalization of this theorem to ω -categorical structures was shown in [17]. We give a new proof here, which derives the theorem from the well-known homomorphism preservation theorem in model-theory. The proof given in [17] is neither long nor very complicated; however, we believe that the proof given below is also interesting, because it demonstrates how universal-algebraic preservation theorems can be derived from classical preservation theorems in model theory.

Theorem 7 (of [17]). *Let \mathbf{A} be an ω -categorical or a finite structure. A relation R has a primitive-positive definition in \mathbf{A} if and only if R is preserved by all polymorphisms of \mathbf{A} .*

The corresponding preservation theorem in model theory is as follows (even though we have not been able to find a reference for this fact, it should be considered to be known). We say that a first-order formula ϕ is *equivalent to ψ modulo* a first-order theory T if $\forall \bar{x}. \phi(\bar{x}) \leftrightarrow \psi(\bar{x})$ holds in all models of T . We also say that $\phi(x_1, \dots, x_l)$ is *preserved* by a homomorphism f from a direct product of models $\Gamma_1, \dots, \Gamma_k$ of T to a model Γ of T if whenever ϕ holds on l -tuples a_1, \dots, a_l in $\Gamma_1, \dots, \Gamma_k$, respectively, then ϕ also holds on $f((a_1^1, \dots, a_l^1), \dots, (a_1^k, \dots, a_l^k))$ in Γ .

Theorem 8. *Let T be a first-order theory. A first-order formula ϕ is equivalent to a primitive positive formula modulo T if and only if ϕ is preserved by all homomorphisms from finite direct products of models of T to models of T .*

The theorem can be shown by a slight modification of the proofs in [46]. This approach relies on several concepts and basic results in model-theory, and is out of the scope of this paper. We recommend Hodges' book [42] as an accessible introduction to model-theoretic preservation theorems.

Proof (of Theorem 7). First, observe that every relation R with a primitive positive definition in Γ is preserved by all polymorphisms of Γ . This can be shown in a straightforward way by induction on the syntactic structure of primitive positive definitions. To prove the converse, let R be a k -ary relation that is preserved by all polymorphisms of Γ . In particular, R is preserved by all automorphisms of Γ , and Theorem 1 implies that R has a first-order definition ϕ in Γ .

Let T be the first-order theory of Γ . We claim that if R is preserved by all polymorphisms of Γ then ϕ is preserved by all homomorphisms from finite direct products of models $\Gamma_1, \dots, \Gamma_l$ of T to a model Γ_0 of T . If we have shown the claim, Theorem 7 follows directly from Theorem 8.

So assume to the contrary that ϕ is not preserved by a homomorphism f from a finite direct product of models $\Gamma_1, \dots, \Gamma_l$ of T to a model Γ_0 of T (note that $\Gamma_1, \dots, \Gamma_l$, and Γ_0 might be uncountable). We prove by a standard application of the downward Löwenheim-Skolem theorem (see e.g. [42]) that ϕ is not preserved by some l -ary polymorphism of Γ , as follows. Let Δ be the disjoint union of $\Gamma_0, \Gamma_1, \dots, \Gamma_l$. Additionally, Δ has a relation S^Δ that denotes the union of the relations defined by ϕ in $\Gamma_0, \Gamma_1, \dots, \Gamma_l$, and has unary relations $P_0^\Delta, P_1^\Delta, \dots, P_l^\Delta$, where P_i^Δ denotes the elements from Γ_i , for $1 \leq i \leq l$. Finally, Δ has an l -ary function g^Δ , which is interpreted as follows. If v_1, \dots, v_l are elements from Δ , where v_i is from Γ_i for $1 \leq i \leq l$, then $g^\Delta(v_1, \dots, v_l)$ is the value of the homomorphism f from $\Gamma_1 \times \dots \times \Gamma_l$ to Γ_0 . Because ϕ is not preserved by f , there are l -tuples t_1, \dots, t_l from S^Δ such that $(f(t_1), \dots, f(t_l))$ is not in S^Δ . For all other l -tuples from Δ , the interpretation of g^Δ is set arbitrarily to some fixed element from Δ . By downward Löwenheim-Skolem, the first-order theory of Δ has a countable model Δ' . Note that $P_i^{\Delta'}$ denotes a countable, but still necessarily an infinite set. It is also easy to see that the structures $\Gamma'_1, \dots, \Gamma'_k$ induced by the sets $P_1^{\Delta'}, \dots, P_k^{\Delta'}$ in Δ' have the same first-order theory as Γ , and by ω -categoricity of Γ are isomorphic to Γ . The function $g^{\Delta'}$ maps tuples v_1, \dots, v_k where $v_i \in \Gamma'_i$ for $1 \leq i \leq k$ to Γ'_0 , and still the restriction of $g^{\Delta'}$ to those tuples violates the relation $S^{\Delta'}$. Hence, the restriction of $g^{\Delta'}$ to these tuples gives rise to a polymorphism of Γ that violates ϕ . \square

8 The Algebra of a Template

In this section we discuss how the algebraic approach to constraint satisfaction might be extended to ω -categorical templates Γ . We have already seen in the previous section that the polymorphisms of Γ characterize primitive positive definability of relations in Γ . Therefore, the algebra defined on the domain D of Γ that contains as operations all the polymorphisms of Γ fully describes the computational complexity of $\text{CSP}(\Gamma)$. The algebra that we just defined for an ω -categorical template Γ will be denoted by \mathbb{A}_Γ .

\mathbb{A}_Γ has several important properties. First of all, its set of operations forms a *clone*, i.e., is closed under compositions and contains the projections [59]. Clones with infinite domains are a subject in its own right [39]. Clones that arise as the set of polymorphisms of a structure are always *locally closed*, a concept which only becomes relevant for infinite domains. We need a couple of definitions.

Let D be an infinite set, and let $\mathcal{O}^{(k)}$ be the set of operations from D^k to D , for $k \geq 1$. The symbol \mathcal{O} denotes $\bigcup_{k=1}^{\infty} \mathcal{O}^{(k)}$. We say that an operation $f \in \mathcal{O}^{(k)}$ is *interpolated* by a set $\mathcal{F} \subseteq \mathcal{O}$ if for every finite subset A of D there is some operation $g \in \mathcal{F}$ such that $f(t) = g(t)$ for every $t \in A^k$. The set of all operations that are interpolated by \mathcal{F} is denoted by $I(\mathcal{F})$. A clone is called a *local clone* (or *locally closed*) if $I(\mathcal{F}) = \mathcal{F}$. The smallest clone that contains \mathcal{F} is the clone *generated* by \mathcal{F} , and denoted by $G(\mathcal{F})$. The smallest *local clone* that contains \mathcal{F} is called the clone *locally generated* by \mathcal{F} , and denoted by $L(\mathcal{F})$.

