Dot-Depth and Monadic Quantifier Alternation over Pictures

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
genehmigte Dissertation

vorgelegt von

Diplom-Informatiker Oliver Matz
aus Neumünster

Berichter: Universitätsprofessor Dr. W. Thomas
Universitätsprofessor Dr. E. Grädel

Tag der mündlichen Prüfung: 14.06.1999

D 82 (Diss. RWTH Aachen)

Contents

Chapter 1

Introduction

The interest of computer scientists in monadic second-order logic has two reasons. Firstly, monadic second-order is a very powerful specification language which subsumes other specification languages (such as temporal logic) in expressive power. In order to achieve progress in automated verification, it might be useful to learn more about the complexity of monadic second-order logic, in particular the feasibility of the alternation of existential and universal quantifications over sets.

The hierarchy of properties expressible with an increasing number of these alternations is called "monadic second-order alternation hierarchy", in contrast to the alternation hierarchy of full second-order logic, where quantifications range over relations of arbitrary arity.

The second of the mentioned reasons is the connection to complexity theory. Ebbinghaus and Flum write in the context also of the monadic second-order alternation hierarchy:

> The question to what extent the [different versions of second-order alternation] hierarchies are proper in the finite is related to important questions of complexity theory. ([EF95])

The questions that the authors allude to deal with the polynomial hierarchy, which by [Fag74] coincides with the full second-order alternation hierarchy. It is a famous open question whether this hierarchy is *strict*, i.e, whether each level allows us to compute or express more properties than the previous one. It is even conceivable that the polynomial hierarchy collapses to its first level, NP, but only if NP is closed under complement.

Studying the full second-order alternation hierarchy is a difficult task, but we can at least investigate what happens when we restrict ourselves to important fragments like the monadic one. Fagin et al. write:

> One way of attacking the difficult question whether NP equals co-NP is to restrict the classes under consideration. [...] The hope is that the restriction to monadic classes will yield more tractable questions and will serve as a training ground for attacking the problems in their full generality. ([FSV95, AFS97])

Some justification for this point of view is given by the following fact, which is claimed to be folklore in [MP96]. It underlines the (in terms of polynomial-time computations) large expressive power of the monadic fragment of second-order logic.

Fact 1.1 ([MP96, AFS98], see also [Sch97a]) *For every level of the polynomial hierarchy there is a problem complete (via polynomial-time reductions) for that level and definable on the same level of the monadic alternation hierarchy.*

Fagin has shown in [Fag75] that the first level of the monadic alternation hierarchy is not closed under complement, thus answering the "monadic analogue" of the question whether NP equals co-NP.

In this thesis we will show that the monadic second-order alternation hierarchy is strict, thus answering the monadic analogue to the question whether the polynomial hierarchy is strict. Moreover, we show that even the alternation of first-order quantifications allows us to express properties beyond any level of the monadic alternation hierarchy.

It turns out that the study of the above problems, which are located in finite model theory, is related to certain questions in formal language theory. Whereas the classical theory of languages concentrates on sets of words, i.e., one-dimensional, usually finite sequences, it is a natural task to investigate what happens if words are replaced by two-dimensional matrices, usually called "pictures". Such a two-dimensional language is called a picture language.

Some questions in this area are of the form: "Which results of the theory of (regular) word languages can be saved in the two-dimensional case?". Most of the answers are negative, i.e., they show that the situation is substantially more difficult for two dimensions. We add another one to these negative answers by showing that there is a starfree picture language that

is not recognizable, although both "starfreeness" and "recognizability" are straightforward adaptions of classical word language theory, where this cannot happen.

The relation to the monadic alternation hierarchy is the following. A picture language is recognizable iff it is definable in the existential fragment of monadic logic, i.e., level 1 of its alternation hierarchy. Our proof for the strictness of this hierarchy yields as a by-product that beyond any of its levels, there is a starfree picture language.

1.1 Alternation as a Complexity Measure

In descriptive complexity theory a general task is to classify properties by the complexity that a formula of a particular logic must have if it describes this property. Examples of "logics" (in a broad sense) where this question is studied are different fragments of the first-order and higher-order logics as well as variants of modal logic like the modal μ-calculus or temporal logic.

But what is a suitable measure for the complexity of a formula? In each of the above logics, a promising approach is to consider the most "powerful" ingredient(s) of the logic and count the number of nestings or alternations. Let us consider some examples of this in detail.

First-Order Logic. When the considered logic is first-order logic, two "good" measures are firstly the nestings of first-order quantifications (the "quantifier depth") and secondly the number of quantifier blocks of a prenex normal form that are either existential or universal ("alternation depth"). As an example, consider a formula of the form $\exists x_1 \exists x_2 \forall y \psi$, where the kernel formula ψ is quantifier-free. The quantifier depth of this formula is 3, whereas the alternation depth is 2. A more refined analysis of the alternation depth is possible: If the first quantifier block of a formula (say in prenex normal form) of alternation depth k is an existential one, i.e., if the formula is of the form

$$\exists \overline{x}_1 \forall \overline{x}_2 \ldots \eth \overline{x}_k \psi$$

for a quantifier free kernel formula ψ and first-order variable tuples $\overline{x}_1, \ldots, \overline{x}_k$ of possibly different length and $\eth \in \{\exists, \forall\}$, depending on whether k is odd or even, then such a formula will be called a Σ^0_k*-formula*. (The superscript 0 indicates the reference to first-order logic.) If, to the contrary, the first quantifier block is universal, then a formula of alternation depth k is called a Π^0_k*-formula.*

By the term "alternation hierarchy" we refer to the sequence of classes of (properties definable by) formulas of increasing alternation depth.

Modal μ-Calculus. In the modal μ-calculus, the most powerful operations are the two dual fixpoint operators μ and ν, thus an analogous approach is to count the number of their alternations. (However, this turns out to be a bit more difficult than in the case of, say, first-order formulas because there is no "prenex normal form" for the modal μ-calculus and thus there are different ways to count the number of alternations. This gives rise to at least three different measures, namely the syntactic, the Emerson-Lei, and the Niwiński alternation depth. See [Bra96].)

Second-Order Logic. When we consider (fragments of) second-order logic, where quantifications are allowed to range over relations (of some fixed arity), then a typical formula of the k-th level is of the form

$$\exists\overline{X}_1\forall\overline{X}_2\dots\mho\overline{X}_k\psi, \tag{1.1}$$

where now ψ is first-order and $\overline{X}_1,\dots,\overline{X}_k$ are tuples (of possibly different length) of relation variables (of possibly different arity). Note that when we adopt this measure, we neglect the first-order quantifications in the kernel formula in order to concentrate on second-order quantifiers, the most powerful ingredient of second-order logic. Again, one may distinguish between Σ^1_k- and Π^1_k-formulas, depending on whether the first quantifier block is existential or universal. (The superscript 1 indicates the reference to second-order logic.)

Let us have a closer look at the quantifier alternation hierarchy of full second-order logic. One observes that it is even possible to neglect first-order quantifications if they are outside the kernel formula. For example, a formula $\forall x\exists X\psi$, where x is a first-order variable, X is a second-order variable of arity r, and ψ is a first-order formula, is equivalent to a formula of the form $\exists X'\forall x\psi'$, where X' is of arity $r+1$. The proof idea is to modify the kernel formula ψ so that ψ' asserts for every $(r+1)$-ary relation X' and every x the same as ψ asserts for the r-ary relation $\{\overline{y} \mid (x,\overline{y}) \in X'\}$ and x.

Repeating this trick one may show that for every k, the set of Σ^1_k-formulas is closed under first-order quantifications, making this a robust class.

[Fag74] shows that each level (except for zero) of the full second-order alternation hierarchy coincides with the corresponding level of the polynomial hierarchy in expressive power, which establishes a link between computational complexity theory and descriptive complexity theory. It is one of the most famous open questions of (descriptive) complexity theory whether this hierarchy is infinite or, to the contrary, the number of second-order quantifier alternations needed to express second-order properties can be limited to some fixed number.

A result of [MP94] says that in full second-order logic, it is not possible to restrict the arity of the second-order variables *and* the alternation depth at the same time without loss of expressiveness.

Monadic Second-Order Logic. This is the fragment of second-order logic that this thesis is chiefly concerned with. The Σ_k-formulas of the alternation hierarchy of this fragment are of the form (1.1), where this time $\overline{X}_1,\dots,\overline{X}_k$ are tuples (of possibly different length) of set variables, i.e., ranging over unary relations, and ψ is first-order.

The central result of this thesis (first published in [MT97]) states that this hierarchy is strict, i.e., the number of monadic second-order quantifier alternations cannot be limited to some fixed number without loss of expressiveness.

What had been known about this question before? It has been shown that a collapse of this hierarchy would have implied a collapse of the full second-order alternations hierarchy because—as mentioned as Fact 1.1—each level of the monadic second-order hierarchy allows to define properties that are complete (wrt. polynomial-time reducibility) for the corresponding level of the polynomial hierarchy. However, without complexity-theoretic assumptions, the only known separation was that Σ_1 and Π_1 differ in expressiveness.

One drawback that limits the impact of our main result, the strictness of the monadic alternation hierarchy, is the observation (published e.g. in [AFS97]) that the levels of this hierarchy are not as robust as those of the full second-order hierarchy with respect to first-order quantifications. Recall that the "trick" of the previous paragraph (for the full second-order logic) involved passing from an r-ary variable to an $(r+1)$-ary one. If the arity is restricted to 1 like in monadic second-order logic, this is not possible. And indeed, an inspection of our hierarchy proof shows that it is the unbounded alternation of first-order quantifications followed by one monadic quantifier block which allows to define properties that cannot be captured by a bounded number of monadic second-order alternations. We will investigate the effect of first-order alternations in more detail.

Recall that for the first-order fragment there is a natural measure for the complexity of formulas that is different from the alternation depth, namely the quantifier depth. This measure has an analogue in the (monadic) second-order logic, too. Chapter 6 deals with the hierarchy questions concerned with this measure.

Other Fragments of Second-Order Logic. Of course, it is straightforward to consider other "bounded arity" fragments of full second-order logic, i.e., where quantification is allowed only over relations up to some fixed arity. However, even for the "binary" fragment, where quantifications range over binary relations, we do not have any hierarchy results, neither for the alternation hierarchy nor for the quantifier depth hierarchy. A possible approach to get at least partial results is to consider fragments of this binary fragment, for example those in which quantifications range over functions or, more restrictive, over permutations. For these classes, some separation results have been proved in [DLS98].

1.2 Separation Results for Monadic Second-Order Logic

In this section we shall try to give a brief summary of the historical development of questions and results treated in this thesis. We will also give motivation. A more precise overview based on detailed definitions will be given in the next chapter.

For particular classes of structures, namely words or trees, every monadic second-order formula is equivalent to an existential one. Moreover, there exists an automaton model that captures this existential fragment of monadic second-order logic in expressive power. This fundamental property can be used to build tools for automated verification.

This as well as the previously mentioned relations to complexity-theoretic questions motivate to ask the following.

Question 1.2 *Does the monadic hierarchy collapse to its existential fragment, i.e., is every monadic formula equivalent to an existential one?*

If not in general (i.e., over arbitrary finite structures), then for which classes of structures?

As mentioned above, the well-known collapse of the monadic alternation hierarchy over word and trees justified some hope that this would happen to other classes of structures, too.

However, in [Fag74] Fagin showed that there is a Σ_1-formula over graphs that is not equivalent to any Π_1-formula, thus answering this question negatively. Fagin even showed this fact for a very restricted class of graphs, namely those that are "disjoint unions" of cycles.

Question 1.3 *Is there a k such that every monadic second-order sentence is equivalent to a Σ_k-sentence, i.e., can the number of quantifier alternations be limited in monadic second-order logic?*

Let $\underline{\Sigma_k}$ and $\underline{\Pi_k}$ denote the class of properties definable by Σ_k- and Π_k-sentences, respectively. Then Question 1.3 can be formulated as: is there some k such that for all l we have $\underline{\Sigma_k} = \underline{\Sigma_{k+l}}$?

In [MT97] this question is answered negatively. The proof uses the class of *grids*. A grid of size (m, n) is, roughly speaking, a finite graph whose vertices are arranged as elements of a $(m \times n)$-matrix and which has two edge relations corresponding to vertical and horizontal successors. More precisely, [MT97] shows the following fact. (See Corollary 2.25 below.)

Fact 1.4 $\underline{\Sigma_k} \subsetneq \underline{\Sigma_{k+1}}$ *for the classes of coloured grids and thus for the class of graphs.*

The non-definability part of this separation result, which usually attracts more attention though it is the easier one, uses automata-theoretic ideas rather than "Ehrenfeucht-Fraïssé games with multiple colouring rounds".

In [Sch97a] this result was strengthened to $\underline{\Sigma_k} \subsetneq \underline{\Sigma_{k+1}}$ for the class of (non-coloured) grids. In this thesis we even show the following. (See Corollary 2.25 and Theorem 2.32.)

Fact 1.5 $\underline{\Sigma_k}$ *is not closed under complement for the class of grids and thus for the class of graphs.*

By Fact 1.5, expressive power increases if we allow a formula on level k to have both an existential and a universal part, combined by some boolean combination. On the other hand, expressive power decreases if we require a formula on level k to be equivalent both to an existential and to a universal formula.

This leads us to the definition of the formula classes $B(\Sigma_k)$ and Δ_k. A boolean combination of Σ_k-formulas is a $B(\Sigma_k)$-formula. A formula that is equivalent[1] both to a Σ_k-formula and to a Π_k-formula is called a Δ_k-formula. The classes of properties definable by sentences in these classes are denoted $\underline{B(\Sigma_k)}$ and $\underline{\Delta_k}$.

Question 1.6 *Is every formula on level $k+1$ equivalent to a boolean combination of formulas on level k? If not in general, then maybe over the class of grids?*

Results from [Sch97a, MT97] show the following fact, strengthening Fact 1.4. (See Corollary 2.31.)

Fact 1.7 $\underline{B(\Sigma_k)} \subsetneq \underline{\Delta_{k+1}}$ *even for the class of grids.*

When investigating the proof for Facts 1.4 and 1.7 it is noticeable that monadic quantifications are needed only in a very limited way. We would like to investigate this in more detail, i.e., we consider the following question.

Question 1.8 *"How much" of* monadic *quantification is really needed to exceed the k-th level of the monadic hierarchy in expressiveness?*

Two results that shed light on this problem are Facts 1.9 and 1.10, both shown in [MT97] (see Corollary 2.25). The first one deals with the *unary transitive closure operator (unary* TC*-operator)*. The unary transitive closure operator is an extension of first-order logic that allows to define the reflexive and transitive closure of a binary relation B on the universe, provided the relation B is definable itself.

Fact 1.9 *Beyond each level in the monadic hierarchy over coloured grids, there is a property expressible by a first-order formula with unary transitive closure operators.*

This is indeed an answer to Question 1.8 because one may replace the unary TC-operator by monadic quantifications, but in general not vice versa.

[1] We will have to make precise later on which notion of equivalence we use here. For the moment, this may be read as "logically equivalent".

Thus the unary TC-operator amounts to a "limited form" of monadic quantification.

First-order quantification may be viewed as such a limited form of monadic quantification, too. As mentioned in the previous section, each level of the full second-order hierarchy is closed under first-order quantifications. This means that, for example, if φ belongs to level k of the full second-order alternation hierarchy, then $\exists x \forall y \varphi$ belongs to the same level.

However, the levels of the monadic hierarchy are *not* closed under first-order quantifications, making this a less robust hierarchy, as pointed out in [AFS97].

Taking up Question 1.8, a natural approach is to ask: Do *first-order* alternations followed by one monadic second-order quantification suffice to leave Σ_k? This is indeed the case; namely we have the following (see Corollary 2.25).

Fact 1.10 *Beyond each level in the monadic hierarchy over coloured grids, there is a property expressed by a formula of the form*

$$\eth_1 x_1 \ldots \eth_n x_n \varphi,$$

where φ is equivalent to a Σ_1-formula and $\eth_1, \ldots, \eth_n$ are existential or universal first-order quantifiers.

The proof shows that there is a formula in the above form whose number of quantifier blocks in the first-order quantifier prefix is $k+1$ and which is not equivalent to a formula in the k-th level.

Both Fact 1.9 and Fact 1.10 can be transferred to the class of graphs, which has been done in [MST99] but not explicitly in [MT97].

Fact 1.10 motivates the following definition: For a class $\mathcal{F}$ of formulas, the *first-order closure of $\mathcal{F}$*, denoted by $FO(\mathcal{F})$, is defined as the smallest class of formulas that contains $\mathcal{F}$ and is closed under first-order quantifications and boolean combinations.

Fact 1.10 shows that, in some sense, first-order alternation is fairly powerful compared to monadic second-order alternation: so powerful that the first-order closure even of the first level allows to define properties beyond any level of the monadic hierarchy. This observation gives rise to several more questions.

Question 1.11 *Is there a level whose first-order closure captures the complete monadic second-order logic in expressiveness?*

In [AFS97] it is proved (based on the growth argument of [MT97]) that $\underline{\Sigma_3} \not\subseteq \underline{FO(\Sigma_1)}$, answering Question 1.11 negatively for the first level. In Corollary 2.31 (see also [Mat98a]) we show:

Fact 1.12 *There is no level of the monadic hierarchy whose first-order closure captures the complete monadic second order logic, more precisely: $\underline{\Delta_{k+2}} \not\subseteq \underline{FO(\Sigma_k)}$ for every $k \geq 1$.*

Besides, the first-order closure of level k is strictly less expressive than the first-order closure of level $k+1$, more precisely: $\underline{FO(\Sigma_k)} \subsetneq \underline{FO(\Delta_{k+1})}$ for every $k \geq 1$.

In contrast to Fact 1.10, which stresses the expressive power of the first-order closure, Fact 1.12 shows its weakness: it is so weak that for any level of the monadic hierarchy, adding two more monadic alternations results in expressive power that cannot be captured by adding any number of first-order alternations.

The second statement of Fact 1.12 might be viewed as a preliminary result for the following question.

Question 1.13 *"How much" can be added to monadic logic so that the strictness of the alternation hierarchy remains valid?*

Fact 1.10 gives a partial answer to Question 1.8 in that it shows that we need only first-order quantifications followed by only one block of monadic quantifiers in order to define properties high in the monadic hierarchy.

We have two more results that suggest that the answer to Question 1.8 is "hardly any".

Fact 1.14 *(cf. Corollary 2.25) In the formulas of Fact 1.10, the monadic quantification is only needed in a restricted way, namely the Σ_1-kernel φ expresses the existence of a tuple of sets that is uniquely determined by a first-order formula. In particular, φ is also equivalent to a Π_1-formula.*

Fact 1.15 *(cf. Theorem 6.7) Over the class of grids, the length of the quantifier block of an existential monadic second-order formula (for example, of the φ from Fact 1.10) can be limited to one.*

Let us have a look at Fact 1.9. Motivated from experience of formal word languages, it is straightforward to expand grids by two more binary relations, namely $\leq_1$, $\leq_2$, denoting the vertical and horizontal partial order of positions, respectively.

Obviously, these two orderings are the reflexive and transitive closure of the two successor relations, so if we add these two partial orderings to first-order logic, it is at most as expressive as (and, as one easily verifies, strictly less expressive than) first-order logic with unary TC-operator. That is why the following result from [Sch97a] strengthens Fact 1.9.

Fact 1.16 *Beyond each level in the monadic hierarchy over coloured grids, there is a property expressed by a first-order formula using the additional two partial orderings.*

In the following subsection, we will inspect this fact in a closer way.

Picture Languages

Coloured grids may be identified with "pictures", i.e., matrices over a finite alphabet. Pictures are the two-dimensional analogue to words, so it is a natural idea to try to transfer definitions and results from word languages to picture languages.

One interesting class of formal word languages is the class of *starfree* word languages. This class of word languages defined by starfree expressions, i.e., regular expressions using concatenation, union, and complement. In other words, the class of starfree languages is the smallest class of word languages that contains all finite ones and is closed under concatenation, union, and complement.

There is a natural adaption of "starfeeness" to picture languages. The difference is that one allows the two concatenation operations of pictures (horizontal and vertical) instead of the single concatenation in the word language case.

It is a well-known fact from word language theory that a word language is starfree iff it is definable by a first-order sentence with ordering. One may ask whether the corresponding statement is true for picture languages. It is easy to see that every starfree picture language is definable by a first-order sentence with the two partial orderings, but [Wil97] shows that the converse is not true:

Fact 1.17 *The class of starfree picture languages is strictly included in the class of picture languages definable by first-order sentences with the horizontal and vertical ordering.*

What does this mean with respect to Question 1.8 and Fact 1.16? It means that Fact 1.16 does not answer the following question.

Question 1.18 *Is every starfree picture language in Σ_1?*

Question 1.18 was open in [GRST96, GR96]. It was even unknown whether every regular picture language is in Σ_1, where "regular" means definable by a regular expression using the two concatenations, union, complement, and two distinct "Kleene stars" that iterate the two concatenations.

In [Mat98b] it is shown that the answer is "no". This might be considered surprising because in word language theory, the starfree languages are a rather small subset of the recognizable languages, and the adaptions of both of these notions to pictures is somewhat straightforward.

This negative answer to Question 1.18 can be strengthend (see Corollary 2.25) to the following fact, which (by Fact 1.17) also strengthens Fact 1.16.

Fact 1.19 *Beyond every level of the monadic hierarchy there is a starfree picture language.*

The minimal number of (e.g. monadic) quantifier alternations is an appropriate measure for the complexity of properties definable in a certain logic. What is an appropriate measure in the case of starfree (word or picture) languages? One measure has been introduced in [BK71] and is called "dot-depth". A starfree word language is of dot-depth k if it can be defined by a starfree expression with at most k "alternations" of concatenations and boolean operations. It is well-known that the dot-depth hierarchy over word languages is strict in the sense that for every k, there is a starfree word language that is of dot-depth $k+1$ but not of dot-depth k.

The notion of dot-depth can be transferred to starfree picture languages in a straightforward way. One may ask whether this hierarchy is also strict over pictures. The answer is trivially "yes". This strictness is an immediate consequence of the strictness in the word language case because every word language may be identified with a picture language, and this identification respects the dot-depth measure. But with regard to Fact 1.19 it seems justified to investigate the relation between dot-depth and quantifier alternation. The proof for Facts 1.19 and 1.10 shows the following result (see Corollary 2.25).

Fact 1.20 *For every k, there is a picture language of dot-depth $k+1$ that is not in level k of the monadic hierarchy.*

This stresses the close connection of first-order quantifier alternation to dot-depth: existential first-order quantification corresponds to concatenation, and negation corresponds to complement. This close correspondence is not a new insight. It has been investigated e.g. in [Tho82].

In the subseqent chapters we are going to prove, among others, the three Facts 1.7, 1.10, and 1.20. This requires giving, for each level k, a "difficult" picture language that is definable by formulas of a certain type as well as by a starfree expression of a certain type. In order to keep this as simple as possible and to avoid double work, we are going to show that this difficult picture language is in some particular master class that allows us to deduce the definability both by formulas of the various required types and by certain starfree expressions. See Definition 2.20 as well as Theorems 2.22 and 2.24.

What Follows

The rest of this thesis is structured as follows: In Sections 2.1–2.3 we will introduce most of the necessary definitions of this thesis. In Section 2.4 we will state all separation results.

Our proofs for these separation results typically consist of two parts. The first part states that a certain class of pictures is definable by a formula of a certain shape. The second part states that it is not definable by a formula of some other shape.

We collect all the "first parts" in Chapter 3 and all "second parts" in Chapter 4. The reader might wish to skip some sections in these two chapters, see Subsection 2.4.4 for hints on shortcuts.

In Chapter 5 we develop an encoding technique that allows us to transfer most of the separation results of the preceding chapters from pictures to other structures, in particular finite graphs.

In Chapter 6 we investigate the quantifier depth hierarchy (rather than the alternation hierarchy), for which we recall two known facts (Facts 6.1 and 6.3) and show one new result (Theorem 6.7).

We give a conclusion in the last chapter, which also contains some open questions that I consider interesting.

Chapter 2

Pictures, Grids, and Alternation—Definitions and Results

In this chapter we will define the previously mentioned notions formally and observe some easy facts. In Section 2.4 we will state the main results of this thesis.

Some Basic Notations

We write $\mathbb{N}$ for the set of natural numbers (without zero). For every $n, m \geq 0$, we write $[n]$ for $\{0, \ldots, n-1\}$ and $[n, m]$ for $\{n, n+1, \ldots, m\}$.

If a binary relation f is a partial function, we use the usual notations, for example $f(m)$ for "the" element n with $(m, n) \in f$ (if it exists). We write $\mathrm{dom} f$ for the set of all m for which $f(m)$ is defined.

We say that a partial function $f : \mathbb{N} \dashrightarrow \mathbb{N}$ "is $g(\mathcal{O}(m))$" if there is a $c \geq 1$ such for almost all $m \in \mathrm{dom} f$ we have $f(m) \leq g(cm)$. We also use the Θ- and the Ω-notation in the usual meaning.

For a set M, the powerset of M is denoted by $\mathcal{P}(M)$.

2.1 Grids and Logic

2.1.1 Grids

Grids are particular relational structures over the signature $\tau_{Grids} = \{S_1, S_2\}$ containing the binary "successor" relation symbols S_1 and S_2. The universe of a grid is of the form $[m]\times[n]$. The *grid of height m and width n* is the τ_{Grids}-structure

$$\underline{[m]\times[n]} := ([m]\times[n], S_1^{m,n}, S_2^{m,n}),$$

where $S_1^{m,n}$ and $S_2^{m,n}$ are the "vertical" and the "horizontal" successor relations on $[m]\times[n]$, containing all pairs $((i,j),(i+1,j))$ and all pairs $((i,j),(i,j+1))$, respectively, from $[m]\times[n]$. Let $\mathrm{size}(\underline{[m]\times[n]}) = (m,n)$.

The expressions *row, column, top, bottom* etc. are interpreted as in the terminology of matrices; e.g. the leftmost column contains exactly the vertices $(i, 0)$ for $0 \leq i \leq m-1$, and the top row consists of all vertices $(0, j)$ for $0 \leq j \leq n-1$.

We adjoin unary predicate symbols $X_1, \ldots, X_t$ to the signature of grids, obtaining the signature $\tau_{Grids_t} = \tau_{Grids} \cup \{X_1, \ldots, X_t\}$. A *$t$-bit grid* (for some $t \geq 0$) is a τ_{Grids_t}-structure, i.e., of the form $R = (\mathrm{dom}R, S_1^R, S_2^R, X_1^R, \ldots, X_t^R)$, whose restriction to τ_{Grids} is a grid. So grids are 0-bit grids.

The classes of grids and t-bit grids will be denoted *Grids* and $Grids_t$, respectively.

A different version of grids is obtained via the signature $\tau_{Grids} \cup \{\leq_1, \leq_2\}$ with binary relation symbols $\leq_1, \leq_2$. In this context, a grid is considered as the structure

$$([m]\times[n], S_1^{m,n}, S_2^{m,n}, \leq_1^{m,n}, \leq_2^{m,n}),$$

where $\leq_1^{m,n} = \{((i,j),(i',j)) \in ([m]\times[n]) \times ([m]\times[n]) \mid i \leq i'\}$, and $\leq_2^{m,n} = \{((i,j),(i,j')) \in ([m]\times[n]) \times ([m]\times[n]) \mid j \leq j'\}$. That means that $\leq_1^{m,n}$ (and $\leq_2^{m,n}$) are the reflexive and transitive closure of $S_1^{m,n}$ (and $S_2^{m,n}$, respectively).

2.1.2 Monadic Second-Order Logic

Let us fix a signature τ. By "signature" we always mean "relational signature". Many notations introduced in this and the next section are relative to τ without mentioning this explicitly. This will not cause any confusion because in most of this thesis, τ will be the signature τ_{Grids} from above.

We use $x, y, x_1, x_2, \ldots$ as first-order variables and $X, Y, \ldots$ as monadic second-order variables (i.e., *set variables*). *Monadic second-order formulas over* τ (*MSO*-formulas for short) are built up as usual from atomic formulas $x = y$, Xy, and $r(x_1, \ldots, x_l)$ (for an l-ary relation symbol r of τ). Formally, if φ and ψ are *MSO*-formulas, x is a first-order variable, and X is a set variable, then $\varphi \vee \psi$, $\neg\varphi$, $\exists x\varphi$, and $\exists X\varphi$ are *MSO*-formulas, too. We may use the other propositional connectives, like $\rightarrow$, $\wedge$, etc., as well as universal quantifications as abbreviations with their usual meaning.

First-order formulas are *MSO*-formulas in which no second-order quantifier occurs. *FO* denotes the class of first-order formulas.

In Subsection 2.2.2 we will give two extensions of the syntax of formulas.

The set of variables occurring free in a formula φ is denoted $\mathit{free}(\varphi)$. If $\mathit{free}(\varphi) \subseteq \{X_1, \ldots, X_t, x_1, \ldots, x_m\}$, we sometimes write $\varphi(X_1, \ldots, X_t, x_1, \ldots, x_m)$. If M is a τ-structure and $U_1, \ldots, U_t$ and $u_1, \ldots, u_m$ are subsets and elements, respectively, of the domain $\mathrm{dom}M$ of M, and if $\varphi(X_1, \ldots, X_t, x_1, \ldots, x_m)$ is a formula over τ, then we write $M \models \varphi[U_1, \ldots, U_t, u_1, \ldots, u_m]$ to indicate that φ is true in M under the assignment that maps X_i to U_i and x_j to u_j for all i, j. If $\mathcal{C}$ is a class of τ-structures and φ is a sentence, i.e., $\mathit{free}(\varphi) = \emptyset$, then we write $\mathit{Mod}_{\mathcal{C}}(\varphi)$ for the set $\{M \in \mathcal{C} \mid M \models \varphi\}$ of models of φ. If $L = \mathit{Mod}_{\mathcal{C}}(\varphi)$, we say that φ *defines* L *(within* $\mathcal{C}$*)*. If $\mathcal{F}$ is a set of formulas, then we call a set $L \subseteq \mathcal{C}$ of structures $\mathcal{F}$*-definable within* $\mathcal{C}$ if there is a sentence $\varphi \in \mathcal{F}$ with $L = \mathit{Mod}_{\mathcal{C}}(\varphi)$.

In the context of t-bit grids, an *MSO*-formula φ over τ_{Grids} with $\mathit{free}(\varphi) \subseteq \{X_1, \ldots, X_t\}$ is considered as a sentence over τ_{Grids_t}. (When doing so, we assume that $X_1, \ldots, X_t$ are not bound in φ.)

2.2 Closures and Alternation

We have motivated in Chapter 1 to measure the complexity of monadic second-order formulas by the alternation of existential and universal set quantifications.

Closure Operations for Formula Classes

We will continue to define the different levels of the monadic quantifier alternation hierarchy. As mentioned before, we will also be interested in first-order alternation and the interference of these two. Besides, we will introduce the "unary TC-operator" (for the transitive closure) and "the monadic

ι-operator" (for unique choice) as syntactic extensions of monadic second-order formulas. Some of our separation results give nontrivial facts about their expressive power in monadic second-order logic.

In order to be able to state and prove all these facts formally, we will have to introduce some more general notations.

For a class $\mathcal{F}$ of formulas, let *co-*$\mathcal{F}$ be the class of formulas $\neg\varphi$ with $\varphi \in \mathcal{F}$.

A class $\mathcal{F}$ of formulas is *closed under (1) negation, (2) disjunction, (3) conjunction, (4) existential first-order quantification, (5) universal first-order quantification, or (6) existential second-order quantification*, respectively, iff for every $\varphi, \varphi' \in \mathcal{F}$ we have

1. $\neg\varphi \in \mathcal{F}$,
2. $(\varphi \vee \varphi') \in \mathcal{F}$,
3. $(\varphi \wedge \varphi') \in \mathcal{F}$,
4. $\exists x\varphi \in \mathcal{F}$ for every first-order variable x,
5. $\forall x\varphi \in \mathcal{F}$ for every first-order variable x,
6. $\exists X\varphi \in \mathcal{F}$ for every second-order variable X,

respectively. $\mathcal{F}$ is closed under *boolean combinations* iff $\mathcal{F}$ is closed under both negation and disjunction. $\mathcal{F}$ is closed under *positive boolean combinations* iff $\mathcal{F}$ is closed under both disjunction and conjunction.

Let $\mathcal{F}$ be a class of formulas. The

1. *boolean closure* of $\mathcal{F}$, denoted $B(\mathcal{F})$,
2. *positive boolean closure* of $\mathcal{F}$, denoted $PB(\mathcal{F})$,
3. *existential first-order closure* of $\mathcal{F}$, denoted $\Sigma_1^0(\mathcal{F})$,
4. *universal first-order closure* of $\mathcal{F}$, denoted $\Pi_1^0(\mathcal{F})$,
5. *existential second-order closure* of $\mathcal{F}$, denoted $\Sigma_1(\mathcal{F})$,
6. *closed existential second-order closure* of $\mathcal{F}$, denoted $\mathit{closed}(\Sigma_1)(\mathcal{F})$,
7. *first-order closure* of $\mathcal{F}$, denoted $FO(\mathcal{F})$,

respectively, is defined as the smallest superclass of $\mathcal{F}$ that is closed under

1. boolean combinations,
2. positive boolean combinations,
3. existential first-order quantifications and positive boolean combinations,
4. universal first-order quantifications and positive boolean combinations,
5. existential second-order quantifications and positive boolean combinations,
6. existential second-order quantifications, positive boolean combinations, and existential and universal first-order quantifications,
7. existential first-order quantifications and boolean combinations,

respectively. In Subsection 2.2.2 we will introduce two more closure operations on formula classes.

Remark 2.1 Let $\mathcal{F}$ be a class of formulas that is closed under positive boolean combinations. Typical formulas in (3) $\Sigma^0_1(\mathcal{F})$, in (4) $\Pi^0_1(\mathcal{F})$, in (5) $\Sigma_1(\mathcal{F})$, in (6) *closed*$(\Sigma_1)(\mathcal{F})$, and in (7) $FO(\mathcal{F})$, respectively, are (up to logical equivalence) of the form

(3) $\exists\overline{x}\psi$,

(4) $\forall\overline{x}\psi$,

(5) $\exists\overline{X}\psi$,

(6) $\exists\overline{X}_1\eth_1\overline{x}_1\exists\overline{X}_2\eth_2\overline{x}_2\ldots\exists\overline{X}_n\eth_n\overline{x}_n\psi$ (where $\eth_1,\ldots,\eth_n \in \{\exists,\forall\}$),

(7) $\exists\overline{x}_1\forall\overline{x}_2\ldots\eth\overline{x}_n\psi$ (for $\eth=\exists$ or $\eth=\forall$),

where ψ is a formula in $\mathcal{F}$ and $\overline{x},\overline{x}_1,\ldots,\overline{x}_n$ are first-order variable tuples and $\overline{X},\overline{X}_1,\ldots,\overline{X}_n$ are second-order variable tuples, both times of possibly different length. □

In Chapter 5 we will use these definitions to introduce a kind of calculus for formula classes involving these closures and show that for such formula classes, separation results can be transferred from one class of structures (like grids) to another (say graphs) by the encoding technique (called "strong first-order reductions") presented in that chapter.

2.2.1 Alternation Hierarchies for Monadic Logic

Now we are ready to define several alternation hierarchies for monadic logic. One benefit of the notions of "closures" introduced above is that these definitions are quite succinct.

Definition 2.2 (Alternation Hierarchies) Let $\mathcal{F}$ be a class of formulas. We define:

$$\Sigma^0_0(\mathcal{F})=\mathcal{F} \quad\text{and}\quad \Sigma^0_{k+1}(\mathcal{F})=\Sigma^0_1(co\text{-}\Sigma^0_k(co\text{-}\mathcal{F})),$$
$$\Sigma_0(\mathcal{F})=\mathcal{F} \quad\text{and}\quad \Sigma_{k+1}(\mathcal{F})=\Sigma_1(co\text{-}\Sigma_k(co\text{-}\mathcal{F}))$$

for every $k\geq 0$. Let $\Pi^0_k(\mathcal{F})=co\text{-}\Sigma^0_k(co\text{-}\mathcal{F})$ and $\Pi_k(\mathcal{F})=co\text{-}(\Sigma_k(co\text{-}\mathcal{F}))$ for every k.

Additionally,

$$closed(\Sigma_0)(\mathcal{F})=\mathcal{F} \quad\text{and}\quad closed(\Sigma_{k+1})(\mathcal{F})=closed(\Sigma_1)(co\text{-}\Sigma_k(co\text{-}\mathcal{F}))$$

for every $k\geq 0$. We write Σ_k, Π_k, and *closed*(Σ_k) instead of $\Sigma_k(FO)$, $\Pi_k(FO)$, *closed*$(\Sigma_k)(FO)$, respectively. □

(Note that the re-definition of $\Sigma^0_1(\mathcal{F})$, $\Sigma_1(\mathcal{F})$, and *closed*$(\Sigma_1)(\mathcal{F})$ is compatible with the original one, provided we identify formulas of the form φ and $\neg\neg\varphi$.)

For the purposes of this chapter and Chapters 3 and 4 it would not make a difference if we replaced the "$co\text{-}\mathcal{F}$" in the above definition by "$\mathcal{F}$" or "$B(\mathcal{F})$". But in Chapter 5 we will use these closures for the mentioned calculus, and then this definition allows for more succinct proofs, see e.g. Remark 5.13.

The class Σ_k of formulas (as well as the class of properties definable by such formulas) is known as "the k-th level of the monadic quantifier alternation hierarchy". There is some arbitrariness in choosing Σ_k to be "the level k". Other possible choices would be Π_k, or $\Sigma_k\cup\Pi_k$, or the boolean closure of Σ_k, or the class Δ_k introduced below. The class of properties definable by formulas in Σ_1, i.e., in the existential fragment of monadic second-order logic, is often called "monadic NP", but we avoid this term because it is not a complexity class.

The class *closed*(Σ_k) has been introduced in [AFS97] and is called "the k-th level of the closed monadic quantifier alternation hierarchy", and the class of properties definable on the first level of this hierarchy is called "closed monadic NP" in [AFS97].

Remark 2.3 Let $\mathcal{F}$ be a class of formulas that is closed under positive boolean combinations. Let $k \geq 0$. A formula in $\Sigma_k^0(\mathcal{F})$ is (up to logical equivalence) of the form $\exists\overline{x}_1\neg\exists\overline{x}_2\neg\ldots\neg\exists\overline{x}_k\varphi$, where $\varphi \in \mathcal{F} \cup co\text{-}\mathcal{F}$ and $\overline{x}_1,\ldots,\overline{x}_k$ are first-order variable tuples of possibly different length. Using the universal quantifier, which we allowed as an abbreviation, such a formula is equivalent to a formula of the form

$$\exists\overline{x}_1\forall\overline{x}_2\ldots\mho\overline{x}_k\varphi',$$

for some $\varphi' \in \mathcal{F}$ and $\mho = \exists$ or $\mho = \forall$, depending on whether k is odd or even. A typical formula in $\Pi_k^0(\mathcal{F})$ is of the form $\neg\exists\overline{x}_1\neg\exists\overline{x}_2\neg\ldots\neg\exists\overline{x}_k\varphi$ for some $\varphi \in \mathcal{F} \cup co\text{-}\mathcal{F}$, i.e., up to logical equivalence of the form

$$\forall\overline{x}_1\exists\overline{x}_2\ldots\mho\overline{x}_k\varphi'$$

for some $\varphi' \in \mathcal{F}$.

Formulas in $\Sigma_k(\mathcal{F})$ and $\Pi_k(\mathcal{F})$ are similar but with monadic second-order quantifications instead of first-order quantifications. The important special case is that $\mathcal{F} = FO$ is the class of first-order formulas, so a Σ_k-formula is of the form

$$\exists\overline{X}_1\neg\exists\overline{X}_2\ldots\neg\exists\overline{X}_k\varphi$$

for a first-order φ and second-order variable tuples $\overline{X}_1,\ldots,\overline{X}_k$. □

Equivalence and Δ

There may be Σ_k-definable properties which are not Π_k-definable and vice versa. There is a notion for those properties, say of a class $\mathcal{C}$ of structures, which are both Σ_k- and Π_k-definable. Usually this class of classes of $\mathcal{C}$-structures is denoted $\Delta_k(\mathcal{C})$ or similarly. However, for our purposes it does not suffice to introduce this notion on a semantic level, i.e., as a class of classes of structures, but we would like to have a notion of "Δ_k-formula". For this aim, we clearly need a notion of equivalence of formulas. The equivalence we use will be coarser than logical equivalence because it is relative to a given class of structures (in this thesis usually grids).

Definition 2.4 Let $\mathcal{C}$ be a set of τ-structures. Two formulas φ, ψ are *equivalent over* $\mathcal{C}$ iff $\varphi \leftrightarrow \psi$ is true in all structures of $\mathcal{C}$ under all assignments. For a set $\mathcal{F}$ of formulas we write $\mathcal{F}^{\mathcal{C}}$ for the set of all formulas that are

equivalent (over $\mathcal{C}$) to a formula in $\mathcal{F}$, and $\underline{\mathcal{F}}(\mathcal{C}) := \{Mod_{\mathcal{C}}(\varphi) \mid \varphi$ is a sentence in $\mathcal{F}\}$.

Let ${\Delta_k}^{\mathcal{C}} = {\Sigma_k}^{\mathcal{C}} \cap {\Pi_k}^{\mathcal{C}}$ for every k, and $\underline{\Delta_k}(\mathcal{C}) = \underline{{\Delta_k}^{\mathcal{C}}}(\mathcal{C})$. □

In any of the above notations, we omit $\mathcal{C}$ when this causes no confusion. (If nothing else is stated, $\mathcal{C}$ is the class of grids.) Consequently, we do not distinguish between a class of formulas and its closure under equivalence over grids—every formula that is equivalent over grids to a, say, Σ_k-formula will be called "Σ_k-formula" itself.

For a class $\mathcal{F}$ of formulas, the Δ_1*-closure of* $\mathcal{F}$ *wrt.* $\mathcal{C}$ is given by

$$\Delta_1^{\mathcal{C}}(\mathcal{F}) = \Sigma_1(B(\mathcal{F}))^{\mathcal{C}} \cap \Pi_1(B(\mathcal{F}))^{\mathcal{C}}.$$

The notion of a "Δ_k-formula" has to be treated carefully because it refers implicitly to a given class of structures. If $\mathcal{D}$ is a superclass of $\mathcal{C}$ over the same signature, there might be a ${\Delta_k}^{\mathcal{C}}$-formula that is not equivalent to any formula in ${\Delta_k}^{\mathcal{D}}$. This observation will be of major importance in Chapter 5, in particular in Section 5.3.

2.2.2 The Unary TC-Operator and the Monadic ι-Operator

Our proof of the fact that the monadic second-order quantifier alternation hierarchy is strict will show that one can define properties "high" in the hierarchy with a very limited form of set quantifications. In order to formulate this observation precisely, we introduce two extensions of the syntax of formulas which may—in some situations—replace set quantifications.

The Unary Transitive Closure Operator

The first one is the "unary transitive closure operator" (TC-operator). If φ is a formula and x, y are first-order variables, then $\mathrm{TC}(x, y, \varphi)$ is a formula whose free variables are x, y, and the free variables of φ. The intended meaning of the formula $\mathrm{TC}(x, y, \varphi)$ is that x and y are in the reflexive and transitive closure of the binary relation given by $\{(x', y') \mid x', y' \text{ satisfy } \varphi\}$.

Formally, if $\varphi(x, y, \overline{Z}, \overline{z})$ is a formula, where $\overline{Z}$ and $\overline{z}$ are first- and second-order variable tuples, respectively, of possibly different lengths, then the satisfaction relation $M \models \mathrm{TC}(x, y, \varphi)[u, v, \overline{W}, \overline{w}]$ holds for two elements u, v, a tuple $\overline{W}$ of subsets, and a tuple $\overline{w}$ of elements of $\mathrm{dom}M$ iff there are elements $u_0, \dots, u_m \in \mathrm{dom}M$ for some $m \geq 0$ such that $u = u_0$, $v = u_m$, and $M \models \varphi[u_i, u_{i+1}, \overline{W}, \overline{w}]$ for all $i \in \{0, \dots, m-1\}$.

Example 2.5 A common example of the use of the unary transitive closure operator is the following. Consider the signature $\{E\}$ with the binary relation symbol E. A directed graph is a finite structure over that signature. The formula $\varphi(x, y) = Exy$ asserts, for two vertices x and y in the domain of some graph, that there is an edge from x to y. The formula $\mathrm{TC}(x, y, \varphi)$ asserts that y is reachable from x via some (possibly empty) path in the graph. □

We remark that the application of a unary TC-operator can always be replaced by a universal set quantification because a formula $\mathrm{TC}(x, y, \varphi)$ is equivalent to the formula $\forall X((Xx \wedge \forall x \forall y(Xx \wedge \varphi \to Xy)) \to Xy)$, which states that every set X that is closed under the φ-relation and contains x also contains y.

In this sense, the unary TC-operator is indeed a limited form of set quantification.

Another Closure Operation of Formula Classes

Let $\mathcal{F}$ be a class of formulas. $\mathcal{F}$ is *closed under application of the unary* TC*-operator* iff for every $\varphi \in \mathcal{F}$ and all first-order variables x, y we have that $\mathrm{TC}(x, y, \varphi)$ is in $\mathcal{F}$.

The *unary* TC*-closure* of $\mathcal{F}$, denoted $FOTC(\mathcal{F})$, is defined as the smallest superclass of $\mathcal{F}$ that is closed under application of the unary TC-operator, existential first-order quantifications, and positive boolean combinations.

The Monadic ι-Operator

The second extension of the syntax of formulas will be described now.

If φ and ψ are formulas and $\overline{X} = (X_1, \dots, X_t)$ is a tuple of set variables that are free in φ, then "$\psi(\iota\overline{X}(\varphi))$" is meant to be a formula in which the free variables are those of φ and those of ψ but not the variables in $\overline{X}$. It shall express that ψ holds for "the unique" tuple $\overline{X}$ that makes φ true. Of course, this only makes sense in case φ uniquely determines a tuple $\overline{X}$, and this depends on the considered class of structures.

Formally, let $\mathcal{C}$ be a class of structures and let $\overline{Y}$ be a tuple of first- and second-order variables such that $\varphi = \varphi(\overline{X}, \overline{Y})$ and $\psi = \psi(\overline{X}, \overline{Y})$, i.e., every free variable of φ or ψ appears either in $\overline{X}$ or in $\overline{Y}$. Let us write $\overline{X} = \overline{X}'$ to abbreviate the first-order formula $\bigwedge_{i=1}^{t}(\forall x(X_i x \leftrightarrow X_i' x))$.

Assume that the formula $\exists!\overline{X}\varphi$, which abbreviates $\exists\overline{X}\forall\overline{X}'(\varphi(X', \overline{Y}) \leftrightarrow \overline{X} = \overline{X}')$, is true in all $\mathcal{C}$-classes under all assignments; then

$$\psi(\iota\overline{X}(\varphi), \overline{Y})$$

is also a formula with free variables among $\overline{Y}$.

The satisfaction relation is defined as follows. For a structure M of $\mathcal{C}$ and a tuple $\overline{V}$ (to be assigned to $\overline{Y}$) of subsets and elements of $\mathrm{dom}M$, we have $M \models \psi(\iota\overline{X}(\varphi), \overline{Y})[V]$ iff $M \models \psi[\overline{U}, \overline{V}]$ for the unique tuple $\overline{U}$ for which $M \models \varphi[\overline{U}, \overline{V}]$.

As stated in Proposition 2.6 below, the application of the unary ι-operator may be replaced by a block of existential or universal set quantifications. In this sense, the ι-operator is a restricted form of second-order quantifications, too.

Similarly as in the case of a "Δ_k-formula", the reader is warned about the implicit reference to a given class $\mathcal{C}$ of structures. In the context of the monadic ι-operator, the question of what is or is not a "formula" depends on $\mathcal{C}$! If $\mathcal{D}$ is a superclass of $\mathcal{C}$ over the same signature, then there might be a some $\psi(\iota\overline{X}(\varphi), \overline{Y})$ which is a formula relative to the class $\mathcal{C}$ but which is no formula relative to the class $\mathcal{D}$ because $\exists!\overline{X}\varphi$ is a valid formula in all $\mathcal{C}$- but not in all $\mathcal{D}$-structures.

Thus we are dealing with a semantic restriction on the syntax of formulas, and we should be careful whenever the ι-operator appears. In order to stress this fact, we will call a formula that uses the ι-operator (and thus matches this restriction) an *allowed formula.*

Fortunately, in Chapters 3 and 4 there will be one fixed class of structures, namely the class of grids, so that there will be no doubt about what is an allowed formula.

Yet Another Closure Operation of Formula Classes

Let $\mathcal{F}$ be a class of formulas and $\mathcal{C}$ be a class of structures over the same signature. Let $\Delta_1^{\mathrm{U},\mathcal{C}}(\mathcal{F})$ be the set of allowed formulas of the form

$$\psi(\iota\overline{X}(\varphi), \overline{Y}),$$

where $\varphi(\overline{X}, \overline{Y})$ is a formula in $B(\mathcal{F})$, the formula $\psi(\overline{X}, \overline{Y})$ is first-order, $\overline{X}$ is a tuple of set variables, and $\overline{Y}$ is a tuple of first-order and/or set variables of possibly different length.

Note that although we drop the superscript $\mathcal{C}$ when it is fixed, it is essential for the definition of $\Delta_1^{\mathrm{U},\mathcal{C}}(\mathcal{F})$.

2.2.3 Some Calculation Rules

Some calculation rules may be deduced from the well-known rules of predicate logic. For example, the following inclusions hold for every formula class $\mathcal{F}$.

$$\begin{array}{ll}
\underline{\Delta_k(\mathcal{F})} \subseteq \underline{\Sigma_k(\mathcal{F})} \subseteq \underline{B(\Sigma_k(\mathcal{F}))} & \text{for all } k \geq 1. \\
\underline{co\text{-}\Sigma_k(\mathcal{F})} = \underline{\Pi_k(co\text{-}\mathcal{F})} & \text{for all } k \geq 0. \\
\underline{\Sigma_1(\Sigma_k(\mathcal{F}))} \subseteq \underline{\Sigma_k(\mathcal{F})} & \text{for all } k \geq 0. \\
\underline{\Sigma_k^0(\mathcal{F})} \subseteq \underline{\Sigma_k(\mathcal{F})} & \text{if } k \geq 0 \text{ and } FO \subseteq \mathcal{F} = B(\mathcal{F}). \\
\bigcup_{k \geq 0} \underline{\Sigma_k^0(\mathcal{F})} = \underline{FO(\mathcal{F})}. &
\end{array}$$

(For the fourth line the crucial point is that first-order quantifications may be replaced by set quantifications relativized by a first-order formula that asserts for a set that it is a singleton.) All of these closure operations are monotone wrt. set inclusion. Some of such rules are applied e.g. in Remark 3.62.

The closure $\Delta_1^{\mathrm{U}}(\mathcal{F})$ of a formula class $\mathcal{F}$ may be understood as a particular form of the Δ_1-closure of $\mathcal{F}$, where the membership in both $\Sigma_1(B(\mathcal{F}))$ and $\Pi_1(B(\mathcal{F}))$ is due to the fact that the tuple $\overline{X}$ of sets that is quantified over is determined uniquely by a formula in the boolean closure of $\mathcal{F}$. This is formalized by the following proposition.

Proposition 2.6 *Let $\mathcal{C}$ be a class of τ-structures and let $\mathcal{F}$ be a class of formulas. Then*

$$\Delta_1^{\mathrm{U},\mathcal{C}}(\mathcal{F}) \subseteq \Delta_1^{\mathcal{C}}(\mathcal{F}).$$

Proof Let $\varphi(\overline{X},\overline{Y})$ be a formula in $B(\mathcal{F})$ and $\psi(\overline{X},\overline{Y})$ be first-order such that $\psi(\iota\overline{X}(\varphi),\overline{Y})$ is allowed. Then this formula is equivalent over $\mathcal{C}$ to both of the formulas

$$\exists\overline{X}(\varphi \wedge \psi) \quad \text{and} \quad \forall\overline{X}(\varphi \rightarrow \psi),$$

which are in $\Sigma_1(B(\mathcal{F}))^{\mathcal{C}}$ and $\Pi_1(B(\mathcal{F}))^{\mathcal{C}}$, respectively. $\square$

2.3 Words and Pictures

Words

The notations we use for words and formal languages, called "word languages" here, are fairly standard. We recall some of them. Let Γ be an *alphabet*, i.e., a finite set of so-called *letters*. A *word* over Γ is a finite sequence of letters from Γ. We identify a letter a with the word of length 1 that just contains the letter a. The empty word (of length 0) will be denoted by ε. The set of all words (or all non-empty words) over Γ will be denoted by Γ^* (or Γ^+, respectively).

If $w = a_0 \dots a_{n-1}$ is a word of length n, we write $w\langle i\rangle$ instead of a_i for every $i \in \{0, \dots, n-1\}$.

