# CONSTRAINT SATISFACTION PROBLEMS FOR REDUCTS OF HOMOGENEOUS GRAPHS

MANUEL BODIRSKY, BARNABY MARTIN, MICHAEL PINSKER, AND ANDRÁS PONGRÁCZ

Abstract. For $n \geq 3$, let $(H_n, E)$ denote the $n$-th Henson graph, i.e., the unique countable homogeneous graph with exactly those finite graphs as induced subgraphs that do not embed the complete graph on $n$ vertices. We show that for all structures $\Gamma$ with domain $H_n$ whose relations are first-order definable in $(H_n, E)$ the constraint satisfaction problem for $\Gamma$ is either in P or is NP-complete.

We moreover show a similar complexity dichotomy for all structures whose relations are first-order definable in a homogeneous graph whose reflexive closure is an equivalence relation.

Together with earlier results, in particular for the random graph, this completes the complexity classification of constraint satisfaction problems of structures first-order definable in countably infinite homogeneous graphs: all such problems are either in P or NP-complete.

## 1. Introduction

1.1. **Constraint satisfaction problems.** A *constraint satisfaction problem* (CSP) is a computational problem in which the input consists of a finite set of variables and a finite set of *constraints*, and where the question is whether there exists a mapping from the variables to some fixed domain such that all the constraints are satisfied. We can thus see the possible constraints as relations on the domain, and in an instance of the CSP, we are asked to assign domain values to the variables such that certain specified tuples of variables become elements of certain specified relations.

When the domain is finite, and arbitrary constraints are permitted, then the CSP is NP-complete. However, when only constraints from a restricted set of relations on the domain are allowed in the input, there might be a polynomial-time algorithm for the CSP. The set of relations that is allowed to formulate the constraints in the input is often called the *constraint language*. The question which constraint languages give rise to polynomial-time solvable CSPs has been the topic of intensive research over the past years. It has been conjectured by Feder and Vardi [FV99] that CSPs for constraint languages over finite domains have a complexity dichotomy: they are either in P or NP-complete. This conjecture remains unsettled, although dichotomy is now known on substantial classes (for example when the domain has at most three elements [Sch78, Bul06] or when the constraint language contains a single binary relation without sources and sinks [HN90, BKN09]). Various methods, combinatorial (graph-theoretic), logical, and universal-algebraic have been brought to bear on this classification

*Date*: February 18, 2016.

The first and fourth author have received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 257039). Manuel Bodirsky has also been supported by the DFG-funded project 'Topo-Klon' (Project number 622397). The second and fourth author have received funding from the EPSRC grant "Infinite domain Constraint Satisfaction Problems", grant no. EP/L005654/1. The third author has received funding from project P27600 of the Austrian Science Fund (FWF).

project, with many remarkable consequences. A conjectured delineation for the dichotomy was given in the algebraic language in [BKJ05].

When the domain is infinite, the complexity of the CSP can be outside NP, and even undecidable [BN06]. But for natural classes of such CSPs there is often the potential for structured classifications, and this has proved to be the case for structures first-order definable over the order $(\mathbb{Q},<)$ of the rationals [BK09] or over the integers with successor [BMM15]. Another classification of this type has been obtained for CSPs where the constraint language is first-order definable over the random (Rado) graph [BP15a], making use of structural Ramsey theory. This paper was titled 'Schaefer's theorem for graphs' and it can be seen as lifting the famous classification of Schaefer [Sch78] from Boolean logic to logic over finite graphs, since the random graph is universal for the class of finite graphs.

1.2. **Homogeneous graphs and their reducts.** The notion of *homogeneity* from model theory plays an important role when applying techniques from finite-domain constraint satisfaction to constraint satisfaction over infinite domains. A relational structure is *homogeneous* if every isomorphism between finite induced substructures can be extended to an automorphism of the entire structure. Homogeneous structures are uniquely (up to isomorphism) given by the class of finite structures that embed into them. The structure $(\mathbb{Q},<)$ and the random graph are among the most prominent examples of homogeneous structures. The class of structures that are definable over a homogeneous structure with finite relational signature is a very large generalisation of the class of all finite structures, and CSPs for those structures have been studied independently in many different areas of theoretical computer science, e.g. in temporal and spatial reasoning, phylogenetic analysis, computational linguistics, scheduling, graph homomorphisms, and many more; see [Bod12] for references.

While homogeneous relational structures are abundant, there are remarkably few countably infinite homogeneous (undirected, irreflexive) *graphs*; they have been classified by Lachlan and Woodrow [LW80]. Besides the random graph mentioned earlier, an example of such a graph is the countable homogeneous *universal triangle-free* graph, one of the fundamental structures that appears in most textbooks in model theory. This graph is the up to isomorphism unique countable triangle-free graph $(H_3,E)$ with the property that for every finite independent set $X \subseteq H_3$ and for every finite set $Y \subseteq H_3$ there exists a vertex $x \in H_3 \setminus (X \cup Y)$ such that $x$ is adjacent to every vertex in $X$ and to no vertex in $Y$.

Further examples of homogeneous graphs are the graphs $(H_3,E)$, $(H_4,E)$, $(H_5,E)$, $\ldots$, called the *Henson graphs*, and their complements. Here, $(H_n,E)$ for $n>3$ is the generalisation of the graph $(H_3,E)$ above from triangles to cliques of size $n$. Finally, the list of Lachlan and Woodrow contains only one more family of infinite graphs, namely the graphs $(C^s_n,E)$ whose reflexive closure $Eq$ is an equivalence relation with $n$ classes of equal size $s$, where $1 \leq n,s \leq \omega$ and either $n$ or $s$ equals $\omega$, as well as their complements. We remark that $(C^s_n,Eq)$ is itself homogeneous and first-order interdefinable with $(C^s_n,E)$, and so we shall sometimes refer to the *homogeneous equivalence relations*.

All countable homogeneous graphs, and even all structures which are first-order definable over homogeneous graphs, are *ω-categorical*, that is, all countable models of their first-order theory are isomorphic. Moreover, all countably infinite homogeneous graphs $\Gamma$ are *finitely bounded* in the sense that the *age* of $\Gamma$, i.e., the class of finite structures that embed into $\Gamma$, can be described by finitely many forbidden substructures. Finitely bounded homogeneous structures also share with finite structures the property of having a finite description: up to isomorphism, they are uniquely given by the finite list of forbidden structures that describes their age. Recent work indicates the importance of finite boundedness for complexity classification [BOP15, BM16], and it has been conjectured that all structures with a first-order definition in a finitely bounded homogeneous structure enjoy a complexity dichotomy, i.e., their CSP is either in P or NP-complete (cf. [BPP14, BOP15]). The structures first-order definable in homogeneous graphs therefore provide the most natural class on which to test further the methods developed in [BP15a] specifically for the random graph.

In this article we obtain a complete classification of the computational complexity of CSPs where all constraints have a first-order definition in one of the Henson graphs. We moreover obtain such a classification for CSPs where all constraints have a first-order definition in a countably infinite homogeneous graph whose reflexive closure is an equivalence relation, expanding earlier results for the special cases of one single equivalence class (so-called equality constraints [BK08]) and infinitely many infinite classes [BW12]. Together with the above-mentioned result on the random graph, this completes the classification of CSPs for constraints with a first-order definition in any countably infinite homogeneous graph, by Lachlan and Woodrow's classification.

Following an established convention [Tho91, BP11], we call a structure with a first-order definition in another structure $\Delta$ a *reduct* of $\Delta$. That is, for us a reduct of $\Delta$ is as the classical definition of a reduct with the difference that we first allow a first-order expansion of $\Delta$. With this terminology, the present article provides a complexity classification of the CSPs for all reducts of countably infinite homogeneous graphs. In other words, for every such reduct we determine the complexity of deciding its *primitive positive theory*, which consists of all sentences which are existentially quantified conjunctions of atomic formulas and which hold in the reduct. We remark that all reducts of such graphs can be defined by quantifier-free first-order formulas, by homogeneity and $\omega$-categoricity.

For reducts of $(H_n, E)$, the CSPs express computational problems where the task is to decide whether there exists a finite graph without any clique of size $n$ that meets certain constraints. An example of a reduct whose CSP can be solved in polynomial time is $(H_n, \neq, \{(x, y, u, v) : E(x, y) \Rightarrow E(u, v)\})$, where $n \geq 3$ is arbitrary. As it turns out, for every CSP of a reduct of a Henson graph which is solvable in polynomial time, the corresponding reduct over the Rado graph, i.e., the reduct whose relations are defined by the same quantifier-free formulas, is also polynomial-time solvable. On the other hand, the CSP of the reduct $(H_n, \{(x, y, u, v) : E(x, y) \vee E(u, v)\})$ is NP-complete for all $n \geq 3$, but the corresponding reduct over the random graph can be decided in polynomial time.

Similarly, for reducts of the graph $(C_n^s, E)$ whose reflexive closure is an equivalence relation with $n$ classes of size $s$, where $1 \leq n, s \leq \omega$, the computational problem is to decide whether there exists an equivalence relation with $n$ classes of size $s$ that meets certain constraints.

1.3. **Results.** Our first result is the complexity classification of the CSPs of all reducts of Henson graphs, showing in particular that a uniform approach to infinitely many "base structures" (namely, the $n$-th Henson graph for each $n \geq 3$) is possible.

**Theorem 1.1.** *Let $n \geq 3$, and let $\Gamma$ be a finite signature reduct of the $n$-th Henson graph $(H_n, E)$. Then* $\mathrm{CSP}(\Gamma)$ *is either in* P *or* NP*-complete.*

We then obtain a similar complexity dichotomy for reducts of homogeneous equivalence relations, expanding earlier results for special cases [BW12, BK08].

**Theorem 1.2.** *Let $(C_n^s, E)$ be an infinite graph whose reflexive closure Eq is an equivalence relation with $n$ classes of size $s$, where $1 \leq n, s \leq \omega$. Then for any finite signature reduct $\Gamma$ of $(C_n^s, E)$, the problem* $\mathrm{CSP}(\Gamma)$ *is either in* P *or* NP*-complete.*

Together with the classification of countable homogeneous graphs, and the fact that the complexity of the CSPs of the reducts of the Rado graph have been classified [BP15a], this completes the CSP classification of reducts of all countably infinite homogeneous graphs, confirming further instances of the open conjecture that CSPs of reducts of finitely bounded homogeneous structures are either in P or NP-complete [BPP14, BOP15].

**Corollary 1.3.** *Let $\Gamma$ be a finite signature reduct of a countably infinite homogeneous graph. Then* $\mathrm{CSP}(\Gamma)$ *is either in* P *or* NP*-complete.*

1.4. **The strategy.** The method we employ follows to a large extent the method invented in [BP15a] for the corresponding classification problem where the 'base structure' is the random graph. The key component of this method is the usage of Ramsey theory (in our case, a result of Nešetřil and Rödl [NR89]) and the concept of *canonical functions* introduced in [BP14]. There are, however, some interesting differences and novelties that appear in the present proof, as we now shortly outline.

1.4.1. *Henson graphs.* When studying the proofs in [BP15a], one might get the impression that the complexity of the method grows with the model-theoretic complexity of the base structure, and that for the random graph we have really reached the limits of bearableness for applying the Ramsey method.

However, quite surprisingly, when we step from the random graph to the graphs $(H_n, E)$, which are in a sense more complicated structures from a model-theoretic point of view[1], the classification and its proof become easier again. It is one of the contributions of the present article to explain the reasons behind this effect. Essentially, certain *behaviours* of canonical functions (cf. Section 2) existing on the random graph can not be realised in $(H_n, E)$. For example the canonical polymorphisms of behaviour "max" (cf. Section 2) play no role for the present classification, but account over the random graph for the tractability of, inter alia, the 4-ary relation defined by the formula $E(x, y) \vee E(u, v)$.

Interestingly, we are able to reuse results about canonical functions over the random graph, since the calculus for composing behaviours of canonical functions is the same for any other structure with the same type space, and in particular the Henson graphs. Via this meta-argument we can, on numerous occasions, make statements about canonical functions over the Henson graphs which were proven earlier for the Rado graph, ignoring completely the actual underlying structure; even more comfortably, we can *a posteriori* rule out some possibilities in those statements because of the $K_n$-freeness of the Henson graphs. Examples of this phenomenon appear in Lemmas 3.8 and 3.9.

On the other hand, along with these simplifications, there are also new additional difficulties that appear when investigating reducts of $(H_n, E)$ and that were not present in the classification of reducts of the random graph, which basically stem from the lower degree of symmetry of $(H_n, E)$ compared to the Rado graph. For example, in expansions of Henson graphs by finitely many constants, not all orbits induce copies of Henson graphs; the fact that the analogous statement does hold for the Rado graph was used extensively in [BP15a], for example in the very technical Proposition 7.18 of that paper.

[1]For example, the random graph has a *simple* theory [TZ12], whereas the Henson graphs are the most basic examples of structures whose theory is *not* simple.

1.4.2. *Equivalence relations.* Similarly to the situation for the equivalence relation with infinitely many infinite classes studied in [BW12], there are two interesting sources of NP-hardness for the reducts $\Gamma$ of other homogeneous equivalence relations: namely, if the equivalence relation is invariant under the polymorphisms of $\Gamma$, then the structure obtained from $\Gamma$ by factoring by the equivalence relation might have a NP-hard CSP, implying NP-hardness for the CSP of $\Gamma$ itself; or, roughly, for a fixed equivalence class the restriction of $\Gamma$ to that class might have a NP-hard CSP, again implying NP-hardness of the CSP of $\Gamma$ (assuming that $\Gamma$ is a *model-complete core*, see Sections 3 and 6). But whereas for the equivalence relation with infinitely many infinite classes both the factor structure and the restriction to a class are again infinite structures, for the other homogeneous equivalence relations one of the two is a finite structure, obliging us to combine results about CSPs of finite structures with those of infinite structures. As it turns out, the two-element case is, not surprisingly, different from the other finite cases and, quite surprisingly, significantly more involved than the other cases.

1.5. **Overview.** This article is organized as follows. Basic notions and definitions, as well as the fundamental facts of the method we are going to use, are provided in Section 2.

Sections 3 to 5 deal with the Henson graphs: Section 3 is complexity-free and investigates the structure of reducts of Henson graphs via polymorphisms and Ramsey theory. In Section 4, we provide hardness and tractability proofs for different classes of reducts. Section 5 contains the proof of Theorem 1.1, and we discuss the complexity classification in more detail, formulating in particular a tractability criterion for CSPs of reducts of Henson graphs.

We then turn to homogeneous equivalence relations in Sections 6 to 8. Similarly to the Henson graphs, the first section (Section 6) is complexity-free and investigates the structure of reducts of homogeneous equivalence relations via polymorphisms and Ramsey theory. Section 7 contains tractability proofs, and Section 8 provides the proof of Theorem 1.2.

We finish this work with further research directions in Section 9.

## 2. Preliminaries

2.1. **General notational conventions.** We use one single symbol, namely $E$, for the edge relation of all homogeneous graphs; since we never consider several such graphs at the same time, this should not cause confusion. Moreover, we use $E$ for the symbol representing the relation $E$, for example in logical formulas. In general, we shall not distinguish between relation symbols and the relations which they denote. The binary relation $N(x, y)$ is defined by the formula $\neg E(x, y) \wedge x \neq y$.

When $E$ is the edge relation of a homogeneous graph whose reflexive closure is an equivalence relation, then we denote this equivalence relation by $Eq$; so $Eq(x, y)$ is defined by the formula $E(x, y) \vee x = y$.

When $t$ is an $n$-tuple, we refer to its entries by $t_1, \ldots, t_n$. When $f : A \to B$ is a function and $C \subseteq A$, we write $f[C]$ for $\{f(a) \mid a \in C\}$.

2.2. **Henson graphs.** For $n \geq 2$, denote the clique on $n$ vertices by $K_n$, and the graph of size $n$ which has no edges by $I_n$. For $n \geq 3$, the graph $(H_n, E)$ is the up to isomorphism unique countable graph which is

- *homogeneous*: any isomorphism between two finite induced subgraphs of $(H_n, E)$ can be extended to an automorphism of $(H_n, E)$, and
- *universal for the class of $K_n$-free graphs*: $(H_n, E)$ contains all finite (in fact, all countable) $K_n$-free graphs as induced subgraphs.