Proposition 2 (see e.g. [59]). *For all $\mathcal{F} \subseteq \mathcal{O}$ we have that $L(\mathcal{F}) = I(G(\mathcal{F}))$.*

We say that an algebra \mathbb{A} is called *oligomorphic* if the unary bijective operations (i.e., the permutations) in \mathbb{A} form an oligomorphic permutation group. Theorem 1 asserts that \mathbb{A}_Γ is oligomorphic if Γ is ω -categorical. Conversely, it is not hard to see that for every oligomorphic algebra \mathbb{A} whose operations form a local clone there is an ω -categorical structure Γ such that $\mathbb{A} = \mathbb{A}_\Gamma$. Such an ω -categorical structure Γ can be obtained by equipping the domain of \mathbb{A} with all relations that are preserved by all operations in \mathbb{A} ; we call this structure the *canonical structure* for \mathbb{A} .

An operation of an oligomorphic clone is called *elementary* if it is locally generated by its permutations. Clearly, for clones with a finite domain, the elementary operations are the operations that are composed of a projection with a permutation. Note that all endomorphisms of a model-complete ω -categorical core are elementary. Conversely, if all endomorphisms of an oligomorphic algebra \mathbb{A} are elementary, then the canonical structure for \mathbb{A} is a core [8].

We now define several other important properties of k -ary operations. A k -ary operation f is

- *idempotent* iff $f(x, \dots, x) = x$;
- *essentially unary* iff there is a unary operation g such that $f(x_1, \dots, x_k) = g(x_i)$ for some $i \in \{1, \dots, k\}$;
- *essential* iff f is not essentially unary;
- *commutative* iff f is binary and $f(x, y) = f(y, x)$;
- a *quasi near-unanimity operation* (short, *qnu-operation*) iff $f(x, \dots, x) = f(x, \dots, x, y) = \dots = f(x, \dots, x, y, x, \dots, x) = \dots = f(y, x, \dots, x)$;
- a *quasi Maltsev operation* iff $k = 3$ and $f(x, y, y) = f(y, y, x) = f(x, x, x)$;

An idempotent quasi near-unanimity and quasi Maltsev operation is known as near-unanimity and Maltsev operation, respectively.

9 The Pseudo-variety of a Template

A class \mathcal{V} of algebras with the same signature τ is called a *pseudo-variety* if \mathcal{V} contains all homomorphic images (Section 2), subalgebras (i.e., substructures as defined in Section 3), and finite direct products (Section 7) of algebras in \mathcal{V} . The difference between pseudo-varieties and varieties is that pseudo-varieties need not be closed under *direct products of arbitrary cardinality* (which we did not define here). The smallest pseudo-variety that contains an algebra \mathbb{A} is called the pseudo-variety *generated* by \mathbb{A} .

The results in this section link the universal-algebraic concept of pseudo-varieties with the model-theoretic concept of primitive positive interpretations. We already mentioned that first-order interpretations are a convenient tool to construct ω -categorical structures from simpler ω -categorical structures. *Primitive positive* interpretations (see Section 3) can be used to study the computational complexity of constraint satisfaction problems.

Proposition 3. *Let Δ and Γ be a structures with finite relational signatures. If there is a primitive positive interpretation of Δ in Γ , then there is a polynomial-time reduction from $\text{CSP}(\Delta)$ to $\text{CSP}(\Gamma)$.*

Proof. Let d be the dimension of the primitive positive interpretation of the τ -structure Δ in the σ -structure Γ , let $\delta(x_1, \dots, x_d)$ be the domain formula, let $h : \delta(\Gamma^d) \rightarrow D(\Delta)$ be the coordinate map, and let $\phi_R(x_1, \dots, x_{dk})$ be the formula for the k -ary relation R from Δ .

Let \mathbf{A} be an instance of $\text{CSP}(\Delta)$, and let $A = \{a_1, \dots, a_n\}$ be the elements of \mathbf{A} . We construct an instance \mathbf{B} of $\text{CSP}(\Gamma)$ as follows. The domain B of \mathbf{B} consists of dn vertices b_1^1, \dots, b_n^d . For all $1 \leq i \leq n$, we impose δ in \mathbf{B} on (b_i^1, \dots, b_i^d) : note that δ is a primitive positive σ -formula, and therefore can be simulated by a conjunction of constraints from τ , possibly with adding new vertices for the existentially quantified variables in the definition. If a tuple $(a_{i_1}, \dots, a_{i_k})$ is contained in a k -ary relation R in \mathbf{A} , then we impose $\phi_R(b_{i_1}^1, \dots, b_{i_1}^d, \dots, b_{i_k}^1, \dots, b_{i_k}^d)$ in \mathbf{B} . Clearly, the structure \mathbf{B} is an instance of $\text{CSP}(\Gamma)$ and can be constructed in polynomial time in the size of \mathbf{A} .

We claim that there is a homomorphism from \mathbf{A} to Δ if and only if there is a homomorphism from \mathbf{B} to Γ . Suppose $f : B \rightarrow D(\Gamma)$ is a homomorphism from \mathbf{B} to Γ . By construction, if $R(a_{i_1}, \dots, a_{i_k})$ holds in \mathbf{A} then $\phi_R((f(b_{i_1}^1), \dots, f(b_{i_1}^d)), \dots, (f(b_{i_k}^1), \dots, f(b_{i_k}^d)))$ holds in Γ . By the definition of interpretations, this is the case if and only if $R(h(f(b_{i_1}^1), \dots, f(b_{i_1}^d)), \dots, h(f(b_{i_k}^1), \dots, f(b_{i_k}^d)))$ holds in Δ . Hence, the mapping $g : A \rightarrow D(\Delta)$ that maps a_i to $h(f(b_i^1), \dots, f(b_i^d))$ is a homomorphism from \mathbf{A} to Δ .