There is a concatenation defined on Γ^*, namely if $u, v \in \Gamma^*$ then $u \cdot v$ is the word over Γ that results by appending v to the right of u. The symbol $\cdot$ is usually omitted. This operation is lifted to languages as usual, that means that if $U, V \subseteq \Gamma^*$, then $UV = U \cdot V = \{uv \mid u \in U, v \in V\}$. This concatenation may be iterated as follows. We let $L^0 = \{\varepsilon\}$ and[1] $L^{k+1} = L \cdot L^k$. Then $L^* = \bigcup_{k \geq 0} L^k$ and $L^+ = \bigcup_{k \geq 1} L^k$. When using these notations, we often identify a word w with the singleton $\{w\}$ so that, e.g., a^*bc^* denotes the language of all words that contain exactly one b, which is preceded by a number of a's and succeeded by a number of c's.

If $w = xyz$ for four (possibly empty) words w, x, y, z, then x, y, z are called *prefix*, *infix*, and *suffix*, respectively, of w. For every word language L, we denote by *pref*(L) (and *pref*$_+(L)$) the set of all prefixes (and all non-empty prefixes, respectively) of words in L.

One notation that is non-standard is the following. For a word language $L \subseteq \Gamma^*$, the *cyclic closure of* L is given by *cycl*$(L) = \{vu \mid u, v \in \Gamma^*, uv \in L\}$.

Pictures

The "two-dimensional" analogue to a word is called a *picture*. Many notations for words can be defined similarly for pictures. We pick out some that are useful for our purposes.

Let $(m,n) \in (\mathbb{N} \times \mathbb{N}) \cup \{(0,0)\}$. Recall that $[n] = \{0, \dots, n-1\}$. A *picture* over Γ of *size* (m,n) is an $m \times n$-matrix over Γ, i.e. a mapping P :

[1] Unfortunately, this notation sometimes conflicts with the k-fold Cartesian product, which is considered a different operation. I hope that the reader can always guess what is meant.

$[m]\times[n] \longrightarrow \Gamma$. For a picture P of size (m,n), we define the *height* of P as $\underline{P} = m$, the *width* of P as $|P| = n$, and $\mathrm{size}(P) = (m,n)$. We write $P\langle i,j\rangle$ (rather than $P(i,j)$) for the component of P at position (i,j). The *empty picture* is the picture with size $(0,0)$ and is denoted by ε.

We use the notations $\Gamma^{*,*}$ (and $\Gamma^{+,+}$) for the set of all (or all non-empty, respectively) pictures over Γ. If $m,n \geq 1$, then $\Gamma^{m,n}$ denotes the set of all pictures of size (m,n) over Γ. If $m \geq 1$ then $\Gamma^{m,+} = \bigcup_{n\geq 1}\Gamma^{m,n}$ and $\Gamma^{m,*} = \Gamma^{m,+} \cup \{\varepsilon\}$.

A set of non-empty pictures is called a *picture language.*

Note that unlike in [GR96] and other publications, the domain of a picture of size (m,n) is of the form $\{0,\dots,m-1\}\times\{0,\dots,n-1\}$ rather than of the form $\{1,\dots,m\}\times\{1,\dots,n\}$. Another difference is that we do not allow the empty picture in "picture languages". See also Remark 3.7.

Definition 2.7 For every non-empty picture P over Γ, let $top(P)$ be the word over Γ in the top row of P, i.e., if P is of size (m,n), then $top(P) = P\langle 0,0\rangle \dots P\langle 0,n-1\rangle$.

As usual, we write $top^{-1}(L)$ for the pre-image of the word language L under this mapping, i.e., $top^{-1}(L)$ is the set of non-empty pictures of Γ whose top row is in L. (We only use this notation when the alphabet Γ is clear from the context.) A picture language of the form $top^{-1}(L)$ (for a word language L) is called *top-pre-image* of L. □

Next we define two partial concatenations for pictures. The first one, called "column concatenation", juxtaposes two pictures next to each other (i.e., concatenates the right column of one picture with the left column of the other) provided they have the same height.

Let P and Q be pictures of size (k,l) and (m,n), respectively. If $k = m \geq 1$, then the column concatenation of P and Q is defined by

$$P \oplus Q : [k]\times[l+n] \longrightarrow \Gamma,\ (i,j) \mapsto \begin{cases} P\langle i,j\rangle & \text{if } j < l, \\ Q\langle i,j-l\rangle & \text{if } j \geq l. \end{cases}$$

Besides, $P \oplus \varepsilon = \varepsilon \oplus P = P$ for every picture P.

These definitions can be extended to sets of pictures as usual, i.e., for $L, M \subseteq \Gamma^{*,*}$ we define $L \oplus M = \{P \oplus Q \mid P \in L, Q \in M\}$. The $\oplus$-symbol is often omitted.

This concatenation can be iterated: For a set $L \subseteq \Gamma^{*,*}$ of pictures we set $L^{1,0} := \{\varepsilon\}$ and $L^{1,k+1} := L^{1,k} \oplus L$. We set $L^{1,*} := \bigcup_{k\geq 0} L^{1,k}$ and $L^{1,+} := \bigcup_{k\geq 1} L^{1,k}$. Since here the first superscript is always 1, we allow to drop it, thus writing L^k instead of $L^{1,k}$ and so on.

A partial row concatenation, denoted $\ominus$, can be defined similarly: for two pictures P, Q of the same width, $P \ominus Q$ is the picture that results from P by appending Q to the bottom. This row concatenation, however, will be used much less frequently than the column concatenation. Its iteration is defined as follows. For a set $L \subseteq \Gamma^{*,*}$ of pictures we set $L^{0,1} := \{\varepsilon\}$ and $L^{k+1,1} = L^{k,1} \ominus L$. We define $L^{*,1} = \bigcup_{k\geq 0} L^{k,1}$ and $L^{+,1} = \bigcup_{k\geq 1} L^{k,1}$.

Words vs. Pictures

A non-empty picture of size (m,n) over alphabet Γ may be viewed as a word over alphabet $\Gamma^{m,1}$ of length n, the so-called *column word.* The set of all column words of pictures of height m of a given picture language L is called the *height-m fragment of L.* These definitions are made formal in Definition 4.1. We will not identify a picture with its column word, but some frequently used notions like "infix" of a picture etc. will advocate this view. For example, if P, Q, R are (possibly empty) pictures for which PQR is defined, then Q is an "infix" and P is a "prefix" of PQR. Furthermore, $cycl(L)$ and $pref_+(L)$ denote the cyclic closure and the set of non-empty prefixes, respectively, of a picture language L.

Conversely, every non-empty word is identified with a picture of height 1 over the same alphabet, and the empty word is identified with the empty picture. Consequently, every word language not containing the empty word is also a picture language.

Coloured Grids vs. Pictures

Recall from Section 2.1 the definition of grids and t-bit grids. To every t-bit grid $([m]\times[n], S_1, S_2, X_1, \dots, X_t)$ we may associate a non-empty picture P over alphabet $\{0,1\}^t$ of the same size, where $P\langle i,j\rangle = (b_1,\dots,b_t) \in \{0,1\}^t$ with $\forall s : b_s = 1 \leftrightarrow (i,j) \in X_s$.

Conversely, from every non-empty picture over alphabet $\{0,1\}^t$ of size (m,n) we may extract an assignment of sets $X_1^P,\dots,X_t^P$ to the set variables $X_1,\dots,X_t$, namely by defining X_s^P to be the set of positions (i,j) for which the s-th component of $P\langle i,j\rangle$ is 1. This way we obtain for every non-empty picture P over $\{0,1\}^t$ the t-bit grid whose associated picture is P.

Considering both t-bit grids and pictures over $\{0,1\}^t$ might seem superfluous or artificial, and indeed we will often allow ourselves to be imprecise in this context. For example, if P is a picture of size (m,n) over $\{0,1\}^t$ and $\varphi(X_1,\ldots,X_t)$ is a formula over signature τ_{Grids}, then we write $P \models \varphi$ instead of $\underline{[m]\times[n]} \models \varphi[X_1^P,\ldots X_t^P]$. We write $Mod_t(\varphi)$ for the set of pictures P over $\Gamma=\{0,1\}^t$ with $P\models\varphi$, i.e., the set of pictures associated to models of φ. If L is a picture language over alphabet Γ and $\mathcal{F}$ a class of formulas over grids, then we say that L is $\mathcal{F}$-definable and write $L \in \underline{\mathcal{F}}$ (or $L\in\underline{\mathcal{F}}(\Gamma)$) iff there is an $\mathcal{F}$-formula $\varphi(X_1,\ldots,X_t)$ over τ_{Grids}, i.e., a sentence φ over τ_{Grids_t}, such that $L = Mod_t(\varphi)$.

However, there is good reason to introduce "pictures" because much terminology from word language theory carries over conveniently then.

"Cyclic"

It turns out that it is sometimes very helpful to think of pictures as if the rightmost column was connected to the leftmost column forming a ring-like stripe. Whenever this ring structure is referred to, we will indicate this by using the word "cyclic". For example, if P, Q, R are pictures for which PQR is defined, then RP will be called a "cyclic infix" of PQR.

Binary Number Representations

Binary number representations play an important rôle for our definability results. That is why we need some convenient notation for it.

Definition 2.8 For every $m,n\geq 0$ with $n\leq 2^m-1$ we write $BIN(n,m)$ for the binary representation of length m of the number n, least significant bit first. Conversely, if w is a word of length m over $\{0,1\}$, then $dual(w)=\sum_{i=0}^{m-1}2^i w\langle i\rangle$ is the number represented by w in binary, least significant bit first. □

It will be technically convenient to consider the binary representation with least significant bit first, i.e., the other way round than usual.

2.3.1 Different Notions of Locality

This subsection is concerned with the transfer of the important notions of locality, locally threshold testability, and recognizability from word languages to picture languages. Let us briefly recall some basic facts about these notions in the world of word languages.

Local Word Languages. A word language L (over some alphabet Γ not containing #) is *local* iff there is a set Δ of words of length 2 over $\Gamma\cup\{\#\}$ such that L is the set of those words w for which all infixes of length 2 of the word $\#w\#$ are in Δ.

Locally Threshold Testable Word Languages. For two words u,w we let $occ(u,w)$ be the number of occurrences of w as an infix of u.

A word language L is *locally threshold testable* iff it is a finite union of equivalence classes of equivalence relations $\cong_{d,t}$, where $u\cong_{d,t}v$ holds for two words u,v and two numbers $d,t\geq 1$ iff for each word w of length $\leq d$ we have $occ(u,w)$ and $occ(v,w)$ are equal or both $\geq t$.

Intuitively, a word language is locally threshold testable iff membership of a word can be decided by counting the numbers of occurrences of infixes of bounded length up to a fixed threshold.

Recognizable Word Languages. An *alphabet projection* is a mapping from one alphabet into another. Such a mapping is lifted to pictures and picture languages as usual.

A word language L (with $\varepsilon\notin L$) is *recognizable* iff one of the following equivalent conditions is true:

1. L is the image of a local word language under some alphabet projection.
2. L is the image of a locally threshold testable word language under some alphabet projection.
3. L is recognizable by a nondeterministic finite automaton.
4. L is definable by a sentence of the existential fragment of monadic second-order logic over word structures (with the successor relation).

Recognizable word languages are also called *regular*. I picked the above choice of characterizations of regularity of word languages because these are the ones that—when adapted to pictures correctly—define the "recognizable picture languages". Other important characterizations of regular word languages are by different versions of regular expressions and by deterministic finite automata. However, these notions fail to generalize to picture languages in such a way that they capture recognizability. See [GR96, Mat95] for details on this subject.

Locality

Now we define when to call a picture language "local" or "generalized local". These definitions will be compatible with our agreement to identify word languages with picture languages all of whose elements have height 1.

Let P be a picture of size (m,n) over Γ. If $0 \le i \le i' \le m-1$ and $0 \le j \le j' \le n-1$, we write $P([i,i']\times[j,j'])$ for the picture P' of size $(i'-i+1, j'-j+1)$ over Γ defined by $P'\langle x,y\rangle = P\langle i+x, j+y\rangle$ for every $(x,y) \in [0,i'-i]\times[0,j'-j]$. The picture P' is called a *subblock of* P.

Let $\#$ be a new symbol not in Γ. By $\hat{P}$ we denote the picture of size $(m+2, n+2)$ over $\Gamma \cup \{\#\}$ that results from P by surrounding it with the symbol $\#$, i.e.,

$$\hat{P}\langle i,j\rangle = \begin{cases} P\langle i-1, j-1\rangle & \text{if } (i,j) \in [1,m]\times[1,n], \\ \# & \text{if } i \in \{0, m+1\} \vee j \in \{0, n+1\}. \end{cases}$$

A picture language L is *local* if there is a set $\Delta \subseteq (\Gamma \cup \{\#\})^{2,2}$ such that L is the set of non-empty pictures P over Γ such that all 2×2-subblocks of $\hat{P}$ are in Δ. In that case we write $\mathcal{L}(\Delta) = L$. We call Δ a *local tiling system* and its elements *tiles*.

A picture language L is *generalized local* if there is a finite set F of "forbidden subblocks", i.e., F is a finite set of pictures of arbitrary size over $\Gamma \cup \{\#\}$ such that L is the set of non-empty pictures P over Γ such that no subblock of $\hat{P}$ is in F.

Obviously, every local picture language is generalized local. A picture language is *recognizable* if it is the image of a local picture language under some alphabet projection. See Section 4.2.

Two simple examples of local pictures languages follow.

Example 2.9 Consider the alphabet $\{a,b\}$ and let Δ be the set of all 2×2-subblocks of the pictures

$$\begin{matrix} \# & \# & \# & \# \\ \# & a & b & \# \\ \# & b & a & \# \\ \# & \# & \# & \# \end{matrix} \quad \text{and} \quad \begin{matrix} \# & \# & \# & \# \\ \# & b & a & \# \\ \# & a & b & \# \\ \# & \# & \# & \# \end{matrix}.$$

Then $\mathcal{L}(\Delta)$ is the set of all non-empty pictures P over $\{a,b\}$ for which the set $P^{-1}(a)$ of a-positions is either the set of positions (i,j) for which $i+j$ is even or the set of positions where this is odd.

$\mathcal{L}\left(\Delta \setminus \left\{ \begin{smallmatrix} \# & \# \\ \# & b \end{smallmatrix}, \begin{smallmatrix} b & \# \\ \# & \# \end{smallmatrix} \right\}\right)$ is the set L_1 of all non-empty pictures P over $\{0,1\}$ for which, firstly, $\underline{P} + |P|$ is even and, secondly, for all $(i,j) \in \mathrm{dom}P$ we have $P\langle i,j\rangle = a$ iff $i+j$ is even.

The set L_2 of non-empty pictures P over $\{a\}$ for which $\underline{P} + |P|$ is even is the image of the local picture language L_1 under the alphabet projection $\{a,b\} \to \{a\}$, $x \mapsto a$. Thus L_2 is recognizable. □

Example 2.10 Consider

$$L = \{P \in \{0,1\}^{+,+} \mid \forall (i,j) \in \mathrm{dom}P : i \neq 0 \to P\langle i,j\rangle = 0\},$$

i.e., L is the set of all non-empty pictures over $\{0,1\}$ that have no 1 in a non-top row. Then L is local because

$$L = \mathcal{L}\left(\left\{ \begin{smallmatrix} a & c \\ b & d \end{smallmatrix} \in \{0,1,\#\}^{2,2} \;\middle|\; b,d \neq 1 \vee a,c = \# \right\}\right).$$

□

Another variant of locality is defined in Subsection 6.2.2.

The following fact is easy to prove.

Remark 2.11 Every generalized local picture language is first-order definable (in the signature with two binary successor relation symbols, but without orderings). □

This is a special case of a more general result of [GRST96] that states that the class of first-order definable picture languages coincides with the class of *locally threshold testable* picture languages. Though this notion is not needed in this thesis, we introduce it here for the interested reader. It is a straightforward adaptation of the corresponding notion for word languages, see page 33.

Definition 2.12 (cf. [GR92, GRST96, GR96]) For two pictures P, R we let $occ(P,R)$ be the number of occurrences of R as a subblock of P. A picture language L is *locally threshold testable* iff it is a finite union of equivalence classes of equivalence relations $\cong_{d,t}$, where $P \cong_{d,t} Q$ holds for two pictures P, Q and two numbers $d, t \ge 1$ iff for each picture R of height and width $\le d$ we have that $occ(P,R)$ and $occ(Q,R)$ are equal or both $\ge t$. □

Cyclic Locality

Recall that the word "cyclic" indicates that one should imagine the rightmost column of a picture as connected to the leftmost column. With this intuition, it is straightforward to deduce a definition of "cyclically local picture language" and the like.

Definition 2.13 (Cyclic Subblocks) Let P and Q be pictures. Q is a *cyclic subblock of* P iff it is a subblock of some picture in $cycl(\{P\})$. A *properly cyclic subblock* of P is a picture of the form $P([i,i']\times[j',n-1]) \oplus P([i,i']\times[0,j])$, where $0 \le i \le i' \le \overline{P}$ and $0 \le j < j' \le |P|$. □

Obviously, a picture Q is a cyclic subblock of P iff it is a subblock or a properly cyclic subblock.

Example 2.14 The words ca as well as ab are (as 1×2-pictures) cyclic subblocks of abc. The word $abca$ is not a cyclic subblock of abc. □

Let P be a picture of size (m,n) over Γ and $\#$ a new symbol not in Γ. By $\tilde{P}$ we denote the picture of size $(m+2,n)$ over $\Gamma\cup\{\#\}$ that results from P by attaching one row of #'s to the top and one to the bottom, i.e.,

$$\tilde{P}\langle i,j\rangle = \begin{cases} P\langle i-1,j\rangle & \text{if } (i,j)\in[1,m]\times[0,n-1], \\ \# & \text{if } i\in\{0,m+1\},\ j\in[0,n-1]. \end{cases}$$

A picture language L is *cyclically local* if there is a set $\Delta \subseteq (\Gamma\cup\{\#\})^{2,2}$ such that L is the set of non-empty pictures P over Γ such that Δ *cyclically tiles* $\tilde{P}$, i.e., such that all cyclic 2×2-subblocks of $\tilde{P}$ are in Δ.

A picture language L is *generalized cyclically local* if there is a finite set F of "forbidden cyclic subblocks", i.e., F is a set of pictures over $\Gamma\cup\{\#\}$ such that L is the set of non-empty pictures P over Γ such that no cyclic subblock of $\tilde{P}$ is in F.

The next example illustrates the definition of cyclic locality. The idea of a counting mechanism, which plays a crucial rôle here, will be important later on.

Example 2.15 For a picture $C \in \{0,1\}^{m,1}$, we have that $dual(C^\top) = \sum_{l=0}^{m-1} C\langle l,0\rangle \cdot 2^l$, i.e., the number represented by C in binary, where the least significant bit is at the top.

Let L be the language of non-empty pictures P over $\{0,1\}$ for which $dual(C'^\top) = dual(C^\top)+1 \bmod 2^{\overline{P}}$ for every two cyclically successive columns C, C' of P. Then L is a cyclically local picture language. To see this, let Δ be the set of cyclic 2×2-subblocks of

$$\begin{matrix} \# & \# & \# & \# & \# & \# & \# & \# \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ \# & \# & \# & \# & \# & \# & \# & \# \end{matrix}.$$

Then L is indeed the set of pictures P for which Δ cyclically tiles $\tilde{P}$. □

The essential point for the proof of the above statement is that when one passes from the binary representation of some number to that of its successor (modulo the 2-power of the length) then this process can be described locally: Flip the least significant digit and then check the other digits in order of increasing significance; every position whose digit changes from zero to one means a "carry", which causes the following digit to be flipped (cf. Lemma 3.26).

In Section 3.3, Definition 3.22, we will come back to this example. We will introduce a picture language that is very similar to the above. In that section we will also prove formally that this cyclic counting process results indeed in a cyclically local picture language.

We conclude with two simple propositions, whose proofs are left to the reader.

Proposition 2.16 *Let Γ be an alphabet. If L is a local (generalized local, cyclically local, generalized cyclically local, recognizable) word language, then $top^{-1}(L)$ is a picture language with the respective property.*

Proposition 2.17 *Let $\alpha : \Gamma \to \Omega$ be a alphabet projection. If $L \subseteq \Omega^{+,+}$ is a local (or cyclically local, or locally threshold testable, respectively) picture language, then so is $\alpha^{-1}(L)$.*

2.3.2 Alternation Hierarchies for Starfree Picture Languages

Starfreeness

In formal word language theory, one important subclass of the class of recognizable word languages is the class of *starfree* word languages. It is the smallest class of word languages (over a fixed alphabet) that contains all finite word languages and is closed under concatenation, union, and complement (relative to the set of all words over the fixed alphabet).

This notion can be easily transferred to picture languages. The set of *starfree picture languages over* Γ (denoted $SF(\Gamma)$) is the smallest class of picture languages over alphabet Γ that contains all finite picture languages and is closed under both concatenations $\ominus$ and $\oplus$ as well as under union and complement (relative to the set of all non-empty pictures over Γ).

Since the set of pictures of height 1 over a fixed alphabet is a starfree picture language, the notion of starfreeness is compatible with our identification of words with pictures of height 1.

The following proposition from [Mat98b] gives an example of a starfree picture language.

Proposition 2.18 *Let Corners be the set of non-empty pictures P over $\{0,1\}$ such that whenever $P\langle i,j\rangle = P\langle i',j\rangle = P\langle i,j'\rangle = 1$ then also $P\langle i',j'\rangle = 1$. (Intuitively: Whenever three corners of a rectangle carry a 1, then also the fourth one does.) Corners is a starfree picture language.*

Proof Let $L := \bigcup (w\,(\{0,1\}^{*,*})\,x) \ominus (\{0,1\}^{*,*}) \ominus (y\,(\{0,1\}^{*,*})\,z)$, where the union ranges over all quadruples $(w,x,y,z) \in \{0,1\}^4$ such that $wxyz \in 1^*01^*$. Then L is the set of all pictures over $\{0,1\}$ such that exactly one of its four corners carries an 0.

Clearly, $(\{0,1\}^{*,*}) \ominus ((\{0,1\}^{*,*})\,L\,(\{0,1\}^{*,*})) \ominus (\{0,1\}^{*,*})$ is the complement of *Corners*, so *Corners* is starfree. (The use of the empty picture is avoidable by Remark 3.7.) □

In Subsection 4.2.1 we will show that this starfree picture language is not recognizable.

There is a certain relation between the starfree word languages and first-order logic: the starfree word languages are exactly those which are definable by a formula in first-order logic with the built-in order relation. This indicates that "starfreeness" is a robust notion for word languages.

However, [Wil97] shows that the notion of "starfreeness" is not as robust for picture languages. The author gives an example of a non-starfree picture language that is definable in first-order logic with the two partial successor relations (cf. Fact 1.17).

It turns out that nevertheless the class of starfree picture languages is rich enough to contain picture languages beyond arbitrarily high levels of

the monadic hierarchy, which is one of our main results. The proof shows that there is a certain relation between the alternation of quantifiers (or, equivalently, of existential quantifiers and negations) on one side and the alternation of concatenation and negations on the other. The investigation of this relation is prepared now.

Alternation

Recall from the introduction that we plan to prove all of our separation results uniformly, both those that involve starfreeness and those that deal with monadic second-oder logic only.

We define two slightly different alternation hierarchies for starfree picture languages. Later we will prove that these hierarchies "cut through" the monadic alternation hierarchy in the sense that in level $k+1$ of these alternation hierarchies for starfree picture languages, there is a picture language that is not on level k of the monadic alternation hierarchy.

Let $\mathcal{L}$ be a class of picture languages over some fixed alphabet. The $\oplus$-*closure*, the $\oplus\ominus$-*closure*, the $\cup\cap$-*closure*, and the *boolean closure of* $\mathcal{L}$ are defined as the smallest superset of $\mathcal{L}$ that is closed under $\oplus$-concatenations, or under both $\ominus$- and $\oplus$-concatenations, or under both union and intersection, or under both union and complement (relative to $\Gamma^{+,+}$), respectively. We denote these classes by $\oplus$-cl($\mathcal{L}$), $\oplus$,$\ominus$-cl($\mathcal{L}$), $\cap$,$\cup$-cl($\mathcal{L}$), and $B(\mathcal{L})$, respectively.

The following definition is a straightforward adaptation of the dot-depth hierarchy of starfree word languages.

Definition 2.19 (Dot-Depth Hierarchy) *dot-depth*$_0(\Gamma)$ is the class of all finite or cofinite picture languages over Γ, and

$$\textit{dot-depth}_{k+1}(\Gamma) = B(\oplus,\ominus\text{-cl}(\textit{dot-depth}_k(\Gamma)))$$

for every $k \geq 0$. □

Intuitively, a picture language over alphabet Γ is in *dot-depth*$_k(\Gamma)$ iff it can be constructed from finite picture languages by k alternations of (row-/column-) concatenation on one side and boolean combinations on the other side.

The dot-depth hierarchy of word languages is a well studied hierarchy (see e.g. [BK71, Pin86]), which is strict and exhausts the class of starfree word languages. The strictness of the dot-depth hierarchy of picture languages can be deduced easily from the strictness of the dot-depth hierarchy of word languages because the dot-depth levels are compatible with our identification of words with pictures of height 1.

Recall that our aim was to find a suitable hierarchy of starfree picture languages that "cuts through" (and allows to deduce the infiniteness of) the monadic hierarchy. For this, the dot-depth hierarchy is not suitable because its first level does not capture first-order expressiveness (in the signature τ_{Grids}). Even if we modified the definition in such a way that level 0 (or level 1) contained exactly the first-order definable (or, by [GRST96] equivalently, the locally threshold testable) picture languages, then still every level would be closed under boolean combinations, which is not the case for the monadic hierarchy.

So if we want to deduce also facts like "$\underline{\Sigma}_k \neq \underline{\Pi}_k$" from the result that our alternation hierarchy of starfree picture languages "cuts through" the monadic hierarchy, we need a version of alternation hierarchy that gives reasonable distinctions between Σ- and Π-branches.

I suggest the following.

Definition 2.20 ((Generalized) Local Alternation Hierarchy) Let Γ be an alphabet.

Let $\Sigma_0^{\mathrm{loc}}(\Gamma)$ denote the smallest class of picture languages over Γ that contains all local and all cyclically local picture languages and all *top*-pre-images of locally threshold testable word languages, and that is closed under boolean combinations.

Let $\Sigma_0^{\mathrm{gloc}}(\Gamma)$ denote the smallest class of picture languages over Γ that contains all generalized local and all generalized cyclically local picture languages and all *top*-pre-images of locally threshold testable word languages, and that is closed under boolean combinations.

For every $k \geq 0$, let

$$\begin{aligned}\Sigma_{k+1}^{\mathrm{loc}}(\Gamma) &= \cap,\cup\text{-cl}(\oplus\text{-cl}(B(\Sigma_k^{\mathrm{loc}})))(\Gamma),\\ \Sigma_{k+1}^{\mathrm{gloc}}(\Gamma) &= \cap,\cup\text{-cl}(\oplus\text{-cl}(B(\Sigma_k^{\mathrm{gloc}})))(\Gamma).\end{aligned}$$

For every $k \geq 0$, let $\Pi_k^{\mathrm{loc}}(\Gamma)$ denote the class of picture language that contains all complements of languages in $\Sigma_k^{\mathrm{loc}}(\Gamma)$, and $\Delta_k^{\mathrm{loc}}(\Gamma) = \Sigma_k^{\mathrm{loc}}(\Gamma) \cap \Pi_k^{\mathrm{loc}}(\Gamma)$, and likewise for the decoration gloc instead of loc.

We omit the explicit mentioning of Γ if it is clear from the context.

Σ_k^{loc} and Σ_k^{gloc} are called the k-th level of the *local alternation hierarchy* and the *generalized local alternation hierarchy*, respectively. □

Obviously, $\Sigma_k^{\mathrm{loc}} \subseteq \Sigma_k^{\mathrm{gloc}}$ for every $k \geq 0$.

Note that in these two hierarchies, the row concatenation is completely disregarded. Thus they certainly do not exhaust the starfree picture languages, i.e., there are starfree picture languages which are in none of the Σ_k^{gloc}. (Just take the transposed of a starfree but not generalized local word language. An example of such a word language is the set of all words over $\{a, b, c\}$ for which for every two occurrences of a's there is a c in between.)

There are several ways to modify the definition of this hierarchy in order to make it look more familiar. Firstly, one could choose to replace the $\oplus$-closure in the definition of Σ_k^{gloc} by the $\oplus\ominus$-closure. Then the resulting hierarchy would exhaust the class of starfree picture languages.

Secondly, level zero could be modified to contain also all first-order definable (i.e., all locally threshold testable) picture languages.

A third possible modification is discussed on page 66.

However, let me assure the reader that there is good reason to consider exactly these two hierarchies from Definition 2.20. The most important one is that by considering these two hierarchies, we can state (and prove) the strongest separation results we have. For example, if we replaced the $\oplus$-closure by the $\oplus\ominus$-closure, then this would clearly make the levels larger and thus our separation results weaker. If we modified the definition of level 0 as indicated above, then we would lose the close connection to the dot-depth hierarchy, which is made explicit in Proposition 2.21.

Proposition 2.21 *$B(\Sigma_k^{\mathrm{gloc}})(\Gamma) \subseteq$ dot-depth$_{k+1}(\Gamma)$ for every $k \geq 0$.*

This proposition is shown in Subsection 3.1.2. In Subsection 3.1.3 we will show that $\Sigma_k^{\mathrm{gloc}} \subseteq \underline{\Sigma_k^0(\Delta_1^{\mathrm{U}})} \subseteq \underline{\Sigma}_k$ for every $k \geq 1$.

Syntax vs. Semantics

Since this thesis concentrates on expressiveness results, I chose not to introduce unnecessary syntactic notions. However, the reader who is used to such notions as "regular" or "starfree expressions" may easily fill this gap. There are a few remarks on the lengths of starfree expressions in Subsection 3.3.1.

One even could have avoided introducing formulas on a syntactic level.

2.4 Separation Results

In this section we present all separation results we have. The proofs of these results will be presented in the following chapters. At the end of this section I give some hints on which parts of these subsequent chapters are relevant for which results.

Recall that if Γ is an alphabet of the form $\{0,1\}^t$, we write $\underline{\Sigma_k}(\Gamma)$ for the class of Σ_k-definable picture languages over Γ.

We have the following main theorem, which says that the local alternation hierarchy for picture languages cuts through the monadic second-order quantifier alternation hierarchy for picture languages. It is shown in Section 3.3 (Corollary 3.38) and Section 4.3 (Corollary 4.14).

Theorem 2.22 *For every $k \geq 1$,*

$$\Pi_k^{\mathrm{loc}}(\{0,1\}^{2k}) \not\subseteq \underline{\Sigma_k}(\{0,1\}^{2k}),$$

i.e., there is a starfree picture language over alphabet $\{0,1\}^{2k}$ in the complement of the k-th level of local alternation hierarchy which is not in the k-th level of the monadic second-order quantifier alternation hierarchy.

This theorem allows us to separate several other classes from $\underline{\Sigma_k}$, for instance all that are depicted in Figure 2.1.[2] In particular, $\Sigma_{k+1}^{\mathrm{loc}}(\{0,1\}^{2k}) \supseteq \Pi_k^{\mathrm{loc}}(\{0,1\}^{2k}) \not\subseteq \underline{\Sigma_k}(\{0,1\}^{2k})$, which can be interpreted in the following way: the alternation of column concatenations and complementation of picture languages is so powerful that it cannot be captured by a smaller number of alternations of monadic quantification.

Theorem 2.22 also implies the following.

Corollary 2.23 *For every $k \geq 1$ there is a formula φ in Π_k over grids which is not equivalent to any formula in Σ_k.*

Proof Let $k \geq 1$. Then $\underline{\Pi_k}(\{0,1\}^{2k}) \supseteq \Pi_k^{\mathrm{gloc}}(\{0,1\}^{2k}) \supseteq \Pi_k^{\mathrm{loc}}(\{0,1\}^{2k}) \not\subseteq \underline{\Sigma_k}(\{0,1\}^{2k})$ by Theorem 3.12. Thus there is a set of $2k$-bit grids definable by a formula $\varphi(X_1, \dots, X_{2k})$ (over signature τ_{Grids}) in Π_k but by no formula in Σ_k. □

[2]The inclusion diagram of Figure 2.1 remains true if the bottom-most class is replaced by Π_k^{loc}.

Recall that we identify a formula $\varphi(X_1, \dots, X_t)$ in the signature τ_{Grids} of grids with a sentence in the signature τ_{Grids_t} of t-bit grids. Thus we have shown that the monadic alternation hierarchy over coloured grids is strict, i.e., for every k there is a t such that Σ_k is properly less expressive than Σ_{k+1} over t-bit grids, which is the main result of [MT97].

But note that Theorem 2.22 and Corollary 2.23 do not limit the size of the alphabet $\{0,1\}^{2k}$ (i.e., the number $2k$ of free set variables). They do not imply that there is a fixed number t of colours such that the monadic alternation hierarchy over t-bit grids is strict.

The subsequent two subsections collect the results we have in the cases $t = 1$ and $t = 0$, respectively. They also give some separations that refer to other fragments of monadic second-order logic.

He who is only interested in separation results over finite structures or over the class of graphs might be satisfied with Corollary 2.23 because in Chapter 5 we will show how this statement can be transferred from formulas over grids to sentences over graphs.

However, it is justified and possible to investigate the situation for the class of grids more precisely. This will also allow us to deduce some more separation results for the class of graphs, for instance that for every k there is a Δ_{k+1}-sentence over graphs which is not equivalent to any $B(\Sigma_k)$-sentence. This has been shown first in [MST99], combining results from [MT97] and [Sch97a].

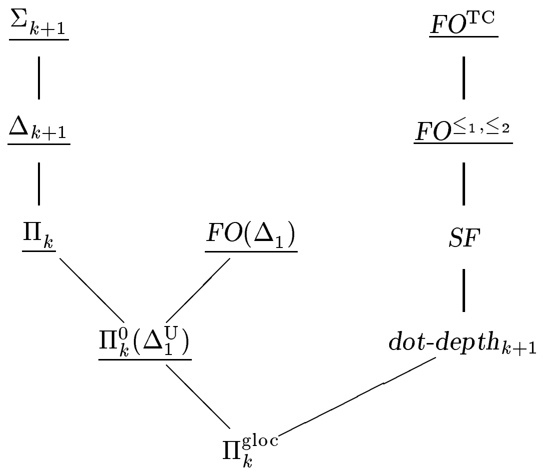
2.4.1 Separation Results for Pictures over a Two-Letter Alphabet

The following central result states that Theorem 2.22 can be saved for the fixed alphabet $\{0,1\}$ if the local alternation hierarchy is replaced by the generalized local alternation hierarchy. It follows from Theorem 2.22 by a standard encoding technique, see Subsection 3.1.1.

Theorem 2.24 *For every $k \geq 1$,*

$$\Pi_k^{\mathrm{gloc}}(\{0,1\}) \not\subseteq \underline{\Sigma_k}(\{0,1\}),$$

i.e., there is a starfree picture language over alphabet $\{0,1\}$ in the complement of the k-th level of the generalized local alternation hierarchy which is not in the k-th level of the monadic second-order quantifier alternation hierarchy.



For the $k \geq 1$ under consideration, the bottom-most class—and therefore all the others—are not contained in $\underline{\Sigma_k}$, provided that the underlying alphabet has at least two symbols. In this diagram lines indicate (not necessarily proper) inclusions. A missing line does not indicate incomparability.

Figure 2.1: Picture language classes separated from $\underline{\Sigma_k}$

Figure 2.1 shows some superclasses of Π^{gloc}_k which are therefore no subsets of $\underline{\Sigma_k}$ either. The inclusions depicted in that figure are either trivial (like $\underline{FO^{\leq_1,\leq_2}} \subseteq \underline{FO^{\mathrm{TC}}}$) or proved in Subsections 3.1.2 and 3.1.3. The inferred separations are summarized in the following corollary.

Corollary 2.25 *For every $k \geq 1$, the following classes contain a picture language over $\{0,1\}$ that is not in $\underline{\Sigma_k}$:*

- *$dot\text{-}depth_{k+1}$ and thus SF, $\underline{FO^{\leq_1,\leq_2}}$, and $\underline{FO^{\mathrm{TC}}}$;*
- *$\underline{\Pi^0_k(\Delta^{\mathrm{U}}_1)}$ and thus $\underline{FO(\Delta_1)}$ as well as $\underline{\Pi_k}$, $\underline{\Sigma_{k+1}}$, and $\underline{\Delta_{k+1}}$.*

The non-containment $\underline{\Sigma_k}(\{0,1\}) \not\subseteq \underline{\Pi_k}(\{0,1\})$ says that in the statement of Corollary 2.23, the number of free variables of φ can be limited to one. Theorem 2.32 will state that it can even be limited to zero.

2.4.2 Separation Results for (Non-Coloured) Grids

In this subsection we shall summarize the separation results we have for pictures over a one-letter alphabet. The investigation of starfree picture languages does not make much sense because the class of starfree picture languages over a singleton alphabet is very poor—in particular, every starfree language over a one-letter alphabet is Σ_1-definable, see e.g. [Mat95]. However, there are non-trivial separations concerning the expressiveness of certain fragments of monadic second-order logic. In this context it is more canonical to refer to "grids" (i.e., particular finite structures) rather than "pictures over a trivial alphabet".

The Growth Technique

Recall that for a set L of grids, size(L) is the set of all (m,n) such that the grid of size (m,n) is in L.

Our separation results for the class of non-coloured grids are of two types. The first one is concerned with classes L of grids where size(L) is a function, i.e., for every m there is exactly one grid of height m in L. Separations of the first type are then witnessed by k-fold exponential functions for increasing k. This has been called "the growth technique" in [AFS97].

The second type of separation results is by classes of grids L where size(L) is a non-functional relation of $\mathbb{N}$. The separation $\underline{\Pi_k} \not\subseteq \underline{\Sigma_k}$ is witnessed by a set L of grids where the function $m \mapsto \min\{n \mid (m,n) \in \text{size}(L)\}$ is $(k+1)$-fold exponential. This is further explained on page 50ff.

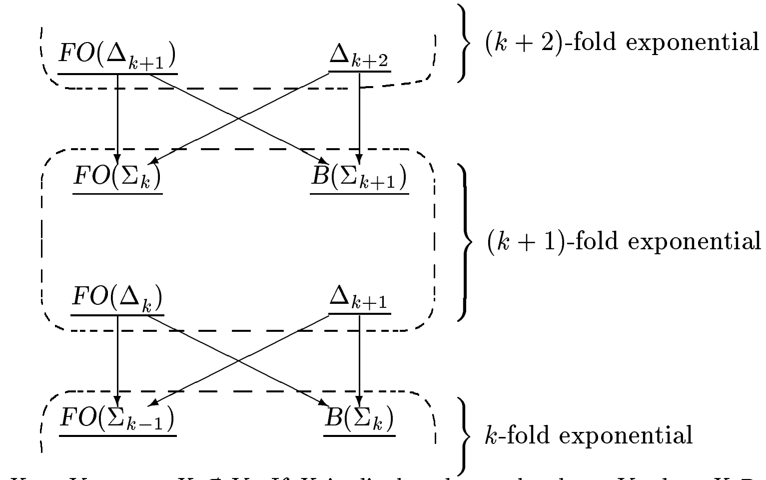
Definition 2.26 A sentence φ over τ_{Grids} *defines* a relation $r \subseteq \mathbb{N}^2$ iff $r = \{(m,n) \mid \underline{[m]\times[n]} \models \varphi\}$. A relation r is $\mathcal{F}$-definable (for a class of formulas $\mathcal{F}$ over τ_{Grids}), if there is a sentence $\varphi \in \mathcal{F}$ that defines r.

Let $f\colon \mathbb{N} \dashrightarrow \mathbb{N}$ be a partial function. Then f is called *at most k-fold exponential* if $f(m)$ is $s_k(\mathcal{O}(m))$, where $s_0(m) = m$ and $s_{k+1}(m) = 2^{s_k(m)}$ for every $m \geq 1$, $k \geq 0$.

We call f *at least k-fold exponential* if f is total and $f(m)$ is $s_k(\Omega(m))$. If f is both at most and at least k-fold exponential, i.e, if f is total and $f(m)$ is $s_k(\Theta(m))$, then we say that f is *k-fold exponential.* □

The following notion is convenient to state some of our results succinctly.

Definition 2.27 (Asymptotic Bounds for Formula Classes) Let $\mathcal{F}$ be a class of formulas over τ_{Grids}. We say that



$X \to Y$ means $X \not\subseteq Y$. If X is displayed exactly above Y, then $X \supseteq Y$ is valid for trivial reasons. Classes encircled by dashed lines allow for the same rate of growth of functions—nothing is claimed about mutual inclusions!

Figure 2.2: Separation results for grids

- "$\mathcal{F}$ is at most k-fold exponential" iff f is at most k-fold exponential for every $\mathcal{F}$-definable function f;
- "$\mathcal{F}$ is at least k-fold exponential" iff f is at least k-fold exponential for some $\mathcal{F}$-definable function f;
- "$\mathcal{F}$ is k-fold exponential" iff $\mathcal{F}$ is both at most and at least k-fold exponential.

$\square$

Example 2.28 In [Gia94], the author shows that Σ_1 is singly (i.e. 1-fold) exponential. This means that every Σ_1-definable function is at most singly exponential, and there is one particular singly exponential, Σ_1-definable function (namely $m \mapsto 2^m$). $\square$

There is a more detailed investigation of the class of Σ_1-definable functions in [Gia94]. It is shown that this class is closed under certain operations and deduced that, for example, every polynomial with non-negative integer coefficients is Σ_1-definable.

We extend Example 2.28 to higher levels of the monadic alternation hierarchy and show that Σ_k is k-fold exponential for every $k \geq 1$. Besides, it turns out that some quite small "fragments" of Σ_k are k-fold exponential, too. Formally, we have the following fact. (It follows from Corollaries 3.49 and 4.22, but a weaker form follows already from the (easier) Corollaries 3.41 and 4.16.)

Theorem 2.29 *For every $k \geq 1$, the following formula classes are k-fold exponential: $\Delta_1^{\mathrm{U}}(\Pi^0_{k-1}(\Delta_1^{\mathrm{U}}))$, $\Delta_1^{\mathrm{U}}(\Pi_{k-1})$, Δ_k, Σ_k, $B(\Sigma_k)$.*

The upper bounds of this theorem, namely that $B(\Sigma_k)$ (and thus the other classes, less expressive classes) are *at most* k-fold exponential, have been shown in [MT97]. The lower bound for $\Delta_1^{\mathrm{U}}(\Pi_{k-1})$ (and thus for Δ_k, Σ_k, $B(\Sigma_k)$) has been shown by Schweikardt (see [Sch97a]), improving the techniques of [MT97], which had only rendered the weaker result that Σ_{2k+3} is at least k-fold exponential. (Actually, in [Sch97a] the author considers only $\Delta_1(\Pi_{k-1})$, but her proof can be modified to show that the outermost quantification is unique and may thus be replaced by a ι-application.)

The following gives asymptotic growth rates for some more formula classes. (It follows from Corollary 3.66 and Theorem 4.25.)

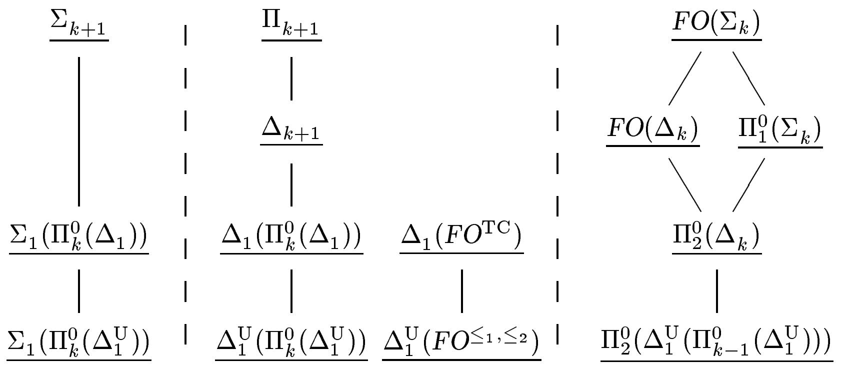
Theorem 2.30 *Let $k \geq 1$. The following classes are $(k+1)$-fold exponential: $\Pi^0_2(\Delta_1^{\mathrm{U}}(\Pi^0_k(\Delta_1^{\mathrm{U}})))$, $\Pi^0_j(\Delta_k)$ (for every $j \geq 2$) and $FO(\Sigma_k)$.*

The upper bound for $FO(\Sigma_k)$ (and thus for the other, smaller classes) and the lower bound for $FO(\Sigma_k)$ have been shown in [Mat98a].

Figure 2.2 illustrates some of the separation results stated in the above two theorems for grid classes separated by asymptotic growth rates.

Referring to Theorem 2.29, the above may be rephrased as follows: *With respect to the growth rate of functions*, the first-order closure is as powerful as 'one more' alternation in the monadic hierarchy. However, this should be read with some caution because the focus of attention to asymptotic growth rate bounds is quite essential here—it is by no means true that any alternation of first-order quantifications can be replaced by one block of monadic quantifications, probably not even for sentences that define functions. See also Corollary 2.31.

It is a challenging unsolved problem to prove that $\underline{\Pi^0_j(\Sigma_k)} \subsetneqq \underline{\Pi^0_{j+1}(\Sigma_k)}$ over grids for every j, k. We only know that this holds for 1-bit grids and $k = 1$.



The bottom-most classes—and therefore all the others—are not contained in $\underline{\Sigma_k}$, not in $\underline{B(\Sigma_k)}$, and neither in $\underline{FO(\Sigma_{k-1})}$.

Figure 2.3: Inclusion diagram for grid classes separated from $\underline{B(\Sigma_k)}$

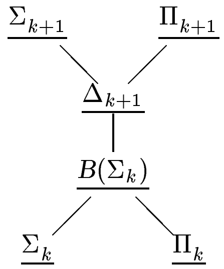
The separations witnessed by fast growing function are summarized in Corollary 2.31.

Corollary 2.31 *For every $k \geq 0$, the following classes of formulas allow to define a set of grids that is not $B(\Sigma_k)$-definable and (if $k \geq 1$) neither $FO(\Sigma_{k-1})$-definable:*

- $\Sigma_1(\Pi^0_k(\Delta^{\mathrm{U}}_1))$ *and thus* $\Sigma_1(\Pi^0_k(\Delta_1))$ *and* Σ_{k+1}*;*
- $\Delta^{\mathrm{U}}_1(\Pi^0_k(\Delta^{\mathrm{U}}_1))$ *and thus* $\Delta_1(\Pi^0_k(\Delta_1))$, Π_{k+1} *and* Δ_{k+1}*;*
- $\Delta^{\mathrm{U}}_1(FO^{\leq_1,\leq_2})$ *and thus* $\Delta^{\mathrm{U}}_1(FO^{\mathrm{TC}})$*;*
- *if $k \geq 1$, also* $\Pi^0_2(\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)))$ *and thus* $\Pi^0_2(\Delta_k)$, $FO(\Delta_k)$, $\Pi^0_1(\Sigma_k)$, $FO(\Sigma_k)$.

Figure 2.3 shows an extraction from the inclusion diagram of those classes of sets of grids (or classes of picture languages over trivial alphabet) that are by Corollary 2.31 not a subset of $\underline{B(\Sigma_k)}$. Note that this is not meant to be a complete inclusion diagram. The meaning of the dashed lines is explained in Subsection 2.4.4.

Among the many separations of the preceding corollary, let me comment on a few which I consider to be of major importance.



Lines indicate proper inclusions.

Figure 2.4: The monadic hierarchy over grids and graphs

The central claim of the preceding theorems is the separation "$\underline{\Sigma_k} \subsetneqq \underline{\Sigma_{k+1}}$", i.e., the "strictness of the monadic hierarchy". It has first been shown for the classes of graphs and coloured grids in [MT97], but the stronger result that this is true even for the class of non-coloured grids is due to [Sch97a].

Schweikardt's results together with [MT97] imply the stronger Fact 1.7, i.e., "$\underline{B(\Sigma_k)} \subsetneqq \underline{\Delta_{k+1}}$ for the class of coloured grids" and thus for the class of graphs. This completes, for the cases of graphs and 1-bit grids, the picture reminding of the classical arithmetical hierarchy. (See Figure 2.4).

The result "$\underline{FO(\Sigma_k)} \subsetneqq \underline{FO(\Sigma_{k+1})}$ for the class of grids", announced in Fact 1.12, may be rephrased informally as follows: The monadic hierarchy quantifier remains strict when one considers the first-order closure of each level. However, this result should not be overestimated because the first-order closure of Σ_k is defined somewhat artificially in the sense that—in contrast to the levels of closed hierarchy—first-order quantifications are added only in the outermost level of second-order quantifications.

The result that "$\underline{\Delta_{k+2}} \not\subseteq \underline{FO(\Sigma_k)}$ for the class of grids", announced in Fact 1.12, too, may be viewed as an indication that the first-order closure is not as powerful as two additional monadic quantifier alternations, whereas the result "$\underline{FO(\Delta_1)} \not\subseteq \underline{\Sigma_k}$" (for $k \geq 1$) from Corollary 2.25 asserts that the first-order closure is so powerful that it lifts even the lowest level of the monadic hierarchy beyond any level of the monadic hierarchy.

$$\underline{\Pi_k}$$
$$|$$
$$\underline{\Pi^0_1(\Delta_k)}$$
$$|$$
$$\underline{\Pi^0_1(\Delta_1(\Pi^0_{k-1}(\Delta_1)))}$$
$$|$$
$$\underline{\Pi^0_1(\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)))}$$

These classes are not contained in $\underline{\Sigma_k}$. Lines indicate (not necessarily proper) inclusions.

Figure 2.5: Inclusion diagram for grid classes separated from $\underline{\Sigma_k}$

Separation Results With Non-Functional Relations

While the previous results imply the situation depicted in Figure 2.4 for the classes of 1-bit grids and thus for the class of graphs, it remained open in [MT97, Sch97a, MST99] how the situation is for the class of non-coloured grids. We know that $\underline{B(\Sigma_k)} \subsetneq \underline{\Delta_{k+1}}$ for the class of grids by Theorem 2.29, but the same theorem tells us that if there is a *function* f witnessing that $\underline{\Sigma_k}$ and $\underline{\Pi_k}$ are incomparable, then the reason is not only the asymptotic growth rate of f.

However, this separation result can be shown using a witness set that is not a functional relation. This shows that Figure 2.4 is correct for grids, too, and it implies that the class of recognizable (i.e., Σ_1-definable) picture languages is not closed under complement even in the case of a trivial alphabet. This disproves a conjecture from [Mat98a]. Precisely, the following is a consequence of Theorem 3.61 and Corollary 4.15.

Theorem 2.32 *For every $k \geq 1$, there is a $(k+1)$-fold function f such that the relation $\{(m,n) \mid f(m) \text{ divides } n\}$ is definable in $\Pi^0_1(\Delta^{\mathrm{U}}_1(\Pi^0_k(\Delta^{\mathrm{U}}_1)))$, in $\Pi^0_1(\Delta_k)$, and in Π_k, but not in Σ_k.*

(See Figure 2.5 for an illustration of these separations.) Note that $\underline{\Pi_k} \not\subseteq \underline{\Sigma_k}$ automatically implies the converse non-inclusion and also $\underline{\Delta_k} \subsetneq \underline{B(\Sigma_k)}$.

2.4.3 Separation Results for Graphs

In Chapter 5 we explain an encoding technique that allows us to transfer the separation results of Corollary 2.25 from 2-bit grids to other structures. We will give such an encoding into undirected graphs and thus conclude the following from Corollaries 2.25 and 2.31.

Theorem 2.33 *Let $k \geq 1$. The following formula classes allow to define (within the class of graphs) a class of connected undirected finite graphs (of bounded degree) that is not definable in Σ_k:*
FO^{TC}, $\Pi^0_{k-1}(\Sigma_1)$, $FO(\Sigma_1)$, Π_k, Σ_{k+1}, Δ_{k+1}.

Besides, the following formula classes allow to define (within the class of graphs) a class of connected undirected finite graphs (of bounded degree) that is not definable in $B(\Sigma_k)$ and neither in $FO(\Sigma_{k-1})$:
Δ_{k+1}, $\Sigma_1(\Pi^0_{k-1}(\Sigma_1))$, $\Pi_1(\Sigma^0_{k-1}(\Pi_1))$.

The reason why not all of the separation results from Corollary 2.25 carry over to graphs arises from the fact that the class of (structures isomorphic to) grids is itself not first-order definable within the class of τ_{Grids}-structures.

However, if we restrict the class of graphs under consideration to connected ones, we may prove a slightly stronger separation.

Theorem 2.34 *Let $k \geq 1$. The following formula classes allow to define (within the class of connected graphs) a class of undirected finite graphs (of bounded degree) that is not definable in Σ_k:*
$\Pi^0_k(\Delta_1)$, $FO(\Delta_1)$.

