Logics with Counting, Auxiliary Relations, and Lower Bounds for Invariant Queries

Leonid Libkin
Bell Laboratories/INRIA
600 Mountain Avenue
Murray Hill, NJ 07974, USA
Email: libkin@research.bell-labs.com

Abstract

We study the expressive power of counting logics in the presence of auxiliary relations such as orders and pre-orders. The simplest such logic, first-order with counting, captures the complexity class TC^0 *over ordered structures. We also consider first-order logic with arbitrary unary quantifiers, and infinitary extensions.*

The main result of the paper is that all the counting logics above, in the presence of pre-orders that are almost-everywhere linear orders, exhibit a very tame behavior normally associated with first-order properties of unordered *structures. This is in sharp contrast with the expressiveness of these logics in the presence of linear orders: such a tame behavior is not the case even for first-order logic with counting, and the most powerful logic we consider can express every property of ordered structures. The results attest to the difficulty of proving separation results for the ordered case, in particular, to proving the separation of* TC^0 *from NP. To prove the main results, we use locality techniques from finite-model theory, modifying the main notions of locality along the way.*

1 Introduction

The main motivation for studying the expressive power of logics on finite structures comes from applications in Complexity Theory and Databases. Many complexity classes have logical characterizations in terms of expressiveness of various extensions of first-order logic (FO) on finite structures, and most traditional database query languages have well-understood logical counterparts. As the expressiveness of FO is quite limited – most notably, FO cannot express nontrivial counting properties and recursive computation, – various extensions are considered in the literature. In this paper, we study logics that extend first-order with a counting mechanism. Typically, this is done by adding counting quantifiers or terms [8, 11, 14, 20, 29].

Several extensions of FO capture familiar complexity classes over finite structures, and most of the capture results assume that the structures are ordered. The intuition behind the introduction of a linear order is that it allows us to simulate encodings of structures on the tape of a Turing machine. While for order-invariant properties it does not matter in which order elements appear on the tape (indeed, properties like connectivity of graphs to do not depend on how graphs are represented), they do appear in *some* order, and one must be able to use this order in logical formulae. Among the best known characterizations of this kind are characterization of PTIME as FO + LFP (least-fixpoint operator) [19, 35], PSPACE as FO + PFP (partial-fixpoint) [35], TC^0 as FO(**C**) (FO with counting quantifiers) [2], all over ordered structures.

Even though the particular ordering does not change the result of formula, the mere presence of an order gives many logics extra power. For example, while FO+LFP and FO+PFP capture PTIME and PSPACE over ordered structures, they possess the 0-1 law over unordered structures [21], meaning that such a simple PTIME property as parity cannot be expressed. The lower bound of Cai, Fürer and Immerman [4] shows that there are PTIME properties of unordered structures not definable even in FO+LFP extended with counting quantifiers. A similar phenomenon is observed for other logics, e.g., FO and FO(**C**) [3, 30].

Our main goal is to study the impact of auxiliary relations, such as orderings, on the expressive power of counting. The primary motivation comes from complexity theory: while good expressivity bounds exists for counting logics, e.g., $\mathrm{FO}(\mathbf{C})$, over unordered structures [8, 23, 24], no nontrivial bounds are known for the ordered case. As we mentioned, $\mathrm{FO}(\mathbf{C})$, over ordered structures, captures TC^0, the class of problems solvable by polynomial-size, constant-depth threshold circuits, under DLOGTIME-uniformity, see [2]. This is an important complexity class: problems such as integer multiplication and division, and sorting belong to it; TC^0 has also been studied in connection with neural nets, cf. [31]. Despite many efforts, the separation $\mathrm{TC}^0 \subset \mathrm{NP}$ has not been proved, and it appears that there are very serious obstacles to proving it using traditional approaches to circuit lower bounds, see [1, 32]. One might thus hope that the approach based on proving expressivity bounds for logics may circumvent the problems raised by [32].

The results we prove apply to a variety of logics, starting with FO and $\mathrm{FO}(\mathbf{C})$, and ending with a logic $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ proposed in [24]. This logic subsumes $\mathrm{FO}(\mathbf{C})$ and all other known pure counting extensions of FO. (When we speak of counting extensions of FO, we mean extensions that only add a counting mechanism, as opposed to those – extensively studied in the literature, see [29] – that add both counting and fixpoint.)

We will show a *dichotomy* of the following kind: with auxiliary relations that are almost-everywhere linear orders, $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ and other counting logics exhibit a very tame behavior, normally associated with FO definable properties. However, when the order is added, this tameness is lost. For example, $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ expresses *every* property of ordered structures. These results further attest to the difficulty of proving separation of TC^0 from other classes.

As our definition of tame behavior we shall use the *bounded number of degrees property*, or *BNDP*, first introduced in [26]. We define it first for mappings Q from graphs to graphs. Such a mapping Q is said to have the BNDP, if there exists a function $f_Q : \mathbb{N} \to \mathbb{N}$ such that whenever the degrees of all nodes in a graph G are at most k, then in $Q(G)$ one finds at most $f_Q(k)$ different degrees. Note a certain asymmetry in this definition: while the assumption is that the degrees in G are below k, the conclusion is that the *number of different degrees* in $Q(G)$ is below $f_Q(k)$.

It is known that over unordered structures FO definable graph queries have the BNDP. This was proved in [26], using Gaifman's locality theorem. More recently, this property was shown to hold in $\mathrm{FO}(\mathbf{C})$ [23] and $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ [24] (again, over unordered structures) and very recently it was proved for FO in the ordered case [13], assuming that queries are order-invariant.

Informally, our main result can be then stated as follows: *In the presence of relations which are almost-everywhere linear orders, invariant queries definable in $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ and other counting logics have the bounded number of degrees property.*

The BNDP gives us easy proofs of expressivity bounds. For example, it is easy to see that transitive closure *trcl* violates the BNDP: if one starts with a graph of a successor relation on an n-element set (i.e., a chain in which all degrees are bounded by 1), in its transitive closure one finds $n+1$ different degrees, showing that f_{trcl} cannot exist. Thus, there are LOGSPACE problems that cannot be expressed in $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ in the presence of auxiliary relations that coincide with linear orders almost everywhere. Note that in a rather ad-hoc way (the proof only works for *trcl*) the inexpressibility of *trcl* in $\mathrm{FO}(\mathbf{C})$ in the presence of such auxiliary relations was proved very recently [27]; from the results here, this will follow as an easy corollary. The paper [27] then raised a natural question: is it possible that $\mathrm{FO}(\mathbf{C})$ has the same power on ordered structures as it has on structures equipped with almost-linear-order preorder relations? A positive answer would imply that the lower bounds of [27] apply to TC^0. However, we shall show (as a corollary of the main result) that the answer to the above question is negative.

To prove the main result, we exploit the locality techniques in finite-model theory. Originated in the work by Hanf [15] and Gaifman [10], they were recently a subject of renewed attention [5, 9, 13, 26, 23, 24, 28, 34]. The BNDP is typically proved by showing that a logic satisfies an analog of either Hanf's or Gaifman's theorem [23]. However, those fail for $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ in the presence of several classes of preorders. Nevertheless, we prove a statement, weaker than Gaifman's theorem, for counting logics in the presence of auxiliary relations, and show that it implies the BNDP.

Organization In Section 2, we give formal definitions of various counting extensions of FO, notions of locality, and definability with auxiliary relations. We also give an example that shows how the presence of auxiliary relations affects expressiveness.

In Section 3, we state the main result and its corollaries, in particular, the above mentioned dichotomy: there is an enormous gain in expressiveness of counting logics, by going from auxiliary relations which almost-everywhere linear orders, to linear orders. We also give an example of failure of Gaifman's locality theorem for $\mathrm{FO}(\mathbf{C})$ in the presence of almost-everywhere linear orders.