Now, suppose that f is a homomorphism from \mathbf{A} to Δ . Since h is a surjective mapping from $\delta(\Gamma)^d$ to Δ , there are elements e_1^1, \dots, e_n^d in Γ such that $h(e_i^1, \dots, e_i^d) = f(a_i)$, for all $i \in \{1, \dots, n\}$. We claim that the mapping $g : B \rightarrow D(\Gamma)$ that maps b_i^j to e_i^j is a homomorphism from \mathbf{B} to Γ . By construction, any constraint in \mathbf{B} comes from a primitive positive formula $\phi_R(b_{i_1}^1, \dots, b_{i_1}^d, \dots, b_{i_k}^1, \dots, b_{i_k}^d)$ that was introduced for a constraint $R(a_{i_1}, \dots, a_{i_k})$ in \mathbf{A} . It therefore suffices to show that $\phi_R(g(b_{i_1}^1), \dots, g(b_{i_1}^d), \dots, g(b_{i_k}^1), \dots, g(b_{i_k}^d))$ holds in Γ . Since f is a homomorphism from \mathbf{A} to Δ , $R(f(a_{i_1}), \dots, f(a_{i_k}))$ holds in Δ . By the choice of e_1^1, \dots, e_n^d , this shows that $R(h(e_{i_1}^1, \dots, e_{i_1}^d), \dots, h(e_{i_k}^1, \dots, e_{i_k}^d))$ holds in Γ . By the definition of interpretations, this is the case if and only if $\phi_R(e_{i_1}^1, \dots, e_{i_1}^d, \dots, e_{i_k}^1, \dots, e_{i_k}^d)$ holds in Γ , which is what we had to show. \square

Section 11 contains a hardness proof that uses this proposition. We now present the mentioned connection between primitive positive interpretations and pseudo-varieties. We have not been able to find the following theorem explicitly in the literature even in the case of finite algebras, and therefore present its proof in full detail.

Theorem 9. *Let Γ be a finite or ω -categorical relational structure. Then a structure Δ has a primitive positive interpretation in Γ if and only if there is an algebra \mathbb{B} in the pseudo-variety generated by \mathbb{A}_Γ such that all operations of \mathbb{B} are polymorphisms of Δ .*

Proof. Let \mathcal{V} be the pseudo-variety generated by \mathbb{A}_Γ . Similarly to the famous HSP theorem for varieties, every algebra in \mathcal{V} is the homomorphic image of a subalgebra of a finite direct product of \mathbb{A}_Γ (we can use the same proof as for the HSP theorem given in [24]).

First assume that there is an algebra \mathbb{B} in \mathcal{V} all of whose operations are polymorphisms of Δ . Then there exists a finite number $d \geq 1$, a subalgebra \mathbb{C} of $(\mathbb{A}_\Gamma)^d$, and a surjective homomorphism h from \mathbb{C} to \mathbb{B} .

We claim that Δ has a first-order interpretation with dimension d in Γ . All operations of \mathbb{A}_Γ preserve C (viewed as a d -ary relation over $D(\Gamma)$), since \mathbb{C} is a subalgebra of $(\mathbb{A}_\Gamma)^d$. By Theorem 7, this implies that C has a primitive positive definition $\delta(x_1, \dots, x_d)$ in Γ , which becomes the domain formula of our interpretation. As coordinate map we choose the mapping h .

If R is a k -ary relation in Δ , let R' the dk -ary relation over $D(\Gamma)$ that contains the dk -tuple $(a_1^1, \dots, a_1^d, \dots, a_k^1, \dots, a_k^d)$ whenever R contains the tuple $(h(a_1^1, \dots, a_1^d), \dots, h(a_k^1, \dots, a_k^d))$. Let f be any operation in \mathbb{A}_Γ . By assumption, the corresponding operation f in \mathbb{B} preserves R . It is easy to verify that then the operation f in \mathbb{A}_Γ preserves R' . Hence, all polymorphisms of Γ preserve R' , and because Γ is ω -categorical, the relation R' has a primitive positive definition ϕ_R in Γ . The formula ϕ_R becomes the defining formula for R in the interpretation of Δ in Γ .

Finally, since h is an algebra homomorphism, the kernel² K of h is a congruence³ of \mathbb{C} . In other words, if $((a_1^1, \dots, a_1^d), (a_2^1, \dots, a_2^d)) \in K$ then $(f(a_1^1, \dots, a_1^d), f(a_2^1, \dots, a_2^d)) \in K$ for all operations f from \mathbb{C} , and hence $((f(a_1^1), \dots, f(a_2^1)), (f(a_2^1), \dots, f(a_2^d))) \in K$ for all operations f from \mathbb{A}_Γ . It follows that K , viewed as a $2d$ -ary relation over $D(\Gamma)$, is preserved by all operations from \mathbb{A}_Γ . Theorem 7 implies that K has a primitive positive definition in Γ . This definition becomes the formula $\phi_=_$ in the interpretation of Δ in Γ . It is straightforward to verify that we have found a primitive positive interpretation of Δ in Γ .

To prove the opposite direction, suppose that Δ has a primitive positive interpretation in an ω -categorical τ -structure Γ . We have to show that \mathcal{V} contains an algebra \mathbb{B} such that all operations in \mathbb{B} are polymorphisms of Δ . Let d be the dimension and δ be the domain formula of the interpretation. Clearly, the set $\delta(\Gamma^d)$ is preserved by all operations in \mathbb{A}_Γ , and therefore induces a subalgebra \mathbb{C} of $(\mathbb{A}_\Gamma)^d$.

We first show that the kernel of the coordinate map h of the interpretation is a congruence ρ of \mathbb{C} . For all d -tuples $\bar{a}, \bar{b} \in C$ the $2d$ -tuple (\bar{a}, \bar{b}) satisfies $\phi_=_$ in Γ if and only if $h(\bar{a}) = h(\bar{b})$. Hence, the restriction of $\phi_=_$ to C defines the kernel of h . Since $\phi_=_$ is

² The *kernel* of a map $f : A \rightarrow B$ is the equivalence relation on A that contains all pairs (x, y) such that $f(x) = f(y)$.

³ A *congruence* of an algebra \mathbb{A} is an equivalence relation that is preserved by all operations in \mathbb{A} ; equivalently, a congruence is the kernel of a homomorphic image of \mathbb{A} .

primitive positive definable in Γ , it is preserved by all polymorphisms of Γ , and therefore is a congruence of $(\mathbb{A}_\Gamma)^d$. By the second isomorphism theorem in universal algebra [24], the restriction of this congruence to C is a congruence ρ of \mathbb{C} .

