THE ROYAL SWEDISH ACADEMY OF SCIENCES

INSTITUT MITTAG-LEFFLER

Auravägen 17, SE-182 60 Djursholm, Sweden
Tel. +46 8 622 05 60 Fax. +46 8 622 05 89
info@mittag-leffler.se www.mittag-leffler.se

Regular Representations of Uniform TC^0

L. Hella, J. Kontinen and K. Luosto

REPORT No. 25, 2009/2010, fall

ISSN 1103-467X
ISRN IML-R- -25-09/10- -SE+fall

Regular Representations of Uniform TC^0*

Lauri Hella Juha Kontinen Kerkko Luosto

Abstract

The circuit complexity class DLOGTIME-uniform AC^0 is known to be a modest subclass of DLOGTIME-uniform TC^0. The weakness of AC^0 is caused by the fact that AC^0 is not closed under restricting AC^0-computable queries into AC^0-computable substrings of the input. Analogously, in descriptive complexity, the logics corresponding to DLOGTIME-uniform AC^0 do not have the relativization property and hence they are not regular. This weakness of DLOGTIME-uniform AC^0 has been elaborated in the line of research on the Crane Beach Conjecture. The conjecture (which was refuted by Barrington, Immerman, Lautemann, Schweikardt and Thérien in [BIL$^+$05]) was that if a language L has a neutral letter, then L can be defined in $\mathrm{FO}_{\mathcal{A}}$, first-order logic with the collection of all arithmetic built-in relations $\mathcal{A}$, iff L can be already defined in $\mathrm{FO}_{\leq}$.

In the first part of this article we consider logics in the range of AC^0 and TC^0. First we formulate a combinatorial criterion for a cardinality quantifier C_S implying that all languages in DLOGTIME-uniform TC^0 can be defined in $\mathrm{FO}_{\leq}(\mathsf{C}_S)$. For instance, this criterion is satisfied by C_S if S is the range of some polynomial with positive integer coefficients of degree at least two. In the second part of the paper we first adapt the key properties of abstract logics to accommodate built-in relations. Then we define the regular interior $\mathcal{R}$-int($\mathcal{L}$) and regular closure $\mathcal{R}$-cl($\mathcal{L}$), of a logic $\mathcal{L}$, and show that the Crane Beach Conjecture can be interpreted as a statement concerning $\mathcal{R}$-int($\mathrm{FO}_{\mathcal{B}}$). By extending the results of [BIL$^+$05], we show that if $\mathcal{B} = \{+\}$, or $\mathcal{B}$ contains only unary relations besides $\leq$, then $\mathcal{R}\text{-int}(\mathrm{FO}_{\mathcal{B}}) \equiv \mathrm{FO}_{\leq}$. In contrast, our results imply that if $\mathcal{B}$ contains $\leq$ and the range of a polynomial of degree at least two, then $\mathcal{R}$-cl($\mathrm{FO}_{\mathcal{B}}$) includes all languages in DLOGTIME-uniform TC^0.

*The first and the second author are affiliated with the EUROCORES LogICCC LINT programme. The second author was supported by grant 127661 of the Academy of Finland. The authors are grateful for the generous support of the Mittag-Leffler Institute, where part of the work reported here was carried out.

1 Introduction

Many circuit complexity classes have been logically characterized by extensions of first-order logic in terms of varying sets of built-in relations and generalized quantifiers. The seminal paper in this area is [BIS90] in which the connection between DLOGTIME-uniformity and $\mathrm{FO}_{\{+,\times\}}$-definability was established, where $\mathrm{FO}_{\{+,\times\}}$ denotes first-order logic with ternary built-in relations $+$ and $\times$. It was also shown in [BIS90] that the languages in DLOGTIME-uniform AC^0 are exactly those definable in $\mathrm{FO}_{\{+,\times\}}$. It is known that the predicates $+$ and $\times$ can be defined in terms of the BIT relation. In fact it was shown in [DDLW98] that BIT alone can define the corresponding canonical ordering, hence $\mathrm{FO}_{\{+,\times\}} \equiv \mathrm{FO}_{\mathrm{BIT}}$ (in the case of a single built-in relation, we drop the set parenthesis in the subscript). A provably larger class of languages, TC^0, is acquired by allowing also majority gates in the circuits. On the logical side, DLOGTIME-uniform TC^0 corresponds to the extension $\mathrm{FO}_{\{+,\times\}}(\mathsf{Maj})$ of first-order logic by the unary majority quantifier Maj with built-in $+$ and $\times$. We refer later to DLOGTIME-uniform AC^0 and DLOGTIME-uniform TC^0 simply as AC^0 and TC^0. It was shown in [BIS90] that the "majority of pairs" Maj^2 can be expressed in terms of Maj with the help of the arithmetic relations. On the other hand, with Maj^2 and order, the arithmetic relations become definable. In [BIS90] it was also asked whether already Maj is enough to define the arithmetic relations. This has been shown not to hold in [Lin95], [Ruh99] and [LMSV01]. On the other hand, in [Luo04], it was observed that the extension of $\mathrm{FO}_{\leq}$ by the general divisibility quantifier D is enough to capture TC^0. Note that, unlike Maj^2, D is a unary quantifier. In [HAB02], TC^0 was shown to include the problems of division and iterated multiplication of binary numbers. These results make essential use of the logical characterization of TC^0 discussed above.

It is interesting to note that, while even non-uniform AC^0 fails to define *Parity* by the famous Theorem of Ajtai [Ajt83] and Furst, Saxe and Sipser [FSS84], it is not known if NP strictly includes TC^0. The weakness of AC^0 is due to the fact that AC^0 is not closed under restricting AC^0-computable queries into AC^0-computable substrings of the input. On the logical side, the logics corresponding to AC^0 do not have the so-called *relativization property* and hence they are not *regular* (for the definition of these concepts, see [Ebb85] and Section 4). In fact, we will show that the only thing that AC^0 lacks, compared to TC^0, is the ability to relativize. The weakness of AC^0, and logics $\mathrm{FO}_{\mathcal{B}}$, where $\mathcal{B}$ is a collection of built-in relations, is also reflected in the fact that $\mathrm{FO}_{\mathcal{B}}$ cannot count cardinalities of sets. This defect can be fixed by extending $\mathrm{FO}_{\mathcal{B}}$ in terms of counting quantifiers $\exists^{=y}x$, resulting with the logic $\mathrm{FOC}_{\mathcal{B}}$ (see, e.g., [Sch05]). On ordered structures, we can equivalently extend $\mathrm{FO}_{\mathcal{B}}$ by Maj or the Härtig quantifier I expressing equicardinality [Luo04]. In the presence of I (respectively, Maj or $\exists^{=y}x$) built-in relations can be replaced by certain (unary) generalized quantifiers an vice verca. Moreover, these quantifiers are universe independent and hence the logic $\mathrm{FO}_{\mathcal{B}}(\mathsf{I})$ is always regular (see Example 4.5). On the other hand, without the quantifier I, these quantifiers can be more expressive than the corresponding built-in relations [Luo04]. It is worth noting that the counting extension $\mathrm{FOC}_{\mathcal{B}} \equiv \mathrm{FO}_{\mathcal{B}}(\mathsf{I})$ of $\mathrm{FO}_{\mathcal{B}}$ is not in general the least regular logic including $\mathrm{FO}_{\mathcal{B}}$, since, e.g., for $S = \{nk \mid k \in \mathbb{N}\}$ we have that

$$\mathrm{FO}_{\{\leq,S\}} \leq \mathrm{FO}_{\leq}(\mathsf{D}_n) < \mathrm{FO}_{\{\leq,S\}}(\mathsf{I}) \equiv \mathrm{FOC}_{\{\leq,S\}},$$

where D_n is the divisibility quantifier corresponding to the set S. Here the logic $\mathrm{FO}_{\leq}(\mathsf{D}_n)$ is the least regular logic including $\mathrm{FO}_{\{\leq,S\}}$ (see Example 5.14), and $\mathrm{FO}_{\leq}(\mathsf{D}_n) < \mathrm{FO}_{\{\leq,S\}}(\mathsf{I})$ follows by Theorem 2.3.

The definability theory of generalized quantifiers has not yet been systematically developed on ordered structures. Some work has been done, e.g., in [Nur00], definability of divisibility quantifiers over ordered structures has been studied. In [Luo04], a systematic study of *cardinality quantifiers* C_S has been conducted with an eye on the possibility to define the quantifier I on ordered structures. Cardinality quantifiers are the simplest kind of unary quantifiers and their definability theory is well understood over unordered structures. In fact, it is known that on unordered structures the quantifier I cannot be defined in terms of any cardinality quantifiers [KV95]. On ordered structures, the situation is very much different. Cardinality quantifiers can be classified into two cases: the quantifier I can be defined in $\mathrm{FO}_{\leq}(\mathsf{C}_S)$ if and only if S is sufficiently non-periodic [Luo04] (see Theorem 2.3). For example, if $S = \{2^n \mid n \in \mathbb{N}\}$ or $S = \mathrm{rg}(P)$, where P is a polynomial with nonnegative integer coefficients of degree at least two, then I can be expressed in $\mathrm{FO}_{\leq}(\mathsf{C}_S)$. In the first part of this paper we build on this classification of cardinality quantifiers. In the main result of Section 3, we formulate a further combinatorial criterion called *pseudolooseness* for $S \subseteq \mathbb{N}$ (see Definition 3.3) which implies that S is non-periodic in the sense of Theorem 2.3, and furthermore implies that $\mathrm{FO}_{\{+,\times\}}(\mathsf{Maj}) \leq \mathrm{FO}_{\leq}(\mathsf{C}_S)$. For instance, if $S = \mathrm{rg}(P)$ is the range of some polynomial P with nonnegative integer coefficients of degree at least two, then this criterion is satisfied, hence

$$\mathrm{FO}_{\leq}(\mathsf{C}_S) \equiv \mathrm{FO}_{\{+,\times\}}(\mathsf{I}) \equiv \mathrm{FO}_{\{+,\times\}}(\mathsf{Maj}), \tag{1}$$

implying that $\mathrm{FO}_{\leq}(\mathsf{C}_S)$ captures TC^0. It is worth noting that if $S = \mathrm{rg}(P)$ and P is of degree one, then all the languages definable in $\mathrm{FO}_{\leq}(\mathsf{C}_S)$ are regular. Also, for any $r > 1$ a real the set $S_r = \{\lfloor x^r \rfloor \mid x \in \mathbb{N}\}$ is pseudoloose, and therefore

$$\mathrm{FO}_{\leq}(\mathsf{C}_{S_r}) \geq \mathrm{FO}_{\{+,\times\}}(\mathsf{Maj}).$$

Interestingly, for $E = \{2^n \mid n \in \mathbb{N}\}$ the non-periodicity of E implies that

$$\mathrm{FO}_{\leq}(\mathsf{C}_E) \equiv \mathrm{FO}_{+}(\mathsf{I}, \mathsf{C}_E)$$

but since E is not pseudoloose, it remains open whether $\mathrm{FO}_{\leq}(\mathsf{C}_E) \equiv \mathrm{FO}_{\{+,\times\}}(\mathsf{Maj})$.

In Section 4 we adapt the familiar properties of *abstract logics* to acommodate also built-in relations. We denote by $\mathcal{L}_{\mathcal{B}}$ the logic $\mathcal{L}$ with built-in relations $\mathcal{B}$. We will usually assume that either $\leq$ is definable in $\mathcal{L}_{\mathcal{B}}$ or $\mathcal{B} = \emptyset$. We say that $\mathcal{L}_{\mathcal{B}}$ is *semiregular*, if it is closed under FO-operations and has the *substitution property*. Also, $\mathcal{L}_{\mathcal{B}}$ is *regular* if it is semiregular and closed under relativization.

Without built-in relations, semiregularity of a logic can be characterized in terms of generalized quantifiers: $\mathcal{L}$ is semiregular if and only if there is a class $\mathcal{Q}$ of quantifiers such that $\mathcal{L} \equiv \mathrm{FO}(\mathcal{Q})$. This characterization remains valid also for logics with built-in relations once the notion of generalized quantifier is adapted to the framework of built-in relations ($\mathcal{B}$-*quantifiers*): any semiregular logic $\mathcal{L}_\mathcal{B}$ with built-in relations $\mathcal{B}$ is equivalent to a logic $\mathrm{FO}_\mathcal{B}(\mathcal{Q})$, where $\mathcal{Q}$ is a class of $\mathcal{B}$-quantifiers. As for regularity, $\mathrm{FO}_\leq$ is a regular logic since the restriction of a linear order to a subset of a model is again a linear order. Note that, while $\mathrm{FO}_\mathcal{B}$ is usually not regular, $\mathrm{FO}_\leq(\mathcal{Q})$ is regular for any class $\mathcal{Q}$ of universe independent $\leq$-quantifiers (see Proposition 4.4). For our purposes, an important consequence of this observation is that $\mathrm{FO}_\leq(\mathsf{C}_S)$ is regular for any cardinality quantifier C_S. Inspired by these observations, we define the notions (adapted from [Luo09]) of *regular interior* $\mathcal{R}\text{-int}(\mathcal{L}_\mathcal{B})$ and *regular closure* $\mathcal{R}\text{-cl}(\mathcal{L}_\mathcal{B})$ of a logic $\mathcal{L}_\mathcal{B}$ with built-in relations. The regular interior of $\mathcal{L}_\mathcal{B}$ is the largest regular logic inside $\mathcal{L}_\mathcal{B}$ and the regular closure of $\mathcal{L}_\mathcal{B}$ is the least regular logic including $\mathcal{L}_\mathcal{B}$.

The Crane Beach Conjecture was a conjecture that arbitrary built-in relations (besides the order) are of no help in defining languages with a neutral letter in first-order logic. A symbol $e \in \Sigma$ is a *neutral letter* for a language $L \subseteq \Sigma^*$ if for all $u, v \in \Sigma^*$ it holds that $uv \in L \iff uev \in L$. In other words, e is a neutral letter for L if inserting or deleting any number of e's in a word does not affect its membership in L. Since the property of having a neutral letter is a language theoretic analogue for the property of being universe independent, it is straightforward to reformulate the Crane Beach Conjecture as a statement concerning the regular interior $\mathcal{R}\text{-int}(\mathrm{FO}_\mathcal{B})$ of $\mathrm{FO}_\mathcal{B}$. We say that a set $\mathcal{B}$ of built-in relations has the Neutral Letter Collapse Property (NLCP) with respect to a class $\mathcal{C}$ of languages if and only if the implication

$$L \text{ is definable in } \mathrm{FO}_\mathcal{B} \implies L \text{ is definable in } \mathrm{FO}_\leq$$

holds for every language $L \in \mathcal{C}$ with a neutral letter. Loosely speaking we will then show that $\mathrm{FO}_\mathcal{B}$ has NLCP with respect to a class $\mathcal{C}$ of languages if and only if $\mathcal{R}\text{-int}(\mathrm{FO}_\mathcal{B})$ collapses to $\mathrm{FO}_\leq$ with respect to definability of languages modulo $\mathcal{C}$. For instance, it was shown in [BIL$^+$05] that $\mathcal{U}$ and $\{+\}$ have NLCP with respect to the class of all languages, where $\mathcal{U}$ contains all unary arithmetical relations together with the order $\leq$. Hence, it directly follows that over word structures the logics $\mathcal{R}\text{-int}(\mathrm{FO}_\mathcal{B})$ and $\mathrm{FO}_\leq$ are equivalent, for $\mathcal{B} \in \{\mathcal{U}, \{+\}\}$. We show that this equivalence actually extends to all vocabularies, that is,

$$\mathcal{R}\text{-int}(\mathrm{FO}_\mathcal{B}) \equiv \mathrm{FO}_\leq, \tag{2}$$

which implies (and explains) the observation that $\mathcal{B}$ has NLCP with respect to the class of all languages.

From the computational perspective, the regular closure $\mathcal{R}\text{-cl}(\mathcal{L})$ of a logic $\mathcal{L}$ is very interesting since most complexity classes are closed under relativization. Our results imply that if $\mathcal{B}$ contains the range of a polynomial of degree at least two, then $\mathcal{R}\text{-cl}(\mathrm{FO}_\mathcal{B}) \geq \mathrm{FO}_{\{+,\times\}}(\mathsf{Maj})$ in contrast to (2). In particular, the regular closure of AC^0 is TC^0.

2 Background and preliminaries

In this section we recall concepts and previous results which are addressed in Section 3. In particular, a more detailed exposition of generalized quantifiers (with built-in relations) is provided in Section 4.

Generalized quantifiers

In this subsection we define the generalized quantifiers dicussed in the introduction. The notion of a generalized quantifier goes back to [Mos57] and [Lin66]. For a more complete account on quantifiers, see [KV95].

