The Expressive Power of CSP-Quantifiers

Lauri Hella
Faculty of Information Technology and Communication Sciences, Tampere University, Finland

Abstract

A generalized quantifier $Q_{\mathcal{K}}$ is called a CSP-quantifier if its defining class $\mathcal{K}$ consists of all structures that can be homomorphically mapped to a fixed finite template structure. For all positive integers $n \geq 2$ and k, we define a pebble game that characterizes equivalence of structures with respect to the logic $L^k_{\infty\omega}(\mathbf{CSP}^+_n)$, where $\mathbf{CSP}^+_n$ is the union of the class $\mathbf{Q}_1$ of all unary quantifiers and the class $\mathbf{CSP}_n$ of all CSP-quantifiers with template structures that have at most n elements. Using these games we prove that for every $n \geq 2$ there exists a CSP-quantifier with template of size $n+1$ which is not definable in $L^\omega_{\infty\omega}(\mathbf{CSP}^+_n)$. The proof of this result is based on a new variation of the well-known Cai-Fürer-Immerman construction.

2012 ACM Subject Classification Theory of computation → Finite Model Theory

Keywords and phrases generalized quantifiers, constraint satisfaction problems, pebble games, finite variable logics, descriptive complexity theory

Digital Object Identifier 10.4230/LIPIcs.CSL.2023.25

1 Introduction

The present paper continues the research line in descriptive complexity theory that was originated in the seminal paper [3] of Cai, Fürer and Immerman. Using a pebble game characterization, introduced in [11], for the infinitary k-variable logic with counting, $C^k_{\infty\omega}$, they proved that there are PTIME-computable properties that are not definable in $C^\omega_{\infty\omega} = \bigcup_{k\in\omega} C^k_{\infty\omega}$. Since fixed point logic with counting, IFPC, is contained in $C^\omega_{\infty\omega}$, they obtained as a corollary that IFPC does not capture PTIME on finite unordered structures. However, perhaps the most important contribution in [3] is the clever technique for constructing pairs $(\mathfrak{A}_k, \mathfrak{B}_k)$ of non-isomorphic structures such that Duplicator has a winning strategy in the pebble game with k pebbles. This *Cai-Fürer-Immerman* (CFI) construction has been later adapted for several different pebble games that characterize various extensions of $C^\omega_{\infty\omega}$.

The counting logic $C^k_{\infty\omega}$ is obtained by adding the counting quantifiers $\exists^{\geq m}x$ ("there are at least m elements x such that") to the corresponding infinitary k-variable logic $L^k_{\infty\omega}$. The counting quantifiers are examples of unary generalized quantifiers as they bind a single variable in the formula that follows. Generalized quantifiers can also bind variables in several formulas simultaneously, as well as several variables in each of the formulas. An example of the former is the Härtig quantifier $I\,x, y\,(\varphi(x), \psi(y))$ stating that "the number of x satisfying φ is the same as the number of y satisfying ψ". As an example of the latter, if $\mathcal{C}$ is the class of connected graphs, then $Q_{\mathcal{C}}$ is a generalized quantifier binding two variables in a formula, and $Q_{\mathcal{C}}xy\,\varphi(x, y)$ has the meaning "the binary relation defined by the formula $\varphi(x, y)$ is the edge relation of a connected graph". More generally, any isomorphism closed class $\mathcal{K}$ of structures in a finite relational vocabulary can be used as an interpretation of a generalized quantifier $Q_{\mathcal{K}}$. The quantifier $Q_{\mathcal{K}}$ is r-ary, if it binds at most r variables in each formula.

Since $C^\omega_{\infty\omega}$ turned out to be too weak to define all PTIME computable properties of finite structures, it was natural to study the expressive power of extensions of $L^\omega_{\infty\omega}$ with quantifiers of arity more than 1. For this purpose we introduced in [10] the *r-bijective k-pebble game* that characterizes equivalence with respect to the logic $L^k_{\infty\omega}(\mathbf{Q}_r)$, the extension of $L^k_{\infty\omega}$ by the class $\mathbf{Q}_r$ of all generalized quantifiers of arity at most r. Furthermore, we proved that, for any positive integer r, there is a PTIME-computable generalized quantifier of arity

$r+1$ which is not definable in $L^{\omega}_{\infty\omega}(\mathbf{Q}_r)$. As a corollary, we got the result that there is no set $\mathbf{Q}$ of generalized quantifiers of bounded arity such that IFP($\mathbf{Q}$) captures PTIME. The proof of these results is based on the n-bijective k-pebble game and a variation of the CFI construction, where the binary edge relation in the so-called "gadget graphs" is replaced by an $r+1$-ary relation. Since $C^{\omega}_{\infty\omega}$ has the same expressive power as $L^{\omega}_{\infty\omega}(\mathbf{Q}_1)$, the results in [10] are a natural extension of those in [3].

More than a decade after the publication of [10], research on extensions of $C^{\omega}_{\infty\omega}$, pebble games and CFI constructions became active again, when Dawar, Grohe, Holm and Laubner [5] introduced the extension IFPR of IFP by rank operators. A rank operator rk_p binds two tuples $\vec{x}$ and $\vec{y}$ of variables in a formula φ, and outputs the rank of the matrix defined by φ over the field F_p, where values of $\vec{x}$ are thought as columns and values of $\vec{y}$ as rows. Dawar and Holm [6] defined a pebble game that characterizes equivalence with respect to the extension of $L^{k}_{\infty\omega}$ with the rank operators, and also the so-called invertible map game, which corresponds to a stronger equivalence. The invertible map game was later shown in [4] to characterize equivalence with respect to the extension of $L^{k}_{\infty\omega}$ by all linear algebraic operators, i.e., operators that are invariant under similarity mappings (with respect to the field F_p considered).

The rank logic IFPR was shown to fall short of capturing PTIME by Grädel and Pakusa [9]; their proof is based on a variation of the CFI construction. They also suggested a stronger version IFPR* of rank logic that allows the parameter p in the operator rk_p to be given by a term. The question whether IFPR* captures PTIME was open for a few years, but it was recently settled in the negative by Lichter ([14]), who used the invertible map game mentioned above to show that a generalized CFI construction produces non-isomorphic structures that cannot be separated in the extension of $L^{k}_{\infty\omega}$ with the stronger rank operator of [9].

Another application of the CFI construction was given by Atserias, Bulatov and Dawar [1], who observed that the CFI structures $\mathfrak{A}_k$ and $\mathfrak{B}_k$ of [3] can be separated by PTIME-computable constraint satisfaction problems that arise from affine systems of equations. The same is true for the CFI construction in [10]: for each r there is a constraint satisfaction problem r-CFI that separates the $L^{k}_{\infty\omega}(\mathbf{Q}_r)$-equivalent structures obtained by the construction. The problem r-CFI is an instance of solving systems of equations over F_2, and hence it is PTIME-computable.

A constraint satisfaction problem CSP($\mathfrak{C}$) with a finite template structure $\mathfrak{C}$ is the class of all finite structures in the same vocabulary as $\mathfrak{C}$ that can be homomorphically mapped to $\mathfrak{C}$. By a CSP-quantifier we mean a generalized quantifier $Q_{\mathcal{K}}$ such that its defining class $\mathcal{K}$ is of the form CSP($\mathfrak{C}$). In this paper we use two numerical parameters for classifying CSP-quantifiers: the *arity* of $Q_{\mathsf{CSP}(\mathfrak{C})}$ is the maximum arity of relations in $\mathfrak{C}$, and the *size* of $Q_{\mathsf{CSP}(\mathfrak{C})}$ is the number of elements in the universe of $\mathfrak{C}$. Thus, the observation above concerning the CFI structures in [10] can be formulated as an *arity hierarchy* for CSP-quantifiers: for any positive integer r, there exists a PTIME-computable CSP-quantifier $Q_{\mathsf{CSP}(\mathfrak{C})}$ of arity $r+1$ and size 2 which is not definable in $L^{\omega}_{\infty\omega}(\mathbf{Q}_r)$, and a fortiori, not definable in the extension of $L^{\omega}_{\infty\omega}$ by all CSP-quantifiers of arity at most r.

In this paper, we turn attention to the size of CSP-quantifers. Given an integer $n \geq 2$, we define $\mathbf{CSP}_n$ to be the class of all CSP-quantifiers of size at most n. We introduce a pebble game CSPG$[\mathfrak{A},\mathfrak{B},n,k]$, and prove that it characterizes equivalence of the structures $\mathfrak{A}$ and $\mathfrak{B}$ with respect to $L^{k}_{\infty\omega}(\mathbf{CSP}^{+}_n)$, where $\mathbf{CSP}^{+}_n$ is the union of $\mathbf{CSP}_n$ and the class of all unary quantifiers $\mathbf{Q}_1$. We also define another pebble game that we call *bijective colouring game*, BCG$[\mathfrak{A},\mathfrak{B},n,k]$. The game BCG has simpler rules than CSPG, but it can still be used for proving undefinability results for $L^{\omega}_{\infty\omega}(\mathbf{CSP}^{+}_n)$: we prove that if Duplicator has a winning strategy in BCG$[\mathfrak{A},\mathfrak{B},n,k]$, then she has one also in CSPG$[\mathfrak{A},\mathfrak{B},n,k]$.

As our main result in the paper, we prove a *size hierarchy* theorem for CSP-quantifiers: for any $n \geq 2$ there exists a CSP-quantifier $Q_{\mathsf{CSP}(\mathfrak{C}_n)}$ of size $n+1$ which is not definable in $L^{\omega}_{\infty\omega}(\mathbf{CSP}^+_n)$. The proof of this result is based on a new variation of the CFI construction that produces for each input graph G two structures $\mathfrak{A}^{\mathrm{ev}}_n(G)$ and $\mathfrak{A}^{\mathrm{od}}_n(G)$ such that Duplicator has a winning strategy in the bijective colouring game $\mathrm{BCG}[\mathfrak{A}^{\mathrm{ev}}_n(G), \mathfrak{A}^{\mathrm{od}}_n(G), n, k]$, assuming that G is large enough. On the other hand, $\mathfrak{A}^{\mathrm{ev}}_n(G) \in \mathsf{CSP}(\mathfrak{C}_n)$ and $\mathfrak{A}^{\mathrm{od}}_n(G) \notin \mathsf{CSP}(\mathfrak{C}_n)$ for all graphs G.

The size hierarchy result differs from the arity hierarchy result in two important aspects. First, the membership problem of $\mathsf{CSP}(\mathfrak{C}_n)$ is NP-complete. It is an open problem whether $\mathsf{CSP}(\mathfrak{C}_n)$ can be replaced in the size hierarchy result by some PTIME-computable $\mathsf{CSP}(\mathfrak{D}_n)$. Second, the arity of $\mathsf{CSP}(\mathfrak{C}_n)$ depends on n; in fact, $\mathfrak{C}_n$ has a single relation which is $3n$-ary. We do not know if the size hierarchy result holds for CSP-quantifiers of arity r for some fixed r. It should also be mentioned here that we do not include vectorization in the definition of logics with generalized quantifiers, unlike is done in many recent papers on the topic. The pebble game CSPG could be adapted for vectorized quantifiers, but using such games would probably be too difficult to handle, as they involve existential second order quantification of ℓ-ary relations for $\ell \geq 2$. Furthermore, it seems quite possible that equivalence with respect to $L^k_{\infty\omega}$ with the second vectorization of CSP-quantifiers is just isomorphism.

The structure of the paper is the following. We explain the necessary background on logics, constraint satisfaction problems and generalized quantifiers in Sections 2 and 3. Section 4 is devoted to the definitions of the pebble games CSPG and BCG, and the corresponding characterizations of equivalence with respect to the logic $L^k_{\infty\omega}(\mathbf{CSP}^+_n)$. In Section 5, we describe our generalized CFI construction, and in Section 6, we introduce the CSP-quantifer that separates the constructed structures $\mathfrak{A}^{\mathrm{ev}}_n(G)$ and $\mathfrak{A}^{\mathrm{od}}_n(G)$. The proof that Duplicator has a winning strategy in the game $\mathrm{BCG}[\mathfrak{A}^{\mathrm{ev}}_n(G), \mathfrak{A}^{\mathrm{od}}_n(G), n, k]$ is given in Section 7. Finally, in Section 8, we give a brief overview of our contributions in the paper, and list some open problems.

2 Preliminaries

In this section we go briefly through definitions of basic notions concerning logics and constraint satisfaction problems. For a more comprehensive exposition on first-order logic and finite variable logic, we refer to the excellent textbooks [7] and [13]. Fo more information on constraint satisfaction problems from a logical point of view we refer to [12]. We start by giving some notational and other conventions.

2.1 Notation and conventions

For a positive integer n, we denote the set $\{1, \ldots, n\}$ by $[n]$. The cardinality of a (finite) set A is denoted by $|A|$. If $f\colon A \to B$ is a function, and $\vec{a} = (a_1, \ldots, a_r) \in A^r$, then we use the shorthand notation $f(\vec{a}) := (f(a_1), \ldots, f(a_r))$.

All vocabularies considered in this paper are *finite* and *relational*, i.e., they are of the form $\tau = \{R_1, \ldots, R_n\}$, where each R_i is a relation symbol. The arity of R_i is denoted by $\mathrm{ar}(R_i)$. A τ-structure is then a tuple $\mathfrak{A} = (A, R_1^{\mathfrak{A}}, \ldots, R_n^{\mathfrak{A}})$, where $R_i^{\mathfrak{A}} \subseteq A^{\mathrm{ar}(R_i)}$ for each $i \in [n]$. All our results are in the context of finite model theory; thus, we will assume without further notice that the universe A of any structure $\mathfrak{A}$ we consider is a finite set.