It is also possible to transfer the separation results that deal with the ι-operator, i.e., one can show that $\underline{\Pi^0_k(\Delta^{\mathrm{U}}_1)} \not\subseteq \underline{\Sigma_k}$ for the class of connected graphs. I refrained from presenting these proofs in this thesis because this would have required considerable technical effort and I consider the transfer of these separations of minor importance.

2.4.4 What Do I Have to Read?

I am aware of the fact that it is quite tedious to read all of this thesis, and probably the reader will consider a lot of results unimportant and will want to read only what is necessary to check the important separation results. Therefore I shall try to indicate short cuts through the thesis wherever possible.

However, the thesis does not present the separation results in order of descending interest, which would certainly help the reader, but rather in an order that simplifies the presentation and minimizes the amount of work.[3] For this reason, all the definability results[4] are collected in Chapter 3, whereas all non-definability results are collected in Chapter 4. In each of these two chapters the last two sections can potentially be skipped. More precisely:

If the reader is interested only in separations $X \not\subseteq Y$, where X is in the left (or middle, respectively) column of Figure 2.3, he or she may skip Sections 3.4 and 3.5 (or just Section 3.5, respectively). He who does not care for the alphabet sizes of the witnesses of the separation results may in any case skip Subsection 3.1.1.

If the reader is interested only in separations $X \not\subseteq Y$, where Y is $\underline{\Sigma_k}$ (or $\underline{B(\Sigma_k)}$, respectively), he or she may skip Sections 4.4 and 4.5 (or just Section 4.5, respectively).

In Chapter 5, which is concerned with the transfer of separation results from grids or coloured grids to other structures, for example finite graphs, the reader might wish to skip Section 5.3, which is only needed for classes that involve the Δ_1-closure operator.

I am afraid that readers who are interested in monadic second-order logic over graphs but not in grids, pictures, and the generalized local alternation hierarchy have no way to avoid reading the parts that deal with these notions.

Chapter 6, which deals with the quantifier depth hierarchy, may be read independently from the rest of the thesis.

[3]That is why we invented the local and the generalized local alternation hierarchy.

[4]The few, easy inclusion results (like $\Pi_k^{\text{gloc}} \subseteq \underline{\Pi_k}$) we need are considered as definability results, too.

Chapter 3

Definability Results

A separation result typically requires two parts. The first one states that a certain property, say of pictures, is a member of a certain class, e.g., expressible by a certain type of formula. The second part states that it is not a member of some other class.

In this chapter we collect all the definability results, i.e., all first parts. These are sometimes referred to as "upper bound proofs", but this term is misleading for our purposes because an essential idea for our definability results is to give *lower* bounds on the asymptotic growth rates of functions whose associated sets of grids are definable in a certain fragment.

Our strategy is the following. In Section 3.3, the most important section of this chapter, we develop a sequence $(Num_k)_{k\geq 1}$ of picture languages for which we will prove membership in Π_{k-1}^{loc}. The membership of Num_k in level k of the dot-depth hierarchy as well as its definability in some other fragments of monadic second-order logic we are interested in then follows easily because Π_k^{loc} is a subset of, e.g., $\underline{\Pi_k}$. These and some other easy inclusions will be presented in Subsection 3.1.3.

In Sections 3.4 and 3.5 we use Num_k to show some additional definability results that allow for a few more separation results. Section 3.4 is required for example for $\underline{B(\Sigma_k)} \subsetneq \underline{\Delta_{k+1}}$. That section is fairly easy compared to Section 3.5, in which the definability parts for example for the separation $\underline{FO(\Delta_k)} \subsetneq \underline{FO(\Sigma_k)}$ are shown.

The family (Num_k) of picture languages has the annoying property that its members are not over the same alphabet. This does not affect the above argumentation, but it is nicer to have, for each level k of the monadic hierarchy, picture languages over the fixed alphabet $\{0,1\}$ that witness the separation of this level from the preceding one. Fortunately, a standard encoding argument, which is presented in Subsection 3.1.1, allows us to obtain such witnesses, thereby passing from some level Σ_k^{loc} of the local hierarchy to level Σ_k^{gloc} of the generalized local alternation hierarchy of picture languages. The reader may safely skip that subsection if he or she is not interested in alphabet sizes.

In Section 3.2 we will introduce some more notation that will be helpful in Sections 3.3, 3.4 and 3.5.

Talking about Formulas. When we talk about formulas, it is often convenient to mix syntax and semantics. For example, when we say "the formula $\varphi(x,y)$ asserts for (a picture P and) a position x that x has some property A, provided that position y has property B", we really mean that for every non-empty picture P and every $u,v \in \mathrm{dom}P$, if v has property B, then $P \models \varphi[u,v]$ iff u has property A.

3.1 Easy Inclusion Results

In the next four sections we will state and prove definability results for the local alternation hierarchy. Before we do that, we collect some easy inclusion results that show how to transfer these definability results to some other classes of picture languages.

In Subsection 3.1.1 we show that in the context of certain definability results of the generalized local alternation hierarchy, the size of the alphabet can be reduced to two.

In Subsection 3.1.2 we show that level k of the generalized local alternation hierarchy is contained in level $k+1$ of the dot-depth hierarchy.

Subsection 3.1.3 shows how to transfer these definability results to classes of picture languages as defined by certain fragments of monadic second-order logic. It states e.g. $\Sigma_k^{\mathrm{gloc}} \subseteq \underline{\Sigma_k}$.

3.1.1 Reduction to Two-Letter Alphabet

With standard techniques we may code alphabets in $\{0,1\}$. This is done in this subsection but it requires the generalized local rather than the local alternation hierarchy.

With respect to the separation $\underline{\Sigma_k} \subsetneq \underline{\Sigma_{k+1}}$ of monadic logic over grids, this has the effect that the length of the outermost quantifier block in formulas witnessing the non-inclusion may be reduced to one.

Throughout this subsection, let $\Gamma = \{0,1\}^k$ and $code : \Gamma \to \{0,1\}^+$ be the mapping defined by $(b_1,\dots,b_k) \mapsto 11b_1 0 b_2 0 \dots 0 b_k$. Let $code$ be extended to a mapping $\Gamma^{+,+} \to \{0,1\}^{+,+}$ and then to a mapping $\mathcal{P}(\Gamma^{+,+}) \to \mathcal{P}(\{0,1\}^{+,+})$ in the straightforward way.

The main content of this subsection is the following.

Theorem 3.1 *Let L be a picture language over Γ. If L is in $\Sigma_k^{\mathrm{gloc}}(\Gamma)$, then $code(L) \in \Sigma_k^{\mathrm{gloc}}(\{0,1\})$.*

Note that it is quite essential that here we refer to the generalized local alternation hierarchy (i.e., the "Σ_k^{gloc}"-version) rather than to the local alternation hierarchy (i.e., the "Σ_k^{loc}"-version). In order to show the above theorem, the following lemma will be useful.

Lemma 3.2 *$code(\Gamma^{+,+})$ is generalized local and thus, in particular, in Π_0^{gloc}.*

Proof This follows from the observation that $code(\Gamma^{+,+})$ contains exactly those non-empty pictures P over $\{0,1\}$ for which $\hat{P}$ does not contain a subblock of the form

- $\#a_0a_1$ with $0 \in \{a_0,a_1\}$,
- $a_0a_1b_1c_1b_2\dots c_{k-1}b_k\#$ with $0 \in \{a_0,a_1\}$,
- $a_0a_1b_1\dots b_{2k-1}a_0'a_1'$ with $a_0 = a_1 = 1 \not\leftrightarrow a_0' = a_1' = 1$,
- $11b_1\dots b_i\#$ with $i < 2k-1$,
- $11b_1c_1b_2\dots c_{k-1}b_k$ with $1 \in \{c_1,\dots,c_{k-1}\}$.

□

The above lemma will help to prove Propositions 3.5 and 3.6, which together show the claim of Theorem 3.1 for the case $k=0$.

Proposition 3.5 can be formulated easier with the following definition.

Definition 3.3 A picture language L over Γ is called *principal* (or *cyclically principal*) if there is a non-empty picture Q over $\Gamma \cup \{\#\}$ such that L is the set of all non-empty pictures P for which Q is a subblock (or a cyclic subblock, respectively). □

Remark 3.4 Every generalized local or generalized cyclically local picture language is a boolean combination of principal or cyclically principal picture languages. □

Proposition 3.5 *Let L be a principal or a cyclically principal picture language over Γ. Then $code(L) \in \Pi_0^{\mathrm{gloc}}$.*

Proof We start by extending *code* to a partial mapping $(\Gamma \cup \{\#\})^{+,+} \longrightarrow \{0,1,\#\}^{+,+}$ in two steps:

Recall that if $Q \in \Gamma^{m,n}$ is a non-empty picture over Γ of size (m,n), with $m,n \geq 1$, then $code(Q) \in \{0,1\}^{m,n(2k+1)}$ is defined as the picture that results from Q by replacing every letter by its *code*-image.

Now if Q is of the form

$$Q = \#^{m+j+k,i} \varobar (\#^{j,n} \ominus Q' \ominus \#^{k,n}) \varobar \#^{m+j+k,l}$$

for some $m,n \geq 1$, $Q' \in \Gamma^{m,n}$, and $i,j,k,l \in \{0,1\}$, i.e., if Q results from a picture Q' over Γ by attaching rows and/or columns of #'s to some of the borders, then

$$code(Q) = \#^{m+j+k,i} \varobar (\#^{j,(2k+1)n} \ominus code(Q') \ominus \#^{k,(2k+1)n}) \varobar \#^{m+j+k,l}$$

i.e., $code(Q)$ results from $code(Q')$ by adding the respective rows and/or columns of the appropriate length or height, respectively.

Now we have extended *code* to a partial mapping $(\Gamma \cup \{\#\})^{+,+} \longrightarrow \{0,1,\#\}^{+,+}$. (The function value is undefined if the #-symbols in the argument do not fill complete rows or columns, or appear at non-border rows or columns.)

Turning to the proof, we consider only the case that L is cyclically principal. The modifications for the case that L is principal are straightforward.

So let Q be a picture over $\Gamma \cup \{\#\}$ such that L is the set of non-empty pictures P over Γ for which $\tilde{P}$ contains Q as a subblock. Let us write L' for the set of non-empty pictures P' over $\{0,1\}$ for which $\tilde{P'}$ contains $code(Q)$ as a cyclic subblock.

We will exploit that

$$code(L) = code(\Gamma^{+,+}) \cap L'. \tag{3.1}$$

The $\subseteq$-direction is trivial. For the converse direction one uses the fact that in the *code*-image of a picture in Γ, two horizontally successive 1's occur exactly at a column whose number is a multiple of $2k+1$.

Since L' is cyclically principal (and hence in Π_0^{gloc}), the claim follows from (3.1) and Lemma 3.2. □

Additionally to generalized cyclic locality, which Proposition 3.5 is concerned with, locally threshold testable word languages are an ingredient for the starfree alternation hierarchy. The following lemma shows that this notion is also respected by the mapping *code*.

Proposition 3.6 *If L is a locally threshold testable word language, then so is $code(L)$.*

Proof Clearly it suffices to consider the case that L is the set of words w for which $\#w\#$ contains some pattern u at least n times as a subblock.

Let L' be the set of non-empty words w over $\{0,1\}$ for which $\#w\#$ contains $code(u)$ at least n times as a subblock. Like in the above proof we have

$$code(L) = code(\Gamma^{+}) \cap L',$$

which shows the claim because $code(\Gamma^{+})$ is locally threshold testable. □

Now we are ready to prove the main result of Subsection 3.1.1.

Proof of Theorem 3.1 We start by noting the following fact: If L, L' are picture languages over Γ, then

$$\begin{aligned} code(L \varobar L') &= code(L) \varobar code(L'), & (3.2)\\ code(L \cup L') &= code(L) \cup code(L'), & (3.3)\\ code(L \cap L') &= code(L) \cap code(L'), & (3.4)\\ code(\Gamma^{+,+} \setminus L) &= code(\Gamma^{+,+}) \setminus code(L). & (3.5) \end{aligned}$$

For the proof one shows the following claims by simultaneous induction on the definition of Σ_k^{gloc}.

1. If L is in $\Sigma_k^{\rm gloc}(\Gamma)$, then $code(L)$ is in $\Sigma_k^{\rm gloc}(\{0,1\})$.

2. If L is in $B(\Sigma_k^{\rm gloc}(\Gamma))$, then $code(L)$ is in $B(\Sigma_k^{\rm gloc}(\{0,1\}))$.

3. If L is in $\oplus$-cl$(B(\Sigma_k^{\rm gloc}(\Gamma)))$, then $code(L) \in \oplus$-cl$(B(\Sigma_k^{\rm gloc}(\{0,1\})))$.

The first two claims of the induction basis are easy by the previous propositions: Let L be a $\Sigma_0^{\rm gloc}$ picture language over Γ. Then L is a boolean combination of picture languages that are (1) generalized local or (2) generalized cyclically local or (3) *top*-pre-images of locally testable word languages. By Propositions 3.5 and 3.6 and Equations (3.3)–(3.5), $code(L)$ is a boolean combination of $\Sigma_0^{\rm gloc}$-languages and thus also in $\Sigma_0^{\rm gloc}$.

Using Lemma 3.2 and the Equations (3.2)–(3.5), the third claim and the induction step are straightforward, . $\square$

The reader who is interested only in monadic second-order logic and not in starfree picture languages may show (by a simpler version of the argumention of this subsection) that Theorem 3.1 remains true if $\Sigma_k^{\rm gloc}$ is replaced by Σ_k. I do not present that proof because we will not need that fact.

3.1.2 Dot-Depth and the Generalized Local Alternation Hierarchy

Now we will prepare the proof of the Proposition 2.21, which states that level k of the generalized local alternation hierarchy is contained in level $k+1$ of the dot-depth hierarchy.

Before we do that, let us make an observation that deals with the empty picture.

Remark 3.7 For convenience and because I preferred not to consider "empty grids", picture languages must not contain the empty picture, and complementation is also relative to the set of non-empty pictures.

However, in proofs and examples like Example 2.18 we will frequently use the empty picture and complementation wrt. $\Gamma^{*,*}$, so it is necessary to note that this does not affect the definition of the dot-depth hierarchy or the class of starfree picture languages.

Formally, if $\textit{dot-depth}'_k(\Gamma) \subseteq \mathcal{P}(\Gamma^{*,*})$ is defined as in Definition 2.19 but with complementation relative to $\Gamma^{*,*}$, then we have that for every $k \geq 0$

and every $L \subseteq \Gamma^{*,*}$ that

$$L \in \textit{dot-depth}'_k(\Gamma) \iff L \setminus \{\varepsilon\} \in \textit{dot-depth}_k(\Gamma),$$

and the same is true for the classes $SF(\Gamma)$ and $\Sigma_k^{\rm gloc}$ etc. The proof is easy, see e.g. [Mat95, Lemma 3.7]. The essential point is that $\oplus$ and $\ominus$ distribute over $\cup$. $\square$

Lemma 3.8 *Every generalized local or generalized cyclically local picture language is of dot-depth* 1.

Proof It suffices to show that for every non-empty picture Q over $\Gamma \cup \{\#\}$, the sets L_Q, M_Q, and N_Q of non-empty pictures P such that Q is a subblock of $\hat{P}$, a properly cyclic subblock of $\tilde{P}$, or a cyclic subblock of $\tilde{P}$, respectively, are of dot-depth one.

So let Q be a non-empty picture over $\Gamma \cup \{\#\}$.

If the set $Q^{-1}(\#)$ of those positions that carry a $\#$ is not a union of rows and columns, or if $\#$ occurs at a non-border position, then Q cannot occur as a subblock of any picture, thus $L_Q = M_Q = N_Q = \emptyset$. So assume that $Q^{-1}(\#)$ consists of a union of none or several of the two border rows and two border columns.

Thus we have to consider sixteen different cases. Firstly, we consider the case that the boundary symbol $\#$ does not occur in Q.

$$\begin{aligned} L_Q &= \Gamma^{*,*} \ominus (\Gamma^{*,*} Q \Gamma^{*,*}) \ominus \Gamma^{*,*}, \\ M_Q &= \Gamma^{*,*} \ominus \left(\bigcup_{\substack{U,V \in \Gamma^{+,+} \\ UV=Q}} V \Gamma^{*,*} U \right) \ominus \Gamma^{*,*}, \\ N_Q &= L_Q \cup M_Q. \end{aligned}$$

This completes the proof for the case that $\#$ does not occur in Q.

Secondly, let the positions of Q that carry a $\#$ be exactly those in the top row of Q. Let Q' result from Q by removing the top row. Then

$$\begin{aligned} L_Q &= (\Gamma^{*,*} Q' \Gamma^{*,*}) \ominus \Gamma^{*,*}, \\ M_Q &= \left(\bigcup_{\substack{U,V \in \Gamma^{+,+} \\ UV=Q'}} V \Gamma^{*,*} U \right) \ominus \Gamma^{*,*}, \\ N_Q &= L_Q \cup M_Q. \end{aligned}$$

Now, thirdly, let the positions of Q that carry a # be exactly those in the left column. Let Q' result from Q by removing the left column. Then

$$\begin{aligned} L_Q &= \Gamma^{*,*} \ominus (Q'\Gamma^{*,*}) \ominus \Gamma^{*,*}, \\ M_Q &= \emptyset, \\ N_Q &= L_Q. \end{aligned}$$

The other thirteen cases are similar. □

The following fact about word languages is easy and stated here without proof.

Proposition 3.9 *Every locally threshold testable word language is of dot-depth 1.*

Lemma 3.10 *If L is a word language of dot-depth 1, then $top^{-1}(L)$ is a picture language of dot-depth 1.*

Proof Fix an alphabet Γ. For a finite or cofinite word language L, we have that $top^{-1}(L) = L \ominus \Gamma^{*,*} \in$ $\oplus$,$\ominus$-cl($dot\text{-}depth_0(\Gamma)$). Thus we have for every word language L in $\oplus$-cl($dot\text{-}depth_0$) that $top^{-1}(L) \in$ $\oplus$-cl($\oplus$,$\ominus$-cl($dot\text{-}depth_0$)) = $\oplus$,$\ominus$-cl($dot\text{-}depth_0$). Thus we have for every word language which is in $dot\text{-}depth_1$ (and thus in $\cap$,$\cup$-cl($\oplus$-cl($dot\text{-}depth_0$))) that $top^{-1}(L) \in$ $\cap$,$\cup$-cl($\oplus$,$\ominus$-cl($dot\text{-}depth_0$)) = $dot\text{-}depth_1$. □

Now we are finally ready to prove the announced connection between the dot-depth hierarchy and the generalized local alternation hierarchy, i.e., that $B(\Sigma_k^{\mathrm{gloc}})(\Gamma) \subseteq dot\text{-}depth_{k+1}(\Gamma)$ for every $k \geq 0$.

Proof of Proposition 2.21 The proof is by induction on k. The case $k = 0$ follows from Propositions 3.8, 3.9 and Lemma 3.10.
So assume the claimed implication is true for some $k \geq 0$. Then $B(\Sigma_{k+1}^{\mathrm{gloc}}) =$ $B(\cap,\cup\text{-cl}(\oplus\text{-cl}(B(\Sigma_k^{\mathrm{gloc}})))) = B(\oplus\text{-cl}(B(\Sigma_k^{\mathrm{gloc}}))) \subseteq B(\oplus\text{-cl}(dot\text{-}depth_{k+1})) =$ $dot\text{-}depth_{k+2}$. □

3.1.3 Generalized Local and Monadic Alternation Hierarchy

In this subsection we will show how to pass from a picture language of some given level of the generalized local alternation hierarchy to a formula of monadic second-order logic of some particular shape.

The following is the main result of this Subsection 3.1.3.

Theorem 3.11 *For every $k \geq 0$, every picture language in Σ_k^{gloc} is definable by a $\Sigma_k^0(\Delta_1^{\mathrm{U}})$-formula.*

Before we turn to the proof of this theorem, let us conclude the following corollary, which will be used in the proofs of Corollaries 3.38 and 3.48.

Corollary 3.12 *Let $k \geq 0$. Every picture language in Σ_k^{gloc} is definable by a formula in $FO^{\leq_1,\leq_2}$, in $\Sigma_k^0(\Delta_1^{\mathrm{U}})$, in $\Sigma_k^0(\Delta_1)$, in $FO(\Delta_1^{\mathrm{U}})$, and in $FO(\Delta_1)$. Every picture language in Π_k^{gloc} is definable in $\Pi_k^0(\Delta_1)$.*

If $k \geq 1$, then every picture language in Σ_k^{gloc} (or Π_k^{gloc}, respectively) is definable in Σ_k (or Π_k, respectively).

Proof We have $\Sigma_k^{\mathrm{gloc}} \subseteq \underline{FO^{\leq_1,\leq_2}}$ by Proposition 3.16, and $\underline{\Sigma_k^{\mathrm{gloc}}} \subseteq \underline{\Sigma_k^0(\Delta_1^{\mathrm{U}})} \subseteq \underline{FO(\Delta_1^{\mathrm{U}})} \subseteq \underline{FO(\Delta_1)}$ by Proposition 2.6 and Theorem 3.11.

If $k \geq 1$, then $\Sigma_k^{\mathrm{gloc}} \subseteq \underline{\Sigma_k^0(\Delta_1)} \subseteq \underline{\Sigma_k(\Delta_1)} \subseteq \underline{\Sigma_k}$ and $\underline{\Pi_k^{\mathrm{gloc}} \subseteq \Pi_k^0(\Delta_1)} \subseteq \underline{\Pi_k}$ by calculation rules of Subsection 2.2.3. □

We prepare the proof of Theorem 3.11 with the following definition. For both $i \in \{1, 2\}$, let $\preceq_i$ be the pre-order on $\mathbb{N} \times \mathbb{N}$ defined by $(m_1, m_2) \preceq_i (n_1, n_2)$ iff $m_i \leq n_i$.

Lemma 3.13 *There are first-order formulas $smaller(x, Z)$ and $larger(x, Z)$ such that $smaller(x, Z)$ asserts that $Z = \{z \mid z \preceq_2 x\}$, and $larger(x, Z)$ asserts that $Z = \{z \mid z \succeq_2 x\}$.*

Proof For $i \in \{1, 2\}$, let $succ_i\text{-}closed(Z) = \forall x \forall y((S_i xy \wedge Zx) \rightarrow Zy)$. It is a first-order formula that asserts that Z is closed under S_i-successors. There are similar formulas $pred_i\text{-}closed(Z)$ (for $i \in \{1, 2\}$) that assert that Z is closed under S_i-predecessors.

The formula $smaller(x, Z) := Zx \wedge succ_1\text{-}closed(Z) \wedge pred_1\text{-}closed(Z) \wedge$ $pred_2\text{-}closed(Z) \wedge \forall y(S_2 xy \rightarrow \neg Zy)$ has the desired property.

The formula $larger(x, Z)$ is defined analogously. □

Throughout Chapters 3 and 6, the following formulas will be useful. The formula $top(x) = \neg\exists y(S_1 yx)$ asserts that position x is in the top row. Similarly, there are first-order formulas $bottom(x)$, $left(x)$, $right(x)$ that assert that x is at the respective border.

Recall the definition of (properly) cyclic subblock from Definition 2.13.

Proposition 3.14 *Let Q be a non-empty picture over $\{0,1\}^t \cup \{\#\}$. Let $\overline{X}$ be the variable tuple $(X_1, \ldots, X_t)$. There are Δ_1^{U}-formulas has-subblock$_Q(\overline{X})$, has-properly-cyclic-subblock$_Q(\overline{X})$, has-cyclic-subblock$_Q(\overline{X})$ that assert for a picture P that Q is a subblock of $\hat{P}$, a properly cyclic subblock of $\tilde{P}$, or a cyclic subblock of $\tilde{P}$, respectively.*

Proof Similar to the proof of Lemma 3.8, we make a case distinction concerning the set $Q^{-1}(\#)$ of positions of Q that contain the boundary symbol #.

In case this set of positions is not a union of top or bottom rows as well as left or right columns, the three formulas $has\text{-}subblock_Q(\overline{X})$, $has\text{-}properly\text{-}cyclic\text{-}subblock_Q(\overline{X})$, $has\text{-}cyclic\text{-}subblock_Q(\overline{X})$ may be chosen as trivial formulas that are always false.

In case $Q^{-1}(\#)$ is a union of top or bottom rows as well as left or right columns, then we have to distinguish sixteen different combinations of these cases.

Let us consider the case that # does not appear in Q. Choose m, n such that Q is of size (m, n). Let $\overline{Y}$ be the variable tuple $(Y_{js})_{j\in[0,n-1],\, s\in[1,t]}$. We need the following first-order formulas.

- For every $k \le n$, the formula $horiz_k(x_0, \ldots, x_{k-1}) = \bigwedge_{j=0}^{k-2} S_2 x_j x_{j+1}$ asserts that $x_0, \ldots, x_{k-1}$ occur one after the other in a row.
- $vertical(x_0, \ldots, x_{m-1}) = \bigwedge_{i=0}^{m-2} S_1 x_i x_{i+1}$, asserts that $x_0, \ldots, x_{m-1}$ occur one after the other in a column.
- $$\begin{aligned} stripes(\overline{Y}, \overline{X}) := {} & \bigwedge_{s=1}^{t} \bigwedge_{j=0}^{n-1} \big(succ_2\text{-}closed(Y_{js}) \wedge pred_2\text{-}closed(Y_{js})\big) \\ & \wedge \forall x_0 \ldots x_{n-1} \big(left(x_0) \wedge horiz_n(x_0, \ldots, x_{n-1}) \\ & \quad \rightarrow \bigwedge_{s=1}^{t} \bigwedge_{j=0}^{n-1} (X_s x_j \leftrightarrow Y_{js} x_j)\big), \end{aligned}$$
  where $succ_2$-closed and $pred_2$-closed are as in the preceding proof. $stripes(\overline{Y}, \overline{X})$ asserts that $Y_{js} = \{(i, j') \in \mathrm{dom}P \mid (i, j) \in X_s^P\}$ for every $s \le t$, $j < n$, i.e., Y_{js} is the union of all those rows whose jth position is contained in X_s.
- For every $(i, j) \in \mathrm{dom}Q$, the formula
  $$\theta_{ij}(\overline{X}, x) = \bigwedge_{\substack{s\in\{1,\ldots,t\} \\ (i,j)\in X_s^Q}} X_s x \wedge \bigwedge_{\substack{s\in\{1,\ldots,t\} \\ (i,j)\notin X_s^Q}} \neg X_s x$$
  asserts that, for all s, the s-th component of $Q\langle i, j\rangle$ is 1 iff $x \in X_s$.

For every $k \le n$, let us write $\overline{x}_k$ for the variable tuple $(x_{ij})_{i\in[m],\, j\in[k]}$.

$$\begin{aligned}
& has\text{-}subblock_Q(\overline{X}) := \\
& \quad \exists \overline{x}_n \big(horiz_n(x_{00}, \ldots, x_{0,n-1}) \wedge \textstyle\bigwedge_{j=0}^{n-1} vertical(x_{0j}, \ldots, x_{m-1,j}) \\
& \quad \wedge \textstyle\bigwedge_{i=0}^{m-1} \bigwedge_{j=0}^{n-1} (\theta_{ij}(X_1, \ldots, X_t, x_{ij}))\big), \\
& has\text{-}properly\text{-}cyclic\text{-}subblock'_Q(\overline{Y}, \overline{X}) := \\
& \quad \textstyle\bigvee_{k=1}^{n-1} \Big(\exists \overline{x}_k \big(horiz_k(x_{00}, \ldots, x_{0,k-1}) \wedge right(x_{0,k-1}) \\
& \quad \wedge \textstyle\bigwedge_{j=0}^{k-1} vertical(x_{0j}, \ldots, x_{m-1,j}) \\
& \quad \wedge \textstyle\bigwedge_{i=0}^{m-1} \bigwedge_{j=0}^{k-1} (\theta_{ij}(X_1, \ldots, X_t, x_{ij})) \\
& \quad \wedge \textstyle\bigwedge_{i=0}^{m-1} \bigwedge_{j=k}^{n-1} (\theta_{ij}(Y_{j1}, \ldots, Y_{jt}, x_{i0}))\big)\Big), \\
& has\text{-}cyclic\text{-}subblock'_Q(\overline{Y}, \overline{X}) := \\
& \quad has\text{-}subblock_Q(\overline{X}) \vee has\text{-}properly\text{-}cyclic\text{-}subblock'_Q(\overline{Y}, \overline{X}).
\end{aligned}$$

Then $has\text{-}subblock_Q(\overline{X})$, $has\text{-}properly\text{-}cyclic\text{-}subblock'_Q(\overline{Y}, \overline{X})$, and $has\text{-}cyclic\text{-}subblock'_Q(\overline{Y}, \overline{X})$ assert for a picture P over $\{0,1\}^t$ that Q is a subblock, a properly cyclic subblock, a cyclic subblock, respectively, of P, provided that for all j, s, the set Y_{js} is the (uniquely determined) union of all those rows whose jth position is contained in X_s.

Thus

$$\begin{aligned} & has\text{-}properly\text{-}cyclic\text{-}subblock_Q(\overline{X}) := \\ & \quad has\text{-}properly\text{-}cyclic\text{-}subblock'_Q(\iota\overline{Y}(stripes(\overline{Y}, \overline{X}), \overline{X})) \end{aligned}$$

is a Δ_1^{U}-formula that asserts that Q is a properly cyclic subblock of P, as required. An analogous construction works for $has\text{-}cyclic\text{-}subblock_Q$.

This completes the proof for the case that the boundary symbol # does not appear in Q. Minor modifications have to be made in the other fifteen cases, as we will sketch now.

In case the positions of Q that contain # are exactly those in the top row (or those in the bottom row, respectively), we remove the variables $x_{00}, \ldots, x_{0,n-1}$ (or the variables $x_{m-1,0}, \ldots, x_{m-1,n-1}$, respectively) from the variable tuple $\overline{x}_n$ and add $top(x_{10})$ (or $bottom(x_{m-2,0})$, respectively) as a conjunct in the scope of the existential quantifier in the definition of *has-subblock*$_Q$. The other two formulas are modified similarly.

In case the positions of Q that contain # are exactly those in the left column (or right column, respectively), we remove the variables $x_{00}, \ldots, x_{m-1,0}$ (or $x_{0,n-1}, \ldots, x_{m-1,n-1}$, respectively) from the variable tuple $\overline{x}_n$ and add $left(x_{01})$ (or $right(x_{0,n-2})$, respectively) in the definition of *has-subblock*$_Q$, and similarly for the other two formulas.

In case the positions of Q that contain # are exactly those in the left column and those in the top row, we combine the modifications of the two cases, and similarly in the ten other combinations of these cases.

This completes the proof of Proposition 3.14. □

Proposition 3.15 *Every picture language in Σ_0^{gloc} is definable in Δ_1^{U}.*

Proof Since $\underline{\Delta_1^{\mathrm{U}}}$ is closed under boolean combinations and by Remark 3.4, it suffices to consider the cases of principal or cyclically principal picture languages and of *top*-pre-image of locally threshold testable word languages.

Every principal or cyclically principal picture language is in $\underline{\Delta_1^{\mathrm{U}}}$ by Proposition 3.14.

Every locally threshold testable word language is known to be first-order definable (in the signature with one successor relation symbol S but without ordering). A first-order sentence φ in this signature translates easily (by replacing S with S_2 and relativizing quantifications to *top*) to a first-order sentence φ' in our signature τ_{Grids} in such a way that φ' is true for those non-empty pictures whose top row fulfills φ . This completes the proof. □

A proof for the following easy fact can be found in [Wil97].

Proposition 3.16 *Every starfree picture language is definable in $FO^{\leq_1,\leq_2}$.*

The crucial idea of the proof is that a concatenation can be imitated by an existential first-order quantification, just like in the well-known word language case.

We will do a similar proof for Theorem 3.11. It will be essential to relativize quantifications in some first-order formula, say φ, in such a way that the resulting formula asserts for the picture left (or right) from a given position the same as φ asserts for the whole picture. This idea is made precise by the following definition.

Definition 3.17 Let $\overline{x}$ be the variable tuple $(x_1, \ldots, x_n)$. Let $\varphi(\overline{x})$, $\varphi'(\overline{x}, z)$ be formulas in the signature τ_{Grids_t}, where $t \geq 0$.

φ' *relativizes* φ *to* $\preceq_2 z$ iff for all non-empty pictures P over $\{0,1\}^t$, all $(i,j) \in \mathrm{dom}P$, and all $x_1, \ldots, x_n \preceq_2 (i,j)$:

$$P \models \varphi'[\overline{x}, (i,j)] \iff P([0, \underline{P} - 1] \times [0, j]) \models \varphi[\overline{x}].$$

φ' *relativizes* φ *to* $\succeq_2 z$ iff for all non-empty pictures P over $\{0,1\}^t$, all $(i,j) \in \mathrm{dom}P$, and all $x_1, \ldots, x_n \succeq_2 (i,j)$:

$$P \models \varphi'[\overline{x}, (i,j)] \iff P([0, \underline{P} - 1] \times [j, |P| - 1]) \models \varphi[\overline{x}].$$

A class of formulas $\mathcal{F}$ is $\preceq_2$*-relativizable* iff for every $\varphi \in \mathcal{F}$ there are $\mathcal{F}$-formulas that relativize φ to $\preceq_2 z$ and to $\succeq_2 z$, respectively. □

Lemma 3.18 *For every $k \geq 0$, the formula class $\Sigma_k^0(\Delta_1^{\mathrm{U}})$ is $\preceq_2$-relativizable.*

Proof We argue by induction on k.

For the induction base $k = 0$, let $\psi(\iota\overline{X}(\varphi(\overline{X}, \overline{Y}), \overline{Y}))$ be some Δ_1^{U}-formula, i.e., $\varphi(\overline{X}, \overline{Y})$, $\psi(\overline{X}, \overline{Y})$ are first-order and $\overline{X}$ is a tuple of set variables, whereas the tuple $\overline{Y}$ may contain first-order as well as set variables. Let Z be a fresh set variable.

Let $\varphi'(\overline{X}, Z, \overline{Y})$ and $\psi'(\overline{X}, Z, \overline{Y})$ result from φ and ψ, respectively, by relativizing first-order quantifications to Z. For every X_i in the tuple $\overline{X}$, let us write "$X_i \subseteq Z$" instead of $\forall x(X_i x \to Z)$. Let $\varphi''(\overline{X}, Z, \overline{Y}, z) = \varphi' \wedge smaller(z, Z) \wedge \bigwedge_i (X_i \subseteq Z)$. Since $\exists!\overline{X}(\varphi)$ is valid, so is $\exists!\overline{X}Z(\varphi'')$. Thus[1] $\psi'(\iota\overline{X}Z(\varphi''), \overline{Y}, z)$ is an allowed Δ_1^{U}-formula. It is easy to see that it relativizes ψ to $\preceq_2 z$. A similar construction can be done to get a Δ_1^{U}-formula that relativizes ψ to $\succeq_2 z$.

For the induction step, let $k \geq 0$ such that $\Sigma_k^0(\Delta_1^{\mathrm{U}})$ is $\preceq_2$-relativizable. Let $\varphi(\overline{x})$ be a formula in $\Sigma_{k+1}^0(\Delta_1^{\mathrm{U}})$. Then $\varphi(\overline{x})$ is equivalent to a formula of the form $\exists\overline{y}(\neg\psi(\overline{x}, \overline{y}))$ for some $\Sigma_k^0(\Delta_1^{\mathrm{U}})$-formula ψ and a first-order variable tuple $\overline{y}$. By induction hypothesis there is a $\Sigma_k^0(\Delta_1^{\mathrm{U}})$-formula $\psi'(\overline{x}, \overline{y}, z)$ that

[1]Note that we added z to the list of potentially free variables in ψ.

relativizes ψ to $\preceq_2 z$. By Lemma 3.13 there is a Δ_1^{U}-formula $\sigma(\overline{y}, z)$ that asserts that $y_i \preceq_2 z$ for all y_i in the tuple $\overline{y}$.

Then $\varphi(\overline{x})$ is relativized to $\preceq_2 z$ by the formula $\exists \overline{y}\,(\sigma(\overline{y}, z) \wedge \neg\psi'(\overline{x}, \overline{y}, z))$. This formula is in $\Sigma^0_{k+1}(\Delta_1^{\mathrm{U}})$ because $\sigma \in \Delta_1^{\mathrm{U}} \subseteq \Pi^0_k(\Delta_1^{\mathrm{U}})$ and $\Pi^0_k(\Delta_1^{\mathrm{U}})$ is closed under conjunction. A formula that relativizes φ to $\succeq z$ can be obtained similarly. This finishes the induction and thus the proof. □

The reader who is not interested in the formula classes $\Sigma^0_k(\Delta_1^{\mathrm{U}})$ may prove (similarly) that the classes $\Sigma^0_k(\Delta_1)$ and (if $k \geq 1$) also Σ_k are $\preceq_2$-relativizable.

Now we can state the lemma that clarifies the mentioned idea that first-order quantifications may replace (column-)concatenations.

Lemma 3.19 *Let $\mathcal{F}$ be a $\preceq_2$-relativizable class of formulas that contains all first-order formulas. If L_1, L_2 are picture languages (over some alphabet $\{0,1\}^t$) and definable in $\mathcal{F}$, then $L_1 \obar L_2$ is definable in $\Sigma^0_1(\mathcal{F})$.*

In particular, if $\mathcal{F}$ is closed under conjunction, disjunction, and existential first-order quantifications, then the set of $\mathcal{F}$-definable picture languages is closed under column concatenation.

Proof Write $\overline{X}$ for the variable tuple $(X_1, \ldots, X_t)$. Let $\varphi_1(\overline{X})$ and $\varphi_2(\overline{X})$ be formulas in $\mathcal{F}$ that define L_1 and L_2, respectively. Let $\varphi_1'(\overline{X}, z)$ and $\varphi_2'(\overline{X}, z)$ be $\mathcal{F}$-formulas that relativize φ_1 to $\preceq_2 z$ and φ_2 to $\succeq_2 z$, respectively. Let $\varphi(\overline{X}) = \exists z, z'(S_2 z z' \wedge \varphi_1'(z) \wedge \varphi_2'(z'))$. Then φ defines $L_1 \obar L_2$. □

We conclude this subsection with the proof of its main result, Theorem 3.11, which states that $\Sigma_k^{\mathrm{gloc}} \subseteq \underline{\Sigma^0_k(\Delta_1^{\mathrm{U}})}$.

Proof of Theorem 3.11 We argue by induction on k. The case $k = 0$ is Proposition 3.15. Assume $\Sigma_k^{\mathrm{gloc}} \subseteq \underline{\Sigma^0_k(\Delta_1^{\mathrm{U}})}$ is true for some k.

Since $\Sigma_{k+1}^{\mathrm{gloc}}$ is a subset of the smallest superset of $B(\Sigma_k^{\mathrm{gloc}})$ that is closed under union, intersection, and column concatenation, the claimed implication follows from the fact that $\Sigma^0_{k+1}(\Delta_1^{\mathrm{U}}) \supseteq B(\Sigma^0_k(\Delta_1^{\mathrm{U}}))$ and that $\Sigma^0_{k+1}(\Delta_1^{\mathrm{U}})$ is closed under disjunction, conjunction, and (in the sense of Lemma 3.19) column concatenation. □

Note that the proof shows that Theorem 3.11 remains true if the inductive definition of the generalized local alternation hierarchy is modified as follows: $\Sigma_{k+1}^{\mathrm{gloc}}$ is the smallest superclass of $B(\Sigma_k^{\mathrm{gloc}})$ that is closed under positive boolean combinations *as well as* $\obar$- and $\ominus$-concatenations.

3.2 Attributed Alphabets

In this section we will introduce some notation that will be helpful later on. Since this notation might seem strange at first sight, I would like to motivate it.

For the syntax of monadic second-order formulas, it is possible to restrict the choice of set variables to a fixed set $\{X_1, X_2, X_3, \ldots\}$, and this is very convenient in arguments *against* definability by certain formulas. In such non-definability proofs it even suffices to consider formulas φ for which there is an i such that $\mathit{free}(\varphi) = \{X_1, \ldots, X_i\}$ because other cases can be reduced to this one by renaming of free variables.

But it will be agreed that it is extraordinarily helpful to allow free variables to be chosen arbitrarily and not only from the "anonymous" set $\{X_1, X_2, X_3, \ldots\}$. This improves readability; it allows us to indicate a particular relationship between free variables by choosing similar, suggestive names. Apart from the fact that variable names like X_{count} and the like are more readable than X_1, X_2, etc., it is very important that we are allowed to define formulas φ, ψ with incomparable sets of free variables because it allows us to combine formulas, say by conjunction, without renaming of bound variables and yet avoiding name clashes.

This chapter will be chiefly concerned with starfree picture languages over alphabets of 0-1-tuples. Suppose we would use only alphabets of the form $\{0,1\}^t$ for some t. This would make notations difficult for the same reasons as in the context of formulas, as I will illustrate in an example. Suppose we want to construct a starfree word language L over $\{0,1\}^3$ from two languages M, N over $\{0,1\}^2$ in such a way that L contains those words over $\{0,1\}^3$ whose letter-wise restriction to the first two components is in M whereas the letter-wise restriction to the second two ones is in N. Then we would have to write $L = (M \otimes \{0,1\}^*) \cap (\{0,1\}^* \otimes N)$, where the $\otimes$ means "letter-wise pairing", see Subsection 6.2.1 for a definition. Note that the understanding is further complicated by the fact that what used to be the first component in (symbols of words in) the language N is the second component in L.

The main idea to avoid such complications is to give names to the components of letters, i.e., to have families over some finite index set as alphabet symbols. Such families are called "data base tuples" in data base theory, and the elements of the index set are called "attributes". Suppose in the above example these "attributed" alphabets are chosen in such a way that the second component in M and the first component in N correspond to the same attributes, i.e., μ, ν, ξ are different attributes, and M and N are over alphabets $\{0,1\}^{\{\mu,\nu\}}$ and $\{0,1\}^{\{\nu,\xi\}}$, respectively. If $a \in \{0,1\}^{\{\mu,\nu\}}$ and $b \in \{0,1\}^{\{\nu,\xi\}}$ are two letters, then the *join* $a \bowtie b$ is defined iff $a(\nu) = b(\nu)$, namely by $(a \bowtie b)(\mu) = a(\mu)$, and $(a \bowtie b)(\nu) = a(\nu) = b(\nu)$, and $(a \bowtie b)(\xi) = b(\xi)$. This join operation is lifted to an operation $\otimes$ on words and pictures in the usual (i.e., letter-wise) way. This operation is in turn lifted to sets of words or pictures, thus we will simply write $L = M \otimes N$ in the above situation.

Formally, let I be a finite set of so-called "attributes". By $\{0,1\}^I$ we denote the set of functions $I \to \{0,1\}$, i.e., $\{0,1\}^I$ is the set of I-indexed families over $\{0,1\}$.

If I, J are attribute sets and $a \in \{0,1\}^I$, $b \in \{0,1\}^J$, and if $a(\mu) = b(\mu)$ for every $\mu \in I \cap J$, then $a \bowtie b$ is the element of $\{0,1\}^{I \cup J}$ with

$$(a \bowtie b)(\mu) = \begin{cases} a(\mu) & \text{if } \mu \in I, \\ b(\mu) & \text{if } \mu \in J. \end{cases}$$

If $a(\mu) \neq b(\mu)$ for some $\mu \in I \cap J$, then "$a \bowtie b$" is not defined.

This (partial) operation is extended to a (total) operation on sets as usual: If $L \subseteq \{0,1\}^I$, $M \subseteq \{0,1\}^J$, then $L \bowtie M = \{a \bowtie b \mid a \in L, b \in M\}$. We will use sets of the form $\{0,1\}^I$ as alphabets. The operation $\bowtie$ is also extended to pictures over such alphabets: If P is a picture over $\{0,1\}^I$ and Q is a picture over $\{0,1\}^J$ of equal size, and if "$(P\langle i,j\rangle) \bowtie (Q\langle i,j\rangle)$" is defined for every position $(i,j) \in \mathrm{dom}P$, then $P \otimes Q$ is defined, namely as the picture of size size(P) over $\{0,1\}^{I \cup J}$ with

$$(P \otimes Q)\langle i,j\rangle = (P\langle i,j\rangle) \bowtie (Q\langle i,j\rangle)$$

for every $(i,j) \in \mathrm{dom}P$. Again, this partial operation is lifted to a total operation on sets as usual.

Some more notation is needed for attributed alphabets. If μ is some attribute and $b \in \{0,1\}$, then $(b)_\mu \in \{0,1\}^{\{\mu\}}$ is defined by $((b)_\mu)(\mu) = b$. This way, $(.)_\mu$ is an alphabet projection from $\{0,1\}$ to $\{0,1\}^\mu$, which is lifted to pictures and picture languages the usual way. If $a \in \{0,1\}^I$ is some I-indexed family and $J \subseteq I$, then $\mathsf{restr}_J(a)$ is the restriction of a to a J-indexed family. Again, $\mathsf{restr}_J(.)$ is an alphabet projection from $\{0,1\}^I$ to $\{0,1\}^J$ and is lifted to pictures and picture languages.

Likewise, we consider for some $\mu \in I$ the alphabet projection $\mathsf{pr}_\mu : \{0,1\}^I \to \{0,1\}$, $a \mapsto a(\mu)$. It can be lifted to pictures, i.e., if P is a picture over $\{0,1\}^I$ and $\mu \in I$, then we write $\mathsf{pr}_\mu(P)$ for the picture of the same size over $\{0,1\}$ with $(\mathsf{pr}_\mu(P))\langle i,j\rangle = (P\langle i,j\rangle)(\mu)$ for every position (i,j) of P, and analogously for picture languages.

Remark 3.20 Let L, M be picture languages over alphabets $\{0,1\}^I$ and $\{0,1\}^J$, respectively, and let $\Omega = \{0,1\}^{I \cup J}$.

Then $L \otimes M = (L \otimes \Omega^{+,+}) \cap (M \otimes \Omega^{+,+})$ and $L \otimes \Omega^{+,+} = \{P \in \Omega^{+,+} \mid \mathsf{restr}_I(P) \in L\}$.

With these identities one easily deduces that the class of local picture languages is closed under $\otimes$ from the fact that it is closed under inverse projections (Proposition 2.17) and intersection. □

We sometimes write $(0)_I$ for the letter $a \in \{0,1\}^I$ with $\mathsf{pr}_\mu(a) = 0$ for all $\mu \in I$.

3.3 Counting Cyclically— The Central Definability Results

In this section we shall introduce the sequence $(\mathit{Num}_k)_k$ of those picture languages that witness almost all non-inclusion results of this thesis. The crucial point is that the pictures of these languages are very "slim" in the sense that the pictures are very wide compared to their height.

For every $k \geq 1$, we define the k-fold exponential function $f_k : \mathbb{N} \to \mathbb{N}$ as follows.

$$\begin{aligned} f_1(m) &= 2^m, \\ f_{k+1}(m) &= f_k(m)2^{f_k(m)}. \end{aligned}$$

The pictures of Num_k will be of size $(m, f_k(m))$ for $m \geq 1$.

Throughout this chapter, we consider the attributes $\mathit{num}\text{-}1, \mathit{num}\text{-}2, \ldots$, $\mathit{end}\text{-}1, \mathit{end}\text{-}2, \ldots$. For every $k \geq 1$ the attribute set I_k is fixed as $I_k = \{\mathit{num}\text{-}1, \ldots, \mathit{num}\text{-}k, \mathit{end}\text{-}1, \ldots, \mathit{end}\text{-}k\}$.

The picture language Num_k will be defined as a certain picture language over alphabet $\{0,1\}^{I_k}$. For every $l \leq k$, the restriction of a picture in Num_k to attribute $\mathit{num}\text{-}l$ will establish a counting mechanism that cyclically enumerates binary numbers of length $f_l(m)$. The projection to attribute $\mathit{end}\text{-}l$ will mark the ends of these binary number representations.

We start with the following auxiliary definition.

Definition 3.21 For every $k \geq 1$, let $\mathit{OnlyTop}_k$ be the set of non-empty pictures P over $\{0,1\}^{\{\mathit{num}\text{-}k,\mathit{end}\text{-}k\}}$ for which $\mathsf{pr}_{\mathit{num}\text{-}k}(P)\langle i,j\rangle = \mathsf{pr}_{\mathit{end}\text{-}k}(P)\langle i,j\rangle = 0$ for all $(i,j) \in \mathrm{dom}P$ with $i \neq 0$. □

In other words, $\mathit{OnlyTop}_k = (L)_{\mathit{num}\text{-}k} \otimes (L)_{\mathit{end}\text{-}k}$ for the local picture language $L = (\{0,1\}^+ \ominus \{0\}^{*,*})$ of Example 2.10. As a consequence of that example, $\mathit{OnlyTop}_k$ is a local picture language.

Now we will define, for every $k, m \geq 1$, a picture P_{km} of height m over alphabet $\{0,1\}^{I_k}$, and then the picture language Num_k over $\{0,1\}^{I_k}$.

Definition 3.22 For every $m \geq 1$ let P_{1m} be the picture of size $(m, f_1(m))$ over $\{0,1\}^{I_1}$ for which

- $\mathsf{pr}_{\mathit{num}\text{-}1}(P_{1m})$ is the picture over $\{0,1\}$ whose j-th column holds the binary representation of j for every $j \in \{0,\ldots,2^m-1\}$ (least significant bit at the top), (cf. Example 2.15),

- $\mathsf{pr}_{\mathit{end}\text{-}1}(P_{1m})$ is the picture over $\{0,1\}$ such that for all $(i,j) \in \mathrm{dom}P = [m]\times[2^m]$

$$\mathsf{pr}_{\mathit{end}\text{-}1}(P_{1m})\langle i,j\rangle = 1 \iff \forall j' \geq j : \mathsf{pr}_{\mathit{num}\text{-}1}(P_{1m})\langle i,j'\rangle = 1.$$

For every $k \geq 1$, $m \geq 1$, let $P_{k+1,m}$ be the picture of size $(m, f_{k+1}(m))$ over alphabet $\{0,1\}^{\{\mathit{num}\text{-}k+1,\mathit{end}\text{-}k+1\}}$ such that

- $\mathsf{pr}_{\mathit{num}\text{-}k+1}(\mathit{top}(P_{k+1,m})) = \mathit{BIN}(0, f_k(m)) \ldots \mathit{BIN}(2^{f_k(m)-1}, f_k(m))$ is the word of length $f_{k+1}(m)$ that consists of the concatenation of the reverse binary representations (of length $f_k(m)$) of the numbers $0,\ldots,2^{f_k(m)}-1$,

- $\mathsf{pr}_{\mathit{end}\text{-}k+1}(\mathit{top}(P_{k+1,m})) = 0^{f_{k+1}(m)-1}1$, and

- $P_{k+1,m} \in \mathit{OnlyTop}_{k+1}$.

For every $k \geq 1$, let

$$\mathit{Num}_k = \bigcup_{m\geq 1} \{P_{1m}\}^+ \otimes \ldots \otimes \{P_{k-1,m}\}^+ \otimes \{P_{k,m}\}.$$

□

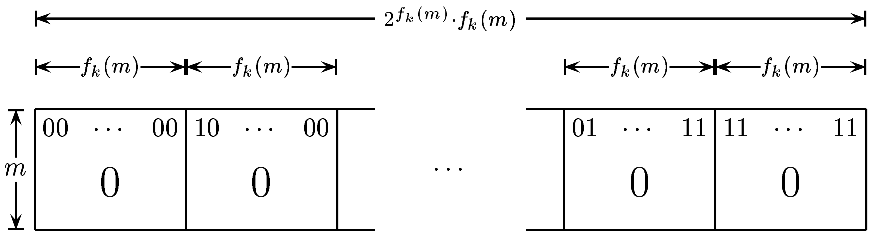


Figure 3.1: The picture $\mathsf{pr}_{\mathit{num}\text{-}k+1}(P_{k+1,m})$

The picture $\mathsf{pr}_{\mathit{num}\text{-}1}(P_{1,3})$ is displayed as (3.6) on page 72. As an illustration of Definition 3.22, Figure 3.1 displays the projection of $P_{k+1,m}$ to its attribute $\mathit{num}\text{-}k+1$. Note that the length of $P_{k+1,m}$ is indeed $2^{f_k(m)} \cdot f_k(m) = f_{k+1}(m)$.

The following is the main theorem of this section. When this theorem is proved, most of the "hard work" of this chapter will have been done. Theorem 3.43 follows quite simply and rest of the hard work is for Theorem 3.61.

These three theorems and the easy inclusion results depicted in Figure 2.1 provide all definability results needed for the separation results announced in Section 2.4.

Theorem 3.23 *$\mathit{cycl}(\mathit{Num}_k^+)$ is in Π^{loc}_{k-1} for every $k \geq 1$.*

Note that this will in particular imply that Num_k is Π^{loc}_{k-1}, which in turn allows us to deduce that the function f_k is Σ_k-definable in the sense of Definition 2.26. Both [MT97] and [Sch97a] construct monadic second-order formulas that define such k-fold exponential function inductively over k. (It is f_k in [Sch97a] and a very similar function in [MT97].) However, both of these inductions require the simultaneous construction of two other formulas, which makes the proof difficult to understand.