The graph $(H_n, E)$ has the *extension property*: for all disjoint finite $U, U' \subseteq H_n$ such that $U$ is not inducing a copy of $K_{n-1}$ in $(H_n, E)$ there exists $v \in H_n$ such that $v$ is adjacent in $(H_n, E)$ to all members of $U$ and to none in $U'$. Up to isomorphism, there exists a unique countably infinite $K_n$-free graph with this extension property, and hence the property can be used as an alternative definition of $(H_n, E)$.

2.3. **Homogeneous equivalence relations.** For $1 \leq n, s \leq \omega$ the graph $(C_n^s, E)$ is the up to isomorphism unique countable graph whose reflexive closure is an equivalence relation with $n$ classes all of which have size $s$. Clearly, $(C_n^s, E)$ is homogeneous and universal in a similar sense as above.

2.4. **Constraint satisfaction problems.** A first-order $\tau$-formula is called *primitive positive* if it is of the form

$$\exists x_1, \ldots, x_n \, (\psi_1 \wedge \cdots \wedge \psi_m)$$

where the $\psi_i$ are *atomic*, i.e., of the form $y_1 = y_2$ or $R(y_1, \ldots, y_k)$ for a $k$-ary relation symbol $R \in \tau$ and not necessarily distinct variables $y_i$.

Let $\Gamma$ be a structure with a finite relational signature $\tau$. The *constraint satisfaction problem for* $\Gamma$, denoted by $\mathrm{CSP}(\Gamma)$, is the computational problem of deciding for a given primitive positive (pp-) $\tau$-sentence $\phi$ whether $\phi$ is true in $\Gamma$. The following lemma has been first stated in [JCG97] for finite domain structures $\Gamma$ only, but the proof there also works for arbitrary infinite structures.

**Lemma 2.1.** *Let $\Gamma = (D, R_1, \ldots, R_\ell)$ be a relational structure, and let $R$ be a relation that has a primitive positive definition in $\Gamma$. Then $\mathrm{CSP}(\Gamma)$ and $\mathrm{CSP}(D, R, R_1, \ldots, R_\ell)$ are polynomial-time equivalent.*

When a relation $R$ has a primitive positive definition in a structure $\Gamma$, then we also say that $\Gamma$ *pp-defines* $R$. Lemma 2.1 enables the so-called *universal-algebraic approach* to constraint satisfaction, as exposed in the following.

2.5. **The universal-algebraic approach.** We say that a $k$-ary function (also called *operation*) $f \colon D^k \to D$ *preserves* an $m$-ary relation $R \subseteq D^m$ if for all $t_1, \ldots, t_k \in R$ the tuple $f(t_1, \ldots, t_k)$, calculated componentwise, is also contained in $R$. If an operation $f$ does not preserve a relation $R$, we say that $f$ *violates* $R$.

If $f$ preserves all relations of a structure $\Gamma$, we say that $f$ is a *polymorphism* of $\Gamma$, and that $f$ *preserves* $\Gamma$. We write $\mathrm{Pol}(\Gamma)$ for the set of all polymorphisms of $\Gamma$. The unary polymorphisms of $\Gamma$ are just the *endomorphisms* of $\Gamma$, and denoted by $\mathrm{End}(\Gamma)$.

The set of all polymorphisms $\mathrm{Pol}(\Gamma)$ of a relational structure $\Gamma$ forms an algebraic object called a *function clone* (see [Sze86], [GP08]), which is a set of finitary operations defined on a fixed domain that is closed under composition and that contains all projections. Moreover, $\mathrm{Pol}(\Gamma)$ is closed in the *topology of pointwise convergence*, i.e., an $n$-ary function $f$ is contained in $\mathrm{Pol}(\Gamma)$ if and only if for all finite subsets $A$ of $\Gamma^n$ there exists an $n$-ary $g \in \mathrm{Pol}(\Gamma)$ which agrees with $f$ on $A$. We will write $\overline{F}$ for the closure of a set of functions on a fixed domain in this topology; so $\overline{\mathrm{Pol}(\Gamma)} = \mathrm{Pol}(\Gamma)$.

When $\Gamma$ is countable and $\omega$-categorical, then we can characterize primitive positive definable relations via $\mathrm{Pol}(\Gamma)$, as follows.

**Theorem 2.2** (from [BN06])**.** *Let $\Gamma$ be a countable $\omega$-categorical structure. Then the relations preserved by the polymorphisms of $\Gamma$ are precisely those having a primitive positive definition in $\Gamma$.*

Theorem 2.2 and Lemma 2.1 imply that if two countable $\omega$-categorical structures $\Gamma, \Delta$ with finite relational signatures have the same clone of polymorphisms, then their CSPs are polynomial-time equivalent. Moreover, if $\text{Pol}(\Gamma)$ is contained in $\text{Pol}(\Delta)$, then $\text{CSP}(\Gamma)$ is, up to polynomial time, at least as hard as $\text{CSP}(\Delta)$.

Note that the *automorphisms* of a structure $\Gamma$ are just the bijective unary polymorphisms of $\Gamma$ which preserve all relations and their complements; the set of all automorphisms of $\Gamma$ is denoted by $\text{Aut}(\Gamma)$. For every reduct $\Gamma$ of a structure $\Delta$ we have that $\text{Pol}(\Gamma) \supseteq \text{Aut}(\Gamma) \supseteq \text{Aut}(\Delta)$. In particular, this is the case for reducts of the homogeneous graphs $(H_n, E)$ and $(C_n^s, E)$. Conversely, it follows from the $\omega$-categoricity of homogeneous graphs $(D, E)$ (in our case, $D = H_n$ or $D = C_n^s$) that every topologically closed function clone containing $\text{Aut}(D, E)$ is the polymorphism clone of a reduct of $(D, E)$.

When $(D, E)$ is a homogeneous graph, and $F$ is a set of functions and $g$ is a function on the domain $D$, then we say that $F$ *generates* $g$ if $g$ is contained in the smallest topologically closed function clone which contains $F \cup \text{Aut}(D, E)$.

We finish this section with a general lemma that we will refer to on numerous occasions; it allows to restrict the arity of functions violating a relation. For a structure $\Gamma$ and a tuple $t \in \Gamma^k$, the *orbit of* $t$ in $\Gamma$ is the set $\{\alpha(t) \mid \alpha \in \text{Aut}(\Gamma)\}$. We also call this the orbit of $t$ with respect to $\text{Aut}(\Gamma)$.

**Lemma 2.3** (from [BK09]). *Let $\Gamma$ be a relational structure. Suppose that $R \subseteq \Gamma^k$ intersects at most $m$ orbits of $k$-tuples in $\Gamma$. If $\text{Pol}(\Gamma)$ contains a function violating $R$, then $\text{Pol}(\Gamma)$ also contains an $m$-ary operation violating $R$.*

2.6. **Canonical functions.** It will turn out that the polymorphisms relevant for the CSP classification show regular behaviour with respect to the underlying homogeneous graph, in a sense that we are now going to define.

**Definition 2.4.** Let $\Delta$ be a structure. The *type* $\text{tp}(a)$ of an $n$-tuple $a$ of elements in $\Delta$ is the set of first-order formulas with free variables $x_1, \ldots, x_n$ that hold for $a$ in $\Delta$. For structures $\Delta_1, \ldots, \Delta_k$ and tuples $a^1, \ldots, a^n \in \Delta_1 \times \cdots \times \Delta_k$, the type of $(a^1, \ldots, a^n)$ in $\Delta_1 \times \cdots \times \Delta_k$, denoted by $\text{tp}(a^1, \ldots, a^n)$, is the $k$-tuple containing the types of $(a_i^1, \ldots, a_i^n)$ in $\Delta_i$ for each $1 \leq i \leq k$.

We bring to the reader's attention the well-known fact that in homogeneous structures, in particular in $(H_n, E)$ and $(C_n^k, E)$, two $n$-tuples have the same type if and only if their orbits coincide.

**Definition 2.5.** Let $\Delta_1, \ldots, \Delta_k$ and $\Lambda$ be structures. A *behaviour* $B$ between $\Delta_1, \ldots, \Delta_k$ and $\Lambda$ is a partial function from the types over $\Delta_1, \ldots, \Delta_k$ to the types over $\Lambda$. Pairs $(s, t)$ with $B(s) = t$ are also called *type conditions.* We say that a function $f \colon \Delta_1 \times \cdots \times \Delta_k \to \Lambda$ *satisfies the behaviour* $B$ if whenever $B(s) = t$ and $(a^1, \ldots, a^n)$ has type $s$ in $\Delta_1, \ldots, \Delta_k$, then the $n$-tuple $(f(a_1^1, \ldots, a_k^1), \ldots, f(a_1^n, \ldots, a_k^n))$ has type $t$ in $\Lambda$. A function $f \colon \Delta_1 \times \cdots \times \Delta_k \to \Lambda$ is *canonical* if it satisfies a behaviour which is a total function from the types over $\Delta_1, \ldots, \Delta_k$ to the types over $\Lambda$.

We remark that since our structures are homogeneous and have only binary relations, the type of an $n$-tuple $a$ is determined by its binary subtypes, i.e., the types of the pairs $(a_i, a_j)$, where $1 \leq i, j \leq n$. In other words, the type of $a$ is determined by which of its components are equal, and between which of its components there is an edge. Therefore, a function

$f\colon (H_n,E)^k \to (H_n,E)$ or $f\colon (C_n^s,E)^k \to (C_n^s,E)$ is canonical iff it satisfies the condition of the definition for types of 2-tuples.

To provide immediate examples for these notions, we now define some behaviours that will appear in our proof as well as in the precise CSP classification. For $m$-ary relations $R_1,\dots,R_k$ over a set $D$, we will in the following write $R_1\cdots R_k$ for the $m$-ary relation on $D$ that holds between $k$-tuples $x^1,\dots,x^m \in D^k$ iff $R_i(x_i^1,\dots,x_i^m)$ holds for all $1 \leq i \leq k$. We start with behaviours of binary injective functions on $(H_n,E)$.

**Definition 2.6.** We say that a binary injective operation $f\colon H_n^2 \to H_n$ is

- *balanced in the first argument* if for all $u,v \in H_n^2$ we have that E=$(u,v)$ implies $E(f(u),f(v))$ and N=$(u,v)$ implies $N(f(u),f(v))$;
- *balanced in the second argument* if $(x,y)\mapsto f(y,x)$ is balanced in the first argument;
- *balanced* if $f$ is balanced in both arguments, and *unbalanced* otherwise;
- *$E$-dominated ($N$-dominated) in the first argument* if for all $u,v \in H_n^2$ with $\neq$=$(u,v)$ we have that $E(f(u),f(v))$ $(N(f(u),f(v)))$;
- *$E$-dominated ($N$-dominated) in the second argument* if $(x,y)\mapsto f(y,x)$ is $E$-dominated ($N$-dominated) in the first argument;
- *$E$-dominated ($N$-dominated)* if it is $E$-dominated ($N$-dominated) in both arguments;
- *of behaviour* min if for all $u,v \in H_n^2$ with $\neq\neq(u,v)$ we have $E(f(u),f(v))$ if and only if EE$(u,v)$;
- *of behaviour* max if for all $u,v \in H_n^2$ with $\neq\neq(u,v)$ we have $N(f(u),f(v))$ if and only if NN$(u,v)$;
- *of behaviour* $p_1$ if for all $u,v \in H_n^2$ with $\neq\neq(u,v)$ we have $E(f(u),f(v))$ if and only if $E(u_1,v_1)$;
- *of behaviour* $p_2$ if $(x,y)\mapsto f(y,x)$ is of behaviour $p_1$;
- *of behaviour projection* if it is of behaviour $p_1$ or $p_2$.

Each of these properties describes the set of all functions of a certain behaviour. We explain this for the first item defining functions which are balanced in the first argument, which can be expressed by the behaviour consisting of the following two type conditions. Let $(u,v)$ be any pair of elements $u,v \in H_n \times H_n$ such that E=$(u,v)$, and let $s$ be the type of the pair $(u,v)$ in $(H_n,E)\times(H_n,E)$. Let $x,y \in H_n$ satisfy $E(x,y)$, and let $t$ be the type of $(x,y)$ in $(H_n,E)$. Then the first type condition is $(s,t)$. Now let $s'$ be the type in $(H_n,E)\times(H_n,E)$ of any pair $(u,v)$, where $u,v \in H_n\times H_n$ satisfy N=$(u,v)$, and let $t'$ be the type in $(H_n,E)$ of any $x,y \in H_n$ with $N(x,y)$. The second type condition is $(s',t')$.

To justify the less obvious names of some of the above behaviours, we would like to point out that a binary injection of behaviour max is reminiscent of the Boolean maximum function on $\{0,1\}$, where $E$ takes the role of 1 and $N$ the role of 0: for $u,v \in H_n^2$ with $\neq\neq(u,v)$, we have $E(f(u),f(v))$ if $u,v$ are connected by an edge in at least one coordinate, and $N(f(u),f(v))$ otherwise. The names "min" and "projection" can be explained similarly.

**Definition 2.7.** An injective ternary function $f\colon H_n^3 \to H_n$ is of behaviour

- *majority* if for all $u,v \in H_n^3$ with $\neq\neq\neq(u,v)$ we have that $E(f(u),f(v))$ if and only if EEE$(u,v)$, EEN$(u,v)$, ENE$(u,v)$, or NEE$(u,v)$;
- *minority* if for all $u,v \in H_n^3$ with $\neq\neq\neq(u,v)$ we have $E(f(u),f(v))$ if and only if EEE$(u,v)$, NNE$(u,v)$, NEN$(u,v)$, or ENN$(u,v)$.

Consider the first item of Definition 2.6. In the sequel, it will be convenient to express the fact that E=$(u,v)$ implies $E(f(u),f(v))$ by writing $f(E,=)=E$. Similarly, the majority behaviour in Definition 2.7 can be expressed by writing $f(E,E,E)=f(E,E,N)=f(E,N,E)=f(N,E,E)=E$ and $f(N,N,N)=f(E,N,N)=f(N,E,N)=f(N,N,E)=N$. This notation, which we are going to use for other type conditions as well, is justified by the fact that the type conditions satisfied by a function induce a partial function from types to types, and that in the case of homogeneous graphs, all that matters is the three types of pairs, given by the relations $E$, $N$, and $=$.

**Definition 2.8.** A ternary canonical injection $f\colon H_n^3 \to H_n$ is *hyperplanely of behaviour projection* iff the functions $(u,v)\mapsto f(c,u,v)$, $(u,v)\mapsto f(u,c,v)$, and $(u,v)\mapsto f(u,v,c)$ are of behavior projection for all $c \in H_n^3$. Similarly other hyperplane behaviours, such as hyperplanely $E$-dominated, are defined.

Note that hyperplane behaviours are defined by conditions for the type functions $f(=,\cdot,\cdot)$, $f(\cdot,=,\cdot)$, and $f(\cdot,\cdot,=)$. For example, hyperplanely $E$-dominated precisely means that

$$f(=,=,\neq)=f(=,\neq,=)=f(\neq,=,=)=E\,.$$

We have not defined any behaviours on the graphs $(C_n^s,E)$, but some of the behaviours for the Henson graphs above, such as behaviour minority, will also play a role for those structures (with the same definition).

2.7. **Ramsey theory.** The next proposition, which is an instance of more general statements from [BP11, BPT13], provides us with the main combinatorial tool for analyzing functions on Henson graphs. Equip $H_n$ with a total order $\prec$ in such a way that $(H_n,E,\prec)$ is homogeneous; up to isomorphism, there is only one such structure $(H_n,E,\prec)$, called the *random ordered $K_n$-free graph*. The order $(H_n,\prec)$ is then isomorphic to the order $(\mathbb{Q},<)$ of the rationals. By [NR89], $(H_n,E,\prec)$ is a *Ramsey structure*, which implies the following proposition – for more details, see the survey [BP11].