Given a τ-structure $\mathfrak{A}$, an *assignment on* $\mathfrak{A}$ is a function $\alpha\colon X \to A$, where $\mathrm{dom}(\alpha) = X$ is some finite set of variables. If y is a variable and $a \in A$, then $\alpha[a/y]$ denotes the modified assignment on $\mathfrak{A}$ with $\mathrm{dom}(\alpha[a/y]) = X \cup \{y\}$ such that $\alpha[a/y](x) = \alpha(x)$ for all $x \in X \setminus \{y\}$, and $\alpha[a/y](y) = a$.

Let $\mathfrak{A}$ and $\mathfrak{B}$ be τ-structures. A *partial isomorphism* from $\mathfrak{A}$ to $\mathfrak{B}$ is a bijection $p\colon C \to D$ such that $C \subseteq A$, $D \subseteq B$, and for all $R \in \tau$ and all $\vec{a} \in C^{\mathrm{ar}(R)}$, $\vec{a} \in R^{\mathfrak{A}} \iff p(\vec{a}) \in R^{\mathfrak{B}}$. We denote the set of all partial isomorphisms from $\mathfrak{A}$ to $\mathfrak{B}$ by $\mathrm{PI}(\mathfrak{A}, \mathfrak{B})$. If α is an assignment on $\mathfrak{A}$ and β is an assignment on $\mathfrak{B}$ such that $\mathrm{dom}(\alpha) = \mathrm{dom}(\beta)$, then we denote the relation $\{(\alpha(x), \beta(x)) \mid x \in \mathrm{dom}(\alpha)\} \subseteq A \times B$ by $\alpha \mapsto \beta$. Thus, $\alpha \mapsto \beta \in \mathrm{PI}(\mathfrak{A}, \mathfrak{B})$ if this relation is a bijection that preserves the relations in τ.

An *n-colouring* of a set T is a function $g\colon T \to [n]$. If $\mathfrak{A}$ is a τ-structure and g is an n-colouring of its universe A, we define $\mathfrak{A}^g$ to be the $\tau \cup \{S_1, \ldots, S_n\}$-structure with the same universe such that $R^{\mathfrak{A}^g} := R^{\mathfrak{A}}$ for all $R \in \tau$, and $S_i^{\mathfrak{A}^g} := \{a \in A \mid g(a) = i\}$ for each $i \in [n]$. Note that if $\mathfrak{B}^h$ is another τ-structure with an n-colouring h, and $p \in \mathrm{PI}(\mathfrak{A}, \mathfrak{B})$, then $p \in \mathrm{PI}(\mathfrak{A}^g, \mathfrak{B}^h)$ if and only if p preserves colouring: $g(a) = h(p(a))$ for all $a \in \mathrm{dom}(p)$.

We define graphs as ordered pairs $G = (V, E)$, where V is a nonempty set of vertices, and E is a set of *unordered pairs* of vertices, called edges; i.e., $E \subseteq \{\{u, v\} \mid u, v \in V,\ u \neq v\}$. Thus, a graph G *is not* a relational structure, but it can be represented as one by setting $\mathfrak{A}_G := (V, R_E)$, where $R_E := \{(u, v) \mid \{u, v\} \in E\}$.

2.2 Logics

First-order logic, FO, is defined as usually (see, e.g., [13]). The infinitary logic $L_{\infty\omega}$ is the extension of FO that allows disjunctions $\bigvee \Psi$ and conjunctions $\bigwedge \Psi$ of arbitrary sets Ψ of formulas. Since we assume that all structures are finite, it suffices to consider infinite disjunctions and conjunctions only over countable sets Ψ.

For each positive integer k, the *infinitary k-variable logic* $L^k_{\infty\omega}$ is the fragment of $L_{\infty\omega}$ consisting of formulas that contain at most k different variables; we assume throughout that the variables allowed in the formulas of $L^k_{\infty\omega}$ come from a fixed set $X_k := \{x_1, \ldots, x_k\}$. The k-variable first-order logic, FO^k, is defined analogously. The *finite variable logic* $L^\omega_{\infty\omega}$ is the union of $L^k_{\infty\omega}$ over positive integers k.

All logics L we consider are extensions of FO, $L^\omega_{\infty\omega}$, or their k-variable fragments with generalized quantifiers. Given a τ-structure $\mathfrak{A}$, an assignment $\alpha\colon X \to A$ on $\mathfrak{A}$, and a τ-formula $\varphi \in L$ with free variables in X, we write $(\mathfrak{A}, \alpha) \models \varphi$, if φ is true in $\mathfrak{A}$ under the interpretation α. If φ is a sentence and α is the empty assignment $\emptyset$, we write $\mathfrak{A} \models \varphi$ instead of $(\mathfrak{A}, \emptyset) \models \varphi$.

2.3 Constraint satisfaction problems

Let τ be a vocabulary, and let $\mathfrak{A}$ and $\mathfrak{B}$ be τ-structures. A function $h\colon A \to B$ is a *homomorphism* $\mathfrak{A} \to \mathfrak{B}$, if the implication

$$\vec{a} \in R^{\mathfrak{A}} \implies h(\vec{a}) \in R^{\mathfrak{B}}$$

holds for every relation symbol $R \in \tau$ and every tuple $\vec{a} \in A^{\mathrm{ar}(R)}$. The *uniform* Constraint Satisfaction Problem (CSP) on τ-structures asks whether there exists a homomorphism $h\colon \mathfrak{A} \to \mathfrak{B}$ for a pair $(\mathfrak{A}, \mathfrak{B})$ of input structures of vocabulary τ. It is well known that this problem is NP-complete, assuming that τ contains at least one relation symbol R with $\mathrm{ar}(R) \geq 2$.

We consider in this paper the *non-uniform* version of CSP, in which the second structure, called the template of the problem, is fixed, and only the first structure is given as input. For each template structure $\mathfrak{C}$ of vocabulary τ, the class of positive instances of the corresponding non-uniform CSP is denoted by $\mathsf{CSP}(\mathfrak{C})$. Thus, $\mathsf{CSP}(\mathfrak{C})$ consists of all τ-structures $\mathfrak{A}$ such that there exists a homomorphism $h\colon \mathfrak{A} \to \mathfrak{C}$.

We classify template structures $\mathfrak{C}$ by two numerical parameters:

- The *arity* of $\mathfrak{C}$ is $\mathrm{ar}(\mathfrak{C}) := \max\{\mathrm{ar}(R) \mid R \in \tau\}$, where τ is the vocabulary of $\mathfrak{C}$.
- The *size* of $\mathfrak{C}$ is $\mathrm{sz}(\mathfrak{C}) := |C|$.

We will henceforth assume without loss of generality that the universe of any template $\mathfrak{C}$ with $\mathrm{sz}(\mathfrak{C}) = n \geq 3$ is $[n]$; in case $\mathrm{sz}(\mathfrak{C}) = 2$, we may also use the Boolean universe $\{0, 1\}$.

As an example, consider the n-colourability problem n-COL of graphs. Clearly a graph G is n-colourable if and only if there is a homomorphism from $\mathfrak{A}_G$ to the structure $\mathfrak{C}_{n\text{-}\mathsf{COL}} := ([n], \{(i, j) \in [n]^2 \mid i \neq j\})$. Thus, we can identify n-COL with $\mathsf{CSP}(\mathfrak{C}_{n\text{-}\mathsf{COL}})$. The arity and size of $\mathfrak{C}_{n\text{-}\mathsf{COL}}$ are 2 and n, respectively. It is well known that for $n \geq 3$, n-COL is NP-complete, while for $n = 2$ it is in $\mathsf{LOGSPACE}$.

Another, highly relevant example is related to the generalization of the CFI construction given in [10]: a close inspection of the structures $\mathbf{A}(\mathbf{G})$ and $\mathbf{B}(\mathbf{G})$ constructed in Section 8 of that paper reveals that they can be separated by a CSP. More precisely, $\mathbf{A}(\mathbf{G}) \in \mathsf{CSP}(\mathfrak{C}_{n\text{-}\mathsf{CFI}})$ and $\mathbf{B}(\mathbf{G}) \notin \mathsf{CSP}(\mathfrak{C}_{n\text{-}\mathsf{CFI}})$, where $\mathfrak{C}_{n\text{-}\mathsf{CFI}} = (\{0, 1\}, R^{\mathrm{ev}}, \{(0, 0), (1, 1)\})$ for $R^{\mathrm{ev}} := \{(b_1, \ldots, b_{n+1}) \mid b_1 + \cdots + b_{n+1} = 0 \mod 2\}$.[1] The arity and size of $\mathfrak{C}_{n\text{-}\mathsf{CFI}}$ are $n + 1$ and 2, respectively. The problem n-$\mathsf{CFI} := \mathsf{CSP}(\mathfrak{C}_{n\text{-}\mathsf{CFI}})$ can be solved by Gaussian elimination, whence it is PTIME-computable.

The complexity of non-uniform CSP has been studied intensively for more than three decades. This is because a large variety of real-world problems can be formulated as CSPs, and hence it is important to understand the borderline between feasible and unfeasible cases. Research on the topic culminated recently in the proof by Bulatov [2] and Zhuk [16] of the Dichotomy Conjecture formulated by Feder and Vardi in [8]: for any template $\mathfrak{C}$, $\mathsf{CSP}(\mathfrak{C})$ is either in PTIME, or NP-complete.

3 Generalized quantifiers

The notion of generalized quantifier was originally defined by Lindström [15] in 1966. Thus, generalized quantifiers are often called *Lindström quantifiers*. We go here quickly through the definitions and notations concerning generalized quantifiers, and then introduce the special case of CSP-quantifiers. For more detailed treatment of quantifiers we refer to [7] and [10].

Let $\tau = \{R_1, \ldots, R_m\}$ be a relational vocabulary, and let $\mathrm{ar}(R_i) = r_i$ for each $i \in [m]$. To any class $\mathcal{K}$ of τ-structures that is closed under isomorphisms, we assign a *generalized quantifier* (or a *Lindström quantifier*) $Q_{\mathcal{K}}$. The extension $L(Q_{\mathcal{K}})$ of a logic L by $Q_{\mathcal{K}}$ is obtained by adding the following rules in the syntax and semantics of L:

- If $\psi_1, \ldots, \psi_m$ are formulas and $\vec{y}_1, \ldots, \vec{y}_m$ are tuples of variables with $|\vec{y}_i| = r_i$ for $i \in [m]$, then $\varphi = Q_{\mathcal{K}}\vec{y}_1, \ldots, \vec{y}_m\, (\psi_1, \ldots, \psi_m)$ is a formula. A variable is free in φ if, for some $i \in [m]$, it is free in ψ_i but does not occur in $\vec{y}_i$.
- $(\mathfrak{A}, \alpha) \models Q_{\mathcal{K}}\vec{y}_1, \ldots, \vec{y}_m\, (\psi_1, \ldots, \psi_m)$ if and only if $(A, \psi_1^{\mathfrak{A},\alpha,\vec{y}_1}, \ldots, \psi_m^{\mathfrak{A},\alpha,\vec{y}_m}) \in \mathcal{K}$, where $\theta^{\mathfrak{A},\alpha,\vec{y}} := \{\vec{a} \in A^r \mid (\mathfrak{A}, \alpha[\vec{a}/\vec{y}]) \models \theta\}$ for a formula θ and an r-tuple $\vec{y}$ of variables.

The extension $L(\mathbf{Q})$ of L by a class $\mathbf{Q}$ of generalized quantifiers is defined by adding the rules above to L for each $Q_{\mathcal{K}} \in \mathbf{Q}$.

[1] This was observed in hindsight after the paper [1]. At the time we wrote [10] we were not familiar with the literature on CSP.

Let $Q_{\mathcal{K}}$ and $Q_{\mathcal{K}'}$ be generalized quantifiers. We say that $Q_{\mathcal{K}}$ is *definable* in $L(Q_{\mathcal{K}'})$ if the defining class $\mathcal{K}$ is definable in $L(Q_{\mathcal{K}'})$, i.e., there is a sentence φ of $L(Q_{\mathcal{K}'})$ such that $\mathcal{K} = \{\mathfrak{A} \mid \mathfrak{A} \models \varphi\}$.

The *type* of the quantifier $Q_{\mathcal{K}}$ is $(r_1, \ldots, r_m)$, and the *arity* of $Q_{\mathcal{K}}$ is $\max\{r_1, \ldots, r_m\}$. For the sake of simplicity, we assume from now onwards that the type of $Q_{\mathcal{K}}$ is *uniform*, i.e., $r_i = r_j$ for all $i, j \in [m]$. This is no loss of generality, since for any quantifier $Q_{\mathcal{K}}$ there is another quantifier $Q_{\mathcal{K}'}$ of uniform type with the same arity such that $Q_{\mathcal{K}}$ is definable in $\mathrm{FO}(Q_{\mathcal{K}'})$ and $Q_{\mathcal{K}'}$ is definable in $\mathrm{FO}(Q_{\mathcal{K}})$.

Furthermore, we restrict the syntactic rule of $Q_{\mathcal{K}}$ by requiring that $\vec{y}_i = \vec{y}_j$ for all $i, j \in [m]$. Then we can denote the formula obtained by applying the rule simply by $\varphi = Q_{\mathcal{K}}\vec{y}\,(\psi_1, \ldots, \psi_m)$. Note however, that this convention disallows formulas of the type $\theta = Qx, y\,(R(x, y), R(y, x))$ in which both x and y remain free even though x is bound in $R(x, y)$ and y is bound in $R(y, x)$, and hence weakens the expressive power of $\mathrm{FO}^k(Q_{\mathcal{K}})$ and $L^k_{\infty\omega}(Q_{\mathcal{K}})$. Fortunately the loss can be compensated by using more variables (e.g., θ is equivalent with $Qz\,(R(z, y), R(z, x))$. Hence the restriction does not affect the expressive power of $\mathrm{FO}(Q_{\mathcal{K}})$ and $L^\omega_{\infty\omega}(Q_{\mathcal{K}})$.