Examples of unary quantifiers of vocabulary $\{U\}$ are the *divisibility quantifier* D_n expressing that the size of a unary relation is divisible by n ($n \in \mathbb{N}$) and, more generally, for a fixed $S \subseteq \mathbb{N}$ the *cardinality quantifier* C_S. For every $\mathfrak{M} \in \mathrm{Str}(\{U\})$,

$$\mathfrak{M} \models \mathsf{C}_S\, x\, (U(x)) \text{ iff } |U^{\mathfrak{M}}| \in S.$$

Note that $\mathsf{D}_n = \mathsf{C}_S$ with $S = n\mathbb{N} = \{nk \mid k \in \mathbb{N}\}$. The quantifier Maj has also vocabulary $\{U\}$:

$$\mathfrak{M} \models \mathsf{Maj}\ x\, (U(x)) \text{ iff } |U^{\mathfrak{M}}| > \mathrm{card}(\mathfrak{M})/2.$$

For the purposes of this paper, *universe independence* (see Section 4) is an important property of the quantifiers C_S. Universe independence, in the case of the quantifiers C_S, means simply that the truth of the formula $\mathsf{C}_S\, x\, (U(x))$ depends only on $|U^{\mathfrak{M}}|$, disregarding $|\,\mathrm{Dom}(\mathfrak{M}) \setminus U^{\mathfrak{M}}|$. Note that the quantifier Maj is an example of a quantifier which is not universe independent.

The *equicardinality* or *the Härtig quantifier* I, and the *general divisibility quantifier* D are quantifiers of vocabulary $\{U, V\}$. For $\mathfrak{M} \in \mathrm{Str}(\{U, V\})$, we have

$$\mathfrak{M} \models \mathsf{I}\ xy\, (U(x), V(y)) \text{ iff } |U^{\mathfrak{M}}| = |V^{\mathfrak{M}}|,$$

and

$$\mathfrak{M} \models \mathsf{D}\ xy\, (U(x), V(y)) \text{ iff } |U^{\mathfrak{M}}| \,\Big|\, |V^{\mathfrak{M}}|.$$

Härtig showed in [Här65] that addition, as a ternary predicate $+$, can be defined in terms of the quantifier I and first-order logic on ordered structures. Indeed, it is easy to verify that $+$ is defined by the formula $\phi(x, y, z)$, where

$$\phi(x, y, z) := \mathsf{I}\, uv(u < x, y < v \leq z).$$

It was observed in [Luo04] that

$$\mathrm{FO}_{\leq}(\mathsf{Maj}) \equiv \mathrm{FO}_{\leq}(\mathsf{I}) \equiv \mathrm{FOC}_{\leq}, \tag{3}$$

where FOC is the extension of FO in terms of counting quantifiers $\exists^{=y}x$ [IL90]. The counting quantifiers work syntactically so that, in the formula $\exists^{=y}x\psi(x, \boldsymbol{z})$, variable x is bound and y is a new free variable. Let $\mathfrak{M}$ be an ordered structure with $\mathrm{Dom}(\mathfrak{M}) = \{0, \ldots, n-1\}$ and $\leq^{\mathfrak{M}}$ the natural order. Then the semantics of $\exists^{=y}x$ is given by

$$\mathfrak{M} \models \exists^{=y} x\psi(x, \boldsymbol{z})[c/y, \boldsymbol{b}/\boldsymbol{z}] \text{ iff } |\{a \in \mathrm{Dom}(\mathfrak{M}) \mid \mathfrak{M} \models \psi[a, \boldsymbol{b}/\boldsymbol{z}]\}| = c.$$

It is interesting to note that, by (3), $\mathrm{FOC}_{\leq}$ can be replaced by either of the quantifier logics $\mathrm{FO}_{\leq}(\mathsf{Maj})$ or $\mathrm{FO}_{\leq}(\mathsf{I})$, which have well-defined counterparts, $\mathrm{FO}(\mathsf{Maj})$ and $\mathrm{FO}(\mathsf{I})$, on unordered structures. Note that FOC does not make sense without order. Further note that, of the quantifiers Maj and I, the latter is also universe independent.

It was shown in [Lin95], [Ruh99] and [LMSV01], that multiplication cannot be defined in the logic $\mathrm{FO}_{\leq}(\mathsf{Maj})$. The following logics are therefore strictly more expressive than the logics in (3):

$$\mathrm{FO}_{\{+,\times\}}(\mathsf{Maj}) \equiv \mathrm{FO}_{\mathrm{BIT}}(\mathsf{Maj}) \equiv \mathrm{FO}_{\leq}(\mathsf{Maj}^2) \equiv \mathrm{FO}_{\leq}(\mathsf{D}). \tag{4}$$

For the first equivalence, i.e., $\mathrm{FO}_{\{+,\times\}} \equiv \mathrm{FO}_{\mathrm{BIT}}$, see [Imm99]. The second equivalence, where Maj^2 denotes the second vectorization of Maj (i.e., majority of pairs), was shown in [BIS90], and the last equivalence is due to [Luo04]. It is worth noting that, unlike Maj^2, D is a unary quantifier. Finally, we note that the logics in (4) are regular logics. This follows by Proposition 4.4 applied to $\mathrm{FO}_{\leq}(\mathsf{D})$.

As a demonstration of the power of regularity, we show that the undefinability of multiplication in any of the logics in (3) is actually a direct corollary of an old result of Krynicki and Lachlan.

Theorem 2.1 ([KL79])**.** *The* $\mathrm{FO}(\mathsf{I})$*-theory of* $\mathfrak{N} = \langle \mathbb{N}, + \rangle$*, i.e., the set of* $\mathrm{FO}(\mathsf{I})$*-sentences true in* $\mathfrak{N}$*, is decidable.*

Corollary 2.2. *Multiplication is not* $\mathrm{FO}(\mathsf{I})$*-definable on finite structures with built-in addition.*

Proof. Suppose towards contradiction that there were a formula $\mu(x, y, z)$ of $\mathrm{FO}(\mathsf{I})$ defining multiplication in finite structures with built-in addition. Working now in the structure $\mathfrak{N} = \langle \mathbb{N}, + \rangle$, we observe that we can interpret the structure $\langle n, + \rangle$ given one parameter $n \in \mathbb{Z}_+$. By regularity of $\mathrm{FO}(\mathsf{I})$, there is a $\mathrm{FO}(\mathsf{I})$-formula $\tilde{\mu}(x, y, z, p)$ of the vocabulary $\{+\}$ such that

$$\mathfrak{N} \models \tilde{\mu}[a/x, b/y, c/z, n/p] \text{ iff } \langle n, + \rangle \models \mu[a/x, b/y, c/z] \text{ and } a, b, c < n,$$

for $a, b, c \in \mathbb{N}$. The latter condition is equivalent to $ab = c < n$ and $a, b < n$. Hence, $\nu(x, y, z)\colon \exists p\, \tilde{\mu}(x, y, z, p)$ is a $\mathrm{FO}(\mathsf{I})$-formula that defines multiplication on $\mathfrak{N}$.

Given a sentence $\phi \in \mathrm{FO}[\{+, \times\}]$, there is an effective way to do the substitution which gives $\phi(\nu/\times) \in \mathrm{FO}(\mathsf{I})[\{+\}]$. Now we have

$$\langle \mathbb{N}, +, \times \rangle \models \phi \text{ iff } \mathfrak{N} \models \phi(\nu/\times),$$

which reduces the FO-theory of $\langle \mathbb{N}, +, \times \rangle$ to the decidable $\mathrm{FO}(\mathsf{I})$-theory of $\mathfrak{N}$. However, this is in contradiction with one of the most fundamental results of mathematical logic that FO-theory of $\langle \mathbb{N}, +, \times \rangle$ is undecidable. □

Cardinality quantifiers expressing equicardinality

In this section we briefly recall the result [Luo04] characterizing cardinality quantifiers C_S which can define the quantifier I on ordered structures. The periodicity of the set S is measured in terms of the following functions f_S and ω_S.

(a) Let $f\colon \mathbb{N} \to \mathbb{N}$ and $\Delta \subseteq \mathbb{N}$ be an interval. Then f is *periodic on* Δ *with period* $\omega \in \mathbb{N}$ if $f(x) = f(x+\omega)$ whenever $x, x+\omega \in \Delta$. A set $S \subseteq \mathbb{N}$ is *periodic on* Δ *with period* ω if its characteristic function χ_S is, i.e., $x, x+\omega \in \Delta$ implies $x \in S \iff x+\omega \in S$.

(b) Let $S \subseteq \mathbb{N}$. The functions $f_S, \omega_S\colon \mathbb{N} \to \mathbb{N}$ are defined in the following way. Let $n \in \mathbb{N}$. Then $f_S(n)$ is the least $\ell \in \mathbb{N}$ such that for some $\omega \in \mathbb{N}$, $0 < \omega \leq \ell$, the set S is periodic on the interval $\{i \in \mathbb{N} \mid \ell - \omega \leq i \leq n - (\ell - \omega)\}$ with period ω. Furthermore, $\omega_S(n)$ is the least $\omega \in \mathbb{N}$ such that S is periodic on the interval $\{i \in \mathbb{N} \mid f_S(n) - \omega \leq i \leq n - (f_S(n) - \omega)\}$ with period ω.

Theorem 2.3 ([Luo04]). *Let $S \subseteq \mathbb{N}$. Then the following are equivalent:*

1. *There are $k, \ell \in \mathbb{N}$ such that $k \cdot f_S(n) \cdot \omega_S(n)^\ell \geq n$, for almost all $n \in \mathbb{N}$.*
2. $\mathrm{FO}_\leq(\mathsf{I}) \leq \mathrm{FO}_\leq(\mathsf{C}_S)$.

In [Luo04], the inequality in condition 1 of Theorem 2.3 was required to hold for every $n \in \mathbb{N}$. It is easy to see that, if a set S satisfies condition 1 with $k, \ell \in \mathbb{N}$, then, by replacing k by a big enough k', the inequality will be satisfied by every $n \in \mathbb{N}$.

Next we apply Theorem 2.3 to certain interesting sets $S \subseteq \mathbb{N}$.

Lemma 2.4 ([Luo04]). *(a) Let $S = \mathrm{rg}(P)$ be the range of some polynomial P with nonnegative integer coefficients of degree at least two. Then the non-periodicity condition 1 of Theorem 2.3 holds and hence the quantifier I is definable in $\mathrm{FO}_\leq(\mathsf{C}_S)$.*

(b) Let $E = \{2^n \mid n \in \mathbb{N}\}$. Then the non-periodicity condition 1 of Theorem 2.3 holds and hence the quantifier I is definable in $\mathrm{FO}_\leq(\mathsf{C}_E)$.

(c) Let $F = \{n! \mid n \in \mathbb{N}\}$. Then the condition 1 of Theorem 2.3 fails, which implies $\mathrm{FO}_\leq(\mathsf{I}) \not\leq \mathrm{FO}_\leq(\mathsf{C}_F)$.

Complexity classes

Next we recall the complexity classes relevant for this article. In this article AC^0 and TC^0 refer to the classes of languages recognized by DLOGTIME-uniform families $(C_n)_{n\in\mathbb{N}}$ of constant depth polynomial-size circuits. For AC^0, the circuit C_n may have NOT and unbounded fan-in AND and OR gates. For TC^0, also unbounded fan-in MAJORITY gates are allowed, which output 1 if and only if at least half of the inputs are 1. The requirement of DLOGTIME-uniformity means that $(C_n)_{n\in\mathbb{N}}$, as a family of directed acyclic graphs, can be recognized by a random access machine in time $O(\log(n))$ (see [Vol99] for details).

In this article we are concerned with the logical counterparts of AC^0 and TC^0:

$$\mathrm{AC}^0 \equiv \mathrm{FO}_{\{+,\times\}}$$
$$\mathrm{TC}^0 \equiv \mathrm{FO}_{\{+,\times\}}(\mathsf{Maj}) \equiv \mathrm{FO}_{\{+,\times\}}(\mathsf{I}).$$

These logical characterizations were proved in the seminal paper [BIS90], where the connection between DLOGTIME-uniformity and $\mathrm{FO}_{\{+,\times\}}$-definability was established.

3 Cardinality quantifiers and TC^0

In this section we obtain new characterizations of TC^0 in terms of certain cardinality quantifiers.

Let $\mathrm{Sq} = \{n^2 \mid n \in \mathbb{N}\}$ be the set of squares and let us consider the quantifier C_{Sq}. By Lemma 2.4, the Härtig quantifier I is definable in $\mathrm{FO}_{\leq}(\mathsf{C}_{\mathrm{Sq}})$. On the other hand, Lynch showed in [Lyn82] that multiplication is already definable in terms of addition and the relation Sq. From this it directly follows that

$$\mathrm{FO}_{\leq}(\mathsf{C}_{\mathrm{Sq}}) \equiv \mathrm{TC}^0, \tag{5}$$

since the quantifier C_{Sq} is easily seen to be expressible in $\mathrm{FO}_{\{+,\times\}}(\mathsf{I})$. In this section we show that, in (5), Sq can be replaced by numerous sets S of natural numbers, among others by any range $S = \mathrm{rg}(P)$ where P is a polynomial P with nonnegative integer coefficients of degree at least two.

We use only elementary methods to achieve our goal. However, we need several steps to complete the proof. The key point is the elementary combinatorial fact that if $(A_i)_{i\in I}$ is a finite disjoint family of finite sets of equal size, then

$$\left|\bigcup_{i\in I} A_i\right| = |I|\,|A_{i_0}|$$

with $i_0 \in I$ arbitrary. Actually, this fact can be construed as a combinatorial definition of multiplication on natural numbers. We shall use cardinality quantifiers C_S for a partial logical implementation of this definition.

A *partial multiplication* is a ternary relation R on natural numbers such that $(a, b, c) \in R$ implies $a \cdot b = c$. We are going to show that under certain circumstances, it is possible to extend a partial multiplication to the multiplication restricted to the universe at hand. To this end, we first analyze first-order ways to extend partial multiplication.

For $k \in \mathbb{N}$ and a partial multiplication R, we define $\gamma(R, k)$ to be the biggest $r \in \mathbb{N}$ such that $r = 0$ or for every $a, b \in \mathbb{N}$ with $a \leq k$ and $b \leq r$, we have $(a, b, ab) \in R$. We fix ternary relation symbols A and M. By a *partial model of arithmetic* we understand a finite $\{A, M\}$-structure $\mathfrak{N}$ such that $\mathrm{Dom}(\mathfrak{N}) = \{0, 1, \ldots, n-1\}$, $A^{\mathfrak{N}}$ is addition restricted to n, i.e., $A^{\mathfrak{N}} = \{(a, b, c) \in \mathrm{Dom}(\mathfrak{N})^3 \mid a + b = c\}$ and $M^{\mathfrak{N}}$ is a partial multiplication.

The mapping γ has the following obvious monotonicity property: If $R \subseteq R' \subseteq \mathbb{N}^3$ and $k, k' \in \mathbb{N}$, $k \geq k'$, then $\gamma(R, k) \leq \gamma(R', k')$. In addition, we observe that if $\mathfrak{N}$ is a partial model of arithmetic with $n = \mathrm{card}(\mathfrak{N})$ and $0 < k < n$, then $\gamma(M^{\mathfrak{N}}, k) \leq \left\lfloor \frac{n-1}{k} \right\rfloor$.

Lemma 3.1. *There exists a first-order formula $\mu(x, y, z)$ such that for any partial model of arithmetic $\mathfrak{N}$, we have that $\mu^{\mathfrak{M}}$ is also a partial multiplication and for all $a, b \in \mathrm{Dom}(\mathfrak{N}) \setminus \{0\}$,*

a) $\gamma(\mu^{\mathfrak{N}}, a) \geq \gamma(M^{\mathfrak{N}}, a)$,

b) $\gamma(\mu^{\mathfrak{N}}, a) \geq b$ *iff* $\gamma(\mu^{\mathfrak{N}}, b) \geq a$ *and* *c) if $a < b \le a^2 + a$, then*

$$\gamma(\mu^{\mathfrak{N}}, b) \ge \min \left\{ \left\lfloor \frac{\gamma(M^{\mathfrak{N}}, a)}{\lfloor (b-1)/a \rfloor} \right\rfloor, \left\lfloor \frac{n-1}{b} \right\rfloor \right\}$$

where $n = \mathrm{card}(\mathfrak{N})$.

Proof. We use the basic algebraic laws of multiplication for extending $M^{\mathfrak{N}}$: The idea is simply that given $x, y \in \mathrm{Dom}(\mathfrak{N})$, if we are able to represent x as $x = tu + t'u'$, then we can calculate $xy = t(uy) + t'(u'y) = tv + t'v'$ with $v = uy$ and $v' = u'y$ provided that the appropriate products are definable by $M^{\mathfrak{N}}$. More formally, consider the formulas

$$\begin{aligned} \alpha(x, t, u, t', u') : & \quad \exists s \exists s' (M(t, u, s) \wedge M(t', u', s') \wedge A(s, s', x)) \\ \mu_-(x, y, z) : & \quad \exists t \exists t' \exists u \exists u' \exists v \exists v' (M(u, y, v) \wedge M(u', y, v') \\ & \quad \wedge \alpha(x, t, u, t', u') \wedge \alpha(z, t, v, t', v')) \end{aligned}$$

and

$$\mu(x, y, z) : \quad \mu_-(x, y, z) \wedge \mu_-(y, x, z).$$

Let $\mathfrak{N}$ be a partial model of arithmetic. For $a, d, d', e'e' \in \mathrm{Dom}(\mathfrak{N})$, it holds that $\mathfrak{N} \models \alpha[a/x, d/t, d'/t', e/u, e'/u']$ iff $a = de + d'e'$. Furthermore, for $a, b, c \in \mathrm{Dom}(\mathfrak{N})$ we have that $\mathfrak{N} \models \mu_-[a/x, b/y, c/z]$ iff there are $d, d', e, e' \in \mathrm{Dom}(\mathfrak{N})$ such that $a = de + d'e'$, $c = deb + d'e'b = ab$ and

$$\{(d, e, de), (d', e', d'e'), (e, b, eb), (e', b, e'b), (de, b, deb), (d'e', b, d'e'b)\} \subseteq M^{\mathfrak{N}}.$$

Hence, $\mu_-^{\mathfrak{N}}$ is a partial multiplication, and by commutativity of multiplication, this holds for $\mu^{\mathfrak{N}}$, too.

