Constraint Satisfaction with Countable Homogeneous Templates

Manuel Bodirsky[1] and Jaroslav Nešetřil[2]

[1] Humboldt Universität zu Berlin, Germany
bodirsky@informatik.hu-berlin.de
[2] Institute for Theoretical Computer Science (ITI)
and Department of Applied Mathematics (KAM),
Charles University, Prague, Czech Republic
nesetril@kam.mff.cuni.cz

Abstract. For a fixed countable homogeneous structure Γ we study the computational problem whether a given finite structure of the same relational signature homomorphically maps to Γ. This problem is known as the constraint satisfaction problem CSP(Γ) for Γ and was intensively studied for finite Γ. We show that - as in the case of finite Γ - the computational complexity of CSP(Γ) for countable homogeneous Γ is determinded by the clone of polymorphisms of Γ. To this end we prove the following theorem which is of independent interest: The primitive positive definable relations over an ω-categorical structure Γ are precisely the relations that are invariant under the polymorphisms of Γ.
Constraint satisfaction with countable homogeneous templates is a proper generalization of constraint satisfaction with finite templates. If the age of Γ is finitely axiomatizable, then CSP(Γ) is in NP. If Γ is a digraph we can use the classification of homogeneous digraphs by Cherlin to determine the complexity of CSP(Γ).

1 Introduction

For a fixed relational structure Γ (called the *template*), the constraint satisfaction problem CSP(Γ) is the following computational problem: Given

[1] Member of the European Graduate Program "Combinatorics, Geometry and Computation", supported by DFG grant GRK 588/2.
[2] Supported by grant LN00A056 of the Czech Ministry of Education.
Both authors supported by the COMBSTRU network of the European Union.

a finite structure S of the same signature as Γ, is there a homomorphism from S to Γ?

Constraint satisfaction problems frequently occur in theoretical computer science, and have attracted much attention for finite templates Γ. It is conjectured that CSP(Γ) has a *dichotomy* in the sense that every constraint satisfaction problem CSP(Γ) for finite structure Γ is either tractable or NP-complete. This is true for templates that are undirected graphs [20], for two element templates [35] or three element templates [6]. It is known that every constraint satisfaction problem is polynomial time equivalent to a digraph-homomorphism problem [17]. There are powerful and well-understood classes of algorithms solving the known tractable constraint satisfaction problems [17], namely group theoretic algorithms and local-consistency based algorithms [14, 21].

But many constraint satisfaction problems in the literature can not be formulated as a constraint satisfaction problem with a finite template. One example is Allen's interval algebra [1] that has applications in temporal reasoning in artificial intelligence. The classification of the tractable and hard subalgebras of Allen's algebra was completed only recently [8, 29], and they also exhibit a complexity dichotomy. Other examples are tree description languages that were introduced in computational linguistics [3, 4, 11]. Even digraph-acyclicity can not be formulated as a constraint satisfaction problem with finite template Γ. However, arbitrary infinite templates Γ might have undecidable constraint satisfaction problems.

We propose to study constraint satisfaction with *countable homogeneous templates*. This can be seen as a strict generalization of constraint satisfaction with finite templates, since every constraint satisfaction problem with a finite template is polynomial-time equivalent to a constraint satisfaction problem with a homogeneous template (see Section 3). Moreover, the constraint satisfaction problems mentioned above can be formulated naturally in this new framework. To prove tractability or hardness of constraint satisfaction problems with homogeneous templates reductions to different hard problems and new algorithms are used, which have not yet been considered for CSP(Γ) with finite Γ. Countable homogeneous structures are intensively studied by model theorists, and they have many remarkable properties. For finite signatures they allow quantifier elimination and are ω-categorical, i.e. their first-order theories have only one countable model up to isomorphism. Countable homogeneous structures have been classified for all digraphs [10].

Adding relations to a template Γ that are *primitive positive* definable over Γ does not change the computational complexity of CSP(Γ). The central theorem [5] here is that a relation is primitive positive definable over a finite relational structure Γ if and only if it is left invariant under the polymorphisms of Γ. This was first used in the context of constraint satisfaction by Jeavons et al. [23], and initiated the algebraic approach to constraint satisfaction, which has successfully been carried further e.g. in [13]. We will generalize this result to ω-categorical structures Γ: A relation is p.p.-definable in Γ if and only if it is invariant under the polymorphisms of Γ.

We can determine the complexity of $\mathrm{CSP}(\Gamma)$ and prove a dichotomy if Γ is a homogeneous graph. Since there are uncountably many countable homogeneous digraphs Γ and uncountably many corresponding constraint satisfaction problems, the class of problems $\mathrm{CSP}(\Gamma)$ contains undecidable problems. However, if we assume that the class of finite induced substructures of a countable homogeneous digraph Γ is finitely axiomatized, we can determine the complexity of $\mathrm{CSP}(\Gamma)$ with only three exceptions ($S(2)$, $S(3)$, and $\mathcal{P}(3)$; see [10]; the discussion of their constraint satisfaction problems will appear elsewhere).