▶ **Definition 1.** *Let r and $n \geq 2$ be positive integers.*

(a) *We denote the class of all generalized quantifiers $Q_{\mathcal{K}}$ of arity at most r by $\mathbf{Q}_r$.*

(b) *A generalized quantifier $Q_{\mathcal{K}}$ is a* CSP-quantifier *if its defining class $\mathcal{K}$ is* $\mathsf{CSP}(\mathfrak{C})$ *for some template structure $\mathfrak{C}$. We will denote $Q_{\mathsf{CSP}(\mathfrak{C})}$ simply by $Q_{\mathfrak{C}}$.*

(c) *We denote the class of all CSP-quantifiers $Q_{\mathfrak{C}}$ such that* $\mathrm{sz}(\mathfrak{C}) \leq n$ *by $\mathbf{CSP}_n$.*

(d) *We write $\mathbf{CSP}^+_n := \mathbf{Q}_1 \cup \mathbf{CSP}_n$.*

For example $Q_{n\text{-}\mathsf{COL}} \in \mathbf{Q}_2 \cap \mathbf{CSP}_n$ for each $n \geq 2$, and $Q_{r\text{-}\mathsf{CFI}} \in \mathbf{Q}_{r+1} \cap \mathbf{CSP}_2$ for each r. As mentioned in the previous section, r-CFI is PTIME-computable, and it separates the structures used in the proof of the main result in [10] (Theorem 8.6). Thus, the corresponding CSP-quantifiers form a strong hierarchy: $Q_{r\text{-}\mathsf{CFI}}$ is not definable in $L^\omega_{\infty\omega}(\mathbf{Q}_r)$ for any $r \geq 2$.

In the present paper we will prove a similar hierarchy result for CSP-quantifiers in terms of the size parameter: for every $n \geq 2$ we define a template structure $\mathfrak{C}_n$ with $\mathrm{sz}(\mathfrak{C}_n) = n + 1$ such that $Q_{\mathfrak{C}_n}$ is not definable in $L^\omega_{\infty\omega}(\mathbf{CSP}^+_n)$. It is well known that *inflationary fixed point logic* IFP is contained in $L^\omega_{\infty\omega}$, and this remains valid in the presence of any additional quantifiers. Moreover, *inflationary fixed point logic with counting* IFPC is contained in $L^\omega_{\infty\omega}(\mathbf{Q}_1)$. Thus, as a corollary, we obtain that $Q_{\mathfrak{C}_n}$ is not definable in $\mathrm{IFPC}(\mathbf{CSP}_n)$.

This corollary is the main reason for considering $\mathbf{CSP}^+_n$ instead of just $\mathbf{CSP}_n$. Without the class $\mathbf{Q}_1$ of all unary quantifiers we would only get the undefinability of $Q_{\mathfrak{C}_n}$ in $\mathrm{IFP}(\mathbf{CSP}_n)$, and $\mathrm{IFP}(\mathbf{CSP}_n)$ is strictly less expressive than $\mathrm{IFPC}(\mathbf{CSP}_n)$. For example, using a straightforward quantifier elimination argument we can show that no sets A and B with at least kn elements can be separated by a sentence of $L^\omega_{\infty\omega}(\mathbf{CSP}_n)$; thus, e.g. even cardinality can not be defined in $\mathrm{IFP}(\mathbf{CSP}_n)$.

Given two structures $\mathfrak{A}$ and $\mathfrak{B}$ of the same vocabulary, and assignments α and β on $\mathfrak{A}$ and $\mathfrak{B}$, respectively, such that $\mathrm{dom}(\alpha) = \mathrm{dom}(\beta)$, we write $(\mathfrak{A}, \alpha) \equiv^k_{\infty\omega,n} (\mathfrak{B}, \beta)$ if the equivalence

$$(\mathfrak{A}, \alpha) \models \varphi \iff (\mathfrak{B}, \beta) \models \varphi$$

holds for all formulas $\varphi \in L^k_{\infty\omega}(\mathbf{CSP}^+_n)$ with free variables in $\mathrm{dom}(\alpha)$. Similarly we write $(\mathfrak{A}, \alpha) \equiv^k_n (\mathfrak{B}, \beta)$ if the equivalence above holds for all $\mathrm{FO}^k(\mathbf{CSP}^+_n)$-formulas φ. If $\alpha = \beta = \emptyset$, we write simply $\mathfrak{A} \equiv^k_{\infty\omega,n} \mathfrak{B}$ instead of $(\mathfrak{A}, \emptyset) \equiv^k_{\infty\omega,n} (\mathfrak{B}, \emptyset)$, and similarly for $\equiv^k_n$.

4 Pebble games

In order to prove undefinability results for the logic $L^k_{\infty\omega}(\mathbf{CSP}^+_n)$, we introduce in this section two pebble games. The first game gives an exact characterization for the equivalence $\equiv^k_{\infty\omega,n}$. The second game corresponds to an at least as strong equivalence, but has simpler rules. Hence we use the second game in the proof of the main result in Section 7.

4.1 Game for CSP-quantifiers

Assume that $\mathfrak{A}$ and $\mathfrak{B}$ are τ-structures for a relational vocabulary τ. Furthermore, assume that α and β are assignments on $\mathfrak{A}$ and $\mathfrak{B}$, respectively, such that $\mathrm{dom}(\alpha) = \mathrm{dom}(\beta) \subseteq X_k$ (recall that $X_k = \{x_1, \ldots, x_k\}$). The *CSP game* for $(\mathfrak{A}, \alpha)$ and $(\mathfrak{B}, \beta)$ is played between *Spoiler* and *Duplicator*, and it has two integer parameters: $n \geq 1$ for the size of templates and $k \geq 1$ for the number of pebbles. The parameters n and k are kept fixed in each play of the game. As the structures $\mathfrak{A}$ and $\mathfrak{B}$ are also fixed in plays, but the assignments α and β are changed, we denote the game by $\mathrm{CSPG}[\mathfrak{A}, \mathfrak{B}, n, k](\alpha, \beta)$, and we use the shorthand notation $\mathrm{CSPG}(\alpha, \beta)$ whenever $\mathfrak{A}$, $\mathfrak{B}$, n and k are clear from the context.

▶ **Definition 2.** *The rules of the game* $\mathrm{CSPG}[\mathfrak{A}, \mathfrak{B}, n, k](\alpha, \beta)$ *are the following:*

(1) *If* $\alpha \mapsto \beta \notin \mathrm{PI}(\mathfrak{A}, \mathfrak{B})$*, then the game ends, and Spoiler wins.*

(2) *If (1) does not hold, there are three types of moves that Spoiler can choose to play:*

- **Bijection move:** *Spoiler starts by choosing a variable* $y \in X_k$*.*
  Assuming that $|A| = |B|$*, Duplicator answers by choosing a bijection* $f\colon A \to B$*. Spoiler completes the round by choosing an element* $a \in A$*. The players continue by playing* $\mathrm{CSPG}(\alpha[a/y], \beta[f(a)/y])$*.*
  On the other hand, if $|A| \neq |B|$*, then the game ends, and Spoiler wins.*
- **Left CSP-quantifier move:** *Spoiler starts by choosing* $r \in [k]$ *and an* r*-tuple* $\vec{y} \in X^r_k$ *of distinct variables and a colouring* $g\colon A \to [n]$*. Duplicator chooses next a colouring* $h\colon B \to [n]$ *such that* $\mathrm{rng}(h) \subseteq \mathrm{rng}(g)$*. Spoiler answers by choosing an* r*-tuple* $\vec{b} \in B^r$*. Duplicator completes the round by choosing an* r*-tuple* $\vec{a} \in A^r$ *such that* $g(a_j) = h(b_j)$ *for all* $j \in [r]$*.*[2] *The players continue by playing* $\mathrm{CSPG}(\alpha[\vec{a}/\vec{y}], \beta[\vec{b}/\vec{y}])$*.*
- **Right CSP-quantifier move:** *Spoiler starts by choosing* $r \in [k]$ *and an* r*-tuple* $\vec{y} \in X^r_k$ *of distinct variables and a colouring* $h\colon B \to [n]$*. Duplicator chooses next a colouring* $g\colon A \to [n]$ *such that* $\mathrm{rng}(g) \subseteq \mathrm{rng}(h)$*. Spoiler answers by choosing an* r*-tuple* $\vec{a} \in A^r$*. Duplicator completes the round by choosing an* r*-tuple* $\vec{b} \in B^r$ *such that* $g(a_j) = h(b_j)$ *for all* $j \in [r]$*. The players continue by playing* $\mathrm{CSPG}(\alpha[\vec{a}/\vec{y}], \beta[\vec{b}/\vec{y}])$*.*

(3) *Duplicator wins the game if Spoiler does not win it in a finite number of rounds.*

We prove now that the CSP game characterizes equivalence with respect to both of the logics $\mathrm{FO}^k(\mathbf{CSP}^+_n)$ and $L^k_{\infty\omega}(\mathbf{CSP}^+_n)$.

▶ **Theorem 3.** *The following conditions are equivalent:*

(1) *Duplicator has a winning strategy in the game* $\mathrm{CSPG}[\mathfrak{A}, \mathfrak{B}, n, k](\alpha, \beta)$*,*

(2) $(\mathfrak{A}, \alpha) \equiv^k_{\infty\omega,n} (\mathfrak{B}, \beta)$*,*

(3) $(\mathfrak{A}, \alpha) \equiv^k_n (\mathfrak{B}, \beta)$*.*

Proof. (1) ⇒ (2): We prove by induction on $\varphi \in L^k_{\infty\omega}(\mathbf{CSP}^+_n)$ that (for any assignments α and β) if Duplicator has a winning strategy in $\mathrm{CSPG}(\alpha, \beta)$, then $(\mathfrak{A}, \alpha) \models \varphi \iff (\mathfrak{B}, \beta) \models \varphi$.

[2] Note that since $\mathrm{rng}(h) \subseteq \mathrm{rng}(g)$, such a tuple $\vec{a}$ always exists.

- If φ is an atomic formula, the claim follows from the fact that Spoiler always wins the game $\mathrm{CSPG}(\alpha,\beta)$ immediately if $\alpha\mapsto\beta\notin\mathrm{PI}(\mathfrak{A},\mathfrak{B})$.
- The cases $\varphi=\neg\psi$, $\varphi=\bigvee\Psi$ and $\varphi=\bigwedge\Psi$ are straightforward.
- Assume next that $\varphi=Qy\,(\psi_1,\ldots,\psi_\ell)$ for some unary generalized quantifer Q. Let Spoiler start the game $\mathrm{CSPG}(\alpha,\beta)$ by a bijection move with variable y. By our assumption Duplicator can answer by a bijection $f\colon A\to B$ such that for any $a\in A$, she has a winning strategy in the continuation $\mathrm{CSPG}(\alpha[a/y],\beta[f(a)/y])$ of the game. By the induction hypothesis we have

  $$(\mathfrak{A},\alpha[a/y])\models\psi_i\iff(\mathfrak{B},\beta[f(a)/y])\models\psi_i.$$

  This means that f is an isomorphism between the structures $(A,\psi_1^{\mathfrak{A},\alpha,y},\ldots,\psi_\ell^{\mathfrak{A},\alpha,y})$ and $(B,\psi_1^{\mathfrak{B},\beta,y},\ldots,\psi_\ell^{\mathfrak{B},\beta,y})$, whence it follows that $(\mathfrak{A},\alpha)\models\varphi\iff(\mathfrak{B},\beta)\models\varphi$.
- Consider finally the case $\varphi=Q_{\mathfrak{C}}\vec{y}\,(\psi_1,\ldots,\psi_\ell)$ for some r-ary CSP-quantifier $Q_{\mathfrak{C}}$ with template $\mathfrak{C}=([n],R_1^{\mathfrak{C}},\ldots,R_\ell^{\mathfrak{C}})$. We start by assuming that $(\mathfrak{A},\alpha)\models\varphi$. Thus, there is a homomorphism g from the structure $(A,\psi_1^{\mathfrak{A},\alpha,\vec{y}},\ldots,\psi_\ell^{\mathfrak{A},\alpha,\vec{y}})$ to $\mathfrak{C}$. Let Spoiler play in the game $\mathrm{CSPG}(\alpha,\beta)$ a left CSP-quantifier move with r, the tuple $\vec{y}\in X_k^r$ and the function g, and let $h\colon B\to[n]$ be the answer of Duplicator given by her winning strategy. We claim that h is a homomorphism $(B,\psi_1^{\mathfrak{B},\beta,\vec{y}},\ldots,\psi_\ell^{\mathfrak{B},\beta,\vec{y}})$ to $\mathfrak{C}$, and consequently $(\mathfrak{B},\beta)\models\varphi$.

  Assume that this is not the case. Then there is $i\in[\ell]$ and a tuple $\vec{b}\in B^r$ such that $\vec{b}\in\psi_i^{\mathfrak{B},\beta,\vec{y}}$ (i.e., $(\mathfrak{B},\beta[\vec{b}/\vec{y}])\models\psi_i$), but $h(\vec{b})\notin R_i^{\mathfrak{C}}$. Let Spoiler play the tuple $\vec{b}$ after Duplicator has played h, and let $\vec{a}\in A^r$ be the answer to this move given by the winning strategy of Duplicator. Then $g(\vec{a})=h(\vec{b})$, and Duplicator has a winning strategy in the continuation $\mathrm{CSPG}(\alpha[\vec{a}/\vec{y}],\beta[\vec{b}/\vec{y}])$ of the game, whence by the induction hypothesis $(\mathfrak{A},\alpha[\vec{a}/\vec{y}])\models\psi_i$, or equivalently, $\vec{a}\in\psi_i^{\mathfrak{A},\alpha,\vec{y}}$. This is in contradiction with the fact that g is a homomorphism, since $g(\vec{a})=h(\vec{b})\notin R_i^{\mathfrak{C}}$. Thus, h is a homomorphism, as we claimed.