The lower bounds are now easy to derive.

a) Clearly, $\mu_-^{\mathfrak{N}} \subseteq \mu^{\mathfrak{N}}$ and we may use representations of the form $a_0 = a_0 \cdot 1 + 0 \cdot 0$ to show that $\gamma(M^{\mathfrak{N}}, a) \le \gamma(\mu_-^{\mathfrak{N}}, a)$. Hence, $\gamma(\mu^{\mathfrak{N}}, a) \ge \gamma(\mu_-^{\mathfrak{N}}, a) \ge \gamma(M^{\mathfrak{N}}, a)$.

b) By symmetry of μ and the definition of γ, this is obvious.

c) Suppose $a, b \in \mathrm{Dom}(\mathfrak{N})$ satisfy $0 < a < b \le a^2 + a$. Put

$$d = \min \left\{ \left\lfloor \frac{\gamma(M^{\mathfrak{N}}, a)}{\lfloor (b-1)/a \rfloor} \right\rfloor, \left\lfloor \frac{n-1}{b} \right\rfloor \right\}.$$

Then $d \le \gamma(M^{\mathfrak{N}}, a) < n$, $bd < n$, and we are to show that $\gamma(\mu^{\mathfrak{N}}, b) \ge d$. We may assume that $d > 0$, which implies $\gamma(M^{\mathfrak{N}}, a) > 0$.

Let $b_0, d_0 \in \mathbb{N}$ where $b_0 \le b$ and $d_0 \le d$. Consider the representation $b_0 = a \cdot q + r \cdot 1$ where $q = \left\lfloor \frac{b_0 - 1}{a} \right\rfloor \le \left\lfloor \frac{a^2 + a - 1}{a} \right\rfloor = a$ and $r \le a$. Since $d_0 \le d \le \gamma(M^{\mathfrak{N}}, a)$, we immediately see that $(q, d_0, qd_0), (1, d_0, d_0), (r, 1, r), (r, d_0, rd_0) \in M^{\mathfrak{N}}$. The critical case is that of the triple (a, qd_0, aqd_0), which is in $M^{\mathfrak{N}}$, because

$$qd_0 \leq \left\lfloor \frac{b-1}{a} \right\rfloor \cdot d \leq \left\lfloor \frac{b-1}{a} \right\rfloor \cdot \left\lfloor \frac{\gamma(M^{\mathfrak{N}}, a)}{\lfloor (b-1)/a \rfloor} \right\rfloor \leq \gamma(M^{\mathfrak{N}}, a).$$

Clearly also $(a, q, aq) \in M^{\mathfrak{N}}$. As in all needed cases we may multiply using $M^{\mathfrak{N}}$, we get $(b_0, d_0, b_0 d_0) \in \mu_-^{\mathfrak{N}} \subseteq \mu^{\mathfrak{N}}$. Consequently, $\gamma(\mu^{\mathfrak{N}}, b) \geq d$. □

Proposition 3.2. *For $k \in \mathbb{Z}_+$, there exists an $\{A, M\}$-formula $\pi(x, y, z)$ of* FO$[\{A, M\}]$ *such that the following condition holds: Let $\mathfrak{N}$ is a partial model of arithmetic with* card$(\mathfrak{N}) = n \geq k^2$. *Suppose that there is $a^* \in$* Dom$(\mathfrak{N})$ *such that*

$$n^{1/k} \leq a^* \leq n^{1-1/k}/k$$

and

$$ka^* \gamma(M^{\mathfrak{N}}, a^*) \geq n.$$

Then π is the multiplication restricted to $n =$ Dom$(\mathfrak{N})$.

Proof. We substitute μ of the preceding lemma repeatedly for M, getting $\pi_0(x, y, z) = M(x, y, z)$ and $\pi_{i+1} = \mu(\pi_i / M)$, for $i \in \mathbb{N}$. We claim that π_{i*} with $i^* = 2\lceil \mathrm{lb}\, k \rceil + 2$ has the required property, where $\mathrm{lb}\, k$ denotes the binary logarithm of k. Suppose $\mathfrak{N}$ satisfies the assumptions of the lemma. Consider the partial model of arithmetic $\mathfrak{N}_i$ with $\mathrm{Dom}(\mathfrak{N}_i) = n$ and $M^{\mathfrak{N}_i} = \pi_i^{\mathfrak{N}}$. Then for every $i \in \mathbb{N}$ and $a, b \in \{1, \ldots, n-1\}$ the preceding lemma gives

a) $\gamma(\pi_{i+1}^{\mathfrak{N}}, a) \geq \gamma(\pi_i^{\mathfrak{N}}, a)$,

b) $\gamma(\pi_{i+1}^{\mathfrak{N}}, a) \geq b$ iff $\gamma(\pi_{i+1}^{\mathfrak{N}}, b) \geq a$ and

c) if $a < b \leq a^2 + a$, then

$$\gamma(\pi_{i+1}^{\mathfrak{N}}, b) \geq \min \left\{ \left\lfloor \frac{\gamma(\pi_i^{\mathfrak{N}}, a)}{\lfloor (b-1)/a \rfloor} \right\rfloor, \left\lfloor \frac{n-1}{b} \right\rfloor \right\}.$$

Furthermore, these inequalities imply for $a \in \{1, \ldots, n-1\}$ and $i \in \mathbb{Z}_+$ that

d) $\gamma(\pi_{i+1}^{\mathfrak{N}}) \geq \min \left\{ 2\gamma(\pi_i^{\mathfrak{N}}, a), \left\lfloor \frac{n-1}{a} \right\rfloor \right\}$.

In order to prove this last statement, put $b = \gamma(\pi_i^{\mathfrak{N}}, a)$ and $c = \min \left\{ 2b, \left\lfloor \frac{n-1}{a} \right\rfloor \right\}$ so that we are to prove $\gamma(\pi_{i+1}^{\mathfrak{N}}, a) \geq c$. If $b = 0$ or $b = \left\lfloor \frac{n-1}{a} \right\rfloor$, then $b = c$ and the inequality follows directly from inequality a. So suppose that $0 < b < \left\lfloor \frac{n-1}{a} \right\rfloor$. Then by symmetry (item b), we have $\gamma(\pi_i^{\mathfrak{N}}, b) \geq a$, and as $0 < b < c \leq b^2 + b$, we have by estimate c that

$$\begin{aligned} \gamma(\pi_{i+1}^{\mathfrak{N}}, c) &\geq \min \left\{ \left\lfloor \frac{\gamma(\pi_i^{\mathfrak{N}}, b)}{\lfloor (c-1)/b \rfloor} \right\rfloor, \left\lfloor \frac{n-1}{c} \right\rfloor \right\} \\ &\geq \min \left\{ \gamma(\pi_i^{\mathfrak{N}}, b), \left\lfloor \frac{n-1}{\left\lfloor \frac{n-1}{a} \right\rfloor} \right\rfloor \right\} \geq a. \end{aligned}$$

Again by symmetry, $\gamma(\pi^{\mathfrak{N}}_{i+1}, a) \geq c$.

Set $a_1 = \min\left\{a^*, \gamma(M^{\mathfrak{N}}, a^*)\right\}$ and recursively $a_{i+1} = \min\{a_i^2, n-1\}$, for $i \in \mathbb{Z}_+$. Then we have that

1. $a_1^2 < n$, 2. $a_1^k \geq n$, 3. $ka_1\gamma(\pi^{\mathfrak{N}}_i, a_1) \geq n$ and 4. $\gamma(\pi^{\mathfrak{N}}_i, a_1) \geq a_1$,

for $i \in \mathbb{Z}_+$. The first inequality follows from $a_1^2 \leq a^*\gamma(M^{\mathfrak{N}}, a^*) \leq a^*\left\lfloor \frac{n-1}{a^*} \right\rfloor < n$. If $a_1 = a^*$, the rest of the inequalities are clear, so assume $a_1 = \gamma(M^{\mathfrak{N}}, a^*)$. Then by the assumptions of the lemma, we have $ka^*a_1 \geq n$ and $ka^* \leq n^{1-\frac{1}{k}}$, so $a_1 \geq n^{1/k}$, implying the second inequality. By a and b, we have $\gamma(\pi^{\mathfrak{N}}_i, a_1) \geq \gamma(M^{\mathfrak{N}}, a_1) \geq a^*$, i.e., last inequality. Consequently, $ka_1\gamma(\pi^{\mathfrak{N}}_i, a_1) \geq ka_1a^* \geq n$.

Consider $f\colon \{a_1, \ldots, n-1\} \to n$,

$$f(x) = \min\left\{ \left\lfloor \frac{\gamma(\mu^{\mathfrak{N}}, a_1)}{\lfloor x/a_1 \rfloor} \right\rfloor, \left\lfloor \frac{n-1}{x} \right\rfloor \right\}.$$

By induction on $i \in \mathbb{Z}_+$, we get for $b \in \{a_1, a_1+1, \ldots, a_i\}$ that $\gamma(\pi^{\mathfrak{N}}_i, b) \geq f(b)$. Indeed, the case $i = 1$ is trivial. Supposing the induction hypothesis holds for i, we need only to prove the inequality $\gamma(\pi^{\mathfrak{N}}_{i+1}, b) \geq f(b)$ for $b \in \{a_{i+1}, \ldots, a_i^2\}$ as $\gamma(\pi^{\mathfrak{N}}_{i+1}, b) \geq \gamma(\pi^{\mathfrak{N}}_i, b)$. Then $a_i < b \leq a_i^2 + a_i$ so by induction hypothesis and inequality c, we have

$$\begin{aligned}
\gamma(\pi^{\mathfrak{N}}_{i+1}, b) &\geq \min\left\{ \left\lfloor \frac{\gamma(\pi^{\mathfrak{N}}_i, a_i)}{\lfloor (b-1)/a_i \rfloor} \right\rfloor, \left\lfloor \frac{n-1}{b} \right\rfloor \right\} \\
&\geq \min\left\{ \left\lfloor \frac{f(a_i)}{\lfloor b/a_i \rfloor} \right\rfloor, \left\lfloor \frac{n-1}{b} \right\rfloor \right\} = \left\lfloor \min\left\{ \frac{f(a_i)}{\lfloor b/a_i \rfloor}, \frac{n-1}{b} \right\} \right\rfloor \\
&= \left\lfloor \min\left\{ \frac{\gamma(\mu^{\mathfrak{N}}, a_1)}{\lfloor b/a_i \rfloor \lfloor a_i/a_1 \rfloor}, \frac{\left\lfloor \frac{n-1}{a_i} \right\rfloor}{\left\lfloor \frac{b}{a_i} \right\rfloor}, \frac{n-1}{b} \right\} \right\rfloor \\
&\geq f(b),
\end{aligned}$$

as for arbitrary $r, s \geq 0$, it holds that $\lfloor rs \rfloor \geq \lfloor r \rfloor \lfloor s \rfloor$.

Obviously $a_i = \min\left\{a_i^{2^{i-1}}, n-1\right\}$, for $i \in \mathbb{Z}_+$. In particular, we get $a_{i^*/2} = a_{\lceil \mathrm{lb}\, k \rceil + 1} = n-1$. Hence, for $b \in \{a_1, a_1+1, \ldots, n-1\}$, it holds that

$$\gamma(\pi^{\mathfrak{N}}_{i^*/2}, b) \geq f(b) \geq \min\left\{ \left\lfloor \frac{n/ka_1}{\lfloor b/a_1 \rfloor} \right\rfloor, \left\lfloor \frac{n-1}{b} \right\rfloor \right\} \geq \left\lfloor \frac{n}{kb} \right\rfloor,$$

as $ka_1\gamma(\mu^{\mathfrak{N}}, a_1) \geq n$. We would like to have the same estimate for $b \in \{1, \ldots, a_1 - 1\}$, too. If $\left\lfloor \frac{n}{kb} \right\rfloor < a_1$, then by natural monotonicity of γ, we have

$$\gamma(\pi^{\mathfrak{N}}_{i^*/2}, b) \geq \gamma(\pi^{\mathfrak{N}}_{i^*/2}, a_1) \geq a_1 > \left\lfloor \frac{n}{kb} \right\rfloor,$$

so suppose $\left\lfloor \frac{n}{kb} \right\rfloor \geq a_1$. Then we may apply the estimate for $\left\lfloor \frac{n}{kb} \right\rfloor$, getting

$$\gamma(\pi^{\mathfrak{N}}_{i^*/2}, \left\lfloor \frac{n}{kb} \right\rfloor) \geq \left\lfloor \frac{n}{k \left\lfloor \frac{n}{kb} \right\rfloor} \right\rfloor \geq b,$$

and by symmetry (item b), $\gamma(\pi^{\mathfrak{N}}_{i^*/2}, b) \geq \left\lfloor \frac{n}{kb} \right\rfloor$.

To finish the proof, we show that for every $b \in \{1, \ldots, n-1\}$, we have $\gamma(\pi^{\mathfrak{N}}, b) = \gamma(\pi_{i^*}, b) = \left\lfloor \frac{n-1}{b} \right\rfloor$, which is equivalent to claiming that $\pi^{\mathfrak{N}}$ is the multiplication restricted to n.

Applying inequality d repeteadly, we get

$$\begin{aligned}\gamma(\pi^{\mathfrak{N}}_{i^*/2+j}, b) &\geq \min\left\{2^j \gamma(\pi^{\mathfrak{N}}_{i^*/2}, b), \left\lfloor \frac{n-1}{b} \right\rfloor\right\} \\ &\geq \min\left\{2^j \left\lfloor \frac{n}{kb} \right\rfloor, \left\lfloor \frac{n-1}{b} \right\rfloor\right\},\end{aligned}$$

for $b \in \{1, \ldots, n-1\}$ and $j \in \mathbb{N}$. Furthermore, for $b \leq n/k$, we get $2^{i^*/2} \lfloor n/kb \rfloor \geq 2k \cdot (n/2kb) = n/b$ and therefore $\gamma(\pi^{\mathfrak{N}}, b) = \left\lfloor \frac{n-1}{b} \right\rfloor$. If $b > n/k \geq k$, then $\left\lfloor \frac{n-1}{b} \right\rfloor \leq n/k$, so $\gamma(\pi^{\mathfrak{N}}, a) = \left\lfloor \frac{n-1}{a} \right\rfloor$ holds for $a = \left\lfloor \frac{n-1}{b} \right\rfloor$. However, clearly $\left\lfloor \frac{n-1}{a} \right\rfloor \geq b$, so by the symmetry property $\gamma(\pi^{\mathfrak{N}}, a) \geq b$ implies $\gamma(\pi^{\mathfrak{N}}, b) \geq a = \left\lfloor \frac{n-1}{b} \right\rfloor$. Hence, $\gamma(\pi^{\mathfrak{N}}, b) = \left\lfloor \frac{n-1}{b} \right\rfloor$. □

Now that possibilities for extending a partial multiplication have been explored, it is relatively easy to extract some sufficient properties of cardinality quantifiers that are needed to define multiplication. To start with, we need some notation.

For an set $S \subseteq \mathbb{N}$, we define mappings $\delta_S \colon S \to \mathbb{Z}_+$ and $\gamma_S \colon \mathbb{Z}_+ \times \mathbb{Z}_+ \to \mathbb{N}$ as follows: For each $n \in S$, $\delta_S(n) = m - n$ where m is the least number in S which is bigger than n if such an m exist, otherwise $\delta_S(n) = \infty$. For every $n, t \in \mathbb{Z}_+$, let $S_{n,t}$ be the set of $m \in S$ with $\delta_S(m) \geq t$ and $m + t < n$. Put $\gamma_S(n, t) = |S_{n,t}|$. For $\chi \colon \mathbb{N} \to \{0, 1\}$ and a binary word $w \in \{0, 1\}^s$, $s \in \mathbb{N}$, put

$$T^{\chi}_{w} = \{m \in \mathbb{N} \mid \forall i \in \{0, \ldots, s-1\} \, (\chi_S(m+i) = w(i))\}.$$

Definition 3.3. Let $S \subseteq \mathbb{N}$, $n \in \mathbb{N}$ and $\varepsilon > 0$. The set S is *ε-loose relative to n* if there exists $t \in \mathbb{Z}_+$ satisfying the inequalities $t\gamma_S(n, t) \geq \varepsilon n$ and $n^{\varepsilon} \leq t \leq n^{1-\varepsilon}$. S is *loose* if there is $\varepsilon > 0$ such that for almost all $n \in \mathbb{Z}_+$ it holds that S in ε-loose relative to n. S is *pseudoloose* if for some $\varepsilon > 0$, we have that for almost all $n \in \mathbb{Z}_+$ there exists a word $w \in \{0, 1\}^s$ such that $s \leq n^{1-\varepsilon}$ and $T^{\chi_S}_{w}$ is ε-loose relative to n.