In the remainder of the paper, we prove the main result. In Section 4, we present two notions of locality that are weaker than the notion corresponding to Gaifman's theorem. We explain the connections between those notions and the BNDP, and show that the main theorem reduces to proving *weak semi-locality* of a logic. In Section 5, we prove weak semi-locality of $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ in the presence of almost-everywhere linear orders, combining the bijective games of [16] and a strategy for the duplicator inspired by [33].

Concluding remarks are given in Section 6. All proofs can be found in the full version [25].

2 Notations

Finite Structures and Logics All structures are assumed to be *finite*. A relational signature σ is a set of relation symbols $\{R_1, ..., R_l\}$, with associated arities $p_i > 0$. For directed graphs, the signature consists of one binary predicate. A σ-structure is $\mathcal{A} = \langle A, R_1^{\mathcal{A}}, \ldots, R_l^{\mathcal{A}}\rangle$, where A is a finite set, and $R_i^{\mathcal{A}} \subseteq A^{p_i}$ interprets R_i. The class of finite σ-structures is denoted by $\mathrm{STRUCT}[\sigma]$. When there is no confusion, we write R_i in place of $R_i^{\mathcal{A}}$. Isomorphism is denoted by $\cong$. The carrier of a structure $\mathcal{A}$ is always denoted by A.

We abbreviate first-order logic by FO, and omit the standard definitions. *FO with counting*, denoted by $\mathrm{FO}(\mathbf{C})$, is a two-sorted logic, with second sort being interpreted as an initial segment of natural numbers. That is, a structure $\mathcal{A}$ is of the form

$$\langle \{v_1, \ldots, v_n\}, \{1, \ldots, n\}, <, \mathrm{BIT}, \underline{1}, \underline{n}, R_1^{\mathcal{A}}, \ldots, R_l^{\mathcal{A}}\rangle.$$

Here the relations $R_i^{\mathcal{A}}$ are defined on the domain $\{v_1, \ldots, v_n\}$, while on the numerical domain $\{1, \ldots, n\}$ one has $\underline{1}, \underline{n}, <$ and the BIT predicate available ($\mathrm{BIT}(i, j)$ iff the ith bit in the binary representation of j is one). This logic also has *counting quantifiers* $\exists ix.\varphi(x)$, meaning that that are at least i elements x that satisfy $\varphi(x)$; here i refers to the numerical domain and x to the domain $\{v_1, \ldots, v_n\}$. These quantifiers bind x but not i. Ternary predicates $+$ and $*$ are definable on the numerical domain [8], as is the quantifier $\exists!ix$ meaning the existence of exactly i elements satisfying a formula. For example, $\exists i \exists j\ [(j+j) = i \wedge \exists! ix.\varphi(x)]$ tests if the number of x satisfying φ is even; this property is not definable in FO alone. We separate first-sort variables from second-sort variables by semicolon: $\varphi(\vec{x}; \vec{\jmath})$.

There are several counting extensions of FO that are more powerful than $\mathrm{FO}(\mathbf{C})$; among them $\mathrm{FO}(\mathbf{Q}_u)$, which is FO extended with *all* unary quantifiers. We refer the reader to [16] for the definition of $\mathrm{FO}(\mathbf{Q}_u)$ and its properties. Here, we mostly work with an even more powerful logic, defined below.

We denote the infinitary logic by $\mathcal{L}_{\infty\omega}$; it extends FO by allowing infinite conjunctions $\bigwedge$ and disjunctions $\bigvee$. Then $\mathcal{L}_{\infty\omega}(\mathbf{C})$ is a two-sorted logic, that extends infinitary logic $\mathcal{L}_{\infty\omega}$. Its structures are of the form $(\mathcal{A}, \mathbb{N})$, where $\mathcal{A}$ is a finite relational structure, and $\mathbb{N}$ is a copy of natural numbers. Assume that every constant $n \in \mathbb{N}$ is a second-sort term. To $\mathcal{L}_{\infty\omega}$, add counting quantifiers $\exists ix$ for every $i \in \mathbb{N}$, and *counting terms:* If φ is a formula and $\vec{x}$ is a tuple of free first-sort variables in φ, then $\#\vec{x}.\varphi$ is a term of the second sort, and its free variables are those in φ except $\vec{x}$. Its interpretation is the number of tuples $\vec{a}$ over the finite first-sort universe that satisfy φ. That is, given a structure $\mathcal{A}$, a formula $\varphi(\vec{x}, \vec{y}; \vec{\jmath})$, $\vec{b} \subseteq A$, and $\vec{\jmath}_0 \subset \mathbb{N}$, the value of the term $\#\vec{x}.\varphi(\vec{x}, \vec{b}; \vec{\jmath}_0)$ is the cardinality of the (finite) set $\{\vec{a} \subseteq A \mid \mathcal{A} \models \varphi(\vec{a}, \vec{b}; \vec{\jmath}_0)\}$. For example, the interpretation of $\#x.E(x, y)$ is the in-degree of node y in a graph with the edge-relation E.

As this logic is too powerful (it expresses every property of finite structures), we restrict it by means of the *rank* of a formulae and terms, denoted by rk. It is defined as quantifier rank (that is, it is 0 for atomic formulae, $\mathsf{rk}(\bigvee_i \varphi_i) = \max_i \mathsf{rk}(\varphi_i), \mathsf{rk}(\neg\varphi) = \mathsf{rk}(\varphi), \mathsf{rk}(\exists x\varphi) = \mathsf{rk}(\exists ix\varphi) = \mathsf{rk}(\varphi) + 1$) but it does not take into account quantification over $\mathbb{N}$: $\mathsf{rk}(\exists i\varphi) = \mathsf{rk}(\varphi)$. Furthermore, $\mathsf{rk}(\#\vec{x}.\psi) = \mathsf{rk}(\psi) + |\vec{x}|$.

Definition 1 (see [24]) *The logic $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ is defined to be the restriction of $\mathcal{L}_{\infty\omega}(\mathbf{C})$ to terms and formulae of finite rank.*

It is known [24] that $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ formulae are closed under Boolean connectives and all quantification, and that every predicate on $\mathbb{N} \times \ldots \times \mathbb{N}$ is definable by a $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ formula of rank 0. Thus, we assume that $+, *, -, \leq$, and in fact *every* predicate on natural numbers is available. Known counting expansions of FO are contained in $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$. That is, for every FO, $\mathrm{FO}(\mathbf{C})$, or $\mathrm{FO}(Q_\mathbf{u})$ formula, there exists an equivalent $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ formula of the same rank. A counting logic of [3] can also be embedded into $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$.

Definability with auxiliary relations An *m-ary query* on σ-structures, Q, is a mapping that associates to each $\mathcal{A} \in \mathrm{STRUCT}[\sigma]$ a structure $\langle A, S\rangle$, where $S \subseteq A^m$. We write $\vec{a} \in Q(\mathcal{A})$ if $\vec{a} \in S$, where $\langle A, S\rangle = Q(\mathcal{A})$. A query Q is definable in a logic $\mathcal{L}$ if there exists an $\mathcal{L}$ formula $\varphi(x_1, \ldots, x_m)$ such that $Q(\mathcal{A}) = \varphi[\mathcal{A}] \stackrel{\text{def}}{=} \langle A, \{\vec{a} \mid \mathcal{A} \models \varphi(\vec{a})\}\rangle$.

Let σ' be a relational signature disjoint from σ. If $\mathcal{A}$ is a σ-structure on a universe A, and $\mathcal{A}'$ is a σ'-structure on A, we use the notation $(\mathcal{A}, \mathcal{A}')$ for the $\sigma \cup \sigma'$-structure on A which inherits the interpretation of σ relational symbols from $\mathcal{A}$, and the interpretation of σ' symbols from $\mathcal{A}'$.