**Proposition 2.9.** *Let $f\colon H_n^k \to H_n$, let $c_1,\dots,c_r \in H_n$, and let $(H_n,E,\prec,c_1,\dots,c_r)$ be the expansion of $(H_n,E,\prec)$ by the constants $c_1,\dots,c_r$. Then*

$$\overline{\{\alpha\circ f\circ(\beta_1,\dots,b_r)\mid \alpha\in \operatorname{Aut}(H_n,E,\prec),\ \beta_1,\dots,\beta_r\in\operatorname{Aut}(H_n,E,\prec,c_1,\dots,c_r)\}}$$

*contains a function $g$ such that*

- *$g$ is canonical as a function from $(H_n,E,\prec,c_1,\dots,c_r)$ to $(H_n,E,\prec)$;*
- *$g$ agrees with $f$ on $\{c_1,\dots,c_r\}^k$.*

*In particular, $f$ generates a function $g$ with these properties.*

Similarly, Ramsey theory allows us to produce canonical functions on $(C_n^s,E)$, expanded with a certain linear order. Equip $C_n^s$ with a total order $\prec$ so that the equivalence classes of $(C_n^s,Eq)$ are *convex* with respect to $\prec$, i.e., whenever $Eq(u,v)$ holds and $u\prec w\prec v$, then $Eq(u,w)$. Moreover, in the case where $s=\omega$, we require the order to be isomorphic to the order of the rational numbers on each equivalence class, and in case where $n=\omega$, we require the order to be isomorphic to the order of the rational numbers between the classes (note that we required convexity). If $n$ is finite, let $P_1,\dots,P_n$ denote predicates such that $P_i$ contains precisely the elements in the $i$-th equivalence relation with respect to $\prec$. Then the structure $(C_n^s,E,\prec,P_1,\dots,P_n)$ is homogeneous, and we have that the precise statement of Proposition 2.9 holds for this structure as well.

If $s$ is finite, we add $s$ unary predicates $Q_1, \ldots, Q_s$ where $Q_i$ contains precisely the $i$-th element for each equivalence class with respect to the order $\prec$. Then $(C_n^s, E, \prec, Q_1, \ldots, Q_s)$ is homogeneous, and again it follows from standard arguments in Ramsey theory that it is a Ramsey structure, so that Proposition 2.9 can be applied also in this case.

## 3. Polymorphisms over Henson graphs

We investigate polymorphisms of reducts of $(H_n, E)$. We start with unary polymorphisms in Section 3.1, obtaining that we can assume that the relations $E$ and $N$ are pp-definable in our reducts. We then turn to binary polymorphisms in Section 3.2, obtaining Proposition 3.10 telling us that we may further assume the existence of a binary injective polymorphism. Building on the results of those sections, we show in Section 3.3 via an analysis of ternary polymorphisms that for any reduct which pp-defines the relations $E$ and $N$, either the polymorphisms preserve a certain relation $H$, or there is a polymorphism of behaviour min (Proposition 3.12).

3.1. **The unary case: model-complete cores.** A countable $\omega$-categorical structure $\Delta$ is called a *model-complete core* if $\mathrm{Aut}(\Delta)$ is dense in $\mathrm{End}(\Delta)$, or equivalently, every endomorphism of $\Delta$ is an elementary self-embedding, i.e., preserves all first-order formulas over $\Delta$. Every countable $\omega$-categorical structure $\Gamma$ is *homomorphically equivalent* to an up to isomorphism unique $\omega$-categorical model-complete core $\Delta$, that is, there exists homomorphisms from $\Gamma$ into $\Delta$ and vice-versa [Bod07]. Since the CSPs of homomorphically equivalent structures are equal, it has proven fruitful in classification projects to always work with model-complete cores. The following proposition essentially calculates the model-complete cores of the reducts of Henson graphs.

**Proposition 3.1.** *Let $\Gamma$ be a reduct of $(H_n, E)$. Then either $\mathrm{End}(\Gamma)$ contains a function whose image induces an independent set, or $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(\Gamma)} = \overline{\mathrm{Aut}(H_n, E)}$.*

*Proof.* Assume that $\mathrm{End}(\Gamma) \neq \overline{\mathrm{Aut}(H_n, E)}$. Then there exists an $f \in \mathrm{End}(\Gamma)$ which violates $E$ or $N$. If $f$ violated $N$ but not $E$, then there would be a copy of $K_n$ in the range of $f$, a contradiction. Thus, we may assume that $f$ violates $E$, i.e., there exist $(u, v) \in E$ such that $(f(u), f(v)) \in N$. By Proposition 2.9, $f$ generates a canonical function $g\colon (H_n, E, \prec, u, v) \to (H_n, E, \prec)$ such that $f(u) = g(u)$ and $f(v) = g(v)$; in fact, since $f$ is unary, we can disregard the order $\prec$ and assume that $g$ is canonical as a function from $(H_n, E, u, v)$ to $(H_n, E)$ [Pon11, Proposition 3.7].

Let $U_{uv} := \{x \in H_n \mid E(u, x) \wedge E(v, x)\}$, $U_{u\overline{v}} := \{x \in H_n \mid E(u, x) \wedge N(v, x)\}$, $U_{\overline{u}v} := \{x \in H_n \mid N(u, x) \wedge E(v, x)\}$ and $U_{\overline{uv}} := \{x \in H_n \mid N(u, x) \wedge N(v, x)\}$. As all four of these sets contain a copy of $I_n$, $N$ is preserved by $g$ on any of these sets.

If $g$ violates $E$ on $U_{\overline{uv}}$, then it generates a function whose image is an independent set. Thus, we may assume that $g$ preserves $E$ on $U_{\overline{uv}}$.

Then $g$ preserves $N$ between $U_{\overline{uv}}$ and any other orbit $X$ of $\mathrm{Aut}(H_n, E, u, v)$, as otherwise the image of an $n$-element induced subgraph of $(H_n, E)$ which consists of an isolated point in $X$ and a copy of $K_{n-1}$ in $U_{\overline{uv}}$ would be isomorphic to $K_n$.

Assume that $g$ violates $E$ between $U_{\overline{uv}}$ and another orbit of $\mathrm{Aut}(H_n, E, u, v)$. Let $A \subseteq H_n$ be finite with an edge $(x, y)$ in $A$. Then there exists an $\alpha \in \mathrm{Aut}(H_n, E)$ such that $\alpha(x) \in X$ and $\alpha[A \setminus \{x\}] \subseteq U_{\overline{uv}}$. The function $(g \circ \alpha) \restriction_A$ preserves $N$, and it maps $(x, y)$ to a non-edge. By an iterative application of this step we can systematically delete all edges of $A$. Hence, by topological closure, $g$ generates a function whose image is an independent set. Thus, we may assume that $g$ preserves $E$ between $U_{\overline{uv}}$ and any other orbit of $\mathrm{Aut}(H_n, E, u, v)$.

Let $X$ and $Y$ be infinite orbits of $\mathrm{Aut}(H_n, E, u, v)$, and assume that $g$ violates $N$ between $X$ and $Y$. There exist vertices $x \in X$ and $y \in Y$, and a copy of $K_{n-2}$ in $U_{\overline{uv}}$ such that $(x, y)$ is the only non-edge in the graph induced by these $n$ vertices. Thus, the $g$-image of this $n$-element set induces a copy of $K_n$, a contradiction. Hence, we may assume that $g$ preserves $N$ on $H_n \setminus \{u, v\}$.

If $g$ violates $E$ on $H_n \setminus \{u, v\}$, then we can systematically delete the edges of any finite subgraph of $(H_n, E)$, and conclude that $g$ generates a function whose image is an independent set. Thus, we may assume that $g$ preserves $E$ on $H_n \setminus \{u, v\}$.

Assume that $g$ violates $E$ between $u$ and $U_{u\overline{v}}$. Given any finite $A \subseteq H_n$ with a vertex $x \in A$, there exists a $\beta \in \mathrm{Aut}(H_n, E)$ such that $\beta(x) = u$ and $\beta[A \setminus \{x\}] \subseteq U_{u\overline{v}} \cup U_{\overline{uv}}$. Hence, $(g \circ \beta) \restriction_A$ preserves $N$, and it maps edges from $x$ to non-edges. Thus, we can systematically delete the edges of $A$, and consequently, $g$ generates a function whose image is an independent set. Hence, we may assume that $g$ preserves $E$ between $u$ and $U_{u\overline{v}}$.

There exists a vertex $x \in U_{\overline{u}v}$ and a copy of $K_{n-2}$ in $U_{u\overline{v}}$ such that $(x, u)$ is the only non-edge in the graph induced by these $n-1$ vertices and $u$. Thus, if $g$ violates $N$ between $\{u\}$ and $U_{\overline{u}v}$, then the $g$-image of this $n$-element set induces a copy of $K_n$, a contradiction. Hence, $g$ preserves $N$ between $\{u\}$ and $U_{\overline{u}v}$.

Similarly, we may assume that $g$ preserves $N$ between $v$ and $U_{u\overline{v}}$. Thus, $g$ preserves $N$. As $g$ deletes the edge between $u$ and $v$, we can systematically delete the edges of any finite subgraph of $(H_n, E)$. Hence, $g$ generates a function whose image is an independent set. ■

In the first case of Proposition 3.1, the model-complete core of the reduct is in fact a reduct of equality. Since the CSPs of reducts of equality have been classified [BK08], the we do not have to consider any further reducts with an endomorphism whose image induces an independent set.

**Lemma 3.2.** *Let $\Gamma$ be a reduct of $(H_n, E)$, and assume that $\mathrm{End}(\Gamma)$ contains a function whose image is an independent set. Then $\Gamma$ is homomorphically equivalent to a reduct of $(H_n, =)$.*

*Proof.* Trivial. ■

In the second case of Proposition 3.1, it turns out that all polymorphisms preserve the relations $E$, $N$, and $\neq$, by the following lemma and Theorem 2.2.

**Lemma 3.3.** *Let $\Gamma$ be such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(H_n, E)}$. Then $E$, $N$, and $\neq$ have primitive positive definitions in $\Gamma$.*

*Proof.* Since $E$ and $N$ are orbits of pairs with respect to $\mathrm{Aut}(H_n, E)$, the primitive positive definability of $E$ and $N$ is an immediate consequence of Theorem 2.2 and Lemma 2.3. The primitive positive formula $\exists z(E(x, z) \wedge N(y, z))$ defines $x \neq y$. ■

Before moving on to binary polymorphisms, we observe the following corollary of Proposition 3.1, first mentioned in [Tho91].

**Corollary 3.4.** *For every $n \geq 3$, the permutation group $\mathrm{Aut}(H_n, E)$ is a maximal closed subgroup of the full symmetric group on $H_n$, i.e., every closed subgroup of the full symmetric group containing $\mathrm{Aut}(H_n, E)$ either equals $\mathrm{Aut}(H_n, E)$ or the full symmetric group.*

*Proof.* The closure $\overline{G}$ of any supergroup $G$ of $\mathrm{Aut}(H_n, E)$ in the set of all unary functions on $H_n$ is a closed transformation monoid, and hence the endomorphism monoid of a reduct of $(H_n, E)$ (cf. for example [BP14]). By Proposition 3.1, it either contains a function whose image induces an independent set, or it equals $\overline{\mathrm{Aut}(H_n, E)}$. In the first case, it is easy to see that $G$ equals the full symmetric group, and in the latter case, $G = \mathrm{Aut}(H_n, E)$. ■

We remark that the automorphism group of the Rado graph has five closed supergroups [Tho91], which leads to more cases in the corresponding CSP classification in [BP15a].

3.2. **Binary polymorphisms.** We investigate binary functions preserving $E$, $N$, and $\neq$. Every unary function gives rise to a binary function by adding a dummy variable; the following definition rules out such "improper" higher-arity functions.

**Definition 3.5.** A finitary operation $f(x_1, \ldots, x_n)$ on a set is *essential* if it depends on more than one of its variables $x_i$.

**Lemma 3.6.** *Let $f\colon H_n^2 \to H_n$ be a binary essential function that preserves $E$, $N$, and $\neq$. Then $f$ generates a binary injection.*

*Proof.* We call sets of the form $(\cdot, x) \subseteq H_n^2$ for any $x \in H_n$ *vertical lines*, and those of the form $(x, \cdot) \subseteq H_n^2$ *horizontal lines.* A line is *good* if the restriction of $f$ to it is injective. A point $(x, y) \in H_n^2$ is *v-good* if $f(x, y) \neq f(x, z)$ for all $y \neq z$. We follow the strategy of the proof of [BP14, Proposition 38]. So let $\Delta$ be the structure with underlying set $H_n$ and whose relations are those preserved by $\{f\} \cup \mathrm{Aut}(H_n, E)$. In particular, $E$, $N$, and $\neq$ are relations of $\Delta$. Then it is well-known (cf. [Sze86]) that $\mathrm{Pol}(\Delta)$ consists precisely of the functions generated by $f$, and so we need to show that $\mathrm{Pol}(\Delta)$ contains a binary injection. By [BP14, Lemma 42] it is enough to show that for all primitive positive formulas $\phi$ over $\Delta$ we have that whenever $\phi \wedge x \neq y$ and $\phi \wedge s \neq t$ are satisfiable in $\Delta$, then the formula $\phi \wedge x \neq y \wedge s \neq t$ is also satisfiable in $\Delta$. Still following the proof of [BP14, Proposition 38] it is enough to show the following claim.

**Claim.** Given two 4-tuples $a = (x, y, z, z)$ and $b = (p, p, q, r)$ in $H_n^4$ such that $x \neq y$ and $q \neq r$, there exist 4-tuples $a'$ and $b'$ of the same type as $a$ and $b$ in $(H_n, E)$ such that $f(a', b')$ is a 4-tuple whose first two coordinates are different and whose last two coordinates are different.

*Proof of Claim.* We may assume that $x \neq z$ and $p \neq q$. We may also assume that $f$ itself is not a binary injection.

Assume without loss of generality that there exist $u_1 \neq u_2, v \in H_n$ such that $f(u_1, v) = f(u_2, v)$. In particular, as $f$ preserves $\neq$, the points $(u_1, v)$ and $(u_2, v)$ are v-good. First set the values $q', z'$ such that $(z', q')$ is v-good. We may assume that for any $x', y' \in H_n$ with $\mathrm{tp}(x', y', z') = \mathrm{tp}(x, y, z)$ and $\mathrm{tp}(p', q') = \mathrm{tp}(p, q)$ we have $f(x', p') = f(y', p')$, otherwise the tuples $a' = (x', y', z', z')$ and $b' = (p', p', q', r')$ are appropriate with any $r' \in H_n$ with $\mathrm{tp}(p', q', r') = \mathrm{tp}(p, q, r)$. Hence, as $f$ preserves $\neq$, all the points $(x', p')$ with $\mathrm{tp}(x', z') = \mathrm{tp}(x, z)$ and $\mathrm{tp}(p', q') = \mathrm{tp}(p, q)$ are v-good. So we obtained that whenever the point $(s, t)$ is v-good, and $s_0, t_0 \in H_n$ are such that $\mathrm{tp}(s, s_0) = \mathrm{tp}(x, z)$ and $\mathrm{tp}(t, t_0) = \mathrm{tp}(p, q)$, then $(s_0, t_0)$ is also v-good, or otherwise we are done. We show that whatever the types $Q_1 = \mathrm{tp}(x, z)$ and $Q_2 = \mathrm{tp}(p, q)$ are, we can reach any point $(s_4, t_4)$ in $H_n^2$ from a given v-good point $(s_0, t_0)$ by at most four such steps. To see this, note that $Q_1$ and $Q_2$ are different from $=$ by assumption. Now let $s_1, s_2, s_3, t_1, t_2, t_3$ be such that

- $s_0, s_1, s_2, s_3, s_4$ are pairwise different except that $s_0 = s_4$ is possible, and

- $t_0, t_1, t_2, t_3, t_4$ are pairwise different except that $t_0 = t_4$ is possible, and
- $(s_0, s_1), (s_1, s_2), (s_2, s_3), (s_3, s_4) \in Q_1$ and all other pairs $(s_i, s_j)$ are in $N$ except that $s_0 = s_4$ is possible, and
- $(t_0, t_1), (t_1, t_2), (t_2, t_3), (t_3, t_4) \in Q_2$ and all other pairs $(t_i, t_j)$ are in $N$ except that $t_0 = t_4$ is possible.