My point is that considering the picture language $\mathit{cycl}(\mathit{Num}_k^+)$ instead of Num_k makes the induction required for the construction a lot simpler.

Contrary to this simplification, the detour over the local alternation hierarchy complicates the construction of such formulas. However, this detour is non-essential. The reader who is interested in definability results only for monadic second-order logic may extract easily an inductive proof that directly constructs, say, Π_k-formulas. That means that the following may be proved via induction over k:

$$\mathit{cycl}(\mathit{Num}_k^+) \in \begin{cases} \underline{\Delta_1} & \text{if } k = 1, \\ \underline{\Pi_{k-1}} & \text{if } k \geq 2. \end{cases}$$

(In the induction step one constructs from a Δ_k-*sentence*(!) for $\mathit{cycl}(\mathit{Num}_k^+)$ a Π_k-sentence for $\mathit{cycl}(\mathit{Num}_{k+1}^+)$.)

The following proposition serves as an induction basis for Theorem 3.23.

Proposition 3.24 *$\mathit{cycl}(\mathit{Num}_1^+)$ is in Π_0^{loc}.*

This proposition is an almost immediate consequence of Example 2.15. The picture language considered in that example is $\mathsf{pr}_{\mathit{num}\text{-}1}(\mathit{cycl}(\mathit{Num}_1^+))$. Nevertheless it is instructive to carry out the proof in more detail because the verification of counting mechanisms in pictures in $\mathit{cycl}(\mathit{Num}_k^+)$ for $k > 1$ will depend on the same ideas.

We start with a definition.

Definition 3.25 Let Allow be the set of $2{\times}2$-pictures $\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}$ over $\{0,1\}$ for which $((a,c) = (1,0)) \leftrightarrow (b \neq d)$. Let Forb $= \{0,1\}^{2,2} \setminus$ Allow. □

In other words, Allow is the set of $2{\times}2$ subblocks pictures that result from column concatenating columns that are successive binary number representations, i.e., Allow the set of cyclic $2{\times}2$-subblocks of the picture

$$\begin{matrix} 0\,1\,0\,1\,0\,1\,0\,1 \\ 0\,0\,1\,1\,0\,0\,1\,1 \\ 0\,0\,0\,0\,1\,1\,1\,1 \end{matrix}\ . \tag{3.6}$$

Counting is a "local process" in the sense that one may determine the new value of every bit by investigating the old value of that bit as well as the new and old values of the previous bit. The latter contains the information about whether a carry influences the current bit. This observation is exploited in the following proof.

Proof of Proposition 3.24 Let $\Gamma = \{0,1\}^{I_1}$.

$$\begin{aligned} \Delta = \ & \left\{ \begin{smallmatrix} \boldsymbol{a} & \boldsymbol{c} \\ \boldsymbol{b} & \boldsymbol{d} \end{smallmatrix} \in \Gamma^{2,2} \ \middle|\ \mathsf{pr}_{\mathit{num}\text{-}1}\begin{pmatrix} \boldsymbol{a} & \boldsymbol{c} \\ \boldsymbol{b} & \boldsymbol{d} \end{pmatrix} \in \text{Allow} \right. \\ & \qquad \left. \wedge \left(\mathsf{pr}_{\mathit{end}\text{-}1}(\boldsymbol{a}) = 1 \leftrightarrow \mathsf{pr}_{\mathit{end}\text{-}1}(\boldsymbol{b}) = \mathsf{pr}_{\mathit{num}\text{-}1}(\boldsymbol{a}) = 1\right) \right\} \\ & \cup \left\{ \begin{smallmatrix} \boldsymbol{a} & \boldsymbol{c} \\ \# & \# \end{smallmatrix} \ \middle|\ (\mathsf{pr}_{\mathit{end}\text{-}1}(\boldsymbol{a}) = 1 \leftrightarrow \mathsf{pr}_{\mathit{num}\text{-}1}(\boldsymbol{a}) = 1) \right\} \\ & \cup \left\{ \begin{smallmatrix} \# & \# \\ \boldsymbol{b} & \boldsymbol{d} \end{smallmatrix} \ \middle|\ \mathsf{pr}_{\mathit{num}\text{-}1}(\boldsymbol{b}) \neq \mathsf{pr}_{\mathit{num}\text{-}1}(\boldsymbol{d}) \right\}. \end{aligned}$$

Let P be a non-empty picture over Γ, say of size (m,n). Let $x_0, x_1, \ldots, x_{n-1}$ be the words of length m over Γ such that x_i is the transposed column i of P, i.e.,

$$x_i = P\langle 0, i\rangle \ldots P\langle m-1, i\rangle.$$

Δ is chosen in such a way that every cyclic $2{\times}2$-subblock of $\tilde{P}$ is in Δ iff

- the words $\mathsf{pr}_{\mathit{num}\text{-}1}(x_0), \ldots, \mathsf{pr}_{\mathit{num}\text{-}1}(x_{n-1})$ fulfill the second property of Lemma 3.26 ("the Cyclic Counting Lemma"), and
- for every x among $x_0, \ldots, x_{n-1}$, the word $\mathsf{pr}_{\mathit{end}\text{-}1}(x)$ is determined by $\mathsf{pr}_{\mathit{end}\text{-}1}(x) = 0^i 1^j$ iff j is maximal such that 1^j is a suffix of $\mathsf{pr}_{\mathit{num}\text{-}1}(x)$.

Thus (by Lemma 3.26) Num_1^+ is the set of non-empty pictures P over Γ such that Δ cyclically tiles $\tilde{P}$, thus Num_1^+ is cyclically local and in particular Π_0^{loc}. □

Before we give the missing Lemma 3.26, I make a historical remark. The observation stated in the above lemma, namely that—speaking informally—counting is a local process, has been one of the two[2] essential improvements of the definability part of the proof of [Sch97a] compared to [MT97] because it made it possible to save quantifier alternations in the monadic second order formulas that define "slim" picture languages. This observation has been used previously in [Gia94] (see Example 2.28) and it can also be found implicit in [Mat95, Example 2.4]. What is new here is to continue the counting process cyclically.

After this historical comment, let us continue with the announced lemma.

[2]The other improvement is to start the induction appropriately, i.e., at $k = 1$ with the exponential function f_1 instead of at $k = 0$ with the identity. This allows to save one quantifier alternation and, moreover, this allows to describe the counting pattern in $FO^{\leq_1, \leq_2}$.

Lemma 3.26 (Cyclic Counting Lemma) *Let $m, n \geq 1$ and $x_0, \ldots, x_{n-1}$ be words of length m over alphabet $\{0,1\}$. Then the following are equivalent:*

1. $\mathit{dual}(x_{(j+1) \bmod n}) = (\mathit{dual}(x_j)+1) \bmod 2^m$ *for every* $j \in \{0, \ldots, n-1\}$;

2. • $x_j\langle 0\rangle \neq x_{(j+1) \bmod n}\langle 0\rangle$ *for every* $j \in \{0, \ldots, n-1\}$.
   • $\begin{pmatrix} x_j\langle i\rangle & x_{(j+1) \bmod n}\langle i\rangle \\ x_j\langle i+1\rangle & x_{(j+1) \bmod n}\langle i+1\rangle \end{pmatrix} \in \mathit{Allow}$ *for all* $j \in \{0, \ldots, n-1\}$, $i \in \{0, \ldots, m-1\}$.

Proof The proof depends on the following fact (cf. [Sch97a, Remark 16] or, more detailed, [Sch97b, Prop. 5.12]. See also Example 2.15.): Let x, y be words of length m over $\{0,1\}$. Then we have the following equivalence chain:
$\mathit{dual}(y) = (\mathit{dual}(x)+1) \bmod 2^m$ iff
$(x\langle 0\rangle \neq y\langle 0\rangle) \wedge \forall i : (x\langle i\rangle, y\langle i\rangle) = (1,0) \leftrightarrow x\langle i+1\rangle \neq y\langle i+1\rangle$ iff
$x\langle 0\rangle \neq y\langle 0\rangle$ and all $2{\times}2$-subblocks of $x^\top \oplus y^\top$ are in Allow. □

Proposition 3.24 proves Theorem 3.23 for the case $k=1$. Our aim is to do the proof of Theorem 3.23 by induction on k. We will sketch the strategy for this informally.

Checking membership in $\mathit{cycl}(\mathit{Num}^+_{k+1})$ for some picture P over $\{0,1\}^{I_{k+1}}$ whose restriction to I_k is in $\mathit{cycl}(\mathit{Num}^+_k)$ means essentially to check that the component corresponding to the attribute $\mathit{num}\text{-}k+1$ enumerates all binary representations of length $f_k(m)$, where m is the height of P. The idea for this is similar to the case $k=1$, namely to exploit the Cyclic Counting Lemma 3.26 in order to verify the counting mechanism locally. The difference is that—unlike in the case $k=1$—corresponding bits of successive binary number representations are no more horizontally neighboured, but at horizontal cyclic distance $f_k(m)$ from each other. The restriction of P to the I_k-components can be used to determine such corresponding positions because if the restriction of P to attribute set I_k is in $\mathit{cycl}(\mathit{Num}^+_k)$, then two positions are at horizontal cyclic distance $f_k(m)$ from each other iff the restriction to I_k-components of the cyclic infix of P between these positions is in $\mathit{cycl}(\mathit{Num}_k)$. Some additional tricks are necessary to control the existence of particular cyclic infixes on the next lemma of the local alternation hierarchy.

We will carry out the sketched proof formally. For this, we need the following technical lemma. Its essential statement is the equivalence of Items 3 and 5, which states that Num^+_{k+1} results from Num^+_k by enlarging the underlying alphabet (i.e. passing from attribute set I_k to I_{k+1}) and then forbidding particular cyclic infixes.

Lemma 3.27 *Let $k \geq 1$, let $\Omega = \{0,1\}^{I_{k+1}}$. Let U_1, U_2, U_3 be the word languages over Ω defined as follows:*

$$\begin{aligned}
U_1 &= \bigcup_{b \in \{0,1\}} (10^*10)_{\mathit{end}\text{-}k} \otimes (\{0,1\}b\{0,1\}^*b)_{\mathit{num}\text{-}k+1} \otimes \Omega^+, \\
U_2 &= \bigcup_{\substack{a\,c \\ b\,d} \in \mathit{Forb}} (0^+10^+\{0,1\})_{\mathit{end}\text{-}k} \otimes (ab\{0,1\}^*cd)_{\mathit{num}\text{-}k+1} \otimes \Omega^+, \\
U_3 &= (0^*1)_{\mathit{end}\text{-}k} \otimes (1^*)_{\mathit{num}\text{-}k+1} \otimes (\sim(0^*1))_{\mathit{end}\text{-}k+1} \otimes \Omega^+ \\
&\cup \ (0^*1)_{\mathit{end}\text{-}k} \otimes (\sim(1^*))_{\mathit{num}\text{-}k+1} \otimes (\sim(0^*))_{\mathit{end}\text{-}k+1} \otimes \Omega^+.
\end{aligned}$$

For every $m \geq 1$ let bin_m abbreviate $\mathit{BIN}(0, f_k(m)) \ldots \mathit{BIN}(2^{f_k(m)-1}, f_k(m))$, i.e., the word of length $f_{k+1}(m)$ that consists of the reverse binary representations (of length $f_k(m)$) of the numbers $0, \ldots, 2^{f_k(m)}-1$.

The following are equivalent for every picture $P \in \mathit{cycl}(\mathit{Num}^+_k) \otimes \Omega^{+,+}$ of height m.

1. $\mathit{top}(P) \in \mathit{cycl}(((\{0,1\}^*1)_{\mathit{end}\text{-}k} \otimes (\mathit{bin}_m)_{\mathit{num}\text{-}k+1})^+) \otimes \Omega^+$,

2. *$\mathit{top}(P)$ has no cyclic infix of length $f_k(m)+2$ in $U_1 \cup U_2$.*

Furthermore, the following are equivalent for every $P \in \mathit{cycl}(\mathit{Num}^+_k) \otimes \mathit{OnlyTop}_{k+1}$ of height m.[3]

3. $P \in \mathit{cycl}(\mathit{Num}^+_{k+1})$,

4. $\mathit{top}(P) \in \mathit{cycl}(((\{0,1\}^*1)_{\mathit{end}\text{-}k} \otimes (\mathit{bin}_m)_{\mathit{num}\text{-}k+1} \otimes (0^*1)_{\mathit{end}\text{-}k+1})^+) \otimes \Omega^+$,

5. *$\mathit{top}(P)$ has no cyclic infix of length $f_k(m)+2$ in $U_1 \cup U_2$ and none of length $f_k(m)$ in U_3.*

Proof Let $P \in \mathit{cycl}(\mathit{Num}^+_k) \otimes \Omega^{+,+}$ and m be the height of P. W.l.o.g. we may assume that $P \in \mathit{Num}^+_k \otimes \Omega^{+,+}$ because validity of Items 1–5 is not affected by passing from P to any picture in $\mathit{cycl}(P)$. There are $n \geq 1$ and

[3] Recall $\mathit{OnlyTop}_{k+1}$ from Definition 3.21.

words $x_0, \ldots, x_{n-1} \in \Omega^+$ such that $top(P) = x_0 \ldots x_{n-1}$ and $\mathsf{pr}_{end\text{-}k}(x_j) = 0^{f_k(m)-1}1$ for every j.

The cyclic infixes of length $f_k(m)+2$ of the word $x_0 \ldots x_{n-1} = top(P)$ can be classified as follows. They are of either one of the following two kinds:

1. ax_jb, where $j < n$ and $a = x_{(j-1) \bmod n}\langle f_k(m)-1\rangle$ is the last letter of $x_{(j-1) \bmod n}$, and $b = x_{(j+1) \bmod n}\langle 0\rangle$ is the first letter of $x_{(j+1) \bmod n}$.
2. $x_j\langle i\rangle \ldots x_j\langle f_k(m)-1\rangle x_{(j+1) \bmod n}\langle 0\rangle \ldots x_{(j+1) \bmod n}\langle i+1\rangle$ for some $j < n$ and $i < f_k(m)-1$.

If x is of the first kind, then $\mathsf{pr}_{end\text{-}k}(x) \in 10^*10$, and thus x is a candidate for U_1. If x is of the second kind, then $\mathsf{pr}_{end\text{-}k}(x) \in 0^+10^+\{0,1\}$, and thus x is a candidate for U_2.

2⇔1 We have the following equivalence chain:

$top(P) \in cycl(((\{0,1\}^*1)_{end\text{-}k} \otimes (bin_m)_{num\text{-}k+1})^+) \otimes \Omega^+$;

⇔ (since $x_0, \ldots, x_{n-1}$ are those cyclic infixes of $top(P) = x_0 \ldots x_{n-1}$ that are in $(0^*1)_{end\text{-}k} \otimes \Omega$):

$$\mathsf{pr}_{num\text{-}k+1}(x_j \ldots x_{n-1}x_0 \ldots x_{j-1}) \in {bin_m}^+ \text{ for some } j;$$

⇔ (by definition of bin_m): for every $j \leq n-1$:

$$dual(\mathsf{pr}_{num\text{-}k+1}(x_{(j+1) \bmod n})) = (dual(\mathsf{pr}_{num\text{-}k+1}(x_j)) + 1) \bmod 2^{f_k(m)};$$

⇔ (by the "cyclic counting Lemma"):

$$\mathsf{pr}_{num\text{-}k+1}(x_j\langle 0\rangle) \neq \mathsf{pr}_{num\text{-}k+1}(x_{(j+1) \bmod n}\langle 0\rangle)$$

for every $j < n$, and

$$\mathsf{pr}_{num\text{-}k+1}\begin{pmatrix} x_j\langle i\rangle & x_{j+1 \bmod n}\langle i\rangle \\ x_j\langle i+1\rangle & x_{j+1 \bmod n}\langle i+1\rangle \end{pmatrix} \in \text{Allow}$$

for every $j < n$, $i < m$;

⇔ (by the preceding classification of infixes of length $f_k(m)+2$): The word $top(P)$ has no cyclic infix x of length $f_k(m)+2$ with

$$\begin{array}{rcl} \mathsf{pr}_{end\text{-}k}(x) & \in & 1 \quad 0^* \quad 10, \\ \mathsf{pr}_{num\text{-}k+1}(x) & \in & \{0,1\}b\{0,1\}^* \ b, \end{array}$$

where $b \in \{0,1\}$, and no cyclic infix x of length $f_k(m)+2$ with

$$\begin{array}{rcl} \mathsf{pr}_{end\text{-}k}(x) & \in & 0 \quad 0^*10^* \ 0\{0,1\}, \\ \mathsf{pr}_{num\text{-}k+1}(x) & \in & ab\{0,1\}^*c \quad d, \end{array}$$

where $\begin{smallmatrix} a & c \\ b & d \end{smallmatrix} \in$ Forb;

⇔ $top(P)$ has no cyclic infix x of length $f_k(m)+2$ in $U_1 \cup U_2$.

3⇒4 We note that

$$top(Num_{k+1}) \subseteq ((0^*1)^*)_{end\text{-}k} \otimes (bin)_{num\text{-}k+1} \otimes (0^*1)_{end\text{-}k+1} \otimes \Omega^+.$$

Apply $.^+$ and then $cycl(.)$ on both sides of this inclusion to obtain the claim.

3⇐4 First we remark that if Q, Q' are two non-empty pictures over $\{0,1\}^{I_k}$ with $QQ' \in cycl(Num_k^+)$ such that the last letter both of $\mathsf{pr}_{end\text{-}k}(top(Q))$ and of $\mathsf{pr}_{end\text{-}k}(top(Q'))$ is "1", then $Q, Q' \in Num_k^+$.

Let $P \in cycl(Num_k^+) \otimes OnlyTop_{k+1}$ fulfill Property 4, $m := \underline{\overline{P}}$. We may assume w.l.o.g. that

$$top(P) \in ((\{0,1\}^*1)_{end\text{-}k} \otimes (bin_m)_{num\text{-}k+1} \otimes (0^*1)_{end\text{-}k+1})^+ \otimes \Omega^{+,+},$$

i.e., there are pictures $P_0, \ldots, P_{n-1}$ such that $P = P_0 \ldots P_{n-1}$ and

$$\begin{array}{rcl} \mathsf{pr}_{end\text{-}k}(top(P_j)) & \in & \{0,1\}^*1, \\ \mathsf{pr}_{num\text{-}k+1}(top(P_j)) & = & bin_m, \\ \mathsf{pr}_{end\text{-}k+1}(top(P_j)) & \in & 0^*1 \end{array}$$

for every j. Since $\mathsf{restr}_{I_k}(P_0 \ldots P_{n-1}) \in cycl(Num_k^+)$ and the last letter of $\mathsf{pr}_{end\text{-}k}(top(P_j))$ is 1 for each j, it follows (by trivial induction and the above remark) that $\mathsf{restr}_{I_k}(P_j) \in Num_k^+$ for every j, and thus $P_j \in Num_k^+ \otimes OnlyTop_{k+1} \otimes top^{-1}((bin_m)_{num\text{-}k+1}) \otimes top^{-1}((0^*1)_{end\text{-}k+1}) = Num_{k+1}$ for every j. This implies $P = P_0 \ldots P_{n-1} \in Num_{k+1}^+ \subseteq cycl(Num_{k+1}^+)$.

4⇔5 Let $P \in Num_k^+ \otimes OnlyTop_{k+1}$. In case P does not fulfill Property 1, then it does not fulfill Properties 2, 4, 5 either, making the equivalence "4⇔5" valid.

So assume P fulfills Property 1. We may assume w.l.o.g. that $top(P) \in ((\{0,1\}^*1)_{end\text{-}k} \otimes (bin_m)_{num\text{-}k+1})^+ \otimes \Omega^+$. There are uniquely determined words $y_1, \dots, y_l$ ($l \geq 1$) with $top(P) = y_1 \dots y_l$ and $\mathsf{pr}_{num\text{-}k+1}(y_j) = bin_m$ and $\mathsf{pr}_{end\text{-}k}(y_j) \in \{0,1\}^*1$ for every $j \leq l$. Let $n = 2^{f_k(m)}$. There are x_{ji} (with $j \in \{1, \dots, l\}$, $i \in \{0, \dots, n-1\}$) such that $y_j = x_{j0} \dots x_{j,n-1}$ and $\mathsf{pr}_{num\text{-}k+1}(x_{ji}) = BIN(i, f_k(m))$ for every j, i. Since every $\mathsf{pr}_{end\text{-}k}(y_j)$ is in $\mathsf{pr}_{end\text{-}k}(top(Num_k^+)) = (0^{f_k(m)-1}1)^+$ and of length $|bin_m| = f_{k+1}(m) = nf_k(m)$, we have $\mathsf{pr}_{end\text{-}k}(x_{ji}) = 0^{f_k(m)-1}1$ for every i, j.

Then we have the following equivalence chain:

P has Property 4;

$\Leftrightarrow$ (by choice of $y_1, \dots, y_l$): $\mathsf{pr}_{end\text{-}k+1}(y_j) \in 0^*1$ for all $j \leq l$;

$\Leftrightarrow$ (by choice of $x_{j0}, \dots x_{j,n-1}$): for all i, j:

$$\mathsf{pr}_{end\text{-}k+1}(x_{ji}) \in \begin{cases} 0^*1 & \text{if } i = n-1, \\ 0^* & \text{if } i < n-1; \end{cases}$$

$\Leftrightarrow$ (since $\mathsf{pr}_{num\text{-}k+1}(x_{ji}) = BIN(i, f_k(m))$ for all i, j) for all i, j:

$$\mathsf{pr}_{end\text{-}k+1}(x_{ji}) \in \begin{cases} 0^*1 & \text{if } \mathsf{pr}_{num\text{-}k+1}(x_{ji}) \in 1^*, \\ 0^* & \text{if } \mathsf{pr}_{num\text{-}k+1}(x_{ji}) \notin 1^*; \end{cases}$$

$\Leftrightarrow$ (since the x_{ji} are exactly those cyclic infixes of $top(P)$ which are in $(0^{f_k(m)-1}1)_{end\text{-}k} \otimes \Omega^+$):

$top(P)$ has no cyclic infix x with $\mathsf{pr}_{end\text{-}k}(x) \in 0^{f_k(m)-1}1$ and

$$\begin{array}{lcll} \mathsf{pr}_{num\text{-}k+1}(x) \in 1^* & \wedge & \mathsf{pr}_{end\text{-}k+1}(x) \notin 0^*1, & \text{or} \\ \mathsf{pr}_{num\text{-}k+1}(x) \notin 1^* & \wedge & \mathsf{pr}_{end\text{-}k+1}(x) \notin 0^*; & \end{array}$$

$\Leftrightarrow$ $top(P)$ has no cyclic infix of length $f_k(m)$ in U_3;

$\Leftrightarrow$ (by assumption and the equivalence "1$\Leftrightarrow$2") P has Property 5.

This completes the proof. □

The important point of Lemma 3.27 is that Num_{k+1}^+ results from Num_k^+ by enlarging the underlying alphabet (i.e., passing from attribute set I_k to I_{k+1}) and then forbidding particular cyclic infixes. The next lemmas

show how this "forbidding" can be done on the next level of the alternation hierarchy for starfree picture languages.

Firstly, we remark that it is easy to forbid certain pictures as (non-cyclic) infixes: Suppose M, K are picture languages over Γ. Then the set of all pictures of M that do not have an infix in K is given by $M \setminus (\Gamma^{*,*}K\Gamma^{*,*})$. If M and K are in Π_{k-1}^{loc}, then this set is in Π_k^{loc}.

The situation is more complicated when we deal with *cyclic* infixes because a cyclic infix consists of a prefix and a suffix and it is not possible to control its length etc. as easily as usual infixes.

However, in our particular case (see Lemma 3.27), the languages Num_k^+ (which will play the role of M) and also the languages $top^{-1}((U_1 \cap \Omega^{f_k(m)+2}) \cup (U_2 \cap \Omega^{f_k(m)+2}))$ and $top^{-1}(U_3 \cap \Omega^{f_k(m)})$ (which will play the rôle of K) have certain properties that can be exploited.

Decomposability

The next definition, remark, and proposition state the required property of $top^{-1}(U_1 \cup U_2)$ and $top^{-1}(U_3)$.

Definition 3.28 (Suffix-Prefix Decomposition) A picture language L over alphabet Γ is Π_0^{loc}*-suffix/prefix-decomposable* (or *sp-decomposable* for short) iff there are $n \geq 1$ and picture languages $U_1, \dots, U_n, V_1, \dots, V_n \in \Pi_0^{\text{loc}}$ such that for all non-empty pictures P, Q over Γ we have $PQ \in L$ iff there is an $i \leq n$ with $P \in V_i$, $Q \in U_i$. □

The following example illustrates this definition.

Example 3.29 The word language $ab\{0,1\}^*$, where $a, b \in \{0,1\}$, is an sp-decomposable word language. To see this, consider

$$\begin{array}{lcllcl} V_1 & = & \{a\}, & U_1 & = & b\{0,1\}^*, \\ V_2 & = & ab\{0,1\}^*, & U_2 & = & \{0,1\}^+. \end{array}$$

Since U_1, U_2, V_1, V_2 are locally threshold testable word languages over $\{0,1\}$ and for all words $v, u \in \{0,1\}^+$ we have $vu \in L \iff \bigvee_i (v \in V_i \wedge u \in U_i)$, this shows sp-decomposability of $ab\{0,1\}^*$. □

The sp-decomposable word languages are characterized by the following remark.

Remark 3.30 One can show that every word language in Π_0^{loc} (and thus, in particular, every locally threshold testable word language) is sp-decomposable.

Every sp-decomposable picture language is in Π_0^{loc}. □

Note that there are Π_0^{loc} picture languages that are not sp-decomposable, for example Num_1.

The following is shown easily, either with Remark 3.30 or directly.

Remark 3.31 If L_1, L_2 are sp-decomposable word languages and α is some alphabet projection, then $\alpha^{-1}(L_1)$, $L_1 \cap L_2$, $L_1 \cup L_2$, and $L_1 \otimes L_2$ are sp-decomposable word languages, and $top^{-1}(L_1)$ is an sp-decomposable picture language. □

We will need the following proposition.

Proposition 3.32 *Let $a, b \in \{0,1\}$. Then all the word languages 10^*, $\{0,1\}b\{0,1\}^*$, 0^+10^*, $0\{0,1\}^*$, $ab\{0,1\}^*$ are sp-decomposable.*

Proof The proof is simple with Remark 3.30 because the above languages are locally threshold testable. Alternatively we may in each case proceed as in Example 3.29. □

Cyclic Infix Property

The following definition and remark state the property of Num_k required in order to forbid cyclic infixes (like in Item 5 of Lemma 3.27) in the next level of the local alternation hierarchy.

Definition 3.33 A picture language N has the cyclic-infix property iff for all pictures P, Q, R for which PQR is defined we have

$$PQR \in cycl(N^*) \land Q \in cycl(N^*) \Longrightarrow PR \in cycl(N^*).$$

□

Proposition 3.34 *Num_k has the cyclic-infix property for every $k \geq 1$.*

The above proposition is easy but nevertheless we will sketch how it can be proved. It becomes a little clearer with the following (purely word language-theoretic) lemma.

Lemma 3.35 *Let Γ be an alphabet and $\$ \notin \Gamma$. Let $w \in \Gamma^*\$$ and $u, v \in (\Gamma \cup \{\$\})^+$ with $uv, v \in cycl(\{w\}^+)$. Then $u \in cycl(\{w\}^+)$.*

Proof The letter \$ occurs both in u and in v. Let x and x' be the unique prefixes of u and v, respectively, that are in $\Gamma^*\$$. Since $v, uv \in cycl(\{w\}^+)$ there is $y \in \Gamma^*$ such that $uv \in xw^*y$ and $v \in x'w^*y$ and $yx = w = yx'$. The latter implies $x = x'$. Thus there is $y' \in \Gamma^*$ with $u \in xw^*y'$ and $y'x = w = yx$, so $y' = y$. Thus $u \in xw^*y' = xw^*y \subseteq cycl(\{w\}^+)$. □

Proof of Proposition 3.34 It clearly suffices to consider the case that $R = \varepsilon$ because otherwise one may pass to RP and ε instead of P and R, respectively, which does not affect the validity of the implication statement. So let $k \geq 1$ and $P, Q \in Num_k^*$ be pictures such that $PQ, Q \in cycl(Num_k^*)$. If $Q = \varepsilon$ or $P = \varepsilon$, the claim is immediate, so we may assume that P and Q are of height $m \geq 1$. We may apply the above lemma to the column word w of the picture of height m in Num_k and the column words u and v of P and Q, respectively. (Choose as \$ the last column of the picture of height m in Num_k. It is the only one whose top position carries a 1 in its *end-k*-component.) We conclude that $Q \in cycl(Num_k^+)$. □

The next lemma shows why the cyclic-infix property is so helpful: it gives us certain control about cyclic infixes from sp-decomposable picture languages.

Lemma 3.36 (Cyclic Infix Lemma) *Let L_1, L_2 be sp-decomposable picture languages, let N have the cyclic-infix property. If $cycl(N^+) \in \Pi_{k-1}^{\text{loc}}$, then*

$$cycl((L_1 \cap cycl(N^+))(L_2 \cap cycl(N^+))) \in \Sigma_k^{\text{loc}}.$$

Proof Choose $U_1, \ldots, U_n, V_1, \ldots, V_n$ according to Definition 3.28 for L_1, and likewise $U'_1, \ldots, U'_m, V'_1, \ldots, V'_m$ for L_2. As an abbreviation, let $M = cycl(N^+)$. Then

$$\begin{aligned} cycl((L_1 \cap M)(L_2 \cap M)) = \; & (L_1 \cap M)(L_2 \cap M) \cup (L_2 \cap M)(L_1 \cap M) \\ & \cup \big(M \cap \textstyle\bigcup_{j \leq m}(V'_j(L_1 \cap M)U'_j)\big) \\ & \cup \big(M \cap \textstyle\bigcup_{i \leq n}(V_i(L_2 \cap M)U_i)\big), \end{aligned}$$

which is clearly in Σ_k^{loc}. □

By Remark 3.20, the above lemma remains true if $\cap$ is replaced by $\otimes$.

In the next lemma we will combine the above lemma and Lemma 3.27 (Items 3 and 5) in order to construct, for a given k, a picture language $K_3 \in \Sigma_k^{\mathrm{loc}}$ such that for every picture P in $\mathit{cycl}(\mathit{Num}_k^+) \otimes \mathit{OnlyTop}_{k+1}$ over $\{0,1\}^{I_{k+1}}$ we have $P \in K_3 \iff P \notin \mathit{Num}_{k+1}$. This will suffice to complete the proof of Theorem 3.23.

Lemma 3.37 *Let $k \geq 1$ such that $\mathit{cycl}(\mathit{Num}_k^+) \in \Pi_{k-1}^{\mathrm{loc}}(\{0,1\}^{I_k})$. Let $\Omega := \{0,1\}^{I_{k+1}}$. Let U_1, U_2, U_3 as in Lemma 3.27.*

For every $a,b,c,d \in \{0,1\}$, there are picture languages L_{bb}, L_{abcd}, L, K_1, K_2, K_3 in $\Sigma_k^{\mathrm{loc}}(\Omega)$ such that for all $P \in \mathit{cycl}(\mathit{Num}_k^+) \otimes \mathit{OnlyTop}_{k+1}$ we have

1. *$P \in L_{bb}$ iff $\mathit{top}(P)$ has a cyclic infix of length $f_k(m)+2$ in*
$$(10^*10)_{\mathit{end}\text{-}k} \otimes (\{0,1\}b\{0,1\}^*b)_{\mathit{num}\text{-}k+1} \otimes \Omega^+.$$
2. *$P \in L_{abcd}$ iff $\mathit{top}(P)$ has a cyclic infix of length $f_k(m)+2$ in*
$$(0^+10^+\{0,1\})_{\mathit{end}\text{-}k} \otimes (ab\{0,1\}^*cd)_{\mathit{num}\text{-}k+1} \otimes \Omega^+.$$
3. *$P \in L$ iff $\mathit{top}(P)$ has a cyclic infix of length $f_k(m)$ in U_3.*
4. *$P \in K_1$ iff $\mathit{top}(P)$ has a cyclic infix of length $f_k(m)+2$ in U_1.*
5. *$P \in K_2$ iff $\mathit{top}(P)$ has a cyclic infix of length $f_k(m)+2$ in U_2.*
6. *$P \in K_3$ iff $\mathit{top}(P)$ has a cyclic infix of length $f_k(m)+2$ in $U_1 \cup U_2$ or one of length $f_k(m)$ in U_3.*

Proof Each of the first three claims is an immediate consequence of Lemma 3.36. We give the choice of L_1 and L_2 in the application of that lemma. The remaining three claims are consequences.

Ad 1. Let $b \in \{0,1\}$. Choose

$$\begin{aligned} L_1 &= \mathit{top}^{-1}((10^*)_{\mathit{end}\text{-}k} \otimes (\{0,1\}b\{0,1\}^*)_{\mathit{num}\text{-}k+1} \otimes \Omega^+),\\ L_2 &= \mathit{top}^{-1}((10\{0,1\}^*)_{\mathit{end}\text{-}k} \otimes (\{0,1\}b\{0,1\}^*)_{\mathit{num}\text{-}k+1} \otimes \Omega^+),\\ L_{bb} &= \mathit{cycl}((L_1 \otimes \mathit{cycl}(\mathit{Num}_k^+))(L_2 \otimes \mathit{cycl}(\mathit{Num}_k^+))). \end{aligned}$$

By Lemma 3.36 and Proposition 3.32, L_{bb} is in Σ_k^{loc}. It remains to show that for every $P \in \mathit{cycl}(\mathit{Num}_k^+) \otimes \mathit{OnlyTop}_{k+1}$ we have that $P \in L_{bb}$ iff $\mathit{top}(P)$ has a cyclic infix of the desired form.

Let $P \in \mathit{cycl}(\mathit{Num}_k^+) \otimes \mathit{OnlyTop}_{k+1} \otimes L_{bb}$. Then there are Q_1, Q_2 such that $P \in \mathit{cycl}(Q_1Q_2)$ and $Q_i \in L_i \otimes \mathit{cycl}(\mathit{Num}_k^+)$ for both $i \in \{1,2\}$. Let $u = \mathit{top}(Q_1)$ and $v = \mathit{top}(Q_2)$. Then

$$\begin{aligned} &\mathsf{pr}_{\mathit{end}\text{-}k}(u) \in 1\ 0\ 0^*, && \mathsf{pr}_{\mathit{end}\text{-}k}(v) \in 1\ 0\{0,1\}^*,\\ &\mathsf{pr}_{\mathit{num}\text{-}k+1}(u) \in \{0,1\}b\{0,1\}^*, && \mathsf{pr}_{\mathit{num}\text{-}k+1}(v) \in \{0,1\}b\{0,1\}^*. \end{aligned}$$

Choose x as the prefix of length 2 of v, then ux is a cyclic infix of $\mathit{top}(P)$ with

$$\begin{aligned} \mathsf{pr}_{\mathit{end}\text{-}k}(ux) &\in 1\ 0^*\ 1\ 0,\\ \mathsf{pr}_{\mathit{num}\text{-}k+1}(ux) &\in \{0,1\}\ b\ \{0,1\}b. \end{aligned}$$

Since $\mathsf{restr}_{I_k}(Q_1) \in \mathit{cycl}(\mathit{Num}_k^+)$ and $\mathsf{pr}_{\mathit{end}\text{-}k}(u) = \mathsf{pr}_{\mathit{end}\text{-}k}(\mathit{top}(Q_1)) \in 0^*10^*$, this implies that $\mathsf{restr}_{I_k}(Q_1) \in \mathit{cycl}(\mathit{Num}_k)$ and thus $|u| = |Q_1| = f_k(m)$, which implies $|ux| = f_k(m)+2$.

The converse direction, i.e., that every picture P in $\mathit{cycl}(\mathit{Num}_k^+) \otimes \mathit{OnlyTop}_{k+1}$ that has an infix of the desired form is in L_{bb}, is shown similarly: Write that infix as ux, where x is of length 2. Choose v such that $\mathit{top}(P) \in \mathit{cycl}(uv)$. The above properties of u, v, x are shown in the converse direction. Finally choose Q_1, Q_2 such that $P \in \mathit{cycl}(Q_1Q_2)$ and $\mathit{top}(Q_1) = u$ and $\mathit{top}(Q_2) = v$, show that $Q_i \in L_i$ and conclude that $P \in L_{bb}$.

Ad 2. For every $a,b,c,d \in \{0,1\}$, choose

$$\begin{aligned} L_1 &= \mathit{top}^{-1}((0^+10^*)_{\mathit{end}\text{-}k} \otimes (ab\{0,1\}^*)_{\mathit{num}\text{-}k+1} \otimes \Omega^+),\\ L_2 &= \mathit{top}^{-1}((0\{0,1\}^*)_{\mathit{end}\text{-}k} \otimes (cd\{0,1\}^*)_{\mathit{num}\text{-}k+1} \otimes \Omega^+),\\ L_{abcd} &= \mathit{cycl}((L_1 \otimes \mathit{cycl}(\mathit{Num}_k^+))(L_2 \otimes \mathit{cycl}(\mathit{Num}_k^+))). \end{aligned}$$

Again, by Lemma 3.36 and Proposition 3.32, L_{abcd} is in Σ_k^{loc}. The proof that L_{abcd} has the property claimed in Item 2 is analogous to the above.

Ad 3. We have $L = \mathit{cycl}((\mathit{top}^{-1}(U_3) \otimes \mathit{cycl}(\mathit{Num}_k^+))(\Omega^{+,+} \otimes \mathit{cycl}(\mathit{Num}_k^+))) \cup \mathit{cycl}((\mathit{top}^{-1}(U_3) \otimes \mathit{cycl}(\mathit{Num}_k^+)))$ is in Σ_k^{loc} (exploit Lemma 3.36 with $L_1 = \mathit{top}^{-1}(U_3)$ and $L_2 = \Omega^{+,+}$) and has the desired property by similar arguments.

Ad 4, 5, and 6. Choose

$$\begin{aligned} K_1 &= L_{00} \cup L_{11}, \\ K_2 &= \bigcup_{\substack{a\ c \\ b\ d} \in \mathrm{Forb}} L_{abcd}, \\ K_3 &= K_1 \cup K_2 \cup L. \end{aligned}$$

Then K_1, K_2, and K_3 are in Σ_k^{loc} and have the desired properties. □

The sixth claim of the above lemma is the key for the following proof of Theorem 3.23, which states that $\mathit{cycl}(\mathit{Num}_k^+)$ is in Π_{k-1}^{loc} for every $k \geq 1$.

Proof (of Theorem 3.23) The proof is by induction. By Proposition 3.24, $\mathit{cycl}(\mathit{Num}_1^+)$ is in Π_0^{loc}, which is the induction basis. So let $k \geq 1$ with $\mathit{cycl}(\mathit{Num}_k^+)$ is in Π_{k-1}^{loc}, choose K_3 as in Lemma 3.37. By Lemma 3.27, we have

$$\mathit{cycl}(\mathit{Num}_{k+1}^+) = (\mathit{cycl}(\mathit{Num}_k^+) \otimes \mathit{OnlyTop}_{k+1}) \setminus K_3.$$

Since $\mathit{cycl}(\mathit{Num}_k^+) \in \Pi_{k-1}^{\mathrm{loc}} \subseteq \Pi_k^{\mathrm{loc}}$ and $K_3 \in \Sigma_k^{\mathrm{loc}}$ and $\mathit{OnlyTop}_{k+1} \in \Pi_0^{\mathrm{loc}}$, it follows that $\mathit{cycl}(\mathit{Num}_{k+1}^+)$ is in Π_k^{loc}, which completes the induction. □

Together with the non-definability results of Chapter 4 (even with the 'easy' Section 4.3), the membership of $\mathit{cycl}(\mathit{Num}_k^+)$ in Π_{k-1}^{loc} already implies interesting separation results, among them $\underline{\Sigma_k} \subsetneq \underline{\Sigma_{k+1}}$ for the class of coloured grids or the class of graphs.

For these separations, we will need the following corollary, which depends on facts that are proved in Section 3.1. For a picture language L, we define $\mathrm{size}(L)$ as $\{\mathrm{size}(P) \mid P \in L\}$.

Corollary 3.38 *For every k, the languages $\mathit{cycl}(\mathit{Num}_k^+)$, Num_k^+, and Num_k are in $\Pi_{k-1}^{\mathrm{loc}}(\{0,1\}^{I_k})$, in $\underline{\Pi_{k-1}^0(\Delta_1^{\mathrm{U}})}$, in $\underline{\Pi_{k-1}^0(\Delta_1)}$, in $\underline{\Pi_{k-1}}$, in $\underline{FO(\Delta_1)}$, and in $\underline{FO^{\leq_1,\leq_2}}$.*

In particular, there are picture languages L_0, L_1 in the classes listed above such that $\mathrm{size}(L_0) = \{(m,n) \mid f_k(m)$ divides $n\}$ and $\mathrm{size}(L_1) = f_k$.

Proof Let $k \geq 1$. We first show membership in Π_{k-1}^{loc}. For $\mathit{cycl}(\mathit{Num}_k^+)$, this is Theorem 3.23, and for the other two languages, this follows from

$$\begin{aligned} \mathit{Num}_k^+ &= \mathit{cycl}(\mathit{Num}_k^+) \otimes \mathit{top}^{-1}((\{0,1\}^*1)_{\mathit{end}\text{-}k}), \\ \mathit{Num}_k &= \mathit{cycl}(\mathit{Num}_k^+) \otimes \mathit{top}^{-1}((0^*1)_{\mathit{end}\text{-}k}), \end{aligned}$$

because $\{0,1\}^*1$ and 0^*1 are locally threshold testable word languages, so that their *top*-pre-images are in Π_0^{loc}, and because Π_{k-1}^{loc} is closed under join.

Membership in the other listed classes follows from Theorem 3.11 and Corollary 3.12. For the second statement let $L_0 = \mathit{Num}_k^+$ and $L_1 = \mathit{Num}_k$. □

Corollary 3.38 provides the definability result for Theorem 2.22, stating that $\Pi_k^{\mathrm{loc}}(\{0,1\}^{2k}) \not\subseteq \underline{\Sigma_k}(\{0,1\}^{2k})$ for all $k \geq 1$. In Corollary 4.14 we will show the corresponding non-definability result.

Using Subsection 3.1.1 we obtain also the following two corollaries.

Corollary 3.39 *Let $k \geq 1$. There is a picture language L over alphabet $\{0,1\}$ in $\Pi_{k-1}^{\mathrm{gloc}}$ such that $\mathrm{size}(L)$ is a k-fold exponential function.*

Proof Let π be a bijection $\{0,1\}^{I_k} \to \{0,1\}^{2k}$. Let $\mathit{code} : \{0,1\}^{2k} \to \{0,1\}^+$ as in Theorem 3.1. By that theorem, $L := \mathit{code}(\pi(\mathit{Num}_k)) \in \Pi_k^{\mathrm{gloc}}(\{0,1\})$. Besides, $\mathrm{size}(L) = \{(m, (4k+1) \cdot f_k(m)) \mid m \geq 1\}$ is indeed a k-fold exponential function. □

Corollary 3.40 *Let $k \geq 1$. There is a k-fold exponential function definable by a sentence of the form $\exists X_1 \varphi(X_1)$, where φ is a formula in $\Pi_{k-1}^0(\Delta_1^{\mathrm{U}})$, in $\Pi_{k-1}^0(\Delta_1)$, in Π_{k-1}, in $FO(\Delta_1)$, and in $FO^{\leq_1,\leq_2}$.*

Proof Choose $\varphi(X_1)$ to define the language from Corollary 3.39, and use Corollary 3.12. □

Recall from Definition 2.27 when we call a formula class "at least k-fold exponential".

Corollary 3.41 *The following formula classes are at least k-fold exponential:*

- $\Sigma_1(\Pi_{k-1}^0(\Delta_1^{\mathrm{U}}))$ *and thus* $\Sigma_1(\Pi_{k-1}^0(\Delta_1))$, $\underline{\Sigma_k}$, $\Sigma_1(FO(\Delta_1))$,
- $\Sigma_1(FO^{\leq_1,\leq_2})$ *and thus* $\Sigma_1(FO^{\mathrm{TC}})$.

Proof The proof is immediate from the preceding corollary. The reader who has skipped the (optional) Subsection 3.1.1 may prove Corollary 3.41 directly from Corollary 3.38 by building, for every k, a sentence

$$\exists X_1 \dots X_{2k} \varphi$$

that defines the k-fold exponential function f_k: choose $\varphi \in \Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)$ defining $\pi(\mathit{Num}_k)$, where π is as in the proof of Corollary 3.39. □

Corollary 3.41 provides the definability result for Theorems 2.24 and 2.29 and thus Corollary 2.25. Corollary 4.14 will state the corresponding non-definability result.

The reader who is interested only in the most important facts or impatient for the first separation result may continue with Chapter 4.

3.3.1 Digression: Length of Starfree Expressions

The notion of "starfreeness" of a picture language may alternatively be introduced via "starfree picture expressions". Starfree picture expressions over alphabet Ω are built as follows: For every $a \in \Omega$, the term a is a starfree picture expression, and if e and f are starfree picture expressions, then so are the terms $(e \mathbin{\varobar} f)$, $(e \ominus f)$, $\sim e$, and $(e \cup f)$.

The semantics is defined in a straightforward manner. Then a picture language is starfree iff it is defined by a starfree expression.

With this syntactic notion one may investigate the *length* $|r|$ of an expression r, i.e., the number of symbols in r.

In [Sto74] the author shows the following surprising fact. There is a sequence $(r_k)_{k\geq 1}$ of starfree word expressions such that the length of the shortest word matching r_k is not elementary in the length of r_k, i.e., not bounded by $s_i(|r|)$ for any fixed i.

A proof of this fact can also be extracted from the constructions of this section. Note that these hardly used the row concatenation $\ominus$. A careful analysis of these constructions—in particular those of Lemmas 3.36, 3.37 and Theorem 3.23—shows that the height-1 fragment of $\mathit{cycl}(\mathit{Num}_k)$ is defined by a starfree word expression r_k of length singly exponential in k, and we know that the length of the only word matching r_k is $f_k(1) \geq s_k(1)$, i.e., more than a tower of 2's of height k.

Concerning the length of starfree *picture* expressions, one may extract from these proofs that there is a sequence $(e_k)_{k\geq 1}$ of expressions such that e_k is of length $2^{\mathcal{O}(k)}$ and defines $\mathit{cycl}(\mathit{Num}_k)$.

3.4 Interrupted Counting—More Definability Results

This section contains definability results that are needed for some stronger separation claims, among them $\underline{\Sigma_k} \subsetneq \underline{\Delta_{k+1}}$ for the class of coloured grids or the class of graphs.

We will show that $\mathit{pref}_+(\mathit{Num}^+_k)$ is in Π^{loc}_{k-1} for every k. With respect to monadic logic this will in particular imply that the k-fold exponential function f_k is $\underline{\Delta_k}$-definable (rather than just $\underline{\Sigma_k}$-definable as shown in the previous section.)

The idea and the justification for the headline of this section is that the counting mechanism in Num^+_k is interrupted somewhere in the middle. In the previous section, the validity of the counting mechanism was essentially checked by excluding particular patterns as cyclic infixes. The trick works as well if that counting mechanism is interrupted somewhere, provided we forbid these patterns only as (non-cyclic) infixes. Thus this section will be quite analogous to the previous one.

The following lemma, for instance, is the counterpart of the "Cyclic Counting Lemma" 3.26 for the case of "interrupted counting". Here we benefit from considering binary number representations with least significant bit first.

Lemma 3.42 (Prefix Counting Lemma) *Let $n \geq 0$ and $x_0, \dots, x_n \in \{0,1\}^+$ with $m = |x_0| = \dots = |x_{n-1}| \geq |x_n| \geq 1$. Then the following are equivalent:*

1. $\mathit{dual}(x_j) = j \bmod 2^{|x_j|}$ *for all* $j \geq n$.
2. • $x_0 \in 0^+$,
   • $x_j\langle 0\rangle \neq x_{j+1}\langle 0\rangle$ *for all* $j \in \{0, \dots n-1\}$,
   • $\begin{pmatrix} x_j\langle i\rangle & x_{j+1}\langle i\rangle \\ x_j\langle i+1\rangle & x_{j+1}\langle i+1\rangle \end{pmatrix} \in \mathit{Allow}$ *for every* $j \in \{0, \dots n-1\}$ *and* $i \in \{0, \dots, |x_{j+1}| - 2\}$.

The proof is as simple as that of Lemma 3.26 and thus omitted.

We wish to show the following counterpart of Theorem 3.23.

Theorem 3.43 $\mathit{pref}_+(\mathit{Num}^+_k) \in \Pi^{\mathrm{loc}}_{k-1}$ *for every* $k \geq 1$.

The induction basis is given by the following counterpart of Theorem 3.24.

Proposition 3.44 *$pref_+(Num_1^+) \in \Pi_0^{\mathrm{loc}}$.*

Proof Let $\Gamma = \{0,1\}^{I_1}$. We will show that $pref_+(Num_1^+)$ is a local picture language. Let $\Delta \subseteq (\Gamma \cup \{\#\})^{2,2}$ as in Proposition 3.24.

Let Δ' be the following subset of $(\Gamma \cup \{\#\})^{2,2}$:

$$\begin{aligned}\Delta' = \quad & \Delta \cup \left\{ \begin{smallmatrix} \# & \boldsymbol{c} \\ \# & \boldsymbol{d} \end{smallmatrix} \;\middle|\; \boldsymbol{c},\boldsymbol{d} \in \{(0)_{num\text{-}1} \bowtie (0)_{end\text{-}1},\ \#\} \right\} \\ & \cup \left\{ \begin{smallmatrix} \boldsymbol{a} & \# \\ \boldsymbol{b} & \# \end{smallmatrix} \;\middle|\; \mathsf{pr}_{end\text{-}1}(\boldsymbol{a}) = 1 \quad\Leftrightarrow\quad \mathsf{pr}_{end\text{-}1}(\boldsymbol{b}) = \mathsf{pr}_{num\text{-}1}(\boldsymbol{a}) = 1 \right\} \\ & \cup \left\{ \begin{smallmatrix} \boldsymbol{a} & \# \\ \# & \# \end{smallmatrix} \;\middle|\; \mathsf{pr}_{end\text{-}1}(\boldsymbol{a}) = 1 \quad\Leftrightarrow\quad \mathsf{pr}_{num\text{-}1}(\boldsymbol{a}) = 1 \right\} \\ & \cup \left\{ \begin{smallmatrix} \# & \# \\ \boldsymbol{b} & \# \end{smallmatrix} \;\middle|\; \boldsymbol{b} \in \Gamma \right\}.\end{aligned}$$

Let P be a non-empty picture over Γ, say of size (m,n). For every $i \leq n-1$, let $x_i = P\langle 0,i\rangle \ldots P\langle m-1,i\rangle$.

Δ' is chosen in such a way that Δ' contains all $2{\times}2$-subblocks of $\hat{P}$ iff

- $\mathsf{pr}_{num\text{-}1}(x_0), \ldots, \mathsf{pr}_{num\text{-}1}(x_{n-1})$ fulfill the second property of the "Prefix Counting Lemma" 3.42, and
- (like in the proof of Proposition 3.24) for every x among $x_0, \ldots, x_{n-1}$, the $\mathsf{pr}_{end\text{-}1}(x)$ is determined by $\mathsf{pr}_{end\text{-}1}(x) = 0^i 1^j$ iff j is maximal such that 1^j is a suffix of $\mathsf{pr}_{num\text{-}1}(x)$.

Thus $\mathcal{L}(\Delta') = pref_+(Num_1)$. □

The following is the counterpart of Lemma 3.27.

Lemma 3.45 *Let $k \geq 1$, let $\Omega = \{0,1\}^{I_{k+1}}$. Let U_1, U_2, U_3 and bin_m as in Lemma 3.27. Let*

$$U_4 = (pref_+(0^*1))_{end\text{-}k} \otimes (\sim 0^*)_{num\text{-}k+1} \otimes \Omega^+.$$

The following are equivalent for every $P \in pref_+(Num_k^+) \otimes \Omega^{+,+}$ of height m.

1. *$top(P) \in pref_+(((\{0,1\}^*1)_{end\text{-}k} \otimes (bin_m)_{num\text{-}k+1})^+) \otimes \Omega^+$;*
2. *$top(P)$ has no infix of length $f_k(m)+2$ in $U_1 \cup U_2$ and no prefix in U_4.*

Furthermore, the following are equivalent for every $P \in pref_+(Num_k^+ \otimes OnlyTop_{k+1})$ of height m.

3. *$P \in pref_+(Num_{k+1}^+)$;*
4. *$top(P) \in pref_+(((\{0,1\}^*1)_{end\text{-}k}) \otimes (bin_m)_{num\text{-}k+1} \otimes (0^*1)_{end\text{-}k+1})^+) \otimes \Omega^+$;*
5. *$top(P)$ has no infix of length $f_k(m)+2$ in $U_1 \cup U_2$ and none of length $f_k(m)$ in U_3 and no prefix in U_4.*

The proof of this lemma is very similar to the proof of Lemma 3.27, except for the difference that the "Counting Lemma" 3.42 is used instead of the "Cyclic Counting Lemma" 3.26.