  By using the right CSP-quantifier move in place of the left CSP-quantifier move, we can prove that $(\mathfrak{B},\beta)\models\varphi$ implies $(\mathfrak{A},\alpha)\models\varphi$. Thus, $(\mathfrak{A},\alpha)\models\varphi\iff(\mathfrak{B},\beta)\models\varphi$, as desired.

The implication (2) ⇒ (3) is trivially true, as $\mathrm{FO}^k(\mathbf{CSP}_n^+)$ is contained in $L^k_{\infty\omega}(\mathbf{CSP}_n^+)$.

(3) ⇒ (1): Observe first that since $\mathfrak{A}$ is finite, there are only finitely many formulas of $\mathrm{FO}^k(\mathbf{CSP}_n^+)$ that are non-equivalent on $\mathfrak{A}$. Thus, for each assignment γ of $\mathfrak{A}$ with $\mathrm{dom}(\gamma)\subseteq X_k$ there is a formula $\Psi_{\mathfrak{A}}^{n,k}(\gamma)\in\mathrm{FO}^k(\mathbf{CSP}_n^+)$ such that $(\mathfrak{A},\gamma)\models\Psi_{\mathfrak{A}}^{n,k}(\gamma)$, and

$(*)$ $(\mathfrak{C},\delta)\models\Psi_{\mathfrak{A}}^{n,k}(\gamma)$ implies $(\mathfrak{A},\gamma)\equiv_n^k(\mathfrak{C},\delta)$ for any structure $\mathfrak{C}$ and assignment δ.

Assume now that $(\mathfrak{A},\alpha)\equiv_n^k(\mathfrak{B},\beta)$. We show that Duplicator can play in the first round of the game $\mathrm{CSPG}(\alpha,\beta)$ in such a way that $(\mathfrak{A},\alpha')\equiv_n^k(\mathfrak{B},\beta')$ holds in the next position (α',β') of the game. Clearly playing this way in all rounds of the game, Duplicator is guaranteed to win, as the condition $(\mathfrak{A},\alpha)\equiv_n^k(\mathfrak{B},\beta)$ implies that $\alpha\mapsto\beta\in\mathrm{PI}(\mathfrak{A},\mathfrak{B})$. There are three cases based on the type of the first move Spoiler chooses:

- Spoiler makes a bijection move, and picks a variable $y\in X_k$. For the sake of simplicity we use below the shorthand notation $\Psi_a:=\Psi_{\mathfrak{A}}^{n,k}(\alpha[a/y])$ for each $a\in A$. The assumption $(\mathfrak{A},\alpha)\equiv_n^k(\mathfrak{B},\beta)$ implies that $(\mathfrak{B},\beta)\models\exists^{=|A|}y\,(y=y)$, and for each $a\in A$, $(\mathfrak{B},\beta)\models\exists^{=m_a}y\,\Psi_a$, where $m_a=|\{a'\in A\mid\Psi_{a'}\Leftrightarrow\Psi_a\}|$. This means that $|B|=|A|$ and $|\{b\in B\mid(\mathfrak{B},\beta[b/y])\models\Psi_a\}|=m_a$ for all $a\in A$, whence there is a bijection $f\colon\ A\to B$ such that for every $a\in A$, $(\mathfrak{B},\beta[f(a)])\models\Psi_a$. Thus, using this bijection f as her answer to the move of Spoiler, Duplicator makes sure that $(\mathfrak{B},\beta[f(a)/y])\models\Psi_a$, and hence by $(*)$, $(\mathfrak{A},\alpha[a/y])\equiv_n^k(\mathfrak{B},\beta[f(a)/y])$ holds in the next position $(\alpha[a/y],\beta[f(a)/y])$ of the game.

- Spoiler makes a left CSP-quantifier move, and chooses $r \in [k]$, an r-tuple $\vec{y} \in X_k^r$ of variables and a colouring $g\colon A \to [n]$. To simplify notation, we denote $\Psi_{\mathfrak{A}}^{n,k}(\alpha[\vec{a}/\vec{y}])$ by $\Psi_{\vec{a}}$ for each $\vec{a} \in A^r$. Let $\mathfrak{C} = ([n], (R_{\vec{a}}^{\mathfrak{C}})_{\vec{a} \in A^r}, R_{\emptyset}^{\mathfrak{C}})$ be the canonical structure arising from the function g and the formulas $\Psi_{\vec{a}}$: $R_{\vec{a}}^{\mathfrak{C}} = \{g(\vec{a}') \mid \Psi_{\vec{a}'} \Leftrightarrow \Psi_{\vec{a}}\}$ and $R_{\emptyset}^{\mathfrak{C}} = \emptyset$. Then by the assumption $(\mathfrak{A}, \alpha) \equiv_n^k (\mathfrak{B}, \beta)$, $(\mathfrak{B}, \beta) \models Q_{\mathfrak{C}}\vec{y}\,((\Psi_{\vec{a}})_{\vec{a} \in A^r}, \bigwedge_{\vec{a} \in A^r} \neg\Psi_{\vec{a}})$. Thus, there is a homomorphism h from the structure $(B, (\Psi_{\vec{a}}^{\mathfrak{B},\beta,\vec{y}})_{\vec{a} \in A^r}, (\bigwedge_{\vec{a} \in A^r} \neg\Psi_{\vec{a}})^{\mathfrak{B},\beta,\vec{y}})$ to $\mathfrak{C}$. Let Duplicator use this function $h\colon B \to [n]$ as her move, and assume that Spoiler chooses next the tuple $\vec{b} \in B^r$. Since $R_{\emptyset}^{\mathfrak{C}} = \emptyset$, there exists $\vec{a} \in A^r$ such that $\vec{b} \in \Psi_{\vec{a}}^{\mathfrak{B},\beta,\vec{y}}$, i.e., $(\mathfrak{B}, \beta[\vec{b}/\vec{y}]) \models \Psi_{\vec{a}}$. Then $h(\vec{b}) \in R_{\vec{a}}^{\mathfrak{C}}$, whence by the definition of $R_{\vec{a}}^{\mathfrak{C}}$, there exists $\vec{a}' \in A^r$ such that $h(\vec{b}) = g(\vec{a}')$ and $\Psi_{\vec{a}'} \Leftrightarrow \Psi_{\vec{a}}$. We let Duplicator use this tuple $\vec{a}'$ as the final step of her move. The next position in the game is then $(\alpha[\vec{a}'/\vec{y}], \beta[\vec{b}/\vec{y}])$, and $(\mathfrak{B}, \beta[\vec{b}/\vec{y}]) \models \Psi_{\vec{a}'}$, whence $(*)$ implies that $(\mathfrak{A}, \alpha[\vec{a}'/\vec{y}]) \equiv_n^k (\mathfrak{B}, \beta[\vec{b}/\vec{y}])$, as desired.
- The case of right CSP-quantifier move is proved in the same way by switching the roles of the structure $(\mathfrak{A}, \alpha)$ and $(\mathfrak{B}, \beta)$. ◀

4.2 Bijective colouring game

As we mentioned in the beginning of Section 4, instead of the CSP game we will use another game with simpler rules in the proof of the size hierarchy theorem in Section 7. We will now introduce this game.

Assume again that $\mathfrak{A}$ and $\mathfrak{B}$ are τ-structures, and α and β are assignments on them with $\mathrm{dom}(\alpha) = \mathrm{dom}(\beta) \subseteq X_k$. Furthermore, let $g\colon A \to [n]$ and $h\colon B \to [n]$ be colourings. We define next the *bijective colouring game* $\mathrm{BCG}(\alpha, \beta, g, h) := \mathrm{BCG}[\mathfrak{A}, \mathfrak{B}, n, k](\alpha, \beta, g, h)$ for $(\mathfrak{A}, g, \alpha)$ and $(\mathfrak{B}, h, \beta)$ with parameters n and k.

▶ **Definition 4.** *The rules of the game* $\mathrm{BCG}[\mathfrak{A}, \mathfrak{B}, n, k](\alpha, \beta, g, h)$ *are the following:*

(1) *If* $\alpha \mapsto \beta \notin \mathrm{PI}(\mathfrak{A}^g, \mathfrak{B}^h)$*, then the game ends, and Spoiler wins.*

(2) *Otherwise Duplicator chooses a bijection* $f : A \to B$ *(if* $|A| \neq |B|$ *the game ends and Spoiler wins). Spoiler can now choose to play one of the two options:*
   - **Element move:** *Spoiler chooses a variable* $y \in X_k$ *and an element* $a \in A$*. The players continue by playing* $\mathrm{BCG}(\alpha[a/y], \beta[f(a)/y], g, h)$*.*
   - **Colouring move:** *Spoiler chooses a function* $g' : A \to [n]$*. The players continue by playing* $\mathrm{BCG}(\alpha, \beta, g', h')$*, where* h' *is the unique function* $B \to [n]$ *such that* $g' = h' \circ f$*.*

(3) *Duplicator wins the game if Spoiler does not win it in a finite number of rounds.*

The following result shows that the bijective colouring game corresponds to an at least as strong equivalence as the CSP game. We leave it as an open problem, whether the converse of this holds.

▶ **Theorem 5.** *If Duplicator has a winning strategy in the game* $\mathrm{BCG}[\mathfrak{A}, \mathfrak{B}, n, k](\alpha, \beta, g, h)$ *for some colourings* g *and* h*, then she has a winning strategy in the game* $\mathrm{CSPG}[\mathfrak{A}, \mathfrak{B}, n, k](\alpha, \beta)$*.*

Proof. Assume that $g\colon A \to [n]$ and $h\colon B \to [n]$ are colourings such that Duplicator has a winning strategy in $\mathrm{BCG}(\alpha, \beta, g, h)$. Then $\alpha \mapsto \beta \in \mathrm{PI}(\mathfrak{A}^g, \mathfrak{B}^h) \subseteq \mathrm{PI}(\mathfrak{A}, \mathfrak{B})$, and hence Spoiler does not win $\mathrm{CSPG}(\alpha, \beta)$ immediately in position (α, β). We will show that Duplicator can play in the game $\mathrm{CSPG}(\alpha, \beta)$ in such a way that the condition

(†) Duplicator has a winning strategy in $\mathrm{BCG}(\gamma, \delta, \tilde{g}, \tilde{h})$ for some $\tilde{g}\colon A \to [n]$, $\tilde{h}\colon B \to [n]$

holds in every position (γ, δ) during the play. Since (†) implies that $\gamma \mapsto \delta \in \mathrm{PI}(\mathfrak{A}, \mathfrak{B})$, playing this way Duplicator is guaranteed to win the game $\mathrm{CSPG}(\alpha, \beta)$.

Assume now that Spoiler and Duplicator have reached in the game $\mathrm{CSPG}(\alpha, \beta)$ a position (γ, δ) such that (†) holds. Let f be the bijection given by the winning strategy of Duplicator in $\mathrm{BCG}(\gamma, \delta, \tilde{g}, \tilde{h})$. We consider now the options Spoiler has for his move in $\mathrm{CSPG}(\gamma, \delta)$.

- Spoiler plays a bijection move, and picks a variable $y \in X_k$. We let Duplicator use the bijection f as her answer. Let Spoiler choose $a \in A$ to complete the round of the game. By the choice of f, Duplicator has then a winning strategy in the game $\mathrm{BCG}(\gamma[a/y], \delta[f(a)/y], \tilde{g}, \tilde{h})$. Thus (†) holds in the next position $(\gamma[a/y], \delta[f(a)/y])$ of the game $\mathrm{CSPG}(\alpha, \beta)$.
- Spoiler plays a right CSP-quantifier move, and picks a tuple $\vec{y} = (y_1, \ldots, y_r) \in X_k^r$ and a colouring $h' \colon B \to [n]$. Duplicator answers this by choosing $g' = h' \circ f$. Note that if Spoiler chooses g' in the game $\mathrm{BCG}(\gamma, \delta, \tilde{g}, \tilde{h})$, then the next position is (γ, δ, g', h'). Thus Duplicator has a winning strategy in $\mathrm{BCG}(\gamma, \delta, g', h')$. Let Spoiler choose next a tuple $\vec{a} = (a_1 \ldots, a_r) \in A^r$. We define the components b_i of the answer $\vec{b} \in B^r$ of Duplicator by induction on $i \in [r]$:
  - Assume that $b_1, \ldots, b_i$, $i < r$, are already defined, and Duplicator has a winning strategy in $\mathrm{BCG}(\gamma[\vec{a}_i/\vec{y}_i], \delta[\vec{b}_i/\vec{y}_i], g', h')$, where $\vec{a}_i := (a_1, \ldots, a_i)$, $\vec{b}_i := (b_1, \ldots, b_i)$ and $\vec{y}_i := (y_1, \ldots, y_i)$. Let f_{i+1} be the bijection given by Duplicators winning strategy in this game. Let Spoiler now play an element move and choose the component a_{i+1} as his move. Then we let $b_{i+1} = f_{i+1}(a_{i+1})$. By the choice of f_{i+1}, Duplicator has a winning strategy in the continuation $\mathrm{BCG}(\gamma[\vec{a}_i a_{i+1}/\vec{y}_i y_{i+1}], [\vec{b}_i b_{i+1}/\vec{y}_i y_{i+1}], g', h')$ of the game.

  Thus, choosing the tuple $\vec{b}$ defined above as her answer to $\vec{a}$, Duplicator guarantees that condition (†) holds in the next position $(\gamma[\vec{a}/\vec{y}], \delta[\vec{b}/\vec{y}])$.
- The case of left CSP-quantifier move is proved in the same way. ◀

▶ **Corollary 6.** *If Duplicator has a winning strategy in the game* $\mathrm{BCG}[\mathfrak{A}, \mathfrak{B}, n, k](\alpha, \beta, g, h)$ *for some colourings g and h, then* $(\mathfrak{A}, \alpha) \equiv^k_{\infty\omega,n} (\mathfrak{B}, \beta)$.