We note that if S is loose, then it is obviously pseudoloose. We are now quite close to fulfilling our goal. Two steps remain: To implement the combinatorial idea, we need to be able to count sizes of sets, i.e., we need $\mathrm{FO}_{\leq}(\mathsf{I}) \leq \mathrm{FO}_{\leq}(\mathsf{C}_S)$. Then we need to show that a sufficient part of the definition of multiplication can be formalized in $\mathrm{FO}_{\leq}(\mathsf{C}_S)$. For the first subgoal, we need some combinatorial analysis of the set S. Write $[a, b]_{\mathbb{N}} = [a, b] \cap \mathbb{N} = \{n \in \mathbb{N} \mid a \leq n \leq m\}$ for $a, b \in \mathbb{N}$ (and similarly for other types of intervals).

Lemma 3.4. *For infinite $S \subseteq \mathbb{N}$ and $n, t \in \mathbb{Z}_+$, we have $t(\gamma_S(n, t) - 1) \leq 2f_S(n)$ or $t \leq \omega_S(n)$.*

Proof. Suppose $t > \omega_S(n)$. Let $S_{n,t}$ be as above and $\Delta = [f_S(n), n - f_S(n)]_{\mathbb{N}}$. By definition of ω_S, it holds that S is periodic on the interval $[f_S(n) - \omega_S(n), n - f_S(n) + \omega_S(n)]$, so for all $a \in S \cap \Delta$ we have $a + \omega_S(n) \in S$ and thus $\delta_S(a) \leq \omega_S(n) < t$, which implies $a \notin S_{n,t}$. Hence, $\Delta \cap S_{n,t} = \emptyset$. Moreover, $[a, a+t[_{\mathbb{N}} \cap \Delta = \emptyset$ for $a \in S_{n,t}$ except for one possible exception which, if exists, is the unique $a \in S_{n,t}$ for which $f_S(n) \in [a, a+t[_{\mathbb{N}}$. Let s be this unique exception if it exists, otherwise pick any s. Because the union $A = \bigcup_{a \in S_{n,t} \smallsetminus \{s\}} [a, a+t[_{\mathbb{N}}$ is disjoint and included in $\{0, \ldots, n-1\} \smallsetminus \Delta$, we get $t(\gamma_S(n,t) - 1) = |A| \leq \{0, \ldots, n-1\} \smallsetminus \Delta \leq 2 f_S(n)$. □

Lemma 3.5. *Let $S \subseteq \mathbb{N}$, $w \in \{0,1\}^s$ and $T = T_w^{\chi_S}$. Then for every $n \in \mathbb{N}$ we have $f_T(n) \leq f_S(n) + s$. Moreover, $\omega_T(n) \leq \omega_S(n)$ provided that $3 f_S(n) + 3s \leq n$.*

Proof. For the first inequality, we may assume that $f_S(n) + s \leq \left\lceil \frac{n}{2} \right\rceil$, but then it is obvious that the periodicity of S on $[f_S(n) - \omega_S(n), n - f_S(n) + \omega_S(n)]_{\mathbb{N}}$ implies the periodicity of T on $\Delta = [f_S(n) + s - \omega_S(n), n - f_S(n) - s + \omega_S(n)]_{\mathbb{N}}$, both with period $\omega_S(n)$, and the inequality follows.

For the second inequality, suppose towards contradiction that $\omega_T(n) > \omega_S(n)$, but $3 f_S(n) + 3s \leq n$. Besides Δ as above, consider $\Delta^* = [f_T(n) - \omega_S(n), n - f_T(n) + \omega_T(n)]_{\mathbb{N}}$. We already know that T is periodic on Δ with period $\omega_S(n)$, so the reason for $\omega_T(n) > \omega_S(n)$ must be that $\Delta \subsetneq \Delta^*$ and on Δ^*, T has not got period $\omega_S(n)$. However, $|\Delta| = n - 2 f_S(n) - 2s + 2\omega_S(n) + 1 \geq f_S(n) + s + 2\omega_S(n) + 1 > f_T(n) + \omega_S(n) \geq \omega_T(n) + \omega_S(n)$, so for $x, x + \omega_S(n) \in \Delta^*$, we may pick $y \in \Delta$ such that also $y + \omega_S(n) \in \Delta$ and $y \equiv x \,(\mathrm{mod}\ \omega_T(n))$. As T has period $\omega_T(n)$ on Δ^* and period $\omega_S(n)$ on Δ, we get $x \in T$ iff $y \in T$ iff $y + \omega_S(n) \in T$ iff $x + \omega_S(n) \in T$. Hence, T has period $\omega_S(n)$ on Δ^* in contradiction with the minimality of $\omega_T(n)$. □

Proposition 3.6. *If S is pseudoloose, then* $\mathrm{FO}_{\leq}(\mathsf{I}) \leq \mathrm{FO}_{\leq}(\mathsf{C}_S)$.

Proof. We need to verify the non-periodicity condition of Theorem 2.3. By definition of pseudolooseness, there exists $\varepsilon > 0$ such that for almost all $n \in \mathbb{Z}_+$ there is $w \in \{0,1\}^2$ with $s \leq n^{1-\varepsilon}$ such that $T_w^{\chi_S}$ is ε-loose relative to n. Fix n for a moment and put $T = T_w^{\chi_S}$. Choose $t \in \mathbb{Z}_+$ such that $t \gamma_T(n,t) \geq \varepsilon n$ and $n^\varepsilon \leq t \leq n^{1-\varepsilon}$. By Lemma 3.4, we get $f_T(n) \geq \frac{1}{2} t(\gamma_T(n,t) - 1) \geq \frac{1}{2}(\varepsilon n - n^{1-\varepsilon})$ or $\omega_T(n) \geq t \geq n^\varepsilon$. In the former case, we have $f_S(n) \geq f_T(n) - s \geq \frac{1}{2}\varepsilon n - \frac{3}{2} n^{1-\varepsilon}$, by Lemma 3.5. In the latter case, we have $\omega_S(n) \geq \omega_T(n) \geq n^\varepsilon$ or $f_S(n) > \frac{1}{3} n - s \geq \frac{1}{3} n - n^{1-\varepsilon}$, again by Lemma 3.5.

These estimates imply that for almost all $n \in \mathbb{N}$, we have $f_S(n) \geq \frac{1}{4}\varepsilon n$ or $\omega_S(n) \geq n^\varepsilon$. Choosing $l \in \mathbb{Z}_+$ with $l\varepsilon \geq 1$ and $k = 4\varepsilon$ we get $k f_S(n) \omega_S(n)^l \geq \max\left\{4l f_S(n), \omega_S(n)^l\right\} \geq \min\left\{(l\varepsilon)n, n^{l\varepsilon}\right\} \geq n$, for almost all $n \in \mathbb{Z}_+$. Hence, nonperiodicity condition of Theorem 2.3 follows. □

Theorem 3.7. *Let $S \subseteq \mathbb{N}$ be a pseudoloose set. Then*

$$\mathrm{FO}_{\leq}(\mathsf{C}_S) \geq \mathrm{TC}^0.$$

Proof. With eye on the characterization $\mathrm{FO}_{\{+,\times\}}(\mathsf{I}) \equiv \mathrm{TC}^0$, we need to show that the Härtig quantifier, addition and multiplication are definable in $\mathrm{FO}_{\leq}(\mathsf{C}_S)$. The definability of I and addition follows from the preceding proposition and Härtig's observation. It remains to be shown that multiplication is definable.

We proceed rather informally. Let $\mathfrak{N}$ be an ordered $\{\leq\}$-structure with $n = \mathrm{card}(\mathfrak{N})$. For notational simplicity, we assume that $\mathrm{Dom}(\mathfrak{N}) = \{0, \ldots, n-1\}$ and $\leq^{\mathfrak{N}}$ is the natural order. Let $\alpha(x, y, z)$ be a formula in $\mathrm{FO}(\mathsf{C}_S)$ defining addition (restricted to the appropriate domain) on ordered structures. Write $S(x)$ for $\mathsf{C}_S\, t(t < x)$; then $S^{\mathfrak{N}} = S \cap \{0, \ldots, n-1\}$. Similarly, there is a formula $\tau(x, \boldsymbol{q})$ of $\mathrm{FO}(\mathsf{C}_S)$ such that if $w \in \{0,1\}^s$, then τ defines $T_w^{\chi_S} \cap \{0, \ldots, n-s-1\}$ with parameters. The parameters $\boldsymbol{q}$ simply point to some occurence of w if such exists, and addition is used to find other occurences.

Given $t \in \mathbb{Z}_+$, consider now $T = T_w^{\chi_S}$ and $T_{n,t} = \{m \in T \mid \delta_T(n) \geq t, m + t < n\}$. It is straightforward to write $\theta(x, t, \boldsymbol{p})$ of $\mathrm{FO}(\mathsf{C}_S)$ such that with appropriate parameters (interpretations of t and $\boldsymbol{p}$), θ defines $T' = T_{n,t} \cap \{0, \ldots, n-s-t\}$. Finally, we write a formula $\nu(x, y, z)$ of $\mathrm{FO}(\mathsf{C}_S)$ defining partial multiplication by expressing the following: Let $a, b, c \in \mathrm{Dom}(\mathfrak{N})$. Then $\mathfrak{N} \models \nu[a/k, b/y, c/z]$ iff there are parameters such that T' as above has an initial segment T'' of size $|T''| = b$, the union $C = \bigcup_{d \in T''} [d, d+a-1[_{\mathbb{N}}$ is disjoint and $c = |C|$. Note that here we use, again, definability of addition and I for defining the appropriate intervals and for comparing sizes.

Next, we estimate the extent of the partial multiplication $\nu^{\mathfrak{N}}$. Given $t \in \mathbb{Z}_+$, we note that we may compute the product ab where $a, b \in \mathrm{Dom}(\mathfrak{N})$ with the idea presented provided that $a < t$ and $b \leq |T'|$. Therefore

$$\gamma(\nu^{\mathfrak{N}}, t) \geq |T'| \geq |T_{n,t}| - \left(\frac{s+t}{t} + 1\right) = \gamma_T(n, t) - s/t - 2.$$

As S is pseudoloose, there is $\varepsilon > 0$ such that for almost all $n \in \mathbb{Z}_+$ we can choose parameters $s, t \in \mathbb{Z}_+$ so that for $\{\leq\}$-structure $\mathfrak{N}$ with $n = \mathrm{card}(\mathfrak{N})$,

$$\gamma(\nu^{\mathfrak{N}}, t) \geq \gamma_T(n, t) - s/t - 2 \geq \frac{\varepsilon n - n^{1-\varepsilon}}{t} - 2 \geq \frac{\varepsilon n - 2n^{1-\varepsilon}}{t}$$

and $n^{\varepsilon} \leq t \leq n^{1-\varepsilon}$. Fixing any $k \in \mathbb{Z}_+$ with $k\varepsilon > 1$, this means that for almost all $n \in \mathbb{Z}_+$ there is $t \in \mathbb{Z}_+$ such that $n^{1/k} \leq t \leq n^{1-1/k}/k$ and $kt\gamma(\nu^{\mathfrak{N}}, t) \geq n$. In other words, the assumptions of Lemma 3.2 are satisfied for $\boldsymbol{u}^{\mathfrak{N}}$ as the partial multiplication relation. Consequently, the formula $\pi^* = \pi(\alpha/A, \nu/M)$ of $\mathrm{FO}(\mathsf{C}_S)$ defines multiplication in sufficiently large ordered structures. Once we fix the finitely many exceptions of small structures, we are done. □

Corollary 3.8. *a) Let* $P\colon \mathbb{N} \to \mathbb{N}$ *be a polynomial function with coefficients in* $\mathbb{N}$ *and* $\deg(P) = k \geq 2$, *and put* $S = \mathrm{rg}(P)$. *Then* $\mathrm{FO}_{\leq}(\mathsf{C}_S) \equiv \mathrm{TC}^0$.

b) Let $r > 1$ *be a real and* $S_r = \{\lfloor x^r \rfloor \mid x \in \mathbb{N}\}$. *Then* $\mathrm{FO}_{\leq}(\mathsf{C}_{S_r}) \geq \mathrm{TC}^0$.

Proof. Clearly, $\mathrm{FO}_{\leq}(\mathsf{C}_S) \leq \mathrm{TC}^0$. By considering the differences of subsequent elements in the sets S and S_r, it is easy to arrive at estimates which show that these are loose. □

4 Built-in relations

Logics with built-in relations

We call any relation S on the natural numbers $\mathbb{N}$ an *arithmetical relation*. If $\mathcal{B}$ is a set of arithmetical relations, then any finite τ-structure $\mathfrak{M}$ can be expanded into a τ-structure *with built-in relations* $\mathcal{B}$ by fixing a bijection f between $\mathrm{Dom}(\mathfrak{M})$ and $n = \{0, \ldots, n-1\} \in \mathbb{N}$: $\mathfrak{M}^f_{\mathcal{B}} = (\mathfrak{M}, (S^f)_{S \in \mathcal{B}})$, where $S^f = \{\boldsymbol{a} \mid f\boldsymbol{a} \in S\}$ for each $S \in \mathcal{B}$. We denote the class of finite τ-structures with built-in relations $\mathcal{B}$ by $\mathrm{Str}_{\mathcal{B}}(\tau)$.

Let $\mathcal{B}$ be a set of arithmetical relations, and let $\mathcal{L}$ be an abstract logic (we refer to [Ebb85] for the definition of abstract logics). We obtain the corresponding logic $\mathcal{L}_{\mathcal{B}}$ *with built-in relations* $\mathcal{B}$ by assuming that all the structures considered have built-in relations $\mathcal{B}$. Thus, for each vocabulary τ, the set $\mathcal{L}_{\mathcal{B}}(\tau)$ of τ-sentences of $\mathcal{L}_{\mathcal{B}}$ is simply $\mathcal{L}(\tau \cup \sigma_{\mathcal{B}})$, where $\sigma_{\mathcal{B}}$ is the vocabulary of the structure $(\mathbb{N}, (S)_{S \in \mathcal{B}})$. Furthermore, the truth relation is inherited from $\mathcal{L}$: for all $\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}}(\tau)$ and $\phi \in \mathcal{L}_{\mathcal{B}}(\tau)$, $\mathfrak{M} \models_{\mathcal{L}_{\mathcal{B}}} \phi \iff \mathfrak{M} \models_{\mathcal{L}} \phi$. In the sequel we will call logics (models) with built-in relations $\mathcal{B}$ just *$\mathcal{B}$-logics* (*$\mathcal{B}$-models*, respectively).

To simplify the presentation, we will use the convention that a *formula* is a sentence in an expanded vocabulary, i.e., we regard variables as constant symbols, and an $\mathcal{L}_{\mathcal{B}}(\tau)$-formula ϕ with free variables $x_0 \ldots, x_{k-1}$ is just a sentence in $\mathcal{L}_{\mathcal{B}}(\tau \cup \{x_0, \ldots, x_{k-1}\})$. Let $\mathfrak{M}$ be a $\mathcal{B}$-model of vocabulay τ, and let ϕ be an $\mathcal{L}_{\mathcal{B}}(\tau)$-formula with free variables $x_0, \ldots, x_{k-1}$. We write $\mathfrak{M} \models \phi[a_0/x_0, \ldots, a_{k-1}/x_{k-1}]$, or more briefly $\mathfrak{M} \models \phi[\boldsymbol{a}/\boldsymbol{x}]$, if $\mathfrak{M}^+ \models \phi$, where $\mathfrak{M}^+$ is the $\tau \cup \{x_0, \ldots, x_{k-1}\}$-expansion of $\mathfrak{M}$ with $a_i = x_i^{\mathfrak{M}^+}$ for each $i < k$. Furthermore, we use the notation $\phi^{\mathfrak{M}}$ for the *relation* $\{\boldsymbol{a} \in \mathrm{Dom}(\mathfrak{M})^k \mid \mathfrak{M} \models \phi[\boldsymbol{a}/\boldsymbol{x}]\}$ *defined by the formula* ϕ in the model $\mathfrak{M}$.