Also by basic universal algebra [24] the *natural map* g that maps an element c from \mathbb{C} to its congruence class c/ρ in the quotient algebra \mathbb{C}/ρ is a surjective homomorphism from \mathbb{C} to \mathbb{C}/ρ . By the first isomorphism theorem, the elements of $\mathbb{B} := \mathbb{C}/\rho$ are in bijective correspondence to the elements of Δ , and we assume without loss of generality that \mathbb{B} and Δ have the same domain. We finally verify that every operation in \mathbb{B} is a polymorphism of Δ : It suffices to prove that every relation R of Δ is preserved by all operations f in \mathbb{B} . Let ϕ_R be the defining τ -formula of R in the primitive positive interpretation of Δ in Γ . The operation f in \mathbb{B} preserves R if and only if the operation f in \mathbb{A}_Γ preserves the relation defined by ϕ_R . Since the operations in \mathbb{A}_Γ are polymorphisms of Γ , and since ϕ_R is a primitive positive τ -formula, the operation f preserves the relation defined by ϕ_R . \square

Theorem 10. *Let Γ be ω -categorical. Then $\text{CSP}(\Gamma)$ is NP-hard if there is an expansion Δ of the model-complete core of Γ by finitely many singleton relations such that $\mathcal{V}(\mathbb{A}_\Delta)$ contains a 2-element algebra where all operations are projections.*

Proof. By Corollary 1 there is a polynomial-time reduction from $\text{CSP}(\Delta)$ to $\text{CSP}(\Gamma)$. Combining Theorem 9 with Proposition 3, we can show that $\text{CSP}(\Delta)$ is NP-hard, by reduction from the well-known NP-complete problem positive 1-in-3-3SAT, which can be formulated as $\text{CSP}(\{0, 1\}, \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\})$. \square

All known ω -categorical templates with an NP-hard CSP satisfy the condition in Theorem 10. It is therefore natural to ask whether the sufficient condition for hardness given in Theorem 10 is also necessary. For finite templates, it has been conjectured that this is true [21].

Question 4. Does Theorem 10 describe exactly the NP-hard CSPs with an ω -categorical template?

10 Tractable Templates

For finite templates Γ , all known tractable problems $\text{CSP}(\Gamma)$ can be solved by a Datalog program [35] or by an algorithm for templates with a Maltsev polymorphism [23, 43] (generalizing group-theoretic algorithms introduced in [35]), or combinations of those [20].

An ω -categorical model-complete core Γ can not have a Maltsev polymorphism, not even a quasi Maltsev polymorphism [25]. This does not imply that group-theoretic algorithms do not occur for ω -categorical templates, as we will see in Subsection 10.2.

Datalog continues to play an important role also for ω -categorical templates. All tractability results known for temporal and spatial reasoning can be formulated with Datalog programs. We will discuss in this section a result showing that every ω -categorical model-complete core with a *quasi near-unanimity operation* can be solved by a Datalog program. In fact, in this case we can use the Datalog program to compute a *globally consistent* instance. We also present examples of templates in temporal and spatial reasoning that have a quasi near-unanimity polymorphism.

For a large family of CSPs we can show tractability because the template has a certain binary bijective polymorphism. The powerful tractability criterion explains several

results in the literature on temporal and spatial reasoning and will be introduced in Subsection 10.2.

We finally present an ω -categorical template whose CSP can be solved by a simple linear-time algorithm, but which can neither be solved by Datalog nor by algebraic algorithms related to Maltsev polymorphisms.

10.1 From Local to Global Consistency

An introduction to Datalog for constraint satisfaction with finite templates can be found in [Cite Bulatov,Krokhin,Larose in this book]. Datalog programs can also be used to solve CSPs with infinite templates in polynomial time: recall that every Datalog program derives on a given finite input structure only a polynomially bounded number of facts. Thus, any Datalog program can be evaluated in polynomial time. An example of a CSP with an infinite template that can be solved by a $(2, 3)$ -Datalog program is $\text{CSP}(\mathbb{Q}, \neq, \leq)$ [60].

Datalog is particularly useful when the template is ω -categorical, since in this case there exist *canonical Datalog programs*, as in the case of finite templates.

Definition 8. *Let Γ be an ω -categorical structure. The canonical (l, k) -Datalog program for $\text{CSP}(\Gamma)$ is a Datalog program Π that contains an IDB for every at most l -ary primitive positive definable relation in Γ (by ω -categoricity of Γ , there are only finitely many such relations). The empty 0-ary relation will be denoted by **false**. The EDBs are precisely the relation symbols in τ . The program Π contains a rule $\psi :- \psi_1, \dots, \psi_j$ if the implication $\psi_1 \wedge \dots \wedge \psi_j \Rightarrow \psi$ contains at most k variables, is valid in Γ , and ψ is of the form $R(y_1, \dots, y_s)$ for an IDB R and $s \leq l$.*

The final stage in the evaluation of a canonical Datalog program on a given instance gives rise to another instance \mathbf{A} of $\text{CSP}(\Gamma')$, where Γ' is the expansion of Γ by all at most l -ary primitive positive definable relations, and where \mathbf{A} contains all the derived tuples from these relations [32].

Proposition 4 (of [11]). *A constraint satisfaction problem $\text{CSP}(\Gamma)$ with an ω -categorical template Γ can be solved with an (l, k) -Datalog program if and only if the canonical (l, k) -Datalog program solves $\text{CSP}(\Gamma)$.*

Global consistency. Some templates Γ have the strong property that the instance computed by the canonical Datalog program is always *globally consistent*, see Definition 9 below.

Definition 9. *Let Γ be a structure with finite relational signature. An instance \mathbf{A} of $\text{CSP}(\Gamma)$ is called strongly k -consistent if for every subset $S = \{v_1, \dots, v_l\}$ of A with $l \leq k$ and every homomorphism h from $\mathbf{A}[\{v_1, \dots, v_{l-1}\}]$ to Γ there exists an extension of h that is a homomorphism from $\mathbf{A}[S]$ to Γ . An instance \mathbf{A} of $\text{CSP}(\Gamma)$ is called globally consistent if it is k -consistent for all $1 \leq k \leq |A|$.*

Note that if Γ has the property that every strongly k -consistent instance of $\text{CSP}(\Gamma)$ is globally consistent, then $\text{CSP}(\Gamma)$ can be solved in polynomial time. As an example, consider again $\text{CSP}(\mathbb{Q}, \leq, \neq)$. It is known that the canonical $(4, 5)$ -Datalog program computes globally consistent instances, but the canonical $(3, 4)$ -Datalog program does not [48]. We will see a universal-algebraic explanation of this fact in the next paragraph.

Quasi near-unanimity functions. We present a universal-algebraic characterization of those ω -categorical model-complete cores where the canonical (l, k) -Datalog program computes globally consistent instances. Recall the definition of *(quasi) near-unanimity functions* given in Section 8. As an example of a near-unanimity function, consider the structure $(\mathbb{Q}, \leq, <)$, and the operation *median*, which is the ternary function that returns the median of its three arguments. More precisely, for three elements x, y, z from \mathbb{Q} , suppose that $\{x, y, z\} = \{a, b, c\}$, where $a \leq b \leq c$. Then $\text{median}(x, y, z)$ is defined to have value b . It is easy to verify that *median* is a ternary near-unanimity function, and that it is a polymorphism of $(\mathbb{Q}, \leq, <)$.