Thus, the bijective colouring game can be used for proving undefinability results for $L^\omega_{\infty\omega}(\mathbf{CSP}^+_n)$ in the usual way: a generalized quantifier $Q_\mathcal{K}$ is not definable in $L^\omega_{\infty\omega}(\mathbf{CSP}^+_n)$ if for all k there are structures $\mathfrak{A}_k \in \mathcal{K}$ and $\mathfrak{B}_k \notin \mathcal{K}$ such that Duplicator has a winning strategy in $\mathrm{BCG}[\mathfrak{A}_k, \mathfrak{B}_k, n, k](\emptyset, \emptyset, g, h)$ for some g and h.

5 Generalized CFI structures

Fix a natural number n, a $3n$-ary relation symbol R_n, and a connected 3-regular ordered graph $G = (V, E, <_G)$. We describe now the details of the construction of generalized CFI structures $\mathfrak{A}^{\mathrm{ev}}_n(G)$ and $\mathfrak{A}^{\mathrm{od}}_n(G)$. We start by defining structures $\mathfrak{A}^v_n = (A^v_n, R^v_n)$ and $\tilde{\mathfrak{A}}^v_n = (A^v_n, \tilde{R}^v_n)$ for $v \in V$ that will be used as building blocks of $\mathfrak{A}^{\mathrm{ev}}_n(G)$ and $\mathfrak{A}^{\mathrm{od}}_n(G)$. However, before this we introduce some useful notation.

We denote the set of all permutations of a set T by $\mathsf{S}(T)$ and the set of even permutations of T by $\mathsf{A}(T)$. If T is an indexed set of the form $\{t_i \mid i \in [n+1]\}$ and $\pi \in \mathsf{S}(T)$, then we use the notation $\vec{\pi}$ for the n-tuple $(\pi(t_1), \ldots, \pi(t_n))$ (note that $\pi(t_{n+1})$ is not included in $\vec{\pi}$, as π is completely determined by the values $\pi(t_i)$ for $i \in [n]$). The *parity* $\mathsf{p}(\pi) \in \{0, 1\}$ of a permutation $\pi \in \mathsf{S}(T)$ is defined as

$$\mathsf{p}(\pi) := \begin{cases} 0, & \text{if } \pi \in \mathsf{A}(T) \\ 1, & \text{if } \pi \notin \mathsf{A}(T). \end{cases}$$

For $a, b \in \{0, 1\}$, we denote their sum modulo 2 by $a \oplus b$.

For $v \in V$, we denote by $\vec{e}(v)$ the tuple (r, s, t) of edges adjacent to a vertex $v \in V$, where the components are listed in the order $<_G$. Furthermore, we denote by $E(v)$ the set $\{r, s, t\}$. For the definition of $\mathfrak{A}_n^v$ and $\tilde{\mathfrak{A}}_n^v$ we also fix distinct elements a_i^e for all $e \in E$ and $i \in [n+1]$.

▶ **Definition 7.**
- *For each $e \in E$, we define $A_n^e := \{a_i^e \mid i \in [n+1]\}$.*
- *For each $v \in V$, we define $A_n^v := A_n^r \cup A_n^s \cup A_n^t$, where $\vec{e}(v) = (r, s, t)$.*
- *Let $v \in V$ and $\vec{e}(v) = (r, s, t)$. We define the $\{R_n\}$-structures $\mathfrak{A}_n^v := (A_n^v, R_n^v)$ and $\tilde{\mathfrak{A}}_n^v := (A_n^v, \tilde{R}_n^v)$ by setting*
  - $R_n^v := \{\vec{\pi}\vec{\rho}\vec{\sigma} \in P_n^v \mid \mathsf{p}(\pi) \oplus \mathsf{p}(\rho) \oplus \mathsf{p}(\sigma) = 0\}$
  - $\tilde{R}_n^v := \{\vec{\pi}\vec{\rho}\vec{\sigma} \in P_n^v \mid \mathsf{p}(\pi) \oplus \mathsf{p}(\rho) \oplus \mathsf{p}(\sigma) = 1\}$,

  where we use the notation
  - $P_n^v := \{\vec{\pi}\vec{\rho}\vec{\sigma} \mid \pi \in \mathsf{S}(A_n^r), \rho \in \mathsf{S}(A_n^s), \sigma \in \mathsf{S}(A_n^t)\}$

We can now state the important characterization of the automorphisms of these structures and isomomorphisms between them. We say that a bijection $f\colon A_n^v \to A_n^v$ *preserves edges* if $f(a_i^e) \in A_n^e$ for all $e \in E(v)$ and $i \in [n]$. Note that if this holds, then $f = \bigcup_{e \in E(v)} f_e$, where $f_e := f \restriction A_n^e$ for each $e \in E(v)$.

▶ **Lemma 8.** *Let $f\colon A_n^v \to A_n^v$ be a bijection, and let $\vec{e}(v) = (r, s, t)$.*

(a) *f is an automorphism of $\mathfrak{A}_n^v$ and $\tilde{\mathfrak{A}}_n^v$ if and only if it* preserves edges *and* $\mathsf{p}(f) := \mathsf{p}(f_r) \oplus \mathsf{p}(f_s) \oplus \mathsf{p}(f_t) = 0$*;*

(b) *f is an isomorphism between $\mathfrak{A}_n^v$ and $\tilde{\mathfrak{A}}_n^v$ if and only if it preserves edges and $\mathsf{p}(f) = 1$.*

Proof. If f does not preserve edges, then at least for two of the edges $e \in E(v)$ there are permutations $\pi \in \mathsf{S}(A_n^e)$ such that $f(\vec{\pi}) \notin \{\vec{\rho} \mid \rho \in \mathsf{S}(A_n^e)\}$. Clearly this means that $\vec{\pi}$ can be extended to a tuple $\vec{a} \in R_n^v$ such that $f(\vec{a}) \notin R_n^v$ and $f(\vec{a}) \notin \tilde{R}_n^v$. Thus, f is neither an automorphism of $\mathfrak{A}_n^v$, nor an isomorphism from $\mathfrak{A}_n^v$ to $\tilde{\mathfrak{A}}_n^v$. Similarly, there is a tuple $\vec{b} \in \tilde{R}_n^v$ extending $\vec{\pi}$ such that $f(\vec{b}) \notin R_n^v$ and $f(\vec{b}) \notin \tilde{R}_n^v$, whence f is neither an automorphism of $\tilde{\mathfrak{A}}_n^v$, nor an isomorphism from $\tilde{\mathfrak{A}}_n^v$ to $\mathfrak{A}_n^v$.

Assume then that f is edge preserving. Clearly in order to determine whether f is an automorphism of $\mathfrak{A}_n^v$ and $\tilde{\mathfrak{A}}_n^v$ (or an isomorphism between them), it suffices to consider tuples of the form $\vec{a} = \vec{\pi}\vec{\rho}\vec{\sigma} \in P_n^v$. Observe that $f(\vec{a}) = f_r(\vec{\pi}) f_s(\vec{\rho}) f_t(\vec{\sigma}) = \overrightarrow{f_r \circ \pi}\,\overrightarrow{f_s \circ \rho}\,\overrightarrow{f_t \circ \sigma}$ and

$$\mathsf{p}(f_r \circ \pi) \oplus \mathsf{p}(f_s \circ \rho) \oplus \mathsf{p}(f_t \circ \sigma) = \mathsf{p}(\pi) \oplus \mathsf{p}(\rho) \oplus \mathsf{p}(\sigma) \oplus \mathsf{p}(f).$$

Note further that P_n^v is the disjoint union of R_n^v and $\tilde{R}_n^v$. Thus, if $\mathsf{p}(f) = 0$, we have

$$\vec{a} \in R_n^v \iff \vec{a} \notin \tilde{R}_n^v \iff f(\vec{a}) \notin \tilde{R}_n^v \iff f(\vec{a}) \in R_n^v,$$

and if $\mathsf{p}(f) = 1$, we have

$$\vec{a} \in R_n^v \iff \vec{a} \notin \tilde{R}_n^v \iff f(\vec{a}) \in \tilde{R}_n^v \iff f(\vec{a}) \notin R_n^v.$$

This completes the proof of both (a) and (b). ◀

We assign next an $\{R_n\}$-structure $\mathfrak{A}_n(G, U)$ to each subset U of V. The CFI structures $\mathfrak{A}_n^{\mathrm{ev}}(G)$ and $\mathfrak{A}_n^{\mathrm{od}}(G)$ are later defined as special cases of $\mathfrak{A}_n(G, U)$.

▶ **Definition 9.** *Let $U \subseteq V$. We define the $\{R_n\}$-structure $\mathfrak{A}_n(G, U) := (A_n(G), R_n(G, U))$ by setting*
- $A_n(G) := \bigcup_{v \in V} A_n^v = \{a_i^e \mid i \in [n+1], e \in E\}$
- $R_n(G, U) := \bigcup_{v \in V} R_n^v(U)$,

where $R_n^v(U) = \tilde{R}_n^v$ if $v \in U$ and $R_n^v(U) = R_n^v$ if $v \notin U$.

▶ **Lemma 10.** *If $U, U' \subseteq V$ are such that $|U| \equiv |U'| \pmod 2$, then $\mathfrak{A}_n(G,U)$ and $\mathfrak{A}_n(G,U')$ are isomorphic.*

Proof. It suffices to prove the following claim:

(∗) If $U, U' \subseteq V$ and $|(U \setminus U') \cup (U' \setminus U)| = 2$, then $\mathfrak{A}_n(G,U) \cong \mathfrak{A}_n(G,U')$.

The lemma follows from this, since by repeated use of (∗) we see that $\mathfrak{A}_n(G,U) \cong \mathfrak{A}_n(G,\emptyset)$ whenever $|U|$ is even, and $\mathfrak{A}_n(G,U) \cong \mathfrak{A}_n(G,\{v\})$ for any $v \in V$ whenever $|U|$ is odd.

To prove (∗), assume that $U, U' \subseteq V$ and $(U \setminus U') \cup (U' \setminus U) = \{u, u'\}$. Since G is connected, there is an E-path P consisting of edges $e_i = \{u_{i-1}, u_i\}$, $i \in [\ell]$, such that $u_0 = u$ and $u_\ell = u'$. For each $e \in E$, we define a function $f_e \colon A_n^e \to A_n^e$ as follows:

- If $e = e_i$ for some $i \in [\ell]$, then $f_e(a_1^e) = a_2^e$, $f_e(a_2^e) = a_1^e$ and $f_e(a_j^e) = a_j^e$ for $j > 2$.
- Otherwise $f_e = \mathrm{id}_{A_n^e}$ (the identity function of A_n^e).

Let $f : A_n(G) \to A_n(G)$ be the function determined by the functions f_e, $e \in E$: $f(a_i^e) = f_e(a_i^e)$ for all $i \in [n+1]$ and $e \in E$. We show next that f is an isomorphism from $\mathfrak{A}_n(G,U)$ to $\mathfrak{A}_n(G,U')$. Clearly it suffices to show that, for each $v \in V$, the restriction f_v of f to the set A_n^v is an isomorphism $\mathfrak{A}_n^v(U) \to \mathfrak{A}_n^v(U')$, where $\mathfrak{A}_n^v(U) := (A_n^v, R_n^v(U))$.

Note first that $f_v = f_r \cup f_s \cup f_t$, where $(r,s,t) = \vec{e}(v)$, whence f_v is an edge preserving bijection. We consider now the following three cases according to the number $n(P,v) := |\{i \in [\ell] \mid e_i \in E(v)\}|$ (note that the case $n(P,v) = 3$ is not possible):

(1) If $n(P,v) = 0$, then clearly $\mathsf{p}(f_v) = 0$ and $\mathfrak{A}_n^v(U) = \mathfrak{A}_n^v(U') \in \{\mathfrak{A}_n^v, \tilde{\mathfrak{A}}_n^v\}$. Thus, f_v is an isomorphism $\mathfrak{A}_n^v(U) \to \mathfrak{A}_n^v(U')$ by Lemma 8(a).

(2) If $n(P,v) = 1$, then clearly $\mathsf{p}(f_v) = 1$ and $v \in \{u, u'\}$, and hence either $\mathfrak{A}_n^v(U) = \mathfrak{A}_n^v$ and $\mathfrak{A}_n^v(U') = \tilde{\mathfrak{A}}_n^v$, or $\mathfrak{A}_n^v(U) = \tilde{\mathfrak{A}}_n^v$ and $\mathfrak{A}_n^v(U') = \mathfrak{A}_n^v$. Thus, the claim follows from Lemma 8(b) .

(3) If $n(P,v) = 2$, then $\mathsf{p}(f_v) = 0$ and $v \notin \{u, u'\}$, whence $\mathfrak{A}_n^v(U) = \mathfrak{A}_n^v(U') \in \{\mathfrak{A}_n^v, \tilde{\mathfrak{A}}_n^v\}$. The claim follows again from Lemma 8(a). ◀

By Lemma 10, there are at most two isomorphism types of the structures $\mathfrak{A}_n(G,U)$. We use $\mathfrak{A}_n^{\mathrm{ev}}(G) := \mathfrak{A}_n(G,\emptyset)$ and $\mathfrak{A}_n^{\mathrm{od}}(G) := \mathfrak{A}_n(G,\{v_0\})$ as representatives of these classes, where $v_0 \in V$ is the smallest vertex with respect to the order $<_G$. To simplify notation, we denote $R_n(G,\emptyset)$ and $R_n(G,\{v_0\})$ simply by R_n^{ev} and R_n^{od}, respectively.