These rules are not in contradiction with the extension property of $(H_n, E)$, thus such vertices exist, and we can propagate the v-good property from $(s_0, t_0)$ to $(s_4, t_4)$. Hence, every point is v-good, and hence, every vertical line is good, or we are done. If $f(u_1, v) = f(u_2, v)$ for all $u_1, u_2, v \in H_n$ with $\text{tp}(u_1, u_2) = \text{tp}(x, y)$, then $f$ would be essentially unary, since $(H_n, E)$ and its complement have diameter 2. As $f$ is a binary essential function, we can choose $x', y', p' \in H_n$ such that $\text{tp}(x', y') = \text{tp}(x, y)$ and $f(x', p') \neq f(y', p')$. By choosing points $z', q', r' \in H_n$ such that $\text{tp}(x', y', z') = \text{tp}(x, y, z)$ and $\text{tp}(p', q', r') = \text{tp}(p, q, r)$ the tuples $a' = (x', y', z', z')$ and $b' = (p', p', q', r')$ are appropriate. ■

**Lemma 3.7.** *Let $k \geq 2$. Every essential function $f \colon H_n^k \to H_n$ that preserves $E$, $N$, and $\neq$ generates a binary injection.*

*Proof.* By [BP14, Lemma 40], every essential operation generates a binary essential operation over the random graph; the very same proof works for the Henson graphs. Therefore, we may assume that $f$ itself is binary. The assertion now follows from Lemma 3.6. ■

**Lemma 3.8.** *Let $f \colon H_n^2 \to H_n$ be a function of behaviour* min *that preserves $E$ and $N$. Then $f$ generates a binary function of behaviour* min *that is $N$-dominated.*

*Proof.* By Proposition 2.9 we have that $f$ generates a binary injection $g$ that is canonical as a function $(H_n, E, \prec)^2 \to (H_n, E, \prec)$; from the first (and stronger) statement, and since composing functions of a certain behaviour with automorphisms yields functions of the same behaviour, we conclude that $g$ can be assumed to have behaviour min.

We now refer to the proof of Theorem 57 in [BP14], observing that the calculus for canonical functions on the Henson graphs is the same as the calculus on the random graph. More precisely, when we compose canonical functions, then we obtain a canonical function, and its behaviour can be calculated by composing the respective behaviours of the composing functions; this is independent of whether the underlying graph is the random graph or a Henson graph. By that theorem, $g$ generates an operation of behaviour min which is $N$-dominated, or one of behaviour min which is balanced. However, binary balanced injections that preserve $E$ do not exist over $(H_n, E)$, as they would introduce a copy of $K_n$. To see this, let $x_1, \ldots, x_{n-1} \in H_n$ be pairwise adjacent vertices in $H_n$. Then $g(x_1, x_1), \ldots, g(x_{n-1}, x_{n-1})$ are pairwise adjacent since $g$ preserves $E$. For the same reason, $E(g(x_i, x_i), g(x_1, x_{n-1}))$ if $i$ is distinct from 1 and from $n-1$. Finally, if $g$ is balanced then $E(g(x_1, x_1), g(x_1, x_{n-1}))$ and $E(g(x_{n-1}, x_{n-1}), g(x_1, x_{n-1}))$. This is in contradiction to the assumption that $(H_n, E)$ is $K_n$-free. We conclude that $g$, and hence also $f$, generates an operation of behaviour min which is $N$-dominated. ■

**Lemma 3.9.** *Let $k \geq 2$, and let $f \colon H_n^k \to H_n$ be an essential function that preserves $E$, $N$, and $\neq$. Then $f$ generates one of the following binary canonical injections:*

- *of behaviour* min *and $N$-dominated*
- *of behaviour $p_1$, balanced in the first, and $N$-dominated in the second argument.*

*Proof.* By Lemma 3.7 we may assume that $k = 2$ and that $f$ is injective. By Proposition 2.9 we have that $f$ generates a binary injection $g$ that is canonical as a $(H_n, E, \prec)^2 \to (H_n, E, \prec)$

function. We can now refer to Theorem 24 in [BP15a] (itself from [BP14]) since the calculus for canonical functions on the Henson graphs is the same as the calculus on the random graph, and conclude that $f$ generates a function of one of the following behaviours.

(1) a canonical injection of behaviour $p_1$ which is balanced;
(2) a canonical injection of behaviour max which is balanced;
(3) a canonical injection of behaviour $p_1$ which is $E$-dominated;
(4) a canonical injection of behaviour max which is $E$-dominated;
(5) a canonical injection of behaviour $p_1$ which is balanced in the first and $E$-dominated in the second argument;
(6) a canonical injection of behaviour min which is balanced;
(7) a canonical injection of behaviour $p_1$ which is $N$-dominated;
(8) a canonical injection of behaviour min which is $N$-dominated;
(9) a canonical injection of behaviour $p_1$ which is balanced in the first and $N$-dominated in the second argument.

For the $K_n$-free graphs, none of the behaviours max, $E$-dominated, or balanced in both arguments can occur since they would introduce a $K_n$. So we are left with items (7) and (8), proving the lemma. ■

We conclude this section by summarising the results we have so far.

**Proposition 3.10.** *Let* $\Gamma$ *be a reduct of* $(H_n, E)$*, where* $n \geq 3$*. Then either*

(1) $\Gamma$ *is homomorphically equivalent to a reduct of* $(H_n, =)$*, or*
(2) $\Gamma$ *pp-defines* $E$*,* $N$*, and* $\neq$*.*

*In the latter case we have that either*

(2a) *every function in* Pol($\Gamma$) *is essentially unary, or*
(2b) Pol($\Gamma$) *contains one of the two binary canonical injections of Lemma 3.9.*

Note that if item (1) holds then CSP($\Gamma$) is either in P or NP-complete [BK08], and if item (2a) holds then CSP($\Gamma$) is NP-complete (Theorem 10 in [BCKvO09]). In case (2b), when Pol($\Gamma$) contains a binary canonical injection of behaviour min which is hyperplanely $N$-constant then CSP($\Gamma$) is in P, as we will show in Section 4.2. It thus remains to further consider the second case of Lemma 3.9. This is the content of the following section.

3.3. **The relation $H$.** We investigate Case (2b) of Proposition 3.10. The following relation characterizes the NP-complete cases in this situation.

**Definition 3.11.** We define a 6-ary relation $H(x_1, y_1, x_2, y_2, x_3, y_3)$ on $H_n$ by

$$\begin{aligned}
&\bigwedge_{i,j\in\{1,2,3\},i\neq j,u\in\{x_i,y_i\},v\in\{x_j,y_j\}} N(u,v)\\
&\wedge \big((E(x_1,y_1)\wedge N(x_2,y_2)\wedge N(x_3,y_3))\\
&\quad\vee\ (N(x_1,y_1)\wedge E(x_2,y_2)\wedge N(x_3,y_3))\\
&\quad\vee\ (N(x_1,y_1)\wedge N(x_2,y_2)\wedge E(x_3,y_3))\big)\ .
\end{aligned}$$

Our goal for this section is to prove the following proposition, which states that if $\Gamma$ is a reduct of $(H_n, E)$ with $E$ and $N$ primitive positive definable in $\Gamma$, then either $H$ has a primitive positive definition in $\Gamma$, in which case CSP($\Gamma$) is NP-complete, or Pol($\Gamma$) has a certain canonical polymorphism which will imply tractability of the CSP. NP-completeness and tractability for those cases will be shown in Section 4.

**Proposition 3.12.** *Let $\Gamma$ be a reduct of $(H_n, E)$ with $E$ and $N$ primitive positive definable in $\Gamma$. Then at least one of the following holds:*

(a) *There is a primitive positive definition of $H$ in $\Gamma$.*
(b) $\mathrm{Pol}(\Gamma)$ *contains a canonical binary injection of behaviour* min.

In the remainder of this section we prove that if $\Gamma$ is a reduct of $(H_n, E)$ with $E$ and $N$ primitive positive definable in $\Gamma$ such that there is no primitive positive definition of $H$ in $\Gamma$, then the second case of Proposition 3.12 applies. Note that under this assumption, $\Gamma$ has a polymorphism that violates $H$.

3.3.1. *First arity reduction: down to ternary.* Our first goal is to prove that we can assume said polymorphism to be a ternary injection.

**Lemma 3.13.** *Let $f\colon H_n^k \to H_n$ be an operation which preserves $E$ and $N$ and violates $H$. Then $f$ generates a ternary injection which shares the same properties.*

*Proof.* Since the relation $H$ consists of three orbits of 6-tuples with respect to $(H_n, E)$, Lemma 2.3 shows that $f$ generates a ternary function that violates $H$, and hence we can assume that $f$ itself is at most ternary. Then $f$ must certainly be essential, since essentially unary operations that preserve $E$ and $N$ also preserve $H$. Applying Proposition 3.10, we get that $f$ generates a binary canonical injection $g$ of type min or $p_1$. In the case of min we are done, since the ternary injection $g(g(x, y), z)$ violates $H$. Now consider the case where $g$ is of type $p_1$. Then

$$h(x, y, z) := g(g(g(f(x, y, z), x), y), z)$$

is injective and violates $H$ – the latter can easily be verified combining the facts that $f$ violates $H$, $g$ is of type $p_1$, and all tuples in $H$ have pairwise distinct entries. ■

3.3.2. *Second arity reduction: down to binary.* Still with the ultimate goal of producing a binary canonical polymorphism of behaviour min, we now show that under the assumptions of the preceding lemma, we also have a binary polymorphism which is not of behaviour projection. We begin by ruling out some ternary behaviours which do play a role on the random graph.

**Lemma 3.14.** *There are no ternary injections of behaviour majority or minority on $(H_n, E)$.*

*Proof.* These could introduce a $K_n$ in the $K_n$-free graph $(H_n, E)$. ■

**Proposition 3.15.** *Let $f\colon H_n^k \to H_n$ be an operation that preserves $E$ and $N$ and violates $H$. Then $f$ generates a binary injection which is not of behaviour projection.*

*Proof.* By Lemma 3.13, we can assume that $f$ is a ternary injection. Because $f$ violates $H$, there are $x^1, x^2, x^3 \in H$ such that $f(x^1, x^2, x^3) \notin H$. In the following, we will write $x_i := (x_i^1, x_i^2, x_i^3)$ for $1 \leq i \leq 6$. So $(f(x_1), \ldots, f(x_6)) \notin H$.

If there exists $\alpha \in \mathrm{Aut}(H_n, E)$ such that $\alpha(x^i) = x^j$ for $1 \leq i \neq j \leq 3$, then $f$ generates a binary injection that still violates $H$; this function cannot be of behaviour projection, and so the proposition follows.

Assuming there is no such automorphism, we will derive a contradiction. In this situation, by permuting arguments of $f$ if necessary, we can assume without loss of generality that

$$\mathrm{ENN}(x_1, x_2),\ \mathrm{NEN}(x_3, x_4),\ \text{and}\ \mathrm{NNE}(x_5, x_6).$$

We set

$$S := \{y \in H_n^3 \mid \mathrm{NNN}(x_i, y) \text{ for all } 1 \leq i \leq 6\}\ .$$

Consider the binary relations $Q_1Q_2Q_3$ on $H_n^3$, where $Q_i \in \{E, N\}$ for $1 \leq i \leq 3$. We claim that for each such relation $Q_1Q_2Q_3$, whether $E(f(u), f(v))$ or $N(f(u), f(v))$ holds for $u, v \in S$ with $Q_1Q_2Q_3(u, v)$ does not depend on $u, v$; that is, whenever $u, v, u', v' \in S$ satisfy $Q_1Q_2Q_3(u, v)$ and $Q_1Q_2Q_3(u', v')$, then $E(f(u), f(v))$ if and only if $E(f(u'), f(v'))$. Note that this is another way of saying that $f$ satisfies some type conditions on $S$. We go through all possibilities of $Q_1Q_2Q_3$.

(1) $Q_1Q_2Q_3 = \mathrm{ENN}$. Let $\alpha \in \mathrm{Aut}(H_n, E)$ be such that $(x_1^2, x_2^2, u_2, v_2)$ is mapped to $(x_1^3, x_2^3, u_3, v_3)$; such an automorphism exists since

$$\mathrm{NNN}(x_1, u), \mathrm{NNN}(x_1, v), \mathrm{NNN}(x_2, u), \mathrm{NNN}(x_2, v)$$

hold, and since $(x_1^2, x_2^2)$ has the same type as $(x_1^3, x_2^3)$, and $(u_2, v_2)$ has the same type as $(u_3, v_3)$. We are done if the operation $g$ defined by $g(x, y) := f(x, y, \alpha(y))$ is not of type projection. Otherwise, $E(g(u_1, u_2), g(v_1, v_2))$ iff $E(g(x_1^1, x_1^2), g(x_2^1, x_2^2))$. Combining this with the equations $(f(u), f(v)) = (g(u_1, u_2), g(v_1, v_2))$ and $(g(x_1^1, x_1^2), g(x_2^1, x_2^2)) = (f(x_1), f(x_2))$, we get that $E(f(u), f(v))$ iff $E(f(x_1), f(x_2))$, and so our claim holds for this case.

(2) $Q_1Q_2Q_3 = \mathrm{NEN}$ or $Q_1Q_2Q_3 = \mathrm{NNE}$. These cases are analogous to the previous case.

(3) $Q_1Q_2Q_3 = \mathrm{NEE}$. Let $\alpha$ be defined as in the first case. Reasoning as above, if the operation defined by $f(x, y, \alpha(y))$ is not of type projection, then one gets that $E(f(u), f(v))$ iff $N(f(x_1), f(x_2))$.

(4) $Q_1Q_2Q_3 = \mathrm{ENE}$ or $Q_1Q_2Q_3 = \mathrm{EEN}$. These cases are analogous to the previous case.

(5) $Q_1Q_2Q_3 = \mathrm{EEE}$ or $Q_1Q_2Q_3 = \mathrm{NNN}$. Trivial since $f$ preserves $E$ and $N$.

We now make another case distinction, based on the fact that $(f(x_1), \ldots, f(x_6)) \notin H$.

(1) Suppose that $E(f(x_1), f(x_2)), E(f(x_3), f(x_4)), E(f(x_5), f(x_6))$. Then by the above $f$ is of behaviour minority on $S$, a contradiction since $S$ induces a copy of $(H_n, E)^3$ and because of Lemma 3.14.

(2) Suppose that $N(f(x_1), f(x_2)), N(f(x_3), f(x_4)), N(f(x_5), f(x_6))$. Then $f$ has behaviour majority on $S$, again contradicting Lemma 3.14.

(3) Suppose that $E(f(x_1), f(x_2)), E(f(x_3), f(x_4)), N(f(x_5), f(x_6))$. Let $e$ be an endomorphism of $(H_n, E, N)$ such that for all $w \in H_n$, all $1 \leq j \leq 3$, and all $1 \leq i \leq 6$ we have that $N(x_i^j, e(w))$. Then $(u_1, u_2, e(f(u_1, u_2, u_3))) \in S$ for all $(u_1, u_2, u_3) \in S$. Hence, by the above, the ternary operation defined by $f(x, y, e(f(x, y, z)))$ is of type behaviour on $S$, a contradiction.

(4) Suppose that $E(f(x_1), f(x_2)), N(f(x_3), f(x_4)), E(f(x_5), f(x_6))$, or $N(f(x_1), f(x_2))$, $E(f(x_3), f(x_4))$, $E(f(x_5), f(x_6))$. These cases are analogous to the previous case.

Each of the cases leads to a contradiction, hence proving the proposition. ■

3.3.3. *Producing* min. By Proposition 3.15, it remains to show the following to obtain a proof of Proposition 3.12.

**Proposition 3.16.** *Let $f\colon H_n^2 \to H_n$ be a binary injection preserving $E$ and $N$ that is not of behaviour projection. Then $f$ generates a binary canonical injection of behaviour* min.

In the remainder of this section we will prove this proposition by Ramsey theoretic analysis of $f$, which requires the following definitions and facts from [BP14] concerning behaviours with respect to the homogeneous expansion of graphs by the total order $\prec$. At this point, it might be appropriate to remark that canonicity of functions on $H_n$, and even the notion of behaviour, does depend on which underlying structure we have in mind, in particular, whether or not we consider the order $\prec$ (which we almost managed to ignore so far). Let us define the following behaviours for functions from $(H_n, E, \prec)^2$ to $(H_n, E)$; we write $\succ$ for the relation $\{(a, b) \mid b \prec a\}$.

**Definition 3.17.** Let $f : H_n^2 \to H_n$ be injective. If for all $u, v \in H_n^2$ with $u_1 \prec v_1$ and $u_2 \prec v_2$

- $E(f(u), f(v))$ if and only if $\mathrm{EE}(u, v)$, then $f$ *behaves like* min *on input* $(\prec, \prec)$.
- $E(f(u), f(v))$ if and only if $E(u_1, v_1)$, then $f$ *behaves like* $p_1$ *on input* $(\prec, \prec)$.
- $E(f(u), f(v))$ if and only if $E(u_2, v_2)$, then $f$ *behaves like* $p_2$ *on input* $(\prec, \prec)$.