The structure (\mathbb{N}, \neq) is an example of a structure that has no near-unanimity polymorphism (exercise!), but has a ternary quasi near-unanimity operation, defined as follows. Consider the structure $(\mathbb{N}, \neq)^3$, where for all $x, y \in \mathbb{N}$ the triples of the form (y, x, x) , (x, y, x) , (x, x, y) , and (x, x, x) are identified. The resulting structure homomorphically maps to (\mathbb{N}, \neq) (Lemma 2), and this homomorphism is a ternary polymorphism of (\mathbb{N}, \neq) and a quasi near-unanimity operation.

Similarly, we can construct a 5-ary quasi near-unanimity polymorphism of (\mathbb{N}, R) where $R = \{(x, y, u, v) \mid x \neq y \vee u \neq v\}$. To see that this structure has no 4-ary quasi near-unanimity, consider the tuples $t_1 = (1, 0, 0, 0)$, $t_2 = (0, 1, 0, 0)$, $t_3 = (0, 0, 1, 0)$, and $t_4 = (0, 0, 0, 1)$. Any 4-ary quasi near-unanimity operation applied to (t_1, t_2, t_3, t_4) yields a tuple (c, c, c, c) for some $c \in \mathbb{N}$. But note that t_1, \dots, t_4 are in R , whereas c is not.

Theorem 11 (of [11]). *An ω -categorical model-complete core Γ has a k -ary quasi near-unanimity polymorphism if and only if every strongly k -consistent instance of $\text{CSP}(\Gamma)$ is globally consistent.*

Examples. A 5-ary quasi near-unanimity polymorphism of (\mathbb{Q}, \leq, \neq) was described in [9]. We give a different construction of such a polymorphism here. Let $f : \mathbb{Q}^{10} \rightarrow \mathbb{Q}$ be an injective operation that preserves \leq . Such an operation exists; we can simply take any operation such that $f(x_1, \dots, x_{10}) < f(y_1, \dots, y_{10})$ if and only if (x_1, \dots, x_{10}) is lexicographically smaller than (y_1, \dots, y_{10}) .

Proposition 5. *Let $h(x, y, z, u, v)$ be the 5-ary operation*

$$f(\text{median}(x, y, z), \text{median}(x, y, u), \text{median}(y, z, v), \dots, \\ \text{median}(y, u, v), \text{median}(z, u, v))$$

(That is, we apply the median to all of the 10 three-element subsets of $\{x, y, z, u, v\}$.) Then h is a 5-ary quasi near-unanimity polymorphism of (\mathbb{Q}, \leq, \neq) .

It can also be shown that the template with the basic relations of the spatial reasoning formalisms RCC-5 (see Section 5) has a 5-ary quasi near-unanimity polymorphism [9].

We would like to remark that the concept of quasi near-unanimity operations is a natural concept also for finite domain constraint satisfaction problem. For example, it allows to formulate the well-known dichotomy of $\text{CSP}(H)$ for undirected graphs H as follows: *$\text{CSP}(H)$ is in P if and only if H has a quasi near-unanimity operation* (unless $P=NP$). This is easy to derive from the well-known theorem that $\text{CSP}(H)$ is NP-hard if H is not bipartite [40].

10.2 Horn Tractability

In this section we describe a powerful tractability criterion, which in particular explains many results in temporal and spatial reasoning. But also CSPs with a group-theoretic algorithm will be treated here.

We say that a relation R has a *quantifier-free Horn definition* in a τ -structure Γ if R can be defined by a quantifier-free τ -formula in conjunctive normal form in which each clause contains at most one positive literal. If \mathbf{A} is a structure, we denote by $\hat{\mathbf{A}}$ the expansion⁴ of \mathbf{A} that also contains the complement for each relation in \mathbf{A} .

Theorem 12 (of [10]). *Let Γ be an ω -categorical homogeneous structure, and let Δ be a structure with a first-order definition in Γ . If Δ has a polymorphism i which is a strong homomorphism from Γ^2 to Γ , then Δ has a quantifier-free Horn definition in Γ . Moreover, if $\text{CSP}(\hat{\Gamma})$ is tractable, then $\text{CSP}(\Delta)$ is tractable as well.*

Note that if i is bijective (in other words, if i is an isomorphism between Γ^2 and Γ) then the binary polymorphism of Δ satisfies the equation

$$i(x, y) = a(i(y, x))$$

for some automorphism a of Δ [10]. In this case, every expansion of Δ by some relation with a first-order definition in Γ that is not equivalent to a quantifier-free Horn formula over Γ has an NP-hard CSP [10]; in this sense, the tractability result is strongest possible. We illustrate the use of this theorem by examples.

Solving Equations over Infinite Vector Spaces. We come back to the application on solving systems of equations over boolean vectors, described in Subsection 5.3. To recall, the template for this CSP had a first-order definition in the infinite-dimensional vector space \mathbb{V} over the 2-element field \mathbb{F}_2 . It is well-known that \mathbb{V}^2 , the direct product of \mathbb{V} with itself, is isomorphic to \mathbb{V} ; similarly, we have that \mathbf{V}^2 (as defined in Subsection 5.3) is isomorphic to \mathbf{V} . Let i be the isomorphism. Note that $\hat{\mathbf{V}} = \mathbf{V}$, since \mathbf{V} already contains the complements of all its relations. It was shown in [10] that $\text{CSP}(\mathbf{V})$ can be solved in polynomial time.

Finally, let Δ be any relational structure where all relations have a first-order definition in Γ which is a quantifier-free Horn formula. It is straightforward to verify that such relations are preserved by i . Now we can apply Theorem 12, and obtain the result that $\text{CSP}(\Delta)$ can be solved in polynomial time as well.

Spatial Reasoning. Theorem 12 can also be applied to study tractable CSPs in spatial reasoning. The countable atomless boolean ring without identity element \mathbb{B} , as introduced in Subsection 5.2, is isomorphic to \mathbb{B}^2 (because \mathbb{B}^2 is again an atomless boolean ring without identity element, and all such boolean rings are isomorphic [1]); let $i : \mathbb{B}^2 \rightarrow \mathbb{B}$ denote this isomorphism. The constraint satisfaction problem for $(\mathbb{B}, \text{DR}, \text{PP}, \neg\text{DR}, \neg\text{PP})$ is in P [58] (in fact, this structure has a 5-ary qunuf, as shown in [9]). Theorem 12 then implies that $\text{CSP}(\Gamma)$ is in P if all relations of Γ have a first-order definition in $(\mathbb{B}, \text{DR}, \text{PP})$ and are preserved by i . These are precisely the relations that have a first-order Horn

⁴ An *expansion* of a structure \mathbf{A} is a structure that is obtained from \mathbf{A} by adding relations and/or functions to \mathbf{A} .

definition in (B, DR, PP) . In particular, we obtain a result by Renz and Nebel [58] who determined a largest tractable fragment of RCC-5. It follows from the result in [44] that this fragment is the unique largest tractable fragment that contains the basic relations DR and PP. However, note that Theorem 12 also applies to relational structures Δ with relations of arbitrary arity.