Each sentence ϕ of a $\mathcal{B}$-logic $\mathcal{L}_{\mathcal{B}}$ *defines* a class K_ϕ of $\mathcal{B}$-models: $K_\phi = \{\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}}(\tau) \mid \mathfrak{M} \models \phi\}$, where τ is the vocabulary of ϕ. The expressive power of $\mathcal{L}_{\mathcal{B}}$ is determined by the collection of classes that are definable by $\mathcal{L}_{\mathcal{B}}$-sentences. Since we want to compare the expressive power of logics with different sets of built-in relations, we generalize the notion of definability as follows. Let $\mathcal{B}$ and $\mathcal{B}^+$ be sets of built-in relations such that $\mathcal{B} \subseteq \mathcal{B}^+$. The class of $\mathcal{B}^+$-models *defined by a sentence* $\phi \in \mathcal{L}_{\mathcal{B}}(\tau)$ is $K_{\phi,\mathcal{B}^+} = \{\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}^+}(\tau) \mid \mathfrak{M} \models \phi\}$. Here we use the standard convention that if $\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}^+}(\tau)$ is a $\mathcal{B}^+$-model and ϕ is a sentence in $\mathcal{L}_{\mathcal{B}}(\tau)$, then $\mathfrak{M} \models \phi$ if and only if $\mathfrak{M}^- \models \phi$, where $\mathfrak{M}^-$ is the $\mathcal{B}$-model obtained from $\mathfrak{M}$ by dropping the built-in relations in $\mathcal{B}^+ \setminus \mathcal{B}$. We say that a class $K \subseteq \mathrm{Str}_{\mathcal{B}^+}(\tau)$ is *definable* in $\mathcal{L}_{\mathcal{B}}$ if $K = K_{\phi,\mathcal{B}^+}$ for some sentence $\phi \in \mathcal{L}_{\mathcal{B}}(\tau)$.

Now we can define the comparison between a $\mathcal{B}$-logic $\mathcal{L}_{\mathcal{B}}$ and a $\widetilde{\mathcal{B}}$-logic $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$: $\mathcal{L}_{\mathcal{B}}$ is *at most as strong* (w.r.t. expressive power) as $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$, in symbols $\mathcal{L}_{\mathcal{B}} \leq \widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$, if every class $K \subseteq \mathrm{Str}_{\mathcal{B} \cup \widetilde{\mathcal{B}}}(\tau)$ which is definable in $\mathcal{L}_{\mathcal{B}}$ is also definable in $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$. Furthermore, we say that $\mathcal{L}_{\mathcal{B}}$ is *strictly weaker* than $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$, $\mathcal{L}_{\mathcal{B}} < \widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$, if $\mathcal{L}_{\mathcal{B}} \leq \widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$ and $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}} \not\leq \mathcal{L}_{\mathcal{B}}$. If $\mathcal{L}_{\mathcal{B}} \leq \widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$ and $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}} \leq \mathcal{L}_{\mathcal{B}}$, we write $\mathcal{L}_{\mathcal{B}} \equiv \widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$, and say that $\mathcal{L}_{\mathcal{B}}$ and $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$ are equivalent.

As an example, consider first order logic with two sets $\mathcal{B}$ and $\widetilde{\mathcal{B}}$ of built-in relations. If $\mathrm{FO}_{\mathcal{B}} \leq \mathrm{FO}_{\widetilde{\mathcal{B}}}$, then for each relation $S \in \mathcal{B}$ there is a formula $\phi_S \in \mathcal{L}_{\widetilde{\mathcal{B}}}(\emptyset)$ such that

ϕ_S defines S in every $\widetilde{\mathcal{B}}$-model $\mathfrak{M}$: $\phi_S^{\mathfrak{M}} = S^f$, where $f : \mathrm{Dom}(\mathfrak{M}) \to |\mathrm{Dom}(\mathfrak{M})|$ is the bijection that determines the built-in relations in $\mathfrak{M}$. This is because the class $K_S = \{\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}\cup\widetilde{\mathcal{B}}}(\{x_0, \dots, x_{k-1}\}) \mid (x_0^{\mathfrak{M}}, \dots, x_{k-1}^{\mathfrak{M}}) \in S^{\mathfrak{M}}\}$ is trivially definable in $\mathrm{FO}_{\mathcal{B}}$, so by assumption, it is also definable in $\mathcal{L}_{\widetilde{\mathcal{B}}}$. The converse is also true: if all relations S in $\mathcal{B}$ are definable by formulas ϕ_S of $\mathrm{FO}_{\widetilde{\mathcal{B}}}$ in this way, then $\mathrm{FO}_{\mathcal{B}} \leq \mathrm{FO}_{\widetilde{\mathcal{B}}}$. This is easy to prove by using the fact that FO allows substituting relation symbols by formulas (see [Ebb85]).

Generalized quantifiers with built-in relations

The notion of generalized quantifier needs to be adapted to the framework of built-in relations. While ordinary quantifiers can be used without problems in connection with $\mathcal{B}$-logics, the characterization of semiregularity for $\mathcal{B}$-logics (see Proposition 4.3) requires the notion of *$\mathcal{B}$-quantifier*. Given a class $K_Q \subseteq \mathrm{Str}_{\mathcal{B}}(\tau)$ which is closed under isomorphisms, a $\mathcal{B}$-quantifier is a syntactic operator which can be used for binding tuples of variables $\boldsymbol{x}_R$ and formulas ψ_R for $R \in \tau$ to obtain a new formula $Q\,(\boldsymbol{x}_R\psi_R)_{R\in\tau}$. The semantics of Q is given by the clause

$$\mathfrak{M} \models Q\,(\boldsymbol{x}_R\psi_R)_{R\in\tau} \iff (\mathrm{Dom}(\mathfrak{M}), (\psi_R^{\mathfrak{M}})_{R\in\tau}, (S^{\mathfrak{M}})_{S\in\mathcal{B}}) \in K_Q.$$

The crucial difference to the standard definition of generalized quantifiers is that the built-in relations are not explicit in the syntax, whence they cannot be bound. This is important, since it is necessary to forbid substituting definable relations in place of them.

As an example of a $\mathcal{B}$-quantifier, consider the class $K_Q \subseteq \mathrm{Str}_{\leq}(\{P_a, P_b\})$ that consists of all models $\mathfrak{M}$ such that $|P_a^{\mathfrak{M}}| = |P_b^{\mathfrak{M}}|$, $\mathrm{Dom}(\mathfrak{M}) = P_a^{\mathfrak{M}} \cup P_b^{\mathfrak{M}}$, and $u <^{\mathfrak{M}} v$ whenever $u \in P_a^{\mathfrak{M}}$ and $v \in P_b^{\mathfrak{M}}$. Thus, a formula $Q\,xy(\phi, \psi)$ says essentially that the formulas ϕ and ψ, together with the built-in order $\leq$, define a string which belongs to the language $L = \{a^n b^n \mid n \in \mathbb{N}\}$.

More generally, given any language $L \subseteq \Sigma^*$, we can define the corresponding *language quantifier*: Q_L is the $\{\leq\}$-quantifier of the vocabulary $\tau_\Sigma = \{P_a \mid a \in \Sigma\}$ such that K_{Q_L} consists of all models $\mathfrak{M} \in \mathrm{Str}_{\{\leq\}}(\tau_\Sigma)$ which encode strings in L.

Any ordinary quantifier Q can also be interpreted as a $\mathcal{B}$-quantifier $Q_{\mathcal{B}}$ simply by expanding the models in K_Q by all possible ways to $\mathcal{B}$-models: let $K_{Q_{\mathcal{B}}} = \{\mathfrak{M}_{\mathcal{B}}^f \mid \mathfrak{M} \in K_Q, f : \mathrm{Dom}(\mathfrak{M}) \to |\mathrm{Dom}(\mathfrak{M})| \text{ is a bijection}\}$. For example, if Q is the divisibility quantifier D_2, then we obtain in this way the language quantifier $(\mathsf{D}_2)_{\mathcal{B}} = Q_L$, where $L = \{w \in \{a,b\}^* \mid |\{i \mid w_i = a\}| \text{ is even}\}$.

Note that since the class $K_{Q_{\mathcal{B}}}$ is invariant with respect to the built-in relations $\mathcal{B}$, the formulas $Q\,(\boldsymbol{x}_R\psi_R)_{R\in\tau}$ and $Q_{\mathcal{B}}\,(\boldsymbol{x}_R\psi_R)_{R\in\tau}$ are always equivalent. Thus, we can say that an ordinary quantifier Q is definable in a $\mathcal{B}$-logic $\mathcal{L}_{\mathcal{B}}$ just in case $Q_{\mathcal{B}}$ is definable in $\mathcal{L}_{\mathcal{B}}$. Similarly, if $\mathcal{B}$ and $\mathcal{B}'$ are two sets of arithmetical relations such that $\mathcal{B} \subset \mathcal{B}'$ (or all the relations $S \in \mathcal{B}$ are definable in $\mathcal{L}_{\mathcal{B}'}$), we can say that a $\mathcal{B}$-quantifier Q is definable in $\mathcal{L}_{\mathcal{B}'}$ if the corresponding expansion of Q to a $\mathcal{B}'$-quantifier $Q_{\mathcal{B}'}$ is definable in $\mathcal{L}_{\mathcal{B}'}$. For example, the definability of a language quantifier Q_L in $\mathrm{FO}_{\{+,\times\}}$ should be understood in this way.

Semiregular $\mathcal{B}$-logics

In the next two subsections, we will introduce the key notions of regularity for logics with built-in relations. These notions are obtained by adapting the usual framework of Abstract Logic (see [Ebb85]). We say that a $\mathcal{B}$-logic $\mathcal{L}_\mathcal{B}$ is *semiregular*, if for every vocabulary τ, all atomic formulas in τ are expressible in $\mathcal{L}_\mathcal{B}(\tau)$, all relations in $\mathcal{B}$ are definable in $\mathcal{L}_\mathcal{B}(\tau)$, $\mathcal{L}_\mathcal{B}$ is closed under Boolean operations and first order quantification, and $\mathcal{L}_\mathcal{B}$ is *closed under substitution*:

(s) If $\psi_R(\boldsymbol{x}_R)$ are $\mathcal{L}_\mathcal{B}(\sigma)$-formulas with $|\boldsymbol{x}_R| = \mathrm{ar}(R)$ for each $R \in \tau$, and ϕ is an $\mathcal{L}_\mathcal{B}(\tau)$-sentence, then there is an $\mathcal{L}_\mathcal{B}(\sigma)$-sentence θ such that

$$\mathfrak{M} \models \theta \iff (\mathrm{Dom}(\mathfrak{M}), (\psi_R^\mathfrak{M})_{R\in\tau}, (S^\mathfrak{M})_{S\in\mathcal{B}}) \models \phi$$

holds for all $\mathfrak{M} \in \mathrm{Str}_\mathcal{B}(\sigma)$.

Note that substituting relations defined by formulas in place of the built-in relations is not included in this definition. This is because of the very idea of built-in relations: they are thought as *constants* that are always present in the models considered.

For logics of the form $\mathrm{FO}_\mathcal{B}(\mathcal{Q})$, substitution of formulas in place of relations can be defined syntactically: let $\phi[(\psi_R/R)_{R\in\tau}]$ be the sentence obtained by replacing each occurrence $R(\boldsymbol{y})$ of R in ϕ by the formula $\psi_R(\boldsymbol{y})$. It straighforward to show that the the equivalence in (s) always holds for $\theta := \phi[(\psi_R/R)_{R\in\tau}]$. Thus, we have

Lemma 4.1. $\mathrm{FO}_\mathcal{B}(\mathcal{Q})$ *is semiregular for any set $\mathcal{B}$ of built-in relations and any class $\mathcal{Q}$ of $\mathcal{B}$-quantifiers.*

Any semiregular $\mathcal{B}$-logic $\mathcal{L}_\mathcal{B}$ is closed under quantification with respect to any $\mathcal{L}_\mathcal{B}$-definable quantifier. We prove next a useful formulatation of this principle.

Lemma 4.2. *Let $\mathcal{Q}$ be a class of $\mathcal{B}$-quantifiers, and let $\mathcal{L}_{\widetilde{\mathcal{B}}}$ be a semiregular $\widetilde{\mathcal{B}}$-logic. If all relations $S \in \mathcal{B}$ and all quantifiers $Q \in \mathcal{Q}$ are definable in $\mathcal{L}_{\widetilde{\mathcal{B}}}$, then* $\mathrm{FO}_\mathcal{B}(\mathcal{Q}) \le \mathcal{L}_{\widetilde{\mathcal{B}}}$.

Proof. Let $\mathcal{B}^+ = \mathcal{B} \cup \widetilde{\mathcal{B}}$. We prove by induction that for every formula $\phi \in \mathrm{FO}_\mathcal{B}(\mathcal{Q})(\tau)$ there is an equivalent formula $\theta \in \mathcal{L}_{\widetilde{\mathcal{B}}}(\tau)$, i.e., $\phi^\mathfrak{M} = \theta^\mathfrak{M}$, for every $\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}^+}(\tau)$. The claim is true for atomic formulas in the vocabulary τ by the definition of semiregularity; for atomic formulas $S(\boldsymbol{x})$ with $S \in \mathcal{B}$, we use the assumption that S is definable in $\mathcal{L}_{\widetilde{\mathcal{B}}}$. The induction steps corresponding to connectives and existential quantifier go through since, by semiregularity, $\mathcal{L}_{\widetilde{\mathcal{B}}}$ is closed with respect to Boolean operations and first order quantification.

Consider then the step corresponding to a quantifier $Q \in \mathcal{Q}$: let $\phi := Q\,(\boldsymbol{x}_R\psi_R)_{R\in\tau}$, and assume that for each ψ_R, $R \in \tau$, there is an equivalent formula $\eta_R \in \mathcal{L}_{\widetilde{\mathcal{B}}}$. By our assumption on $\mathcal{Q}$, there is a sentence $\chi \in \mathcal{L}_{\widetilde{\mathcal{B}}}(\tau)$ such that $K_{Q,\mathcal{B}^+} = \{\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}^+}(\tau) \mid \mathfrak{M} \models \chi\}$. Since $\mathcal{L}_{\widetilde{\mathcal{B}}}$ is closed under substitution, there is a sentence $\theta \in \mathcal{L}_{\widetilde{\mathcal{B}}}$ such that for every $\widetilde{\mathcal{B}}$-model $\mathfrak{N}$,

$$\mathfrak{N} \models \theta \iff (\mathrm{Dom}(\mathfrak{N}), (\eta_R^\mathfrak{N})_{R\in\tau}, (S^\mathfrak{N})_{S\in\widetilde{\mathcal{B}}}) \models \chi.$$

Clearly the same equivalence holds also for all $\mathcal{B}^+$-models. Let $\mathfrak{M}$ be a $\mathcal{B}^+$-model of vocabulary τ. As $\eta_R^{\mathfrak{M}} = \psi_R^{\mathfrak{M}}$ for each $R \in \tau$, $(\mathrm{Dom}(\mathfrak{M}), (\eta_R^{\mathfrak{M}})_{R\in\tau}, (S^{\mathfrak{M}})_{S\in\mathcal{B}^+}) \models \chi$ is equivalent to $\mathfrak{M} \models Q\,(\boldsymbol{x}_R\psi_R)_{R\in\tau}$. Thus we conclude that ϕ is equivalent with θ. $\square$

Without built-in relations, semiregularity of a logic can be characterized in terms of generalized quantifiers: $\mathcal{L}$ is semiregular if and only if there is a class $\mathcal{Q}$ of quantifiers such that $\mathcal{L} \equiv \mathrm{FO}(\mathcal{Q})$. This characterization remains valid in the framework of logics with built-in relations, once we replace ordinary quantifiers with the appropriate $\mathcal{B}$-quantifiers:

Proposition 4.3. *A logic $\mathcal{L}_{\mathcal{B}}$ with built-in relations is semiregular if and only if there is a class $\mathcal{Q}$ of $\mathcal{B}$-quantifiers such that $\mathcal{L}_{\mathcal{B}} \equiv \mathrm{FO}_{\mathcal{B}}(\mathcal{Q})$.*

Proof. We prove the implication from left to right; the other implication follows directly from Lemma 4.1. Thus, assume that $\mathcal{L}_{\mathcal{B}}$ is semiregular. We will show that $\mathcal{L}_{\mathcal{B}} \equiv \mathrm{FO}_{\mathcal{B}}(\mathcal{Q})$, where $\mathcal{Q}$ is the class of all $\mathcal{B}$-quantifiers which are definable in $\mathcal{L}_{\mathcal{B}}$. Note first that $\mathcal{L}_{\mathcal{B}} \leq \mathrm{FO}_{\mathcal{B}}(\mathcal{Q})$. Indeed, if K is class of $\mathcal{B}$-structures which is definable in $\mathcal{L}_{\mathcal{B}}$, then $K = K_Q$ for a quantifier $Q \in \mathcal{Q}$, whence K is trivially definable in $\mathrm{FO}_{\mathcal{B}}(\mathcal{Q})$. On the other hand, since all the relations in $\mathcal{B}$ and all the quantifiers in $\mathcal{Q}$ are definable in $\mathcal{L}_{\mathcal{B}}$, we have $\mathrm{FO}_{\mathcal{B}}(\mathcal{Q}) \leq \mathcal{L}_{\mathcal{B}}$ by Lemma 4.2. $\square$

Regular $\mathcal{B}$-logics

Let $\mathcal{B}$ be a set of arithmetical relations. We will assume in the sequel that $\mathcal{B}$ contains the usual order $\leq$ of natural numbers. Without this assumption the definition of regularity would become problematic.