The paper is organized as follows. We first give some background on relational homogeneous structures. In the next section on combinatorial constraint satisfaction problems we explain the rôle of *primitive positive definability* in constraint satisfaction. We give a characterization of primitive positive definability on homogeneous structures in Section 5 after introducing the necessary tools from universal algebra in Section 4. We end with a catalog of homogeneous relational structures and a discussion of their constraint satisfaction problems.

2 Background

A *relational signature* τ is a (in this paper always at most countable) set of *relation symbols* R_i, each associated with an *arity* k_i. A *(relational) structure* Γ *over relational signature* τ (also called τ*-structure*) is a set D_Γ (the *domain*) together with a relation $R_i \subseteq D_\Gamma^{k_i}$ for each relation symbol of arity k_i. For simplicity we denote both a relation symbol and its corresponding relation with the same symbol. For a τ-structure Γ and $R \in \tau$ it will also be convenient to say that $R(u_1, \ldots, u_k)$ *holds in* Γ if $(u_1, \ldots, u_k) \in R$. We sometimes use the shortened notation $\overline{x}$ for a vector $x_1, \ldots, x_n$ of any length.

A first-order formula φ over the signature τ is said to be *primitive positive* (we say φ is a *p.p.-formula*, for short) if it is of the form

$$\exists \overline{x}(\varphi_1(\overline{x}) \wedge \cdots \wedge \varphi_k(\overline{x})) \ .$$

where $\varphi_1, \ldots, \varphi_k$ are atomic formulas. (For an introduction to first order logic and model theory see [22].) Let Γ be a relational structure of signature τ. Then a p.p.-formula φ over τ with k free variables defines a k-ary relation $R \subseteq D_\Gamma^k$: the relation R is the set of all tuples satisfying the formula φ in Γ. Equivalently, R is contained in $\langle \Gamma \rangle_{pp}$ if and only if there exists a finite relational τ-structure S containing k designated vertices $x_1, \ldots, x_k$ such that

$$R = \left\{ \big(f(x_1), \ldots, f(x_k)\big) \mid f \colon S \to \Gamma \text{ homomorphism} \right\} .$$

We call these relations *p.p.-definable*, and denote the relational structure that contains all such relations for a given Γ by $\langle \Gamma \rangle_{pp}$. Likewise, the larger set of all first order definable relations is denoted by $\langle \Gamma \rangle_{fo}$.

A relational structure Γ is called *homogeneous* (in the literature also *ultrahomogeneous*) if every partial isomorphism between two finite substructures can be extended to an automorphism of Γ. Prominent examples of countable homogeneous structures are the *Rado graph* $\mathbf{R}$ and the dense linear order $(\mathbb{Q}, <)$. The Rado graph can be defined as the unique (up to isomorphism) model of the almost-sure theory of finite random graphs. Homogeneous structures have been classified for graphs [26], for tournaments, for posets [36], and finally digraphs [10] (there are continuum many homogeneous digraphs). For homogeneous structures with arbitrary relational signatures a classification is not yet known.

The *age* $\mathrm{Sub}(\Gamma)$ of a relational structure Γ over τ is the set of all finite structures over τ that (isomorphically) embed in Γ. An important property of countable homogeneous structures is their characterization by *amalgamation classes*. A class of finite structures $\mathcal{C}$ is an *amalgamation class* if $\mathcal{C}$ is nonempty, closed under isomorphism and taking induced substructures, and has the *amalgamation property*. The amalgamation property says that for all $A, B_1, B_2 \in \mathcal{C}$ and embeddings $e : A \to B_1$ and $f : A \to B_2$ there exists $C \in \mathcal{C}$ and embeddings $g : B_1 \to C$ and $h : B_2 \to C$ such that $ge = hf$.

Theorem 1 (Fraïssé [18]). *A countable class $\mathcal{C}$ of finite relational structures with countable signature is the age of a unique (up to isomorphism) countable homogeneous structure if and only if $\mathcal{C}$ is an amalgamation class.*

If $\mathcal{C}$ is an amalgamation class, we call the corresponding countable homogeneous structure the *Fraïssé-limit* of $\mathcal{C}$. By definition amalgamation classes can be defined by a set of forbidden induced finite substructures. For a set of finite structures $\mathcal{N}$ over τ we denote by Forb($\mathcal{N}$) the set of finite structures S over τ such that no structure in $\mathcal{N}$ is embeddable in S. We say that a class of finite structures $\mathcal{C}$ over τ is finitely axiomatizable if there exists a first order formula φ over τ such that for all τ-structures A we have $A \in \mathcal{C}$ if and only if A is a model of φ. By compactness it follows that an amalgamation class $\mathcal{C}$ is finitely axiomatizable if and only if $\mathcal{C} = Forb(\mathcal{N})$ for some *finite* set of forbidden induced substructures $\mathcal{N}$.

A homogeneous structure Γ over a finite signature is ω-categorial, i.e. every countable structure satisfying the same first order formulas as Γ is isomorphic to Γ. We will later use that fact that a ω-categorical structure has for a fixed number k of variables only finitely many pairwise inequivalent formulas $\varphi(x_1, \ldots, x_k)$. Another important and well-known property of homogeneous structures of finite signature is that they allow for *quantifier elimination* (see [22]). This means that every first-order formula with k free variables is over Γ equivalent to a quantifier-free formula with k free variables.