Let $\mathcal{C}$ be a class of σ'-structures, with σ and σ' being disjoint. Let $\mathcal{A} \in \mathrm{STRUCT}[\sigma]$. A formula $\varphi(\vec{x})$ in the language of $\sigma \cup \sigma'$ is called *$\mathcal{C}$-invariant on $\mathcal{A}$* if for any two $\mathcal{C}$ structures $\mathcal{A}'$ and $\mathcal{A}''$ on A we have $\varphi[(\mathcal{A}, \mathcal{A}')] = \varphi[(\mathcal{A}, \mathcal{A}'')]$. Associated with such a formula is the following m-ary query (where $m = |\vec{x}|$):

$$Q_\varphi^w(\mathcal{A}) = \begin{cases} \varphi[(\mathcal{A}, \mathcal{A}')], & \varphi \text{ is } \mathcal{C}\text{-invariant on } \mathcal{A} \\ \emptyset, & \text{otherwise.} \end{cases}$$

where $\mathcal{A}'$ is any structure from $\mathcal{C}$ on A. We use the notation $(\mathcal{L} + \mathcal{C})_w$ to denote all queries defined in such a way when φ ranges over formulae of $\mathcal{L}$.

A formula φ is *$\mathcal{C}$-invariant* if it is $\mathcal{C}$-invariant on every structure. With such a φ, we associate a query Q_φ given by $Q_\varphi(\mathcal{A}) = \varphi[(\mathcal{A}, \mathcal{A}')]$ where $\mathcal{A}'$ is a structure from $\mathcal{C}$ on A. The class of all such queries is denoted by $\mathcal{L} + \mathcal{C}$. Clearly,

$$\mathcal{L} + \mathcal{C} \quad \subseteq \quad (\mathcal{L} + \mathcal{C})_w$$

We thus shall aim to establish expressivity bounds for $(\mathcal{L} + \mathcal{C})_w$.

When $\mathcal{C}$ is the class of order relations, we shall write $<$ instead of $\mathcal{C}$. The capture results for complexity classes deal with the classes of queries of the form $\mathcal{L}+<$; for example, uniform TC^0 equals $\mathrm{FO}(\mathbf{C})+<$ [2]. While queries in $\mathcal{L}+<$ are independent of a particular order relation used, the mere presence of such a relation can have an impact on the expressivity of a logic.

We give an example for $\mathrm{FO}(\mathbf{C})$. Assume that σ has one binary and unary relation, i.e. its structures are graphs with a selected subset of nodes. Let Q_0 be the following Boolean query [3]: given such a structure $\langle A, E, X\rangle$, where $A \neq \emptyset$, $E \subseteq A^2$ and $X \subseteq A$, return true iff E is an equivalence relation, and the number of distinct sizes of E-classes equals $|X|$. It is known that Q_0 is not expressible in $\mathrm{FO}(\mathbf{C})$ [3]. However, it *is* expressible in $\mathrm{FO}(\mathbf{C})+<$. Indeed, the equivalence relation $x\theta y$ iff the E-equivalence classes of x and y have the same cardinality is definable in $\mathrm{FO}(\mathbf{C})$. Thus, in $\mathrm{FO}(\mathbf{C})$ one defines the set of smallest (wrt $<$) elements of each such class, and then compares, in $\mathrm{FO}(\mathbf{C})$, the size of this set to X. The two are the same iff the value of Q_0 is true. Note that *any* linear order suffices to express this query.

Thus, $\mathrm{FO}(\mathbf{C}) \subset \mathrm{FO}(\mathbf{C})+<$. Since the latter captures uniform TC^0, this means that there are problems in TC^0 not definable in $\mathrm{FO}(\mathbf{C})$ over unordered structures. It is also known that $\mathrm{FO} \subset \mathrm{FO}+<$. We shall see later that this continues to be true for other counting logics.

Bounded number of degrees property (BNDP) If $\mathcal{A} \in \mathrm{STRUCT}[\sigma]$, and R_i is of arity p_i, then $\mathit{degree}_j(R_i^{\mathcal{A}}, a)$ for $1 \leq j \leq p_i$ is the number of tuples $\vec{a}$ in $R_i^{\mathcal{A}}$ having a in the jth position. In the case of directed graphs, this gives us the usual notions of in- and out-degree. By $\mathit{deg\_set}(\mathcal{A})$ we mean the set of all degrees realized in $\mathcal{A}$, and $\mathit{deg\_count}(\mathcal{A})$ stands for the cardinality of $\mathit{deg\_set}(\mathcal{A})$. We use the notation $\mathrm{STRUCT}_k[\sigma]$ for $\{\mathcal{A} \in \mathrm{STRUCT}[\sigma] \mid \mathit{deg\_set}(\mathcal{A}) \subseteq \{0, 1, \ldots, k\}\}$.

Definition 2 (see [26, 5, 23]) *An m-ary query Q, $m \geq 1$, is said to have the* bounded number of degrees property[1] *, or* BNDP*, if there exists a function $f_Q : \mathbb{N} \to \mathbb{N}$ such that $\mathit{deg\_count}(Q(\mathcal{A})) \leq f_Q(k)$ for every $\mathcal{A} \in \mathrm{STRUCT}_k[\sigma]$.* □

The BNDP is very easy to use for proving expressivity bounds [26]. For example, it is very easy to verify that (deterministic) transitive closure violates the BNDP.

Locality All existing proofs of the BNDP establish first that a logic is *local*. We now define this concept. Given a structure $\mathcal{A}$, its *Gaifman graph* [7, 10, 9] $\mathcal{G}(\mathcal{A})$ is defined as $\langle A, E\rangle$ where (a, b) is in E iff there is a tuple $\vec{c} \in R_i^{\mathcal{A}}$ for some i such that both a and b are in $\vec{c}$. The distance $d(a, b)$ is defined as the length of the shortest path from a to b in $\mathcal{G}(\mathcal{A})$; we assume $d(a, a) = 0$. If $\vec{a} = (a_1, \ldots, a_n)$ and $\vec{b} = (b_1, \ldots, b_m)$, then $d(\vec{a}, \vec{b}) = \min_{ij} d(a_i, b_j)$. Given $\vec{a}$ over A, its

[1] This property was formerly known as the bounded degree property, or the BDP, see [5, 17, 24, 26, 27, etc]. However, many found the name confusing, as the property refers to the *number* of degrees in the output being bounded, rather than the degrees themselves. Following a suggestion by Neil Immerman, we decided to change the name from BDP to BNDP.

r-sphere $S_r^{\mathcal{A}}(\vec{a})$ is $\{b \in A \mid d(\vec{a}, b) \leq r\}$. Its *r-neighborhood* $N_r^{\mathcal{A}}(\vec{a})$ is defined as a structure $N_r^{\mathcal{A}}(\vec{a})$

$$\langle S_r^{\mathcal{A}}(\vec{a}), R_1^{\mathcal{A}} \cap S_r^{\mathcal{A}}(\vec{a})^{p_1}, \ldots, R_k^{\mathcal{A}} \cap S_r^{\mathcal{A}}(\vec{a})^{p_l}, a_1, \ldots, a_n\rangle$$

in the signature that extends σ with n constant symbols. That is, the carrier of $N_r^{\mathcal{A}}(\vec{a})$ is $S_r^{\mathcal{A}}(\vec{a})$, the interpretation of the σ-relations is inherited from $\mathcal{A}$, and the n extra constants are the elements of $\vec{a}$. If $\mathcal{A}$ is understood, we write $S_r(\vec{a})$ and $N_r(\vec{a})$.