10.3 Other Tractable Templates

All known tractable problems of the form $CSP(\Gamma)$ for templates Γ with a finite domain can be solved by Datalog, or by an algebraic algorithm, or combinations and extensions of these approaches. For infinite templates Γ , there are new kinds of algorithms to solve $CSP(\Gamma)$ in polynomial time. We present a simple example of such an algorithm, and give some details here, because the algorithm is a prototype for several more advanced algorithms [4, 16].

In Section 5, we have already mentioned the problem $CSP(\mathbb{Q}, R^{\min})$, which we call the *min-ordering problem*: in this problem we are given a set A of variables, and a set R^{\min} of triples on these variables. We want to find a mapping $\alpha : A \rightarrow \mathbb{Q}$ such that for each triple (x, y, z) either $\alpha(x) > \alpha(y)$ or $\alpha(x) > \alpha(z)$. It can be shown [14] that there is no Datalog program that solves this problem. However, here is a simple linear-time algorithm for $CSP(\mathbb{Q}, R^{\min})$.

The algorithm we present relies on the fact that R^{\min} has the binary operation *min* as a polymorphism, where $\min(x, y)$ is by definition the minimum of x and y . The operation *min* is an example of a *semilattice operation*, which is an associative, commutative, and idempotent operation. For finite structures it is known that if a template Γ has a semilattice polymorphism, then $CSP(\Gamma)$ is tractable by a Datalog program. This result does not carry over to ω -categorical templates (we have already mentioned that (\mathbb{Q}, R^{\min}) has a semilattice polymorphism, but cannot be solved by a Datalog program).

Let \mathbf{A} be an instance of the min-ordering problem that has a solution α . It will be easy to generalize our algorithm to all templates where all relations have a first-order definition in $(\mathbb{Q}, <)$ and are preserved by *min*. The algorithm was found in the course of the classification of the complexity of $CSP(\Gamma)$ for all Γ with a first-order definition in $(\mathbb{Q}, <)$, and is one of the simplest cases in this classification [13, 14].

Definition 10. *A set S of variables in an instance \mathbf{A} of the min-ordering problem is called free if the instance has a solution α where $\alpha(x) \leq \alpha(y)$ for all $y \in A$ and $x \in S$.*

If x is from a free set of variables S , then R^{\min} cannot contain a triple (x, y, z) where x appears in the first entry, because then either $\alpha(y)$ or $\alpha(z)$ are strictly smaller than $\alpha(x)$, a contradiction. We say that a variable $x \in A$ is *blocked* if R^{\min} contains a tuple (x, y, z) where x appears in the first entry (and *unblocked* otherwise).

Lemma 6. *If all variables in an instance of the min-ordering problem are blocked, then the instance is unsatisfiable.*

Proof. Suppose the instance has a solution, then there must exist a free set of variables. But as we have seen, free variables cannot be blocked. \square

It turns out that if an instance of the min-ordering problem is satisfiable, then the set of *all* min-candidates is free. This follows from the correctness of the following algorithm.

```

Min-Ordering( $\mathbf{A}$ )
Input: A structure  $(A, R^{\min})$  with a ternary relation  $R^{\min}$ .

If  $A = \emptyset$  then accept
else
  Compute the set  $S$  of min-candidates;
  If  $S = \emptyset$  then reject;
  else Min-Ordering( $\mathbf{A}[A \setminus S]$ )
end if

```

Fig. 1. The Min-Ordering algorithm.

Proposition 6. *The Min-Ordering procedure given in Figure 1 correctly decides the min-ordering problem in linear time in the input size.*

Proof. Let \mathbf{A} be an instance of the min-ordering problem. If at some point during the execution of the Min-Ordering procedure on \mathbf{A} we recursively apply the min-ordering procedure to a substructure \mathbf{B} of \mathbf{A} and do not find an unblocked variable, then Lemma 6 implies that \mathbf{B} and therefore also \mathbf{A} does not have a solution.

Otherwise, it is clear that \mathbf{A} has a solution if the set of variables is empty. So, suppose the algorithm finds a non-empty set S . Inductively assume that $\mathbf{A}[A \setminus S]$ has a solution α . Then the extension α' of α that maps all variables in S to a value smaller than the value of $\alpha(y)$ for all other variables $y \in A \setminus S$ is clearly a solution to \mathbf{A} . Hence, the algorithm is correct. It is not difficult to see that it can be implemented such that it has a linear worst-case running time. \square

A similar, but more complicated algorithm based on finding a free set of variables can be used to solve the branching-time consistency problem (and therefore also the rooted triple consistency problem, as we have noticed before) in quadratic time [16].

11 Classification for Equality Constraint Languages

In this section, we show how to use the concepts we have seen so far to describe a full classification for a very restricted, but important class of ω -categorical templates. We study countably infinite relational structures that are first-order definable in $(\mathbb{N}, =)$. Equivalently, we study relational structures where all relations have a definition by a boolean combination of atomic formulas of the form $x = y$. Such templates are called *equality constraint languages* [12]. Example 1 in Section 2 is an example of an equality constraint language. Another example is (\mathbb{N}, S) where

$$S = \{(x, y, z) \in \mathbb{N}^3 \mid x = y = z \vee x \neq y \neq z \neq x\}.$$

It is not hard to see that equality constraint languages are precisely those templates that are preserved by *all* permutations of the domain.

Theorem 13 (of [12]). *Let Γ be an equality constraint language. If Γ is preserved by a unary constant operation, or by a binary injective operation, then $\text{CSP}(\Gamma)$ is in P .*

Otherwise, all polymorphisms of Γ are essentially unary and preserve \neq , and $\text{CSP}(\Gamma)$ is NP-complete.

We can answer Question 4 positively in case that Γ is an equality constraint language. In fact, Theorem 13 has the following reformulation using pseudo-varieties and equations.

Theorem 14. *Let Γ be an equality constraint language. The either $\mathcal{V}(\mathbb{A}_\Gamma)$ contains a 2-element algebra where all operations are projections, and $\text{CSP}(\Gamma)$ is NP-complete, or \mathbb{A}_Γ contains operations a, f satisfying $f(x, y) = a(f(y, x))$.*

Proof. If \mathbb{A}_Γ contains operations a, f satisfying $f(x, y) = a(f(y, x))$ then this equation holds for all members of $\mathcal{V}(\mathbb{A}_\Gamma)$, and since the projections do not satisfy this equation, the two cases mentioned in the theorem are indeed disjoint.