3 Combinatorial Constraint Satisfaction

Let Γ be an arbitrary structure with relational signature τ - also called the *template*. Then the constraint satisfaction problem CSP(Γ) is the following computational problem:

Given: A finite τ-structure S.
Question: Is there some homomorphism from S to Γ?

Formally, we denote by CSP(Γ) the set of all finite τ-structures that homomorphically map to Γ. For finite Γ we can assume without loss of generality that Γ is a *core*, i.e. all endomorphisms of Γ are automorphisms. If Γ is a core, adding all the singleton relations to Γ does not change the complexity of CSP(Γ) (as stated in [7]). In this case Γ becomes a homogeneous relational structure. Therefore constraint satisfaction with homogeneous templates can be seen as a generalization of constraint satisfaction with finite templates.

All constraint satisfaction problems with finite Γ are clearly contained in NP. If the age of a relational homogeneous structure Γ of finite signature is finitely axiomatizable then CSP(Γ) is also contained in NP. To see this, suppose we are given an instance S of CSP(Γ). An algorithm can then guess the image of S under a homomorphism, and verify that the image belongs to the age of Γ in polynomial time using the finite axiomatization. Thus we have

Proposition 1. *Let Γ be a countable homogeneous relational structure of finite signature τ with a finitely axiomatizable age. Then CSP(Γ) in in NP.*

Note that we need the axiomatizability assumption in Proposition 1 as there exist homogeneous Γ such that CSP(Γ) is undecidable, see Section 6. In analogy with the dichotomy conjecture of Feder and Vardi [17], we can make the following conjecture.

Conjecture 1 (Dichotomy). Let Γ be a countable homogeneous relational structures with a finitely axiomatizable age. Then the class of constraint satisfaction problems CSP(Γ) has a dichotomy.

For both finite and infinite Γ, the following simple lemma explains the relevance of p.p.-definable relations in constraint satisfaction. Suppose we extend a relational structure Γ by a p.p.-definable relation R. This does not change the computational complexity of the corresponding constraint satisfaction problem, since we can replace every occurence of R in an instance of CSP(Γ) by the τ-structure that defines R.

Lemma 1. *Let Γ be a τ-structure and let Γ' be the extension of this structure by a relation R that is p.p.-definable over Γ. Then* CSP(Γ) *is polynomial-time equivalent to* CSP(Γ').

In the next section we introduce the algebraic notions that will be needed to characterize p.p.-definability.

4 The Clone of Polymorphisms

In this section, D will stand for a countable set and O for the set of *finitary operations* on D, i.e., functions from D^k to D for finite k. We say that $f \in O$ *preserves* a k-ary relation $R \subseteq D^k$ if R is a subalgebra of $(D, f)^k$. An operation that preserves all relations of a relational structure Γ is called a *polymorphism* of Γ. The set of all k-ary polymorphisms of Γ is denoted by $Pol^{(k)}(\Gamma)$, and we write $Pol(\Gamma)$ for the set of all finitary polymorphisms $Pol(\Gamma) = \bigcup_{i=1} Pol^{(i)}(\Gamma)$.

The notion of a *product* of relational structures allows an equivalent definition of polymorphisms, relating polymorphisms to homomorphisms. The *(categorical- or cross-) product* $\Gamma_1 \times \Gamma_2$ of two relational τ-structures Γ_1 and Γ_2 is a τ-structure on the domain $D_{\Gamma_1} \times D_{\Gamma_2}$. For all relations $R \in \tau$ the relation $R\big((x_1, y_2), \ldots, (x_k, y_k)\big)$ holds in $\Gamma_1 \times \Gamma_2$ iff $R(x_1, \ldots, x_k)$ holds in Γ_1 and $R(y_1, \ldots, y_k)$ holds in Γ_2. Comparing the corresponding definitions we see that a k-ary polymorphism f of a relational structure is a homomorphism from $\Gamma^k = \Gamma \times \ldots \times \Gamma$ to Γ, i.e., for an m-ary relation R in τ, if $R(x_1, \ldots, x_m)$ holds in Γ^k then $R\big(f(x_1), \ldots, f(x_m)\big)$ holds in Γ.

An operation π is a *projection* (or a *trivial polymorphism*) if for all n-tuples, $\pi(x_1, \ldots, x_n) = x_i$ for some fixed $i \in \{1, \ldots, n\}$. The *composition* of a k-ary operation f and k operations $g_1, \ldots, g_k$ of arity n is an n-ary operation defined by

$$f(g_1, \ldots, g_k)(x_1, \ldots, x_n) = f\big(g_1(x_1, \ldots, x_n), \ldots, g_k(x_1, \ldots, x_n)\big) \quad .$$

A *clone* F is a set of operations from O that is closed under composition and that contains all projections. We write D_F for the *domain* D of the clone F. For a set of operations F from O we write $\langle F \rangle$ for the smallest clone containing all operations in F (the clone *generated* by F). Observe that $Pol(\Gamma)$ is a clone with the domain D_Γ.