If $\mathcal{A}, \mathcal{B} \in \mathrm{STRUCT}[\sigma]$, and there is an isomorphism $N_r^{\mathcal{A}}(\vec{a}) \to N_r^{\mathcal{B}}(\vec{b})$ (that sends $\vec{a}$ to $\vec{b}$), we write $\vec{a} \approx_r^{\mathcal{A},\mathcal{B}} \vec{b}$. If $\mathcal{A} = \mathcal{B}$, we write $\vec{a} \approx_r^{\mathcal{A}} \vec{b}$.

Definition 3 (cf. [23]) *An m-ary query Q is called* local *if there exists a number $r \geq 0$ such that, for any structure $\mathcal{A}$ and any $\vec{a}, \vec{b} \in A^m$*

$$\vec{a} \approx_r^{\mathcal{A}} \vec{b} \quad \textit{implies} \quad \vec{a} \in Q(\mathcal{A}) \;\; \textit{iff} \;\; \vec{b} \in Q(\mathcal{A}).$$

The minimum such r is called the locality rank *of Q, and is denoted by $\mathsf{lr}(Q)$.* □

It follows from Gaifman's theorem [10] that every FO-definable query is local; moreover, if Q is definable by a formula $\varphi(\vec{x})$, then $\mathsf{lr}(Q) \leq (7^{\mathsf{qr}(\varphi)} - 1)/2$. It was shown in [23, 24] that every $\mathrm{FO}(Q_{\mathbf{u}})$, $\mathrm{FO}(\mathbf{C})$, and $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$-definable query is local; furthermore, $\mathsf{lr}(Q) \leq 2^{\mathsf{rk}(\varphi)}$ [24].

Fact 1 (see [5]) *Every local query has the bounded number of degrees property.* □

Thus, without auxiliary relations, queries such as transitive closure cannot be expressed in $\mathrm{FO}(\mathbf{C})$ and even in $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$.

3 Main results

We need to define structures that are "as close as possible" to linear orders. We use the approach of [27]: take a linear order, and replace a small portion of it at the end by a preorder whose equivalence classes have size 2. See Figure 1 for a picture.

Formally, let $g : \mathbb{N} \to \mathbb{R}$ be a nondecreasing function[2]. Define $<_{\upharpoonright g}$ as the class of binary relations (A, R) such that there exists a partition $A = B \cup C$ with $|B| \geq n - g(n)$ and the following properties:

- R restricted to B is a linear order.
- R restricted to C is a preorder where every equivalence class has at most two elements.
- For any $b \in B$ and $c \in C$, $(b, c) \in R$.
- For any $b \in B$ and $c \in C$, $(c, b) \notin R$

Proviso: When we deal with queries in $\mathcal{L} + \mathcal{C}$ and $(\mathcal{L} + \mathcal{C})_w$, which are defined on structures $(\mathcal{A}, \mathcal{A}')$, $\mathcal{A}' \in \mathcal{C}$, all locality concepts (neighborhoods, degrees, etc) refer only to the σ-structure $\mathcal{A}$, and not to the auxiliary structure $\mathcal{A}'$ from $\mathcal{C}$.

Theorem 1 *Let $g : \mathbb{N} \to \mathbb{R}$ be a nondecreasing function that is not bounded by a constant. Then every query in $(\mathcal{L}^*_{\infty\omega}(\mathbf{C}) + <_{\upharpoonright g})_w$ has the bounded number of degrees property.*

That is, with auxiliary structures arbitrarily close to linear orders, the most powerful of counting logics, $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$, still exhibits the very tame behavior typical for FO queries over unordered structures.

Corollaries With g as above, the (deterministic) transitive closure, and, more generally, problems complete for classes DLOGSPACE and above it under first-order reductions, are not definable in any of the counting logics we consider, even in the presence of relations from $<_{\upharpoonright g}$. That is,

Corollary 1 *Let $g : \mathbb{N} \to \mathbb{R}$ be a nondecreasing function that is not bounded by a constant. Then every query in $(\mathrm{FO}(Q_{\mathbf{u}}) + <_{\upharpoonright g})_w$, $(\mathrm{FO}(\mathbf{C}) + <_{\upharpoonright g})_w$, $\mathcal{L}^*_{\infty\omega}(\mathbf{C}) + <_{\upharpoonright g}$, $\mathrm{FO}(Q_{\mathbf{u}}) + <_{\upharpoonright g}$, or $\mathrm{FO}(\mathbf{C}) + <_{\upharpoonright g}$ has the BNDP.*

The following corollaries demonstrate the enormous gain in expressiveness by going from "almost orders" to orders. By a *colored* graph we mean a structure of the signature $(E, U_1, \ldots, U_m)$ where E is binary, and U_is are unary. That is, it is a graph with a few selected subsets of nodes. A colored graph query is a binary query Q on colored graphs; that is, it returns graphs. The *hardness* of such a query is defined as the function $\mathcal{H}_Q : \mathbb{N} \to \mathbb{N}$ where $\mathcal{H}_Q(n)$ is $\max\{\mathit{deg\_count}(Q(\mathcal{A}))\}$ with $\mathcal{A}$ ranging over structures with $|A| = n$ and E being a successor relation.

Recall that $\mathit{deg\_count}(\cdot)$ is the cardinality of the set of all degrees realized in a structure. That is, the hardness shows how complex the output might look like if the

[2]One can deal with functions $g : \mathbb{N} \to \mathbb{N}$ as well; however, as in many examples we use $\log_2$, we prefer to have $\mathbb{R}$ as the range.

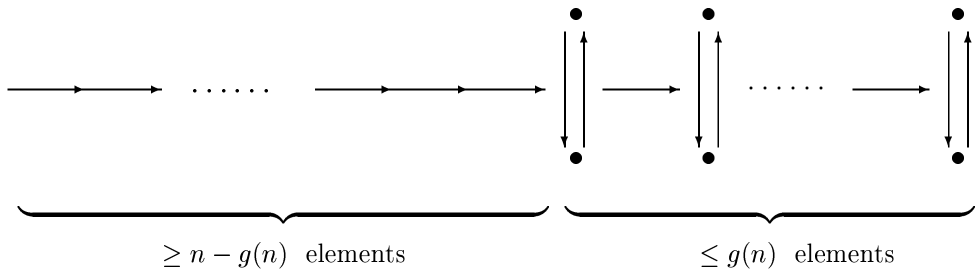

Figure 1. A relation from $<_{\restriction_g}$

input is a successor relation with a few colored subsets. Note that $0 \leq \mathcal{H}_Q(n) \leq n+1$. Since every property of ordered structures is definable in $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ [24], we obtain the following dichotomy result:

Corollary 2
- *Let $g : \mathbb{N} \to \mathbb{R}$ be any nondecreasing function that is not bounded by a constant. Let Q be a colored graph query in $\mathcal{L}^*_{\infty\omega}(\mathbf{C})+<_{\restriction_g}$. Then there exists a constant C such that $\mathcal{H}_Q(n) < C$ for all n.*
- *For any function $f : \mathbb{N} \to \mathbb{N}$ such that $0 \leq f(n) \leq n+1$, there exists a colored graph query Q in $\mathcal{L}^*_{\infty\omega}(\mathbf{C})+<$ such that $\mathcal{H}_Q = f$.*

Thus, dropping a tiny portion of linear order (e.g., $\log\log\ldots\log n$ elements) accounts for the increase in hardness from constant to arbitrary one!