6 Separating the CFI structures

We prove next that $\mathfrak{A}_n^{\mathrm{ev}}(G)$ and $\mathfrak{A}_n^{\mathrm{od}}(G)$ are indeed non-isomorphic. More precisely, we show that $\mathfrak{A}_n^{\mathrm{ev}}(G)$ and $\mathfrak{A}_n^{\mathrm{od}}(G)$ are separated by a constraint satisfaction problem $\mathsf{CSP}(\mathfrak{C}_n)$. The template structure $\mathfrak{C}_n$ is defined as follows:

▶ **Definition 11.** *We define $\mathfrak{C}_n := (C_n, \mathcal{R}_n)$ to be the $\{R_n\}$-structure such that*

- $C_n := [n+1]$, *and*
- $\mathcal{R}_n = \{\vec{\pi}\vec{\rho}\vec{\sigma} \mid \pi, \rho, \sigma \in \mathsf{S}[n+1],\ \mathsf{p}(\pi) \oplus \mathsf{p}(\rho) \oplus \mathsf{p}(\sigma) = 0\}$.

Here we consider C_n as the indexed set $\{c_i \mid i \in [n+1]\}$, where $c_i = i$ for each $i \in [n+1]$.

▶ **Proposition 12.** $\mathfrak{A}_n^{\mathrm{ev}}(G) \in \mathsf{CSP}(\mathfrak{C}_n)$, *but* $\mathfrak{A}_n^{\mathrm{od}}(G) \notin \mathsf{CSP}(\mathfrak{C}_n)$.

Proof. Given a permutation $\pi \in \mathsf{S}(A_n^e)$, $e \in E$, we denote its natural projection to $[n+1]$ by π_*: i.e., $\pi_*(i) = j$ if and only if $\pi(a_i^e) = a_j^e$. Note that clearly $\mathsf{p}(\pi_*) = \mathsf{p}(\pi)$.

We show that the function $g\colon A_n(G) \to C_n$ such that $g(a_i^e) = i$ for all $e \in E$ and $i \in [n+1]$, is a homomorphism $\mathfrak{A}_n^{\mathrm{ev}}(G) \to \mathfrak{C}_n$. Thus, assume that $\vec{a} \in R_n^{\mathrm{ev}}$. Then $\vec{a} \in R_n^v$ for some $v \in V$, whence there are permutations $\pi \in \mathsf{S}(A_n^r)$, $\rho \in \mathsf{S}(A_n^s)$ and $\sigma \in \mathsf{S}(A_n^t)$, such that $\mathsf{p}(\pi) \oplus \mathsf{p}(\rho) \oplus \mathsf{p}(\sigma) = 0$ and $\vec{a} = \vec{\pi}\vec{\rho}\vec{\sigma}$, where $\vec{e}(v) = (r,s,t)$. Then we have $g(\vec{a}) = g(\vec{\pi})g(\vec{\rho})g(\vec{\sigma}) = \vec{\pi}_*\vec{\rho}_*\vec{\sigma}_* \in \mathcal{R}_n$, since $\mathsf{p}(\pi_*) \oplus \mathsf{p}(\rho_*) \oplus \mathsf{p}(\sigma_*) = \mathsf{p}(\pi) \oplus \mathsf{p}(\rho) \oplus \mathsf{p}(\sigma) = 0$.

To prove that $\mathfrak{A}_n^{\mathrm{od}}(G) \notin \mathsf{CSP}(\mathfrak{C}_n)$, assume towards a contradiction that $h\colon A_n(G) \to C_n$ is a homomorphism $\mathfrak{A}_n^{\mathrm{od}}(G) \to \mathfrak{C}_n$. For each $e \in E$, we let $\chi_e\colon [n+1] \to [n+1]$ be the function such that $\chi_e(i) := h(a_i^e)$.

We observe first that χ_e is a bijection for all $e \in E$. Indeed, if χ_e is not a bijection, then there are $i, j \in [n+1]$, such that $i \neq j$ and $\chi_e(i) = \chi_e(j)$. Let $\pi \in \mathsf{S}(A_n^e)$ be a permutation such that a_i^e and a_j^e occur in $\vec{\pi}$. Then there are edges e' and e'' and permutations $\rho \in \mathsf{S}(A_n^{e'})$ and $\sigma \in \mathsf{S}(A_n^{e''})$ such that $\vec{a} \in R_n^{\mathrm{od}}$, where $\vec{a}$ is either $\vec{\pi}\vec{\rho}\vec{\sigma}$, $\vec{\rho}\vec{\pi}\vec{\sigma}$, or $\vec{\rho}\vec{\sigma}\vec{\pi}$ (depending on the order between e, e' and e''). On the other hand, $h(\vec{a}) \notin \mathcal{R}_n$, since $h(\vec{\pi})$ is not of the form $\vec{\eta}$ for any permutation $\eta \in \mathsf{S}[n+1]$. Hence h is not a homomorphism against our assumption.

Observe next that if $\pi \in \mathsf{S}(A_n^e)$, then $h(\pi(a_i^e)) = \chi_e(\pi_*(i))$ for all $i \in [n+1]$, whence $h(\vec{\pi}) = \overrightarrow{\chi_e \circ \pi_*}$. In particular, $h(\vec{\iota}) = \vec{\chi}_e$ for the identity permutation ι of A_n^e. Consider now a vertex $v \in V \setminus \{v_0\}$ with $\vec{e}(v) = (e_1, e_2, e_3)$ and the tuple $\vec{a} = \vec{\iota}_1\vec{\iota}_2\vec{\iota}_3$, where $\iota_j \in \mathsf{S}(A_n^{e_j})$ is the identity permutation of $A_n^{e_j}$ for $j \in [3]$. Then $\vec{a} \in R_n^{\mathrm{od}}$, whence we must have $h(\vec{a}) = \vec{\chi}_{e_1}\vec{\chi}_{e_2}\vec{\chi}_{e_3} \in \mathcal{R}_n$, or equivalently, $\mathsf{p}(\chi_{e_1}) \oplus \mathsf{p}(\chi_{e_2}) \oplus \mathsf{p}(\chi_{e_3}) = 0$.

On the other hand, if $\vec{e}(v_0) = (d_1, d_2, d_3)$, $\iota_j \in \mathsf{S}(A_n^{d_j})$ is the identity permutation of $A_n^{d_j}$ for $j \in [2]$, and $\pi \in \mathsf{S}(A_n^{d_3}) \setminus \mathsf{A}(A_n^{d_3})$, then $\vec{a} = \vec{\iota}_1\vec{\iota}_2\vec{\pi} \in R_n^{\mathrm{od}}$. Hence we must have $h(\vec{a}) = \vec{\chi}_{d_1}\vec{\chi}_{d_2}\overrightarrow{\chi_{d_3} \circ \pi_*} \in \mathcal{R}_n$, or equivalently, $\mathsf{p}(\chi_{d_1}) \oplus \mathsf{p}(\chi_{d_2}) \oplus \mathsf{p}(\chi_{d_3}) \oplus \mathsf{p}(\pi_*) = 0$. Since $\mathsf{p}(\pi_*) = 1$, it follows that $\mathsf{p}(\chi_{d_1}) \oplus \mathsf{p}(\chi_{d_2}) \oplus \mathsf{p}(\chi_{d_3}) = 1$.

For each $v \in V$, let $O(h, v)$ be the number $|\{e \in E(v) \mid \mathsf{p}(\chi_e) = 1\}|$. By the observations above, we see that $O(h, v)$ is even for all $v \in V \setminus \{v_0\}$ and $O(h, v_0)$ is odd. Thus we see that the sum $O(h) := \sum_{v \in V} O(h, v)$ of these numbers is odd. However this is impossible, since clearly $O(h) = 2 \cdot |\{e \in E \mid \mathsf{p}(\chi_e) = 1\}|$, as each $e \in E$ is adjacent to exactly two vertices. ◀

▶ **Proposition 13.** $\mathsf{CSP}(\mathfrak{C}_n)$ *is* NP*-complete.*

Proof. We give a reduction from 3-colourability to $\mathsf{CSP}(\mathfrak{C}_2)$; it is easy to generalize this reduction to the case of $\mathsf{CSP}(\mathfrak{C}_n)$ with $n > 2$. Given a graph $G = (V, E)$, we define an $\{R_2\}$-structure $\mathfrak{D} = (D, R_2^{\mathfrak{D}})$ as follows:

- $D := V \cup \{(u, v, i) \mid \{u, v\} \in E,\ i \in [2]\}$,
- $R_2^{\mathfrak{D}} := \{(u, v, u, v, (u, v, 1), (u, v, 2)) \mid \{u, v\} \in E\}$.

We show that G is 3-colourable if and only if $\mathfrak{D} \in \mathsf{CSP}(\mathfrak{C}_2)$.

Assume first that $h\colon V \to [3]$ is a 3-colouring of G. Let g be the extension of h to the set D obtained by setting $g(u, v, i) = i$ for all $\{u, v\} \in E$ and $i \in [2]$. If $\vec{a} = (u, v, u, v, (u, v, 1), (u, v, 2)) \in R_2^{\mathfrak{D}}$, then $g(\vec{a}) = (h(u), h(v), h(u), h(v), 1, 2) = \vec{\pi}\vec{\pi}\vec{\iota}$, where $\pi \in \mathsf{S}[3]$ is the permutation such that $\pi(1) = h(u)$ and $\pi(2) = h(v)$, and $\iota \in \mathsf{S}[3]$ is the identity permutation. Thus we see $g(\vec{a}) \in \mathcal{R}_2$, as clearly $\mathsf{p}(\pi) \oplus \mathsf{p}(\pi) \oplus \mathsf{p}(\iota) = 0$. Hence g is a homomorphism $\mathfrak{D} \to \mathfrak{C}_2$.

On the other hand, if $g\colon D \to [3]$ is a homomorphism $\mathfrak{D} \to \mathfrak{C}_2$, then necessarily $g(u) \neq g(v)$ whenever $\{u, v\} \in E$. Thus, the restriction of g to V is a 3-colouring of G. ◀

We do not know whether the structures $\mathfrak{A}_n^{\mathrm{ev}}(G)$ and $\mathfrak{A}_n^{\mathrm{od}}(G)$ can be separated by a PTIME-computable CSP-quantifier, but, in any case, there is a PTIME-property that separates them.

▶ **Proposition 14.** *There exists a* PTIME*-computable class* $\mathcal{K}$ *of* $\{R_n\}$*-structures that is closed under isomorphisms such that* $\mathfrak{A}_n^{\mathrm{ev}}(G) \in \mathcal{K}$ *and* $\mathfrak{A}_n^{\mathrm{od}}(G) \notin \mathcal{K}$.

Proof. Let $\mathcal{K}$ consist of all structures that are isomorphic to $\mathfrak{A}_n^{\mathrm{ev}}(G)$ for some ordered connected 3-regular graph $G = (V, E, <_G)$. Then $\mathfrak{A}_n^{\mathrm{ev}}(G) \in \mathcal{K}$ by definition, and $\mathfrak{A}_n^{\mathrm{od}}(G) \notin \mathcal{K}$ by Lemma 10 and Proposition 12.

We need to show that the membership problem of $\mathcal{K}$ is in PTIME. Let $\mathfrak{B} = (B, R_n^{\mathfrak{B}})$ be an $\{R_n\}$-structure. We describe now an algorithm for deciding whether $\mathfrak{B} \in \mathcal{K}$.

(1) We define first a binary relation P on B by the condition

$$(b, b') \in P \iff \text{there are tuples } \vec{b}_1, \vec{b}_2, \vec{b}_3 \in B^n \text{ such that } \vec{b}_1\vec{b}_2\vec{b}_3 \in R_n^{\mathfrak{B}}, \\ b \text{ occurs in } \vec{b}_i \text{ and } b' \text{ occurs in } \vec{b}_j \text{ for some } 1 \leq i \leq j \leq 3.$$

If $\mathfrak{B}$ is in the class $\mathcal{K}$, then the transitive closure $\preceq$ of P is a linear pre-order of B (i.e., it is transitive and satisfies the dichotomy law: $b \preceq b'$ or $b' \preceq b$ for all $b, b' \in B$). Thus, we compute the transitive closure $\preceq$ of P and reject $\mathfrak{B}$ if it is not a linear pre-order.

(2) Next we define the equivalence relation $\sim$ that corresponds to $\preceq$:

$$b \sim b' \iff b \preceq b' \text{ and } b' \preceq b.$$

We check that every $\sim$-equivalence class $[b] \in B/{\sim}$ has exactly $n + 1$ elements; if this is not the case, then clearly $\mathfrak{B}$ is not in $\mathcal{K}$, and hence $\mathfrak{B}$ is rejected.

(3) We define an ordered graph $G = (V, E, <_G)$ related to $\mathfrak{B}$ as follows:
- $V := \{[b] \cup [c] \cup [d] \mid [b], [c], [d] \in B/{\sim},\ [b]^n \times [c]^n \times [d]^n \cap R_n^{\mathfrak{B}} \neq \emptyset\}$,
- $E := \{\{u, v\} \mid u, v \in V, u \neq v, u \cap v \neq \emptyset\}$,
- $<_G$ is the strict version of the lexicographic linear order on V obtained from the linear pre-order $\preceq$ (note that $\preceq$ is a linear order on $B/{\sim}$).

If $\mathfrak{B}$ is isomorphic to $\mathfrak{A}_n^{\mathrm{ev}}(G')$ for some $G' = (V', E', <_{G'})$, then clearly $G \cong G'$, and the mapping $f(\{u, v\}) = u \cap v$ defines a bijection $E \to B/{\sim}$. Thus, $\mathfrak{B}$ is rejected if G is not an ordered connected 3-regular graph, or if f is not a bijection $E \to B/{\sim}$.