Moreover, $Pol(\Gamma)$ is also closed under interpolation: We say that an operation $f \in O$ is an *interpolation* of a subset F of O if for every finite subset B of D there is some operation $g \in \langle F \rangle$ such that $f|_B = g|_B$ (f restricted to B equals g restricted to B, i.e., $f(a) = g(a)$ for every $a \in B^k$). The set of interpolations of F is called the *local closure* of F. If the maximal arity of Γ is bounded, $Pol(\Gamma)$ is also locally closed.

The converse was proved by Rosenberg and Schweigert:

Proposition 2 (Rosenberg and Schweigert [33]). *A set $F \subseteq O$ of operations is locally closed if and only if $X = Pol(\Gamma)$ for some relational structure Γ of bounded maximal arity.*

Together with a theorem of Larose and Tardif [27] on infinite graphs this is one of the few known results on infinite structures and their polymorphisms. Many results on clones in general can be found in [38]. Local clones on infinite domains were introduced and studied by Rosenberg and Schweigert in [33, 34], where they studied *locally maximal* clones to characterize *local maximality* of a clone.

Important properties of operations in a clone: a k-ary operation f is *conservative* iff $f(x_1, \ldots, x_k) \in \{x_1, \ldots, x_k\}$ for all $x_1, \ldots, x_k \in D$; it is *idempotent* iff $f(x, \ldots, x) = x$ for all $x \in D$. An operation f is called *essentially unary* iff there is a unary operation f_0 such that $f(x_1, \ldots, x_k) = f_0(x_i)$ for some $i \in \{1, \ldots, k\}$. A relational structure Γ is called *projective*, iff all idempotent polymorphisms of Γ are projections, and *strongly projective*, iff all polymorphisms of Γ are projections [32].

Let F be a (local) clone with domain D. Then $R \subseteq D^m$ *is invariant under* F, if every $f \in F$ preserves R. We denote by $Inv(F)$ the relational structure containing the set of all relations left invariant under F. A fundamental result of Bodnarčuk et al. [5] (other presentations can be found in [12, 31]) says that for arbitrary finite relational structures Γ the p.p.-definable relations can be characterized as the invariants of the polymorphisms of Γ.

Theorem 2 (Bodnarčuk et al. [5]). *Let Γ be a finite relational structure. Then*

$$\langle \Gamma \rangle_{pp} = Inv(Pol(\Gamma)) \ .$$

The proof of Theorem 2 also shows that it is decidable, whether for a given finite relational structure Γ a given relation R is p.p.-definable or not. Generalizations of Theorem 2 and the related Galois correspondence were also studied for infinite domains. For arbitrary relational structures Γ the set of relations $Inv(Pol(\Gamma))$ was characterized with local closure operators on relational algebras in [37] (see also [30], page 32).

In the next section we will show that for countable homogeneous structures Γ any first-order definable relation is in $\langle \Gamma \rangle_{pp}$ if and only if it is left invariant under all polymorphisms of finite arity. But first we note that the following is well-known for arbitrary cardinalities of the domain.

Proposition 3 (see e.g. [31]). *Let Γ be a relational structure. Then*

$$\langle \Gamma \rangle_{pp} \subseteq Inv(Pol(\Gamma)) \ .$$

Proof. Let R be a relation in $\langle \Gamma \rangle_{pp}$. We prove that $R \in Inv(Pol(\Gamma))$ by induction on the length of a defining p.p.-formula φ. The claim is true for $\varphi = R(x_1, \ldots, x_n)$. For $\varphi = \exists x . \varphi'$ we observe that every polymorphism that is left invariant by φ' also leaves φ invariant. The same holds for $\varphi = \varphi_1 \wedge \varphi_2$. □

For a structure with a countable domain the inclusion of Proposition 3 might be strict. Consider for instance the following relational structure $\Gamma =$

$(\mathbb{N}; R_1, R_2, R_3)$ on the natural numbers communicated to the authors by Ferdinand Börner. We show that $Inv(Pol(\Gamma))$ contains relations that are not p.p.-definable.

$$\begin{aligned} R_1 &= \{(a,b,c,d) \mid a = b \text{ or } c = d,\ a,b,c,d \in \mathbb{N}\} \\ R_2 &= \{(0)\} \\ R_3 &= \{(a, a+1) \mid a \in \mathbb{N}\} \end{aligned}$$

Every function preserving R_1 is essentially unary. If f is unary and preserves R_2 then $f(0) = 0$. Furthermore, if f preserves R_3 we have $f(a+1) = f(a)+1$ for all a, and inductively follows $f(a) = a$. Therefore $Pol(\Gamma)$ is the set of all projections. Every projection preserves all relations, but even the unary first-order definable relation $\{x \mid x = 1 \vee x = 3\}$ is not p.p.-definable.

For ω-categorical structures Γ the situation looks better: It is known that the first-order definable relations are precisely the relations that are invariant under the automorphisms of Γ, i.e. $\langle \Gamma \rangle_{fo} = Inv(Aut(\Gamma))$ (see e.g. [22]). We will prove a corresponding theorem for primitive positive definability in the next section.

5 A Characterization of Primitive Positive Definability

We characterize the primitive positive first-order definable relations over an ω-categorical structure Γ by the polymorphisms of Γ of finite arity.