$\mathrm{FO}(\mathbf{C})$ also admits this kind of dichotomy, as there exists a colored graph query Q definable in $\mathrm{FO}(\mathbf{C})+<$ such that $\mathcal{H}_Q(n) \geq \log n$ [17]. In particular, there are problems in uniform TC^0 that cannot be expressed in $\mathrm{FO}(\mathbf{C})+<_{\restriction_g}$. Moreover, it is known that there are uniform AC^0 (that is, $\mathrm{FO}(\mathrm{BIT})+<$) queries that violate the BNDP ([12], see also [6]). Hence, we obtain:

Corollary 3 $\mathrm{AC}^0 \not\subseteq (\mathcal{L}^*_{\infty\omega}(\mathbf{C})+<_{\restriction_g})_w$. $\square$

Corollary 3 also answers an open question from [27]. While [27] showed that *trcl* $\notin \mathrm{FO}(\mathbf{C})+<_{\restriction_g}$, it was left open if $\mathrm{FO}(\mathbf{C})+<_{\restriction_g}= \mathrm{TC}^0$ for some function g as in Theorem 1. Now we have:

Corollary 4 *Let $g : \mathbb{N} \to \mathbb{R}$ be as in Theorem 1, and $\mathcal{L}$ be $\mathrm{FO}(\mathbf{C})$, or $\mathrm{FO}(Q_\mathbf{u})$, or $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$. Then $\mathcal{L}+<_{\restriction_g} \neq \mathcal{L}+<$. Furthermore, $\mathrm{FO}(\mathbf{C}) \subset \mathrm{FO}(\mathbf{C})+<_{\restriction_g}$.*

Note that the presence of some form of counting is essential in these results: it was shown recently [13] that every query in $\mathrm{FO}+<$ has the BNDP.

Outline of the proof of Theorem 1 All proofs of the BNDP that are currently known derive it from locality of queries. Unfortunately, we cannot use this method as queries in $(\mathcal{L}^*_{\infty\omega}(\mathbf{C})+<_{\restriction_g})_w$ need not be local.

Proposition 1 *Let $g(n) < \frac{\log n}{\log\log n}$ be nondecreasing, and not bounded by a constant. Then there exist nonlocal queries in $(\mathcal{L}^*_{\infty\omega}(\mathbf{C})+<_{\restriction_g})_w$.*

Proof sketch: We construct a query Q definable by a formula $\varphi(x)$, and a sequence of structures $\mathcal{A}_n$, $n \in \mathbb{N}$, with an n-element universe, so that for each n large enough, there are two points a, b in $\mathcal{A}_n$ with isomorphic r-neighborhoods, and $(\mathcal{A}_n, P) \models \varphi(a) \wedge \neg\varphi(b)$ for any $P \in <_{\restriction_g}$, where r increases with n.

The signature σ consists of three unary relations U_1, U_2 and C, and one binary relation E. We use P for the auxiliary relation from $<_{\restriction_g}$. Let $l(n) = \lfloor \frac{\log(n-\log n)}{g(n)+1} \rfloor$. In $\mathcal{A}_n$, U_1 has cardinality $M_n = l(n)(g(n)+1) \leq \log(n - \log n)$, and U_2 is its complement. The unary relation C is interpreted as a two-element subset of U_2. Let E' be defined on U_1 as a disjoint union of $g(n)+1$ successor relations of length $l(n)$ each. For each such successor relation E'_i, $i = 1, \ldots, g(n)+1$, let c_i be the node at the distance $\lfloor l(n)/3 \rfloor$ from the start node, and d_i be the node at the distance $\lfloor 2 \cdot l(n)/3 \rfloor$ from the start node. Let $C^{\mathcal{A}_n} = \{a, b\}$. We then define $E^{\mathcal{A}_n} = E' \cup \bigcup_{i=1}^{g(n)+1} \{(a, c_i), (b, d_i)\}$.

We next show that there exists a formula $\beta(x, y)$ in $\mathrm{FO}(\mathbf{C})$ such that $\beta(x, y)$ implies $x, y \in C$ and $(\mathcal{A}_n, P) \models \beta(a, b)$ and $(\mathcal{A}_n, P) \models \neg\beta(b, a)$ for any interpretation of P as a relation from $<_{\restriction_g}$. This will clearly suffice, as a and b have isomorphic neighborhoods of radius $O(l(n))$.

The formula $\beta(x,y)$ is defined as $C(x) \wedge C(y) \wedge \exists u,v.(E(x,u) \wedge E(y,v) \wedge \gamma(u,v))$ where $\gamma(u,v)$ holds iff there is an E-path from u to v all of whose nodes are in P_1, the linear order part of P. That γ can be expressed follows from two observations: first, there are sufficiently many successor relations in E for one of them to be totally contained in P_1, and second, on that successor relation, one can use the order part of P to code monadic second-order using counting, as it was done in [17]. See [25] for details. $\Box$

Proposition 1 provides the first nontrivial example that separates the notion of locality and the BNDP. Now one needs a different technique to prove Theorem 1. We introduce this technique in two steps. In the next section, we consider two ways of weakening the notion of locality, and we show that one of them, weak semi-locality, implies the BNDP. In Section 5, we show how the bijective games [16] can be used to prove weak semi-locality of $(\mathcal{L}^*_{\infty\omega}(\mathbf{C}) + <_{\restriction_g})_w$ queries.

4 Weak locality

To define locality of a query, we considered the equivalence relation $\vec{a} \approx^{\mathcal{A}}_r \vec{b}$ iff $N^{\mathcal{A}}_r(\vec{a}) \cong N^{\mathcal{A}}_r(\vec{b})$. We now consider two refinements that lead to weaker notions of locality. First, we write $\vec{a} \leftrightsquigarrow\!\!\!\!\leftrightsquigarrow^{\mathcal{A}}_r \vec{b}$ if $\vec{a} \approx^{\mathcal{A}}_r \vec{b}$ and $S^{\mathcal{A}}_r(\vec{a}) \cap S^{\mathcal{A}}_r(\vec{b}) = \emptyset$.

For the other refinement, consider a partition $\mathcal{I} = (I_1, I_2)$ of the set $\{1,\ldots,n\}$. Given $\vec{x} = (x_1,\ldots,x_n)$, we denote by $\vec{x}^{\mathcal{I}}_1$ and $\vec{x}^{\mathcal{I}}_2$ the subtuples of $\vec{x}$ that consist of those components whose indices belong to I_1 or I_2, respectively. For example, if $n = 4$ and $\mathcal{I} = (\{1,3\},\{2,4\})$, then $\vec{x}^{\mathcal{I}}_1 = (x_1,x_3)$ and $\vec{x}^{\mathcal{I}}_2 = (x_2,x_4)$. We then write $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_r \vec{b}$, for $\vec{a},\vec{b} \in A^n$, if there exists a partition $\mathcal{I} = (I_1,I_2)$ of $\{1,\ldots,n\}$ such that

- $\vec{a}^{\mathcal{I}}_1 \approx^{\mathcal{A}}_r \vec{b}^{\mathcal{I}}_1$;
- $\vec{a}^{\mathcal{I}}_2 = \vec{b}^{\mathcal{I}}_2$;
- $S^{\mathcal{A}}_r(\vec{a}^{\mathcal{I}}_1)$, $S^{\mathcal{A}}_r(\vec{a}^{\mathcal{I}}_2)$, $S^{\mathcal{A}}_r(\vec{b}^{\mathcal{I}}_1)$ are disjoint.

Clearly, $\vec{a} \leftrightsquigarrow\!\!\!\!\leftrightsquigarrow^{\mathcal{A}}_r \vec{b}$ implies $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_r \vec{b}$ (by taking I_2 to be empty), and $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_{r+1} \vec{b}$ implies $\vec{a} \approx^{\mathcal{A}}_r \vec{b}$.