Analogously, we define behaviour on input $(\prec, \succ)$ using pairs $u, v \in V^2$ with $u_1 \prec v_1$ and $u_2 \succ v_2$.

**Proposition 3.18.** *Let $f : H_n^2 \to H_n$ be an injection which is canonical as a function from $(H_n, E, \prec)^2$ to $(H_n, E, \prec)$ and suppose $f$ preserves $E$ and $N$. Then it behaves like* min*, $p_1$ or $p_2$ on input $(\prec, \prec)$ (and similarly on input $(\prec, \succ)$).*

*Proof.* By definition of the term canonical; one only needs to enumerate all possible types of pairs $(u, v)$, where $u, v \in H_n^2$ and recall that $(H_n, E)$ does not contain any clique of size $n$, which makes some behaviours impossible to be realized by $f$. ■

**Definition 3.19.** If an injection $f : H_n^2 \to H_n$ behaves like $X$ on input $(\prec, \prec)$ and like $Y$ on input $(\prec, \succ)$, where $X, Y \in \{\min, p_1, p_2\}$, then we say that $f$ is of *behaviour $X/Y$*.

In the following lemmas, we show that every injective canonical binary function which behaves differently on input $(\prec, \prec)$ and on input $(\prec, \succ)$ generates a function which behaves the same on both inputs, allowing us to ignore the order again.

**Lemma 3.20.** *Suppose that $f : H_n^2 \to H_n$ is injective and canonical from $(H_n, E, \prec)^2$ to $(H_n, E, \prec)$, and suppose that it is of type $\min / p_i$ or of type $p_i / \min$, where $i \in \{1, 2\}$. Then $f$ generates a binary injection of type* min*.*

*Proof.* Since the calculus for behaviours on the Henson graphs is the same as that on the random graph, the same proof as in [BP15a] works. ■

**Lemma 3.21.** *No binary injection $f : H_n^2 \to H_n$ can have behaviour $p_1/p_2$.*

*Proof.* Such a behaviour would introduce a $K_n$ in a $K_n$-free graph. ■

Having ruled out some behaviours without constants, we finally introduce constants to the language to prove Proposition 3.16.

*Proof of Proposition 3.16.* Let $f$ be given. By Lemma 3.9 we have that $f$ generates a binary canonical injection $g$ of type projection or min. In the latter case we are done, so consider the first case. We claim that then $f$ also generates a (not necessarily canonical) binary injection $h$ of type min. We can then consider $h(g(x, y), g(y, x))$ which is still of type min and in addition canonical, and the proposition follows.

To prove our claim, we use Proposition 3.15 to observe that $f$ generates a binary injection $t$ which is not of behaviour projection. Fix a finite set $\{c_1, \ldots, c_m\} \subseteq H_n$ on which the latter fact is witnessed. Invoking Proposition 2.9, we may henceforth assume that $t$ is canonical as a function from $(H_n, E, \prec, c_1, \ldots, c_m)^2$ to $(H_n, E, \prec)$.

Of the structure $(H_n, E, \prec, c_1, \ldots, c_m)$, consider the orbit

$$O := \{v \in H_n \mid N(v, c_i) \text{ and } v \prec c_i \text{ for all } 1 \leq i \leq m\},$$

Then $O$ induces a structure isomorphic to $(H_n, E, \prec)$, as it satisfies the extension property for totally ordered $K_n$-free graphs: the same extensions can be realized in $O$ as in $(H_n, E, \prec)$. Therefore, by Lemma 3.18, $t$ has one of the there mentioned behaviors. By Lemmas 3.21 and 3.20, we may assume that $t$ behaves like a projection, say $p_1$, on $O$, for otherwise we are done.

Let $u \in O^2$ and $v \in (H_n \setminus \{c_1, \ldots, c_m\})^2$ satisfy $\neq\neq (u, v)$; we claim that $t$ behaves like $p_1$ or like min on $\{u, v\}$. Otherwise we must have $\mathrm{NE}(u, v)$, and $t$ behaves like $p_2$ on $\{u, v\}$. Pick $q_1, \ldots, q_{n-1} \in O^2$ forming a clique in the first coordinate, an independent set in the second coordinate, and such that the type of $(q_i, v)$ equals the type of $(u, v)$. Then by canonicity, the image of $\{q_1, \ldots, q_{n-1}, v\}$ under $t$ forms a clique of size $n$, a contradiction.

We next claim that for all $u, v \in (H_n \setminus \{c_1, \ldots, c_m\})^2$ with $\neq\neq (u, v)$ we have that $t$ must behave like $p_1$ or like min on $\{u, v\}$. Otherwise we must have $\mathrm{NE}(u, v)$, and $t$ behaves like $p_2$ on $\{u, v\}$. Pick $q_1, \ldots, q_{n-2} \in O^2$ forming a clique in the first coordinate, an independent set in the second coordinate, and adjacent to $u$ and $v$ only in the first coordinate. Applying $t$ we get a clique of size $n$, a contradiction.

By applying the same argument again, we now get that $t$ must behave like $p_1$ or like min on $\{u, v\}$ for all $u, v \in H_n^2$ with $\neq\neq (u, v)$ (picking $q_1, \ldots, q_{n-2} \in H_n^2$ this time, but with the same properties relative to $u, v$).

Somewhere $t$ does not behave like $p_1$ but like min, and so by a standard iterating argument it generates a binary injection of type min. ■

## 4. CSPs over Henson graphs

4.1. **Hardness of $H$.** We now show that any reduct of $(H_n, E)$ which has $H$ among its relations, and hence by Lemma 2.1 every reduct which pp-defines $H$, has an NP-hard CSP. While it would be possible to show NP-hardness of $\mathrm{CSP}(H_n, H)$ directly by reduction of, say, the NP-hard problem positive 1-in-3-SAT, we will use results from [BP15b], and in fact a recent strengthening thereof from [BOP15], to prove hardness more elegantly via a structural property of $\mathrm{Pol}(H_n, H)$.

**Definition 4.1.** Let $\Gamma$ be a structure. A *projective clone homomorphism* of $\Gamma$ is a mapping from $\mathrm{Pol}(\Gamma)$ onto its projections which

- preserves arities;
- fixes each projection;
- preserves composition.

A *projective strong h1 clone homomorphism* of $\Gamma$ is a mapping as above, where the third condition is weakened to preservation of composition with projections.

**Theorem 4.2** (from [BOP15]). *Let $\Gamma$ be a countable $\omega$-categorical structure in a finite relational language which has a uniformly continuous strong h1 clone homomorphism. Then* $\mathrm{CSP}(\Gamma)$ *is NP-hard.*

**Proposition 4.3.** *The structure $(H_n, H)$ has a uniformly continuous strong h1 clone homomorphism. Consequently,* $\mathrm{CSP}(H_n, H)$ *is NP-hard.*

*Proof.* Note that $H$ consists of three orbits of 6-tuples with respect to $(H_n, E)$. Let $a^1, a^2, a^3 \in H$ be representatives of those three orbits. We claim that whenever $f \in \text{Pol}(H_n, H)$ is ternary, and $b^1, b^2, b^3 \in H$ are so that the matrices $(b^1, b^2, b^3)$ and $(a^1, a^2, a^3)$ are isomorphic in the sense that $E(b^j_i, b^l_k)$ iff $E(a^j_i, a^l_k)$ for all $1 \leq i, k \leq 6$ and all $1 \leq j, l \leq 3$, then $f(b^1, b^2, b^3)$ and $f(a^1, a^2, a^3)$ are isomorphic, i.e., have the same type.

To see the claim, let $c^1, c^2, c^3 \in H$ be so that $(c^1, c^2, c^3)$ is isomorphic to $(b^1, b^2, b^3)$ and $(a^1, a^2, a^3)$, and such that no entry of any $c^i$ is adjacent to any component of any $b^j$ or $a^j$. If $f(b^1, b^2, b^3)$ and $f(a^1, a^2, a^3)$ do not have the same type, then one of them does not have the same type as $f(c^1, c^2, c^3)$; say without loss of generality this is the case for $f(a^1, a^2, a^3)$. Again without loss of generality, this is witnessed on the first two components of the 6-tuples $f(c^1, c^2, c^3)$ and $f(a^1, a^2, a^3)$. For $1 \leq i \leq 3$, consider the 6-tuple $d^i := (c^i_1, c^i_2, a^i_3, \ldots, a^i_6)$, i.e., in $a^i$ we replace the first two components by the components from $c^i$. Then $d^i \in H$, but $f(d^1, d^2, d^3) \notin H$, a contradiction.

It follows from the claim that every ternary $f \in \text{Pol}(H_n, H)$ satisfies the three type conditions $E(f(u), f(v))$ whenever $\text{ENN}(u, v)$, $N(f(u), f(v))$ whenever $\text{NEN}(u, v)$, and $N(f(u), f(v))$ whenever $\text{NNE}(u, v)$, or a similar set of type conditions where the special role of the first coordinate is taken by one of the other two coordinates. The mapping which sends every ternary $f \in \text{Pol}(H_n, H)$ to the ternary projection onto the special coordinate of $f$ is a strong h1 clone homomorphism from the ternary functions of $\text{Pol}(H_n, H)$, and is uniformly continuous since the value of every $f$ under the mapping can be seen on any test matrix $(a^1, a^2, a^3)$ as above. It is easy to see and well-known that any such mapping from the ternary functions of a function clone extends to the entire clone. ∎

4.2. **Tractability of** min**.** It remains to prove that if a reduct $\Gamma$ of $(H_n, E)$ with finite relational signature has a polymorphism which is of behaviour min and $N$-dominated, then $\text{CSP}(\Gamma)$ is in P. Observe that any such polymorphism is an embedding of $(H_n, E)^2$ into $(H_n, E)$. We apply Theorem 4.4 below for the structure $\Delta := (H_n, E)$. The following notation is used there: for a structure $\Delta$, we write $\hat{\Delta}$ for the expansion of $\Delta$ by the inequality relation $\neq$ and by the complement $\hat{R}$ of each relation $R$ in $\Delta$.

**Theorem 4.4** (Proposition 14 in [BCKvO09])**.** *Let $\Delta$ be an $\omega$-categorical structure, and let $\Gamma$ be a structure with a first-order definition in $\Delta$. If $\Gamma$ has a polymorphism $e$ which is an embedding of $\Delta^2$ into $\Delta$, and if $\text{CSP}(\hat{\Delta})$ is in P, then $\text{CSP}(\Gamma)$ is in P as well.*

Hence, it remains to show that the CSP for $\hat{\Delta} = (H_n, E, \hat{E}, \neq)$ can be solved in polynomial time. But this is easy: an instance of this CSP is satisfiable if and only if there are no variables $x_1, \ldots, x_n$ such that

- $E(x_i, x_j)$ is in the input for all distinct $i, j \in \{1, \ldots, n\}$ (in particular, the statement for $x_1 = \cdots = x_n$ implies that the input does not contain constraints of the form $E(x, x)$),
- there are no constraints of the form $x_1 \neq x_1$, and
- there are no constraints of the form $E(x_1, x_2)$ and $\hat{E}(x_1, x_2)$.

Since $n$ is fixed, it is clear that these conditions can be checked in polynomial time.

## 5. Summary for the Henson graphs

5.1. **Proof of the complexity dichotomy.** We are ready to prove the dichotomy for the CSPs of reducts of Henson graphs.

*Proof of Theorem 1.1.* Let $\Gamma$ be a reduct of $(H_n, E)$. If $\mathrm{End}(\Gamma)$ contains a function whose image is an independent set, then $\mathrm{CSP}(\Gamma)$ equals the CSP for a reduct of $(H_n, =)$ by Lemma 3.2, and such CSPs are either in P or NP-complete [BK08]. Otherwise, $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(H_n, E)}$ by Proposition 3.1. Lemma 3.3 shows that $E$, $N$, and $\neq$ are pp-definable in $\Gamma$. If also the relation $H$ is pp- definable in $\Gamma$, then $\mathrm{CSP}(\Gamma)$ is NP-hard by Proposition 4.3; it is in NP since $\Gamma$ is a reduct of a finitely bounded homogeneous structure. So let us assume that this is not the case; then Proposition 3.12 shows that $\mathrm{Pol}(\Gamma)$ contains a canonical binary injection of behavior min. By Lemma 3.8, we can assume that this polymorphism is $N$-dominated. Polynomial-time tractability of $\mathrm{CSP}(\Gamma)$ then follows from the argument given in Section 4.2. ∎

5.2. **Discussion.** We can restate Theorem 1.1 in a more detailed fashion as follows.

**Theorem 5.1.** *Let $\Gamma$ be a reduct of a Henson graph $(H_n, E)$. Then one of the following holds.*

(1) $\Gamma$ *has an endomorphism inducing an independent set, and is homomorphically equivalent to a reduct of* $(H_n, =)$.
(2) $\mathrm{Pol}(\Gamma)$ *has a uniformly continuous projective clone homomorphism.*
(3) $\mathrm{Pol}(\Gamma)$ *contains a binary canonical injection which is of behavior* min *and $N$-dominated.*

*Items (2) and (3) cannot simultaneously hold, and when $\Gamma$ has a finite relational signature, then* (2) *implies NP-completeness and (3) implies tractability of its CSP.*

The first statement follows directly from the proof of Theorem 1.1, with the additional observation that the strong h1 clone homomorphism defined in Proposition 4.3 is in fact a clone homomorphism. When (3) holds for a reduct, then (2) cannot hold, because (3) implies the existence of $f(x,y) \in \mathrm{Pol}(\Gamma)$ and $\alpha \in \overline{\mathrm{Aut}(\Gamma)}$ such that $f(x,y) = \alpha f(y,x)$ holds, and equation impossible to satisfy by projections. In fact, by further analyzing case (1), one can easily show that it also implies either (2) or (3), so that we have the following.

**Corollary 5.2.** *For every reduct $\Gamma$ of a Henson graph $(H_n, E)$, precisely one of the following holds.*

- $\mathrm{Pol}(\Gamma)$ *has a uniformly continuous projective clone homomorphism.*
- $\mathrm{Pol}(\Gamma)$ *contains* $f(x,y) \in \mathrm{Pol}(\Gamma)$ *and* $\alpha \in \overline{\mathrm{Aut}(\Gamma)}$ *such that* $f(x,y) = \alpha f(y,x)$.

*When $\Gamma$ has a finite relational signature, then the first case implies NP-completeness and the second case implies tractability of its CSP.*

## 6. Polymorphisms over homogeneous equivalence relations

We now investigate polymorphisms of reducts of the graphs $(C_n^s, E)$, for $2 \leq n, s \leq \omega$, with precisely one of $n, s$ equal to $\omega$. Recall from the preliminaries that we write $Eq$ for the reflexive closure of $E$.

Similarly to the case of the Henson graphs, we start with unary polymorphisms in Section 6.1, reducing the problem to model-complete cores.

We then turn to higher-arity polymorphisms; here, the organization somewhat differs from the case of the Henson graphs. The role of the NP-hard relation $H$ from the Henson graphs is now taken by the two sources of NP-hardness mentioned in the introduction: the first source being that factoring by the equivalence relation $Eq$ yields a structure with an NP-hard problem, and the second source being that restriction to some equivalence class yields a structure with an NP-hard problem. In Section 6.2, we show that in fact, one of the two sources always applies for model-complete cores when $2 < n < \omega$ or $2 < s < \omega$. Consequently, only the higher-arity polymorphisms of the reducts of $(C_2^\omega, E)$ and $(C_\omega^2, E)$ require deeper investigation using Ramsey theory; this will be dealt with in Sections 6.3 and 6.4, respectively.

### 6.1. **The unary case: model-complete cores.**

**Proposition 6.1.** *Let $\Gamma$ be a reduct of $(C_n^s, E)$, where $1 \leq n, s \leq \omega$, and at least one of $n, s$ equals $\omega$. Then $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(\Gamma)} = \overline{\mathrm{Aut}(C_n^s, E)}$, or $\mathrm{End}(\Gamma)$ contains an endomorphism onto a clique or an independent set.*

*Proof.* Assume that $\mathrm{End}(\Gamma) \neq \overline{\mathrm{Aut}(C_n^s, E)}$, so there is an endomorphism $f$ of $\Gamma$ violating either $E$ or $N$.