Now, suppose that all polymorphisms of \mathbb{A}_Γ are essentially unary and preserve \neq . Let $f : D(\Gamma)^2 \rightarrow \{0, 1\}$ be such that (x, y) is mapped to 0 for $x = y$ and to 1 for $x \neq y$. It is straightforward to verify that f is a surjective algebra homomorphism from \mathbb{A}_Γ^2 to an algebra \mathbb{B} where all operations are projections. Theorem 10 implies that $\text{CSP}(\Gamma)$ is NP-complete.

If \mathbb{A}_Γ contains an operation that does not preserve \neq , or is not essentially unary, then Theorem 13 shows that \mathbb{A}_Γ contains a constant operation or a binary injective operation. It is easy to see that in both cases there is a permutation a and a binary operation f in \mathbb{A}_Γ such that $f(x, y) = a(f(y, x))$. \square

12 Outlook

In the opinion of the author, there are three important directions of future research on CSPs with infinite templates.

1. **Apply the universal-algebraic approach to unify and generalize existing CSP complexity results in the application areas.** Similarly as for the mentioned classification for equality constraint languages, it would be interesting to have a complete classification for templates for reasoning over time points, time intervals (i.e., the generalization of Allen's interval algebra to constraint languages with arbitrary, and not only binary constraints), for reasoning over partially ordered time, branching time, and spatial reasoning with RCC-5, and for other important qualitative reasoning formalisms that have been studied in the literature. I am optimistic that with the universal-algebraic approach a classification in the style of Theorem 13 can be obtained in all these areas. In particular, I expect that in this way we discover new and interesting tractable languages.
2. **Further develop the universal-algebraic theory for ω -categorical templates.**
 - (a) When is it useful to work with varieties instead of pseudo-varieties? For finite templates Γ , the pseudo-variety and the variety of the Γ contain the same finite algebras. Birkhoff's theorem thus asserts that the computational complexity of $\text{CSP}(\Gamma)$ is captured by the set of equations satisfied by the operations in \mathbb{A}_Γ . Indeed, also for ω -categorical templates several tractability criteria can be formulated by equations satisfied by the polymorphisms of the template (see Section 10).

- (b) Develop tame congruence theory for oligomorphic algebras \mathbb{A} . An interesting approach might be to apply classical tame congruence theory to the finite algebras in the pseudo-variety generated by \mathbb{A} , and to study the implications for the algebra \mathbb{A} (recall the examples in Section 11).
 - (c) Suppose we have classified the computational complexity for all templates with a first-order definition in another ω -categorical structure Γ . Can we lift this classification to a classification for all structures with a first-order interpretation in Γ ?
 - (d) Follow the successful lines in model-theory where ω -categorical structures were classified under certain additional assumptions, for example strict minimality [26, 61], finite homogeneity or stability [28, 51], or smooth approximability [27, 45]. These classifications are usually up to inter-definability, sometimes even bi-interpretability, so even if a classification is complete in the model-theoretic sense, quite some effort might be necessary to classify the complexity of the corresponding CSPs. Finally, we would like to mention that many of the ω -categorical templates in Artificial Intelligence have an automorphism group that is a *Jordan group* [25, 53], for which strong classification results are known [2, 3].
3. **Clarify which constraint satisfaction problems in NP can be formulated with an ω -categorical template.** We have seen an example that does not have such a formulation in Section 5. But we have also seen that all CSPs in MSNP do have a formulation with an ω -categorical template. Is the same true for all problems in SNP? Which constraint satisfaction problems in NP are in SNP?

Acknowledgements. I would like to thank the referee for his critical and helpful comments, all my CSP co-authors for the cooperation: Hubie Chen, Victor Dalmau, Jan Kára, Martin Kutz, Jaroslav Nešetřil, Michael Pinsker, Timo von Oertzen.

References

1. A. Abian. Categoricity of denumerable atomless boolean rings. *Studia Logica*, 30(1):63–67, 1972.
2. S. Adeleke and P. M. Neumann. Structure of partially ordered sets with transitive automorphism groups. *AMS Memoir*, 57(334), 1985.
3. S. A. Adeleke and D. Macpherson. Classification of infinite primitive jordan permutation groups. *Proceedings of the London Mathematical Society*, s3-72(1):63–123, 1996.
4. A. Aho, Y. Sagiv, T. Szymanski, and J. Ullman. Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions. *SIAM Journal on Computing*, 10(3):405–421, 1981.
5. J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.
6. B. L. Bauslaugh. The complexity of infinite h-coloring. *J. Comb. Theory, Ser. B*, 61(2):141–154, 1994.
7. M. Bodirsky. Cores of countably categorical structures. *Logical Methods in Computer Science (LMCS)*, 2007. DOI: 10.2168/LMCS-3(1:2).
8. M. Bodirsky and H. Chen. Oligomorphic clones. *Algebra Universalis*, 57(1):109–125, 2007.
9. M. Bodirsky and H. Chen. Qualitative temporal and spatial reasoning revisited. In *16th EACSL Annual Conference on Computer Science and Logic (CSL’07)*, 2007.
10. M. Bodirsky, H. Chen, J. Kara, and T. von Oertzen. Maximal infinite-valued constraint languages. *Submitted. A preliminary version appeared at ICALP’07*, 2007.