Theorem 3. *Let Γ be an ω-categorical structure with relational signature τ. Then a relation R on Γ is invariant under the polymophisms of Γ if and only if R is p.p.-definable, i.e.,*

$$\langle \Gamma \rangle_{pp} \;=\; Inv(Pol(\Gamma)).$$

Proof. We already stated in Proposition 3 that the p.p.-definable relations over Γ are invariant under the polymophisms of Γ.

For the converse, let R be a k-ary relation from $Inv(Pol(\Gamma))$. Note that R is first-order definable in Γ: This follows already from ω-categoricity of Γ and the fact that Γ and $Inv(Pol(\Gamma))$ have the same automorphism group. The relation R is a union of *finitely* many orbits of the automorphism group of Γ (this is a consequence of ω-categoricity), it can be defined by a disjunction φ of τ-formulas that define these orbits. Let $M_1, \ldots, M_w$ be the (satisfiable) monomials in this disjunction, and let $x_1, \ldots, x_k$ be the variables of the monomials.

We have to construct a finite τ-structure Q with designated vertices $v_1, \ldots, v_k$ such that

$$R = \{(f(v_1), \ldots, f(v_k)) \mid f \colon Q \to \Gamma \text{ homomorphism}\}.$$

The idea is to first consider an *infinite* τ-structure with this property, namely the categorial product Γ^w, and then to apply König's Lemma to prove the existence of a suitable finite substructure (i.e., a compactness argument).

For each monomial $M_j \in M_1, \ldots, M_w$ of φ we find a substructure $a_1^j, \ldots, a_k^j$ of Γ, such that $a_1^j, \ldots, a_k^j$ satisfies M_j in Γ. Let $b_1, b_2, \ldots$ be an enumeration of the w-tuples in D_Γ^w, starting with $b_i = (a_1^i, \ldots, a_w^i)$ for $1 \leq i \leq k$. Since R is invariant under all polymorphisms, all homomorphisms from Γ^w to Γ map $b_1, \ldots, b_k$ to a tuple satisfying one of the monomials $M_1, \ldots, M_w$.

We claim that there is a finite substructure Q of Γ^w with this property. Assume for contradiction that all finite substructures of Γ^w containing $b_1, \ldots, b_k$ have a homomorphism mapping $b_1, \ldots, b_k$ to a tuple not satisfying φ. We call such a homomorphism a *bad* homomorphism. We now construct a bad homomorphism from Γ^w to Γ, i.e. the images of $b_1, \ldots, b_k$ do not satisfy φ. This contradicts the fact that R is invariant under all polymorphisms.

To this end, consider the following infinite but finitely branching tree of finite substructures of Γ. The root of the tree is the empty substructure. Now assume inductively that on level $i \geq 0$ of the tree we have structures P with i vertices, such that for every i-type over Γ (see [22]) there is precisely one substructure of Γ with vertices $(u_1, \ldots, u_i)$ in the tree that have this type. Such vertices exist since Γ is ω-categorical. Moreover we only find a finite number of structures on level i. We will now show how to extend this tree to the level $i+1$. Consider some $i+1$-type over Γ. We find an extension P' of one of the structures P of the tree on level i that realizes that type. Then we make P' adjacent to P, and continue like this for all $i+1$ types over Γ.

By our assumption we will find a bad homomorphism to Γ for each finite substructure of Γ^w containing $b_1, \ldots, b_k$. The image of every such homomorphism has the same type as some node in the above tree. By König's Lemma we will find an infinite chain in this tree such that every substructure of Γ on this chain is the image of a bad homomorphism. Hence there also exists an homomorphism from Γ^w to this structure such that the image of $b_1, \ldots, b_k$ do not satisfy φ.

We proved by contradiction that there must be a finite substructure Q containing the vertices $b_1, \ldots, b_k$ of Γ^k such that all homomorphisms from Q to Γ map $b_1, \ldots, b_k$ to a tuple satisfying φ. Conversely, every mapping $f : Q \to \Gamma$ such that the tuple $(f(b_1), \ldots, f(b_k))$ satisfies in Γ the monomial M_j can be extended to a homomorphism $f : \Gamma^k \to \Gamma$. To see this note that both $a^j_1, \ldots, a^j_k$ and $(f(b_1), \ldots, f(b_k))$ satisfy M_j and that both lie in the same orbit of $Aut(\Gamma)$. Thus we can choose f to be the jth projection combined with the automorphism sending $(a^j_1, \ldots, a^j_k)$ to $(f(b_1), \ldots, f(b_k))$. This completes the proof. □

The clone of polymorphisms of an infinite structure is usually a very complicated object. However for homogeneous structures of finite signature we have the following:

Proposition 4. *Let Γ be a homogeneous structure with finite relational signature. Then the polymorphisms of Γ are locally generated by countably many polymorphisms of Γ and the automorphism group $Aut(\Gamma)$.*

Proof. Let $u_1, u_2, \ldots, u_k$ be m-tuples from Γ, and let $v_i = g(u_i)$ for some m-ary polymorphism g of Γ. Depending on m and k there are only finitely many isomorphism types of the substructure of Γ induced by the elements of $u_1, u_2, \ldots, u_k$ and the elements $v_1, \ldots, v_k$. Let F be a set of polymorphisms of arity k containing a polymorphism g for each of these isomorphism types in Γ.