Definition 4 *An m-ary query Q on σ-structures is called* weakly local *if there exists a number $r \in \mathbb{N}$ such that for any $\mathcal{A} \in \mathrm{STRUCT}[\sigma]$ and any $\vec{a},\vec{b} \in A^m$, $\vec{a} \leftrightsquigarrow\!\!\!\!\leftrightsquigarrow^{\mathcal{A}}_r \vec{b}$ implies $\vec{a} \in Q(\mathcal{A})$ iff $\vec{b} \in Q(\mathcal{A})$.*

A query Q is said to be weakly semi-local *if there exists a number $r \in \mathbb{N}$ such that for any $\mathcal{A} \in \mathrm{STRUCT}[\sigma]$ and any $\vec{a},\vec{b} \in A^m$, $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_r \vec{b}$ implies $\vec{a} \in Q(\mathcal{A})$ iff $\vec{b} \in Q(\mathcal{A})$.*

Proposition 2 *Every local query is weakly semi-local, and every weakly semi-local query is weakly local. There exist queries that are weakly local but not weakly semi-local, and there exist queries that are weakly semi-local but not local.* $\Box$

We study these notions because they are easier to prove than the BNDP, and we will see that the BNDP can be derived from them. The notion of weak locality is particularly simple: the only difference between it and locality is the disjointness of neighborhoods. However, it only gives us a partial result:

Proposition 3 *a) Let Q be a binary weakly local query (i.e., the output is a graph). Then Q has the bounded number of degrees property.*

b) For every $m > 2$, there exists an m-ary weakly local query that does not have the bounded number of degrees property. $\Box$

Combined with the results of Section 5, that would be sufficient to derive Theorem 1 for queries that return graphs. However, for arbitrary queries, we need the more involved notion of weak semi-locality:

Theorem 2 *Every weakly semi-local query has the bounded number of degrees property.*

Proof sketch. For an m-ary query Q on σ-structures, let r witness its weak semi-locality. For each $k > 0$, we show how to find a number $M = M(\sigma,m,r,k)$ such that, whenever $\mathcal{A} \in \mathrm{STRUCT}_k[\sigma]$, $N^{\mathcal{A}}_r(a) \cong N^{\mathcal{A}}_r(b)$ and the isomorphism type of $N^{\mathcal{A}}_r(a)$ is realized at least M times in $\mathcal{A}$, then for each fixed $i \leq m$, $degree_i(a) = degree_i(b)$ in $Q(\mathcal{A})$. From this we can calculate $f_Q(k)$ and derive the BNDP. See [25] for details. $\Box$

To incorporate the information about the function g, we modify the definition as follows: $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_{g,r} \vec{b}$ if $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_r \vec{b}$ and $\left|S^{\mathcal{A}}_r(\vec{a}) \cup S_r(\vec{b})\right| \leq g(|A|)$. Then a query Q is *g-weakly semi-local* if there exists an $r \in \mathbb{N}$ such that $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_{g,r} \vec{b}$ implies $\vec{a} \in Q(\mathcal{A})$ iff $\vec{b} \in Q(\mathcal{A})$. The following is easily derived from Theorem 2.

Corollary 5 *Let $g : \mathbb{N} \to \mathbb{R}$ be nondecreasing and not bounded by a constant. Then every g-weakly semi-local query has the BNDP.* $\Box$

5 Games and weak semi-locality

The goal of this section is to prove the g-weak semi-locality of queries in $(\mathcal{L}^*_{\infty\omega}(\mathbf{C})+<_{\restriction g})_w$. We do this by using *bijective games* of [16].

The game is played by two players, called the spoiler and the duplicator, on two structures $\mathcal{A}, \mathcal{B} \in \text{STRUCT}[\sigma]$. For the n-round game, in each round $i = 1, \ldots, n$, the duplicator selects a bijection $f_i : A \to B$, where B is the carrier of $\mathcal{B}$, and the spoiler selects a point $a_i \in A$ (if $card(A) \neq card(B)$, then the spoiler immediately wins). The duplicator wins after n rounds if the relation $\{(a_i, f_i(a_i)) \mid 1 \leq i \leq n\}$ is a partial isomorphism $\mathcal{A} \to \mathcal{B}$; otherwise the spoiler wins. If the duplicator has a winning strategy in the n-move bijective game on $\mathcal{A}$ and $\mathcal{B}$, we write $\mathcal{A} \equiv^{bij}_n \mathcal{B}$. We write $(\mathcal{A}, \vec{a}) \equiv^{bij}_n (\mathcal{B}, \vec{b}))$ if the duplicator has a winning strategy in the n-move bijective game that starts with the position $(\vec{a}, \vec{b})$. This condition implies that for a FO (or FO$(Q_{\mathbf{u}})$) formula $\varphi(\vec{x})$ of quantifier rank n, $\mathcal{A} \models \varphi(\vec{a})$ iff $\mathcal{B} \models \varphi(\vec{b})$ [16]. We extend this to $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$. Note that the lemma below follows from a slightly more general result of [18], but it also has a simple direct proof, see [25].

Lemma 1 *Let $\varphi(x_1, \ldots, x_m)$ be a $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ formula in the language of σ, with all free variables of the first sort. Let $(\mathcal{A}, \vec{a}) \equiv^{bij}_{\mathsf{rk}(\varphi)} (\mathcal{B}, \vec{b})$, where $\vec{a} \in A^m, \vec{b} \in B^m$. Then $\mathcal{A} \models \varphi(\vec{a})$ iff $\mathcal{B} \models \varphi(\vec{b})$.* $\square$

The following is the key lemma, which is proved by a technique reminiscent of that in [33], extended to deal with bijective games.

Lemma 2 *Let $g : \mathbb{N} \to \mathbb{R}$ be nondecreasing and not bounded by a constant. For any $\mathcal{A}$, $m > 0$, $\vec{a}, \vec{b} \in A^m$, and $n > 0$, if $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_{g,2^n} \vec{b}$, then there exists a preorder P on A such that $P \in <_{\restriction g}$ and*

$$(\mathcal{A}, P, \vec{a}) \quad \equiv^{bij}_n \quad (\mathcal{A}, P, \vec{b})$$

Proof sketch. Let $r = 2^n$ and $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_{g,r} \vec{b}$. Let $\mathcal{I} = (I_1, I_2)$ be a partition witnessing that. We assume without loss of generality that I_1 is nonempty and equals $\{1, \ldots, l\}$, $l \leq m$. Let $\vec{a}' = (a_1, \ldots, a_l)$, $\vec{b}' = (b_1, \ldots, b_l)$, and $\vec{c} = (a_{l+1}, \ldots, a_m) = (b_{l+1}, \ldots, b_m)$. Then $\vec{a}' \mathrel{\substack{\leftrightsquigarrow \\ \leftrightsquigarrow}}^{\mathcal{A}}_r \vec{b}'$, $S^{\mathcal{A}}_r(\vec{a}'\vec{b}') \cap S^{\mathcal{A}}_r(\vec{c}) = \emptyset$, and $\left|S^{\mathcal{A}}_r(\vec{a}'\vec{b}'\vec{c})\right| \leq g(|A|)$.

We now construct P. Let A_0 be $S^{\mathcal{A}}_r(\vec{a}') - \{a_1, \ldots, a_l\}$. Pick any ordering $\prec_1$ on $S^{\mathcal{A}}_r(\vec{a}')$ such that $a_1 \prec_1 a_2 \prec_1 \ldots \prec_1 a_l$ and further, for any $a \in S^{\mathcal{A}}_r(\vec{a}') - \{a_1, \ldots, a_l\}$ we have $a_i \prec_1 a$, for each $i = 1, \ldots, l$, and for any $a', a'' \in S^{\mathcal{A}}_r(\vec{a}') - \{a_1, \ldots, a_l\}$, $d(a', \vec{a}') < d(a'', \vec{a}')$ implies $a' \prec_1 a''$.

Let h be an isomorphism of $N^{\mathcal{A}}_r(\vec{a})$ onto $N^{\mathcal{A}}_r(\vec{b})$. Define, on $S^{\mathcal{A}}_r(\vec{b}')$, an ordering $\prec_2$ by letting $b' \prec_2 b''$ iff $h^{-1}(b') \prec_1 h^{-1}(b'')$. Clearly, the initial fragment of $\prec_2$ is $(b_1, \ldots, b_l)$, and it respects the distance to $\vec{b}'$: $d(b', \vec{b}') < d(b'', \vec{b}')$ implies $b' \prec_2 b''$.