(4) For each $e \in E$, let $\{b_1^e, \ldots, b_{n+1}^e\}$ be an arbitrary enumeration of the equivalence class $f(e)$. For each $v \in V$, let $\mathfrak{B}_v$ be the structure $(v, R_n^{\mathfrak{B}} \cap v^{3n})$, and let $h_v \colon v \to A_n^v$ be the function $h_v(b_i^e) = a_i^e$ for each $e \in E$ such that $f(e) \subseteq v$ and each $i \in [n + 1]$. Define the sets $U^+, U^- \subseteq V$ as follows:
- $U^+ := \{v \in V \mid h_v \text{ is an isomorphism } \mathfrak{B}_v \to \mathfrak{A}_n^v\}$,
- $U^- := \{v \in V \mid h_v \text{ is an isomorphism } \mathfrak{B}_v \to \tilde{\mathfrak{A}}_n^v\}$.

If $\mathfrak{B}$ is in $\mathcal{K}$, then $U^+ \cup U^- = V$. Thus, $\mathfrak{B}$ is rejected, if $U^+ \cup U^- \neq V$. On the other hand, if $U^+ \cup U^- = V$, then the function $h = \bigcup_{v \in V} h_v$ is an isomorphism $\mathfrak{B} \to \mathfrak{A}_n(G, U^-)$. By Lemma 10, $\mathfrak{A}_n(G, U^-) \cong \mathfrak{A}_n^{\mathrm{ev}}$ if and only if $|U^-|$ is even. Thus, $\mathfrak{B}$ is accepted if $|U^-|$ is even, and rejected otherwise.

Clearly each of the steps (1)–(4) of the algorithm can be realized in polynomial time. In particular, since $|v| = 3(n + 1)$ for all $v \in V$ and n is a constant, the sets U^+ and U^- can be computed in polynomial time with respect to $|B|$. ◀

7 Winning the bijective colouring game

Our aim is to prove that Duplicator has a winning strategy in the bijective colouring game $\mathrm{BCG}[\mathfrak{A}_n^{\mathrm{ev}}(G), \mathfrak{A}_n^{\mathrm{od}}(G), n, k]$ provided that G is large enough with respect to the parameter k. Here being large enough is defined in terms of a game, $\mathrm{CR}_k(G)$, that is a minor variation of the *cops&robber game* introduced in [10].

The game $\mathrm{CR}_k(G)$ uses G as a board, and is played between two players, Cop and Robber. Positions of the game are pairs (F, u), where $F \subseteq E$ is such that $|F| \leq k$ and $u \in V$. If $E(u) \subseteq F$, then the game ends immediately, and Cop wins. If this is not the case, then the players play a round as follows:

- Robber chooses a path $P\colon u_0, \ldots, u_\ell$ from $u = u_0$ to some vertex $u' = u_\ell$ such that $\{u_{i-1}, u_i\} \notin F$ for all $i \in [\ell]$.
- Cop chooses an edge $e \in E$ and a set $F' \subseteq F \cup \{e\}$ such that $|F'| \leq k$.
- The next position of the game is (F', u').

Robber wins the game if Cop does not win it in a finite number of rounds.

The intuitive idea of the game is that Cop has k cop pebbles that he moves on the edges of G, and Robber moves one pebble on the vertices of G. Cop tries to capture Robber by surrounding her pebble by his cop pebbles. In each round, Robber can escape along a path that does not contain cop pebbles, and after that Cop is allowed to either add one unused cop pebble on the board, or move one cop pebble to a new position. Robber wins if she can escape forever.

▶ **Definition 15.** *Let $F \subseteq E$ be a set of edges such that $|F| \leq k$. We denote by $\mathrm{W}_k(F, G)$ the set of all vertices $u \in V$ such that Robber has a winning strategy in the game $\mathrm{CR}_k(G)$ starting from position (F, u).*

If Robber has a winning strategy in the game $\mathrm{CR}_k(G)$ from a given position, then she can choose a move in the first round in such a way that she has a winning strategy from the next position after Cop's move. We formulate this simple principle in terms of the winning set $\mathrm{W}_k(F, G)$.

▶ **Lemma 16.** *Let $F \subseteq E$ be a set of edges such that $|F| \leq k$, and let $u \in V$. If $u \in \mathrm{W}_k(F, G)$, then there exists an E-path $u_0 \ldots, u_\ell$ from $u = u_0$ to a vertex $u' = u_\ell$ such that $\{u_{i-1}, u_i\} \notin F$ for all $i \in [\ell]$, and $u' \in \mathrm{W}_k(F', G)$ for all $e \in E$ and $F' \subseteq F \cup \{e\}$ such that $|F'| \leq k$.*

In order to win the bijective colouring game for $\mathfrak{A}_n^{\mathrm{ev}}(G)$ and $\mathfrak{A}_n^{\mathrm{od}}(G)$, Duplicator has to use *edge preserving* bijections $f\colon A_n(G) \to A_n(G)$, i.e., bijections f such that $f(a_i^e) \in A_n^e$ for all $e \in E$ and $i \in [n+1]$. We denote the restriction of an edge preserving f to the set A_n^e by f_e. Note that $f_e \in \mathsf{S}(A_n^e)$ for each $e \in E$, and f is the disjoint union of the bijections f_e. Furthermore, for each $v \in V$ we denote the restriction of f to the set A_n^v by f_v. Thus, if $\vec{e}(v) = (r, s, t)$, then $f_v = f_r \cup f_s \cup f_t$.

We formulate next another important property of bijections that Duplicator will use in her strategy.

▶ **Definition 17.** *An edge preserving bijection $f\colon A_n(G) \to A_n(G)$ is* good *if there is exactly one vertex $v \in V$ such that the following holds:*

(TW) *either $v \neq v_0$ and $\mathsf{p}(f_v) = 1$, or $v = v_0$ and $\mathsf{p}(f_v) = 0$.*

If in addition $g, h\colon A_n(G) \to [n]$ are colourings and $g = h \circ f$, then f is good for g and h.

We denote the set of all good bijections by $\mathrm{GB}(G)$, and the set of all bijections that are good for g and h by $\mathrm{GB}_{gh}(G)$. Furthermore, if $f \in \mathrm{GB}(G)$, then we denote the unique vertex v such that (TW) holds by $\mathrm{tw}(f)$.

The intuition behind condition (TW) is that $\mathrm{tw}(f)$ (the "twist of f") is the only vertex v such that f_v does not preserve the relation R_n.

Note that for any $v \in V$ there are bijections $f \in \mathrm{GB}(G)$ such that $\mathrm{tw}(f) = v$. In the case $v = v_0$, this is witnessed by the identity function $\mathrm{id}_{A_n(G)}$ of $A_n(G)$. For $v \neq v_0$, this follows from the proof of Lemma 10: in the case $U = \{v_0\}$ and $U' = \{v\}$ the isomorphism f from $\mathfrak{A}(G, U)$ to $\mathfrak{A}(G, U')$ constructed in the proof is in $\mathrm{GB}(G)$ and clearly $\mathrm{tw}(f) = v$.

It is also crucial in Duplicator's strategy that the partial functions $p = \alpha \mapsto \beta$ corresponding to positions (α, β, g, h) of the game $\mathrm{BCG}[\mathfrak{A}_n^{\mathrm{ev}}(G), \mathfrak{A}_n^{\mathrm{od}}(G), n, k]$ are restrictions of bijections f that are good for g and h, and whose twist $\mathrm{tw}(f)$ is far enough from $\mathrm{dom}(p)$.

▶ **Definition 18.** *A partial function $p\colon A_n(G) \to A_n(G)$ is* good for g and h *if* $|\mathrm{dom}(p)| \leq k$ *and there exists a bijection $f \in \mathrm{GB}_{gh}(G)$ such that $p \subseteq f$ and $\mathrm{tw}(f) \in \mathrm{W}_k(F^p, G)$, where $F^p := \{e \in E \mid \mathrm{dom}(p) \cap A_n^e \neq \emptyset\}$. We denote the set of all partial functions that are good for g and h by $\mathrm{GP}_{gh}(G)$.*

We prove next that all good partial functions are partial isomorphisms.

▶ **Lemma 19.** *If $p \in \mathrm{GP}_{gh}(G)$, then $p \in \mathrm{PI}(\mathfrak{A}_n^{\mathrm{ev}}(G)^g, \mathfrak{A}_n^{\mathrm{od}}(G)^h)$.*

Proof. Let f be a bijection in GB_{gh} such that $p \subseteq f$ and $\mathrm{tw}(f) \in \mathrm{W}_k(F^p, G)$. Then $g = h \circ f$, whence $g(a) = h(p(a))$ for all $a \in \mathrm{dom}(p)$. Furthermore, since f satisfies the condition (TW) and $\mathrm{tw}(f) \in \mathrm{W}_k(F^p, G)$, we have $\mathsf{p}(f_v) = 0$ for every $v \in V \setminus \{v_0\}$ such that $\mathrm{dom}(p)^{3n} \cap P_n^v \neq \emptyset$, and $\mathsf{p}(f_{v_0}) = 1$ if $\mathrm{dom}(p)^{3n} \cap P_n^{v_0} \neq \emptyset$. Thus, the claim follows from Lemma 8. ◀

The next lemma is the key for defining Duplicator's strategy in the bijective colouring game $\mathrm{BCG}[\mathfrak{A}^{\mathrm{ev}}(G), \mathfrak{A}^{\mathrm{od}}(G)]$: it shows that if $p = \alpha \mapsto \beta$ is a good partial function for g and h, then there is a suitable good bijection f' that Duplicator can play in the position (α, β, g, h).

▶ **Lemma 20.** *Let $g, h\colon A_n(G) \to [n]$ be colourings. Assume that $p \in \mathrm{GP}_{gh}(G)$. Then there exists $f' \in \mathrm{GB}_{gh}(G)$ such that $p \subseteq f'$ and $\mathrm{tw}(f') \in \mathrm{W}_k(F', G)$ for any $e \in E$ and $F' \subseteq F^p \cup \{e\}$ with $|F'| \leq k$.*

Proof. By the definition of GP_{gh} there is a bijection $f \in \mathrm{GB}_{gh}(G)$ such that $p \subseteq f$ and $\mathrm{tw}(f) \in \mathrm{W}_k(F^p, G)$. Hence, by Lemma 16, there exists an E-path $u_0, \ldots, u_\ell$ from $\mathrm{tw}(f) = u_0$ to some vertex $u' = u_\ell$ such that $e_i := \{u_{i-1}, u_i\} \notin F^p$ for all $i \in [\ell]$ and $u' \in \mathrm{W}_k(F', G)$ for any $e \in E$ and $F' \subseteq F^p \cup \{e\}$ with $|F'| \leq k$. We define the bijection f' we are looking for as the union of component bijections $f'_e : A_n^e \to A_n^e$.

For all edges $e \in E$ not on the path P, we let $f'_e := f_e$. Consider then an edge e on the path P. By the pigeon hole principle, there are elements $a, b \in A_n^e$ such that $a \neq b$ and $g(a) = g(b)$. We define now f'_e as follows:

$$f'_e(c) := \begin{cases} f_e(c), & \text{if } c \notin \{a, b\} \\ f_e(b) & \text{if } c = a \\ f_e(a) & \text{if } c = b. \end{cases}$$

Note that since $g = h \circ f$, we have $h(f_e(a)) = h(f_e(b))$, whence $g(c) = h(f'_e(c))$ for all $c \in A_n^e$. Note further that $\mathsf{p}(f'_e) = \mathsf{p}(f_e) \oplus 1$.

Clearly f' is an edge preserving bijection $A_n(G) \to A_n(G)$, and since $g(c) = h(f'_e(c))$ holds for all $e \in E$ and $c \in A_n^e$, we have $g = h \circ f'$. Moreover, since F^p contains no edges of the path P, we have $p = f \restriction \mathrm{dom}(p) = f' \restriction \mathrm{dom}(p) \subseteq f'$.

To show that $f' \in \mathrm{GB}_{gh}(G)$, we observe next that

- $\mathsf{p}(f'_v) = \mathsf{p}(f_v)$ for all $v \notin \{u_0, \ldots, u_\ell\}$,
- $\mathsf{p}(f'_{u_i}) = \mathsf{p}(f'_{e_i}) \oplus \mathsf{p}(f'_{e_{i+1}}) \oplus \mathsf{p}(f_e) = (\mathsf{p}(f_{e_i}) \oplus 1) \oplus (\mathsf{p}(f_{e_{i+1}}) \oplus 1) \oplus \mathsf{p}(f_e) = \mathsf{p}(f_{u_i})$ for each $i \in [\ell - 1]$, where e is the third edge adjacent to u_i, and
- $\mathsf{p}(f'_{u_i}) = \mathsf{p}(f'_{e_j}) \oplus \mathsf{p}(f_e) \oplus \mathsf{p}(f_{e'}) = (\mathsf{p}(f_{e_j}) \oplus 1) \oplus \mathsf{p}(f_e) \oplus \mathsf{p}(f_{e'}) = \mathsf{p}(f_{u_i}) \oplus 1$ for $(i, j) \in \{(0, 1), (\ell, \ell)\}$, where e and e' are the other two edges adjacent to u_i.

Thus we see that there is a unique $v \in V$ such that (TW) holds for f', and this unique vertex is $\mathrm{tw}(f') = u'$. Finally, by the choice of u', we have $\mathrm{tw}(f') \in \mathrm{W}_k(F', G)$ for any $e \in E$ and $F' \subseteq F^p \cup \{e\}$ with $|F'| \leq k$. ◀

Putting together the previous lemmas we can now prove that the strategy based on good bijections and good partial functions guarantees a win for Duplicator in the bijective colouring game.