Now let f be an m-ary polymorphism of Γ. We show that f is locally generated by $F \cup Aut(\Gamma)$. Let $B \subseteq D_\Gamma$ a set of cardinality k. By the definition of F, the restriction $f|_B$ is isomorphic to the restriction of one of the operations $g \in F$. Let π be the isomorphism. By homogenity of Γ, π can be extended to an automorphism π' of Γ. The identity $f|_B = \pi'(g)|_B$ implies that f is in the local closure of $Aut(\Gamma) \cup F$. □

6 A Catalog of Homogeneous Templates

We consider various homogeneous structures, some of their polymorphisms and their corresponding constraint satisfaction problems. In particular we look at the binary structures from the classification project for countable homogeneous structures. We end with a brief discussion of two applications with larger signatures.

The countable homogeneous tournaments. We start with the homogeneous tournaments, which have been classified by Lachlan [25]. There are a few types only: The oriented cycle C_3, the dense linear order $(\mathbb{Q}, <)$, the dense local order $S(2)$, and the generic tournament for the set of all finite tournaments.

The problem $\mathrm{CSP}(C_3)$ is known to be tractable. The constraint satisfaction problem of the dense linear order $(\mathbb{Q}, <)$ is computationally equivalent to the problem whether a given digraph D is acyclic. This tractable problem can not be formulated as a constraint satisfaction problem with a finite template. Note that the relational structure $(\mathbb{Q}, <)$ is not projective, e.g. $x, y \mapsto \max(x, y)$ is a polymorphism. The homogeneous tournament which is the Fraïssé-limit of all finite tournaments has a trivial constraint-satisfaction problem: Every finite tournamant homomorphically maps to it. Thus the only interesting remaining case is the *dense linear order* $S(2)$; see [10]).

The countable homogeneous graphs. Lachlan and Woodrow [26] showed that every infinite such graph is either the Rado-graph, the *generic* K_n-free graph, a homogeneous equivalence relation, or a complement of these.

The Rado graph $\mathbf{R}$ is the Fraïssé-limit of the class of all graphs, therefore the constraint satisfaction problem for the Rado graph is trivial: Every graph can be homomorphically mapped to it. The automorphism group of $\mathbf{R}$ has a rich structure (see e.g. [9] for an overview). Łuczak and Nešetřil [28] showed that the Rado graph as well as the generic K_n-free graphs are projective. This is in interesting opposition to the finite case, where projectivity of a core of cardinality at least three implies NP-hardness for the corresponding constraint satisfaction problem, which can be seen using Theorem 2 by reduction of k-colorability. Again the constraint satisfaction problem is easy, since every graph which does not contain the K_n as a subgraph embeds into the generic K_n-free graph. The only interesting remaining case are the equivalence relations with n equivalence classes. Here we find a homomorphism to a graph of size n, in which case the problem reduces to a constraint satisfaction problem with a finite template, which is NP-hard for $n \geq 3$.

Remark: It is perhaps interesting to note that we can use the same results to give a new and short proof that almost all constraint satisfaction problems are NP-complete, if the template is a finite undirected graph. The fact that almost all graphs are strongly projective (Nešetřil and Łuczak [28]) combined with Theorem 2 shows that almost all graphs can simulate the inequality-relation. This implies NP-hardness of the constraint satisfaction problem on a domain of size at least three. Note that we did not use the involved proof of the dichotomy for graphs in [20].

Countable Homogeneous Digraphs. The countable homogeneous digraphs have been classified by Cherlin [10], and there are uncountably many. But the classification shows that the age of all but a countable well-uderstood class of homogeneous digraphs has the strong *free amalgamation property.* Therefore it is easy to see that CSP(Γ) is the set of all (weak) subgraphs of Γ, and that the set of constraint satisfaction problems is also uncountable. Thus there are homogeneous digraphs with an undecidable constraint satisfaction problem. However, restricting to the homogeneous structures that have a finitely axiomatizable age, it is easy to determine the complexity of their constraint satisfaction problems, with the exception of CSP($S(2)$, CSP($S(3)$) and CSP($\mathcal{P}(3)$) (see [10]).

Tree Descriptions. The following constraint satisfaction problem was studied in [11]. Given: a finite structure S over the signature $\tau = \{\longrightarrow, \perp\}$ containing two binary relation symbols. Question: can we find a rooted forest F on the vertices of S such that every edge from $\longrightarrow$ lies in the transitive closure of F, and every edge $\perp$ does not? Let us call such τ-structures S *solvable.*

Using the notion of *dense trees* (see [15]) we can formulate this problem (and related problems) as a constraint satisfaction problem CSP(Γ) for appropriate Γ. This means, it is possible to find an ω-categorical structure Γ such that CSP(Γ) contains precisely the solvable τ-structures. The problem can be decided by a polynomial time algorithm [4]. The graph algorithm presented there can be generalized to various constraint satisfaction problems for tree-like structures. Note that every ω-categorical structure Γ can be made homogeneous by expanding the signature and Γ by some first-order definable relations. In the case of countable dense trees this is possible by two additional ternary relations [16].