The following lemma is analogous to Lemma 3.36.

Lemma 3.46 *Let N such that $cycl(N^+) \in \Pi_{k-1}^{\mathrm{loc}}(\Omega)$. Let $L_1, L_2 \in \Pi_0^{\mathrm{loc}}(\Omega)$. Then there is a picture language L in $\Sigma_k^{\mathrm{loc}}(\Omega)$ such that for a non-empty picture P in $pref_+(N^+)$ we have that P has an infix in $(L_1 \cap cycl(N^+))L_2$ iff $P \in L$.*

Proof Simply choose $L = \Omega^{*,*}(L_1 \cap cycl(N^+))L_2\Omega^{*,*}$. Then $L \in \Sigma_k^{\mathrm{loc}}$. □

The proof of Lemma 3.46 is much simpler than that of Lemma 3.36 because it is so much simpler to exclude a particular pattern as a (non-cyclic) *infix* than as a *cyclic infix*.

Lemma 3.47 *Let $k \geq 1$ and $\Omega := \{0,1\}^{I_{k+1}}$. There is a picture language K_4 in $\Sigma_1^{\mathrm{loc}}(\Omega)$ such that for all $P \in pref_+(Num_k^+) \otimes OnlyTop_{k+1}$ we have $P \in K_4$ iff $top(P)$ has a prefix in U_4, where U_4 is as in Lemma 3.45.*

Proof Choose $K_4 = top^{-1}(U_4) \oplus \Omega^{*,*}$. Since U_4 is locally threshold testable, K_4 is in $\oplus$-cl$(\Pi_0^{\mathrm{loc}}) \subseteq \Sigma_1^{\mathrm{loc}}(\Omega)$. □

Now we are ready to do the proof of the main result of this section. It is very analogous to that of Theorem 3.23.

Proof of Theorem 3.43 By induction. The induction basis is provided by Proposition 3.44.

For the induction step, let U_1, U_2, U_3 as in Lemma 3.27, and let $k \geq 1$ with $pref_+(Num_k^+) \in \Pi_{k-1}^{\mathrm{loc}}$. Like in Lemma 3.37 one shows (using Lemma 3.46 instead of Remark 3.34 and Lemma 3.36) that there is a $K_3 \in \Pi_k^{\mathrm{loc}}(\{0,1\}^{I_{k+1}})$

such that $P \in K_3$ iff $top(P)$ has an infix of length $f_k(m)+2$ in $U_1 \cup U_2$ or one of length $f_k(m)$ in U_3. Choose K_4 as in the above lemma. By Lemma 3.45, Items 3–5,

$$pref_+(Num^+_{k+1}) = (pref_+(Num^+_k) \otimes OnlyTop_{k+1}) \setminus (K_3 \cup K_4) \in \Pi^{\mathrm{loc}}_k.$$

□

Like the previous section we close this one with a few corollaries that depend on results of Section 3.1.

Corollary 3.48 *Let $k \geq 1$, $\Omega = \{0,1\}^{I_k}$. There are picture languages L_0, L_1, L_2, L_3, L_4 in $\Pi^{\mathrm{loc}}_k(\Omega)$ (and thus in $\underline{\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)}$, in $\underline{\Pi^0_{k-1}(\Delta_1)}$, in $\underline{\Pi_{k-1}}$, in $\underline{FO(\Delta_1)}$, and in $\underline{FO^{\leq_1,\leq_2}}$) such that*

$$\begin{aligned}
\mathrm{size}(L_0) &= \{(m,n) \mid f_k(m) \text{ divides } n\},\\
\mathrm{size}(L_1) &= \{(m,n) \mid n = f_k(m)\},\\
\mathrm{size}(L_2) &= \{(m,n) \mid f_k(m) \nmid n\},\\
\mathrm{size}(L_3) &= \{(m,n) \mid n \neq f_k(m)\},\\
\mathrm{size}(L_4) &= \{(m,n) \mid n < f_k(m)\}.
\end{aligned}$$

Besides, there are no distinct pictures $P, P' \in L_i$ with $\mathrm{size}P = \mathrm{size}P'$.

Proof Choose

$$\begin{aligned}
L_0 &= pref_+(Num^+_k) \otimes top^{-1}((\{0,1\}^*1)_{end\text{-}k}) \otimes \Omega^+,\\
L_1 &= pref_+(Num^+_k) \otimes top^{-1}((0^*1)_{end\text{-}k}) \otimes \Omega^+,\\
L_2 &= pref_+(Num^+_k) \otimes top^{-1}((\{0,1\}^*0)_{end\text{-}k}) \otimes \Omega^+,\\
L_3 &= pref_+(Num^+_k) \otimes top^{-1}((\sim(0^*1))_{end\text{-}k}) \otimes \Omega^+,\\
L_4 &= pref_+(Num^+_k) \otimes top^{-1}((0^+)_{end\text{-}k}) \otimes \Omega^+.
\end{aligned}$$

Membership of L_0, L_1, L_2, L_3, L_4 in Π^{loc}_{k-1} follows from Theorem 3.43 and the fact that the word languages $\{0,1\}^*0$, $\sim(0^*1)$, and 0^+ are locally threshold testable, so that their top-pre-images are Π^{loc}_0.

Membership in the other listed classes follows from Theorem 3.11 and Corollary 3.12. The additional claim is immediate from the fact that there are no distinct pictures $P, P' \in pref_+(Num^+_k)$ with $\mathrm{size}P = \mathrm{size}P'$. □

Similar to Corollary 3.39 one may prove that the above corollary remains valid if $\{0,1\}^{I_k}$ is replaced by $\{0,1\}$ and f_k is replaced by the function $m \mapsto (4k+1)f_k(m)$.

The following is similar to Corollary 3.41. It provides the first half of the proof of Theorem 2.29, i.e., the fact that for every k, the class $\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1))$ (and hence the other, larger formula classes of that theorem) are at least k-fold exponential.

Corollary 3.49 *Let $k \geq 1$. The formula classes $\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1))$ and $\Delta_1(FO^{\leq_1,\leq_2})$ are at least k-fold exponential, and thus so are $\Delta_1(\Pi^0_{k-1}(\Delta_1))$, $\Delta_1(FO(\Delta_1))$, Δ_k, and $\Delta_1(FO^{\mathrm{TC}})$.*

Proof Let $\overline{X}$ be the variable tuple $(X_1, \dots, X_{2k})$. Let $\pi : \{0,1\}^{I_k} \to \{0,1\}^{2k}$ be a bijection. By Theorem 3.43 and 3.12, we may choose $\varphi(\overline{X})$ in $\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)$ and in $FO^{\leq_1,\leq_2}$ such that φ defines $\pi(pref_+(Num^+_k))$. We choose a first-order formula $\psi(\overline{X})$ such that ψ defines $\pi(top^{-1}(0^*1)_{end\text{-}k})$. Then $\psi(\iota\overline{X}\varphi(\overline{X}))$ is (by the additional claim of Corollary 3.48) an allowed $\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1))$-formula, and it asserts for the grid of size (m,n) that $n = f_k(m)$.

Membership in the other classes follows from the calculation rules of Subsection 2.2.3. □

Like Corollary 3.40 one may show that there even is a formula of the form $\psi(\iota X_1\varphi(X_1))$, i.e., with the ι-operator ranging over one set variable only, that defines a k-fold exponential function and is in $\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1))$.

Corollary 3.49 provides the definability result for Theorem 2.29. The corresponding non-definability results can be found in the next chapter, namely Corollaries 4.16, 4.22.

3.5 Twin Counting—Yet More Definability Results

In this section we will develop some definability results that allow for even more separation results than those of Section 3.3, among them $\underline{\Sigma_k} \neq \underline{\Pi_k}$ over *non-coloured* grids and $\underline{FO(\Sigma_k)} \subsetneq \underline{FO(\Delta_{k+1})}$ for the classes of grids or of graphs.

A preliminary communication of the results of this section has appeared as [Mat98a].

These definability results depend on some different approach unlike Section 3.4, which was very analogous to Section 3.3. This time we leave the detour of the starfree alternation hierarchy somewhat earlier so that the main theorem states the definability by monadic second-order formulas rather than membership in some level of the starfree alternation hierarchy.

However, a similarity between Sections 3.3 and 3.4 on one hand and Section 3.5 on the other is that we construct "slim" picture languages by exploiting counting mechanisms. But this time there are two of them in the same picture and two patterns of the two mechanisms belong together like twins, which explains the title of this section.

Recall the definition of the functions $f_k : \mathbb{N} \to \mathbb{N}$.

$$\begin{aligned} f_1(m) &= 2^m \\ f_{k+1}(m) &= f_k(m) 2^{f_k(m)}. \end{aligned}$$

Furthermore, let $g_k(m) = \mathrm{lcm}\{f_k(m)+1, \dots, 2f_k(m)\}$ for every m and every k. We will show in Proposition 3.67 that g_k is $(k+1)$-fold exponential.

The main result of this section is Theorem 3.61, page 106. It states that, for every k, there are two picture languages L_k, M_k definable in certain fragments of the first-order closure of Δ_k and such that $\mathrm{size}(L_k) = \{(m,n) \mid g_k(m) \text{ divides } n\}$ and $\mathrm{size}(M_k) = g_k$. The picture language L_k is even definable in Π_k. In Chapter 4 we will show that L_k and M_k are not definable in Σ_k and neither in $FO(\Sigma_{k-1})$ because g_k is a $(k+1)$-fold exponential function.

This implies the separation results $\underline{FO(\Sigma_k)} \subsetneq \underline{FO(\Delta_k)}$ and $\underline{\Sigma_k} \neq \underline{\Pi_k}$.

3.5.1 Attributed Alphabets and Free Variables

On page 31 we defined how to associate a non-empty picture over $\{0,1\}^t$ to a t-bit grid and then what we call the picture language over $\{0,1\}^t$ *defined by* a formula with free variables among $X_1, \dots, X_t$. Since now we wish to construct formulas defining picture languages over attributed alphabets, we require the (straightforward) adaptions of these definitions to this case.

Let I be a set of attributes, let $(X_\mu)_{\mu\in I}$ be a tuple of monadic second-order variables. An *I-coloured grid* is a grid expanded by monadic predicates $(X_\mu)_{\mu\in I}$. To every I-coloured grid we associate a non-empty picture over $\{0,1\}^I$. We say that this picture fulfills a formula $\varphi((X_\mu)_{\mu\in I})$ iff the I-coloured grid it is associated to makes (with the implicitly given assignment) φ true. A formula $\varphi((X_\mu)_{\mu\in I})$ *defines* the picture language of those pictures over $\{0,1\}^I$ that fulfill φ.

It would have been a natural idea to let the attributes themselves serve as free set variables, but this conflicts with our conventions to use capital letters for second-order variables.

3.5.2 The Twin Language

Recall the attribute sets $I_k = \{\mathit{num}\text{-}1, \dots, \mathit{num}\text{-}k, \mathit{end}\text{-}1, \dots, \mathit{end}\text{-}k\}$ for every k. Let act (for "active") be a fresh attribute not in any of the I_k.

For this subsection we fix some $k \geq 1$.

Let ι be some bijection defined on $I_k \cup \{\mathit{act}\}$ such that $I_k \cup \{\mathit{act}\}$ and its image under ι are disjoint. We use ι also to denote the mapping $\{0,1\}^{I_k\cup\{\mathit{act}\}} \to \{0,1\}^{\iota(I_k\cup\{\mathit{act}\})}$ defined by $\iota(a) : \iota(\mu) \mapsto a(\mu)$ for all $\mu \in I_k \cup \{\mathit{act}\}$. This alphabet projection is further extended to pictures and picture languages in the usual way.

We turn to the definition of the twin language. We proceed in several steps.

$$\begin{aligned} L_k &= \mathit{Num}_k \otimes (1^{+,+})_{\mathit{act}}, \\ S &= (0^{+,1})_{I_k\cup\{\mathit{act}\}}. \end{aligned}$$

So L_k is a picture language over alphabet $\{0,1\}^{I_k\cup\{\mathit{act}\}}$ whose restriction to I_k is the counting pattern introduced in Definition 3.22. The picture language S is over the same alphabet and contains all columns (i.e., pictures of length 1) that have, at each position, a 0 in every component.

For all $m \geq 1$ and all $j \in \{1, \dots, f_k(m)\}$ let

$$\begin{aligned} \mathit{twin}_{k,m,j} &= \{P \in (L_k \obar S^j) \otimes \iota(S^j \obar L_k) \mid \overline{\underline{P}} = m\}, \\ \mathit{ptwin}_{k,m,j} &= \{P \in \mathit{pref}_+(L_k) \otimes \iota(S^+) \mid \overline{\underline{P}} = m\} \\ &\quad \cup \{P \in (L_k \obar S^+) \otimes \iota(S^j \obar \mathit{pref}_+(L_k)) \mid \overline{\underline{P}} = m\}, \\ \mathit{xtwin}_k &= \bigcup_{m\geq 1} \Big(\Big(\bigcup_j \mathit{twin}_{k,m,j} \Big)^* \obar \Big(\bigcup_j \mathit{ptwin}_{k,m,j} \Big) \Big), \\ \mathit{iso\text{-}xtwin}_k &= \bigcup_{m\geq 1} \Big(\bigcup_j (\mathit{twin}_{k,m,j})^* \obar \mathit{ptwin}_{k,m,j} \Big). \end{aligned}$$

These picture languages are over alphabet $\{0,1\}^{\{I_k\cup\iota(I_k)\cup\{\mathit{act},\iota(\mathit{act})\}}$. Figure 3.2 illustrates a member of $\mathit{twin}_{k,m,j}$ (together with two other members of $\mathit{ptwin}_{k,m,j}$). Roughly speaking, a picture in $\mathit{twin}_{k,m,j}$ results by joining two copies (over disjoint alphabets) of the picture P_{km} of height m in Num_k

in such a way that these copies may overlap. Additionally, the components associated with the two copies of attribute *act* mark those positions that belong to the respective copy of P_{km}, i.e., where the respective counting process is "active". In the position in the areas to the right and left, in which one counting process is not active, the components corresponding to its copy of the alphabet have a 0. Note that there is no area in which both counting processes are non-active because a picture of height m in L_k has length $f_k(m)$. Let us call the width j of each of the areas to the right and left, where one counting process is not active, "offset". Note that the width of the picture of height m in $twin_{k,m,j}$ is $f_k(m)+j$.

A picture in $ptwin_{k,m,j}$ ("p" standing for "prefix") is of one of two different types (see last two boxes in Figure 3.2). Either it results from a non-empty prefix of the picture over alphabet $\{0,1\}^{I_k}$ of height m in Num_k by joining with a picture of zeros over a disjoint copy of that alphabet, or it results from a picture in $twin_{k,m,j}$ by clipping off something—possibly empty—in the rightmost area. In both cases each of the two copies of the "*act*"-component carries a 1 whenever the respective counting process is active. This way, for every $j \in \{1,\dots,f_k(m)\}$ and every $n \in \{1,\dots,f_k(m)+j\}$ there is exactly one picture of size (m,n) in $ptwin_{k,m,j}$.

A picture in $xtwin_k$ is a concatenation of several pictures from $twin_{k,m,j}$ (for the same height m, but possibly different "offsets" j) and one picture from $ptwin_{k,m,j}$.

A picture in $iso\text{-}xtwin_k$ is a picture in $xtwin_k$ where the offsets j are equal in all of the subblocks from $twin_{k,m,j}$ and $ptwin_{k,m,j}$.

The following theorem will imply all other definability results of this section.

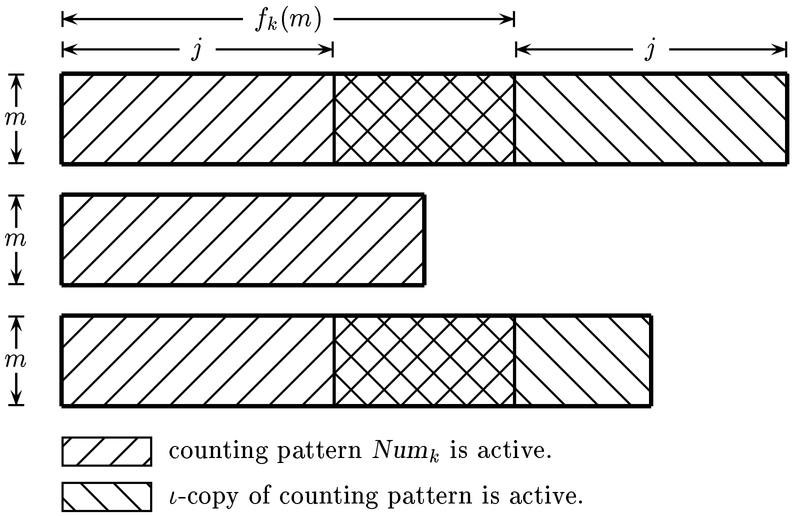
Theorem 3.50 *$iso\text{-}xtwin_k$ is definable in $\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)$.*

We will prove Theorem 3.50 for the cases $k=1$ and $k>1$ separately.

How to Prove Theorem 3.50, case $k=1$.

In this part we assume that $k=1$.

We have reached one of the few points where our notations for attributed alphabets, which are motivated in the beginning of this chapter, make things more complicated rather than easier. That is why we will sometimes use the following abbreviation. Consider the alphabet $\{0,1\}^{\{act,\iota(act)\}}$. In the



The first picture is in $twin_{k,m,j}$, all three are in $ptwin_{k,m,j}$.

Figure 3.2: The twin language

following, we will sometimes write $\binom{i}{j}$ instead of $(i)_{\mathrm{act}} \otimes (j)_{\iota\{\mathrm{act}\}}$ for $i,j \in \{0,1\}$, as if $\{0,1\}^{\{act,\iota(act)\}}$ was equal to the alphabet $\{0,1\}\times\{0,1\}$.

Remark 3.51 Let M be the following word language over $\{0,1\}^{\{act,\iota(act)\}}$ (in the just explained short-hand notation).

$$\left(\left(\binom{1}{0}^{+}\binom{1}{1}^{*}\binom{0}{1}^{+}\right)^{*}\binom{1}{0}^{*}\right)\setminus\{\varepsilon\}.$$

Then M is local. □

Remark 3.52 Let M be the local word language over $\{0,1\}^{\{act,\iota(act)\}}$ defined in Remark 3.51. Let

$$M' = top^{-1}(M) \cap \left(\bigcup_a a^{+,1}\right)^{+},$$

where the union ranges over $a \in \{0,1\}^{\{act,\iota(act)\}}$. Then M' is the set of non-empty pictures over $\{0,1\}^{\{act,\iota(act)\}}$ whose top row is in M and in which two positions have the same letter if they are in the same column. Clearly, M' is local by Remark 3.51. In the above short-hand notation, $M' = \left(\left(\left(\binom{1}{0}^{+,+}\binom{1}{1}^{*,*}\binom{0}{1}^{+,+}\right)^{*}\binom{1}{0}^{*,*}\right) \setminus \{\varepsilon\}.$ □

Let us use the abbreviation $\mathbf{0} = (0)_{num\text{-}1} \bowtie (0)_{end\text{-}1} \bowtie (0)_{act}$.

Proposition 3.53 *The picture language* $N = \left(\mathbf{0}^{*,*} \oplus (L_1 \oplus \mathbf{0}^{+,+})^* \oplus pref(L_1)\right) \setminus \{\varepsilon\}$ *is local.*

Proof Let

$$\begin{aligned} \mathbf{1} &= (1)_{num\text{-}1} \bowtie (1)_{end\text{-}1} \bowtie (1)_{act}, \\ a_0 &= (0)_{num\text{-}1} \bowtie (0)_{end\text{-}1} \bowtie (1)_{act}. \end{aligned}$$

We use the alphabet projection $\varphi : \{0,1\}^{I_k} \cup \{\#\} \to \{0,1\}^{I_k \cup \{act\}} \cup \{\#\}$,

$$a \mapsto \begin{cases} \# & \text{if } a = \#, \\ a \bowtie (1)_{act} & \text{else.} \end{cases}$$

Let Δ' be the set of 2×2-pictures over $\Gamma \cup \{\#\}$ defined in the proof of Proposition 3.44, i.e., $\mathcal{L}(\Delta') = pref_+(Num_1^+)$. (Then $\mathcal{L}(\varphi(\Delta')) = pref_+(L_1)$.)

Let Δ'' result from $\varphi(\Delta')$ by removing $\begin{smallmatrix} \mathbf{1} & a_0 \\ \# & \# \end{smallmatrix}$ and adding the 2×2-subblocks of $\begin{smallmatrix} \# & \# & \# \\ \mathbf{1} & \mathbf{0} & a_0 \\ \mathbf{1} & \mathbf{0} & a_0 \\ \# & \# & \# \end{smallmatrix}$ and $\begin{smallmatrix} \# & \# & \# & \# \\ \# & \mathbf{0} & \mathbf{0} & \# \\ \# & \mathbf{0} & \mathbf{0} & \# \\ \# & \# & \# & \# \end{smallmatrix}$. Then $\mathcal{L}(\Delta'') = N$. □

Proposition 3.54 *$xtwin_1$ is local.*

Proof We claim

$$xtwin_1 = N \otimes \iota(N) \otimes M', \tag{3.7}$$

where M' is chosen as in Remark 3.52 and N is chosen as in the previous proposition. Then we are finished because the class of local picture languages is closed under intersection (and thus, by Remark 3.20, under join).

As indicated earlier, we write $\binom{i}{j}$ instead of $(i)_{act} \bowtie (j)_{\iota(act)}$ and $\Omega = \{0,1\}^{\{I_k \cup \iota(I_k) \cup \{act, \iota(act)\}}$ as an abbreviation

For the inclusion "$\subseteq$" in (3.7), let $P \in xtwin_1$, say of height m. By definition, there are $n \geq 0$ and pictures $P_0, \ldots, P_n$ such that $P = P_0 \ldots P_n$ and $P_0, \ldots, P_{n-1} \in \bigcup_j twin_{1,m,j}$ and $P_n \in ptwin_{1,m,j}$. We show that

1. $\mathsf{restr}_{\{act,\iota(act)\}}(top(P)) \in M$, where M is as in Remark 3.51;
2. $\mathsf{restr}_{I_k \cup \{act\}}(P) \in N$;
3. $\mathsf{restr}_{\iota(I_k \cup \{act\})}(P) \in \iota(N)$.

Then we are finished with the inclusion "$\subseteq$" in (3.7) because membership of $\mathsf{restr}_{\{act,\iota(act)\}}(P)$ in M' follows from Item 1 and the fact that $\mathsf{pr}_{act}(P)$ and $\mathsf{pr}_{\iota(act)}(P)$ are in $(1^{+,1} \cup 0^{+,1})^+$ by Items 2 and 3, respectively.

Let us turn to the proof of Items 1, 2, 3. For every $j < n$ we have

$$\mathsf{restr}_{\{act,\iota(act)\}}(top(P_j)) \in \binom{1}{0}^+\binom{1}{1}^*\binom{0}{1}^+, \tag{3.8}$$

$$\mathsf{restr}_{I_1 \cup \{act\}}(P_j) \in L_1 \oplus \mathbf{0}^{+,+}, \tag{3.9}$$

$$\mathsf{restr}_{\iota(I_1 \cup \{act\})}(P_j) \in \iota(\mathbf{0}^{+,+} \oplus L_1). \tag{3.10}$$

Concerning P_n, we distinguish the cases (1) that $P_n \in pref_+(L_1) \otimes \iota(S^+)$ and (2) that $P_n \in (L_1 \oplus S^+) \otimes \iota(S^j \oplus pref_+(L_1))$ for some $j \leq f_1(m)$. In the first case we have that

$$\begin{aligned} \mathsf{restr}_{\{act,\iota(act)\}}(top(P_n)) &\in \binom{1}{0}^+, \\ \mathsf{restr}_{I_k \cup \{act\}}(P_n) &\in pref_+(L_1), \\ \mathsf{restr}_{\iota(I_k \cup \{act\})}(P_n) &\in \iota(\mathbf{0}^{+,+}), \end{aligned}$$

which implies together with (3.8)–(3.10) Items 1, 2, 3. In the second case we have that

$$\begin{aligned} \mathsf{restr}_{\{act,\iota(act)\}}(top(P_n)) &\in \binom{1}{0}^+\binom{1}{1}^*\binom{0}{1}^+, \\ \mathsf{restr}_{I_k \cup \{act\}}(P_n) &\in L_1 \oplus \mathbf{0}^{+,+}, \\ \mathsf{restr}_{\iota(I_k \cup \{act\})}(P_n) &\in \iota(\mathbf{0}^{+,+} \oplus pref_+(L_1)), \end{aligned}$$

which implies Items 1, 2, 3 again. This completes the proof of the inclusion "$\subseteq$" in (3.7).

For the converse inclusion, let $P \in N \otimes \iota(N) \otimes M'$. There are $n \geq 0$ and pictures $P_0, \ldots, P_n$ such that $P = P_0 \ldots P_n$ and

$$\mathsf{restr}_{\{act,\iota(act)\}}(top(P_j)) \in \binom{1}{0}^{+}\binom{1}{1}^{*}\binom{0}{1}^{+}$$

for all $j < n$, and $\mathsf{restr}_{\{act,\iota(act)\}}(top(P_n)) \in \binom{1}{0}^{*}$.

By choice on N we have $P \neq \varepsilon$. We distinguish between two cases: Either $P_n = \varepsilon$ or $P_n \neq \varepsilon$.

In case $P_n = \varepsilon$ one shows $P_0, \ldots, P_{n-2} \in \bigcup_j twin_{1,m,j}$ and $P_{n-1} \in (L_1 \oplus S^+) \otimes \iota(S^j \oplus pref_+(L_1)) \subseteq ptwin_{1,m,j}$ for some j, thus $P \in xtwin_1$.

In case $P_n \neq \varepsilon$ one shows $P_0, \ldots, P_{n-1} \in \bigcup_j twin_{1,m,j}$ and $P_n \in pref_+(L_1) \otimes \iota(S^+) \subseteq ptwin_{1,m,j}$, thus also $P \in xtwin_1$. □

We use the previous proposition to prove the following, which is the claim of Theorem 3.50 for the case $k = 1$.

Proposition 3.55 *iso-xtwin*$_1$ *is definable in* Δ_1^{U}.

Proof There is a first-order formula $less(X_{leq}, X_{\iota(act)})$ such that *less* asserts for a grid of height m that

$$X_{leq} = \{(0,j) \mid (0,0), \ldots, (0,j) \notin X_{\iota(act)}\},$$

i.e., that X_{leq} is the set of all top-row positions up to (but not including) the first position in $X_{\iota(act)}$. Such a formula *less* is given by

$$\begin{aligned} less(X_{leq}, X_{\iota(act)}) = \;& \forall x \big(X_{leq}x \leftrightarrow \\ & \big(top(x) \wedge \neg X_{\iota(act)}x \wedge \forall x'(S_2x'x \rightarrow X_{leq}x')\big)\big). \end{aligned}$$

There is a first-order formula $maxless(X_{leq}, X_{maxleq})$ that asserts that X_{maxleq} contains those positions that are in a column whose top row position is in X_{leq} and has no right successor in X_{leq}.

There is a first-order formula $offsets(X_{offs}, X_{\iota(act)})$ that asserts that X_{offs} contains those top row positions that are not in $X_{\iota(act)}$ but whose right successor is.

There is a first-order formula $transport(X_{row}, Y, X_{num\text{-}1})$ that asserts that X_{row} is a union of rows with $X_{row} \cap Y = X_{num\text{-}1} \cap Y$.

Choose

$$\begin{aligned} \varphi(\overline{X}) = \;& less(X_{leq}, X_{\iota(act)}) \wedge maxless(X_{leq}, X_{maxleq}) \\ & \wedge transport(X_{row}, X_{maxleq}, X_{num\text{-}1}). \end{aligned}$$

Then for every $I_1 \cup \{\iota(act)\}$-coloured grid R there is exactly one tuple $\overline{X} = (X_{leq}, X_{maxleq}, X_{offs}, X_{row})$ of subsets of dom$R$ that fulfills φ because $less(X_{leq}, X_{\iota(act)}) \wedge maxless(X_{leq}, X_{maxleq})$ determines the column that X_{maxleq} contains (it is the first hatched one in Figure 3.3), and thus $transport(X_{row}, X_{maxleq}, X_{num\text{-}1})$ determines X_{row} uniquely.

Choose[4]

$$\psi(\overline{X}, (X_\mu)_{\mu \in (I_1 \cup \iota(I_1) \cup \{act, \iota(act)\})}) = transport(X_{row}, X_{offs}, X_{num\text{-}1}).$$

Then $\psi(\iota\overline{X}(\varphi), (X_\mu))$ is an allowed Δ_1^{U}-formula which asserts for a $(I_1 \cup \iota(I_1) \cup \{act, \iota(act)\})$-coloured grid whose associated picture P is in $xtwin_1$ that P is in *iso-xtwin*$_1$ because it asserts that the colouring of those rows that precede the rows in which the second copy of the counting process starts are equal. See Figure 3.3.

Since $xtwin_1$ is local by the previous Proposition, there is a first-order formula π that asserts for a $(I_1 \cup \iota(I_1) \cup \{act, \iota(act)\})$-coloured grid that its associated picture is in $xtwin_1$. Thus $\pi \wedge \psi(\iota\overline{X}(\varphi), (X_\mu))$ defines *iso-xtwin*$_1$. This formula is Δ_1^{U} because Δ_1^{U} is closed under intersection.

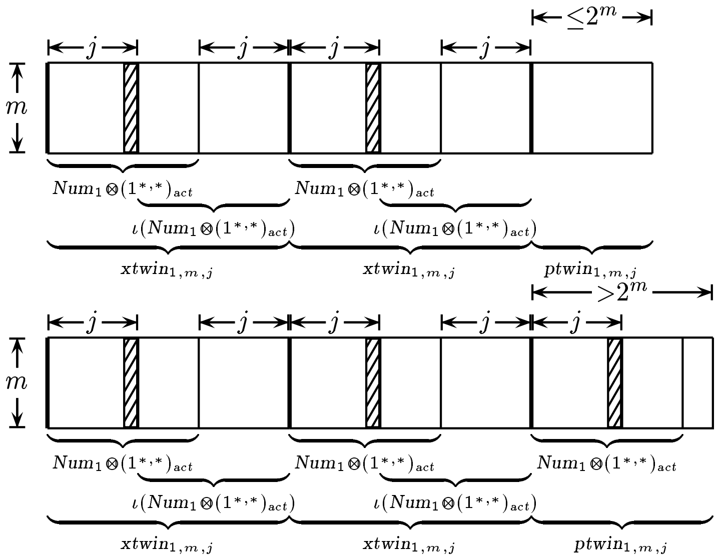
This completes the proof of Proposition 3.55 and thus the proof of Theorem 3.50 for the case $k = 1$. □

How to Prove Theorem 3.50, case $k > 1$

For the proof of Theorem 3.50 we will need the following proposition, which will be proved in Subsection 3.5.3. Recall that $(0)_I$ denotes the letter in $\{0,1\}^I$ all of whose components are 0.

Proposition 3.56 $cycl(L_k \oplus (0)^{*,*}_{I_k \cup \{act\}})$ *is in* Π^{loc}_{k-1}.

[4] List of potentially free variables is enlarged for formal reasons. Note that the two tuples are disjoint.



The hatched columns are coloured identically. This ensures that the offset is always the same j.

Figure 3.3: Two pictures in *iso-xtwin*$_1$.

Proof of Theorem 3.50, case $k \geq 2$. We use the following abbreviations: $\mathbf{0} = (0)_{I_k \cup \{act\}}$, the mapping $\varphi : \{0,1\}^{I_k} \to \{0,1\}^{I_k \cup \{act\}}$, $a \mapsto a \bowtie (1)_{act}$, and $\Omega = \{0,1\}^{I_k \cup \iota(I_k) \cup \{act, \iota(act)\}}$, and $L = L_k = Num_k \otimes (1^{+,+})_{act}$.

Let M' be the local picture languages of Remark 3.52. For a picture language U over some alphabet Γ we will write $\overline{U}$ for the complement $\Gamma^{+,+} \setminus U$ of U in $\Gamma^{+,+}$.

We wish to show that *iso-xtwin*$_k$ is the set of all non-empty pictures P over Ω that fulfill the following properties:

1. $\mathsf{restr}_{\{act, \iota(act)\}}(P) \in M'$.
2. P has no prefix in
$$\begin{aligned} M_1 &= (1^{+,+})_{act} \otimes \overline{pref_+(L)} \otimes \Omega^{+,+}, \\ M_2 &= (1^{*,*}0^{+,1})_{act} \otimes (\overline{L} \oplus \Gamma^{+,1}) \otimes \Omega^{+,+}. \end{aligned}$$
3. P has no infix in
$$\begin{aligned} N_1 &= ((0)_{act} \bowtie \Omega)^{+,1} \big((1^{+,+})_{act} \otimes \overline{L} \otimes \Omega^{+,+}\big) ((0)_{act} \bowtie \Omega)^{+,1}, \\ N_2 &= ((0)_{\iota(act)} \bowtie \Omega)^{+,1} \big((1^{+,+})_{\iota(act)} \otimes \iota(\overline{L}) \otimes \Omega^{+,+}\big) ((0)_{\iota(act)} \bowtie \Omega)^{+,1}, \\ N_3 &= ((0)_{\iota(act)} \bowtie \Omega)^{+,1} \big((1^{+,+}0^{+,+})_{\iota(act)} \otimes \overline{cycl(\varphi(L)\mathbf{0}^{*,*})} \otimes \Omega^{+,+}\big) \\ &\quad ((1)_{\iota(act)} \bowtie \Omega)^{+,1}. \end{aligned}$$
4. P has no suffix in
$$\begin{aligned} K_1 &= ((0)_{act} \bowtie \Omega)^{+,1} \big((1^{+,+})_{act} \otimes \overline{pref_+(L)} \otimes \Omega^{+,+}\big), \\ K_2 &= ((0)_{\iota(act)} \bowtie \Omega)^{+,1} \big((1^{+,+})_{\iota(act)} \otimes \iota(\overline{pref_+(L)}) \otimes \Omega^{+,+}\big). \end{aligned}$$

Before we verify that these Properties 1–4 characterize the pictures in *iso-xtwin*$_k$, let us note how this implies that *iso-xtwin*$_k \in \underline{\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)}$. By choice of L and Theorems 3.23, 3.43 and Proposition 3.56, the languages $pref_+(L)$, L, $\iota(L)$, $cycl(\varphi(L)\mathbf{0}^{*,*})$, and $\iota(pref_+(L))$ are Π^{loc}_{k-1}, thus their complements are $\Sigma^{\mathrm{loc}}_{k-1}$ and hence in $\underline{\Sigma^0_{k-1}(\Delta^{\mathrm{U}}_1)}(\Omega)$. Since $\underline{\Sigma^0_{k-1}(\Delta^{\mathrm{U}}_1)}$ is closed under $\otimes$ and (by Lemma 3.19) under column concatenation, the picture languages $M_1, M_2, N_1, N_2, N_3, K_1, K_2$ are in $\underline{\Sigma^0_{k-1}(\Delta^{\mathrm{U}}_1)}$. Since $\underline{\Sigma^0_{k-1}(\Delta^{\mathrm{U}}_1)}$ is closed under union and column concatenation, the picture languages $(M_1 \cup M_2)\Omega^{*,*}$, $\Omega^{*,*}(N_1 \cup N_2 \cup N_3)\Omega^{*,*}$, and $\Omega^{*,*}(K_1 \cup K_2)$ are in $\underline{\Sigma^0_{k-1}(\Delta^{\mathrm{U}}_1)}$.

Thus the picture language $\textit{iso-xtwin}_k = (M' \otimes \Omega^{+,+}) \setminus \big((M_1 \cup M_2)\Omega^{*,*} \cup \Omega^{*,*}(N_1 \cup N_2 \cup N_3)\Omega^{*,*} \cup \Omega^{*,*}(K_1 \cup K_2)\big)$ is in $\Pi^0_{k-1}(\underline{\Delta^{\mathrm{U}}_1})$.

It is easy to verify that every picture P in $\textit{iso-xtwin}_k$ fulfills Properties 1–4. Let us consider the converse direction. Let P be a non-empty picture over Ω with Properties 1–4. Again, we sometimes write $\binom{i}{j}$ instead of $(i)_{\mathrm{act}} \otimes (j)_{\iota(\mathrm{act})}$ for every $i, j \in \{0,1\}$. By Property 1 there are $n \geq 1$ and pictures $P_1, \ldots, P_n$ such that $P = P_1 \ldots P_n$ and

$$\mathsf{restr}_{\{act,\iota(act)\}}(P_j) \in \binom{1}{0}^{+,+}\binom{1}{1}^{*,*}\binom{0}{1}^{+,+}$$

for all $j \leq n-1$ and

$$\mathsf{restr}_{\{act,\iota(act)\}}(P_n) \in \binom{1}{0}^{+,+}\binom{1}{1}^{*,*}\binom{0}{1}^{+,+} \cup \binom{1}{0}^{+,+}.$$

Thus for every $j \leq n$ there are P'_j, P''_j, P'''_j such that

- $P = P'_j P''_j P'''_j$,
- $\mathsf{restr}_{\{act,\iota(act)\}}(P'_j) \in \big((1)_{act} \bowtie (0)_{\iota(act)}\big)^{+,+}$,
- $\mathsf{restr}_{\{act,\iota(act)\}}(P''_j) \in \big((1)_{act} \bowtie (1)_{\iota(act)}\big)^{*,*}$,
- $\mathsf{restr}_{\{act,\iota(act)\}}(P'''_j) \in \big((0)_{act} \bowtie (1)_{\iota(act)}\big)^{*,*}$,
- $P'''_j = \varepsilon \implies j = n, P''_n = \varepsilon$.

We have that

$$\forall j \leq n : P'''_j \neq \varepsilon \Longrightarrow P'_j P''_j \in (1^{+,+})_{\mathrm{act}} \otimes L \otimes \Omega^{+,+}. \tag{3.11}$$

For $j = 1$ this follows because P has no prefix in M_2. For $j > 1$ this follows because P has no infix in N_1. We also have

$$\forall j < n : P''_j P'''_j \in \iota((1^{+,+})_{\mathrm{act}} \otimes L) \otimes \Omega^{+,+}, \tag{3.12}$$

because P has no infix in N_2. Besides,

$$P'''_n \neq \varepsilon \Longrightarrow P''_n P'''_n \in \iota((1^{+,+})_{\mathrm{act}} \otimes \textit{pref}_+(L)) \otimes \Omega^{+,+}, \tag{3.13}$$

because P has no suffix in K_2. Moreover,

$$P'''_n = \varepsilon \Longrightarrow P_n \in (1^{+,+})_{\mathit{act}} \otimes \textit{pref}_+(L) \otimes \Omega^{+,+}. \tag{3.14}$$

In case $n = 1$ this is true because P has no prefix in M_1. In case $n \geq 2$ this is true because P has no suffix in K_1. Finally,

$$\forall j \leq n : P'''_j \neq \varepsilon \Longrightarrow |P'_j| = |P'_1|. \tag{3.15}$$

This is shown by induction on j. It is immediate for the case $j = 1$. So let $j < n$ with $|P'_j| = |P'_1|$ and $P'''_{j+1} \neq \varepsilon$. Since P has no infix in N_3 we have that $P''_j P'''_j P'_{j+1} \in \textit{cycl}(\varphi(L)\mathbf{0}^{*,*}) \otimes \Omega^{+,+}$. Since $\mathsf{pr}_{act}(P''_j), \mathsf{pr}_{act}(P'_{j+1}) \in 1^{*,*}$ and $\mathsf{pr}_{act}(P'''_j) \in 0^{*,*}$, this implies $P''_j P'_{j+1} \in \textit{cycl}(\varphi(L)) \otimes \Omega^{+,+}$, thus $|P''_j| + |P'_{j+1}| = f_k(m)$, where $m = \overline{\underline{P}}$. By (3.11) it follows $|P'_j| + |P''_j| = f_k(m) = |P''_j| + |P'_{j+1}|$, thus $|P'_{j+1}| = |P'_j|$. So by induction hypothesis $|P'_{j+1}| = |P'_1|$. This completes the induction and thus the proof for (3.15).

Now we are ready to show that $P \in \textit{iso-xtwin}_k$. Choose $j = |P'_1|$. By (3.11), (3.12), (3.15) and choice of j we have that $P_j \in \textit{twin}_{k,m,j}$ for every $j \leq n-1$. If $P'''_n \neq \varepsilon$, then $P_n \in \textit{ptwin}_{k,m,j}$ by (3.11), (3.13), (3.15) and choice of j; if $P'''_n = \varepsilon$, then $P_n \in \textit{ptwin}_{k,m,j}$ (independently from the choice of j by (3.14)).

This implies $P = P_1 \ldots P_n \in (\textit{twin}_{k,m,j})^* \textit{ptwin}_{k,m,j} \subseteq \textit{iso-xtwin}_k$. This completes the proof of Theorem 3.50. □

3.5.3 Inserting a Sequence of Zero Columns

In this subsection we will prove Proposition 3.56. First we shall sketch informally what has to be done. Let $\alpha(\textit{Num}_k) = (\textit{Num}_k \otimes (1^{+,+})_{act} \oplus (0^{+,+})_{I_k \cup \{act\}})$. We wish to show that $\textit{cycl}(\alpha(\textit{Num}_k))$ is in Π^{loc}_{k-1}, exploiting that $\textit{cycl}(\textit{Num}_k)$ is Π^{loc}_{k-1}.

Intuitively, the pictures in $\textit{cycl}(\alpha(\textit{Num}_k))$ are those that result from a picture P in $\textit{cycl}(\textit{Num}_k)$ by joining an additional component *act* that is always 1, and then cyclically inserting an infix from $(0^{+,+})_{I_k \cup \{act\}}$ between the two columns of P that mark the "end" and the "beginning" of the corresponding $\textit{Num}_k$-picture in $\textit{cycl}(\{P\})$.

Our strategy to prove Proposition 3.56 is to give a sufficient criterion for such a mapping α to respect the local hierarchy. For this purpose we abstract somewhat from the particular situation we have. This a point where we really need the *local* rather than just the generalized local alternation hierarchy.

It would be nice to find a simplification of the proof of Proposition 3.56.

Let I be a finite set of attributes, $act \notin I$, and $\Gamma = \{0,1\}^I$, $\Omega = \{0,1\}^{I\cup\{act\}}$. Let $c, c', d, d' \in \Gamma$ and $C = c \ominus c'^{*,1}$ and $D = d \ominus d'^{*,1}$.

Define the alphabet projection $\varphi : \Gamma \to \Omega$, $a \mapsto a \bowtie (1)_{act}$ as in the proofs of Proposition 3.55 and Theorem 3.50. We lift φ to pictures and picture languages as usual. Let $\mathbf{0} \in \Omega$ be defined by $\mathsf{pr}_\mu(\mathbf{0}) = 0$ for all $\mu \in I \cup \{act\}$, again like in the above proofs.

For every $P \in \Gamma^{+,+}$ let

$$\begin{aligned}\alpha(P) = \ & \{\varphi(P)\} \\ & \cup \bigcup\{\varphi(P'C) \oplus \mathbf{0}^{*,*} \oplus \varphi(DP'') \mid P \in P'CDP''\} \\ & \cup \{\mathbf{0}^{*,*} \oplus \varphi(DP'C) \oplus \mathbf{0}^{*,*} \mid P \in DP'C\}.\end{aligned}$$

Proposition 3.57 *1. $\alpha(L_1 \cup L_2) = \alpha(L_1) \cup \alpha(L_2)$ for all picture languages $L_1, L_2 \subseteq \Gamma^{+,+}$, and the same is true for $\cap$ instead of $\cup$.*

2. $\alpha(\Gamma^+ \setminus L) = \alpha(\Gamma^+) \setminus \alpha(L)$ for every word language $L \subseteq \Gamma^+$.

3. $\alpha(\Gamma^{+,+} \setminus L) = \alpha(\Gamma^{+,+}) \setminus \alpha(L)$ for every picture language $L \subseteq \Gamma^{+,+}$.

4. $\alpha(L_1 \ldots L_n) = (\mathbf{0}^{*,*}\alpha(L_1)\mathbf{0}^{*,*} \ldots \mathbf{0}^{*,*}\alpha(L_n)\mathbf{0}^{*,*}) \cap \alpha(\Gamma^{*,*})$ for all picture languages $L_1, \ldots, L_n \subseteq \Gamma^{+,+}$.

Lemma 3.58 *Let $L \subseteq \Gamma^+$ be a word language over Γ. If L is locally threshold testable, then so is $\alpha(L)$. In particular, $\alpha(top^{-1}(L))$ is in Π_0^{loc} for every locally threshold testable word language L.*

Proof We show:

1. $\alpha(\Gamma^+)$ is locally threshold testable.
2. If is of the form $L = \{w \in \Gamma^+ \mid x$ has t occurrences of the subblock $x\}$ for some $x \in \Gamma^+$ and some $t \geq 1$, then $\alpha(L)$ is locally threshold testable.
3. $\alpha(x\Gamma^*)$, $\alpha(\Gamma^* x)$ are locally threshold testable for every $x \in \Gamma^+$.

Ad 1: $\alpha(\Gamma^+) = \varphi(\Gamma^+) \cup \varphi(\Gamma^* c)\mathbf{0}^*\varphi(d\Gamma^*) \cup \mathbf{0}^*\varphi(d\Gamma^* c)\mathbf{0}^*$ is easily seen to be locally threshold testable because $\mathbf{0}$ is a letter that does not appear in φ-images.

Ad 2: Let $x \in \Gamma^+$, $t \geq 1$, and L as above. Then $\alpha(L)$ is the set of all words $w \in \Omega^+$ for which

- $w \in \alpha(\Gamma^+)$ and w has t different occurrences of $\varphi(x)$, or
- there are x', x'' with $x = x'cdx''$ and $w \in \varphi(\Gamma^* x'c)\mathbf{0}^+\varphi(dx''\Gamma^*)$ and w has $t-1$ different occurrences of $\varphi(x)$, or
- there is x' with $dx'c$ and $w \in \mathbf{0}^*\varphi(dx'c)\mathbf{0}^*$ and w has $t-1$ different occurrences of $\varphi(x)$.

It follows that $\alpha(L)$ is locally threshold testable.

Ad 3: The proofs are similar to that of Item 2 but easier.

The first claim of Lemma 3.58 follows from Items 1–3 and Proposition 3.57, Item 2 by induction.

The second claim follows from the first one, $\alpha(top^{-1}(L)) = \alpha(\Gamma^{+,+}) \cap top^{-1}(\alpha(L))$ for every word language L, and the fact that $\alpha(\Gamma^{+,+})$ is in Π_0^{loc}. □

Lemma 3.59 *Let $L \subseteq \Gamma^{+,+}$. If L is local or cyclically local, then $\alpha(L)$ is in Π_0^{loc}.*

Proof We only consider the case that L is local, the other case in analogous. Choose $\Delta \subseteq (\Gamma \cup \{\#\})^{2,2}$ such that $\mathcal{L}(\Delta) = L$. We extend φ to $\Gamma \cup \{\#\}$ by letting $\varphi(\#) = \#$. Let Δ' contain firstly $\varphi(\Delta)$, secondly all $2{\times}2$-subblocks of

$$\begin{array}{cccc} \# & \# & \# & \# \\ \varphi(c) & \mathbf{0} & \mathbf{0} & \varphi(d) \\ \varphi(c') & \mathbf{0} & \mathbf{0} & \varphi(d') \\ \varphi(c') & \mathbf{0} & \mathbf{0} & \varphi(d') \\ \# & \# & \# & \# \end{array} \quad \text{or} \quad \begin{array}{ccc} \# & \# & \# \\ \# & \mathbf{0} & \# \\ \# & \mathbf{0} & \# \\ \# & \# & \# \end{array}$$

and thirdly, in case $\begin{array}{cc} c & d \\ \# & \# \end{array} \in \Delta$, all $2{\times}2$-subblocks of $\begin{array}{ccc} \varphi(c) & \mathbf{0} & \varphi(d) \\ \# & \# & \# \end{array}$.
Then $\alpha(L)$ is the set of all pictures in $\mathcal{L}(\Delta')$ that are in $\varphi(\Gamma^{+,+})\mathbf{0}^{*,*}\varphi(\Gamma^{+,+})$ or in $\mathbf{0}^{*,*}\varphi(\Gamma^{+,+})\mathbf{0}^{*,*}$, but not in $\mathbf{0}^{+,+}$, i.e.,

$$\alpha(L) = \mathcal{L}(\Delta') \cap top^{-1}\Big(\big(\varphi(\Gamma)^+\mathbf{0}^*\varphi(\Gamma)^+ \cup \mathbf{0}^*\varphi(\Gamma)^+\mathbf{0}^*\big) \setminus \mathbf{0}^+\Big).$$

Thus $\alpha(L)$ is in Π_0^{loc}. □

Lemma 3.60 *For every $k \geq 0$, if $L \in \Pi_k^{\mathrm{loc}}(\Gamma)$, then $\alpha(L) \in \Pi_k^{\mathrm{loc}}(\Omega)$.*

Proof We show by simultaneous induction on the definition on the local hierarchy that

1. if $L \in \Sigma_k^{\mathrm{loc}}$, then $\alpha(L) \in \Sigma_k^{\mathrm{loc}}$,
2. if $L \in B(\Sigma_k^{\mathrm{loc}})$, then $\alpha(L) \in B(\Sigma_k^{\mathrm{loc}})$,
3. if $L \in \obar\text{-cl}(B(\Sigma_k^{\mathrm{loc}}))$, then $\alpha(L) = L' \cap \alpha(\Gamma^{+,+})$ for some $L' \in \obar\text{-cl}(B(\Sigma_k^{\mathrm{loc}}))$.

Claims 1 and 2 of the induction basis $k = 0$ follow from Lemmas 3.58 and 3.59 using Proposition 3.57 (Items 1 and 3).

Claim 3 (both of the induction basis and the induction step) follows from Proposition 3.57, Item 4. Claims 1 and 2 of the induction step follow from Proposition 3.57, Items 1 and 3. This completes the induction.

Proposition 3.60 follows from Item 1 by Lemma 3.57, Item 3. □

Now we have collected all lemmas to present the proof of the result of this Subsection.

Proof of Proposition 3.56 Choose I from the beginning of this subsection to be I_k. Let C and D be the set of columns occurring as the last column or the first column, respectively, of pictures in Num_k, i.e., $C = c \ominus (c')^{*,1}$, where $c, c' \in \{0,1\}^{I_k}$ such that $\mathsf{pr}_\mu(c) = 1$ for all $\mu \in I_k$ and $\mathsf{pr}_\mu(c') = 1$ for both $\mu \in \{num\text{-}1, end\text{-}1\}$ and $\mathsf{pr}_\mu(c') = 0$ for all $\mu \in I_k \setminus \{num\text{-}1, end\text{-}1\}$; and $D = d^{+,1}$, where $d \in \{0,1\}^{I_k}$ such that $\mathsf{pr}_\mu(d) = 0$ for all $\mu \in I_k$.

Proposition 3.56 follows from Lemma 3.60 and Theorem 3.23 and the observation that

$$cycl((Num_k \otimes (1^{+,+})_{act}) \obar \mathbf{0}^{*,*}) = \alpha(cycl(Num_k)).$$

This completes the proof and thus fills the gap in the proof of Theorem 3.50. □

3.5.4 Assembling the Formulas

In this subsection we shall prove the following main result of this section.

Theorem 3.61 *For every $k \geq 1$, there is a $\Pi_1^0(\Delta_1^{\mathrm{U}}(\Pi_{k-1}^0(\Delta_1^{\mathrm{U}})))$-sentence π and a $\Pi_2^0(\Delta_1^{\mathrm{U}}(\Pi_{k-1}^0(\Delta_1^{\mathrm{U}})))$-sentence π' that assert for the grid $\underline{[m]\times[n]}$ that $g_k(m) \mid n$ and that $g_k(m) = n$, respectively.*

Before we begin to develop the proof, we note that the formula classes of this theorem are fragments of two more important (and easier to understand) fragments, which allows to deduce Corollary 3.66.

Remark 3.62 For every $k \geq 1$,

- $\underline{\Delta_1^{\mathrm{U}}(\Pi_{k-1}^0(\Delta_1^{\mathrm{U}}))} \subseteq \underline{\Delta_k}$,
- $\underline{\Pi_1^0(\Delta_k)}, \underline{\Pi_2^0(\Delta_k)} \subseteq \underline{FO(\Delta_k)}$,
- $\underline{\Pi_1^0(\Delta_k)} \subseteq \underline{\Pi_k}$,
- $\underline{\Pi_2^0(\Delta_k)} \subseteq \underline{\Pi_1^0(\Sigma_k)}$.

□

Thus the sentence π from Theorem 3.61 is in Π_k and both the sentences π and π' are in the first-order closure of Δ_k.