**Case 0.** If $n = 1$ or $s = 1$ then the statement is trivial.

**Case 1.** If $n = s = \omega$, so $Eq$ has infinitely many infinite classes, we can refer to [BW12].

**Case 2.** Assume that $n < \omega$ and $s = \omega$.

Suppose that $f$ violates $Eq$ and preserves $N$; then clearly, iterating applications of automorphisms of $(C_n^\omega, E)$ and $f$, we can send any finite subset of $C_n^\omega$ to an independent set in $(C_n^\omega, E)$, contradicting that the number of equivalence classes is the fixed finite number $n$.

If $f$ preserves both $Eq$ and $N$, then there exist $a, b$ with $E(a, b)$ and $f(a) = f(b)$. Via a standard iterative argument, one then sees that $f$ generates a function whose range is an independent set.

Therefore, it remains to consider the case where $f$ violates $N$. Fix $u, v \in C_n^\omega$ with $N(u, v)$ and $Eq(f(u), f(v))$.

By Proposition 2.9 and the remark thereafter, we may assume that $f$ is canonical as a $(C_n^\omega, E, \prec, u, v) \to (C_n^\omega, E, \prec)$ function. Clearly, that function cannot violate $Eq$ anywhere, as otherwise canonicity would imply an infinite independent set in $(C_n^\omega, E)$. Thus, $Eq$ is preserved and $N$ violated, and a standard iterative argument shows that $f$ generates a function whose range is contained in a single equivalence class.

**Case 3.** Assume that $s < \omega$ and $n = \omega$.

Suppose that $f$ violates $N$ and preserves $Eq$; then clearly, $f$ generates a mapping onto a clique. If it preserves both $Eq$ and $N$, then as above, $f$ generates a function whose range is an independent set.

Therefore, we may assume that $f$ violates $Eq$. Fix $u, v \in C_\omega^s$ with $E(u, v)$ such that $N(f(u), f(v))$. By Proposition 2.9 and the remark thereafter, we may assume that $f$ is canonical as a $(C_\omega^s, E, \prec, u, v) \to (C_\omega^s, E, \prec)$ function. We then cannot have $a, b \in C_\omega^s$ with $N(a, b)$ and $E(f(a), f(b))$, because the equivalence classes are finite and because of canonicity. Hence, all $a, b \in C_\omega^s$ with $N(a, b)$ satisfy $f(a) = f(b)$ or $N(f(a), f(b))$. But then using that we have $E(u, v)$ and $N(f(u), f(v))$, we see that $f$ generates a function on whose range the relation $E$ never holds, and which therefore induces an independent set. ■

In the following sections, we investigate essential polymorphisms of reducts $\Gamma$ of $(C_n^s, E)$ which are model-complete cores, i.e., $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_n^s, E)}$. The following proposition implies that in that situation, the equivalence relation $Eq$ is invariant under $\mathrm{Pol}(\Gamma)$.

**Proposition 6.2.** *Let $\Gamma$ be a reduct of $(C_n^s, E)$, where $1 \leq n, s \leq \omega$. If $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_n^s, E)}$, then $E$, $N$ and $Eq$ are preserved by the polymorphisms of $\Gamma$.*

*Proof.* By Lemma 2.3, the condition $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C_n^s, E)}$ implies that all polymorphisms of $\Gamma$ preserve $E$ and $N$, and hence also $Eq$ since $Eq(x, y)$ has the primitive positive definition $\exists z\, (E(x, z) \wedge E(z, y))$. ■

Therefore, in the above situation $Eq$ is an equivalence relation which is invariant under $\mathrm{Pol}(\Gamma)$, and so $\mathrm{Pol}(\Gamma)$ acts naturally on the equivalence classes of $Eq$. Moreover, if we fix any $c \in C^s_n$ and expand the structure $\Gamma$ by the constant $c$, then the equivalence class $C$ of $c$ has a primitive positive definition in that expansion $(\Gamma, c)$, since $Eq$ and $c$ do. Hence, $C$ is invariant under $\mathrm{Pol}(\Gamma, c)$, and so $\mathrm{Pol}(\Gamma, c)$ acts naturally on $C$ via restriction. In the following sections, we analyze these actions.

6.2. **The case $2 < n < \omega$ or $2 < s < \omega$.** It turns out that in these cases, one of the sources of hardness always applies. We will use the following fact about function clones on a finite domain.

**Proposition 6.3** (from [HR94]). *Every function clone on a finite domain of at least three elements which contains all permutations as well as an essential function contains a unary constant function.*

**Proposition 6.4.** *Let $\Gamma$ be a reduct of $(C^\omega_n, E)$, where $2 < n < \omega$, such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^\omega_n, E)}$. Then the action of $\mathrm{Pol}(\Gamma)$ on the equivalence classes of $Eq$ has no essential and no constant operation.*

*Proof.* The action has no constant operation because $N$ is preserved. Therefore, it cannot have an essential operation either, by Proposition 6.3. ■

**Proposition 6.5.** *Let $\Gamma$ be a reduct of $(C^s_\omega, E)$, where $2 < s < \omega$, such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^s_\omega, E)}$. Then for any $c \in C^s_\omega$, the action of $\mathrm{Pol}(\Gamma, c)$ on the equivalence class of $c$ has no essential and no constant operation.*

*Proof.* The action has no constant operation because $E$ is preserved. Therefore, it cannot have an essential operation either, by Proposition 6.3. ■

6.3. **The case of two infinite classes: $n = 2$ and $s = \omega$.** The following proposition states that either one of the two sources of hardness applies, or $\mathrm{Pol}(\Gamma)$ contains a ternary canonical function with a certain behaviour.

**Proposition 6.6.** *Let $\Gamma$ be a reduct of $(C^\omega_2, E)$ such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^\omega_2, E)}$. Then one of the following holds:*

- *the action of $\mathrm{Pol}(\Gamma)$ on the equivalence classes of $Eq$ has no essential function;*
- *the action of $\mathrm{Pol}(\Gamma, c)$ on the equivalence class of $c$ has no essential function, for some $c \in C^\omega_2$;*
- *$\mathrm{Pol}(\Gamma)$ contains a canonical ternary injection of behaviour minority which is hyperplanely of behaviour $E$-dominated projection.*

To prove the proposition, we need to recall a special case of Post's classical result about function clones acting on a two-element set, as well as a result on function clones on a countable set which contain all permutations. Comparing this statement with Proposition 6.3 sheds light on why the case of this section is more involved than the cases of the preceding section.

**Proposition 6.7** (Post [Pos41]). *Every function clone with domain $\{0, 1\}$ containing both permutations of $\{0, 1\}$ as well as an essential function contains a unary constant operation or the ternary addition modulo 2.*

**Proposition 6.8** (from [BK08]). *Every closed function clone on a countably infinite set which contains all permutations as well as an essential operation contains a binary injection.*

*Proof of Proposition 6.6.* Denote the two equivalence classes of $Eq$ by $C_0$ and $C_1$. Suppose that the first statement of the proposition does not hold. Then by Proposition 6.7, the action of $\mathrm{Pol}(\Gamma)$ on $\{C_0, C_1\}$ contains a unary constant operation, or a function which behaves like ternary addition modulo 2. The first case is impossible since the unary functions in $\mathrm{Pol}(\Gamma)$ preserve $N$, so the latter case holds and $\mathrm{Pol}(\Gamma)$ contains a ternary function $g$ which acts like $x + y + z$ modulo 2 on the classes.

Suppose now in addition that the second statement of the proposition does not hold either. Fix some $c \in C_2^\omega$. Then note that the action of $\mathrm{Pol}(\Gamma, c)$ on the class $c$ of $C$ contains all permutations of $C$. By Proposition 6.8, this implies that this action contains a binary injection. Therefore, in $\mathrm{Pol}(\Gamma)$ we have functions $f_0, f_1$ which are a binary injection on the classes $C_0$ and $C_1$, respectively.

We claim that there is a single function $f \in \mathrm{Pol}(\Gamma)$ which is injective on each class. To see this, assume that both $f_0, f_1$ preserve the classes, which can be achieved by composing them with automorphisms. If $f_0$ is essential on $C_1$, then by Proposition 6.8 together with all functions in $\mathrm{Pol}(\Gamma)$ which fix the classes, it generates a function which is injective on $C_1$; this function is then injective on both classes $C_0, C_1$. So assume that $f_0$ is not essential on $C_1$, say without loss of generality that it depends injectively on the first coordinate. Then $f_0(f_1(x, y), f_0(x, y))$ is injective on both classes.

By Proposition 2.9, we may assume that $f$ is canonical as a function from $(C_2^\omega, E, \prec)^2$ to $(C_2^\omega, E, \prec)$. We claim that $f$ is also canonical as a function from $(C_2^\omega, E)^2$ to $(C_2^\omega, E)$. To prove this, it suffices to show that if $u, v, u', v' \in C_2^\omega \times C_2^\omega$ are so that $(u, v)$ and $(u', v')$ have the same type in $(C_2^\omega, E) \times (C_2^\omega, E)$, then $(f(u), f(v))$ and $(f(u'), f(v'))$ have the same type in $(C_2^\omega, E)$. By canonicity of $f$ as a function from $(C_2^\omega, E, \prec)^2$ to $(C_2^\omega, E, \prec)$, we may assume that $EqEq(u, u')$ and $EqEq(v, v')$ in order to test this. Since $E$ is preserved, we have $Eq(f(u), f(u'))$ and $Eq(f(v), f(v'))$, and so $Eq(f(u), f(v))$ implies $Eq(f(u'), f(v')))$ and vice-versa, by the transitivity of $Eq$. Failure of canonicity can therefore only happen if $Eq(f(u), f(v))$ and $Eq(f(u'), f(v')))$, and precisely one of $f(u) = f(v)$ and $f(u') = f(v')$ holds, say without loss of generality the former. But then picking any $v'' \in C_2^\omega \times C_2^\omega$ distinct from $v$ such that $EqEq(v, v'')$ and such that the type of $(u, v)$ equals the type of $(u, v'')$ in $(C_2^\omega, E, \prec) \times (C_2^\omega, E, \prec)$ shows that $f(v) = f(u) = f(v'')$ by canonicity, contradicting the fact that $f$ is injective on each equivalence class.

We analyze the behavior of the canonical function $f\colon (C_2^\omega, E)^2 \to (C_2^\omega, E)$. Because $E$ and $N$ are preserved, we have $f(E, E) = E$ and $f(N, N) = N$. Because $f$ is injective on the classes, and because $Eq$ is preserved, we have $f(=, E) = f(E, =) = E$

Either $f(\cdot, N) = N$ or $f(N, \cdot) = N$. Otherwise, there exist $Q, P \in \{E, =\}$ such that $f(Q, N) \neq N$ and $f(N, P) \neq N$. Pick $u, v, w \in (C_2^\omega)^2$ such that $QN(u, v), NP(v, w)$, and $NN(u, w)$. Then $Eq(f(u), f(w))$ and $N(f(u), f(w))$, a contradiction.

Assume henceforth without loss of generality that $f(N, \cdot) = N$. Then $f(P, N) \neq N$ for $P \neq N$, because there are only two equivalence classes. Moreover, $f(E, N) = =$ or $f(=, N) = =$ implies that $f$ is not injective on the classes, so we have have $f(E, N) = f(=, N) = E$.

Summarizing, $f$ is a binary injection of behaviour $p_1$, balanced in the first argument, and $E$-dominated in the second argument.

Set $q(x, y, z) := f(x, f(y, z))$. Set $h(x, y, z) := g(q(x, y, z), q(y, z, x), q(z, x, y))$. We have $h(E, E, E) = E$, $h(N, N, N) = N$, $h(=, =, =) = =$. Moreover, $h$ is injective.

We have $h(E, E, N) = g(E, E, N) = N$, and by symmetry, $h(E, N, E) = h(N, E, E) = N$. We have $h(E, N, N) = g(E, N, N) = E$, and by symmetry, $h(N, E, N) = h(N, N, E) = E$. Thus, $h$ has behaviour minority.

We have $h(=, E, E) = g(E, E, E) = E$. We have $h(=, =, E) = g(E, E, E) = E$. Thus $h(P, Q, R) = E$ whenever $P, Q, R \in \{=, E\}$ are not all equal to $=$.

We have $h(=, N, N) = g(N, N, N) = N$. We have $h(=, =, N) = g(N, E, N) = E$. We have $h(=, E, N) = g(E, E, N) = N$. We have $h(=, N, E) = g(N, N, E) = E$.

Summarizing, $h$ is hyperplanely an $E$-dominated projection. ■

6.4. **The case of infinitely many classes of size two: $n = \omega$ and $s = 2$.** As in the preceding section, we show that either one of the two sources of hardness applies, or $\mathrm{Pol}(\Gamma)$ contains a ternary canonical function of a certain behaviour.

**Proposition 6.9.** *Let $\Gamma$ be a reduct of $(C^2_\omega, E)$ such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^2_\omega, E)}$. Then one of the following holds:*

- *the action of $\mathrm{Pol}(\Gamma)$ on the equivalence classes of $Eq$ has no essential function;*
- *the action of $\mathrm{Pol}(\Gamma, c)$ on the equivalence class of $c$ has no essential function, for some $c \in C^2_\omega$;*
- *$\mathrm{Pol}(\Gamma)$ contains a ternary canonical function $h$ such that $h(N, \cdot, \cdot) = h(\cdot, N, \cdot) = h(\cdot, \cdot, N) = N$ which behaves like a minority on $\{E, =\}$.*

To prove the proposition, we are again going to make use of Propositions 6.7 and 6.8, and the following lemma.

**Lemma 6.10.** *Let $\Gamma$ be a reduct of $(C^2_\omega, E)$ such that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^2_\omega, E)}$. If $\mathrm{Pol}(\Gamma)$ contains a ternary function which behaves like $x + y + z$ modulo 2 on some equivalence class, then it contains a ternary function which behaves like $x + y + z$ modulo 2 on all equivalence classes.*

*Proof.* Let $C_0, C_1, \ldots$ be the equivalence classes of $Eq$. Since $\mathrm{Pol}(\Gamma)$ is topologically closed, and since all classes are finite, a standard compactness argument implies that it is sufficient to show that for all $0 \leq n < \omega$, $\mathrm{Pol}(\Gamma)$ contains a function $g_n$ which behaves like $x + y + z$ modulo 2 on each class $C_0, \ldots, C_n$. We show this by induction over $n$.

For $n = 0$, the statement follows from the assumption of the lemma. Now suppose we have shown it for $n$. By the assumption that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(D, E)}$, we may assume that $g_n(x, x, x) = x$ for all $x \in C_0 \cup \cdots \cup C_{n+1}$, and in particular $g_n$ preserves each of the classes $C_1, \ldots, C_{n+1}$.

Assume first that $g_n$ is not essential on $C_{n+1}$; by composing it with an automorphism of $(C^2_\omega, E)$, we may assume it behaves like a projection, without loss of generality to the first coordinate, on $C_{n+1}$. Let $g'_n \in \mathrm{Pol}(\Gamma)$ be a ternary function which has the properties of $g_n$, but with the roles of $C_n$ and $C_{n+1}$ switched. Then $g_{n+1}(x, y, z) := g_n(g'_n(x, y, z), g'_n(y, z, x), g'_n(z, x, y))$ has the desired property.

Next assume that $g_n$ is essential on $C_{n+1}$. Then by Proposition 6.7, a function $h$ that behaves like $x + y + z$ modulo 2 on $C_{n+1}$ can be written as a finite composition of ternary projections, an automorphism of $(C^2_\omega, E)$ which fixes all elements of $C_0 \cup \cdots \cup C_n$ and flips the elements of $C_{n+1}$, and $g_n$. It is easy to see that the restriction of this function $h$ to any of the classes $C_0, \ldots, C_n$ either acts like $x + y + z$ modulo 2 or like a projection. Hence, iterating the preceding case we obtain the desired function. ■

*Proof of Proposition 6.9.* Suppose that none of the first two items hold. Then by Proposition 6.8, $\mathrm{Pol}(\Gamma)$ we contains a binary function $f$ acting injectively on the classes; moreover, using Proposition 6.7 we see that $\mathrm{Pol}(\Gamma)$ contains a ternary function which acts like $x+y+z$ modulo 2 on some equivalence class. Hence, by Lemma 6.10 it contains a ternary function $g$ which behaves like $x+y+z$ modulo 2 on all equivalence classes.