We say that a $\mathcal{B}$-logic $\mathcal{L}_{\mathcal{B}}$ is *regular* if it is semiregular and it is closed under relativization. In the context of $\mathcal{B}$-logics, the definition of *relativization* $\mathfrak{M}|U$ of a model $\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}}(\tau)$ to a subset $U \subseteq \mathrm{Dom}(\mathfrak{M})$ has to be adapted a bit. The universe and the relations in τ are defined as usually: $\mathrm{Dom}(\mathfrak{M}|U) = U$, and $R^{\mathfrak{M}|U} = R^{\mathfrak{M}} \cap U^{\mathrm{ar}(R)}$ for each $R \in \tau$. However, the built-in relations $S^{\mathfrak{M}|U}$ cannot be defined in the same way as restrictions of $S^{\mathfrak{M}}$ to the set U: if U is not an initial segment of $\mathrm{Dom}(\mathfrak{M})$, then $S^{\mathfrak{M}} \cap U^{\mathrm{ar}(S)}$ is not necessarily isomorphic with $S \cap |U|^{\mathrm{ar}(S)}$, as should be the case by the definition of $\mathcal{B}$-models.

On the other hand, the restriction of the linear order $\leq^{\mathfrak{M}}$ of $\mathfrak{M}$ to the set U is a linear order of U, and hence it is isomorphic with $\leq \cap\, |U|^2$. Thus, it it natural to define $\leq^{\mathfrak{M}|U} = \leq^{\mathfrak{M}} \cap\, U^2$. The other built-in relations on $\mathfrak{M}|U$ are then defined in a canonical way: $S^{\mathfrak{M}|U} = S^f$, where $f : U \to |U|$ is the unique bijection such that $u \leq^{\mathfrak{M}|U} v \iff f(u) \leq f(v)$ for all $u, v \in U$.

With this definition, we say that the logic $\mathcal{L}_{\mathcal{B}}$ is *closed under relativization* if the following holds:

(r) If $\psi(x)$ is an $\mathcal{L}_{\mathcal{B}}(\tau)$-formula with one free variable, and ϕ is an $\mathcal{L}_{\mathcal{B}}(\tau)$-sentence, then there is an $\mathcal{L}_{\mathcal{B}}(\tau)$-sentence θ such that

$$\mathfrak{M} \models \theta \iff (\mathfrak{M}|\psi^{\mathfrak{M}}) \models \phi$$

holds for all $\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}}(\tau)$.

In the case of $\mathrm{FO}_{\leq}$, relativization can be defined syntactically in the same way as for FO without built-in relations. Thus, $\mathrm{FO}_{\leq}$ is regular. We will prove a more general result: any quantifier extension $\mathrm{FO}_{\leq}(\mathcal{Q})$ of $\mathrm{FO}_{\leq}$ is regular, provided that all the quantifiers $Q \in \mathcal{Q}$ admit relativization. A simple sufficient condition for this is universe independence. A $\mathcal{B}$-quantifier Q of vocabulary τ is *universe independent* if

$$\mathfrak{M} \in K_Q \iff \mathfrak{M}|U \in K_Q$$

whenever $\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}}(\tau)$ and $U \subseteq \mathrm{Dom}(\mathfrak{M})$ is such that $R^{\mathfrak{M}} \subseteq U^{\mathrm{ar}(R)}$ for each $R \in \tau$.

Proposition 4.4. $\mathrm{FO}_{\leq}(\mathcal{Q})$ *is regular for any class* $\mathcal{Q}$ *of universe independent* $\leq$*-quantifiers.*

Proof. The *relativization* $\phi|\psi$ of a formula $\phi \in \mathrm{FO}_{\leq}(\mathcal{Q})(\tau)$ with respect to a formula $\psi \in \mathrm{FO}_{\leq}(\mathcal{Q})(\tau)$ with one free variable is defined inductively as follows:

- $\phi|\psi := \phi$ for atomic τ-formulas ϕ,
- $(x \leq y)|\psi := x \leq y$,
- $(\neg\phi)|\psi := \neg\, \phi|\psi$,
- $(\phi \wedge \theta)|\psi := \phi|\psi \wedge \theta|\psi$,
- $(\exists y\, \phi)|\psi := \exists y\, (\psi(y) \wedge \phi|\psi)$
- $(Q\, (\boldsymbol{x}_R \eta_R)_{R\in\tau_Q})|\psi := Q\, (\boldsymbol{x}_R(\psi^{k_R}(\boldsymbol{x}_R) \wedge \eta_R|\psi))_{R\in\tau_Q}$ for $Q \in \mathcal{Q}$.

Here $k_R = \mathrm{ar}(R)$ and $\psi^{k_R}(\boldsymbol{x}_R)$ is the conjunction of $\psi(x_i)$ over all components x_i of $\boldsymbol{x}_R$. We prove by induction on ϕ that for all models $\mathfrak{M}$ and all tuples $\boldsymbol{a}$ of elements in $\psi^{\mathfrak{M}}$

$$\mathfrak{M} \models (\phi|\psi)[\boldsymbol{a}/\boldsymbol{x}] \iff (\mathfrak{M}|\psi^{\mathfrak{M}}) \models \phi[\boldsymbol{a}/\boldsymbol{x}].$$

In particular, if ϕ is a sentence, the equivalence in condition (r) holds for $\theta := \phi|\psi$.

The claim for atomic τ-formulas and the induction steps for connectives and existential quantifier are proved exactly as for FO without built-in relations. The claim for atomic formulas of the form $x \leq y$ follows from the fact that $\leq^{\mathfrak{M}|\psi^{\mathfrak{M}}}$ is the restriction of $\leq^{\mathfrak{M}}$ to the set $\psi^{\mathfrak{M}}$. Consider then the induction step for a quantifier $Q \in \mathcal{Q}$. Let $\mathfrak{M}$ and $\boldsymbol{a}$ be fixed, and assume that ϕ is of the form $Q\, (\boldsymbol{x}_R \eta_R)_{R\in\tau_Q}$. Then $\phi|\psi := Q\, (\boldsymbol{x}_R \chi_R)_{R\in\tau_Q}$, where χ_R is the formula $\psi^{k_R}(\boldsymbol{x}_R) \wedge \eta_R|\psi$ for each $R \in \tau_Q$. To simplify notation, we denote $\mathfrak{M}|\psi^{\mathfrak{M}}$ by $\mathfrak{N}$, and the expansions of $\mathfrak{M}$ and $\mathfrak{N}$ by the constants $\boldsymbol{a}$ by $\mathfrak{M}^+$ and $\mathfrak{N}^+$, respectively. By induction hypothesis, we have

$$\mathfrak{M} \models (\eta_R|\psi)[\boldsymbol{a}/\boldsymbol{x}, \boldsymbol{b}/\boldsymbol{x}_R] \iff \mathfrak{N} \models \eta_R[\boldsymbol{a}/\boldsymbol{x}, \boldsymbol{b}/\boldsymbol{x}_R]$$

for all $R \in \tau_Q$ and for all tuples $\boldsymbol{b}$ in $\psi^{\mathfrak{M}}$, and hence

$$\chi_R^{\mathfrak{M}^+} = (\psi^{\mathfrak{M}})^{k_R} \cap (\eta_R|\psi)^{\mathfrak{M}^+} = \eta_R^{\mathfrak{N}^+}$$

for all $R \in \tau_Q$. Now we get the following chain of equivalences:

$$\begin{aligned} \mathfrak{M} \models (\phi|\psi)[\boldsymbol{a}/\boldsymbol{x}] \quad &\iff \quad (\mathrm{Dom}(\mathfrak{M}), (\chi_R^{\mathfrak{M}^+})_{R\in\tau_Q}, \leq^{\mathfrak{M}}) \in K_Q \\ &\iff \quad (\psi^{\mathfrak{M}}, (\chi_R^{\mathfrak{M}^+})_{R\in\tau_Q}, \leq^{\mathfrak{N}}) \in K_Q \\ &\iff \quad \mathfrak{N} \models \phi[\boldsymbol{a}/\boldsymbol{x}]. \end{aligned}$$

Here the second equivalence is true since Q is universe independent, and clearly $\chi_R^{\mathfrak{M}^+} \subseteq (\psi^{\mathfrak{M}})^{k_R}$ for each $R \in \tau_Q$. □

Note that the proof of Proposition 4.4 cannot be extended to $\mathrm{FO}_{\mathcal{B}}(\mathcal{Q})$ if $\mathcal{B}$ contains an arithmetical relation S such that $S^{\mathfrak{M}|U} \neq S^{\mathfrak{M}} \cap U^{\mathrm{ar}(S)}$ for some model $\mathfrak{M}$ and some subset U of $\mathrm{Dom}(\mathfrak{M})$. In fact, as we shall see in Section 5, $\mathrm{FO}_{\mathcal{B}}$ is usually not regular if $\mathcal{B}$ contains such a relation.

Example 4.5. The logic $\mathrm{FO}_{\mathcal{B}}(\mathsf{I})$ is regular for any set $\mathcal{B}$ of built-in relations such that $\leq\, \in \mathcal{B}$. This follows from the fact that for any arithmetical relation $S \subseteq \mathbb{N}^k$, the corresponding universe independent quantifier B_S with defining class

$$K_{\mathsf{B}_S} = \{\mathfrak{M} \in \mathrm{Str}_{\leq}(\{P_0, \ldots, P_{k-1}\}) \mid (|P_0^{\mathfrak{M}}|, \ldots, |P_{k-1}^{\mathfrak{M}}|) \in S\}$$

is definable in $\mathrm{FO}_{\{\leq,S\}}(\mathsf{I})$ (see [Luo04]). Conversely, the built-in relation S is clearly definable in $\mathrm{FO}_{\leq}(\mathsf{I}, \mathsf{B}_S)$. Hence, in fact $\mathrm{FO}_{\mathcal{B}}(\mathsf{I}) \equiv \mathrm{FO}_{\leq}(\mathsf{I}, \mathcal{Q}_{\mathcal{B}})$, where $\mathcal{Q}_{\mathcal{B}}$ is the set of all quantifiers B_S for $S \in \mathcal{B}$.

5 Regular interior and closure of AC^0

There is a notable feature that separates the circuit complexity class AC^0 from TC^0: the former has a definite weakness against "padding" of input strings when computing natural cardinality properties. Indeed, the property of strings being of even length is trivially in AC^0, while, by the famous theorem of Ajtai [Ajt83] and Furst, Saxe and Sipser [FSS84], the property of binary strings of having an even number of 1's is not in AC^0. This weakness of AC^0 can be formulated in a precise way in terms of the logic $\mathrm{FO}_{\{+,\times\}}$ capturing it: $\mathrm{FO}_{\{+,\times\}}$ is not closed under relativization, and so it is not a regular $\{+,\times\}$-logic. On the other hand, the logics capturing TC^0, like $\mathrm{FO}_{\{+,\times\}}(\mathsf{I})$, are regular.

Thus, from a logical perspective, we can say that TC^0 is a better behaving class than AC^0. But regularity is also a very natural requirement from the computational point of view. Just note that a logic $\mathcal{L}_\mathcal{B}$ capturing a complexity class C is closed under substitution if and only if C is closed under composition of queries. Similarly, relativization property for $\mathcal{L}_\mathcal{B}$ translates to the requirement that C is closed under restricting C-computable queries to C-computable subsets of input structures.

In this section, we will study two ways of addressing the weakness of AC^0. The first one is to look for a largest possible fragment of AC^0, which is regular. This leads us to the notion of regular interior of a logic. The second alternative is to look for a minimal regular extension of AC^0. For this purpose, we adapt the notion of regular closure used in the area of Abstract Logic to the case of $\mathcal{B}$-logics.

Regular interior and regular closure

For any semiregular logic $\mathcal{L}_\mathcal{B}$ there is a largest regular logic that is contained in $\mathcal{L}_\mathcal{B}$. We call it the *regular interior* of $\mathcal{L}_\mathcal{B}$, and denote it by $\mathcal{R}\text{-int}(\mathcal{L}_\mathcal{B})$. Regular interior was introduced for logics without built-in relations in [Luo09].

The definition of regular interior is based on the notion of universe independence:

Definition 5.1. Let $\mathcal{L}_\mathcal{B}$ be a semiregular $\mathcal{B}$-logic. We set $\mathcal{R}\text{-int}(\mathcal{L}_\mathcal{B}) := \mathrm{FO}_\leq(\mathcal{Q}_u)$, where $\mathcal{Q}_u$ is the class of all universe independent $\mathcal{B}$-quantifiers which are definable in $\mathcal{L}_\mathcal{B}$.

Before showing that $\mathcal{R}\text{-int}(\mathcal{L}_\mathcal{B})$ has the desired properties, we introduce an auxiliary notion, and prove a couple of lemmas.

There is a canonical way of obtaining a universe independent quantifier from any given quantifier: the *regularization* Q^{reg} of a $\mathcal{B}$-quantifier Q is the $\mathcal{B}$-quantifier of vocabulary $\sigma = \tau_Q \cup \{P\}$ with P a new unary relation symbol, having defining class

$$K_{Q^{\mathrm{reg}}} = \{\mathfrak{M} \in \mathrm{Str}_\mathcal{B}(\sigma) \mid (\mathfrak{M}|P^\mathfrak{M}) \restriction \tau_Q \in K_Q\}.$$

Here $(\mathfrak{M}|P^\mathfrak{M}) \restriction \tau_Q$ is the reduct of $\mathfrak{M}|P^\mathfrak{M}$ to the vocabulary $\tau_Q \cup \sigma_\mathcal{B}$.

Lemma 5.2. *Q^{reg} is universe independent for any $\mathcal{B}$-quantifier Q.*

Proof. Assume that $\mathfrak{M}$ is a $\mathcal{B}$-model of vocabulary $\tau_Q \cup \{P\}$ and U is a subset of $\mathrm{Dom}(\mathfrak{M})$ such that $R^{\mathfrak{M}} \subseteq U^{\mathrm{ar}(R)}$ for all $R \in \tau_Q \cup \{P\}$. In particular $P^{\mathfrak{M}} \subseteq U$, whence clearly $(\mathfrak{M}|U)|P^{\mathfrak{M}} = \mathfrak{M}|P^{\mathfrak{M}}$, and so we have the chain of equivalences

$$\mathfrak{M} \in K_{Q^{\mathrm{reg}}} \iff (\mathfrak{M}|P^{\mathfrak{M}}) \restriction \tau_Q \in K_Q \iff ((\mathfrak{M}|U)|P^{\mathfrak{M}}) \restriction \tau_Q \in K_Q \iff \mathfrak{M}|U \in K_{Q^{\mathrm{reg}}}.$$

□

It is easy to see that Q is always definable in $\mathrm{FO}_{\leq}(Q^{\mathrm{reg}})$: indeed, for any $\mathcal{B}$-model $\mathfrak{M}$, we have $\mathfrak{M} \in K_Q \iff \mathfrak{M} \models Q^{\mathrm{reg}}(\boldsymbol{x}_R \psi_R)_{R \in \tau \cup \{P\}}$, where $\psi_R := R(\boldsymbol{x})$ for $R \in \tau_Q$ and $\psi_P := (x = x)$. The converse direction does not hold in general, but it becomes true in the context of a regular $\mathcal{B}$-logic.

Lemma 5.3. *Let $\mathcal{L}_{\mathcal{B}}$ be a regular $\mathcal{B}$-logic, and let Q be a B-quantifier. If Q is definable in $\mathcal{L}_{\mathcal{B}}$, then Q^{reg} is also definable in $\mathcal{L}_{\mathcal{B}}$.*

Proof. Assume that ϕ is a sentence in $\mathcal{L}_{\mathcal{B}}(\tau_Q)$ that defines the class K_Q. Since $\mathcal{L}_{\mathcal{B}}$ is closed under relativization, there is an $\mathcal{L}_{\mathcal{B}}$-sentence θ of vocabulary $\tau_Q \cup \{P\}$ such that the equivalence

$$\mathfrak{M} \models \theta \iff (\mathfrak{M}|P^{\mathfrak{M}}) \models \phi$$

holds for all $\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}}(\tau_Q \cup \{P\})$. Note further that

$$(\mathfrak{M}|P^{\mathfrak{M}}) \models \phi \iff (\mathfrak{M}|P^{\mathfrak{M}}) \restriction \tau_Q \in K_Q,$$

since P does not occur in the sentence ϕ. This means that θ defines the class $K_{Q^{\mathrm{reg}}}$. □

Now we are ready to prove that the definition of regular interior works as intended:

Proposition 5.4. *$\mathcal{R}$-int$(\mathcal{L}_{\mathcal{B}})$ is the largest sublogic of $\mathcal{L}_{\mathcal{B}}$ that is regular.*

Proof. Since each quantifier in $\mathcal{Q}_u$ and the linear order $\leq$ are definable in $\mathcal{L}_{\mathcal{B}}$, and $\mathcal{L}_{\mathcal{B}}$ is semiregular, it follows from Lemma 4.2 that $\mathcal{R}\text{-int}(\mathcal{L}_{\mathcal{B}}) = \mathrm{FO}_{\leq}(\mathcal{Q}_u) \leq \mathcal{L}_{\mathcal{B}}$. Moreover, $\mathcal{R}\text{-int}(\mathcal{L}_{\mathcal{B}})$ is regular by Proposition 4.4.