Together with the non-definability results Corollary 4.15 and Theorem 4.25, the above implies (among others) the following separation results over the class of grids: $\underline{\Pi_k} \not\subseteq \underline{\Sigma_k}$ and $\underline{\Pi_2^0(\Delta_{k+1})} \not\subseteq \underline{FO(\Delta_k)}$ for all $k \geq 1$. The first non-inclusion solves an open problem of [MT97, Sch97a].

Again, we fix some $k \geq 1$.

Lemma 3.63 *There is a first-order formula $mleq(X, Y)$ that asserts for a grid $\underline{[m]\times[n]}$ and two sets $X, Y \subseteq [m]\times[n]$ that X is the unique set of top row positions $\{0\} \times \{1, \ldots, j\}$, where $j \leq n-1$ is maximal with $(\{0\} \times \{1, \ldots, j-1\}) \cap Y = \emptyset$.*

The proof is easy, cf. proof of Proposition 3.55, formula *less*.

Proposition 3.64 *For every $m, n \geq 1$ and every $j \in \{1, \ldots, f_k(m)\}$ there is exactly one picture of size (m, n) in $(\boldsymbol{twin}_{k,m,j})^* \obar \boldsymbol{ptwin}_{k,m,j}$.*

This is immediate by definition of $\boldsymbol{twin}_{k,m,j}$ and $\boldsymbol{ptwin}_{k,m,j}$.

Lemma 3.65 *Let us consider the attribute set $J = I_k \cup \iota(I_k) \cup \{\mathrm{act}, \iota(\mathrm{act}), \mathrm{leq}, \mathrm{offs}\}$. We write $\overline{X}$ to abbreviate the variable tuple $(X_\mu)_{\mu \in J}$.*

There is a formula $\varphi(\overline{X}, x)$ in $\Pi_{k-1}^0(\Delta_1^{\mathrm{U}})$ such that for all $m, n \geq 1$ and all $x \in [m]\times[n]$ there is exactly one picture P of size (m, n) over $\{0,1\}^J$ such that $P, x \models \varphi$.

Moreover, we have the following for this picture P and this position x: if $x = (i, j) \in \{0\} \times \{1, \ldots, f_k(m)\}$, then $\mathsf{restr}_{I_k \cup \iota(I_k) \cup \{act, \iota(act)\}}(P) \in (\boldsymbol{twin}_{k,m,j})^* \obar \boldsymbol{ptwin}_{k,m,j}$; and $x \notin \{0\} \times \{1, \ldots, f_k(m)\}$ iff $(0,0) \in X_{\mathrm{offs}}^P$.

Proof By Theorem 3.50, there exists a $\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)$-formula $\mathit{iso\text{-}xtwin}_k(\overline{X})$ that asserts for a picture P over $\{0,1\}^J$ that $\mathsf{restr}_{I_k\cup\iota(I_k)\cup\{act,\iota(act)\}}(P) \in \mathit{iso\text{-}xtwin}_k$.

The first-order formula $\psi(X_{\mathit{offs}},x) = \mathit{top}(x) \wedge X_{\mathit{offs}}(x) \wedge \forall x'(S_2xx' \to \neg X_{\mathit{offs}}x')$ asserts that x is the minimal top row position of $X_{\iota(\mathrm{act})}$, provided that $\mathit{mleq}(X_{\mathit{offs}}, X_{\iota(\mathrm{act})})$ holds.

Let $\varphi_1(\overline{X},x)$ be the following $\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)$-formula.

$$\varphi_1 = \mathit{iso\text{-}xtwin}_k \wedge \mathit{mleq}(X_{\mathit{offs}}, X_{\iota(\mathrm{act})}) \wedge \mathit{mleq}(X_{\mathit{leq}}, X_{\mathit{end\text{-}k}}) \wedge \psi.$$

For every $m,n \geq 1$ and every $x \in [m]\times[n]$, we have the following: $x \in \{0\}\times\{1,\ldots,f_k(m)\}$ iff there is a tuple $\overline{X} = (X_\mu)_{\mu\in J}$ of subsets of $[m]\times[n]$ such that the associated picture fulfills φ_1, and in this case $\overline{X}$ is determined uniquely by x (see Proposition 3.64) in such a way that the associated picture is in $\mathit{iso\text{-}xtwin}_k$.

We conclude from Theorems 3.43 and 3.11 that there is a formula $\mathit{pmarknum}_k(\overline{X})$ in $\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)$d that asserts for a non-empty picture P that $\mathsf{restr}_{I_k}(P) \in \mathit{pref}_+(\mathit{Num}^+_k)$. Let $\psi'(\overline{X})$ be a first-order formula that asserts that $X_\mu = \mathrm{dom}P$ for all $\mu \in \iota(I_k)\cup\{act,\iota(act),\mathit{offs}\}$.

Let $\pi(X_{\mathit{leq}},x) = \neg X_{\mathit{leq}}x \wedge \forall x'(S_2x'x \to \neg X_{\mathit{leq}}x')$. The first-order formula π asserts for a set $X_{\mathit{leq}} = \{0\}\times\{1,\ldots,j-1\}$ that $x \notin \{0\}\times\{1,\ldots,j\}$.

Let

$$\varphi_2(\overline{X},x) = \mathit{pmarknum}_k \wedge \mathit{mleq}(X_{\mathit{leq}}, X_{\mathit{end\text{-}k}}) \wedge \pi \wedge \psi'.$$

For every $m,n \geq 1$ and every $x \in [m]\times[n]$, we have the following: $x \notin \{0\}\times\{1,\ldots,f_k(m)\}$ iff there is a tuple $\overline{X} = (X_\mu)_{\mu\in J}$ of subsets of $[m]\times[n]$ such that the associated picture fulfills φ_2, and in this case $\overline{X}$ is determined uniquely (independently from x). The latter is because the conjunct $\mathit{pmarknum}_k$ determines the choice of X_μ for every attribute $\mu \in I_k$, the conjunct $\mathit{mleq}(X_{\mathit{leq}}, X_{\mathit{end\text{-}k}})$ determines the choice of X_{leq} (namely to be $\{0\}\times\{1,\ldots,\max\{n-1,f_k(m)-1\}\}$), and ψ' determines the choice of X_μ for the remaining attributes $\mu \in \iota(I_k)\cup\{act,\iota(act),\mathit{offs}\}$ (namely to $\mathrm{dom}P$).

Let $\varphi = \varphi_1 \vee \varphi_2$. Then φ has the desired property. □

In the following proof we use the mapping $\mathit{col} : \mathbb{N}^2 \to \mathbb{N}$, defined by $\mathit{col}(i,j) = i$. For the position x of some grid, $\mathit{col}(x)$ is the column number of x.

Proof of Theorem 3.61 Let $\psi(\overline{X},x,y) = \exists z(\mathit{top}(z) \wedge \mathit{left}(z) \wedge X_{\mathit{offs}}z) \vee X_{\iota(\mathit{end\text{-}k})}y$. Let $\varphi(\overline{X},x,y)$ as in the above proposition.[5] Then $\varrho(x,y) = \psi(\iota\overline{X}(\varphi(\overline{X},x,y)),x,y)$ is an allowed $\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1))$-formula that asserts the following for a picture P of size (m,n) and $y \in \{0\}\times[n]$:

$$x \notin \{0\}\times\{1,\ldots,f_k(m)\} \;\vee\; f_k(m) + \mathit{col}(x) \mid \mathit{col}(y) + 1.$$

The latter is true because in a picture of size (m,n) in $(\mathit{twin}_{k,m,j})^* \mathbin{⦶} \mathit{ptwin}_{k,m,j}$ (where $1 \leq j \leq f_k(m)$), the component corresponding to the attribute $\iota(\mathit{end\text{-}k})$ is 1 exactly in those positions y that are in the top row and fulfill $f_k(m) + j \mid \mathit{col}(y) + 1$.

Let $\sigma(y) = \forall x(\varrho(x,y))$. Then $\sigma(y)$ is a $\Pi^0_1(\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)))$-formula that asserts for a position y in the top row that

$$\forall j \in \{1,\ldots,f_k(m)\} : f_k(m) + j \mid \mathit{col}(y) + 1,$$

which is equivalent to

$$g_k(m) = \mathrm{lcm}\{f_k(m)+1,\ldots,2f_k(m)\} \mid \mathit{col}(y) + 1.$$

Choose

$$\begin{aligned} \pi &= \forall y(\mathit{top}(y) \wedge \mathit{right}(y) \to \sigma(y)), \\ \pi' &= \forall y(\mathit{top}(y) \to (\mathit{right}(y) \leftrightarrow \sigma(y))). \end{aligned}$$

Then π and π' assert for the grid $\underline{[m]\times[n]}$ that n is a (or the least, respectively) number with $g_k(m) \mid n$.

The sentence π is indeed in $\Pi^0_1(\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)))$, whereas the sentence π' is in $\Pi^0_1(B(\Pi^0_1(\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1))))) \subseteq \Pi^0_2(\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)))$, as required. □

The sentence π from Theorem 3.61 shows (together with Proposition 3.67, Item 4) the definability result of Theorem 2.32. The corresponding non-definability part is Corollary 4.15.

From Theorem 3.61 we conclude the following.

Corollary 3.66 *For every $k \geq 1$, the formula class $\Pi^0_2(\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)))$ is at least $(k{+}1)$-fold exponential, and thus so are $\Pi^0_2(\Delta_k)$, $\Pi^0_1(\Sigma_k)$, $FO(\Sigma_k)$.*

[5] Note that we add x and y in the list of potentially free variables of ψ and φ, respectively, for formal reasons.

Proof Take $\pi' \in \Pi^0_2(\Delta^{\mathrm{U}}_1(\Pi^0_{k-1}(\Delta^{\mathrm{U}}_1)))$ from Theorem 3.61. Membership in the other classes follows from Remark 3.62. $\square$

This corollary provides the definability result for Theorem 2.30. The corresponding non-definability result (i.e., the upper bound for the asymptotic growth rate) is stated in Theorem 4.25.

3.5.5 A Number-Theoretic Proposition

In this subsection we present the missing proof of the fact that for every $k \geq 1$, the function g_k from Theorem 3.61 is indeed $(k+1)$-fold exponential.

For every $m \geq 1$, let

$$\begin{aligned} F(m) &= \max\{\mathrm{lcm}\{x_1,\dots,x_k\} \mid k \geq 1, \sum_i x_i = m\}, \\ G(m) &= \mathrm{lcm}\{m+1, m+2, \dots, 2m\}, \\ H(m) &= 2^{\sqrt{m \log m}}. \end{aligned}$$

Recall that $s_0(m) = m$ and $s_{k+1}(m) = 2^{s_k(m)}$ for every $m \geq 1$, $k \geq 0$; and $g_k(m) = G(f_k(m))$ for every $k, m \geq 1$.

We have $f_k(m)$ is $s_k(\Theta(m))$. It is easy to show that $F(m) \leq 2^m$ for every $m \geq 1$, but this is a very rough bound. [Lan09] and [Sza80] (see also [Chr86]) show very precise approximations for F. We will only need the fact that $F(m)$ is $H(\Theta(m))$.

Proposition 3.67 *Let $k \geq 1$.*

1. *$G(m)$ is $s_1(\Omega(m))$.*
2. *$G(m)$ is $s_1(m^{\mathcal{O}(1)})$.*
3. *$F \circ f_k$ is $(k+1)$-fold exponential.*
4. *g_k is $(k+1)$-fold exponential.*

Proof For the first claim, let $m \geq 1$. Let $p_1, \dots, p_k$ be the distinct prime numbers $< m$, and $p = \prod_i p_i$, and $l_i = \max\{l \mid p_i^l < m\}$ for every $i \leq k$. Then $\mathrm{lcm}\{1,\dots,m-1\} = \prod_i p_i^{l_i} = \frac{1}{p}\prod_i p_i^{l_i+1} \geq \frac{m^k}{p}$. On the other hand, $\mathrm{lcm}\{1,\dots,m-1\} \geq \prod_i p_i = p$. Thus $(\mathrm{lcm}\{1,\dots,m-1\})^2 \geq p \cdot \frac{m^k}{p} = m^k$.

k is $\Omega(m/\log m)$, thus $G(m) = \mathrm{lcm}\{m+1,\dots,2m\} \geq \mathrm{lcm}\{1,\dots,m-1\} \geq m^{k/2}$ is $m^{\Omega(m/\log m)} = 2^{\Omega(m)}$.

For the third claim we need that $F(m)$ is $H(\Theta(m))$. Besides, $s_1(m^{1/2}) = 2^{\sqrt{m}} \leq H(m) \leq s_1(m)$ for all m, thus $H(m)$ and $F(m)$ are $s_1(m^{\Theta(1)})$. This implies the second and third claim: $G(m) = \mathrm{lcm}\{m+1,\dots,2m\} \leq F(\sum_{i=m+1}^{2m} i)$ is $F(\mathcal{O}(m^2)) = H(\mathcal{O}(m^2)) \leq s_1(m^{\mathcal{O}(1)})$; and $F(f_k(m))$ is $F(s_k(\Theta(m))) = s_1((s_k(\Theta(m)))^{\Theta(1)}) = s_{k+1}(\Theta(m))$.

The first and second claim imply that $G(m)$ is $s_1(m^{\Theta(1)})$. Thus $g_k(m) = G(f_k(m))$ is $G(s_k(\Theta(m))) = s_{k+1}(\Theta(m))$, i.e. the fourth claim. $\square$

The proof of the first claim is by Thomasz Schoen (private communication).

Chapter 4

Non-Definability Results

In the previous chapter we have developed definability results, i.e., we have described particular picture languages like Num_k that are definable in certain classes of formulas, e.g. level k of the monadic alternation hierarchy.

Our aim was to develop separation results, so it remains to present the corresponding non-definability results. This will be done in this chapter. For example, in Section 4.3 we will show that Num_k is not definable in level $k-1$ of the monadic hierarchy. This part of a proof is often referred to as "lower bound" proof, but we avoid this phrase because in our case, the non-definability result (at least for the class of non-coloured grids) establishes an *upper bound* on the growth rates of functions that are definable in, say, Σ_k.

The reader who is interested only in the most important facts may skip Sections 4.4 and 4.5, where we extend our techniques to show non-definability in the boolean closure and in the first-order closure, respectively, of e.g. some level of the monadic hierarchy.

As a digression, in Subsection 4.2.1 we illustrate the automata-based non-expressibility proofs for the existential fragment (i.e., level 1 of the alternation hierarchy) by some examples.

In Sections 4.1 and 4.2 we shall give some more definitions relevant for this chapter.

4.1 The Automaton Method

The main idea for our non-definability proofs is to invoke standard pumping techniques for finite automata on words. We will use an automata-theoretic characterization of $\underline{\Sigma_1}$ (from [GRST96]) over pictures to pass from a formula φ of monadic second-order logic over pictures to a family $(\mathfrak{A}_m)_{m\geq 1}$ of nondeterministic finite automata (NFAs). The NFA $\mathfrak{A}_m$ recognizes the set of those words of columns of height m that satisfy (when considered as pictures of height m) the formula φ. The important point is that if φ is in Σ_k, then $(\mathfrak{A}_m)_m$ can be chosen such that the number of states in $\mathfrak{A}_m$ is k-fold exponential in m. This follows from the above characterization and the observation that, in the world of automata, existential set quantification and complement correspond to projection and complementation, respectively.

Finite Automata—Some More Notation

We use standard notations for nondeterministic finite automata on words. An NFA over alphabet Γ is a quintuple $(Q, \Gamma, q_0, \Delta, F)$, where Q is a finite set of states, $q_0 \in Q$ is an initial state, $F \subseteq Q$ is a set of final states, and $\Delta \subseteq Q \times \Gamma \times Q$ is a transition relation. If $\mathfrak{A}$ is such an NFA, q, p are states of $\mathfrak{A}$, and w is a word over Γ, we write $\mathfrak{A} : q \xrightarrow{w} p$ iff there is a path from q to p labeled w. If $P \subseteq Q$ we write $\mathfrak{A} : q \xrightarrow{w} P$ iff there is a $p \in P$ with $\mathfrak{A} : q \xrightarrow{w} p$. The set of (non-empty) words *recognized by* $\mathfrak{A}$ is $\{w \in \Gamma^+ \mid q_0 \xrightarrow{w} F\}$. (The reference to non-empty words has technical reasons.)

$\mathfrak{A}$ is called *deterministic* (*DFA* for short) iff for every $(q, a) \in Q \times \Gamma$ there is at most one $p \in Q$ with $(q, a, p) \in \Delta$.

If $R \subseteq \mathbb{N} \times \mathbb{N}$ and $m \geq 1$, we write $R(m)$ for $\{n \mid (m, n) \in R\}$.

From Pictures to Words

For our non-definability results for picture languages we will use standard combinatorial methods for NFAs on words. Here we introduce notations that perform the link.

Definition 4.1 The *column word* of a non-empty picture, say of size (m, n), over alphabet Γ is a word over $\Gamma^{m,1}$, namely

$$\textit{column-word}\begin{pmatrix} a_{00} & \cdots & a_{0,n-1} \\ \vdots & & \vdots \\ a_{m-1,0} & \cdots & a_{m-1,n-1} \end{pmatrix} = \begin{pmatrix} a_{00} \\ \vdots \\ a_{m-1,0} \end{pmatrix} \cdots \begin{pmatrix} a_{0,n-1} \\ \vdots \\ a_{m-1,n-1} \end{pmatrix}.$$

For a picture language $L \subseteq \Gamma^{+,+}$ and an integer $m \geq 1$, the *height-m fragment of L*, denoted by $L[m]$, is the set of all column words of pictures of height m in L.

Let $t \geq 1$. The column word of a t-bit grid R is the column word of the picture over $\{0,1\}^t$ associated to R. For a class M of t-bit grids, the *height-m fragment of M* (denoted $M[m]$) is the height-m fragment of the set of pictures over $\{0,1\}^t$ associated to grids from M. □

In Section 4.5.2, Definition 4.37, we will slightly generalize this definition.

The notions of locality on words and on pictures are related in the following sense.

Proposition 4.2 *If L is a local picture language over alphabet Γ and $m \geq 1$, then $L[m]$ is a local word language over $\Gamma^{m,1}$.*

Proof Let Δ be a local tiling system with $\mathcal{L}(\Delta) = L$. Then a non-empty word w over $\Gamma^{m,1}$ is in $L[m]$ iff

all 2×2-subblocks of $\begin{array}{cc}\# & \#\\ \vdots & \overline{b}\\ \# & \#\end{array}$ and $\begin{array}{cc}\# & \#\\ \overline{a} & \overline{b}\\ \# & \#\end{array}$ and $\begin{array}{cc}\# & \#\\ \overline{a} & \vdots\\ \# & \#\end{array}$, respectively, are in Δ

for the infix $\#\overline{b}$ of $\#w$ (for all infixes $\overline{a}\,\overline{b}$ of w, or for the infix $\overline{a}\#$ of $w\#$, respectively), where $\overline{a}, \overline{b} \in \Gamma^{m,1}$.

This means that $L[m]$ is a local word language. □

4.2 Recognizable Picture Languages

A class of picture languages that is quite robust is the class of *recognizable* picture languages. Recall its definition from Section 2.3.1:

Definition 4.3 A picture language is *recognizable* if it is the image of a local picture language under some alphabet projection. □

The reason why the class of recognizable picture languages is of interest for us is the following theorem from [GRST96].

Theorem 4.4 *([GRST96]) Let L be a picture language. The following are equivalent:*

1. *L is recognizable.*

2. *L is Σ_1-definable.*

The direction 1 to 2 is easy and not needed here. The converse direction is the core argument for all of our non-definability results. It is proven in two steps: firstly, the authors show (using Ehrenfeucht-Fraïssé games) that every first-order definable picture language is locally threshold testable (see Definition 2.12). Secondly, it is shown that every locally threshold testable picture language is the projection of some local picture language. Then Theorem 4.4 follows.

One may view the representation of a Σ_1-definable picture language by a local tiling system plus an alphabet projection as some kind of "normal form" for Σ_1-formulas, where the kernel formula is not allowed to be an arbitrary first-order formula but may only assert the non-existence of particular local neighbourhoods. A normal form is often helpful in non-definability proofs. Here, we use tiling systems because they allow easy combinatorial arguments. See Theorem 4.5 below.

We remark that there are several other characterizations of recognizable picture languages, see e.g. [GR96]. Most of them, for example the one by "on-line tessellation automata" or the one by projections of "domino-local" picture languages, are quite simple consequences of the definition of recognizability. The latter notion will be explained in Subsection 6.2.2 and used in Chapter 6.

Furthermore, let me remark that in [GRST96, GR96] the authors use a slightly different signature for formulas over pictures than we introduced in Section 2.1. We use that same signature in Chapter 6 and compare these two in Subsection 6.2.1.

Theorem 4.5 *Let L be a recognizable picture language. There is a $c \geq 1$ such that for every $m \geq 1$ there is an NFA with c^m states that recognizes $L[m]$.*

Proof Let $t \geq 0$. Let Ω be the alphabet of L. There exists a local tiling system Δ over some alphabet Γ and a projection $\alpha\colon \Gamma \to \Omega$ such that $\alpha(\mathcal{L}(\Delta)) = L$.

Let $c := |\Gamma| + 1$. For every $m \geq 1$, the word language $\mathcal{L}(\Delta)[m]$ over alphabet $\Gamma^{m,1}$ is (by Proposition 4.2) local and therefore recognized by some deterministic finite automaton (DFA) with $|\Gamma^{m,1}| + 1 \leq c^m$ states. Since $\alpha(\mathcal{L}(\Delta)[m]) = L[m]$ for every m, the well-known construction for projections of word languages gives, for every m, an NFA with $|\Gamma^{m,1}| + 1 \leq c^m$ states that recognizes $L[m]$. $\square$

Since the shortest word accepted by an NFA cannot be longer than its number of states, the following fact from [Gia94] is a consequence of the above Theorem. Recall from Definition 2.26 when to call a relation on natural numbers Σ_1-definable.

Corollary 4.6 *If R is a Σ_1-definable relation, then there is a $c \geq 1$ such that for all $m \geq 1$ with $R(m) \neq \emptyset$ we have $\min R(m) \leq c^m$. In particular, every Σ_1-definable function is at most singly exponential.*

In the next section we will generalize the above argument, i.e., Theorem 4.5 and Corollary 4.6, from recognizable (i.e., from Σ_1-definable) picture languages to Σ_k-definable picture languages. In the next but one section we will then generalize the second statement of Corollary 4.6 (but not the first) to the boolean closure of Σ_1.

From Theorem 3.61 we may conclude the following.

Corollary 4.7 $\underline{\Sigma_1} \neq \underline{\Pi_1}$ *over grids.*

Proof By Theorem 3.61 and Remark 3.62, the relation $\{(m,n) \mid g_1(m)$ divides $n\}$ is Π_1-definable. Since $\min R(m) = g_1(m)$ is 2-fold exponential by Proposition 3.67, R is not Σ_1-definable. $\square$

In the remainder of this section we will present some facts about recognizable picture languages that are not needed for the subsequent sections. The reader who is interested only in the important facts may skip this subsection.

4.2.1 Digression: Examples for Non-Recognizable Picture Languages

In this subsection, which comprises the essential contributions of [Mat98b], we shall give some examples of non-recognizable picture languages. This solves a few open problems of [GR96, Mat95], which are explained at the end of this subsection. This subsection is not required for the proof of our main theorems, and in fact these main theorems strengthen the results of this subsection. However, it might be interesting from a methodical point of view.

The following simple fact was stated for example in [Bir96, GS96].

Lemma 4.8 *Let $n \geq 1$, $L \subseteq \Gamma^+$ be recognizable by an NFA with n states. Let $M \subseteq \Gamma^* \times \Gamma^*$ such that*

$$\begin{aligned} \forall (u,v) \in M \quad &: \quad uv \in L, \\ \forall (u,v),(u',v') \in M \quad &: \quad (u,v) \neq (u',v') \to \{uv', u'v\} \not\subseteq L. \end{aligned}$$

Then $|M| \leq n$.

This lemma allows us to prove lower bounds on the number of states that an NFA needs to recognize a particular word language.

Theorem 4.5 and Lemma 4.8 immediately imply the following result. Intuitively, it says that for a recognizable picture language, there is no more than exponentially much space to pass information from one side of the picture to the other. This way, it is a tool to show non-recognizability of a picture language.

Lemma 4.9 *Let $L \subseteq \Gamma^{+,+}$ be a recognizable picture language. Let $(M_m)_{m\geq 1}$ be a sequence such that $M_m \subseteq \Gamma^{m,+} \times \Gamma^{m,+}$ for all m, and*

$$\begin{aligned} \forall (P,Q) \in M_m \quad &: \quad PQ \in L, \\ \forall (P,Q),(P',Q') \in M_m \quad &: \quad (P,Q) \neq (P',Q') \to \{PQ', P'Q\} \not\subseteq L. \end{aligned}$$

Then $|M_m|$ is $2^{\mathcal{O}(m)}$.

Proof By Theorem 4.5 there is a c such that for every $m \geq 1$ there is an NFA with c^m states that recognizes $L[m]$. By Lemma 4.8, $|M_m| = |\{(\mathit{column\text{-}word}(P), \mathit{column\text{-}word}(Q)) \mid (P,Q) \in M_m\}| \leq c^m$. $\square$

An application of the above lemma is the following.

Proposition 4.10 *([GRST96]) Let L be the set of pictures over $\{0,1\}$ of the form PP where P is a square. Then L is not recognizable.*

Proof For every $m \geq 1$ let $M_m := \{(P,P) \mid P \in \Gamma^{m,m}\}$. We have for all squares P, P' that $PP' \in L \iff P = P'$, so $(M_m)_{m\geq 1}$ has the property of Lemma 4.9. But $|M_m| = 2^{m^2}$ is not $2^{\mathcal{O}(m)}$, therefore L is not recognizable. $\square$

In [GRST96] the non-recognizability of the above language was shown using essentially the same argument. The complement of L is recognizable, so a corollary is that the class of recognizable picture languages over alphabet $\{0,1\}$ is not closed under complement. (Our Theorem 4.7 states that this is not the case even for a trivial alphabet either.)

We give another, slightly more complex example of an application of Lemma 4.9.

Proposition 4.11 *Let Corners be as in Proposition 2.18, i.e., the set of pictures P over $\{0,1\}$ such that whenever $P(i,j) = P(i',j) = P(i,j') = 1$ then also $P(i',j') = 1$. Then Corners is not recognizable.*

Proof Let $m \geq 1$. For every permutation η on $[m]$ let P_η be the picture of size $(2m, m)$ with

$$P_\eta\langle i,j\rangle = \begin{cases} 1 & \text{if } i = j \vee i = \eta(j) + m, \\ 0 & \text{else.} \end{cases}$$

Let M_m be the set of all pairs (P_η, P_η) where η is a permutation of $[m]$. Since $P_\eta P_{\eta'} \in \mathit{Corners} \iff \eta = \eta'$ for all permutations η, η', the family $(M_m)_{m\geq 1}$ has the property of Lemma 4.9. Since $|M_m| = m!$ is not $2^{\mathcal{O}(m)}$ it follows that *Corners* is not recognizable. $\square$

Propositions 2.18 and 4.11 imply the following consequence:

Corollary 4.12 *There is a starfree, non-recognizable picture language over a two-letter alphabet.*

This corrects a mistake in [Mat95] and answers a couple of open questions of [GR96, Mat95], which we report now.

In [GR96], the authors consider different variants of "regularity" for picture languages. In their notation, the class RE of regular picture languages is the smallest class that contains all finite languages and is closed under the two concatenations, the two iterated concatenations (i.e., the two corresponding versions of the "Kleene star"), union, and complement. The class CFRE of "complementation-free" regular languages is defined similarly, but with closure under intersection instead of complement. The class of starfree picture languages (*SF* in our notation), which is closed under the two concatenations, union, and complement, is denoted SFRE there. The class of recognizable picture languages is denoted REC there. All these classes refer to some fixed alphabet.

It is shown in [GR92] that every "complementation-free" regular language is recognizable, namely by showing that the class of recognizable picture languages is closed under the two concatenations, the two iterated concatenations, union, and intersection.

The above corollary answers the following two open questions in [GR96] negatively.

- Is every regular picture recognizable?
- Is every regular picture language complementation-free?

Our main result Corollary 2.25 strengthens the negative answer to the first question essentially. It states that for every k, there is a starfree picture language (even in level $k+1$ of the generalized local alternation hierarchy) that is not in level k of the monadic alternation hierarchy (and thus in particular not in its first level, which corresponds to the class of recognizable picture languages).

Note that the above facts are not true for a singleton alphabet. In that case $\mathit{SF} \subsetneq \text{CFRE} = \text{RE} \subsetneq \text{REC}$.

4.3 Non-Definability in Σ_k

In this section we will continue the "main stream" of this thesis. We shall extend the technique of the last section to formulate a technique that allows us to show non-definability of a picture language in Σ_k for $k \geq 1$. Then we will apply this technique to the picture languages from Chapter 3, in particular the picture language Num_{k+1} from Section 3.3. Together with the definability results of that section we will obtain some important separation results, among them $\underline{\Pi_k} \not\subseteq \underline{\Sigma_k}$ for pictures.

Recall the definition of $\mathit{Mod}_t(\varphi)$ from page 31.

Theorem 4.13 *Let $k \geq 1$, $t \geq 0$, and $\varphi(X_1, \ldots, X_t)$ be a Σ_k-formula over τ_{Grids}. Let $\Gamma = \{0,1\}^t$. There exists $c \geq 1$ such that for every $m \geq 1$ there is an NFA with $s_{k-1}(c^m)$ states that recognizes the word language $\mathit{Mod}_t(\varphi)[m]$ over alphabet $\Gamma^{m,1}$.*

Proof The proof is by induction on k. The induction base $k = 1$ is given by Theorem 4.5 together with Theorem 4.4.

For the induction step, assume we have the claimed implication for some k. Let φ be a Σ_{k+1}-formula with free variables among $X_1, \dots, X_t$. W.l.o.g. we may assume that there is some $r \geq 0$ and a Σ_k-formula ψ with free variables among $X_1, \dots, X_{t+r}$ such that $\varphi = \exists X_{t+1} \dots \exists X_{t+r} \neg\psi$, i.e., $Mod_t(\varphi)$ is a projection of the complement of the picture language $Mod_{t+r}(\psi)$. So for every $m \geq 1$, the word language $Mod_t(\varphi)[m]$ is a projection of the complement of the word language $Mod_{t+r}(\psi)[m]$.

By induction hypothesis there is a $c \geq 0$ such that for every $m \geq 1$ there is an NFA with $s_{k-1}(c^m)$ states that recognizes the word language $Mod_{t+r}(\psi)[m]$. Hence (by the well-known constructions for complementation and projection for NFA) the word language $Mod_t(\varphi)[m]$ is recognized by an NFA with $2^{s_{k-1}(c^m)} = s_k(c^m)$ states.

This completes the induction step. □

The above theorem is from [MT97]. In [Sch97a], this result is improved in three ways. Firstly, it works also if the signature is enlarged by the vertical partial ordering $\leq_1$ (but not by the horizontal one). Secondly, there is an explicit bound on the constant c, and thirdly this bound depends only on the first-order quantifier-depth (the depth of nested first-order quantifiers) of the first-order kernel of φ rather than on the whole kernel.

Corollary 4.14 *Let $k \geq 1$, $t \geq 0$. Let L be a Σ_k-definable picture language over alphabet $\{0,1\}^t$. The partial function $f : \mathbb{N} \dashrightarrow \mathbb{N}$, $m \mapsto \min(\{n \mid \exists P \in L : \mathrm{size}P = (m,n)\}$ is at most k-fold exponential.*

Proof Let $n \geq 1$ and $\varphi(X_1, \dots, X_t)$ be a Σ_n-formula that defines L. By Theorem 4.13 there is a c such that for all $m \geq 1$ there is an NFA with $s_{n-1}(c^m)$ states that accepts $L[m]$. Since the shortest word accepted by an NFA (if it exists) cannot be longer than number of states of this NFA, this implies that $f(m) = \min\{|w| \mid w \in L[m]\} \leq s_{n-1}(c^m) = s_n(m \log_2 c)$ for all $m \in \mathrm{dom} f$. □

The following two corollaries are analogous to Corollary 4.6. The first one is inferred from the above in the case[1] $t = 0$ and provides for a non-definability result that allows us to separate $\underline{\Sigma_k}$ from $\underline{\Pi_k}$.

Corollary 4.15 *Let R be Σ_k-definable relation, $k \geq 1$. Then the partial function $\mathbb{N} \dashrightarrow \mathbb{N}$, $m \mapsto \min(R(m))$ is at most k-fold exponential.*

[1] We adopt the view that $\{0,1\}^0$ is some singleton.

Theorem 3.61, Remark 3.62, and the above imply Theorem 2.32. In particular, one may show the separation $\underline{\Sigma_k} \neq \underline{\Pi_k}$ for the class of grids and every $k \geq 1$ by considering the relation $\{\overline{(m,n) \mid g_k(m) \text{ divides } n}\}$.

The following is an immediate consequence of Corollary 4.15.

Corollary 4.16 *For every $k \geq 1$, every Σ_k-definable function is at most k-fold exponential.*

Together with Corollaries 3.41 and 3.49, this implies the following.

Corollary 4.17 *Let $k \geq 1$. The following formula classes are k-fold exponential: $\Delta_1^{\mathrm{U}}(\Pi_{k-1}^0(\Delta_1^{\mathrm{U}}))$, $\Delta_1^{\mathrm{U}}(\Pi_{k-1})$, $\Sigma_1(\Pi_{k-1}^0(\Delta_1^{\mathrm{U}}))$, Σ_k, and Δ_k.*

This implies the separation $\underline{\Sigma_k} \subsetneq \underline{\Sigma_{k+1}}$. The definability results of Section 3.4 imply that even $\underline{\Sigma_k} \subsetneq \underline{\Delta_{k+1}}$. In the next section we will show how to prove the yet stronger fact $\underline{B(\Sigma_k)} \subsetneq \underline{\Delta_{k+1}}$.

4.4 Arguing Against the Boolean Closure

In this section we wish to extend the technique we used to show non-definability in Σ_k as to work also for the boolean closure of Σ_k. How can this be done? Let us have a look at Corollary 4.15. It says that if R is a Σ_k-definable relation on $\mathbb{N}$, then the partial function $m \mapsto \min R(m)$ is at most k-fold exponential. Can we conclude that this is true also for $B(\Sigma_k)$-definable relations? No, we cannot, because Theorem 3.61 and Remark 3.62 show that there *is* a $(k+1)$-fold exponential function g_k such that the relation $R = \{(m,n) \mid g_k(m) \text{ divides } n\}$ is definable in Π_k.

But what about Corollary 4.16? It states that every Σ_k-definable function is at most k-fold exponential. We will show that *this* statement is true also for $B(\Sigma_k)$ instead of Σ_k. However, how can this be shown? We cannot conclude this fact from Corollary 4.16 because there might be a bunch of relations on $\mathbb{N}$ whose boolean combination is a function, and our knowledge about relations is not large enough to conclude that this function is at most k-fold exponential. Instead we will use Theorem 4.13 in order to investigate periodicity properties of Σ_k-definable relations more precisely. A deep theorem of Chrobak ([Chr86]) will help to exhibit a periodicity property of Σ_k-definable relations that is closed under boolean combinations and still strong enough to prove that every $B(\Sigma_k)$-definable function is at most k-fold exponential.

4.4.1 Automata and Periodicity

We need some preparations on word languages over a singleton alphabet. For a one-letter alphabet a word is identified with its length and hence a word language is identified with a subset of $\mathbb{N} \cup \{0\}$. Consequently, if $L \subseteq \mathbb{N} \times \mathbb{N}$ and $m \geq 1$, we let $L[m] = \{n \in \mathbb{N} \mid (m,n) \in L\}$.

Definition 4.18 Let $n \geq 0$. A set $N \subseteq \mathbb{N} \cup \{0\}$ is *periodic at* n (short: n*-periodic*) iff there is an integer $p \geq 1$ such that $\forall l \geq n(l \in N \leftrightarrow l+p \in N)$. □

Note that every boolean combination of n-periodic sets of integers is also n-periodic. (As period p take the least common multiple of the periods of the given sets.) It is well-known that the regular languages over a one-letter alphabet are exactly those sets of integers that are n-periodic for some n.

The following Theorem of Chrobak ([Chr86]) states that any NFA with n states over a singleton alphabet can be transformed into an equivalent DFA whose transition structure is a loop together with a path of length at most $(n+2)^2$ from the initial state to this loop.

Theorem 4.19 *(see [Chr86, Lemma 4.3, Theorem 4.4]) Let $N \subseteq \mathbb{N}$ be recognizable by some n-state NFA. Then there are some $k \leq (n+2)^2$ and an integer p such that N is recognized by a DFA $\mathfrak{B}$ with states $0,\ldots,(k-1)+p$ such that $\mathfrak{B}$ reaches the state $k + ((l-k) \mod p)$ after reading an input of length $l \geq k$.*

In [Chr86], Chrobak states only that the above k can be chosen in $\mathcal{O}(n^2)$, which would also suffice for our purposes, but his proof shows the present stronger claim. Besides, he also gives an asymptotic bound for p, but we do not need that here. In Subsection 4.5.1 we present a proof of Theorem 4.19 that is very similar to [Chr86].

Corollary 4.20 *Let $N \subseteq \mathbb{N}$ be recognizable by some n-state NFA. Then N is $(n+2)^2$-periodic.*

Proof Let $\mathfrak{B}$, k, p be as in the above theorem. For every $l \geq k$, the automaton reaches the same state on inputs of length either l or $l+p$. Thus $l \in N \iff l+p \in N$ for every $l \geq k$. □

Now we will apply Theorem 4.13 to sentences and hence to singleton alphabets.

Theorem 4.21 *Let $k \geq 1$ and φ be a $B(\Sigma_k)$-sentence. There is a $c \geq 1$ such that for every $m \geq 1$ the set $Mod_0(\varphi)[m]$ is $s_{k-1}(c^m)$-periodic.*

Proof Assume that φ is a boolean combination of the Σ_k-sentences $\varphi_1,\ldots,\varphi_r$. By Theorem 4.13, there are $c_1,\ldots,c_r$ such that for every $m \geq 1$ and every $i \leq r$, the set $Mod_0(\varphi_i)[m] = \{n \mid \underline{[m]\times[n]} \models \varphi_i\}$ is recognized by some NFA with $s_{k-1}(c_i^m)$ states. Let $c = \max\{(c_1+2)^2,\ldots,(c_r+2)^2\}$.

Since $(s_{k-1}(c_i^m)+2)^2 \leq s_{k-1}((c_i+2)^{2m}) \leq s_{k-1}(c^m)$ for every $m \geq 1$ and every $i \leq r$, Corollary 4.20 implies that $Mod_0(\varphi_i)[m]$ is $s_{k-1}(c^m)$-periodic for every i. For every $m \geq 1$, the set $Mod_0(\varphi)[m]$ is a boolean combination of the sets $Mod_0(\varphi_1)[m],\ldots,Mod_0(\varphi_r)[m]$ and hence also $s_{k-1}(c^m)$-periodic. □

Now we are ready to prove the announced upper bound for the asymptotic growth rate of $B(\Sigma_k)$-definable functions (rather than that of Σ_k-definable ones as in Corollary 4.16).

Corollary 4.22 *Every $B(\Sigma_k)$-definable function is at most k-fold exponential.*

Proof Let f be definable by a $B(\Sigma_k)$-sentence. By the above theorem, there is a c such that for all m, the singleton $\{f(m)\}$ is $s_{k-1}(c^m)$-periodic. This implies that $f(m) \leq s_{k-1}(c^m) = s_k(m \log_2 c)$ for all m. □

Note that Corollary 4.22 generalizes Corollary 4.16, but not Corollary 4.15. Corollaries 4.22 and 3.41 imply the following.

Corollary 4.23 *For every k, the formula class $B(\Sigma_k)$ is k-fold exponential.*

Now we can conclude some stronger separation results than at the end of Section 4.3, namely we have that the classes listed in Corollaries 3.49 and 3.66 allow to define a set of grids that is not $B(\Sigma_{k-1})$-definable. In particular we have the following corollary.

Corollary 4.24 $\underline{\Delta_{k+1}}, \underline{\Pi^0_1(\Sigma_k)} \not\subseteq \underline{B(\Sigma_k)}$ *for every $k \geq 1$.*

Now we have proved the most important separation results for grids. The reader might wish to skip the next section and continue with Chapter 5, where these results are transferred to other structures.

4.5 Arguing Against the First-Order Closure

The results of this section can be found as preliminary communications in [Mat98a]. We wish to prove facts similar to Corollaries 4.16 and 4.22. The following is the main result of this section.

Theorem 4.25 *For all $k \geq 1$, every $FO(\Sigma_k)$-definable function is at most $(k+1)$-fold exponential.*

This result might be interpreted as follows: *With respect to the rate of growth of definable functions*, the first-order closure is as powerful as "one more" set quantifier alternation and as powerful as the universal first-order closure because we have the same asymptotically tight bound $s_{k+1}(\mathcal{O}(m))$ for functions definable in Σ_{k+1}, $FO(\Sigma_k)$, and $\Pi^0_1(\Sigma_k)$.

The concept to show this bound is similar to that of the previous section. Again, the key idea is to exhibit a particular periodicity property of Σ_k-definable picture languages and to show that it is not destroyed by the closure under consideration (boolean closure in the previous section and first-order closure in this one). However, things are more complicated here because is does not suffice to consider periodicity properties of relations on $\mathbb{N}$ (or, equivalently, properties of non-coloured grids), but we need to consider properties of classes of "pebbled" grids. These are typical models of grid formulas with free first-order variables but without free set variables. This concept is now explained in an example.

Suppose we have a $FO(\Sigma_k)$-formula $\exists x_1 \forall x_2 \varphi(x_1, x_2)$, where φ is in Σ_k. Then we may associate to φ the set $Mod(\varphi)$ of "pebbled grids", i.e., grids with two differently marked positions. If m is, again, a fixed height, we may pass to the "height-m-fragment" $Mod(\varphi)[m]$ as before. But this time, the word language $Mod(\varphi)[m]$ will not contain words over a *singleton* alphabet, but words in which all but one or two positions (corresponding to the columns of two marked grid positions) carry the same letter. Again, $Mod(\varphi)[m]$ is recognizable by an NFA whose state set size is at most k-fold exponential in m. Since in input words for this NFA all but two positions carry the same letter, there are similar periodicity properties as for unary NFA. We will exhibit what happens to these periodicity properties when taking the first-order quantifications into account.

Theorem 4.25 was shown for the special case $k = 1$ in [AFS97] using Ehrenfeucht-Fraïssé games that are specifically tailored for $FO(\Sigma_1)$. Our method to argue against the first-order closure results from the one in

[AFS97] by extracting the periodicity technique from the description of Duplicator's winning strategy (thus "de-gamifying" this proof). This is combined with the automata-theoretic argument Theorem 4.13 against definability in Σ_k.

Some More Notation

The following definitions will be needed later. Let M, N be sets, $R \subseteq M \times N$ a relation, and $L \subseteq N$. Then

$$\langle R \rangle L = \{m \in M \mid \exists n \in L : (m, n) \in R\}.$$

Pebbled Grids. Let $t, r \geq 0$. An *r-pebbled t-bit grid* is a structure in the signature $\tau_{Grids_t} \cup \{x_1, \ldots, x_r\} = \{S_1, S_2, X_1, \ldots, X_t, x_1, \ldots, x_r\}$, where $X_1, \ldots, X_t$ are unary predicate symbols and $x_1, \ldots, x_r$ are constant symbols, such that the restriction to $\tau_{Grids} = \{S_1, S_2\}$ is a grid. The set of all r-pebbled t-bit grids is denoted $Grids_{t,r}$.

An r-pebbled t-bit grid is a *model* of a formula $\varphi(X_1, \ldots, X_t, x_1, \ldots, x_r)$ if it makes φ true with the implicitly given assignment. The class of r-pebbled t-bit grids that are models of φ is denoted $Mod_{t,r}(\varphi)$.

4.5.1 Automata and Periodicity wrt. a Certain Letter

In this subsection we will extend the definitions and results of Subsection 4.4.1.

Let Γ be an alphabet and $a \in \Gamma$ be some fixed letter. Let $n \geq 0$ and $p \geq 1$. A word language L over alphabet Γ is *periodic (with threshold n, period p, wrt. a)* iff for every $u, v \in \Gamma^*$: $ua^n v \in L \iff ua^{n+p} v \in L$. By *$(p, n)$-periodic* we mean "periodic with threshold n and period p". The explicit mentioning of a will be dropped if it is clear from the context. A set $N \subseteq \mathbb{N} \cup \{0\}$ will be called (p, n)-periodic iff $\{a^n \mid n \in N\}$ is (p, n)-periodic, i.e., if $n' \in N \leftrightarrow n' + p \in N$ for every $n' \geq n$.

This notion of periodicity is a refinement of that used in the previous section in the sense that it firstly is relative to a certain letter and it secondly takes the length p of the period into account.

The following definition will prove to be useful.

Definition 4.26 Let Γ, Ω be alphabets, $\alpha_0, \alpha_1 \subseteq \Gamma \times \Omega$. The relation $R_{(\alpha_0,\alpha_1)} \subseteq \Gamma^+ \times \Omega^+$ associated to (α_0, α_1) is given by

$$\{((a_1 \dots a_n), (a'_1 \dots a'_n)) \mid n \geq 0 \wedge \exists i((a_i, a'_i) \in \alpha_1 \wedge \forall j \neq i((a_j, a'_j) \in \alpha_0))\}.$$

□

In other words, if $(u, v) \in R_{(\alpha_0,\alpha_1)}$, then v results from u by replacing one letter according to α_1 and all the others according to α_0. In our applications, α_0 will be total, functional, and injective, so that this replacement amounts to marking a position and renaming of the other letters.

With this definition we can mimic the effect of first-order quantifications on the level of height-m fragments of grid languages. The reader who wants a better motivation for this definition *now* may read Subsections 4.5.2 and 4.5.3 up to Lemma 4.42 and then continue here.

Remark 4.27 Let $p \geq 1$, $n \geq 0$.

1. Every (p, n)-periodic word language is (p', n')-periodic for every multiple p' of p and every $n' \geq n$.
2. Every boolean combination of (p, n)-periodic word languages is (p, n)-periodic.
3. If $N \subseteq \mathbb{N}$ is (p, n)-periodic, and $n' \geq n$, then $\{k \mid k + n' \in N\}$ is $(p, 0)$-periodic.

□

We are interested in an asymptotic bound for the "threshold" n of the periodicity of fixed height fragments definable by nested first-order quantifications. (Again, we refer to Lemma 4.42 for a precise exposition of the relation to first-order quantification.) The next lemma shows why the length p of the period is essential for this.

Lemma 4.28 *Let Γ, Ω be alphabets, $\alpha_0, \alpha_1 \subseteq \Gamma \times \Omega$ and α_0 total, functional and injective, $(a, a') \in \alpha_0$. Let $R = R_{(\alpha_0,\alpha_1)}$. Let $L \subseteq \Omega^*$ be (p, n)-periodic wrt. a'. Then $\langle R \rangle L$ is $(p, 2n + p)$-periodic wrt. a.*

Proof We have to show that $ua^{2n+p}v \in \langle R \rangle L \iff ua^{2n+2p}v \in \langle R \rangle L$ for every $u, v \in \Sigma^*$.

So let $ua^{2n+p}v \in \langle R \rangle L$. There is a word in L that results from $ua^{2n+p}v$ by replacing one letter at a single position according to α_1 and all the others according to α_0. We will make a case distinction whether this position is in the "middle part" or not.

Formally: There are $u', x', v' \in \Omega^*$ such that $|u| = |u'|$, $|v| = |v'|$, and $|x'| = 2n + p$ such that $ua^{2n+p}v \; R \; u'x'v' \in L$.

In case $x' = a'^{2n+p}$ we have $ua^{2n+2p}v \; R \; u'a'^{2n+2p}v' \in L$ because L is (p, n)-periodic.

In case $x' = a'^i a_1 a'^j$ for some a_1 with $(a, a_1) \in \alpha_1$ and $i + j + 1 = 2n + p$ we have $i \geq n$ or $j \geq n$. Assume that $i \geq n$ (the other case is analogous). Then $ua^{2n+2p}v \; R \; u'a'^{i+p}a_1a'^j v' \in L$.

So in every case $ua^{2n+2p}v \in \langle R \rangle L$.

Conversely, let $ua^{2n+2p}v \in \langle R \rangle L$. There are $u', v', x' \in \Omega^*$ with $|u| = |u'|$, $|v| = |v'|$, $|x'| = 2n + 2p$, and $ua^{2n+2p}v \; R \; u'x'v' \in L$.

In case $x' = a'^{2n+2p}$ we have $ua^{2n+p}v \; R \; u'a'^{2n+p}v' \in L$ as before.

In case $x' = a'^i a_1 a'^j$ for some a_1 with $(a, a_1) \in \alpha_1$ and $i + j + 1 = 2n + 2p$ we have $i \geq n + p$ or $j \geq n + p$. Assume that $i \geq n + p$ (the other case is analogous). Then $ua^{2n+p}v \; R \; u'a'^{i-p}a_1a'^j v' \in L$.

So in every case $ua^{2n+p}v \in \langle R \rangle L$. This completes the proof. □

The essential point of the above lemma is that when passing from L to $\langle R \rangle L$, the period p remains the same whereas the threshold where this period starts increases by a summand p as well as by a constant factor. We will need this lemma in a situation where the period p is exponentially larger than the threshold n, so that a (single or boundedly often repeated) application of this lemma results in a singly exponential increase of that threshold.

The argument of the above proof is extracted from Duplicator's winning strategy in the proof of [AFS97, Lemma 12.6].

Definition 4.29 ([Chr86]) Let $\mathfrak{A}$ be a unary NFA, i.e. with singleton input alphabet. By the words *path*, *cycle*, *strongly connected component (SCC)* etc. we refer to the graph structure of $\mathfrak{A}$. Let p be the number of states on cycles and n be the number of states not on cycles. Then $S(\mathfrak{A}) := (p, n)$. □

Note that in the above definition, $p \geq 1$ and $n \geq 0$, provided $\mathfrak{A}$ recognizes infinitely many inputs.

Remark 4.30 Let $n \geq 0$ and $p \geq 1$. A word language $L \subseteq \{a\}^+$ is (p,n)-periodic iff there is a DFA $\mathfrak{A}$ with $S(\mathfrak{A}) = (p,n)$ that recognizes L. □

As explained in the introduction of this section, we have to investigate periodicity properties of NFAs whose input consists of words most of whose letters are identical. Our aim is to show that when an NFA reads such an input, then in the "gaps" between the designated positions it behaves very much like a unary NFA. This is stated in Lemma 4.32, which is prepared by the following definition.

Definition 4.31 Let $a \in \Gamma$. Let $\mathfrak{A} = (Q, \Gamma, q_0, \Delta, F)$ and $\mathfrak{A}' = (Q', \{a\}, q_0', \Delta', \emptyset)$ be NFAs. Then $\mathfrak{A}'$ *simulates* $\mathfrak{A}$ *wrt.* a iff there is a family $(F_{q,q'})_{q,q' \in Q}$ of subsets of Q' such that

$$\forall k \geq 0: \quad \mathfrak{A} : q \xrightarrow{a^k} q' \iff \mathfrak{A}' : q_0' \xrightarrow{a^k} F_{q,q'}. \tag{4.1}$$

□

Lemma 4.32 *Let $\mathfrak{A}$ be an NFA with n states and a be a letter from its input alphabet. Then there is $p \leq n$ and a unary NFA $\mathfrak{A}'$ with $S(\mathfrak{A}') = (p, n^2 + n)$ such that (1) $\mathfrak{A}'$ simulates $\mathfrak{A}$ wrt. a, and (2) no state of $\mathfrak{A}'$ on a cycle has more than one successor.*

After the proof this lemma, we will show how to derive a proof of Theorem 4.19. The proof of Lemma 4.32 is almost the same as for [Chr86, Lemma 4.3]. The essential observation is that the transition structure of the NFA constructed there does not depend on the initial and final state, so it can be carried out uniformly for all pairs of initial/final state. The original proof, too, gives these accurate bounds on the numbers of states (rather than the asymptotic ones stated in [Chr86, Lemma 4.3]).

The proof of Lemma 4.32 depends on the following number-theoretic fact which is stated in [Chr86] and follows from results of [Bra42, Liu64].

Lemma 4.33 *Let $a_1, \ldots, a_k \in [1, n]$. Let $X = \{\sum_{i=1}^k a_i x_i \mid x_1, \ldots, x_k \geq 0\}$. Then $X \cap \{x \mid x \geq n^2\}$ is an arithmetic progression with period $\gcd\{a_1, \ldots, a_k\}$. In particular, X is $(\gcd\{a_1, \ldots, a_k\}, n^2)$-periodic.*

Proof of Lemma 4.32 Since the transitions with labels different from a are of no concern, we may assume that $\mathfrak{A} = (Q, \{a\}, q_0, \Delta, F)$.

Let $T_1, \ldots, T_k$ be the SCCs of $\mathfrak{A}$. For every $i \leq k$ let y_i be the greatest common divisor of the length of all cycles of T_i. Let $m = n^2 + n$.