Let P_0 be an arbitrary linear ordering on $A - S^{\mathcal{A}}_r(\vec{a}'\vec{b}')$. Intuitively, P is P_0 followed by a preorder obtained by putting together $\prec_1$ and $\prec_2$, and tying them by h. Formally, $(x, y) \in P$ iff

$$\begin{cases} x, y \notin S^{\mathcal{A}}_r(\vec{a}'\vec{b}') \text{ and } (x, y) \in P_0, \text{ or} \\ x \notin S^{\mathcal{A}}_r(\vec{a}'\vec{b}') \text{ and } y \in S^{\mathcal{A}}_r(\vec{a}'\vec{b}'), \text{ or} \\ x \in S^{\mathcal{A}}_r(\vec{a}'), y \in S^{\mathcal{A}}_r(\vec{a}') \text{ and } x \prec_1 y, \text{ or} \\ x \in S^{\mathcal{A}}_r(\vec{b}'), y \in S^{\mathcal{A}}_r(\vec{b}') \text{ and } x \prec_2 y, \text{ or} \\ x \in S^{\mathcal{A}}_r(\vec{a}'), y \in S^{\mathcal{A}}_r(\vec{b}') \text{ and } h(x) \prec_2 y, \text{ or} \\ x \in S^{\mathcal{A}}_r(\vec{b}'), y \in S^{\mathcal{A}}_r(\vec{a}') \text{ and } x \prec_2 h(y) \end{cases}$$

It easily follows from $\vec{a}' \mathrel{\substack{\leftrightsquigarrow \\ \leftrightsquigarrow}}^{\mathcal{A}}_{g,r} \vec{b}'$ that $P \in <_{\restriction g}$.

Our next claims give a winning strategy for the duplicator in the bijective game on $\mathcal{A}_{\vec{a}} = (\mathcal{A}, P, \vec{a})$ and $\mathcal{A}_{\vec{b}} = (\mathcal{A}, P, \vec{b})$. Note that the universe of both structures is the same, A, and in the game the spoiler selects points in A, and the duplicator select bijections $f : A \to A$.

Define a binary relation H on $S^{\mathcal{A}}_r(\vec{a}'\vec{b}')$ by letting $(x, y) \in H$ iff $x = h(y)$ or $y = h(x)$. We show that the duplicator can play in such a way that, if $\vec{x} = (x_1, \ldots, x_n)$ and $\vec{y} = (y_1, \ldots, y_n)$ are points played on $\mathcal{A}_{\vec{a}}$ and $\mathcal{A}_{\vec{b}}$ respectively after n rounds, then there exists a set $J \subseteq \{1, \ldots, n\}$ with the following properties. (1) If $j \in J$, then $(x_j, y_j) \in H$. (2) If $j \notin J$, then $x_j = y_j$. (3) $\vec{a}'\vec{x}^J \approx^{\mathcal{A}}_0 \vec{b}'\vec{y}^J$, where $\vec{x}^J$ is the subtuple of $\vec{x}$ that consists of the component of $\vec{x}$ whose indices are in J, and likewise for $\vec{y}^J$. (4) $d_{\mathcal{A}}(\vec{a}'\vec{x}^J, \vec{x}^{\overline{J}}) > 1$, and $d_{\mathcal{A}}(\vec{b}'\vec{y}^J, \vec{x}^{\overline{J}}) > 1$, where $d_{\mathcal{A}}$ is the distance in $\mathcal{G}(\mathcal{A})$, and $\vec{x}^{\overline{J}}$ consists of the components of $\vec{x}$ whose indices are not in $\vec{x}^J$.

This suffices to show that the duplicator wins. For this we need to establish $\vec{a}'\vec{c}\vec{x} \approx^{\mathcal{A}}_0 \vec{b}'\vec{c}\vec{y}$, and furthermore, show that the mapping F induced by these two tuples preserves P. The latter is clear though as for any $v = F(u)$, either $u = v$ or $(u, v) \in H$, by construction, and thus P is preserved. To see that $\vec{a}'\vec{c}\vec{x} \approx^{\mathcal{A}}_0 \vec{b}'\vec{c}\vec{y}$, notice that $\vec{a}'\vec{x}^J \approx_0 \vec{b}'\vec{y}^J$ by (3), and by (4) and the definition of $\vec{c}$, $d_{\mathcal{A}}(\vec{a}'\vec{x}^J, \vec{c}\vec{x}^{\overline{J}}) > 1$, and $d_{\mathcal{A}}(\vec{b}'\vec{y}^J, \vec{c}\vec{x}^{\overline{J}}) > 1$. Thus no σ-relation can have a tuple containing an element of $\vec{a}'\vec{x}^J$ and an element of $\vec{c}\vec{x}^{\overline{J}}$, or an element of $\vec{b}'\vec{y}^J$ and an element of $\vec{c}\vec{x}^{\overline{J}}$. This suffices to conclude that $\vec{a}'\vec{c}\vec{x} \approx_0^{\mathcal{A}} \vec{b}'\vec{c}\vec{y}$, and thus the duplicator wins the n-round game, provided (1)-(4) hold.

To prove that the duplicator can play as required, we use a strategy somewhat similar to the one used in [33] for ordinary (not bijective) games. Details can be found in [25]. $\square$

We now put these two lemmas together to show

Theorem 3 *Let g be nondecreasing and not bounded by a constant, and let Q be an m-ary query in $(\mathcal{L}^*_{\infty\omega}(\mathbf{C}) + <_{\restriction_g})_w$. Then Q is g-weakly semi-local.*

Proof: Let Q be definable by $\varphi(x_1, \ldots, x_m)$, where φ is a $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ formula in the language of σ and an extra symbol S for the auxiliary preorder. Let $\mathcal{A}$ be a σ-structure, with $\vec{a}, \vec{b} \in A^m$ and $\vec{a} \leftrightsquigarrow^{\mathcal{A}}_{g,2^n} \vec{b}$, where $n = \mathsf{rk}(\varphi)$. Assume that φ is $<_{\restriction_g}$-invariant on $\mathcal{A}$. Let P_0 be a preorder on A, such that $P_0 \in <_{\restriction_g}$. Let $\vec{a} \in Q(\mathcal{A}) = \varphi[(\mathcal{A}, P_0)]$. Choose P to be the preorder given by Lemma 2. Due to the invariance of φ, $\vec{a} \in \varphi[(\mathcal{A}, P)]$; that is, $(\mathcal{A}, P) \models \varphi(\vec{a})$. By Lemmas 2 and 1, $(\mathcal{A}, P) \models \varphi(\vec{b})$, and again by invariance $(\mathcal{A}, P_0) \models \varphi(\vec{b})$. Thus, $\vec{b} \in \varphi[(\mathcal{A}, P_0)] = Q(\mathcal{A})$. This proves g-weak semi-locality of Q. $\square$

Corollary 6 *Let $\precsim_2$ be the class of preorders in which every equivalence class has size at most 2. Then every query definable in $\mathcal{L}^*_{\infty\omega}(\mathbf{C}) + \precsim_2$ is weakly semi-local, and has the BNDP.* $\square$

Proof of Theorem 1 Let Q be in $(\mathcal{L}^*_{\infty\omega}(\mathbf{C}) + <_{\restriction_g})_w$. By Theorem 3, it is g-weakly semi-local. By Corollary 5, it has the BNDP. $\square$