It remains to prove that $\mathcal{R}\text{-int}(\mathcal{L}_{\mathcal{B}})$ contains all regular sublogics of $\mathcal{L}_{\mathcal{B}}$. Thus, assume that $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}} \leq \mathcal{L}_{\mathcal{B}}$ is regular, and let Q be a $\widetilde{\mathcal{B}}$-quantifier which is definable in $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$. Since $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$ is regular, Q^{reg} is definable in $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$ by Lemma 5.3. Furhtermore, since $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}} \leq \mathcal{L}_{\mathcal{B}}$, Q^{reg} is definable in $\mathcal{L}_{\mathcal{B}}$ as well. By Lemma 5.2, the quantifier Q^{reg} is universe independent, whence it is in the class $\mathcal{Q}_u$. As observed above, Q is definable in $\mathrm{FO}_{\leq}(Q^{\mathrm{reg}})$, whence we conclude that Q is definable in $\mathcal{R}\text{-int}(\mathcal{L}_{\mathcal{B}})$. □

On the other hand, every semiregular $\mathcal{B}$-logic $\mathcal{L}_{\mathcal{B}}$ can be extended to a regular $\mathcal{B}$-logic. In fact, there is a least regular extension $\mathcal{R}\text{-cl}(\mathcal{L}_{\mathcal{B}})$ of $\mathcal{L}_{\mathcal{B}}$, which we call the *regular closure* of $\mathcal{L}_{\mathcal{B}}$. The definition of regular closure uses the notion of regularization of quantifiers:

Definition 5.5. Let $\mathcal{L}_{\mathcal{B}}$ be a $\mathcal{B}$-logic, and let $\mathcal{Q}$ be the class of all $\mathcal{B}$-quantifiers which are definable in $\mathcal{L}_{\mathcal{B}}$. We set $\mathcal{R}\text{-cl}(\mathcal{L}_{\mathcal{B}}) := \mathrm{FO}_{\leq}(\mathcal{Q}^{\mathrm{reg}})$, where $\mathcal{Q}^{\mathrm{reg}} = \{Q^{\mathrm{reg}} \mid Q \in \mathcal{Q}\}$.

Just like in the case of ordinary logics and generalized quantifiers, we can prove that $\mathcal{R}\text{-cl}(\mathcal{L}_\mathcal{B})$ is regular, and there is no regular $\mathcal{B}$-logic strictly in-between $\mathcal{L}_\mathcal{B}$ and $\mathcal{R}\text{-cl}(\mathcal{L}_\mathcal{B})$.

Proposition 5.6. *Let $\mathcal{L}_\mathcal{B}$ be a semiregular $\mathcal{B}$-logic. Then $\mathcal{R}\text{-cl}(\mathcal{L}_\mathcal{B})$ is the least extension of $\mathcal{L}_\mathcal{B}$ that is regular.*

Proof. Let $\mathcal{Q}$ be the class of $\mathcal{B}$-quantifiers which are definable in $\mathcal{L}_\mathcal{B}$. Since Q is definable in $\mathrm{FO}_\leq(Q^{\mathrm{reg}})$ for each $Q \in \mathcal{Q}$, it follows that $\mathcal{L}_\mathcal{B} \leq \mathrm{FO}_\leq(\mathcal{Q}^{\mathrm{reg}})$. Thus, $\mathcal{R}\text{-cl}(\mathcal{L}_\mathcal{B})$ is indeed an extension of $\mathcal{L}_\mathcal{B}$. Furthermore, since all the quantifiers in $\mathcal{Q}^{\mathrm{reg}}$ are universe independent, $\mathcal{R}\text{-cl}(\mathcal{L}_\mathcal{B})$ is regular by Proposition 4.4.

To complete the proof, we assume that $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$ is a regular $\widetilde{\mathcal{B}}$-logic such that $\mathcal{L}_\mathcal{B} \leq \widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$. Then each quantifier $Q \in \mathcal{Q}$ is definable in $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$. Since $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$ is regular, it follows from Lemma 5.3 that all quantifiers Q^{reg} in $\mathcal{Q}^{\mathrm{reg}}$ are also definable in $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$. Since $\widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$ is semiregular, and the order $\leq$ is definable in it, by Lemma 4.2 we have $\mathrm{FO}_\leq(\mathcal{Q}^{\mathrm{reg}}) \leq \widetilde{\mathcal{L}}_{\widetilde{\mathcal{B}}}$. Thus we conclude that $\mathcal{R}\text{-cl}(\mathcal{L}_\mathcal{B})$ is the smallest regular extension of $\mathcal{L}_\mathcal{B}$. □

Regular interior and Crane Beach Conjecture

The weakness of AC^0 that we discussed above has earlier inspired researchers to formulate the so-called *Crane Beach Conjecture*[1] (CBC) (see [BIL$^+$05]). The formulation of CBC is based on the notion of neutral letter. A symbol $e \in \Sigma$ is a *neutral letter* for a language $L \subseteq \Sigma^*$ if for all $u, v \in \Sigma^*$ it holds that $uv \in L \iff uev \in L$. In other words, e is a neutral letter for L if inserting or deleting any number of e's in a word does not affect its membership in L.

CBC is the statement that if a language with a neutral letter is definable in first-order logic with arbitrary built-in relations, then it is already definable in first-order logic with linear order as the only built-in relation. The general form of the conjecture was shown to be false in [BIL$^+$05]. However, the paper [BIL$^+$05] also provides some interesting restricted cases in which the conjecture is true.

To formulate these positive results, we will say that a set $\mathcal{B}$ of built-in relations has the *Neutral Letter Collapse Property*[2] (NLCP) with respect to a class $\mathcal{C}$ of languages, if the following holds for every language $L \in \mathcal{C}$ with a neutral letter:

If Q_L is definable in $\mathrm{FO}_\mathcal{B}$, then Q_L is already definable in $\mathrm{FO}_\leq$.

With this terminology, the relevant positive results from [BIL$^+$05] can be stated as follows:

Theorem 5.7. *([BIL$^+$05]). Let $\mathcal{U}$ be the set of all unary arithmetical relations together with the order $\leq$, and let $\mathcal{A}$ be the set of all arithmetical relations.*
(a) The set $\mathcal{U}$ has NLCP with respect to the class of all languages.
(b) The set $\{+\}$ has NLCP with respect to the class of all languages.
(c) The set $\mathcal{A}$ has NLCP with respect to the class of all languages in a binary alphabet.

[1] Named after the location of an attempt to prove it.

[2] This notion is similar to the Crane Beach Property formulated in [LTT06], but not equivalent.

We will next show that NLCP can be formulated in terms of the notion of regular interior. This is not surprising once we notice that the property of having a neutral letter is a language theoretic analogue for the property of a quantifier being universe independent. For the statement of the result, we need the following concept: given a language $L \subseteq \Sigma^*$ and a symbol $e \notin \Sigma$, define the *neutral letter extension* $N(L)$ of L to be the unique language in the alphabet $\Sigma \cup \{e\}$ such that $N(L) \cap \Sigma^* = L$ and e is a neutral letter for $N(L)$.

Lemma 5.8. *A set $\mathcal{B}$ of built-in relations has NLCP with respect to a class $\mathcal{C}$ of languages if and only if the implication*

$$Q_L \textit{ is definable in } \mathcal{R}\text{-int}(\mathrm{FO}_{\mathcal{B}}) \implies Q_L \textit{ is definable in } \mathrm{FO}_{\leq}$$

holds for every language L such that $N(L) \in \mathcal{C}$.

Proof. Assume that L is a language such that $N(L) \in \mathcal{C}$, and consider the corresponding language quantifiers Q_L and $Q_{N(L)}$. It is straightforward to show that $Q_{N(L)}$ is definable by the regularization Q_L^{reg} of Q_L, and vice versa. Moreover, by Lemma 5.3, for any regular $\mathcal{B}$-logic $\mathcal{L}_{\mathcal{B}}$, Q_L is definable in $\mathcal{L}_{\mathcal{B}}$ if and only if Q_L^{reg} is definable in $\mathcal{L}_{\mathcal{B}}$. In particular, this holds for the logics $\mathcal{R}$-int($\mathrm{FO}_{\mathcal{B}}$) and $\mathrm{FO}_{\leq}$. Finally, observe that since Q^{reg} is universe independent, it is definable in $\mathrm{FO}_{\mathcal{B}}$ if and only if it is definable in $\mathcal{R}$-int($\mathrm{FO}_{\mathcal{B}}$). The claim follows from these equivalences. □

Note that if $\mathcal{C}$ is the class of all languages, then $L \in \mathcal{C}$ if and only if $N(L) \in \mathcal{C}$. Thus, by Theorem 5.7 (a), if $\mathcal{U}$ is the set of all unary arithmetical relations together with the order, $\mathcal{R}$-int($\mathrm{FO}_{\mathcal{U}}$) collapses to $\mathrm{FO}_{\leq}$ if we consider only definability on word models. Similarly, by Theorem 5.7 (b), $\mathcal{R}$-int(FO_{+}) collapses to $\mathrm{FO}_{\leq}$ on word models. As a matter of fact, we can prove a much stronger result:

Theorem 5.9. *Let $\mathcal{U}$ be as in Proposition 5.7.*
(a) $\mathcal{R}$-int($\mathrm{FO}_{\mathcal{U}}$) $\equiv \mathrm{FO}_{\leq}$.
(b) $\mathcal{R}$-int(FO_{+}) $\equiv \mathrm{FO}_{\leq}$.

Proof. (Sketch) (a) The proof of Theorem 5.7 (a) (Theorem 4.1 in [BIL+05]) is based on an Ehrenfeucht-Fraïssé game argument. More precisely, it is shown there that if two word models $\mathfrak{M}$ and $\mathfrak{N}$ are equivalent in $\mathrm{FO}_{\leq}$ up to quantifier rank k, for large enough k, then for any finite set $\mathcal{B}$ of unary built-in relations, $\mathfrak{M}$ and $\mathfrak{N}$ can be embedded in larger word models $\mathfrak{M}^*$ and $\mathfrak{N}^*$ which are $\mathrm{FO}_{\{\leq\}\cup\mathcal{B}}$-equivalent up to a given quantifier rank r. Moreover, $\mathfrak{M}^*$ and $\mathfrak{N}^*$ are obtained by adding neutral letters in between the elements of $\mathfrak{M}$ and $\mathfrak{N}$, whence $\mathfrak{M}^*$ is in the language considered if and only if $\mathfrak{M}$ is, and the same holds for $\mathfrak{N}^*$ and $\mathfrak{N}$. A close inspection of the Ehrenfeucht-Fraïssé game used in the proof reveals that all the steps go through for arbitrary structures $\mathfrak{M}$ and $\mathfrak{N}$ in place of word models. In this way we see that any universe independent quantifier which is definable in $\mathrm{FO}_{\mathcal{U}}$ is already definable in $\mathrm{FO}_{\leq}$.

Clause (b) is proved in the same way as clause (a). In [Sch07], Schweikardt proves a similar transfer result for winning strategies in Ehrenfeucht-Fraïssé games showing that structures which are $\mathrm{FO}_{\leq}$-equivalent up to quantifier rank k can be embedded into structures which are FO_{+}-equivalent up to quantifier rank r. The result holds for arbitrary structures, whence we can again prove that any universe independent quantifier definable in FO_{+} is definable in $\mathrm{FO}_{\leq}$. $\square$

It is worth noting that while these results seem to indicate that the regular interior of AC^0 is quite weak, a counterexample for the Crane Beach Conjecture given in [BIL+05] shows that AC^0 does not entirely collapse to $\mathrm{FO}_{\leq}$. Indeed, the counterexample shows that there is a language L (in a ternary alphabet) such that Q_L is definable in $\mathcal{R}$-int($\mathrm{FO}_{\{+,\times\}}$), but not in $\mathrm{FO}_{\leq}$. In particular, it is not possible to generalize the third positive result in Theorem 5.7 in the same way as we did for the other two cases in Theorem 5.9.

Note however, that the counterexample quantifier Q_L is not order-invariant, whereas all the quantifiers corresponding to languages in a binary alphabet with a neutral letter are order-invariant; in fact, they are easily seen to be equivalent with cardinality quantifiers C_S. This raises the question, whether a correct generalization of Theorem 5.7 (c) would be that all order-invariant language quantifiers definable in $\mathcal{R}$-int($\mathrm{FO}_{\mathcal{A}}$) are already definable in $\mathrm{FO}_{\leq}$. In any case, we will show in the next subsection that order-invariance in itself is not the key for obtaining a collapse result.

Regular interior of AC^0

In this subsection we show that there is an order-invariant quantifier Q which is definable in $\mathcal{R}$-int($\mathrm{FO}_{\{+,\times\}}$) but not in $\mathrm{FO}_{\leq}$. For the definition of Q, we fix for each $n > 0$ a set A_n such that $|A_n| = n$ and $A_n \cap \mathcal{P}(A_n) = \emptyset$, and let $\mathfrak{A}_n \in \mathrm{Str}(\{E\})$ be the structure such that $\mathrm{Dom}(\mathfrak{A}_n) = A_n \cup (\mathcal{P}(A_n) \setminus \{\emptyset\})$ and $E^{\mathfrak{A}_n}$ is the membership relation between the sets A_n and $\mathcal{P}(A_n) \setminus \{\emptyset\}$.

Definition 5.10. We define Q to be the quantifier with defining class

$$K_Q = \{\mathfrak{M} \in \mathrm{Str}_{\leq}(\{E\}) \mid \exists n (n \in \mathrm{Sq} \setminus \{0\} \text{ and } (\mathfrak{M}|U) \restriction \{E\} \cong \mathfrak{A}_n)\}, \tag{6}$$

where $U = \mathrm{dom}(\mathrm{E}^{\mathfrak{M}}) \cup \mathrm{rg}(\mathrm{E}^{\mathfrak{M}})$.

We will start by showing that Q can be defined in $\mathcal{R}$-int($\mathrm{FO}_{\{+,\times\}}$). Note first that, for $\mathfrak{M} \in \mathrm{Str}_{\leq}(\{E\})$, the formulas $\delta(x) := \exists y E(x,y)$ and $\rho(x) := \exists y E(y,x)$ define the sets $\mathrm{dom}(\mathrm{E}^{\mathfrak{M}})$ and $\mathrm{rg}(E^{\mathfrak{M}})$ in $\mathfrak{M}$, respectively. Let χ be the conjunction of the following $\mathrm{FO}_{\leq}(\{E\})$-sentences:

- $\exists x\, \delta(x) \wedge \forall x(\delta(x) \rightarrow \neg\rho(x))$
- $\forall x \forall y((\rho(x) \wedge \rho(y) \wedge \forall z(E(z,x) \leftrightarrow E(z,y))) \rightarrow x = y)$
- $\forall x \forall y(\delta(x) \rightarrow \exists z \forall w(E(w,z) \leftrightarrow (w = x \vee E(w,y)))$.

It is straightforward to verify that the equivalence

$$\mathfrak{M} \models \chi \iff (\mathfrak{M}|U) \restriction \{E\} \cong \mathfrak{A}_n \text{ for some } n > 0$$

holds for all $\mathfrak{M} \in \mathrm{Str}_{\leq}(\{E\})$. We still need to express the condition $n \in \mathrm{Sq}$ in the definition of K_Q. To do this, we apply the so-called polylogarithmic counting ability of $\mathrm{FO}_{\{+,\times\}}$:

Theorem 5.11. *[ABO84, FKPS85, DGS86, WWY92] The logic* $\mathrm{FO}_{\{+,\times\}}$ *can count up to* lb^k *for any* $k \in \mathbb{N}$*. In other words, for every* k*, there is a formula* $\sigma_k(x) \in \mathrm{FO}_{\{+,\times\}}(\{P\})$*, where* P *is unary, such that for all* $\mathfrak{M} \in \mathrm{Str}_{\{+,\times\}}(\{P\})$

$$\mathfrak{M} \models \sigma_k[a/x] \iff |\{b \in \mathrm{Dom}(\mathfrak{M}) \mid b <^{\mathfrak{M}} a\}| = |P| \leq \mathrm{lb}^k\, n,$$

where $n = |\operatorname{Dom}(\mathfrak{M})|$.

Let $\theta(x)$ be the $\mathrm{FO}_{\{+,\times\}}(\{E\})$-formula obtained from $\sigma_1(x)$ by substituting the formula δ in place of the relation symbol P. Furthermore, let ψ be the $\mathrm{FO}_{\{+,\times\}}(\{E\})$-sentence $\exists x(\theta(x) \wedge \exists y(x = y \times y))$. By Theorem 5.11, $\mathfrak{M} \models \psi$ if and only if $|\mathrm{dom}(\mathrm{E}^{\mathfrak{M}})| \in \mathrm{Sq}$ and $|\mathrm{dom}(\mathrm{E}^{\mathfrak{M}})| \leq \mathrm{lb}(|\operatorname{Dom}(\mathfrak{M})|)$. Note that the latter condition is automatically true if $\mathfrak{M} \models \chi$. Thus, we conclude that the quantifier Q is defined by the $\mathrm{FO}_{\{+,\times\}}(\{E\})$-sentence $\chi \wedge \psi$. Finally, observe that the quantifier Q is universe independent, whence it is contained in $\mathcal{R}\text{-int}(\mathrm{FO}_{\{+,\times\}})$.