▶ **Theorem 21.** *Assume that $\mathrm{W}_k(\emptyset, G) \neq \emptyset$. Then there are colourings $g, h\colon A_n(G) \to [n]$ such that Duplicator has a winning strategy in the game $\mathrm{BCG}[\mathfrak{A}_n^{\mathrm{ev}}(G), \mathfrak{A}_n^{\mathrm{od}}(G), n, k](\emptyset, \emptyset, g, h)$.*

Proof. Let g_0 and h_0 be the trivial colourings $A_n(G) \to [n]$ defined by $g_0(a) = h_0(a) = 1$ for all $a \in A_n(G)$. We will show that Duplicator has a winning strategy in the game $\mathrm{BCG}[\mathfrak{A}_n^{\mathrm{ev}}(G), \mathfrak{A}_n^{\mathrm{od}}(G), n, k](\emptyset, \emptyset, g_0, h_0)$. It suffices to show that Duplicator can guarantee that the condition

$$(*) \quad \alpha \mapsto \beta \in \mathrm{GP}_{gh}(G)$$

holds in every position (α, β, g, h) during the play. Indeed, by Lemma 19, this implies that $\alpha \mapsto \beta \in \mathrm{PI}(\mathfrak{A}^{\mathrm{ev}}(G)^g, \mathfrak{A}^{\mathrm{od}}(G)^h)$, and hence Spoiler cannot win the game.

Since $\mathrm{W}_k(\emptyset, G) \neq \emptyset$, and $g_0 = h_0 \circ f$ holds for any bijection $f : A_n(G) \to A_n(G)$, there is a bijection $f_0 \in \mathrm{GB}_{gh}(G)$ such that $\mathrm{tw}(f_0) \in \mathrm{W}_k(\emptyset, G)$. Clearly $\emptyset \mapsto \emptyset = \emptyset \subseteq f_0$, and $F^{\emptyset} = \emptyset$. Thus f_0 witnesses the fact that $\emptyset \mapsto \emptyset \in \mathrm{GP}_{g_0 h_0}(G)$, whence condition $(*)$ holds for the starting position $(\emptyset, \emptyset, g_0, h_0)$ of the game.

Assume then that Duplicator and Spoiler have reached a position (α, β, g, h) such that condition $(*)$ holds. Let $p := \alpha \mapsto \beta$. By Lemma 20 there exists a bijection $f' \in \mathrm{GB}_{gh}(G)$ such that $p \subseteq f'$ and

$$(\dagger) \quad \mathrm{tw}(f') \in \mathrm{W}_k(F', G) \text{ for any } e \in E \text{ and } F' \subseteq F^p \cup \{e\} \text{ with } |F'| \leq k.$$

Let Duplicator play f' as her move. We consider separately the two options Spoiler has for his answer:

(1) Spoiler plays an element move by choosing a variable $y \in X_k$ and an element $a \in A_n(G)$. The next position of the game is $(\alpha', \beta', g', h')$, where $\alpha' = \alpha[a/y]$, $\beta' = \beta[f'(a)/y]$, $g' = g$ and $h' = h$. Thus, $f' \in \mathrm{GB}_{g'h'}(G) = \mathrm{GB}_{gh}(G)$. Let $p' := \alpha' \mapsto \beta'$. Since $p \subseteq f'$ and $p' \subseteq p \cup \{(a, f'(a))\}$, we have $p' \subseteq f'$. Furthermore, $|F^{p'}| \leq k$ and $F^{p'} \subseteq F^p \cup \{e\}$, where e is the edge such that $a \in A_n^e$, whence by condition $(\dagger)$, we have $\mathrm{tw}(f') \in \mathrm{W}_k(F^{p'}, G)$. Thus f' witnesses that condtion $(*)$ holds in the position $(\alpha', \beta', g', h')$.

(2) Spoiler plays a colouring move by choosing a function $g'\colon A_n(G) \to [n]$. The next position is $(\alpha', \beta', g', h')$, where $\alpha' = \alpha$, $\beta' = \beta$ and $h'\colon A_n(G) \to [n]$ is the unique function such that $g' = h' \circ f'$. Since $f' \in \mathrm{GB}(G)$ and $g' = h' \circ f'$, we have $f' \in \mathrm{GB}_{g'h'}(G)$. Let $p' := \alpha' \mapsto \beta'$. Then $p' = p \subseteq f'$ by the choice of f'. Using condition $(\dagger)$ in the special case $F' = F^p = F^{p'}$, we see that $\mathrm{tw}(f') \in \mathrm{W}_k(F^{p'}, G)$. Thus we conclude that $\alpha' \mapsto \beta' = p' \in \mathrm{GP}_{g'h'}(G)$, as desired. ◀

By Proposition 12, $\mathfrak{A}_n^{\mathrm{ev}}(G) \in \mathsf{CSP}(\mathfrak{C}_n)$ and $\mathfrak{A}_n^{\mathrm{od}}(G) \notin \mathsf{CSP}(\mathfrak{C}_n)$. Furthermore, it is straightforward to show that for every k there is a graph G such that $\mathrm{W}_k(\emptyset, G) \neq \emptyset$.[3] Hence, by Theorem 21 and Corollary 6, for every k there are structures $\mathfrak{A}_k \in \mathsf{CSP}(\mathfrak{C}_n)$ and $\mathfrak{B}_k \notin \mathsf{CSP}(\mathfrak{C}_n)$ such that $\mathfrak{A}_k \equiv_{\infty\omega,n}^k \mathfrak{B}_k$. Thus we obtain the following size hierarchy result for CSP-quantifiers.

[3] Such graphs were constructed in Section 8 of [10] for the similar cops&robber game, which is more difficult for Robber to win.

▶ **Corollary 22.** *For any integer $n \geq 2$, there is a CSP-quantifier $Q_{\mathfrak{C}}$ with* $\mathrm{sz}(\mathfrak{C}) = n+1$ *which is not definable in* $L^{\omega}_{\infty\omega}(\mathbf{CSP}^{+}_{n})$.

Although by Proposition 13, $\mathsf{CSP}(\mathfrak{C}_n)$ is NP-complete, we can still use the structures $\mathfrak{A}^{\mathrm{ev}}_n(G)$ and $\mathfrak{A}^{\mathrm{od}}_n(G)$ for proving that PTIME cannot be captured by adding CSP-quantifiers with bounded size templates to inflationary fixed point logic with counting. A suitable PTIME-computable property separating $\mathfrak{A}^{\mathrm{ev}}(G)$ and $\mathfrak{A}^{\mathrm{od}}(G)$ is provided by Proposition 14.

▶ **Corollary 23.** *For any integer $n \geq 2$, there is a* PTIME*-computable quantifier $Q_{\mathcal{K}}$ which is not definable in* $\mathrm{IFPC}(\mathbf{CSP}_n)$.

8 Conclusion

In this paper, we introduce two pebble games for extensions of the infinitary k-variable logic $L^k_{\infty\omega}$ by CSP-quantifiers. We prove that the first of these games, $\mathrm{CSPG}[\mathfrak{A}, \mathfrak{B}, n, k]$ characterizes equivalence of $\mathfrak{A}$ and $\mathfrak{B}$ with respect to $L^k_{\infty\omega}(\mathbf{CSP}^{+}_{n})$, where $\mathbf{CSP}^{+}_{n}$ is the union of the class of all CSP-quantifiers $\mathsf{CSP}(\mathfrak{C})$ with the size of $\mathfrak{C}$ at most n and the class $\mathbf{Q}_1$ of all unary quantifiers. Furthermore, we prove that the second game, BCG, corresponds to at least as strong equivalence as CSPG: if Duplicator has a winning strategy in $\mathrm{BCG}[\mathfrak{A}, \mathfrak{B}, n, k]$, then she has one in $\mathrm{CSPG}[\mathfrak{A}, \mathfrak{B}, n, k]$.

As our main contribution in the paper, we prove a size hierarchy result for CSP-quantifiers: for all $n \geq 2$ there is a CSP-quantifier $Q_{\mathfrak{C}_n}$ of size $n+1$ which is not definable in $L^{\omega}_{\infty\omega}(\mathbf{CSP}^{+}_{n})$. For proving this, we introduce a new variation of the CFI construction, and use the game BCG for showing that the structures $\mathfrak{A}^{\mathrm{ev}}_n(G)$ and $\mathfrak{A}^{\mathrm{od}}_n(G)$ obtained in the construction are equivalent with respect to $L^k_{\infty\omega}(\mathbf{CSP}^{+}_{n})$.

We conclude the paper by listing some open problems:

(1) The CSP-quantifiers $Q_{\mathfrak{C}_n}$ that we use in the proof of the hierarchy result, Corollary 22, are NP-complete. Is it possible to separate the structures $\mathfrak{A}^{\mathrm{ev}}_n(G)$ and $\mathfrak{A}^{\mathrm{od}}_n(G)$ by some PTIME-computable CSP-quantifier?

(2) The arity of the CSP-quantifier $Q_{\mathfrak{C}_n}$ is $3n$. Is it possible to find templates $\mathfrak{D}_n$, $n \geq 2$, such that $Q_{\mathfrak{D}_n}$ is not definable in $L^{\omega}_{\infty\omega}(\mathbf{CSP}^{+}_{n})$ and $\mathrm{ar}(\mathfrak{D}_n) = r$ for some constant r?

(3) We do not allow vectorization (i.e., interpreting ℓ-tuples as elements for some ℓ) in the definition of $L(Q_{\mathcal{K}})$, unlike is done in many other recent papers on the topic (e.g., the papers [5, 6, 9, 14, 4] on rank logic). It is not difficult to define a version on the CSP game that works for the ℓth vectorizations of quantifiers in $\mathbf{CSP}^{+}_{n}$. However, using such games would probably be extremely hard (if not impossible), since they involve existential second-order quantification of ℓ-ary relations. Furthermore, it seems quite plausible that for any $\ell, n \geq 2$, equivalence with respect to the extension of $L^k_{\infty\omega}$ by all ℓth vectorizations of the quantifiers in $\mathbf{CSP}^{+}_{n}$ is just isomorphism for large enough k. The challenge is to prove or disprove this.

(4) What is the relationship between $L^{\omega}_{\infty\omega}(\mathbf{CSP}^{+}_{n})$ and the extension $\mathrm{LA}^{\omega}(Q)$ of $L^{\omega}_{\infty\omega}$ with all linear algebraic operators (see [4])? Does equivalence with respect to one of these logics imply equivalence with respect to the other? We believe that the answer to the latter question is "no". This is because $\mathrm{LA}^{\omega}(Q)$ is closed under vectorizations, but it is not plausible that $L^{\omega}_{\infty\omega}(\mathbf{CSP}^{+}_{n})$ is, and closing it with respect to vectorizations could well lead to a logic whose equivalence is isomorphism (see the previous item).

References

1 Albert Atserias, Andrei A. Bulatov, and Anuj Dawar. Affine systems of equations and counting infinitary logic. *Theor. Comput. Sci.*, 410(18):1666–1683, 2009. doi:10.1016/j.tcs.2008.12.049.

2 Andrei A. Bulatov. A dichotomy theorem for nonuniform csps. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 319–330. IEEE Computer Society, 2017. `doi:10.1109/FOCS.2017.37`.

3 Jin-yi Cai, Martin Fürer, and Neil Immerman. An optimal lower bound on the number of variables for graph identifications. *Comb.*, 12(4):389–410, 1992. `doi:10.1007/BF01305232`.

4 Anuj Dawar, Erich Grädel, and Wied Pakusa. Approximations of isomorphism and logics with linear-algebraic operators. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, volume 132 of *LIPIcs*, pages 112:1–112:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. `doi:10.4230/LIPIcs.ICALP.2019.112`.

5 Anuj Dawar, Martin Grohe, Bjarki Holm, and Bastian Laubner. Logics with rank operators. In *Proceedings of the 24th Annual IEEE Symposium on Logic in Computer Science, LICS 2009, 11-14 August 2009, Los Angeles, CA, USA*, pages 113–122. IEEE Computer Society, 2009. `doi:10.1109/LICS.2009.24`.

6 Anuj Dawar and Bjarki Holm. Pebble games with algebraic rules. In Artur Czumaj, Kurt Mehlhorn, Andrew M. Pitts, and Roger Wattenhofer, editors, *Automata, Languages, and Programming – 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part II*, volume 7392 of *Lecture Notes in Computer Science*, pages 251–262. Springer, 2012. `doi:10.1007/978-3-642-31585-5_25`.

7 Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite model theory.* Perspectives in Mathematical Logic. Springer, 1995.

8 Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28(1):57–104, 1998.

9 Erich Grädel and Wied Pakusa. Rank logic is dead, long live rank logic! In Stephan Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic, CSL 2015, September 7-10, 2015, Berlin, Germany*, volume 41 of *LIPIcs*, pages 390–404. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015. `doi:10.4230/LIPIcs.CSL.2015.390`.

10 Lauri Hella. Logical hierarchies in PTIME. *Inf. Comput.*, 129(1):1–19, 1996. `doi:10.1006/inco.1996.0070`.

11 Neil Immerman and Eric Lander. Describing graphs: A first-order approach to graph canonization. In Alan L. Selman, editor, *Complexity Theory Retrospective: In Honor of Juris Hartmanis on the Occasion of His Sixtieth Birthday, July 5, 1988*, pages 59–81. Springer New York, New York, NY, 1990. `doi:10.1007/978-1-4612-4478-3_5`.

12 Phokion G. Kolaitis and Moshe Y. Vardi. A logical approach to constraint satisfaction. In Nadia Creignou, Phokion G. Kolaitis, and Heribert Vollmer, editors, *Complexity of Constraints – An Overview of Current Research Themes [Result of a Dagstuhl Seminar]*, volume 5250 of *Lecture Notes in Computer Science*, pages 125–155. Springer, 2008. `doi:10.1007/978-3-540-92800-3_6`.

13 Leonid Libkin. *Elements of Finite Model Theory.* Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004. `doi:10.1007/978-3-662-07003-1`.

14 Moritz Lichter. Separating rank logic from polynomial time. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29 – July 2, 2021*, pages 1–13. IEEE, 2021. `doi:10.1109/LICS52264.2021.9470598`.

15 Per Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195, 1966. `doi:10.1111/j.1755-2567.1966.tb00600.x`.

16 Dmitriy Zhuk. A proof of CSP dichotomy conjecture. In Chris Umans, editor, *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 331–342. IEEE Computer Society, 2017. `doi:10.1109/FOCS.2017.38`.