6 Conclusion

We have shown that queries definable in counting logics $\mathrm{FO}(\mathbf{C})$, $\mathrm{FO}(Q_{\mathbf{u}})$ and $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$, in the presence of relations from the class $<_{\restriction_g}$ have the bounded number of degrees property. In other words, even extremely powerful counting logics in the presence of relations which are almost-everywhere linear orders have a very tame behavior. The situation changes drastically when $<_{\restriction_g}$ is replaced by a linear order: for example, $\mathcal{L}^*_{\infty\omega}(\mathbf{C}) + <$ expresses every query on ordered structures. Some motivation for this study stems from a result in [27] that showed, in a rather ad hoc way, that transitive closure is not definable in $\mathrm{FO}(\mathbf{C}) + <_{\restriction_g}$. As $\mathrm{FO}(\mathbf{C}) + <$ captures uniform TC^0, one may wonder if the techniques could extend from $<_{\restriction_g}$ to $<$. In fact, [27] did not resolve the problem whether $\mathrm{FO}(\mathbf{C}) + <_{\restriction_g} \neq \mathrm{FO}(\mathbf{C}) + <$, thus leaving open the possibility that the two may coincide. We showed here that this is not the case. The results in this paper provide further evidence that it is very hard to separate TC^0 from other classes, e.g., NP. Unlike previous results of this kind [32], we showed inherent limitations of the current techniques in descriptive complexity, based on the structure of the auxiliary relations.

The techniques of this paper cannot be straightforwardly extended to prove separation results in the ordered case. The logic $\mathcal{L}^*_{\infty\omega}(\mathbf{C})$ is very powerful, as it expresses every property of natural numbers, and all other known counting extensions of FO can be embedded into it. We also relied on bijective games to prove the main result. However, bijective games characterize expressiveness of a logic which defines all queries on ordered finite structures. Thus, in the ordered case one cannot use the generic techniques from [16, 23, 24, 28] that apply to a variety of counting logics.

It was shown in [8] that if there is a proof of inexpressibility of some property in $\mathrm{FO}(\mathbf{C}) + <$, then there must be a proof of that based on the counting games of [20]. The counting game is weaker than the bijective game; on the other hand, it does not have the inherent limitations of the latter in the ordered case. Thus, a possible way of proving a separation result may be to modify the locality techniques to work with the counting, rather than bijective, games.

Another approach would be to modify the ordered conjecture of [22] to include counting. Namely, such a modified conjecture would say that there is no unbounded class of ordered structures on which $\mathrm{FO}(\mathbf{C})$ captures polynomial time. One reason to consider this is that there are strong indications that for FO the conjecture holds [22]. With counting, however, one has to be careful: by considering the class of linear orders and adding unary quantifiers which test for polynomial time properties of cardinalities, one obtains a counting logic for which the conjecture fails. However, $\mathrm{FO}(\mathbf{C})$ has rather limited arithmetic, and perhaps an attempt to understand why it fails to capture polynomial time on various classes of structures may lead to a better understanding of its structural properties which are not shared by other counting logics.

Acknowledgements I thank Lauri Hella for a number of very fruitful discussions, and Martin Grohe, Moshe Vardi, Limsoon Wong and Mihalis Yannakakis for comments and suggestions.

The first draft of the proof of the main theorem (for graph queries, using weak locality) was produced in cafe *La Rotonde* on February 21, 1998; the quality of coffee was certainly a contributing factor.

References

[1] E. Allender. Circuit complexity before the dawn of the new millennium. In *Proc. 16th Conf. on Foundations of Software Technology and Theoretical Computer Science*, Springer LNCS vol. 1180, 1996, 1–18.

[2] D.A. Barrington, N. Immerman, H. Straubing. On uniformity within NC^1. *JCSS*, 41:274–306,1990.

[3] M. Benedikt, H.J. Keisler. Expressive power of unary counters. *ICDT'97*, pages 291–305.

[4] J. Cai, M. Fürer and N. Immerman. On optimal lower bound on the number of variables for graph identification. *Combinatorica*, 12 (1992), 389–410.

[5] G. Dong, L. Libkin and L. Wong. Local properties of query languages. *TCS*, to appear. Extended abstract in *ICDT'97*, pages 140–154.

[6] A. Durand, C. Lautemann and M. More. Counting results in weak formalisms. Manuscript, 1998.

[7] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer Verlag, 1995.

[8] K. Etessami. Counting quantifiers, successor relations, and logarithmic space, *JCSS*, 54 (1997), 400–411.

[9] R. Fagin, L. Stockmeyer and M. Vardi, On monadic NP vs monadic co-NP, *Information and Computation*, 120 (1995), 78–92.

[10] H. Gaifman. On local and non-local properties, *Proceedings of the Herbrand Symposium, Logic Colloquium '81*, North Holland, 1982.

[11] E. Grädel and M. Otto. Inductive definability with counting on finite structures. In *CSL'92*, pages 231–247.

[12] M. Grohe and L. Hella. Personal communication, 1998.

[13] M. Grohe and T. Schwentick. Locality of order-invariant first-order formulas. In *MFCS'98*, pages 437–445.

[14] S. Grumbach and C. Tollu. On the expressive power of counting. *TCS* 149 (1995), 67–99.

[15] W. Hanf. Model-theoretic methods in the study of elementary logic. In J.W. Addison et al, eds, *The Theory of Models*, North Holland, 1965, pages 132–145.

[16] L. Hella. Logical hierarchies in PTIME. *Information and Computation*, 129 (1996), 1–19.

[17] L. Hella, L. Libkin and J. Nurmonen. Notions of locality and their logical characterizations over finite models. *J. Symb. Logic*, to appear.

[18] L. Hella, L. Libkin , J. Nurmonen and L. Wong. Logics with aggregate operators. This volume.

[19] N. Immerman. Relational queries computable in polynomial time. *Information and Control*, 68 (1986), 86–104.

[20] N. Immerman and E. Lander. Describing graphs: A first order approach to graph canonization. In *"Complexity Theory Retrospective"*, Springer Verlag, Berlin, 1990.

[21] Ph. Kolaitis, M. Vardi. Infinitary logic and 0-1 laws. *Inf.& Comput.*, 98 (1992), 258–294.

[22] Ph. Kolaitis and M. Vardi. Fixpoint logic vs. infinitary logic in finite-model theory. In *LICS'92*, pages 46–57.

[23] L. Libkin. On the forms of locality over finite models. In *LICS'97*, pages 204–215.

[24] L. Libkin. On counting logics and local properties. In *LICS'98*, to pages 501–512.

[25] L. Libkin. Logics with counting, auxiliary relations, and lower bounds for invariant queries. Bell Labs Technical Memo, 1998.

[26] L. Libkin, L. Wong. Query languages for bags and aggregate functions. *JCSS* 55 (1997), 241–272. Extended abstract in *PODS'94*.

[27] L. Libkin and L. Wong. Unary quantifiers, transitive closure, and relations of large degree. In *STACS'98*, Springer LNCS 1377, pages 183–193.

[28] J. Nurmonen. On winning strategies with unary quantifiers. *J. Logic and Computation*, 6 (1996), 779–798.

[29] M. Otto. *Bounded Variable Logics and Counting: A Study in Finite Models*. Springer Verlag, 1997.

[30] M. Otto. ϵ-Logic is more expressive than first-order on finite structures. Manuscript, 1997.

[31] I. Parberry and G. Schnitger. Parallel computation and threshold functions. *JCSS* 36 (1988), 278–302.

[32] A. Razborov and S. Rudich. Natural proofs. *JCSS* 55 (1997), 24–35.

[33] T. Schwentick. On winning Ehrenfeucht games and monadic NP. *Annals of Pure and Applied Logic*, 79 (1996), 61–92.

[34] T. Schwentick and K. Barthelmann. Local normal forms for first-order logic with applications to games and automata. In *STACS'98*, Springer LNCS 1377, 1998, pages 444–454.

[35] M. Vardi. The complexity of relational query languages. In *STOC'82*, pages 137–146.