Next we show that Q is not definable in $\mathrm{FO}_{\leq}$. Towards a contradiction, assume that Q can be defined in $\mathrm{FO}_{\leq}$ by a sentence η. We will show that then the language $L = \{w \in \{a\}^* \mid |w| \in \mathrm{Sq}\}$ is definable in $\mathrm{MSO}_{\leq}$ contradicting the fact that all MSO-definable languages are regular and L is not.

For each $n > 0$, let A_n and $\mathfrak{A}_n$ be as above, and let $\mathfrak{M}_n$ be the $\mathrm{Str}_{\leq}(\{E\})$-structure such that $\mathrm{Dom}(\mathfrak{M}_n) = \mathrm{Dom}(\mathfrak{A}_n)$, $E^{\mathfrak{M}_n} = E^{\mathfrak{A}_n}$ and the ordering $\leq^{\mathfrak{M}_n}$ satisfies the condition:

- $\leq^{\mathfrak{M}_n} \cap (\mathcal{P}(A_n) \setminus \{\emptyset\})^2$ is the lexicographic ordering of subsets of A_n induced by $\leq^{\mathfrak{M}_n} \cap A_n^2$,
- $a \leq^{\mathfrak{M}_n} b$ for all $a \in A_n$ and $b \in (\mathcal{P}(A_n) \setminus \{\emptyset\})$.

Furthermore, for each $n > 0$, we let $\mathfrak{N}_n \in \mathrm{Str}_{\leq}\{P_a\}$ be the word model of length n with $\mathrm{Dom}(\mathfrak{N}_n) = P_a^{\mathfrak{N}_n} = A_n$ and $\leq^{\mathfrak{N}_n} = \leq^{\mathfrak{M}_n} \cap A_n^2$.

We will show that any $\mathrm{FO}_{\leq}(\{E\})$-sentence ϕ can be translated into a sentence $\phi^* \in \mathrm{MSO}_{\leq}(\{\mathrm{P_a}\})$ such that for all n

$$\mathfrak{M}_n \models \phi \iff \mathfrak{N}_n \models \phi^*.$$

By the assumption, we have $\mathfrak{M}_n \models \eta$ if and only if $|A_n| \in \mathrm{Sq}$. Therefore, η^* will then define the language $L = \{w \in \{a\}^* \mid |w| \in \mathrm{Sq}\}$. This will be the desired contradiction.

The idea of the translation $\phi \mapsto \phi^*$ is simple: each element $a \in \mathrm{Dom}(\mathfrak{M}_n) \setminus \mathrm{Dom}(\mathfrak{N}_n)$ is also a subset of $\mathrm{Dom}(\mathfrak{N}_n)$. Thus, first-order variables and quantification over elements in $\mathrm{Dom}(\mathfrak{M}_n) \setminus \mathrm{Dom}(\mathfrak{N}_n)$ can be replaced by monadic second-order variables and quantification.

We will now describe the technical details of the translation. First, we assign to each first-order variable x a corresponding monadic second-order variable X. For translating formulas with free variables, we need to keep track of those variables which should be translated into the corresponding second-order variables. Thus, we define a translation $T_S : \mathrm{FO}_{\leq}(\{E\}) \to \mathrm{FO}_{\leq}(\{P_a\})$ for each set S of first-order variables by simultaneous induction:

$$
\begin{aligned}
T_S(x = y) &:= \begin{cases} x = y & \text{if } x, y \notin S \\ \forall z(X(z) \leftrightarrow Y(z)) & \text{if } x, y \in S \\ x \neq x & \text{otherwise} \end{cases} \\
T_S(x \leq y) &:= \begin{cases} x \leq y & \text{if } x, y \notin S \\ x = x & \text{if } x \notin S,\ y \in S \\ x \neq x & \text{if } x \in S,\ y \notin S \\ X \leq Y & \text{if } x, y \in S \end{cases} \\
T_S(E(x, y)) &:= \begin{cases} Y(x) & \text{if } x \notin S,\ y \in S \\ x \neq x & \text{otherwise} \end{cases} \\
T_S(\neg\phi) &:= \neg T_S(\phi) \\
T_S(\phi \wedge \psi) &:= T_S(\phi) \wedge T_S(\psi) \\
T_S(\exists x\, \phi) &:= \exists x\, T_S(\psi) \vee \exists X\, T_{S \cup \{x\}}(\psi)
\end{aligned}
$$

Above, $X \leq Y$ denotes the formula which defines the lexicographic ordering of subsets induced by $\leq$.

By induction on the construction of $\phi \in \mathrm{FO}_{\leq}(\{E\})$ one can now prove that, for all $n > 0$, and $\boldsymbol{a} = (a_1, \ldots, a_k) \in A_n^k)$ and $\boldsymbol{b} = (b_1, \ldots, b_l) \in (\mathcal{P}(A_n) \setminus \{\emptyset\})^l$ it holds that

$$\mathfrak{M}_n \models \phi[\boldsymbol{a}/\boldsymbol{x}, \boldsymbol{b}/\boldsymbol{y}] \iff \mathfrak{N}_n \models T_S(\phi)[\boldsymbol{a}/\boldsymbol{x}, \boldsymbol{b}/\boldsymbol{Y}]$$

whenever $x_i \notin S$ for each component x_i of $\boldsymbol{x}$ and $y_j \in S$ for each component y_j of $\boldsymbol{y}$. In particular, if ϕ is a sentence, we have $\mathfrak{M}_n \models \phi \iff \mathfrak{N}_n \models T_\emptyset(\phi)$. Thus, defining $\phi^* := T_\emptyset(\phi)$ we get the desired translation.

We may now define $\Theta^* := \Theta' \wedge \forall x P_a x$. By the above, Θ^* defines the language $L = \{a^n \mid n \in \mathrm{Sq}\}$, which is a contradiction. We may conclude that Q is not definable in $\mathrm{FO}_{\leq}$.

Regular closure of AC^0

The gap between $\mathcal{R}$-int($\mathrm{FO}_{\mathcal{B}}$) and $\mathcal{R}$-cl($\mathrm{FO}_{\mathcal{B}}$) can be seen as a measure for the irregularity of $\mathrm{FO}_{\mathcal{B}}$: the larger the gap is, the more irregular $\mathrm{FO}_{\mathcal{B}}$ is. We will next show that in the case of $\mathcal{B}$ consisting of a suitable unary relation and the order $\leq$, this gap is extremely large. We have already seen that for such a $\mathcal{B}$, $\mathcal{R}$-int$(\mathrm{FO}_{\mathcal{B}}) \equiv \mathrm{FO}_{\leq}$. For the other direction, we have

Theorem 5.12. *Let $\mathcal{B}$ a set of built-in relations such that $\mathcal{B}$ contains a pseudoloose set S. Then $\mathcal{R}$-cl$(\mathrm{FO}_{\mathcal{B}}) \equiv \mathrm{FO}_{\mathcal{B}}(\mathsf{I})$. Moreover, $\mathrm{TC}^0 \leq \mathcal{R}$-cl$(\mathrm{FO}_{\mathcal{B}})$.*

Proof. Let Q be the $\mathcal{B}$-quantifier with empty vocabulary such that $K_Q = \{\mathfrak{M} \in \mathrm{Str}_{\mathcal{B}}(\emptyset) \mid |\,\mathrm{Dom}(\mathfrak{M})| - 1 \in S\}$. Then $\mathfrak{M} \models Q$ if and only if max $\in \mathrm{rg}(P)$, where max is the largest element of $\mathfrak{M}$ with respect to $\leq^{\mathfrak{M}}$. Thus, Q is definable in $\mathrm{FO}_{\mathcal{B}}$, and so Q^{reg} is definable in $\mathcal{R}$-cl$(\mathrm{FO}_{\mathcal{B}})$. It is easy to see that the cardinality quantifier C_S is definable by Q^{reg}, whence by Propostion 3.6, the Härtig quantifier I is definable in $\mathcal{R}$-cl$(\mathrm{FO}_{\mathcal{B}})$. This shows that $\mathrm{FO}_{\mathcal{B}}(\mathsf{I}) \leq \mathcal{R}$-cl$(\mathrm{FO}_{\mathcal{B}})$. On the other hand, it is easy to show that $\mathrm{FO}_{\mathcal{B}}(\mathsf{I})$ is always regular, whence it necessarily contains the least regular extension $\mathcal{R}$-cl$(\mathrm{FO}_{\mathcal{B}})$ of $\mathrm{FO}_{\mathcal{B}}(\mathsf{I})$. Finally, since C_S is definable in $\mathcal{R}$-cl$(\mathrm{FO}_{\mathcal{B}})$, the second claim follows directly from Theorem 3.7. □

As an immediate corollary, we get that the regular closure of $\mathrm{FO}_{\{+,\times\}}$ is equal to TC^0. This is because for example the pseudoloose set $\mathrm{Sq} = \{n^2 \mid n \in \mathbb{N}\}$ is definable in $\mathrm{FO}_{\{+,\times\}}$. We formulate the result in terms of the circuit complexity classes:

Corollary 5.13. $\mathcal{R}$-cl$(\mathrm{AC}^0) \equiv \mathrm{TC}^0$. □

We will close this section by some examples illustrating the gap between regular interior and regular closure.

Example 5.14. (a) Let $S = \mathrm{rg}(P)$ for some polynomial with integer coefficients and degree at least 2. Then by Theorem 5.9 (a), $\mathcal{R}$-int$(\mathrm{FO}_{\{\leq,S\}}) \equiv \mathrm{FO}_{\leq}$, and by Theorem 5.12, $\mathcal{R}$-cl$(\mathrm{FO}_{\{\leq,S\}}) \equiv \mathrm{FO}_{\{+,\times\}}(\mathsf{I}) \equiv \mathrm{TC}^0$.

(b) By Theorem 5.9 (b), $\mathcal{R}$-int$(\mathrm{FO}_{+}) \equiv \mathrm{FO}_{\leq}$. On the other hand, it is not too difficult to see that $\mathcal{R}$-cl$(\mathrm{FO}_{+}) \equiv \mathrm{FO}_{\leq}(\mathsf{I})$.

(c) Let $S_n = n\mathbb{N} = \{nk \mid k \in \mathbb{N}\}$ for each $n \in \mathbb{N}$. Again, by Theorem 5.9 (a), $\mathcal{R}$-int$(\mathrm{FO}_{\{\leq,S_n\}}) \equiv \mathrm{FO}_{\leq}$. For the other direction we can prove that $\mathcal{R}$-cl$(\mathrm{FO}_{\{\leq,S_n\}}) \equiv \mathrm{FO}_{\leq}(\mathsf{D}_n)$.

Conclusion

We conclude the paper with open questions for future research. Our first open question concerns the set $E = \{2^n \mid n \in \mathbb{N}\}$. Note that E is not pseudoloose but by b) of Lemma 2.4 it holds that

$$\mathrm{FO}_{\leq}(\mathsf{C}_E) \equiv \mathrm{FO}_{+}(\mathsf{I}, \mathsf{C}_E) \leq \mathrm{FO}_{\{+,\times\}}(\mathsf{Maj}).$$

It is an open question whether the latter two logics are equivalent. More generally, we ask if there is $S \subseteq \mathbb{N}$ such that

$$\mathrm{FO}_{\leq}(\mathsf{I}) < \mathrm{FO}_{\leq}(\mathsf{C}_S) < \mathrm{FO}_{\leq}(\mathsf{D}) \equiv \mathrm{FO}_{\{+,\times\}}(\mathsf{Maj}).$$

The last question concerns the counterexample [BIL$^+$05] for the Crane Beach Conjecture showing that the regular interior of AC^0 does not entirely collapse to $\mathrm{FO}_{\leq}$. This counterexample is a language which is not order-invariant. In fact, as far as we know, it is an open question whether the order-invariant version of NLCP holds for AC^0 with respect to the class of all languages, i.e., is every order-invariant $\mathrm{FO}_{\{+,\times\}}$-definable language with a neutral letter already definable in $\mathrm{FO}_{\leq}$. This question can be equivalently formulated as follows: is every $\mathrm{FO}_{\{+,\times\}}$-definable unary quantifier Q already definable in FO.

References

[ABO84] Miklos Ajtai and Michael Ben-Or. A theorem on probabilistic constant depth computations. In *STOC '84: Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 471–474, New York, NY, USA, 1984. ACM.

[Ajt83] Miklós Ajtai. Σ_1^1 formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.

[BIL+05] David A. Mix Barrington, Neil Immerman, Clemens Lautemann, Nicole Schweikardt, and Denis Thérien. First-order expressibility of languages with neutral letters or: The Crane Beach conjecture. *J. Comput. System Sci.*, 70(2):101–127, 2005.

[BIS90] David A. Mix Barrington, Neil Immerman, and Howard Straubing. On uniformity within NC^1. *Journal of Computer and System Sciences*, 41:274–306, 1990.

[DDLW98] Anuj Dawar, Kees Doets, Steven Lindell, and Scott Weinstein. Elementary properties of the finite ranks. *MLQ Math. Log. Q.*, 44(3):349–353, 1998.

[DGS86] Larry Denenberg, Yuri Gurevich, and Saharon Shelah. Definability by constant-depth polynomial-size circuits. *Information and Control*, 70(2/3):216–240, 1986.

[Ebb85] Heinz-Dieter Ebbinghaus. Extended logics: the general framework. In *Model-theoretic logics*, Perspect. Math. Logic, pages 25–76. Springer, New York, 1985.

[FKPS85] Ronald Fagin, Maria M. Klawe, Nicholas J. Pippenger, and Larry Stockmeyer. Bounded-depth, polynomial-size circuits for symmetric functions. *Theor. Comput. Sci.*, 36(2-3):239–250, 1985.

[FSS84] Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17:13–27, 1984.

[HAB02] William Hesse, Eric Allender, and David A. Mix Barrington. Uniform constant-depth threshold circuits for division and iterated multiplication. *J. Comput. System Sci.*, 65(4):695–716, 2002.

[Här65] Klaus Härtig. Über einen Quantifikator mit zwei Wirkungsbereichen. In *Colloq. Found. Math., Math. Machines and Appl. (Tihany, 1962)*, pages 31–36. Akad. Kiadó, Budapest, 1965.

[IL90] Neil Immerman and Eric Lander. Describing graphs: a first-order approach to graph canonization. In *Complexity theory retrospective*, pages 59–81. Springer, New York, 1990.

[Imm99] Neil Immerman. *Descriptive Complexity*. Graduate Texts in Computer Science. Springer Verlag, New York, 1999.

[KL79] Michał Krynicki and Alistair H. Lachlan. On the semantics of the Henkin quantifier. *J. Symbolic Logic*, 44(2):184–200, 1979.

[KV95] Phokion G. Kolaitis and Jouko A. Väänänen. Generalized quantifiers and pebble games on finite structures. *Ann. Pure Appl. Logic*, 74(1):23–75, 1995.

[Lin66] Per Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195, 1966.

[Lin95] Steven Lindell. Monadic counting does not suffice. Manuscript, 1995.

[LMSV01] Clemens Lautemann, Pierre McKenzie, Thomas Schwentick, and Heribert Vollmer. The descriptive complexity approach to LOGCFL. *Journal of Computer and Systems Sciences*, 62(4):629–652, 2001.

[LTT06] Clemens Lautemann, Pascal Tesson, and Denis Thérien. An algebraic point of view on the Crane Beach property. In *Computer science logic*, volume 4207 of *Lecture Notes in Comput. Sci.*, pages 426–440. Springer, Berlin, 2006.

[Luo04] Kerkko Luosto. Equicardinality on linear orders. In *Proc. 19th IEEE Symp. on Logic in Computer Science*, pages 458–465, 2004.

[Luo09] Kerkko Luosto. Classifying unary quantifiers. Preprint 496, Department of Mathematics and Statistics, University of Helsinki, 2009.

[Lyn82] James F. Lynch. Complexity classes and theories of finite models. *Mathematical Systems Theory*, 15(2):127–144, 1982.

[Mos57] Andrzej Mostowski. On a generalization of quantifiers. *Fund. Math.*, 44:12–36, 1957.

[Nur00] Juha Nurmonen. Counting modulo quantifiers on finite structures. *Inform. and Comput.*, 160(1-2):62–87, 2000.

[Ruh99] Matthias Ruhl. Counting and addition cannot express deterministic transitive closure. In *Proc. 14th IEEE Symp. on Logic in Computer Science*, pages 326–334. 1999.

[Sch05] Nicole Schweikardt. Arithmetic, first-order logic, and counting quantifiers. *ACM Trans. Comput. Log.*, 6(3):634–671, 2005.

[Sch07] Nicole Schweikardt. An ehrenfeucht-fraïssé game approach to collapse results in database theory. *Inf. Comput.*, 205(3):311–379, 2007.

[Vol99] Heribert Vollmer. *Introduction to Circuit Complexity – A Uniform Approach*. Texts in Theoretical Computer Science. Springer Verlag, Berlin Heidelberg, 1999.

[WWY92] Ingo Wegener, Norbert Wurm, and Sang-Zin Yi. Symmetric functions in AC^0 can be computed in constant depth with very small size. In *Poceedings of the London Mathematical Society symposium on Boolean function complexity*, pages 129–139, New York, NY, USA, 1992. Cambridge University Press.