Choose $\mathfrak{A}' = (Q', \{a\}, 0, \Delta', \emptyset)$, where

$$\begin{aligned} Q' &= \{0, \ldots, m-1\} \cup \bigcup_{i \leq k} (\{i\} \times \{0, \ldots, y_i - 1\}), \\ \Delta' &= \{(i, a, i+1) \mid 0 \leq i < m-1\} \\ &\quad \cup \{((i,j), a, (i, (j+1) \mod y_i)) \mid i \leq k \wedge j < y_i\} \\ &\quad \cup \{(m-1, a, (i, 0)) \mid i \leq k\}. \end{aligned}$$

Then one has

$$\text{for every } x < m, \qquad \mathfrak{A}' : 0 \xrightarrow{a^x} q \text{ iff } q = x, \tag{4.2}$$

$$\text{for every } x \geq m, \qquad \mathfrak{A}' : 0 \xrightarrow{a^x} (i, s) \text{ iff } s = ((x - m) \mod y_i). \tag{4.3}$$

The number of $\mathfrak{A}'$-states on cycles is $p := \sum_i y_i \leq \sum_i |T_i| = |\bigcup_i T_i| \leq n$, and the number of states that are not on cycles is $m = n^2 + n$. It remains to be shown that $\mathfrak{A}'$ simulates $\mathfrak{A}$ wrt. a.

Choose $F_{q,q'} = \{x < m \mid \mathfrak{A} : q \xrightarrow{a^x} q'\} \cup \{(i, s) \in Q \mid \mathfrak{A} : q \xrightarrow{a^{m+s}} q', i \leq k, s < y_i\}$ for every $q, q' \in Q$. We have to show that $(F_{q,q'})_{q,q'}$ fulfills (4.1). We prepare this with some definitions.

For $q, q' \in Q$, a (q, q')-superpath is a sequence of the form

$$\alpha = P_1 D_1 P_2 D_2 \ldots P_t D_t P_{t+1},$$

where

- $D_1, \ldots, D_t$ are SCCs of $\mathfrak{A}$,
- $P_1, \ldots, P_{t+1}$ are paths in $\mathfrak{A}$ whose inner points do not belong to SCCs,
- the first vertex of P_1 is q, the last vertex of P_{t+1} is q',
- for every $i \in \{1, \ldots, t\}$, the last vertex of P_i belongs to D_i,
- for every $i \in \{2, \ldots, t+1\}$, the first vertex of P_i belongs to D_{i-1}.

For every superpath α in $\mathfrak{A}$, let $L_\alpha = \{x \mid x \text{ is the length of a path in } \alpha\}$.

For every SCC D of $\mathfrak{A}$ let $\Pi_{q,q'}(D)$ be the set of (q, q')-superpaths α such that D is the last SCC of α.

Let $i \in \{1, \ldots, k\}$. Let $\alpha \in \Pi_{q,q'}(T_i)$ and let $a_1, \ldots, a_l$ be the lengths of cycles (in the SCCs) of α. Then $L_\alpha = \{\sum_{j=1}^{l} a_j x_j \mid x_j \geq 0\}$ is $(\gcd\{a_1, \ldots, a_l\}, n^2)$-periodic (by Lemma 4.33) and hence (because y_i is a multiple of $\gcd\{a_1, \ldots, a_l\}$) also (y_i, n^2)-periodic. By Remark 4.27, $\{x \geq 0 \mid x + m \in L_\alpha\}$ is $(y_i, 0)$-periodic for every $i \leq k$ and $\alpha \in \Pi_{q,q'}(T_i)$.

Now we show (4.1). For every $x \in \{0, \ldots, m-1\}$ we have (by (4.2)) $\mathfrak{A}' : 0 \xrightarrow{a^x} F_{q,q'}$ iff $\mathfrak{A} : q \xrightarrow{a^x} q'$.

If $x \geq m$, then for every $i \leq k$ and every $\alpha \in \Pi_{q,q'}(T_i)$ one has $x \in L_\alpha$ iff $x - m \in \{x' \geq 0 \mid x' + m \in L_\alpha\}$ iff (because that set is $(y_i, 0)$-periodic) $((x-m) \mod y_i) + m \in L_\alpha$.

Thus for every $x \geq m$ and every $q, q' \in Q$ we have the following equivalence chain: $\mathfrak{A} : q \xrightarrow{a^x} q'$ iff $\exists i \leq k\, \exists \alpha \in \Pi_{q,q'}(T_i) : x \in L_\alpha$ iff (by the previous sentence) $\exists i \leq k\, \exists \alpha \in \Pi_{q,q'}(T_i) : ((x-m) \mod y_i) + m \in L_\alpha$ iff $\exists i \leq k : \mathfrak{A} : q \xrightarrow{a^{(x-m) \bmod y_i + m}} q'$ iff (by (4.3) and definition of $F_{q,q'}$) $\exists i, s : \mathfrak{A}' : 0 \xrightarrow{a^x} (i, s) \wedge (i, s) \in F_{q,q'}$ iff $\mathfrak{A}' : 0 \xrightarrow{a^x} F_{q,q'}$.

Thus $\mathfrak{A}'$ simulates $\mathfrak{A}$ wrt. a. This completes the proof. □

Here is the proof announced in Section 4.4.1.

Proof of Theorem 4.19 Let $\mathfrak{A} = (Q, \{a\}, q_0, \Delta, F)$ be an NFA with n states. Choose $(Q', \{a\}, 0, \Delta', \emptyset)$ according to Lemma 4.32, and $(F_{q,q'})_{q,q' \in Q}$ according to Definition 4.31. Then $\mathfrak{A}' = (Q', \{a\}, 0, \Delta', \bigcup_{f \in F} F_{q_0,f})$ is an NFA equivalent to $\mathfrak{A}$. Let p be the least common multiple of the lengths of cycles in $\mathfrak{A}'$. Let $k = n^2 + n$.

Choose $\mathfrak{B}$ with states $\{0, \ldots, (k-1) + p\}$ such that $\mathfrak{B}$ reaches state $k + ((l-k) \mod p)$ after reading an input of length $l \geq k$ and state l after reading an input of length $l < k$. The set of final states of $\mathfrak{B}$ is given by $\{l \leq k - 1 + p \mid l \in N\}$. It is straightforward to check that $\mathfrak{B}$ is equivalent to $\mathfrak{A}$. □

Recall the definition of F from Subsection 3.5.5, in particular that $F(m) \leq 2^m$ for all m.

Corollary 4.34 *Let $\mathfrak{A}$ be an NFA with n states and a be a letter from its input alphabet. Then there is $p \in \{1, \ldots, F(n)\}$ such that for all states q, q' of $\mathfrak{A}$,*

$$\mathfrak{A} : q \xrightarrow{a^{(n+1)^2}} q' \iff \mathfrak{A} : q \xrightarrow{a^{(n+1)^2+p}} q'.$$

Proof By Lemma 4.32 there is an NFA $\mathfrak{A}'$ and $p' \leq n$ with $S(\mathfrak{A}') = (p', n^2 + n)$ such that no state on a cycle of $\mathfrak{A}'$ has more than one successor and $\mathfrak{A}'$ simulates $\mathfrak{A}$ wrt. a. Choose a family $(F_{q,q'})_{q,q' \in Q}$ of sets of $\mathfrak{A}'$-states according to Definition 4.31, i.e.

$$\forall q, q' \in Q\, \forall k \geq 0 : \quad \mathfrak{A} : q \xrightarrow{a^k} q' \iff \mathfrak{A}' : q_0' \xrightarrow{a^k} F_{q,q'}.$$

Let $p \geq 1$ be the least common multiple of the length of cycles in $\mathfrak{A}'$. Since the cycles in $\mathfrak{A}'$ have pairwise disjoint state sets, $p \leq F(p') \leq F(n)$. Every path through $\mathfrak{A}'$ of length $\geq (n+1)^2 > n^2 + n$ passes through a state on a cycle, whose length divides p. This implies the claim. □

Now we will state the main result of this Subsection 4.5.1.

Theorem 4.35 *Let $a \in \Gamma$. For every word language $L \subseteq \Gamma^+$ that is recognized by an NFA with n states, there is a $p \in \{1 \ldots, F(n)\}$ such that L is $(p, (n+1)^2)$-periodic wrt. a.*

Proof Let $\mathfrak{A}$ be an NFA with n states that recognizes L. Choose $p \leq F(n)$ according to Corollary 4.34. Then $ua^{(n+1)^2}v \in L$ iff $\exists q_1, q_2 \in Q\, \exists q_3 \in F : q_0 \xrightarrow{u} q_1 \xrightarrow{a^{(n+1)^2}} q_2 \xrightarrow{v} q_3$ iff $\exists q_1, q_2 \in Q\, \exists q_3 \in F : q_0 \xrightarrow{u} q_1 \xrightarrow{a^{(n+1)^2+p}} q_2 \xrightarrow{v} q_3$ iff $ua^{(n+1)^2+p}v \in L$. □

The following corollary strengthens Corollary 4.20. It follows from Theorem 4.35 and Remark 4.30, and it will not be needed later.

Corollary 4.36 *Let $\mathfrak{A}$ be a unary NFA with n-states. Then there is a $p \leq F(n)$ and a DFA $\mathfrak{A}'$ with $S(\mathfrak{A}') = (p, (n+1)^2)$ that simulates $\mathfrak{A}$.*

4.5.2 From Words to Pebbled Grids

Like in Section 4.1, we will associate a "column word" to every (this time r-pebbled) coloured grid in order to exploit the periodicity arguments from the previous subsection.

The following definition is a slight generalization of Definition 4.1.

Definition 4.37 Let R be an r-pebbled t-bit grid of size (m, n). Then the *column word of R* is a word of length n over alphabet $(\{0,1\}^{t+r})^m$ defined as

$$\textit{column-word}(R) = \begin{pmatrix} \overline{a}_{11} \\ \vdots \\ \overline{a}_{m1} \end{pmatrix} \cdots \begin{pmatrix} \overline{a}_{1n} \\ \vdots \\ \overline{a}_{mn} \end{pmatrix},$$

where for every $(i,j) \in [m]\times[n]$, the word $\overline{a}_{ij} = a_{ij}^1 \dots a_{ij}^t a_{ij}^{t+1} \dots a_{ij}^{t+r}$ over $\{0,1\}$ is such that

- $a_{ij}^s = 1$ iff $(i,j) \in X_s^R$ for all $s \in \{1,\dots,t\}$,
- $a_{ij}^{t+s} = 1$ iff $(i,j) = x_s^R$ for all $s \in \{1,\dots r\}$.

For a set L of r-pebbled t-bit grids and $m \geq 1$, the *height-m fragment of L* is given by $L[m] = \{\textit{column-word}(R) \mid R \text{ is of height } m \text{ and } R \in L\}$. □

Proposition 4.38 *If $L \subseteq \textit{Grids}_{0,r}[m]$ is (p,n)-periodic, then so is its complement $\textit{Grids}_{0,r}[m] \setminus L$.*

Proof Immediate by the closure of (p,n)-periodic sets under boolean combinations (Remark 4.27) and the observation that $\textit{Grids}_{0,r}[m]$ is $(1,0)$-periodic. □

4.5.3 Periodicity Properties of $\boldsymbol{FO}(\Sigma_k)$-Definable Relations

The following corollary is similar to (and, as we will see, a consequence of) Theorem 4.13.

Corollary 4.39 *Let $k \geq 1$, $r \geq 0$, and $\varphi(x_1,\dots,x_r)$ be a Σ_k-formula over $\tau_{\textit{Grids}}$. There is a $c \geq 1$ such that for every $m \geq 1$ there is an $s_{k-1}(c^m)$-state NFA that recognizes the word language $\textit{Mod}_{0,r}(\varphi)[m]$.*

The proof of this corollary is immediate from Theorem 4.13 when we have the following lemma, which is easy to prove.

Lemma 4.40 *Let $k \geq 0$. For every Σ_k-formula $\varphi(x_1,\dots,x_r)$ there is a Σ_k-formula $\varphi'(X_1,\dots,X_r)$ with*

$$\textit{Mod}_{r,0}(\varphi') = \{(R,\{u_1\},\dots,\{u_r\}) \mid (R,u_1,\dots,u_r) \in \textit{Mod}_{0,r}(\varphi)\},$$

In particular, $\textit{Mod}_{r,0}(\varphi')[m] = \textit{Mod}_{0,r}(\varphi)[m]$ for every m.

The following definition mimics the effect of first-order quantifications on the level of formal languages.

Definition 4.41 Let $m \geq 1$, $r \geq 0$. Let $\Gamma = (\{0,1\}^r)^{m,1}$ and $\Omega = (\{0,1\}^{r+1})^{m,1}$. We define the following subsets of $\Gamma \times \Omega$:

$$\alpha_0 = \left\{ \left(\begin{pmatrix} \overline{a_1} \\ \vdots \\ \overline{a_m} \end{pmatrix}, \begin{pmatrix} \overline{a_1}0 \\ \vdots \\ \overline{a_m}0 \end{pmatrix} \right) \;\middle|\; \forall i : \overline{a_i} \in \{0,1\}^r \right\},$$

$$\alpha_1 = \left\{ \left(\begin{pmatrix} \overline{a_1} \\ \vdots \\ \overline{a_m} \end{pmatrix}, \begin{pmatrix} \overline{a_1}b_1 \\ \vdots \\ \overline{a_m}b_m \end{pmatrix} \right) \;\middle|\; b_1 \dots b_m \in 0^*10^*,\ \forall i : \overline{a_i} \in \{0,1\}^r \right\}.$$

The *position choice relation for height m and index r* is the relation $R_{(\alpha_0,\alpha_1)} \subseteq \Gamma^+ \times \Omega^+$ associated to (α_0,α_1) in the sense of Definition 4.26. □

Note that in the above definition, α_0 is total, functional, and injective.

The next lemma states that, roughly speaking, the application of the $\langle R \rangle$-modality for the position choice relation R for some fixed height does the same as an existential first-order quantification.

Lemma 4.42 *Let $m \geq 1$. Let $\Gamma = (\{0,1\}^r)^{m,1}$ and $\Omega = (\{0,1\}^{r+1})^{m,1}$. Let $\varphi(x_1,\dots,x_r)$ be a first-order formula. Let R be the position choice relation for height m and index r. Then*

$$\textit{Mod}_{0,r}(\exists x_{r+1}\varphi)[m] = \langle R \rangle(\textit{Mod}_{0,r+1}(\varphi)[m]).$$

Proof Firstly, let $w \in \textit{Mod}_{0,r}(\exists x_{r+1}\varphi)[m]$, say of length n. For every $(i,j) \in [m]\times[n]$ there is $\overline{a_{ij}} \in \{0,1\}^r$ such that

$$w = \begin{pmatrix} \overline{a_{00}} \\ \vdots \\ \overline{a_{m-1,0}} \end{pmatrix} \cdots \begin{pmatrix} \overline{a_{0,n-1}} \\ \vdots \\ \overline{a_{m-1,n-1}} \end{pmatrix}.$$

There exists a $(0,r)$-grid $A = (\underline{[m]\times[n]}, x_1,\dots,x_r)$ such that $A \models \exists x_{r+1}\varphi$ and w is the column word of A (in the sense of Definition 4.37). Thus there is $x_{r+1} = (\hat{\imath},\hat{\jmath}) \in [m]\times[n]$ with $A,(\hat{\imath},\hat{\jmath}) \models \varphi$. For every $(i,j) \in [m]\times[n]$ we define $\overline{a'_{ij}} \in \{0,1\}^{r+1}$ as follows.

$$\overline{a'_{ij}} = \begin{cases} \overline{a_{ij}}1 & \text{if } (i,j) = (\hat{\imath},\hat{\jmath}), \\ \overline{a_{ij}}0 & \text{else.} \end{cases}$$

Then $w = \begin{pmatrix}\overline{a_{00}}\\ \vdots \\ \overline{a_{m-1,0}}\end{pmatrix} \cdots \begin{pmatrix}\overline{a_{0,n-1}}\\ \vdots \\ \overline{a_{m-1,n-1}}\end{pmatrix} \; R \; \begin{pmatrix}\overline{a'_{00}}\\ \vdots \\ \overline{a'_{m-1,0}}\end{pmatrix} \cdots \begin{pmatrix}\overline{a'_{0,n-1}}\\ \vdots \\ \overline{a'_{m-1,n-1}}\end{pmatrix} \in$ $Mod_{0,r}(\varphi)[m]$, thus $w \in \langle R\rangle(Mod_{0,r+1}(\varphi)[m])$.

The converse direction is similar. □

The above lemma justifies our interest in Definitions 4.26, 4.41. The reason why I chose to present them and Lemma 4.28 earlier is that the latter is purely word-theoretic and can be understood without any knowledge of grids and logic.

Now we combine the above lemma and Lemma 4.28 to state what a single first-order quantification can do to the periodicity of a fixed height fragment of some set of grids.

Lemma 4.43 *Let $m \geq 1$, $r \geq 0$, and $\eth \in \{\exists, \forall\}$, and $\varphi(x_1, \ldots, x_{r+1})$ be a formula. If $Mod_{0,r+1}(\varphi)[m]$ is (p,n)-periodic, then $Mod_{0,r}(\eth x_{r+1}\varphi)[m]$ is $(p, 2n+p)$-periodic.*

Proof In case $\eth = \exists$ the claim is immediate from Lemmas 4.42 and 4.28.

In case $\eth = \forall$ we have that $Mod_{0,r}(\neg\varphi)[m]$ is (p,n)-periodic by Proposition 4.38. Thus $Mod_{0,r}(\exists x_{r+1}\neg\varphi)[m]$ is $(p, 2n+p)$-periodic, and thus (by Proposition 4.38) so is $Mod_{0,r}(\forall x_{r+1}\varphi)[m] = Mod_{0,r}(\neg\exists x_{r+1}\neg\varphi)[m]$. □

Since we are interested in a non-definability argument for the first-order closure of a given class of formulas, we have to apply the above Lemma in iteration. This will be done now.

Lemma 4.44 *Let $m \geq 1$. For every $r \geq 0$ and every $\eth_1, \ldots, \eth_r \in \{\exists, \forall\}$, if $\varphi(x_1, \ldots, x_r)$ is a formula such that $Mod_{0,r}(\varphi)[m]$ is (p,n)-periodic, then $Mod_{0,0}(\eth_1 x_1 \ldots \eth_r x_r \varphi)[m]$ is $(p, 2^r(n+p) - p)$-periodic.*

Proof The proof is by induction on r. The case $r = 0$ is trivial because $2^0(n+p) - p = n$. So assume the claimed implication is true for some r. Let $\eth_1, \ldots, \eth_{r+1} \in \{\exists, \forall\}$ and $\varphi(x_1, \ldots, x_{r+1})$ be a formula such that $Mod_{0,r+1}(\varphi)[m]$ is (p,n)-periodic. By Lemma 4.43, $Mod_{0,r}(\eth x_{r+1}\varphi)[m]$ is $(p, 2n+p)$-periodic. By induction hypothesis, $Mod_{0,0}(\eth_1 x_1 \ldots \eth_{r+1} x_{r+1}\varphi)[m]$ is $(p, 2^r(2n+p)+p) - p))$-periodic, i.e., $(p, 2^{r+1}(n+p) - p)$-periodic. This completes the induction. □

The above lemma is [Mat98a, Lemma 38].

Now we will apply the above in the situation where φ is a boolean combination of Σ_k-formulas. In the introduction to this section, I explained our interest in automata running on inputs almost all of whose letters are identical. I mentioned the idea that on such an input an NFA has similar periodicity properties as a unary NFA. However, the truth is a little more complicated, as the proof of the next lemma shows. The point is that we need the fact that if we have several NFAs, then we can find an appropriate periodicity property that is common for these automata by building their cross-product first.

Lemma 4.45 *Let $k \geq 1$ and $r \geq 0$ and $\varphi(x_1, \ldots, x_r)$ be a $B(\Sigma_k)$-formula, and $\eth_1, \ldots, \eth_r \in \{\exists, \forall\}$. Then there is some $d \geq 0$ and such that for every $m \geq 1$ there is some $p \in \{1, \ldots, s_{k+1}(dm)\}$ such that*

$$Mod_{0,0}(\eth_1 x_1 \ldots \eth_r x_r \varphi)[m]$$

is $(p, s_{k+1}(dm))$-periodic.

Proof Assume that $\varphi(x_1, \ldots, x_r)$ is a boolean combination of the Σ_k-formulas $\varphi_1, \ldots, \varphi_j$. By Corollary 4.39, for every $i \leq j$ there is some constant c_i such that for every $m \geq 1$ there is a $s_k(c_i m)$-state NFA that recognizes the word language $Mod_{0,r}(\varphi_i)[m]$.

Choose $c := \max\{c_1, \ldots, c_j\}$ and d such that $2^r\big((s_k(cjm)+1)^2 + s_{k+1}(cjm)\big) \leq s_{k+1}(dm)$ for every $m \geq 1$.

Let $m \geq 1$. By choice of c there is, for every i, an NFA $\mathfrak{A}_i = (Q_i, (\{0,1\}^r)^m, q_{0,i}, \Delta_i, F_i)$ with $s_k(cm)$ states that recognizes $Mod_{0,r}(\varphi_i)[m]$. Build the usual "cross-product automaton" $\mathfrak{A} = (Q_1 \times \ldots \times Q_j, (\{0,1\}^r)^m, (q_{01}, \ldots, q_{0j}), \Delta, \emptyset)$ with state set of size $n := (s_k(cm))^j \leq s_k(cjm)$. Choose p with $1 \leq p \leq F(n) \leq 2^n \leq s_{k+1}(cjm) \leq s_{k+1}(dm)$ according to Corollary 4.34, i.e., such that for every $q, q' \in Q$:

$$\mathfrak{A} : q \xrightarrow{a^{(n+1)^2}} q' \iff \mathfrak{A} : q \xrightarrow{a^{(n+1)^2+p}} q'.$$

Then $Mod_{0,r}(\varphi)[m]$ is $(p, (n+1)^2)$-periodic. By Lemma 4.44, the set $Mod_{0,0}(\eth_1 x_1 \ldots \eth_r x_r \varphi)[m]$ is $(p, (2^r((n+1)^2+p)) - p)$-periodic. The claim follows by choice of d. □

Now we can prove Theorem 4.25, which states that for every $k \geq 1$, the class $FO(\Sigma_k)$ is at most $(k+1)$-fold exponential in the sense of Definition 2.27.

Proof of Theorem 4.25 Let $f : \mathbb{N} \dashrightarrow \mathbb{N}$ be definable by a $FO(\Sigma_k)$-sentence ψ. Clearly, we may assume that ψ is of the form $\eth_1 x_1 \ldots \eth_r x_r \varphi$, where $r \geq 0$ and φ is a $B(\Sigma_k)$-formula.

For every $m \geq 1$, we have $\{a^n \mid f(m) = n\} = L_f[m] = Mod_{0,0}(\psi)[m]$, where a is the element of $(\{0\}^r)^m$.

Choose d as in the above lemma, i.e., such that for all $m \in \mathrm{dom} f$ there is a $p \geq 1$ such that the set $\{f(m)\}$ is $(p, s_{k+1}(dm))$-periodic. Then this implies that $f(m) \leq s_{k+1}(dm)$ for all $m \in \mathrm{dom} f$. $\square$

As a consequence of Theorem 4.25, all subclasses of $FO(\Sigma_k)$ are at most $(k+1)$-fold exponential. Together with Corollary 3.66 this implies the following corollary, which gives tight asymptotic growth rate bounds for some formula classes.

Corollary 4.46 *For every $k \geq 1$, the formula classes of Corollary 3.66 are $(k+1)$-fold exponential, among them $\Pi^0_j(\Sigma_k)$ (for $j \geq 1$) and $FO(\Sigma_k)$.*

Together with Corollary 4.23, this implies Theorem 2.29. In particular, we now have all separation claims from Corollary 2.31.

Chapter 5

From Grids to Graphs

In the previous two chapters we have concentrated on the class of t-bit grids and have shown separation results like $\underline{\Sigma_k} \underset{\neq}{\subseteq} \underline{\Sigma_{k+1}}$ or $\underline{FO^{\mathrm{TC}}} \not\subseteq \underline{\Sigma_k}$.

In this chapter we will show how to transfer such separation results from one class $\mathcal{C}$ of structures, say grids, to a different class $\mathcal{C}'$, possibly over a different signature. Basically, a $\mathcal{C}$-structure M has to be encoded in a $\mathcal{C}'$-structure in such a way that any formula talking M can be transformed into a formula that says the same about the encoding of M and, conversely, every formula talking about a $\mathcal{C}'$-structure M' can be transformed into a formula that says the same about the $\mathcal{C}$-structure whose encoding M' is, if such a $\mathcal{C}$-structure exists. Moreover, these translations of formulas have to respect the formula classes under consideration, i.e., a Σ_k-formula over the signature of $\mathcal{C}$ has to be translated into a Σ_k-formula over the signature of $\mathcal{C}'$.

In the first section we shall introduce the notion of *strong first-order reductions*, which is a particular kind of mapping from one class $\mathcal{C}$ of structures into another class $\mathcal{C}'$. If such a reduction exists, this implies that every $\mathcal{C}$-structure can be encoded in a $\mathcal{C}'$-structure the way described above.

Strong first-order reductions are a variant of "interpretations", see e.g. [Rab77, Hog93]. The name is chosen because they are even more similar to the "monadic first-order reductions" of [Cos93]. They first have been defined in [MST99], although such a technique has been used implicitly in [MT97]. (Probably similar encoding techniques have been used often without giving them a particular name.)

In [MST99] it is explained why the translation of formulas respects all formula classes under consideration in that paper, i.e., Σ_k, Π_k, $B(\Sigma_k)$,

$\Sigma_k(FO^{TC})$, $FO(\Sigma_k)$, $closed(\Sigma_k)$, Δ_k.

We will exhibit on an abstract level a class of formula classes for which this transfer works. Informally, these are those formula classes that result from the class of first-order formulas by repeated application and combination of the closure operations introduced in Section 2.2, namely the positive boolean closure, the existential first-order closure, the existential second-order closure, the Δ_1-closure, and the unary TC-closure. (And, additionally, the mapping $\mathcal{F} \mapsto co\text{-}\mathcal{F}$, which is not really a closure operation.)

This motivates the definition of "abstract formula class transformers (AFCT) given by closure properties". These are, essentially, the mappings obtained by repeated applications of the listed closures. For example, the mapping $\mathcal{F} \mapsto \Sigma_k(\mathcal{F})$ is an AFCT given by closure properties.

It turns out that if Δ_k (with $k \geq 1$) is involved, the situation is a bit more complicated. I order to enable the transfer of separation results for such formula classes, the existence of a strong first-order reduction suffices only if the reduction fulfills a particular additional requirement. This is because the definitions of formula classes that involve the Δ_k-closure are relative to a fixed class of structures (rather than just to its signature). That is why we distinguish between "syntactic" and "semantic" abstract formula class transformers (see Sections 5.2 and 5.3, respectively): semantic ones transform formula classes whose definition is relative to (the equivalence induced by) a given class of structures.

In Section 5.4 we will then give some examples of strong first-order reductions, in particular from the class of coloured grids into the class of directed graphs as well as from the class of directed graphs into the class of undirected graphs.

This will allow us to deduce Theorems 2.33 and 2.34, i.e., our separation results for formula classes over graphs, from Theorem 2.25, i.e., our separation results for grids.

A Word About Notation. Since we have concentrated on the class of (possibly coloured) grids in the previous chapters, there was no need to mention dependencies on the signature τ_{Grids}. Now we write FO_τ for the set of all first-order formulas in a given signature τ, making this dependence explicit.

5.1 Strong First-Order Reductions

We introduce the notion of "strong first-order reducibility" as a tool to transfer separation results from one class of structures to another.

A strong first-order reduction is a certain kind of function that codes a structure M over some signature τ into another structure $\Phi(M)$ over a possibly different signature τ' in such a way that a formula φ over signature τ can be translated into a formula φ' over τ' (and vice versa) such that φ and φ' are "of the same shape" (for instance in the same level of the monadic hierarchy) and φ' asserts for $\Phi(M)$ the same as φ asserts for M. The domain of $\Phi(M)$ will be of the form $\{1, \dots, n\} \times \mathrm{dom}(M)$.

Definition 5.1 Let $\mathcal{C}$, $\mathcal{C}'$ be classes of structures over the relational signatures τ and τ', respectively. Let $n \geq 1$.

A *strong first-order reduction from $\mathcal{C}$ to $\mathcal{C}'$ with rank n* is an injective mapping $\Phi : \mathcal{C} \to \mathcal{C}'$ such that:

1. For every structure $M \in \mathcal{C}$, the universe of $\Phi(M)$ is given by $[1, n] \times \mathrm{dom} M$, i.e., a disjoint union of n copies of the universe of M.

2. There is a first-order formula $\psi(x_1, \dots, x_n)$ over τ' such that for all structures M in $\mathcal{C}$, all $u_1, \dots, u_n \in \mathrm{dom} M$, and all $i_1, \dots, i_n \leq n$:
$$\Phi(M) \models \psi[(i_1, u_1), \dots, (i_n, u_n)] \quad \text{iff} \quad \forall j \leq n : i_j = j \wedge u_j = u_1.$$

3. For every relation symbol r' from τ', say of arity l, and every $\kappa : [1, l] \to [1, n]$ there is a first-order formula $\varphi_\kappa^{r'}(x_1, \dots, x_l)$ over τ such that for all structures M in $\mathcal{C}$ and all $u_1, \dots, u_l \in \mathrm{dom} M$ we have
$$M \models \varphi_\kappa^{r'}[u_1, \dots, u_l] \quad \text{iff} \quad \Phi(M) \models r'[(\kappa(1), u_1), \dots, (\kappa(l), u_l)].$$

4. For every relation symbol r from τ, say of arity l, there is a first-order formula $\varphi^r(x_1, \dots, x_l)$ over τ' such that for all structures M in $\mathcal{C}$ and all $u_1, \dots, u_l \in \mathrm{dom} M$ we have
$$M \models r[u_1, \dots, u_l] \quad \text{iff} \quad \Phi(M) \models \varphi^r[(1, u_1), \dots, (1, u_l)].$$

$\mathcal{C}$ is *strongly first-order reducible* to $\mathcal{C}'$ iff there is a strong first-order reduction from $\mathcal{C}$ to $\mathcal{C}'$. □

Examples of strong first-order reductions are given in Section 5.4.

The main results of this section are Corollaries 5.18 and 5.26. The former implies those separation claims of Theorem 2.33 that do not involve the Δ_1-closure, whereas the latter implies Theorem 2.33 as is.

It is instructive to have a look at an important special case of Corollary 5.18, which can be understood now.

Fact 5.2 *If $\underline{\Sigma_k}(\mathcal{C}) \subsetneq \underline{\Sigma_{k+1}}(\mathcal{C})$ for some k and a class $\mathcal{C}$ of structures that is strongly first-order reducible to $\mathcal{C}'$, then also $\underline{\Sigma_k}(\mathcal{C}') \subsetneq \underline{\Sigma_{k+1}}(\mathcal{C}')$.*

This is a consequence of the following special case of Theorem 5.16.

Fact 5.3 *Let $\mathcal{C}$, $\mathcal{C}'$ be classes of structures, let $\Phi : \mathcal{C} \to \mathcal{C}'$ be a strong first-order reduction, and let $L \subseteq \mathcal{C}$.*

Then L is Σ_k-definable within the class of $\mathcal{C}$-structures iff $\Phi(L)$ is Σ_k-definable within the class of $\Phi(\mathcal{C})$-structures.

The monadic first-order reductions of [Cos93] are defined in a more syntactic flavour. More importantly, the definition of "monadic first-order reductions" lacks Item 4 of Definition 5.1. As a consequence, under the weaker assumption that $\Phi(L)$ is monadic first-order reducible to L, one can only conclude the directions "$\Leftarrow$" of Fact 5.3. (This is formulated in [Cos93] for the class Σ_1, i.e. the existential fragment of monadic second-order logic.) This direction of the proof is prepared in Subsection 5.1.2 by the notion of back-realizability.

The other, easier direction "$\Rightarrow$" is prepared in Subsection 5.1.1 by the notion of forth-realizability.

5.1.1 Forth-Realizability

In this subsection, let τ and τ' be relational signatures and let $\mathcal{C}$ and $\mathcal{C}'$ be classes of τ-structures and τ'-structures, respectively, and $\Phi : \mathcal{C} \to \mathcal{C}'$ be a strong first-order reduction. We show how a first-order formula φ over τ can be translated into a first-order formula φ' over τ' such that φ asserts for a $\mathcal{C}$-structure M the same as φ' asserts for $\Phi(M)$.

For this purpose, we will replace the atomic subformulas of φ by formulas that exist because of Item 4 in Definition 5.1, and we will relativize first-order quantifications to the first of the n disjoint copies of the universe. We could also relativize the second-order quantifications to (subsets of) the first copy, but this is not necessary because the relativization of first-order quantifications implies that the value of second-order variables only matters on the first copy.

Although this proof is simple, we are going to carry it out in detail. For this purpose, the notion of forth-realizability wrt. a given strong first-order reduction Φ is introduced.

Definition 5.4 Let n be the rank of Φ. Let φ and φ' be formulas over τ and τ', respectively, with the same free variables, say $X_1, \dots, X_t, x_1, \dots, x_m$. The formula φ' *forth-realizes* φ *wrt.* Φ iff for all structures M in $\mathcal{C}$ and all $U_1, \dots, U_t \subseteq \mathrm{dom} M$, $u_1, \dots, u_m \in \mathrm{dom} M$ we have

$$\begin{array}{ll} & M \models \varphi[U_1, \dots, U_t, u_1, \dots, u_m] \\ \iff & \Phi(M) \models \varphi'[\{1\} \times U_1, \dots, \{1\} \times U_t, (1, u_1), \dots, (1, u_m)]. \end{array}$$

If $\mathcal{F}'$ is a class of τ'-formulas, we say that a formula φ over τ is *forth-realizable in $\mathcal{F}'$ wrt.* Φ iff there is a formula in $\mathcal{F}'$ that forth-realizes φ wrt. Φ. □

Sometimes we will omit the mentioning of Φ if it is clear from the context.

We will use the following lemma to show Proposition 5.6, i.e., that every first-order formula over τ is forth-realized wrt. Φ by some first-order formula over τ', provided that Φ is a strong first-order reduction from a class of τ-into a class of τ'-structures. Furthermore, this lemma will be helpful later on when particular closure operations on formula classes are considered.

Lemma 5.5 *Let $\mathcal{F}'$ be a class of τ'-formulas. Let φ and ψ be formulas over τ that are forth-realizable in $\mathcal{F}'$ wrt. Φ.*

1. *The formula $\neg\varphi$ is forth-realizable in co-$\mathcal{F}'$ wrt. Φ.*
2. *The formula $\varphi \vee \psi$ is forth-realizable in $\vee$-cl$(\mathcal{F}')$ wrt. Φ.*
3. *The formula $\varphi \wedge \psi$ is forth-realizable in $\wedge$-cl$(\mathcal{F}')$ wrt. Φ.*
4. *The formula $\exists x \varphi$ is forth-realizable in $\Sigma^0_1(\mathcal{F}')$ wrt. Φ.*
5. *The formula $\mathrm{TC}(y, z, \varphi)$ (where y, z are distinct) is forth-realizable in FOTC$(\mathcal{F}')$ wrt. Φ.*
6. *The formula $\exists X \varphi$ is forth-realizable in $\Sigma_1(\mathcal{F}')$ wrt. Φ.*

Proof Let φ' and ψ' be $\mathcal{F}'$-formulas that forth-realize φ and ψ, respectively. Then the formulas $\neg\varphi$, $\varphi \vee \psi$, and $\varphi \wedge \psi$ are forth-realized by $\neg\varphi'$, $\varphi' \vee \psi'$, and $\varphi' \wedge \psi'$, respectively.

Let $\sigma'(x) = \exists x_2 \ldots \exists x_n \psi(x, x_2, \ldots, x_n)$, where ψ is chosen according to Definition 5.1. For an element x of the universe of a structure $\Phi(M)$ in the image of Φ, the formula $\sigma'(x)$ asserts that $x \in \{1\} \times \mathrm{dom}M$.

The formula $\exists x\varphi$ is forth-realized by $\exists x(\sigma'(x) \wedge \varphi')$.

The formula $\mathrm{TC}(y, z, \varphi)$ is forth-realized by $\mathrm{TC}(y, z, \sigma'(y) \wedge \varphi')$.

The formula $\exists X_t\varphi$ is forth-realized by $\exists X_t\varphi'$. (Another possible choice would be $\exists X_t(\varphi' \wedge \forall x(X_t x \rightarrow \sigma'(x)))$.)

This completes the proof of Claims 1–6 because the above formulas are in the respective classes. $\square$

Proposition 5.6 *Every FO_τ-formula is forth-realizable in $FO_{\tau'}$ wrt. Φ.*

Proof We show that the atomic τ-formulas are forth-realizable in FO_τ.

The atomic τ-formulas Xx and $x = y$ are forth-realized by themselves. For every l-ary relation symbol r in τ, there exists an FO_τ-formula that forth-realizes $r(x_1, \ldots, x_l)$ by Item 4 from Definition 5.1.

Now the claim follows by induction on the complexity of FO_τ-formulas, using Items 1–4 of the preceding lemma. $\square$

It is not difficult to show that the above proposition remains true if the classes of first-order formulas over τ and τ' are replaced by the classes of, say, Σ_k-formulas over τ and τ', respectively. However, we will delay this simple induction and do it in Section 5.2 together with the converse translation of formulas, which the next subsection is concerned with.

5.1.2 Back-Realizability

Again, we will fix a strong first-order reduction Φ from the class $\mathcal{C}$ of τ-structures into the class $\mathcal{C}'$ of τ'-structures. Let $n \geq 1$ be the rank of Φ.

We will introduce back-realizability as the counterpart of forth-realizability. If φ' is a *sentence* over signature τ', then back-realizability in a class $\mathcal{F}$ of τ-formulas means that there is a sentence $\varphi \in \mathcal{F}$ that asserts the same for $\mathcal{C}$-structures as φ' asserts for Φ-images of $\mathcal{C}$-structures.

However, in order to perform induction on formulas we need to have the notion of back-realizability also for formulas with free variables. Since the domain of the Φ-image of a $\mathcal{C}$-structure M is n times larger than the domain of M, this makes this notion (and the corresponding proofs) substantially more difficult than in the preceding subsection.

Definition 5.7 Let $\varphi(X_1, \ldots, X_t, x_1, \ldots, x_m)$ be a formula over τ' and $\varphi_\kappa(\overline{X_1}, \ldots, \overline{X_t}, x_1, \ldots, x_m)$ a formula over τ for every $\kappa : [1, m] \rightarrow [1, n]$. (Here, the $\overline{X_i}$ are variable tuples of length n.)

The family $(\varphi_\kappa)_\kappa$ *back-realizes* φ *(wrt. Φ)* iff for all $\mathcal{C}$-structures M, all n-tuples $\overline{U_1}, \ldots, \overline{U_t}$ of subsets of $\mathrm{dom}M$, all $u_1, \ldots, u_m \in \mathrm{dom}M$, and all κ,

$$\begin{aligned} &M \models \varphi_\kappa[\overline{U_1}, \ldots, \overline{U_t}, u_1, \ldots, u_m] \iff \\ &\Phi(M) \models \varphi[\textstyle\bigcup_{i \leq n}(\{i\} \times U_{1i}), \ldots, \bigcup_{i \leq n}(\{i\} \times U_{ti}), (\kappa(1), u_1), \ldots, (\kappa(m), u_m)]. \end{aligned}$$

Let $\mathcal{F}$ be a class of formulas over τ. A formula $\varphi(X_1, \ldots, X_t, x_1, \ldots, x_m)$ over τ' is *back-realizable in $\mathcal{F}$ wrt. Φ* iff there is a family $(\varphi_\kappa(\overline{X_1}, \ldots, \overline{X_t}, x_1, \ldots, x_m))_\kappa$ of formulas in $\mathcal{F}$ that back-realizes φ wrt. Φ. $\square$

Lemma 5.8 *Let $\mathcal{F}$ be a class of τ-formulas. Let $\varphi(X_1, \ldots, X_t, x_1, \ldots, x_m)$ and $\psi(X_1, \ldots, X_t, x_1, \ldots, x_m)$ (with $t, m \geq 0$) be τ'-formulas that are back-realizable in $\mathcal{F}$ wrt. Φ.*

1. *The formula $\neg\varphi$ is back-realizable in co-$\mathcal{F}$ wrt. Φ.*
2. *The formula $\varphi \vee \psi$ is back-realizable in $\vee$-cl($\mathcal{F}$) wrt. Φ.*
3. *The formula $\varphi \wedge \psi$ is back-realizable in $\wedge$-cl($\mathcal{F}$) wrt. Φ.*
4. *The formula $\exists x_m \varphi$ is back-realizable in $\Sigma^0_1(\mathcal{F})$ wrt. Φ.*
5. *The formula $\mathrm{TC}(x_p, x_q, \varphi)$ (where $p, q \leq m$ are distinct) is back-realizable in FOTC($\mathcal{F}$) wrt. Φ.*
6. *The formula $\exists X_t \varphi$ is back-realizable in $\Sigma_1(\mathcal{F})$ wrt. Φ.*

Proof Let $(\varphi_\kappa)_{\kappa:[1,m]\rightarrow[1,n]}$ and $(\psi_\kappa)_{\kappa:[1,m]\rightarrow[1,n]}$ be families of $\mathcal{F}$-formulas (with free variables among $X_{11}, \ldots, X_{tn}, x_1, \ldots, x_m$) that back-realize φ and ψ, respectively, wrt. Φ.

Ad 1, 2, 3: Obviously, $(\neg\varphi_\kappa)_\kappa$ back-realizes $\neg\varphi$, and $(\varphi_\kappa \vee \psi_\kappa)_\kappa$ back-realizes $\varphi \vee \psi$, and $(\varphi_\kappa \wedge \psi_\kappa)_\kappa$ back-realizes $\varphi \wedge \psi$.

Ad 4: For all $\kappa : [1, m-1] \rightarrow [1, n]$ and all $i \leq n$ we define $\kappa(m \mapsto i)$ to be the extension of κ that maps m to i.

Then $\big(\exists x_m(\bigvee_{i\leq n}\varphi_{(\kappa(m\mapsto i))})\big)_{\kappa:[1,m-1]\to[1,n]}$ is a family of $\Sigma^0_1(\mathcal{F})$-formulas that back-realizes $\exists x_m\varphi$.

Ad 5: Let $\overline{X}$ abbreviate the tuple $(X_{11},\dots,X_{tn})$. W.l.o.g. assume $p=1$ and $q=2$. Let $\kappa:[1,m]\to[1,n]$. We define τ-formulas $\sigma^{\kappa,l}_{ij}$ for $i,j\in[1,n]$ and $l\in[0,n]$ inductively over l as follows (reminding of the standard construction for regular expressions from NFA):

$$\sigma^{\kappa,0}_{ij}(\overline{X},\overline{x}) = \begin{cases} \varphi_{\kappa(1\mapsto i)(2\mapsto j)}(\overline{X},\overline{x}) & \text{if } i\neq j,\\ \varphi_{\kappa(1\mapsto i)(2\mapsto j)}(\overline{X},\overline{x})\vee(x_1=x_2) & \text{else,}\end{cases}$$

$$\begin{aligned}\sigma^{\kappa,l}_{ij}(\overline{X},\overline{x}) = \sigma^{\kappa,l-1}_{ij}\vee\exists z,z'(\ & \sigma^{\kappa,l-1}_{il}(\overline{X},x_1,z,x_3,\dots,x_m)\\ &\wedge \mathrm{TC}(x_1,x_2,\sigma^{\kappa,l-1}_{ll})(\overline{X},z,z',x_3,\dots,x_m)\\ &\wedge \sigma^{\kappa,l-1}_{lj}(\overline{X},z',x_2,x_3,\dots,x_m)).\end{aligned}$$

Then $\sigma^{\kappa,l}_{ij}$ is a *FOTC*($\mathcal{F}$)-formula for every i,j,l.

Let $M\in\mathcal{C}$ and let $\overline{U_1},\dots,\overline{U_t}$ be n-tuples of subsets of dom$M$ and let us write U_j instead of $\bigcup_{i\leq n}(\{i\}\times U_{ji})$ for every $j\leq t$. Then one shows, inductively over l, the following property for every l,i,j:

$M\models\sigma^{\kappa,l}_{ij}[\overline{U_1},\dots,\overline{U_t},u_1,\dots,u_m]$ iff there is a sequence α in $[1,n]\times$ dom$M$ that starts in (i,u_1), ends in (j,u_2), all intermediate components are in $[1,l]\times$ dom$M$, and if $(\nu,u),(\nu',u')$ are successive components of α, then $\Phi(M)\models\varphi[U_1,\dots,U_t,(\nu,u),(\nu',u'),(\kappa(3),u_3),\dots,(\kappa(m),u_m)]$.

Let $\varrho_\kappa=\sigma^{\kappa,n}_{\kappa(1),\kappa(2)}$. Then $M\models\varrho_\kappa[\overline{U_1},\dots,\overline{U_t},u_1,\dots,u_m]$ iff $\Phi(M)\models$ $\mathrm{TC}(x_1,x_2,\varphi)[U_1,\dots,U_t,(\kappa(1),u_1),\dots,(\kappa(m),u_m)]$.

Thus the family $(\varrho_\kappa)_\kappa$ of *FOTC*($\mathcal{F}$)-formulas back-realizes $\mathrm{TC}(x_1,x_2,\varphi)$.

Ad 6: We claim that the family $\big(\exists X_{t1}\dots\exists X_{tn}(\varphi_\kappa)\big)_\kappa$ of $\Sigma_1(\mathcal{F})$-formulas back-realizes $\exists X_t\varphi$.

Let $\overline{U_1},\dots,\overline{U_{t-1}}$ be n-tuples of subsets of dom$M$ and $u_1,\dots,u_m\in$ dom$M$ and $\kappa:[1,m]\to[1,n]$. Then we have the following equivalence chain.

$M\models(\exists X_{t1}\dots\exists X_{tn}(\varphi_\kappa))[\overline{U_1},\dots,\overline{U_{t-1}},u_1,\dots,u_m]$ iff there exist sets $U_{t1},\dots,U_{tn}\subseteq$ dom$M$ such that $M\models\varphi_\kappa[\overline{U_1},\dots,\overline{U_t},u_1,\dots,u_m]$ iff (by assumption on φ) there exist $U_{t1},\dots,U_{tn}\subseteq$ dom$M$ such that $\Phi(M)\models$ $\varphi[\bigcup_{i\leq n}(\{i\}\times U_{1i}),\dots,\bigcup_{i\leq n}(\{i\}\times U_{ti}),(\kappa(1),u_1),\dots,(\kappa(m),u_m)]$ iff (since every set $U\subseteq$ dom$\Phi(M)$ can be written as $\bigcup_{i\leq n}(\{i\}\times U_i)$ for some subsets U_i of dom$M$) $\Phi(M)\models(\exists X_t\varphi)[\bigcup_{i\leq n}(\{i\}\times U_{1i}),\dots,\bigcup_{i\leq n}(\{i\}\times U_{t-1,i}),$ $(\kappa(1),u_1),\dots,(\kappa(m),u_m)]$. □

Proposition 5.9 *Every $FO_{\tau'}$-formula is back-realizable in FO_τ wrt. Φ.*

Proof Let n be the rank of Φ. We argue by induction on the complexity of an $FO_{\tau'}$-formula $\varphi(X_1,\dots,X_t,x_1,\dots,x_m)$. Let us consider the atomic formulas first.

The formula X_ix_j (where $i\leq t$, $j\leq m$) is back-realized by $(X_{i,\kappa(j)}x_j)_\kappa$ because for every M, all n-tuples $\overline{U_1},\dots,\overline{U_k}$ of subsets of dom$M$, every $u_1,\dots,u_m\in$ dom$M$, and every $\kappa:[1,m]\to[1,n]$ one has the following equivalence chain: $\Phi(M)\models(X_ix_j)[\bigcup_{l\leq n}(\{l\}\times U_{1l}),\dots,\bigcup_{l\leq n}(\{l\}\times U_{tl}),$ $(\kappa(1),u_1),\dots,(\kappa(m),u_m)]$ iff $(\kappa(j),u_j)\in\bigcup_l(\{l\}\times U_{il})$ iff $u_j\in U_{i,\kappa(j)}$ iff $M\models(X_{i,\kappa(j)}x_j)[\overline{U_1},\dots,\overline{U_t},u_1,\dots,u_m]$.

A formula $x_i{=}x_j$ (where $i,j\leq m$) is back-realized by $(\varphi_\kappa)_\kappa$ with

$$\varphi_\kappa=\begin{cases} x_i{=}x_j & \text{if } \kappa(i)=\kappa(j),\\ \neg(x_i{=}x_i) & \text{else.}\end{cases}$$

A family of formulas that back-realizes $r'(y_1,\dots,y_l)$ for some l-ary relation symbol r' of τ' and variables $y_1,\dots,y_l$ among $x_1,\dots,x_m$ is provided by Definition 5.1, Item 3.

Now let φ be a $FO_{\tau'}$-formula and assume as induction hypothesis that every less complex formula is back-realizable in FO_τ. If φ is atomic, we have just shown that φ is back-realizable in FO_τ, and if φ is of the form $\neg\varphi'$, $\varphi'\vee\varphi''$, or $\exists x\varphi'$, then the preceding lemma shows that φ is back-realizable in FO_τ because $co\text{-}FO_\tau,\vee\text{-}cl(FO_\tau),\Sigma^0_1(FO_\tau)=FO_\tau$. This completes the induction. □

5.2 Syntactic Abstract Formula Class Transformers

In the previous section we have shown, roughly speaking, how a first-order formula over τ can be translated into a first-order formula over τ' and vice versa, in such a way that these formulas assert the same for a structure in $\mathcal{C}$ and its image under a strong first-order reduction $\Phi:\mathcal{C}\to\mathcal{C}'$. In [MT97, MST99] it is argued that this translation can be lifted from first-order to Σ_k, Π_k, FO^{TC} and so on. In this thesis our aim is a detailed definition of a class of formula classes for which this translation works. As a consequence, these will be formula classes for which separation results can be transferred using our technique of strong first-order reductions.

The first observation is that, strictly speaking, the classes of Σ_k-formulas over different signatures differ. Thus when we want to lift results like "Σ_k is less expressive than Σ_{k+1}" from one class of structures to another class over a different signature, we have to exploit some relation between the distinct classes of Σ_k-formulas over these distinct signatures. Recall that we introduced $\Sigma_k(.)$ as a closure operation, i.e., a particular transformation, on formula classes. ("Σ_k" is only an abbreviation for $\Sigma_k(FO)$.) This transformation is independent of any particular signature.

This motivates to base our arguments on transformations of formula classes.

Definition 5.10 A *syntactic abstract formula class transformer (syntactic AFCT)* is a mapping $\mathfrak{A}$ that maps every set of formulas into a set of formulas, in such a way that for every signature τ, if $\mathcal{F}$ is a set of τ-formulas, then so is $\mathfrak{A}(\mathcal{F})$. $\square$

Example 5.11 By $\Sigma_1^0(.)$ we denote the syntactic AFCT given by $\mathcal{F} \mapsto \Sigma_1^0(\mathcal{F})$. Similarly for $\Sigma_1(.)$, $PB(.)$, *co-*., and $FOTC(.)$. The identity mapping *id* on the class of formula classes is also a syntactic AFCT. $\square$

The transformations $\Sigma_k(.)$, $closed(\Sigma_k)(.)$ and the like, which we are interested in, result from the syntactic AFCTs given in Remark 5.11 by application of particular operations, which we will now define formally.

Definition 5.12 If $\mathfrak{A}$ and $\mathfrak{B}$ are syntactic AFCT, then we define

- $\mathfrak{A} \circ \mathfrak{B}$ by $(\mathfrak{A} \circ \mathfrak{B})(\mathcal{F}) = \mathfrak{A}(\mathfrak{B}(\mathcal{F}))$ for every set $\mathcal{F}$ of formulas.
- $\mathfrak{A} \sqcup \mathfrak{B}$ to be the syntactic AFCT such that for every set $\mathcal{F}$ of formulas, $(\mathfrak{A} \sqcup \mathfrak{B})(\mathcal{F})$ is the smallest superclass $\mathcal{F}'$ of $\mathcal{F}$ with $\mathcal{F}' \supseteq \mathfrak{A}(\mathcal{F}') \cup \mathfrak{B}(\mathcal{F}')$.

$\square$

Typically, a syntactic AFCT closes a formula class under certain operations like boolean combinations or existential first-order quantifications. The idea is that if $\mathfrak{A}$ and $\mathfrak{B}$ are syntactic AFCTs, then $\mathfrak{A} \sqcup \mathfrak{B}$ applies the two closures at the same time.

The syntactic AFCT *co-*. is an exception in the sense that *co-*$\mathcal{F}$ is in general not a superset of $\mathcal{F}$.

Remark 5.13 $B(.) = PB(.) \sqcup co\text{-}.$ and $FO(.) = \Sigma_1^0(.) \sqcup B(.)$.