Allen's Interval Algebra and its Fragments. Consider as a base set D the open intervals on the rational numbers, and the following binary relations on these intervals: Let $x = (x^-, x^+)$ and $y = (y^-, y^+)$ be open intervals. We define

- The interval x *precedes* y, $x \,\mathsf{p}\, y$, iff $x^+ < y^-$.
- The interval x *overlaps* y, $x \,\mathsf{o}\, y$, iff $x^- < y^- < x^+$ and $x^+ < y^+$.

- The interval x is *during* y, $x \mathsf{d} y$, iff $y^- < x^-$ and $x^+ < y^+$.
- The interval x *starts* y, $x \mathsf{s} y$, iff $x^- = z^-$ and $x^+ > y^-$.
- The interval x *finishes* y, $x \mathsf{f} y$, iff $x^+ = y^+$ and $x^- > y^-$.
- The interval x *meets* y, $x \mathsf{m} y$, iff $x^+ = y^-$.
- The interval x *equals* y, $x \equiv y$, iff $x^- = y^-$ and $x^+ = y^+$.

Proposition 5. *For any set of relations derived from* $\mathsf{p}, \mathsf{o}, \mathsf{d}, \mathsf{s}, \mathsf{m}, \mathsf{f}$ *and* $\equiv$ *by union and complementation that contains the base relations* $\mathsf{p}, \mathsf{o}, \mathsf{d}, \mathsf{s}, \mathsf{m}, \mathsf{f}$ *and* $\equiv$*, the corresponding class of all finite interval structures with these relations is an amalgamation class.*

Proof. It suffices to verify the amalgamation property. Let B_1 and B_2 be two sets of intervals on the rational number, and suppose a set of intervals A embeds via e_1 and e_2 into B_1 and B_2. The partial isomorphism between $e_1(A)$ and $e_2(A)$ corresponds to a partial isomorphism of the respective rational numbers of the intervals. By homogenity of the rational numbers, we can extend this isomorphism to an automorphism f of the rational numbers and hence an automorphism of the set of all intervals. Then f and the identity are the solution of our amalgamation problem: They are embedings of B_1 and B_2 into the set of intervals $f(B_1) \cup B_2$ such that $fe_1 = e_2$. □

The constraint satisfaction problems for the respective countable homogeneous relational structures have a dichotomy [8]. Whereas for finite templates all known hard constraint satisfaction problems are hard because they can express the relation one-in-three-sat, here the problem of *betweennness* [19] is used to prove hardness.

7 Related Work

We want to relate our work to previous unifying approaches to constraint satisfaction. The literature on *combining constraint solving* [2, 24] has a broader view on constraint satisfaction, and also uses tools from universal algebra. However they are concerned mainly with decidability questions of more expressive constraint languages.

Various logical formalisms have been proposed to formulate constraint satisfaction problems as the model-checking problems of certain higher-order logics [17]. On of them is the class SNP, the class of existential second-order formulas Φ with a universal first-order part, which is using the relation symbols of some given signature τ and the existentially quantified relation symbols. One of the results in [17] says that every problem in NP is equivalent to a problem in SNP, even if the relation symbols from τ occur only negatively in Φ (in which case the class is called *monotone* SNP). To answer the question whether the model checking problem of a given monotone SNP formula can be described as a constraint satisfaction problem with a countable homogeneous structure, the following problem posed by Cherlin [10] is of importance:

Problem 1. Let τ be a relational signature and $\mathcal{N}$ a finite set of finite τ-structures. Give a good criterion for Forb($\mathcal{N}$) to be an amalgamation class.

8 Conclusion

Constraint satisfaction problems with countable homogeneous templates cover several classes of constraint satisfaction problems that were investigated in the literature. Examples are tree description languages and subalgebras of Allen's interval algebra. Using the classification of homogeneous digraphs we can determine the complexity of the constraint satisfaction problems for homogeneous graphs and tournaments. The complexity of the constraint satisfaction problems where the homogeneous digraph has a finitely axiomatizable age will appear in the full version of this paper.

For larger signatures the classification of homogeneous structures is a very difficult task. To study the complexity of a constraint satisfaction problem CSP(Γ) it is useful to know whether a given first-order relation is p.p.-definable over Γ. In particular for large signatures this can be difficult. In Section 5 on page 9 we show that p.p.-definability of a relation is characterized by a countable set of polymorphisms and the automorphisms of Γ. We ask the following question:

Problem 2. Let $\mathcal{N}$ be a finite set of relational structures over signature τ such that Forb($\mathcal{N}$) is the age of a homogeneous structure Γ. Given a first-order formula φ over τ, is it decidable whether φ is on Γ equivalent to a primitive positive formula?

The techniques to describe p.p.-definability might be applied to simplify the technical and intricate proofs in the classification of the tractable fragments of Allen's interval algebra.