By Proposition 2.9, we may assume that $f$ is canonical as a function from $(C^2_\omega, E, \prec)^2$ to $(C^2_\omega, E, \prec)$. As in Proposition 6.6, this implies that $f$ is also canonical as a function from $(C^2_\omega, E)^2$ to $(C^2_\omega, E)$.

We first claim that $f$ behaves like the binary minimum function on $\{Eq, N\}$, where $Eq$ takes the role of the larger element. To see this, let $q \in (C^2_\omega)^2$, and let $Q \subseteq (C^2_\omega)^2$ be infinite such that $EN(q', q)$ for all $q' \in Q$. If $Eq(f(q), f(q'))$ for some $q, q' \in Q$, then all of $Q$ is sent into a single equivalence class, contradicting the fact that $f$ acts injectively on the classes. Hence, $f(E, N) = N$, and by an identical argument for the other types we obtain the claim.

On each class, $f$ preserves $E$, and hence we claim it must behave like a projection on the two elements. Call the two elements of a class 0 and 1. We know that $f$ sends $\{0,1\}^2$ into a single class, without loss of generality itself. We also have $E(f(0,0), f(1,1))$, so without loss of generality $f(0,0) = 0, f(1,1) = 1$. We also know $E(f(0,1), f(1,0))$. Either $f(0,1) = 1, f(1,0) = 0$, then $f$ acts like the second projection; or $f(0,1) = 0, f(1,0) = 1$, then $f$ acts like the first projection. Say now without loss of generality it behaves like the first projection, i.e., $f(E,=) = E$ and $f(=,E) = =$.

The function $q(x,y,z) := f(x, f(y,z))$ satisfies $q(N,\cdot,\cdot) = q(\cdot,N,\cdot) = q(\cdot,\cdot,N) = N$, and $q(P,Q,R) = P$ when $P,Q,R \in \{E,=\}$.

Set $h(x,y,z) := g(q(x,y,z), q(y,z,x), q(z,x,y))$. Then $h(N,\cdot,\cdot) = h(\cdot,N,\cdot) = h(\cdot,\cdot,N) = q(N,N,N) = N$. When $P,Q,R \in \{E,=\}$, then $h(P,Q,R) = g(P,Q,R)$; thus, it behaves like a minority on $\{E,=\}$. ■

## 7. Polynomial-time tractable CSPs over homogeneous equivalence relations

7.1. **Two infinite classes.** We now assume that $Eq$ is an equivalence relation on $C^\omega_2$ with two infinite classes, and suppose that $\Gamma$ is preserved by a canonical injective edge-minority $h$ that is hyperplanely of type $E$-dominated projection. Our algorithm for $\mathrm{CSP}(\Gamma)$ is an adaptation of an algorithm for GraphSAT problems [BP15a].

We first reduce $\mathrm{CSP}(\Gamma)$ to the CSP of a structure that we call the *injectivization* of $\Gamma$, which can then be reduced to a CSP over a Boolean domain.

**Definition 7.1.** A tuple is called *injective* if all its entries have pairwise distinct entries. A relation is called *injective* if all its tuples are injective. A structure is called *injective* if all its relations are injective.

**Definition 7.2.** We define *injectivizations* for relations, atomic formulas, and structures.

- Let $R$ be any relation. Then the *injectivization of* $R$, denoted by $\mathrm{inj}(R)$, is the (injective) relation consisting of all injective tuples of $R$.
- Let $\phi(x_1, \dots, x_n)$ be an atomic formula in the language of $\Gamma$, where $x_1, \dots, x_n$ is a list of the variables that appear in $\phi$. Then the *injectivization of* $\phi(x_1, \dots, x_n)$ is the formula $R^{\mathrm{inj}}_\phi(x_1, \dots, x_n)$, where $R^{\mathrm{inj}}_\phi$ is a relation symbol which stands for the injectivization of the relation defined by $\phi$.

```
// Input: An instance Φ of CSP(Γ) with variables V
While Φ contains a constraint φ that implies x = y for x, y ∈ V do
    Replace each occurrence of x by y in Φ.
    If Φ contains a false constraint then reject
Loop
Accept if and only if inj(Φ) has a solution in inj(Γ) (Proposition 7.4)
if and only if the Boolean equation system Boole(Φ) has a solution (Proposition 7.6).
```

FIGURE 1. Algorithm for CSP(Γ) when Γ is preserved by a ternary injection of type minority which is hyperplanely E-dominated projection.

- The *injectivization* of a relational structure Γ, denoted by inj(Γ), is the relational structure with the same domain as Γ whose relations are the injectivizations of the atomic formulas over Γ, i.e., the relations $R_\phi^{\text{inj}}$.

To state the reduction to the CSP of an injectivization, we also need the following operations on instances of CSP(Γ). Here, it will be convenient to view instances of CSP(Γ) as primitive positive $\tau$-sentences.

**Definition 7.3.** Let Φ be an instance of CSP(Γ). Then the *injectivization of* Φ, denoted by inj(Φ), is the instance $\psi$ of CSP(inj(Γ)) obtained from $\phi$ by replacing each conjunct $\phi(x_1, \dots, x_n)$ of Φ by $R_\phi^{\text{inj}}(x_1, \dots, x_n)$.

We say that a constraint in an instance of CSP(Γ) is *false* if it defines an empty relation in Γ. Note that a constraint $R(x_1, \dots, x_k)$ might be false even if the relation $R$ is non-empty (simply because some of the variables from $x_1, \dots, x_k$ might be equal). The proof of the following statement is identical to the proof for the random graph instead of $(C_2^\omega, Eq)$ in [BP15a].

**Proposition 7.4.** *Let* Γ *be preserved by a binary injection* $f$ *of type E-dominated projection. Then* CSP(Γ) *can be reduced to* CSP(inj(Γ)) *in polynomial time via the algorithm shown in Figure 1.*

To reduce the CSP for injective structures to Boolean CSPs, we need the following definitions. Let $t$ be a $k$-tuple of distinct vertices of $(C_2^\omega, E)$, and let $q$ be $\binom{k}{2}$. Then Boole($t$) is the $q$-tuple $(a_{1,2}, a_{1,3}, \dots, a_{1,k},\ a_{2,3}, \dots, a_{k-1,k}) \in \{0,1\}^q$ such that $a_{i,j} = 0$ if $N(t_i, t_j)$ and $a_{i,j} = 1$ if $E(t_i, t_j)$. If $R$ is a $k$-ary injective relation, then Boole($R$) is the $q$-ary Boolean relation $\{\text{Boole}(t) \mid t \in R\}$. Note that if an injective relation $R$ is preserved by a ternary operation of type minority, then $B := \text{Boole}(R)$ is preserved by the ternary minority function. It is well-known that $B$ then has a definition by a set of linear equations over $\{0,1\}$.

**Definition 7.5.** Let Φ is an instance of Γ with variables $V$. Then Boole(Φ) is the linear equation system with variables $\binom{V}{2}$ (that is, two-element subsets $\{u, v\}$ of $V$, denoted by $uv$) that contains

(1) for each conjunct $\phi(x_1, \dots, x_k)$ of Φ all linear equations with variables $\binom{\{x_1,\dots,x_k\}}{2}$ that define Boole($R_\phi^{\text{inj}}$), and
(2) all equations of the form $xy + yz + xz = 1$ for $a, b, c \in V$.

**Proposition 7.6.** *The formula* inj(Φ) *is satisfiable over* inj(Γ) *if and only if* Boole(Φ) *is satisfiable over* $\{0,1\}$.

*Proof.* Let $V$ be the variables of inj(Φ) so that $\binom{V}{2}$ are the variables of Boole(Φ). First suppose that inj(Φ) has a solution $s\colon V \to C_2^\omega$; we may choose $s$ injective. Then $s'\colon \binom{V}{2} \to \{0,1\}$ defined by $s'(xy) := 0$ if $N(s(x),s(y))$ and $s'(xy) := 1$ if $E(s(x),s(y))$ is a solution to Boole(Φ). Conversely, if $s'\colon \binom{V}{2} \to \{0,1\}$ is a solution to Boole(Φ), then define $s\colon V \to C_2^\omega$ as follows. Choose $x \in V$ and $v \in C_2^\omega$ arbitrarily, and define $s(x) := v$. For any $y \in V \setminus \{x\}$, if $s'(xy) = 1$, then pick $u \in C_2^\omega$ with $E(u,v)$ and if $s'(xy) = 0$, then pick $y \in C_2^\omega$ with $N(u,v)$; in both cases, choose values from $C_2^\omega$ that are distinct from all previously picked values from $C_2^\omega$. We claim that $s$ satisfies all conjuncts $\phi$ of inj(Φ). Let $R$ be the relation defined by $\phi$; then it suffices to show that $s$ satisfies all expressions of the form $E(x_1,y_1) \oplus \cdots \oplus E(x_k,y_k)$ or $\neg(E(x_1,y_1) \oplus \cdots \oplus E(x_k,y_k))$ that correspond to the Boolean equations defining $\text{Boole}(R_\phi^{\text{inj}})$. But

$$\begin{aligned}
&E(s(x_1),s(y_1)) \oplus \cdots \oplus E(s(x_k),s(y_k)) \\
\Leftrightarrow\ & (s'(xx_1)+s'(xy_1)=1) \oplus \cdots \oplus (s'(xx_k)+s'(xy_k)=1) && \text{(by definition of } s\text{)} \\
\Leftrightarrow\ & s'(x_1y_1) \oplus \cdots \oplus s'(x_ky_k)) && \text{(by (2) in Definition 7.5)}
\end{aligned}$$

which is true because $s'$ satisfies the equations from (1) of Definition 7.5. ■

**Corollary 7.7.** *Let* Γ *be preserved by a ternary injection of type minority* $h$ *which is hyperplanely an* $E$*-dominated projection. Then* CSP(Γ) *can be solved in polynomial time.*

*Proof.* Note that $h(x,x,y)$ is $E$-dominated of type projection. So the statement is a consequence of Proposition 7.4 and 7.6 (the algorithm can be found in Figure 1). ■

7.2. **Infinitely many classes of size two.** Let Γ be a reduct of $(C_\omega^2, Eq)$ where $Eq$ is an equivlence relation with infinitely many classes of size two such that Pol(Γ) contains a ternary canonical function $h$ such that

$$h(N,\cdot,\cdot) = h(\cdot,N,\cdot) = h(\cdot,\cdot,N) = N$$

which behaves like a minority on $\{=,E\}$.

**Proposition 7.8.** *A relation with a first-order definition in* $(C_\omega^2, Eq)$ *is preserved by* $h$ *if and only if it can be defined by a conjunction of formulas of the form*

$$N(x_1,y_1) \vee \cdots \vee N(x_k,y_k) \vee Eq(z_1,z_2) \tag{1}$$

*for* $k \geq 0$, *or of the form*

$$N(x_1,y_1) \vee \cdots \vee N(x_k,y_k) \vee (|\{i \in S : x_i \neq y_i\}| \equiv_2 p) \tag{2}$$

*where* $p \in \{0,1\}$ *and* $S \subseteq \{1,\ldots,k\}$.

The proof is inspired from a proof for tractable phylogeny constraints [BJP16].

*Proof.* For the backwards implication, it suffices to verify that formulas of the form in the statement are preserved by $h$. Let $o,p,q \in R$, and let $r := h(o,p,q)$. Assume that $R$ has a definition by a formula $\phi$ of the form as described in the statement. Suppose for contradiction that $r$ does not satisfy $\phi$. For any conjunct of $\phi$ of the form $N(x_1,y_1) \vee \cdots \vee N(x_k,y_k) \vee \theta$, the tuple $r$ must therefore satisfy $Eq(x_1,y_1) \wedge \cdots \wedge Eq(x_k,y_k)$. Since $h$ has the property that $h(N,\cdot,\cdot) = h(\cdot,N,\cdot) = h(\cdot,\cdot,N) = N$, this means that each of $o$, $p$, and $q$ also satisfies this formula. This in turn implies that $o$, $p$, and $q$ must satisfy the formula $\theta$. It suffices to prove that $r$ satisfies $\theta$, too, since this contradicts the assumption that $r$ does not satisfy $\phi$. Suppose first that $\theta$ is of the form $Eq(z_1, z_2)$. In this case, $r$ must also satisfy $Eq(z_1, z_2)$ since $h$ preserves $Eq$. So assume that $\theta$ is of the form $|\{i \in S : x_i \neq y_i\}| \equiv_2 p$ for $S \subseteq \{1, \dots, k\}$ and $p \in \{0, 1\}$. Since each of $o$, $p$, $q$ satisfies this formula and $h$ behaves like a minority on $\{E, =\}$, we have that $r$ satisfies this formula, too.

For the forwards implication, let $R$ be a relation with a first-order definition in $(C^2_\omega, Eq)$ that is preserved by $h$. Define $\sim$ to be the equivalence relation on $(C^2_\omega)^n$ where $a \sim b$ iff $Eq(a_i, a_j) \Leftrightarrow Eq(b_i, b_j)$ for all $i, j \leq n$. Note that $h$ preserves $\sim$. For $a \in (C^2_\omega)^n$, let $R_a$ be the relation that contains all $t \in R$ with $t \sim a$. Let $\psi_a$ be the formula

$$\bigwedge_{i<j\leq n, Eq(a_i,a_j)} Eq(x_i, x_j)$$

and $\psi'_a$ be the formula

$$\bigwedge_{i<j\leq n, N(a_i,a_j)} N(x_i, x_j)\,.$$

Note that $t \in (C^2_\omega)^n$ satisfies $\psi_a \wedge \psi'_a$ if and only if $t \sim a$, and hence a tuple from $R$ is in $R_a$ if and only it satisfies $\psi_a \wedge \psi'_a$.

Pick representatives $a_1, \dots, a_m$ for all orbits of $n$-tuples in $R$.

**Claim 1.** $\bigvee_{i\leq m}(\psi_{a_i} \wedge \psi'_{a_i})$ is equivalent to a conjunction of formulas of the form (1) from the statement.

Rewrite the formula into a formula $\psi_0$ in conjunctive normal form of minimal size where every literal is either of the form $Eq(x, y)$ or of the form $N(x, y)$. Suppose that $\psi_0$ contains a conjunct with literals $Eq(a, b)$ and $Eq(c, d)$. Since $\psi_0$ is of minimal size there exists $r \in (C^2_\omega)^n$ that satisfies $Eq(a, b)$ and none of the other literals in the conjunct, and similarly there exists $s \in (C^2_\omega)^n$ that satisfies $Eq(c, d)$ and none of the other literals. By assumption, $r \sim r' \in R$ and $s \sim s' \in R$. Since $R$ is preserved by $h$, we have $t' := h(r', s', s') \in R$. Then $t \sim t'$ since $h$ preserves $\sim$, and hence $t$ satisfies $\psi_0$. But $t$ satisfies none of the literals in the conjunct, a contradiction. Hence, all conjuncts of $\psi_0$ have form (1) from the statement.

Let $t \in (C^2_\omega)^n$, set $l := \binom{n}{2}$, and let $i_1j_1, \dots, i_lj_l$ be an enumeration of $\binom{\{1,\dots,n\}}{2}$. The tuple $b \in \{0, 1\}^{\binom{n}{2}}$ with $b_s = 1$ if $t_{i_s} \neq t_{j_s}$ and $b_s = 0$ otherwise is called the *split vector* of $t$. We associate to $R_a$ the Boolean relation $B_a$ consisting of all split vectors of tuples in $R_a$. Since $R$ and $R_a$ are preserved by $h$, the relation $B_a$ is preserved by the Boolean minority operation, and hence has a definition by a Boolean equation system. Therefore, there exists a conjunction $\theta$ of equations of the form $|\{s \in S : x_{i_s} = y_{j_s}\}| \equiv_2 p$, $p \in \{0, 1\}$ such that $\theta \wedge \psi_a \wedge \psi'_a$ defines $R_a$.

**Claim 2.** The following formula $\phi$ defines $R$:

$$\phi := \psi_0 \wedge \bigwedge_{a\in\{a_1,\dots,a_m\}} (\neg\psi_a \vee \theta_a)$$

It is straightforward to see that this formula can be rewritten into a formula of the form as required in the statement.