11. M. Bodirsky and V. Dalmau. Datalog and constraint satisfaction with infinite templates. In *Proceedings of the 23rd International Symposium on Theoretical Aspects of Computer Science (STACS'06)*, LNCS 3884, pages 646–659. Springer Verlag, 2006.
12. M. Bodirsky and J. Kára. The complexity of equality constraint languages. *Accepted for publication in Theory of Computing Systems*, 2007. A conference version of the paper appeared in the proceedings of the International Computer Science Symposium in Russia (CSR'06).
13. M. Bodirsky and J. Kára. The complexity of temporal constraint satisfaction problems. Preprint, 2007.
14. M. Bodirsky and J. Kára. A fast algorithm and lower bound for temporal reasoning. Preprint, 2007.
15. M. Bodirsky and M. Kutz. Pure dominance constraints. In *Proceedings of STACS'02*, pages 287–298, 2002.
16. M. Bodirsky and M. Kutz. Determining the consistency of partial tree descriptions. *Artificial Intelligence*, 171:185–196, 2007.
17. M. Bodirsky and J. Nešetřil. Constraint satisfaction with countable homogeneous templates. In *Proceedings of CSL'03*, pages 44–57, Vienna, 2003.
18. V. G. Bodnarčuk, L. A. Kalužnin, V. N. Kotov, and B. A. Romov. Galois theory for post algebras, part I and II. *Cybernetics*, 5:243–539, 1969.
19. M. Broxvall and P. Jonsson. Point algebras for temporal reasoning: Algorithms and complexity. *Artif. Intell.*, 149(2):179–220, 2003.
20. A. Bulatov. A graph of a relational structure and constraint satisfaction problems. In *In Proceedings of the 19th IEEE Annual Symposium on Logic in Computer Science (LICS'04)*, Turku, Finland, 2004.
21. A. Bulatov and P. Jeavons. Algebraic structures in combinatorial problems. Technical report MATH-AL-4-2001, Technische Universität Dresden, submitted to International Journal of Algebra and Computing, 2001.
22. A. Bulatov, A. Krokhin, and P. G. Jeavons. Classifying the complexity of constraints using finite algebras. *SIAM Journal on Computing*, 34:720–742, 2005.
23. A. A. Bulatov and V. Dalmau. A simple algorithm for Mal'tsev constraints. *SIAM J. Comput.*, 36(1):16–27, 2006.
24. S. Burris and H. Sankappanavar. *A Course in Universal Algebra*. Springer Verlag, Berlin, 1981.
25. P. J. Cameron. *Oligomorphic Permutation Groups*. Cambridge Univ. Press, 1990.
26. G. Cherlin, L. Harrington, and A. Lachlan. \aleph_0 -categorical, \aleph_0 -stable structures. *Annals of Pure and Applied Logic*, 28:103–135, 1985.
27. G. Cherlin and E. Hrushovski. *Finite Structures with Few Types*. Princeton University Press, 2003.
28. G. Cherlin and A. H. Lachlan. Stable finitely homogeneous structures. *TAMS*, 296:815–850, 1986.
29. D. Cohen, P. Jeavons, P. Jonsson, and M. Koubarakis. Building tractable disjunctive constraints. *Journal of the ACM*, 47(5):826–853, 2000.
30. M. Droste. Structure of partially ordered sets with transitive automorphism groups. *AMS Memoir*, 57(334), 1985.
31. I. Düntsch. Relation algebras and their application in temporal and spatial reasoning. *Artificial Intelligence Review*, 23:315–357, 2005.
32. H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer, 1999. 2nd edition.
33. D. Evans. Examples of \aleph_0 -categorical structures. In 'Automorphisms of first-order structures', R. Kaye and H.D. Macpherson, Oxford University Press, pages 33–72, 1994.
34. T. Feder, P. Hell, and B. Mohar. Acyclic homomorphisms and circular colorings of digraphs. *SIAM J. Discrete Math.*, 17(1):161–169, 2003.
35. T. Feder and M. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory. *SIAM Journal on Computing*, 28:57–104, 1999.

36. R. Fraïssé. *Theory of Relations*. North-Holland, 1986.
37. M. Garey and D. Johnson. *A guide to NP-completeness*. CSLI Press, 1978.
38. D. Geiger. Closed systems of functions and predicates. *Pacific Journal of Mathematics*, 27:95–100, 1968.
39. M. Goldstern and M. Pinsker. A survey of clones on infinite sets. *Preprint, arXiv:0701030*, 2007.
40. P. Hell and J. Nešetřil. On the complexity of H-coloring. *Journal of Combinatorial Theory, Series B*, 48:92–110, 1990.
41. R. Hirsch. Expressive power and complexity in algebraic logic. *Journal of Logic and Computation*, 7(3):309 – 351, 1997.
42. W. Hodges. *A shorter model theory*. Cambridge University Press, 1997.
43. P. M. Idziak, P. Markovic, R. McKenzie, M. Valeriote, and R. Willard. Tractability and learnability arising from algebras with few subpowers. In *LICS'07*, pages 213–224, 2007.
44. P. Jonsson and T. Drakengren. A complete classification of tractability in RCC-5. *J. Artif. Intell. Res.*, 6:211–221, 1997.
45. W. M. Kantor and H. D. M. M. W. Liebeck. \aleph_0 -categorical structures smoothly approximated by finite substructures. *Proc. London Math. Soc.*, 59:439–463, 1989.
46. J. Keisler. Reduced products and Horn classes. *Trans. Amer. Math. Soc.*, 117:307–328, 1965.
47. O. Klíma, P. Tesson, and D. Thérien. Dichotomies in the complexity of solving systems of equations over finite semigroups. *Theory Comput. Syst.*, 40(3):263–297, 2007.
48. M. Koubarakis. Tractable disjunctions of linear constraints: Basic results and applications to temporal reasoning. *Theoretical Computer Science*, 266:311–339, 2001.
49. A. A. Krokhin, P. Jeavons, and P. Jonsson. Reasoning about temporal relations: The tractable subalgebras of Allen’s interval algebra. *JACM*, 50(5):591–640, 2003.
50. G. Kun. Constraints, mmsnp, and expander relational structures. *arXiv:0706.1701*, 2007.
51. A. H. Lachlan. Stable finitely homogeneous structures: A survey. In *Algebraic Model Theory, NATO ASI Series*, volume 496, pages 145–159, 1996.
52. R. E. Ladner. On the structure of polynomial time reducibility. *JACM*, 22(1):155–171, 1975.
53. D. Macpherson. A survey of jordan groups. In *‘Automorphisms of first-order structures’, R. Kaye and H.D. Macpherson, Oxford University Press*, pages 73–110, 1994.
54. Y. Matiyasevich. Enumerable sets are diophantine. *Doklady Akademii Nauk SSSR*, 191:279–282, 1970.
55. B. Nebel and H.-J. Bürckert. Reasoning about temporal relations: A maximal tractable subclass of Allen’s interval algebra. *JACM*, 42(1):43–66, 1995.
56. M. S. Paterson and M. N. Wegman. Linear unification. *Journal of Computer and System Sciences*, 16:158–167, 1978.
57. B. Poizat. *A Course in Model Theory: An Introduction to Contemporary Mathematical Logic*. Springer, 2000.
58. J. Renz and B. Nebel. On the complexity of qualitative spatial reasoning: A maximal tractable fragment of the region connection calculus. *Artif. Intell.*, 108(1-2):69–123, 1999.
59. A. Szendrei. *Clones in universal Algebra*. Seminaire de mathematiques superieures. Les Presses de L’Universite de Montreal, 1986.
60. P. van Beek and R. Cohen. Exact and approximate reasoning about temporal relations. *Computational Intelligence*, 6:132–144, 1990.
61. B. Zilber. *Uncountable categorical theories, Tranlations of Mathematical Monographs*, volume 117. Amer. Math. Soc., 1993.