References

1. J. F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983.
2. F. Baader and K. Schulz. Combining constraint solving. *H. Comon, C. March, and R. Treinen, editors, Constraints in Computational Logics*, 2001.
3. M. Bodirsky, D. Duchier, J. Niehren, and S. Miele. An efficient algorithm for weakly normal dominance constraints. 2002.
4. M. Bodirsky and M. Kutz. Pure dominance constraints. In *Proceedings of the 19th Annual Symposium on Theoretical Aspects of Computer Science (STACS'02)*, pages 287–298, Antibes - Juan le Pins, 2002.
5. 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.
6. A. Bulatov. Tractable constraint satisfaction problems on a 3-element set. *Research Report*, 2002.
7. A. Bulatov, A. Krokhin, and P. G. Jeavons. Classifying the complexity of constraints using finite algebras. *submitted*, 2003.
8. A. Bulatov, A. Krokhin, and P. Jonsson. Reasoning about temporal relations: The maximal tractable subalgebras of allen's interval algebra. *Technical report PRG-RR-01-12, Oxford University, submitted for publication*, 2001.
9. P. J. Cameron. The random graph. *R.L.Graham and J. Nesetril, editors, The Mathematics of Paul Erdõs*, 1996.
10. G. Cherlin. The classification of countable homogeneous directed graphs and countable homogeneous n-tournaments. *AMS Memoir*, 131(621), January 1998.
11. T. Cornell. On determining the consistency of partial descriptions of trees. In *32nd Annual Meeting of the Association for Computational Linguistics (ACL'94)*, pages 163–170, 1994.
12. V. Dalmau. Computational complexity of problems over generalized formulas. PhD-thesis at the Departament de Llenguatges I Sistemes Informatics at the Universitat Politecnica de Catalunya, 2000.
13. V. Dalmau. A new tractable class of constraint satisfaction problems. In *6th International Symposium on Mathematics and Artificial Intelligence*, 2000.
14. R. Dechter. Local and global relational consistency. *Journal of Theoretical Computer Science*, 1996.
15. M. Droste. Structure of partially ordered sets with transitive automorphism groups. *AMS Memoir*, 57(334), September 1985.
16. M. Droste, W. Holland, and D. Macpherson. Automorphism groups of infinite semilinear orders (i). *Proc. London Math. Soc.*, 58:454–478, 1989.
17. T. Feder and M. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28:57–104, 1999.
18. R. Fraïssé. *Theory of Relations*. North-Holland, 1986.

19. Garay and Johnson. *A Guide to NP-completeness.* CSLI Press, Stanford, 1978.
20. P. Hell and J. Nešetřil. On the complexity of H-coloring. *Journal of Combinatorial Theory, Series B*, 48:92–110, 1990.
21. P. Hell, J. Nešetřil, and X. Zhu. Duality and polynomial testing of tree homomorphisms. *Trans. Amer. Math. Soc.*, 348(4):1281–1297, 1996.
22. W. Hodges. *A shorter model theory.* Cambridge University Press, 1997.
23. P. Jeavons, D. Cohen, and M. Gyssens. Closure properties of constraints. *Journal of the ACM*, 44(4):527–548, 1997.
24. H. Kirchner and C. Ringeissen. Combining symbolic constraint solvers on algebraic domains. *Journal of Symbolic Computation*, 18(2):113–155, 1994.
25. A. H. Lachlan. Countable homogeneous tournaments. *Trans. Amer. Math. Soc.*, 284:431–461, 1984.
26. A. H. Lachlan and R. Woodrow. Countable ultrahomogeneous undirected graphs. *Trans. Amer. Math. Soc.*, 262(1):51–94, 1980.
27. B. Larose and C. Tardif. Strongly rigid graphs and projectivity. *Multiple-Valued Logic 7*, pages 339–361, 2001.
28. T. Łuczak and J. Nešetřil. Projective graphs. *Submitted*, 2003.
29. B. Nebel and H.-J. Bürckert. Reasoning about temporal relations: A maximal tractable subclass of Allen's interval algebra. *Journal of the ACM*, 42(1):43–66, 1995.
30. R. Pöschel. A general galois theory for operations and relations and concrete characterization of related algebraic structures. *Technical Report of Akademie der Wissenschaften der DDR (Berlin)*, 1980.
31. R. Pöschel and L. A. Kalužnin. *Funktionen- und Relationenalgebren.* DVW, 1979.
32. I. G. Rosenberg. Strongly rigid relations. *Rocky Mountain Journal of Mathematics*, 3(4):631–639, 1973.
33. I. G. Rosenberg and D. Schweigert. Local clones. *J. Algorithms, Languages and Combinatorics*, 5(4):421–455, 2000.
34. I. G. Rosenberg and D. Schweigert. Locally maximal clones II. *J. Algorithms, Languages and Combinatorics*, 5(4):421–455, 2000.
35. T. J. Schaeffer. The complexity of satisfiability problems. *In Proc. 10th ACM Symp. on Theory of Computing*, pages 216–226, 1978.
36. J. H. Schmerl. Countable homogeneous partially ordered sets. *Algebra Universalis*, 9:317–321, 1979.
37. L. Szabó. Concrete representation of relatied structures of universal algebras. *Acta Sci. Math. (Szeged)*, 40:175–184, 1978.
38. A. Szendrei. *Clones in universal Algebra.* Seminaire de mathematiques superieures. Les Presses de L'Universite de Montreal, 1986.