To prove the claim, we first show that every $t \in R$ satisfies $\phi$. Clearly, $t$ satisfies $\psi_0$. Let $a \in \{a_1, \dots, a_m\}$ be arbitrary; we have to verify that $t$ satisfies $\neg\psi_a \vee \theta_a$. If there are indices $i, j \in \{1, \dots, n\}$ such that $N(t_i, t_j)$ and $Eq(a_i, a_j)$, then $t$ satisfies $\neg\psi_a$. We are left with the case that for all $i,j \in \{1,\dots,n\}$ if $Eq(a_i,a_j)$ then $Eq(t_i,t_j)$. In order to show that $t$ satisfies $\theta_a$, it suffices to show that there exists a $t' \in R_a$ such that for all $i,j \leq n$ with $Eq(a_i,a_j)$ we have $t_i = t_j$ iff $t'_i = t'_j$. Note that $t' := h(a,a,t) \sim a$ since $h(N,\cdot,\cdot) = h(\cdot,N,\cdot) = h(\cdot,\cdot,N) = h$. Moreover, $t' \in R$ and thus $t' \in R_a$. Finally, for all $i,j \leq n$ with $Eq(a_i,a_j)$ we have $t_i = t_j$ iff $t'_i = t'_j$ because $h$ behaves as a minority on $\{E,=\}$. Hence, $t$ satisfies $\phi$.

Next, we show that every tuple $t$ that satisfies $\phi$ is in $R$. Since $t$ satisfies $\psi_0$ we have that $t \sim a$ for some $a \in \{a_1,\dots,a_m\}$. Thus, $t \models \psi_a \wedge \psi'_a$. By assumption, $t$ satisfies $\neg\psi_a \vee \theta_a$ and hence $t \models \theta_a$. Therefore, $t \in R_a$ and in particular $t \in R$. ■

**Proposition 7.9.** *There is a polynomial-time algorithm that decides whether a given set $\Phi$ of formulas as in the statement of Proposition 7.8 is satisfiable.*

*Proof.* Let $X$ be the set of variables that appear in $\Phi$. Create a graph $G$ with vertex set $X$ that contains an edge between $z_1$ and $z_2$ if $\Phi$ contains a formula of the form $Eq(z_1,z_2)$. Eliminate all literals of the form $N(x_i,y_i)$ in formulas from $\Phi$ when $x_i$ and $y_i$ lie in the same connected component of $G$. Repeat this procedure until no more literals get removed.

We then create a Boolean equation system $\Psi$ with variable set $\binom{X}{2}$ as follows. For each formula $|\{i \in S \mid x_i \neq y_i\}| \equiv_2 p$ we add the Boolean equation $\sum_{i \in S} x_i y_i = p$. We additionally add for all $xy, yz, xz \in \binom{X}{2}$ the equation $xy + yz = xz$. If the resulting equation system $\Psi$ does not have a solution over $\{0,1\}$, reject the instance. Otherwise accept.

To see that this algorithm is correct, observe that the literals that have been removed in the first part of the algorithm are false in all solutions, so removing them from the disjunctions does not change the set of solutions.

If the algorithm rejects, then there is indeed no solution to $\Phi$. To see this, suppose that $s\colon C^2_\omega \to C^2_\omega$ is a solution to $\Phi$. Define $b\colon \binom{X}{2} \to \{0,1\}$ as follows. Note that for every variable $\{x_i,y_i\}$ that appears in some Boolean equation in $\Psi$, a literal $N(x_i,y_i)$ has been deleted in the first phase of the algorithm, and hence we have $Eq(s(x_i),s(y_i))$. Define $s'(x_i y_i) := 1$ if $s(x_i) \neq s(y_i)$ and $s'(x_i y_i) := 0$ otherwise. Then $s'$ is a satisfying assignment for $\Psi$.

We still have to show that there exists a solution to $\Phi$ if the algorithm accepts. Let $s'\colon \binom{X}{2} \to \{0,1\}$ be a solution to $\Psi$. For each connected component $C$ in the graph $G$ at the final stage of the algorithm we pick two values $a_C, b_C \in C^2_\omega$ such that $Eq(a_C,b_C)$, and such that $N(a_C,d)$ and $N(a_C,d)$ for all previously picked values $d \in C^2_\omega$. Moreover, for each connected component $C$ of $G$ we pick a representative $r_C$. Define $s(r_C) := a_C$, and for $x \in C$ define $s(x) := a_C$ if $s'(xr_C) = 0$, and $s(x) := b_C$ otherwise.

Then $s$ satisfies all formulas in $\Psi$ that still contain disjuncts of the form $N(x_i,y_i)$, since these disjuncts will be satisfied by $s$. Formulas of the form $|\{i \in S : x_i \neq y_i\}| \equiv_2 p$ are satisfied, too, since $x_i$ and $y_i$ lie in the same connected component $C$, and hence $s(x_i) \neq s(y_i)$ iff $s'(xr_C) \neq s'(y_i r_C)$, which is the case iff $s'(xr_C) + s'(y_i r_C) = s'(x_i y_i) = 1$ because of the additional equations we have added to $\Psi$. Therefore, $|\{i \in S : x_i = y_i\}| \equiv_2 p$ iff $\sum_{i \in S} s'(x_i y_i) = p$. ■

**Corollary 7.10.** *Let $\Gamma$ be a reduct of $(C^2_\omega, Eq)$ with finite signature and such that* $\mathrm{Pol}(\Gamma)$ *contains the operation $h$. Then* $\mathrm{CSP}(\Gamma)$ *is in P.*

*Proof.* Direct consequence of Proposition 7.8 and Proposition 7.9. ■

## 8. Summary for the homogeneous equivalence relations

**Proposition 8.1.** *Let $\Gamma$ be a finite signature reduct of $(C^s_n, E)$, where either $3 \leq n < \omega$ or $3 \leq s < \omega$, and either $n$ or $s$ equals $\omega$. Then one of the following holds.*

1. *$\Gamma$ is homomorphically equivalent to a reduct of $(C^s_n, =)$, and $\mathrm{CSP}(\Gamma)$ is in P or NP-complete by [BK08].*
2. *$\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^s_n, E)}$, and $\mathrm{CSP}(\Gamma)$ is NP-complete.*

*Proof.* If $\Gamma$ has an endomorphism whose image is a clique or an independent set, then $\Gamma$ is homomorphically equivalent to a reduct of $(C^s_n, =)$ and the complexity classification is known from [BK08]. Otherwise, courtesy of Propositions 6.1 and 6.2, we may assume that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^s_n, E)}$, and that there is a pp-definition of $E$, $N$, and $Eq$ in $\Gamma$.

In the first case, that $Eq$ has a finite number $n$ of classes, we use Proposition 6.4 to see that the action of $\mathrm{Pol}(\Gamma)$ on the classes of $Eq$ has no essential and no constant operation. It follows that this action has a continuous projective clone homomorphism as in Definition 4.1, and hence so does $\mathrm{Pol}(\Gamma)$ itself, implying NP-completeness of $\mathrm{CSP}(\Gamma)$.

In the second case, that $Eq$ has classes of finite size $s \geq 3$, we use Proposition 6.5 to see that the action of $\mathrm{Pol}(\Gamma, c)$ on the equivalence class of any $c \in C^s_n$ has no essential and no constant operation, and hence has a continuous projective clone homomorphism. Consequently, so does $\mathrm{Pol}(\Gamma, c)$, implying NP-completeness of the CSP of $(\Gamma, c)$. Because $\Gamma$ is a model-complete core, this implies NP-completeness of $\mathrm{CSP}(\Gamma)$ [Bod07]. ∎

**Proposition 8.2.** *Suppose $\Gamma$ is a finite signature reduct of $(C^\omega_2, E)$; then either $\mathrm{CSP}(\Gamma)$ is in P or it is NP-complete.*

*Proof.* As in Proposition 8.1 we may assume that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^\omega_2, E)}$, and that $E$, $N$, and $Eq$ are pp-definable. We apply Proposition 6.6. The first two cases from that proposition imply NP-completeness of $\mathrm{CSP}(\Gamma)$, as in the proof of Proposition 8.1. The third case yields tractability as detailed in Section 7.1. ∎

**Proposition 8.3.** *Suppose $\Gamma$ is a finite signature reduct of $(C^2_\omega, E)$; then either $\mathrm{CSP}(\Gamma)$ is in P or it is NP-complete.*

*Proof.* As in Proposition 8.1 we may assume that $\mathrm{End}(\Gamma) = \overline{\mathrm{Aut}(C^2_\omega, E)}$, and that $E$, $N$, and $Eq$ are pp-definable. We apply Proposition 6.9. As before, the first two cases imply NP-completeness of $\mathrm{CSP}(\Gamma)$. The third case from Proposition 6.9 yields tractability as detailed in Theorem 7.2. ∎

Summarizing, we obtain a proof of Theorem 1.2.

*Proof of Theorem 1.2.* The statement follows from the preceding three propositions, together with [BW12] (for $C^\omega_\omega$) and [BK08] (for $C^1_\omega$ and $C^\omega_1$). ∎

## 9. Outlook

We have classified the computational complexity of CSPs for reducts of the infinite homogeneous graphs. Our proof shows that the scope of the classification method from [BP15a] is much larger than one might expect at first sight. The general research goal here is to identify larger and larger classes of infinite-domain CSPs where systematic complexity classification is possible; a general dichotomy conjecture is open for CSPs of reducts of finitely bounded homogeneous structures [BPP14, BOP15]. The next step in this direction might be to show a general complexity dichotomy for reducts of homogeneous structures whose age is finitely bounded and has the *free amalgamation property* (the Henson graphs provide natural examples for such structures). The present paper indicates that this problem might be within reach.

## References


[BCKvO09] Manuel Bodirsky, Hubie Chen, Jan Kára, and Timo von Oertzen. Maximal infinite-valued constraint languages. *Theoretical Computer Science (TCS)*, 410:1684–1693, 2009. A preliminary version appeared at ICALP'07.

[BJP16] Manuel Bodirsky, Peter Jonsson, and Trung Van Pham. The complexity of phylogeny constraint satisfaction. In *Proceedings of the Symposium on Theoretical Aspects of Computer Science (STACS)*, 2016. Preprint arXiv:1503.07310.

[BK08] Manuel Bodirsky and Jan Kára. The complexity of equality constraint languages. *Theory of Computing Systems*, 3(2):136–158, 2008. A conference version appeared in the proceedings of Computer Science Russia (CSR'06).

[BK09] Manuel Bodirsky and Jan Kára. The complexity of temporal constraint satisfaction problems. *Journal of the ACM*, 57(2):1–41, 2009. An extended abstract appeared in the Proceedings of the Symposium on Theory of Computing (STOC'08).

[BKJ05] Andrei A. Bulatov, Andrei A. Krokhin, and Peter G. Jeavons. Classifying the complexity of constraints using finite algebras. *SIAM Journal on Computing*, 34:720–742, 2005.

[BKN09] Libor Barto, Marcin Kozik, and Todd Niven. The CSP dichotomy holds for digraphs with no sources and no sinks (a positive answer to a conjecture of Bang-Jensen and Hell). *SIAM Journal on Computing*, 38(5), 2009.

[BM16] Manuel Bodirsky and Antoine Mottet. Reducts of finitely bounded homogeneous structures, and lifting tractability from finite-domain constraint satisfaction. Submitted. Preprint available under ArXiv:1601.04520, 2016.

[BMM15] Manuel Bodirsky, Barnaby Martin, and Antoine Mottet. Constraint satisfaction problems over the integers with successor. In *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part I*, pages 256–267, 2015.

[BN06] Manuel Bodirsky and Jaroslav Nešetřil. Constraint satisfaction with countable homogeneous templates. *Journal of Logic and Computation*, 16(3):359–373, 2006.

[Bod07] Manuel Bodirsky. Cores of countably categorical structures. *Logical Methods in Computer Science*, 3(1):1–16, 2007.

[Bod12] Manuel Bodirsky. Complexity classification in infinite-domain constraint satisfaction. Mémoire d'habilitation à diriger des recherches, Université Diderot – Paris 7. Available at arXiv:1201.0856, 2012.

[BOP15] Libor Barto, Jakub Opršal, and Michael Pinsker. The wonderland of reflections. Preprint arXiv:1510.04521, 2015.

[BP11] Manuel Bodirsky and Michael Pinsker. Reducts of Ramsey structures. *AMS Contemporary Mathematics, vol. 558 (Model Theoretic Methods in Finite Combinatorics)*, pages 489–519, 2011.

[BP14] Manuel Bodirsky and Michael Pinsker. Minimal functions on the random graph. *Israel Journal of Mathematics*, 200(1):251–296, 2014.

[BP15a] Manuel Bodirsky and Michael Pinsker. Schaefer's theorem for graphs. *Journal of the ACM*, 62(3):Article no. 19, 1–52, 2015. A conference version appeared in the Proceedings of STOC 2011, pages 655–664.

[BP15b] Manuel Bodirsky and Michael Pinsker. Topological Birkhoff. *Transactions of the American Mathematical Society*, 367:2527–2549, 2015.

[BPP14] Manuel Bodirsky, Michael Pinsker, and András Pongrácz. Projective clone homomorphisms. Preprint arXiv:1409.4601, 2014.

[BPT13] Manuel Bodirsky, Michael Pinsker, and Todor Tsankov. Decidability of definability. *Journal of Symbolic Logic*, 78(4):1036–1054, 2013. A conference version appeared in the Proceedings of LICS 2011, pages 321–328.

[Bul06] Andrei A. Bulatov. A dichotomy theorem for constraint satisfaction problems on a 3-element set. *Journal of the ACM*, 53(1):66–120, 2006.

[BW12] Manuel Bodirsky and Michał Wrona. Equivalence constraint satisfaction problems. In *Proceedings of Computer Science Logic*, volume 16 of *LIPICS*, pages 122–136. Dagstuhl Publishing, September 2012.

[FV99] Tomás Feder and Moshe Y. 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.

[GP08] Martin Goldstern and Michael Pinsker. A survey of clones on infinite sets. *Algebra Universalis*, 59:365–403, 2008.

[HN90] Pavol Hell and Jaroslav Nešetřil. On the complexity of H-coloring. *Journal of Combinatorial Theory, Series B*, 48:92–110, 1990.

[HR94] L. Haddad and Ivo G. Rosenberg. Finite clones containing all permutations. *Canadian Journal of Mathematics*, 46(5):951–970, 1994.

[JCG97] Peter Jeavons, David Cohen, and Marc Gyssens. Closure properties of constraints. *Journal of the ACM*, 44(4):527–548, 1997.

[LW80] Alistair H. Lachlan and Robert E. Woodrow. Countable ultrahomogeneous undirected graphs. *Transactions of the AMS*, 262(1):51–94, 1980.

[NR89] Jaroslav Nešetřil and Vojtěch Rödl. The partite construction and Ramsey set systems. *Discrete Mathematics*, 75(1-3):327–334, 1989.

[Pon11] András Pongrácz. Reducts of the Henson graphs with a constant. Preprint, 2011.

[Pos41] Emil L. Post. The two-valued iterative systems of mathematical logic. *Annals of Mathematics Studies*, 5, 1941.

[Sch78] Thomas J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the Symposium on Theory of Computing (STOC)*, pages 216–226, 1978.

[Sze86] Ágnes Szendrei. *Clones in universal algebra*. Séminaire de Mathématiques Supérieures. Les Presses de l'Université de Montréal, 1986.

[Tho91] Simon Thomas. Reducts of the random graph. *Journal of Symbolic Logic*, 56(1):176–181, 1991.

[TZ12] Katrin Tent and Martin Ziegler. *A course in model theory*. Lecture Notes in Logic. Cambridge University Press, 2012.

Institut für Algebra, TU Dresden, 01062 Dresden, Germany
*E-mail address*: Manuel.Bodirsky@tu-dresden.de
*URL*: http://www.math.tu-dresden.de/~bodirsky/

School of Science and Technology, Middlesex University, The Burroughs, London NW4 4BT, United Kingdom
*E-mail address*: B.Martin@mdx.ac.uk
*URL*: http://www.eis.mdx.ac.uk/staffpages/barnabymartin/

Department of Algebra, MFF UK, Sokolovska 83, 186 00 Praha 8, Czech Republic
*E-mail address*: marula@gmx.at
*URL*: http://dmg.tuwien.ac.at/pinsker/

Department of Algebra and Number Theory, University of Debrecen, 4032 Debrecen, Egyetem square 1, Hungary
*E-mail address*: pongracz.andras@science.unideb.hu