$\Pi_1(.) = co\text{-}. \circ \Sigma_1(.) \circ co\text{-}.$ and $\Pi_1^0(.) = co\text{-}. \circ \Sigma_1^0(.) \circ co\text{-}.$. The operation $\sqcup$ is associative. If we identify formula classes whose elements are equal modulo the equivalence induced by "φ is equivalent to $\neg\neg\varphi$", then we have the equation $closed(\Sigma_1)(.) = \Sigma_1(.) \sqcup \Sigma_1^0(.) \sqcup \Pi_1^0(.)$. $\square$

The idea how to obtain the syntactic AFCT we are interested in is, of course, to apply the above operations in iteration.

Definition 5.14 The class of syntactic *AFCT given by closure properties* (or *syntactic cpAFCT* for short) is defined as follows: *id*, $\Sigma_1^0(.)$, $\Sigma_1(.)$, $PB(.)$, *co-*., and $FOTC(.)$ are syntactic cpAFCT. If $\mathfrak{A}$, $\mathfrak{B}$ are syntactic cpAFCT, then so are $\mathfrak{A} \circ \mathfrak{B}$ and $\mathfrak{A} \sqcup \mathfrak{B}$. $\square$

Examples of syntactic cpAFCTs are $\Sigma_k(.)$, $\Sigma_k^0(.)$, $closed(\Sigma_k)(.)$, but also ones that do not have particular names, for instance the mapping $\mathcal{F} \mapsto FO(\Sigma_k(FOTC(\mathcal{F})))$ or the mapping $\mathcal{F} \mapsto FO(\Sigma_1(FO(\Sigma_1(\mathcal{F}))))$.

We wish to show that for those formula classes that result from the class of first-order formulas by application of a syntactic cpAFCT, separation results can be transferred by means of strong first-order reductions. This is formalized in the following definition and theorem, which is the main result of this section.

Definition 5.15 Let τ be a signature, $\mathcal{C}$ a class of τ-structures. For every syntactic AFCT $\mathfrak{A}$, we let

$$\underline{\mathfrak{A}}(\mathcal{C}) = \{Mod_{\mathcal{C}}(\varphi) \mid \varphi \in \mathfrak{A}(FO_\tau)\}.$$

$\square$

This is compatible with our notation $\underline{\Sigma_k}(\mathcal{C})$, cf. Definition 2.4.

Theorem 5.16 *Let $\mathcal{C}$ and $\mathcal{C}'$ be classes of structures over the signatures τ and τ', respectively. Let Φ be a strong first-order reduction from $\mathcal{C}$ to $\mathcal{C}'$, and let $L \subseteq \mathcal{C}$. Then for every syntactic cpAFCT $\mathfrak{A}$ we have*

$$L \in \underline{\mathfrak{A}}(\mathcal{C}) \iff \Phi(L) \in \underline{\mathfrak{A}}(\Phi(\mathcal{C})).$$

Theorem 5.16 will be a simple consequence of the Theorem 5.17, which is prepared by the following definition.

Let $\mathcal{F}$ and $\mathcal{F}'$ be formula classes over τ and τ', respectively. The pair $(\mathcal{F}, \mathcal{F}')$ is *bi-realizable* iff every $\mathcal{F}$-formula is forth-realizable in $\mathcal{F}'$ and every $\mathcal{F}'$-formula is back-realizable in $\mathcal{F}$ wrt. Φ.

Theorem 5.17 *Let Φ be a strong first-order reduction from the class $\mathcal{C}$ of τ-structures into the class $\mathcal{C}'$ of τ'-structures.*

1. *$(FO_\tau, FO_{\tau'})$ is bi-realizable wrt. Φ.*
2. *Let $\mathcal{F}$ and $\mathcal{F}'$ be formula classes over τ and τ', respectively, such that $(\mathcal{F}, \mathcal{F}')$ is bi-realizable wrt. Φ. If $\mathfrak{A} \in \{co\text{-}., \vee\text{-}cl(.), \wedge\text{-}cl(.), \Sigma_1^0(.), FOTC(.), \Sigma_1(.)\}$, then $(\mathfrak{A}(\mathcal{F}), \mathfrak{A}(\mathcal{F}')))$ is bi-realizable wrt. Φ.*
3. *For every syntactic cpAFCT $\mathfrak{A}$ and all formula classes $\mathcal{F}$ and $\mathcal{F}'$ over τ and τ', respectively, where $(\mathcal{F}, \mathcal{F}')$ is bi-realizable wrt. Φ, we have that $(\mathfrak{A}(\mathcal{F}), \mathfrak{A}(\mathcal{F}'))$ is bi-realizable wrt. Φ.*

Proof The first claim is immediate from Lemmas 5.6, 5.9.

Ad 2: Case $\mathfrak{A} = co\text{-}.$: Every $co\text{-}\mathcal{F}$-formula is forth-realizable in $co\text{-}\mathcal{F}'$ by Lemma 5.5, Item 1. Every $co\text{-}\mathcal{F}'$-formula is back-realizable in $co\text{-}\mathcal{F}'$ by Lemma 5.8, Item 1. Thus $(co\text{-}\mathcal{F}, co\text{-}\mathcal{F}')$ is bi-realizable.

Let us consider the case that $\mathfrak{A} = \vee\text{-}cl(.)$. We apply the fixpoint induction principal to the monoton functor f that maps a class $\mathcal{G}$ of τ-formulas to the class $\{\varphi \vee \psi \mid \varphi, \psi \in \mathcal{G}\} \cup \mathcal{F}$. Let G be the class of those formula classes $\mathcal{G}$ each of whose elements is forth-realizable in $\vee\text{-}cl(\mathcal{F}')$.

Firstly, $\emptyset \in G$. Secondly, whenever $\mathcal{G} \in G$, then $f(\mathcal{G}) \in G$ (by Lemma 5.5, Item 2). Thus $\vee\text{-}cl(\mathcal{F})$ (the smallest fixpoint of f) is in G, i.e., every formula in $\vee\text{-}cl(\mathcal{F})$ is forth-realizable in $\mathcal{F}'$.

Using Lemma 5.8, Item 2, instead, one shows analogously that every formula in $\vee\text{-}cl(\mathcal{F}')$ is back-realizable in $\mathcal{F}$. That means that $(\mathcal{F}, \mathcal{F}')$ is bi-realizable.

In the remaining four cases one shows the claim analogously: One exploits Lemma 5.5, Items 3, 4, 5, and 6 in order to show that every formula in $\wedge\text{-}cl(\mathcal{F})$, in $\Sigma_1^0(\mathcal{F})$, in $FOTC(\mathcal{F})$, or in $\Sigma_1(\mathcal{F})$, respectively, is forth-realizable in the respective class of τ'-formulas. One exploits Lemma 5.8, Items 3, 4, 5, and 6 to show that every formula in $\wedge\text{-}cl(\mathcal{F}')$, in $\Sigma_1^0(\mathcal{F}')$, in $FOTC(\mathcal{F}')$, or in $\Sigma_1(\mathcal{F}')$, respectively, is back-realizable in the respective class of τ-formulas.

Ad 3: The proof is by induction on the definition of syntactic cpAFCT. If $\mathfrak{A}$ is the identity, the claim is immediate, and if it is among $co\text{-}.$, $\vee\text{-}cl(.)$, $\wedge\text{-}cl(.)$, $\Sigma_1^0(.)$, $FOTC(.)$, $\Sigma_1(.)$, we apply Claim 2 of this theorem.

If $\mathfrak{A} = \mathfrak{B} \circ \mathfrak{C}$ for some syntactic cpAFCTs $\mathfrak{B}$ and $\mathfrak{C}$, and if $(\mathcal{F}, \mathcal{F}')$ is bi-realizable formula classes, then $(\mathfrak{C}(\mathcal{F}), \mathfrak{C}(\mathcal{F}'))$ is bi-realizable by induction hypothesis on $\mathfrak{C}$, and thus the pair $(\mathfrak{A}(\mathcal{F}), \mathfrak{A}(\mathcal{F}')) = (\mathfrak{B}(\mathfrak{C}(\mathcal{F})), \mathfrak{B}(\mathfrak{C}(\mathcal{F}')))$ is bi-realizable by induction hypothesis on $\mathfrak{B}$.

If $\mathfrak{A} = \mathfrak{B} \sqcup \mathfrak{C}$ for two syntactic cpAFCTs $\mathfrak{B}$, $\mathfrak{C}$, then we may assume by induction hypothesis that for every bi-realizable pair $(\mathcal{G}, \mathcal{G}')$ of formula classes, both the pair $(\mathfrak{B}(\mathcal{G}), \mathfrak{B}(\mathcal{G}))$ and the pair $(\mathfrak{C}(\mathcal{G}), \mathfrak{C}(\mathcal{G}))$ are bi-realizable. If, furthermore, $(\mathcal{F}, \mathcal{F}')$ is a bi-realizable pair of formula classes, we may apply the fixpoint induction principal to the monoton functor that maps a class $\mathcal{G}$ of τ-formulas to $\mathfrak{B}(\mathcal{G}) \cup \mathfrak{C}(\mathcal{G}) \cup \mathcal{F}$ to show that every formula in $\mathfrak{A}(\mathcal{F})$ is forth-realizable in $\mathfrak{A}(\mathcal{F}')$, and argue similarly to show that every formula in $\mathfrak{A}(\mathcal{F}')$ is back-realizable in $\mathfrak{A}(\mathcal{F})$. □

Now we are ready to prove Theorem 5.16. We exploit Theorem 5.17 for sentences.

Proof of Theorem 5.16. Let $\mathfrak{A}$ be a syntactic cpAFCT.

Ad "⇒": Let $L \in \underline{\mathfrak{A}}(\mathcal{C})$. Let φ be an $\mathfrak{A}(FO_\tau)$-sentence that defines L, i.e., $L = Mod_{\mathcal{C}}(\varphi)$. By Proposition 5.6 and Theorem 5.17, the τ-sentence φ is forth-realizable in $\mathfrak{A}(FO_{\tau'})$, i.e., there is an $\mathfrak{A}(FO_{\tau'})$-sentence φ' such that

$$M \models \varphi \iff \Phi(M) \models \varphi'$$

for every structure $M \in \mathcal{C}$. Thus $\Phi(L) = \Phi(Mod_{\mathcal{C}}(\varphi)) = Mod_{\Phi(\mathcal{C})}(\varphi') \in \underline{\mathfrak{A}}(\Phi(\mathcal{C}))$.

Ad "⇐": Assume $\Phi(L) \in \underline{\mathfrak{A}}(\Phi(\mathcal{C}))$. Choose a sentence $\varphi \in \mathfrak{A}(FO_{\tau'})$ that defines $\Phi(L)$ within $\Phi(\mathcal{C})$, i.e., $Mod_{\Phi(\mathcal{C})}(\varphi) = \Phi(L)$.

By Proposition 5.9 and Theorem 5.17, the τ'-sentence φ is back-realizable in $\mathfrak{A}(FO_\tau)$. That means (by Definition 5.7 applied to the case of sentences, i.e., $t = m = 0$ and κ being the empty mapping) that there is an $\mathfrak{A}(FO_\tau)$-sentence $\varphi_\emptyset$ such that

$$M \models \varphi_\emptyset \iff \Phi(M) \models \varphi$$

for every structure $M \in \mathcal{C}$. Since Φ is injective, this implies $L = \Phi^{-1}(\Phi(L)) = Mod_{\mathcal{C}}(\varphi_\emptyset) \in \underline{\mathfrak{A}}(\mathcal{C})$. □

We can conclude the following fact, which allows us to transfer separation results.

Corollary 5.18 *Let $\mathcal{C}$, $\mathcal{C}'$ be classes of relational structures and Φ a strong first-order reduction from $\mathcal{C}$ to $\mathcal{C}'$. Then*

$$\underline{\mathfrak{A}}(\mathcal{C}) \not\subseteq \underline{\mathfrak{B}}(\mathcal{C}) \implies \underline{\mathfrak{A}}(\mathcal{C}') \not\subseteq \underline{\mathfrak{B}}(\mathcal{C}')$$

for all syntactic cpAFCT.

Proof Let $\mathfrak{A}$ and $\mathfrak{B}$ be syntactic cpAFCTs. Let $L \in \underline{\mathfrak{A}}(\mathcal{C}) \setminus \underline{\mathfrak{B}}(\mathcal{C})$. By Theorem 5.16, $\Phi(L) \in \underline{\mathfrak{A}}(\Phi(\mathcal{C})) \setminus \underline{\mathfrak{B}}(\Phi(\mathcal{C}))$. Thus there is a sentence $\varphi \in \mathfrak{A}(FO_{\tau'})$ such that for all $\psi \in \mathfrak{B}(FO_{\tau'})$ we have $(\Phi(L) =)\ Mod_{\Phi(\mathcal{C})}(\varphi) \neq Mod_{\Phi(\mathcal{C})}(\psi)$ and hence (since $\Phi(\mathcal{C}) \subseteq \mathcal{C}'$) also $Mod_{\mathcal{C}'}(\varphi) \neq Mod_{\mathcal{C}'}(\psi)$. This implies $Mod_{\mathcal{C}'}(\varphi) \in \underline{\mathfrak{A}}(\mathcal{C}') \setminus \underline{\mathfrak{B}}(\mathcal{C}')$. □

In Section 5.4 we will show that, for every $t \geq 0$, there is a strong first-order reduction from the class of t-bit grids into the class of undirected graphs (with degree ≤ 5). This allows us to deduce those separations of Theorem 2.33 that do not involve Δ_1, in particular $\underline{B(\Sigma_k)} \subsetneq \underline{\Sigma_{k+1}}$ for the class of graphs. This proof can be found at the end of the next section.

5.3 Semantic Abstract Formula Class Transformers

The technique developed in the preceding Section 5.2 does not quite work when the closure operations $\Delta_1(.)$ and $\Delta_1^{\mathrm{U}}(.)$ are involved because these two transformations of formula classes are no syntactic AFCT in the sense of Definition 5.10. This is due to the fact that what is or is not a formula in, say, ${\Delta_1}^{\mathcal{C}}$ depends on the class $\mathcal{C}$ of structures. We have just allowed to omit this superscript in case no ambiguity arises. See note of page 27.

However, we may extend our concept to work also for AFCT that involve such dependencies on the class of structures (and are called "semantic" rather than "syntactic").

The reader who is not interested in such a detailed analysis of the monadic hierarchy may skip to the proof of Theorem 2.33 on page 155.

Definition 5.19 A *semantic abstract formula class transformer (semantic AFCT)* is a mapping that maps a pair $(\mathcal{F}, \mathcal{C})$, where $\mathcal{F}$ is a set of formulas and $\mathcal{C}$ is a set of structures over the same signature, into a set of formulas over the same signature, in such a way that $\mathfrak{A}(\mathcal{F}, \mathcal{C})$ is closed under equivalence wrt. $\mathcal{C}$. □

The following definition justifies the view that syntactic AFCT are just a special case of semantic AFCT. Recall that for a formula class $\mathcal{F}$ and a class of structures $\mathcal{C}$ over the same signature, the closure of $\mathcal{F}$ under $\mathcal{C}$-equivalence is denoted $\mathcal{F}^{\mathcal{C}}$.

Definition 5.20 If $\mathfrak{A}$ is a syntactic AFCT, we may *associate* to $\mathfrak{A}$ the semantic AFCT

$$\mathfrak{A}(.)^{\cdot} : (\mathcal{F}, \mathcal{C}) \mapsto \mathfrak{A}(\mathcal{F})^{\mathcal{C}}.$$

By $\Delta_1^{\cdot}(.)$ we denote the semantic AFCT $\mathfrak{A}$ that maps $(\mathcal{F}, \mathcal{C})$ to $\Delta_1^{\mathcal{C}}(\mathcal{F})$. □

When no conflict arises, we write $\mathfrak{A}$ instead of $\mathfrak{A}(.)^{\cdot}$, thus using the same notation for a syntactic and its associated semantic AFCT.

The concept of "being given by closure properties" (see Definition 5.22 below) is analogous and compatible to syntactic AFCT, but with the additional "basic" semantic AFCT:

$$\delta : (\mathcal{F}, \mathcal{C}) \mapsto PB(\mathcal{F})^{\mathcal{C}} \cap PB(co\text{-}\mathcal{F})^{\mathcal{C}}.$$

Definition 5.21 If $\mathfrak{A}$, $\mathfrak{B}$ are semantic AFCT, then we define

- $\mathfrak{A} \circ \mathfrak{B}$ to be the semantic AFCT with $(\mathfrak{A} \circ \mathfrak{B})(\mathcal{F}, \mathcal{C}) = \mathfrak{A}(\mathfrak{B}(\mathcal{F}, \mathcal{C}), \mathcal{C})$ for every class $\mathcal{F}$ of formulas and every class $\mathcal{C}$ of structures over the same signature as $\mathcal{F}$;
- $\mathfrak{A} \sqcup \mathfrak{B}$ to be the semantic AFCT such that for every $\mathcal{F}$, $\mathcal{C}$, the set $(\mathfrak{A} \sqcup \mathfrak{B})(\mathcal{F}, \mathcal{C})$ is the smallest superclass $\mathcal{F}'$ of $\mathcal{F}$ with $\mathcal{F}' \supseteq \mathfrak{A}(\mathcal{F}', \mathcal{C}) \cup \mathfrak{B}(\mathcal{F}', \mathcal{C})$.

□

An important application of the δ-Operation, which is available for *semantic* AFCT only, is the following: $\Delta_k^{\cdot}(.) = \delta \circ \Sigma_k(.)^{\cdot}$ for every $k \geq 1$.

Definition 5.22 The class of semantic AFCT *given by closure properties* (or *semantic cpAFCT* for short) is defined as follows: δ as well as every semantic AFCT associated to a basic syntactic AFCT from Definition 5.14 is a semantic cpAFCT. If $\mathfrak{A}$ and $\mathfrak{B}$ are semantic cpAFCT, then so are $\mathfrak{A} \circ \mathfrak{B}$ and $\mathfrak{A} \sqcup \mathfrak{B}$. □

We observe that if a syntactic AFCT is given by closure properties, then so is its associated semantic AFCT. An example of a semantic cpAFCT is $FO(\Delta_1(.))^{\cdot}$. All the formula classes given in Theorems 2.33 and 2.34 result from the class of first-order formulas by application of some semantic cpAFCT.

Definition 5.23 (cf. Definition 5.15) Let τ be a signature, $\mathcal{C}$ a class of τ-structures. For every semantic AFCT, we let $\underline{\mathfrak{A}}(\mathcal{C}) = \{Mod_{\mathcal{C}}(\varphi) \mid \varphi \in \mathfrak{A}(FO_\tau, \mathcal{C})\}$. □

Let $\mathcal{C}$ be a class of structures and $L \subseteq \mathcal{C}$. Then iso(L) denotes the class of $\mathcal{C}$-structures isomorphic to a structure in L. We use this notation only if $\mathcal{C}$ is clear from the context.

We say that a strong first-order reduction $\Phi\colon \mathcal{C} \to \mathcal{C}'$ *respects isomorphism* if for every $M_1, M_2 \in \mathcal{C}$ we have that M_1 and M_2 are isomorphic iff $\Phi(M_1)$ and $\Phi(M_2)$ are.

We wish to show the following theorem, which essentially states that Theorem 5.16 can be saved for semantic cpAFCT if the image of $\mathcal{C}$ under Φ is (up to isomorphism of $\mathcal{C}'$-structures) definable in the respective class of formulas.

Theorem 5.24 *(cf. Theorem 5.16) Let $\mathcal{C}$, $\mathcal{C}'$ be classes of structures over signatures τ and τ', respectively. Let $\Phi\colon \mathcal{C} \to \mathcal{C}'$ be a strong first-order reduction that respects isomorphism, and let $L \subseteq \mathcal{C}$ be closed under $\mathcal{C}$-isomorphism. Let $\mathfrak{A}$ be a semantic cpAFCT.*

$$\text{iso}(\Phi(\mathcal{C})) \in \underline{\mathfrak{A}}(\mathcal{C}') \wedge L \in \underline{\mathfrak{A}}(\mathcal{C}) \implies \text{iso}(\Phi(L)) \in \underline{\mathfrak{A}}(\mathcal{C}'), \tag{5.1}$$

$$L \in \underline{\mathfrak{A}}(\mathcal{C}) \impliedby \text{iso}(\Phi(L)) \in \underline{\mathfrak{A}}(\mathcal{C}'). \tag{5.2}$$

Again, we prepare the proof of the above theorem.

Theorem 5.25 *(cf. Theorem 5.17) Let Φ be a strong first-order reduction from the class $\mathcal{C}$ of τ-structures into the class $\mathcal{C}'$ of τ'-structures.*

For every semantic cpAFCT $\mathfrak{A}$ and all formula classes $\mathcal{F}$ and $\mathcal{F}'$ over τ and τ', respectively, we have:

1. *if every $\mathcal{F}$-formula is forth-realizable in $\mathcal{F}'$ and* iso$(\Phi(\mathcal{C})) \in \mathfrak{A}(\mathcal{F}', \mathcal{C}')$, *then every $\mathfrak{A}(\mathcal{F}, \mathcal{C})$-formula is forth-realizable in $\mathfrak{A}(\mathcal{F}', \mathcal{C}')$.*
2. *if every $\mathcal{F}'$-formula is back-realizable in $\mathcal{F}$, then every $\mathfrak{A}(\mathcal{F}', \mathcal{C}')$-formula is back-realizable in $\mathfrak{A}(\mathcal{F}, \mathcal{C})$.*

Proof For both claims we argue by induction on the definition of semantic cpAFCT.

Ad 1: Let $\mathfrak{A}$ be a semantic cpAFCT. If $\mathfrak{A}$ is associated to a syntactic cpAFCT, then the claim follows from Theorem 5.17. If $\mathfrak{A}$ is of the form $\mathfrak{B} \circ \mathfrak{C}$ or $\mathfrak{B} \sqcup \mathfrak{C}$, then we may use the induction hypothesis and argue as in the proof of Theorem 5.17. In particular, the induction step works for the case $\mathfrak{A} = PB(.)^{\cdot}$.

It remains to consider the case that $\mathfrak{A} = \delta$. Let $\mathcal{F}, \mathcal{F}'$ be classes of formulas over τ and τ', respectively. Assume that every $\mathcal{F}$-formula is forth-realizable in $\mathcal{F}'$ and iso$(\Phi(\mathcal{C})) \in \mathfrak{A}(\mathcal{F}', \mathcal{C}')$. Let $\varphi \in \delta(\mathcal{F}, \mathcal{C}) = PB(\mathcal{F})^{\mathcal{C}} \cap PB(co\text{-}\mathcal{F})^{\mathcal{C}}$. We have to show that φ is forth-realizable in $\delta(\mathcal{F}', \mathcal{C}')$.

There are $\varphi_+ \in PB(\mathcal{F}^{\mathcal{C}})$ and $\varphi_- \in PB((co\text{-}\mathcal{F})^{\mathcal{C}})$, both equivalent to φ over $\mathcal{C}$. We have just observed that every $PB(\mathcal{F}^{\mathcal{C}})$-formula is forth-realizable in $PB(\mathcal{F}'^{\mathcal{C}'})$. Thus also every $PB((co\text{-}\mathcal{F})^{\mathcal{C}})$-formula is forth-realizable in $PB((co\text{-}\mathcal{F}')^{\mathcal{C}'})$.

Thus there are formulas $\varphi'_+ \in PB(\mathcal{F}'^{\mathcal{C}'})$ and $\varphi'_- \in PB((co\text{-}\mathcal{F}')^{\mathcal{C}'})$ both of which forth-realize φ and are thus equivalent over $\Phi(\mathcal{C})$ (but maybe not over $\mathcal{C}'$).

Since iso$(\Phi(\mathcal{C})) \in \delta(\mathcal{F}', \mathcal{C}')$ there is a sentence $\pi' \in PB(\mathcal{F}')^{\mathcal{C}'} \cap PB(co\text{-}\mathcal{F}')^{\mathcal{C}'}$ with $Mod_{\mathcal{C}'}(\pi') = $ iso$(\Phi(\mathcal{C}))$. The formulas $\varphi'_+ \wedge \pi$ and $\varphi_- \wedge \pi$ are equivalent over $\mathcal{C}'$. (To see this, let $M' \in \mathcal{C}'$. If M is isomorphic to a Φ-image of a $\mathcal{C}$-structure, then M fulfills $(\varphi'_+ \wedge \pi) \leftrightarrow (\varphi_- \wedge \pi)$, and if M is not, then π' is invalid in M and thus M fulfills neither $(\varphi'_+ \wedge \pi)$ nor $(\varphi_- \wedge \pi)$.)

Obviously, $\varphi'_+ \wedge \pi' \in PB(\mathcal{F}')^{\mathcal{C}'}$ and $\varphi'_- \wedge \pi' \in PB(co\text{-}\mathcal{F}')^{\mathcal{C}'}$, hence $\varphi'_+ \wedge \pi' \in PB(\mathcal{F}')^{\mathcal{C}'} \cap PB(co\text{-}\mathcal{F}')^{\mathcal{C}'} = \delta(\mathcal{F}', \mathcal{C}')$. Besides, $\varphi'_+ \wedge \pi'$ forth-realizes φ (because it is equivalent to φ'_+ over $\Phi(\mathcal{C})$). Thus φ is forth-realizable in $\delta(\mathcal{F}', \mathcal{C}')$.

Ad 2: Let $\mathcal{F}$, $\mathcal{F}'$ be formula classes such that every $\mathcal{F}'$-formula is back-realizable in $\mathcal{F}$. We have to show that every formula in $\mathfrak{A}(\mathcal{F}', \mathcal{C}')$ is back-realizable in $\mathfrak{A}(\mathcal{F}, \mathcal{C})$. Like above it suffices to consider the case that $\mathfrak{A} = \delta$.

Let $\varphi' \in \delta(\mathcal{F}', \mathcal{C}') = PB(\mathcal{F}')^{\mathcal{C}'} \cap PB(co\text{-}\mathcal{F}')^{\mathcal{C}'}$. There are $\varphi'_+ \in PB(\mathcal{F}')^{\mathcal{C}'}$ and $\varphi'_- \in PB(co\text{-}\mathcal{F}')^{\mathcal{C}'}$ both equivalent to φ' over $\mathcal{C}'$ (in particular equivalent over $\Phi(\mathcal{C}) \subseteq \mathcal{C}'$). Again we may apply the induction hypothesis to $PB(.)^{\cdot}$. Thus there are families $(\varphi_{+,\kappa})_\kappa$ and $(\varphi_{-,\kappa})_\kappa$ of formulas in $PB(\mathcal{F}^{\mathcal{C}})$ and $PB(co\text{-}\mathcal{F}^{\mathcal{C}})$, respectively, that back-realize φ'_+ and φ'_-, respectively. Then both families back-realize φ'. Hence $\varphi_{+,\kappa}$ and $\varphi_{-,\kappa}$ are equivalent over $\mathcal{C}$ for every κ. Thus $(\varphi_{+,\kappa})_\kappa$ is a family of $PB(\mathcal{F})^{\mathcal{C}} \cap PB(co\text{-}\mathcal{F})^{\mathcal{C}}$-formulas, so φ' is back-realizable in $\delta(\mathcal{F},\mathcal{C})$. □

Proof of Theorem 5.24 For the first implication, assume $L \in \underline{\mathfrak{A}}(\mathcal{C})$, i.e., there is a sentence $\varphi \in \mathfrak{A}(FO_\tau,\mathcal{C})$ with $Mod_{\mathcal{C}}(\varphi) = L$. Assume, moreover, that $\mathrm{iso}(\Phi(\mathcal{C})) \in \underline{\mathfrak{A}}(\mathcal{C}')$. By Proposition 5.6 and Theorem 5.25, the sentence φ is forth-realizable in $\mathfrak{A}(FO_{\tau'},\mathcal{C}')$ wrt. Φ. That means that there is a sentence φ' in $\mathfrak{A}(FO_{\tau'},\mathcal{C}')$ with $\mathrm{iso}(\Phi(L)) = Mod_{\mathcal{C}'}(\varphi')$, hence $\mathrm{iso}(\Phi(L)) \in \underline{\mathfrak{A}}(\mathcal{C}')$.

For the second implication statement, assume that $\mathrm{iso}(\Phi(L)) \in \underline{\mathfrak{A}}(\mathcal{C}')$. There is a sentence $\varphi' \in \mathfrak{A}(FO_{\tau'},\mathcal{C}')$ with $Mod_{\mathcal{C}'}(\varphi') = \mathrm{iso}(\Phi(L))$. By Proposition 5.9 and Theorem 5.25, φ' is back-realizable in $\mathfrak{A}(FO_\tau,\mathcal{C})$. That means that there is a sentence $\varphi_\emptyset \in \mathfrak{A}(FO_\tau,\mathcal{C})$ such that (the family containing only) $\varphi_\emptyset$ back-realizes φ'. Then $L = \mathrm{iso}(L) = \{M \in \mathcal{C} \mid \Phi(M) \in \mathrm{iso}(\Phi(L))\} = Mod_{\mathcal{C}}(\varphi_\emptyset) \in \underline{\mathfrak{A}}(\mathcal{C})$. (The first two equalities are by assumption on L and Φ, respectively, the third is by choice of $\varphi_\emptyset$. □

Corollary 5.26 *(cf. Corollary 5.18) Let $\mathcal{C}$, $\mathcal{C}'$ be classes of relational structures and $\Phi\colon \mathcal{C} \to \mathcal{C}'$ be a strong first-order reduction that respects isomorphism. Let $\mathfrak{A}$, $\mathfrak{B}$ be semantic cpAFCT such that* $\mathrm{iso}(\Phi(\mathcal{C})) \in \mathfrak{A}(\mathcal{C}')$. *Then*

$$\underline{\mathfrak{A}}(\mathcal{C}) \not\subseteq \underline{\mathfrak{B}}(\mathcal{C}) \implies \underline{\mathfrak{A}}(\mathcal{C}') \not\subseteq \underline{\mathfrak{B}}(\mathcal{C}').$$

Proof Let $L \in \underline{\mathfrak{A}}(\mathcal{C}) \setminus \underline{\mathfrak{B}}(\mathcal{C})$. We apply Theorem 5.24, namely (5.1) to $\mathfrak{A}$ and the contraposition of (5.2) to $\mathfrak{B}$, and conclude $\Phi(L) \not\in \underline{\mathfrak{A}}(\mathcal{C}') \setminus \underline{\mathfrak{B}}(\mathcal{C}')$. □

Recall that our aim was to show Theorems 2.33 and 2.34. The former states the following non-inclusions for the class of graphs:

$$\begin{array}{rcl} \underline{FO^{\mathrm{TC}}},\ \underline{\Pi^0_{k-1}(\Sigma_1)},\ \underline{FO(\Sigma_1)},\ \underline{\Pi_k},\ \underline{\Sigma_{k+1}},\ \underline{\Delta_{k+1}} & \not\subseteq & \underline{\Sigma_k}, \\ \underline{\Delta_{k+1}},\ \underline{\Sigma_1(\Pi^0_{k-1}(\Sigma_1))},\ \underline{\Pi_1(\Sigma^0_{k-1}(\Pi_1))} & \not\subseteq & \underline{B(\Sigma_k)},\ \underline{FO(\Sigma_{k-1})}, \end{array}$$

and the latter states the following for the class of *connected* graphs:

$$\underline{\Pi^0_k(\Delta_1)}, \underline{FO(\Delta_1)} \not\subseteq \underline{\Sigma_k}.$$

Now we can present the proofs of these theorems, using the existence of the announced strong first-order reductions, which are presented in the next section.

Proof of Theorem 2.33 From Corollary 2.25 we know that

$$\underline{FO^{\mathrm{TC}}} \not\subseteq \underline{\Sigma_k} \quad\text{and}\quad \underline{\Pi^0_{k-1}(\Sigma_1)} \supseteq \underline{\Pi^0_k(\Delta_1)} \not\subseteq \underline{\Sigma_k}$$

for the class of 1-bit grids. From Corollary 5.18 and Proposition 5.29 we conclude that $\underline{FO^{\mathrm{TC}}}$, $\underline{\Pi^0_{k-1}(\Sigma_1)} \not\subseteq \underline{\Sigma_k}$ also for the class of graphs (with in- and out-degree ≤ 4). Invoking Corollary 5.18 again, this time with Proposition 5.30, we conclude that the above separations are true also for the class of undirected graphs (with degree ≤ 5).

By calculation rules as presented in Subsection 2.2.3, we infer the separation claims $\underline{FO(\Sigma_1)}, \underline{\Pi_k}, \underline{\Sigma_{k+1}}, \underline{\Delta_{k+1}} \not\subseteq \underline{\Sigma_k}$.

Turning to the second statement of Theorem 2.33, we conclude from Corollary 2.31 (by calculation rules) that

$$\underline{\Delta_{k+1}},\ \underline{\Sigma_1(\Pi^0_{k-1}(\Sigma_1))},\ \underline{\Pi_1(\Sigma^0_{k-1}(\Pi_1))} \not\subseteq \underline{B(\Sigma_k)},\ \underline{FO(\Sigma_{k-1})}$$

for the class of grids. From Corollary 5.26 and Propositions 5.29 and 5.30 we conclude that this is true also for the class of graphs and undirected graphs. (In all cases except for the Δ_{k+1}, the easier Corollary 5.18 instead of Corollary 5.26 will do and the additional claim of Proposition 5.29 is not needed either.) □

Proof of Theorem 2.34 From Corollary 2.25 we know that $\underline{\Pi^0_k(\Delta_1)} \not\subseteq \underline{\Sigma_k}$ for the class of 1-bit grids. From Propositions 5.29 and 5.30 we conclude that this is true also for the class connected graphs and connected undirected graphs with the respective bound on the degree. □

In the next section we will prove the announced Propositions 5.29 and 5.30. For the strong version of Proposition 5.29, as required for the transfer of separation claims that involve Δ_k, we will need the following corollary of Theorem 5.16.

Corollary 5.27 *Let $\mathcal{C}$ and $\mathcal{C}'$ be classes of structures over the signatures τ and τ', respectively. Let Φ be a strong first-order reduction from $\mathcal{C}$ to $\mathcal{C}'$, and let $L \subseteq \mathcal{C}$. Let $\mathfrak{A}$ be a syntactic cpAFCT such that $\mathfrak{A}(FO_{\tau'})$ is closed under conjunction. Assume* $\mathrm{iso}(\Phi(\mathcal{C})) \in \underline{\mathfrak{A}}(\mathcal{C}')$ *and* $L \in \underline{\mathfrak{A}}(\mathcal{C})$. *Then* $\mathrm{iso}(\Phi(L)) \in \underline{\mathfrak{A}}(\mathcal{C}')$.

Proof By assumption, there is $\pi' \in \mathfrak{A}(FO_{\tau'})$ with $Mod_{\mathcal{C}'}(\pi') = \text{iso}(\Phi(\mathcal{C}))$. By Theorem 5.16 we have $\Phi(L) \in \underline{\mathfrak{A}}(\Phi(\mathcal{C}))$, thus there is $\varphi' \in \mathfrak{A}(FO_{\tau'})$ with $Mod_{\Phi(\mathcal{C})}(\varphi') = \Phi(L)$.
Then $\text{iso}(\Phi(L)) = Mod_{\text{iso}(\Phi(\mathcal{C}))}(\varphi') = Mod_{\mathcal{C}'}(\varphi' \wedge \pi') \in \underline{\mathfrak{A}}(\mathcal{C}')$. □

5.4 Examples of Strong First-Order Reductions

In this section we will give some examples of strong first-order reductions, namely from the class of grids to the class of finite directed graphs, from the class of t-bit grids to the class of finite directed graphs (for every t), and from the class of finite directed graphs to the class of finite undirected graphs.

In all cases the image of the reduction is Π_1-definable.

The approach is always the same: For the strong first-order reduction from $\mathcal{C}$ to $\mathcal{C}'$ we introduce a $\mathcal{C}'$-gadget that replaces each node of a $\mathcal{C}$-structure. The crucial point is that in the Φ-image of a structure, there are no induced subgraphs isomorphic to the respective gadgets except for the gadgets themselves. That is why a first-order formula ψ (for Item 2 of Definition 5.1) can be found that allows to fix these gadgets (simply by describing them).

5.4.1 From Grids to Directed Graphs

In this subsection we shall define a first-order reduction from the class of grids into the class of graphs. Recall that by "graph" we always mean "finite, directed graph", i.e., a graph is a finite structure in the signature $\{E\}$, where E is binary. A graph G is called *undirected* if the edge relation E^G is irreflexive and symmetric. We use all the usual notions like "path", "degree" etc.

We define for every grid R the graph $\Phi(R) := (\{1,2\} \times \text{dom}R, E)$ with

$$\begin{array}{rl} E = & \{((1,x),(1,x)) \mid x \in \text{dom}R\} \\ \cup & \{((1,x),(2,x)) \mid x \in \text{dom}R\} \\ \cup & \{((2,x),(1,y)) \mid (x,y) \in S_1^R\} \\ \cup & \{((2,x),(2,y)) \mid (x,y) \in S_2^R\}. \end{array}$$

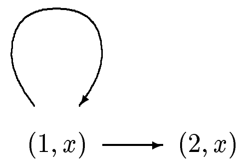


Figure 5.1: Gadget to encode grid node x in directed graphs

Figure 5.1 illustrates a "gadget" that replaces a grid node. The S_1- (respectively S_2-) edges of a grid are then coded by graph edges from 2- to 1-nodes (to 2-nodes, respectively).

For this mapping we claim the following.

Proposition 5.28 *Φ is a strong first-order reduction from the class of grids into the class of graphs (with in-degree ≤ 2 and out-degree ≤ 2).*

Moreover, its image $\Phi(Grids)$ is (up to graph isomorphism) Π_1-definable within the class of all graphs, and even first-order definable within the class of connected graphs.

The additional statement is relevant only in the context of Corollary 5.26, i.e., for separations involving the Δ_1-closure.

Proof of Proposition 5.28 We give the formulas required to verify that Φ is indeed a strong first-order reduction. For Item 2 of Definition 5.1 we use $\psi(x_1,x_2) = Ex_1x_1 \wedge Ex_1x_2 \wedge \neg Ex_2x_2$. Then the formulas for Item 3 are:

$$\begin{array}{rcl} \varphi^E_{(1,1)}(x_1,x_2) & = & x_1{=}x_2, \\ \varphi^E_{(1,2)}(x_1,x_2) & = & x_1{=}x_2, \\ \varphi^E_{(2,1)}(x_1,x_2) & = & S_1x_1x_2, \\ \varphi^E_{(2,2)}(x_1,x_2) & = & S_2x_1x_2. \end{array}$$

The formulas for Item 4 of Definition 5.1 are:

$$\begin{array}{rcl} \varphi^{S_1}(x_1,y_1) & = & \exists x_2 \exists y_2(\psi(x_1,x_2) \wedge \psi(y_1,y_2) \wedge Ex_2y_1), \\ \varphi^{S_2}(x_1,y_1) & = & \exists x_2 \exists y_2(\psi(x_1,x_2) \wedge \psi(y_1,y_2) \wedge Ex_2y_2). \end{array}$$

One easily verifies the properties required for a strong first-order reduction, completing the proof of the first statement in Proposition 5.28.

For the second statement we have to show that the class $\Phi(\mathit{Grids})$ of those graphs that are the Φ-image of a grid is definable firstly by a Π_1-formula (within the class of graphs) and secondly by a first-order formula within the class of connected graphs.

Let us call a τ_{Grids}-structure R *irreflexive* if S_1^R and S_2^R are irreflexive. We note that the given first-order reduction Φ is in fact a strong first-order reduction from the class of irreflexive τ_{Grids}-structures into the class of graphs, and its image is first-order definable within the class of graphs. Besides, the image $\Phi(R)$ of an irreflexive τ_{Grids}-structure R is connected iff R is.

We show:

1. The class of grids is (up to isomorphism) first-order definable within the class of connected, irreflexive τ_{Grids}-structures.
2. The image under Φ of the class of connected, irreflexive τ_{Grids}-structures is (up to isomorphism) first-order definable within the class of connected graphs.
3. The same image is Π_1-definable within the class of graphs.

Then we are finished because of Corollary 5.27: From Items 1 and 2 we infer the first-order definability of $\Phi(\mathit{Grids})$ within the class of connected graphs. And from Items 1 and 3 we infer the Π_1-definability within the class of all graphs.

Ad 1: Choose a first-order formula to assert for a structure M in τ_{Grids} that, firstly, S_1 and S_2 are partial injective functions that commute; secondly that the "borders" look like in a grid, i.e. if $\{i,j\} = \{1,2\}$ and whenever $S_i xy$, then x has a S_j-predecessor (-successor, respectively) iff y has; and thirdly that there is exactly one vertex that has no successor (predecessor, respectively) at all.

Ad 2: The following first-order sentence does the trick.

$$\begin{aligned}\varphi = {} & \forall x(Exx \to \exists! y(Exy \wedge \neg(x=y)) \wedge \exists y(Exy \wedge \neg Eyy)) \wedge \\ & \forall y(\neg Eyy \to \exists! x(Exy \wedge Exx)).\end{aligned}$$

It is easy to see that every Φ-image of an irreflexive τ_{Grids}-structure fulfills φ. For the converse direction, let $G = (V^G, E^G)$ be a graph fulfilling φ. Choose

$$\begin{aligned} V_1 &= \{x \in V^G \mid (x,x) \in E^G\}, \\ V_2 &= \{x \in V^G \mid (x,x) \notin E^G\}. \end{aligned}$$

Since G fulfills φ, there is a bijection $\alpha\colon V_1 \to V_2$ such that for every $x \in V_1$, the only two edges leaving x lead to x and $\alpha(x)$. Then G is isomorphic to the irreflexive τ_{Grids}-structure (V_1, S_1, S_2), where

$$\begin{aligned} S_1 &= \{(x,y) \mid (\alpha(x), y) \in E^G\}, \\ S_2 &= \{(x,y) \mid (\alpha(x), \alpha(y)) \in E^G\}. \end{aligned}$$

Ad 3: We build the conjunction of the previous sentence φ with a Π_1-sentence ψ that expresses connectedness. Choose ψ to assert that for every set X of nodes for which neither X nor its complement are empty, there is an edge linking a node inside X with one outside X. $\square$

5.4.2 From Coloured Grids to Directed Graphs

Now we shall strengthen the result of the preceding subsection by giving a strong first-order reduction that even works in the presence of "colours". For this strong first-order reduction from t-bit grids (for some $t \geq 0$) to directed graphs, we need gadgets similar to the ones above but with the additional feature of coding some t-bit colour.

Let $M = \{(1,1),(2,1)\} \cup \{(i,i+1) \mid 1 \leq i \leq t+1\}$. For every t-bit grid R let $\Phi(R)$ be the graph with domain $[1,t+2] \times \mathrm{dom}R$ and edge set $\{((i,x),(j,x)) \mid (i,j) \in M, x \in \mathrm{dom}R\} \cup \{((i+2,x),(i+2,x)) \mid x \in X_i^R\} \cup \{((2,x),(1,y)) \mid (x,y) \in S_1^R\} \cup \{((2,x),(2,y)) \mid (x,y) \in S_2^R\}$.

Figure 5.2 illustrates this gadget for the case $k=3$.

Proposition 5.29 *Φ is a strong first-order reduction from the class of t-bit grids into the class of graphs (with in-degree ≤ 3 and out-degree ≤ 4)*

Moreover, its image $\Phi(\mathit{Grids}_t)$ is (up to isomorphism) Π_1-definable within the class of all graphs, and even first-order definable within the class of connected graphs.

Proof Let $\psi(x_1,\dots,x_{t+2}) = \bigwedge_{(i,j)\in M} Ex_ix_j$. Then the formulas for Item 3 of Definition 5.1 are:

$$\varphi^E_{(i,j)}(x_1,x_2) = \begin{cases} x_1{=}x_2 & \text{if } j=i+1 \vee i=j=1, \\ S_1x_1x_2 \vee x_1{=}x_2 & \text{if } (i,j)=(2,1), \\ S_2x_1x_2 & \text{if } (i,j)=(2,2), \\ x_1{=}x_2 \wedge X_{i-2}x_1 & \text{if } i=j\geq 3, \\ \neg(x_1{=}x_1) & \text{else.} \end{cases}$$

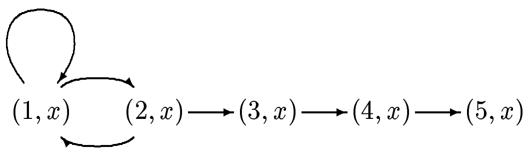


Figure 5.2: Gadget to encode 3-bit grid node x in directed graphs

The formulas for Item 4 of Definition 5.1 are:

$$\begin{aligned}
\varphi^{S_1}(x_1,y_1) &= \exists x_2\dots x_{t+2}\,\exists y_2\dots y_{t+2}\\
&\quad (\psi(x_1,\dots,x_{t+2})\wedge\psi(y_1,\dots,y_{t+2})\wedge Ex_2y_1),\\
\varphi^{S_2}(x_1,y_1) &= \exists x_2\dots x_{t+2}\exists y_2\dots y_{t+2}\\
&\quad (\psi(x_1,\dots,x_{t+2})\wedge\psi(y_1,\dots,y_{t+2})\wedge Ex_2y_2),\\
\varphi^{X_i}(x_1) &= \exists x_2\dots x_{t+2}(\psi(x_1,\dots,x_{t+2})\wedge Ex_{i+2}x_{i+2}).
\end{aligned}$$

Using these formulas it is straightforward to verify that Φ is indeed a strong first-order reduction. The additional claim of Proposition 5.29 is shown in the same way as for the reduction from grids to graphs. $\Box$

5.4.3 From Directed Graphs to Undirected Graphs

We will show that there is a strong first-order reduction from directed graphs to undirected graphs. For this purpose we define, for every directed graph G, the undirected graph $\Phi(G)$ with domain $[1,6]\times \mathrm{dom}G$ whose edges are as in Figure 5.3 plus an edge between $(1,x)$ and $(6,y)$ whenever (x,y) is an edge in G.

Formally: Let $M=\{\{1,2\},\{2,3\},\{3,4\},\{4,5\},\{5,2\},\{3,5\},\{5,6\}\}$ and $N=\{\{1,3\},\{1,4\},\{1,5\},\{2,4\},\{2,6\},\{3,6\},\{4,6\}\}$. The edge set of $\Phi(G)$ is given as $\{((i,x),(j,y)) \mid (x=y\wedge\{i,j\}\in M)\vee((i,j)=(1,6)\wedge(x,y)\in E^G)\vee((i,j)=(6,1)\wedge(y,x)\in E^G)\}$.

Proposition 5.30 *Φ is a strong first-order reduction from the class of graphs into the class of undirected graphs. Its image is (up to isomorphism) first-order definable. Besides, the Φ-image of a graph whose in- and out-degree are both ≤ 4 has degree ≤ 5.*

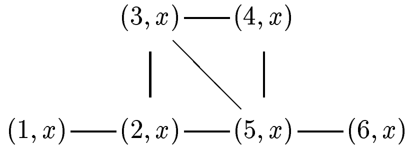


Figure 5.3: Gadget to encode graph node x in undirected graphs

Proof The formulas required to show that Φ is indeed a strong first-order reduction are the following: For Item 2 of Definition 5.1 let

$$\begin{aligned}
\psi(x_1,\dots,x_6) = {}& \forall z\,(Ezx_4\to(x_3=z\vee x_5=z))\\
&\wedge \textstyle\bigwedge_{\{i,j\}\in M}(Ex_ix_j)\wedge\bigwedge_{\{i,j\}\in N}(\neg Ex_ix_j).
\end{aligned}$$

ψ fixes indeed the six nodes of a gadget (in the sense of Item 2 of Definition 5.1). To verify this, let us consider an induced subgraph U of an undirected graph $\Phi((V,E))$ such that U satisfies ψ, say consisting of vertices $u_1,\dots,u_6$ as interpretations of $x_1,\dots,x_6$, respectively. If U constitutes a gadget of $\Phi((V,E))$, i.e., if $(u_1,\dots,u_6)=((1,u),\dots,(6,u))$ for some $u\in V$, then we are done. Otherwise, the vertices u_2,u_3,u_4,u_5 have to be in $\{1,6\}\times V$, i.e., "endpoints" of at least two different gadgets. Thus in $\Phi((V,E))$, the node u_4 is adjacent with a vertex in $\{2,5\}\times V$ different from u_3 and u_5, contradicting the condition ψ.

The following formulas show Item 3 of Definition 2:

$$\varphi^E_{(i,j)}(x,y)=\begin{cases} x=y & \text{if } \{i,j\}\in M,\\ Exy & \text{if } (i,j)=(1,6),\\ Eyx & \text{if } (j,i)=(1,6),\\ \neg(x=x) & \text{else.}\end{cases}$$

For Item 4 of Definition 2 we need:

$$\begin{aligned}
\varphi^E(x,y) = {}& \exists x_2\dots x_6\,\exists y_1\dots y_5(\\
&\psi(x,x_2,\dots,x_6)\wedge\psi(y_1,\dots,y_5,y)\wedge Exy).
\end{aligned}$$

These formulas show that Φ is a strong first-order reduction from the class of graphs to the class of undirected graphs.

For the second claim of Proposition 5.30 we need a first-order formula that asserts for an undirected graph that it consists of disjoint copies of the gadget in Figure 5.3, possibly linked in the way described above. This is done by the following formula:

$$
\begin{aligned}
&\forall x \forall y \exists x_1 \dots x_6 \exists y_1 \dots y_6 \Big(\psi(x_1, \dots, x_6) \wedge \psi(y_1, \dots, y_6) \\
&\wedge \big(\textstyle\bigvee_{i \le 6} x = x_i \big) \wedge \big(\textstyle\bigvee_{i \le 6} y = y_i \big) \\
&\wedge \Big(Exy \rightarrow \big((x_1 = y_1) \vee (x = x_1 \wedge y = y_6) \vee (x = x_6 \wedge y = y_1) \big) \Big) \Big),
\end{aligned}
$$

where ψ is as above.

The claim about the degree is immediate. □

Chapter 6

Quantifier Depth

In the previous chapters we considered the quantifier alternation hierarchy of monadic second-order logic, which measures the complexity of formulas by the alternation of quantifiers.

In this chapter we will concentrate on the quantifier depth, which is a distinct measure for the complexity of formulas. The quantifier depth of a formula is the maximal number of nested quantifiers. For a formula in prenex normal form, this is the length of the quantifier prefix.

We distinguish a variant of the quantifier depth designed for monadic second-order logic. For this measure, we count the number of nestings of set quantifiers and neglect the alternation inside first-order subformulas. This variant is called *set quantifier depth.*. As [Ott95] has shown, this measure gives rise to a hierarchy of expressiveness, even within the existential fragment of monadic second-order logic.

However, the situation looks different for particular, less general classes of structures. For example, the monadic second-order quantifier depth hierarchy of the existential fragment of monadic second-order logic over words collapses to its first level, that means:

Fact 6.1 ([Tho82]) *Every sentence of existential monadic second-order logic over words is equivalent to a sentence of the form $\exists X \varphi(X)$ where φ is first-order.*

Note that this statement remains true even if the word "existential" is removed. A word language is definable by such a formula iff it is regular.

Our aim is to transfer the above fact (and its proof) to pictures instead of words.

6.1 More Definitions and Results

In this section we present results that concern the set quantifier depth. These and the corresponding notions have not been mentioned in Chapter 2, which puts the emphasis on alternation hierarchies.

The following definitions are relative to some fixed signature τ.

Definition 6.2 The *set quantifier depth* $sqd(\varphi)$ of a monadic second-order formula φ is defined as follows: If φ is atomic, then $sqd(\varphi) = 0$. Besides, $sqd(\varphi \vee \psi) = \max\{sqd(\psi), sqd(\psi)\}$, $sqd(\neg\varphi) = sqd(\varphi)$, $sqd(\exists X \varphi) = sqd(\varphi) + 1$ for every set variable X, and

$$sqd(\exists x \varphi) = \begin{cases} 0 & \text{if } sqd(\varphi) = 0 \\ sqd(\varphi) + 1 & \text{else,} \end{cases}$$

for every first-order variable x_0.
A formula is of *existential quantifier depth* k iff it is of the form $\exists X_1 \dots X_k(\varphi)$ for a first-order φ. □

If a monadic second-order formula φ is of the form $Q_1 X_1 \dots Q_k X_k \psi$, where $Q_1, \dots, Q_k \in \{\exists, \forall\}$ and ψ is first-order, then φ is of set quantifier depth k.

There is a trivial relationship between set quantifier depth and alternation depth: If a formula has set quantifier depth k, then it is (up to logical equivalence) in $B(\Sigma_k)$. Thus the strictness of the monadic alternation hierarchy implies the infiniteness of the quantifier depth hierarchy. The following result of Otto shows that even inside the existential fragment of monadic second-order logic, a deeper nesting of quantifiers gives more expressive power.

Fact 6.3 ([Ott95]) *The existential quantifier depth hierarchy over graphs is infinite.*

Otto's proof can easily be modified (using Ehrenfeucht-Fraïsé games with multiple colouring rounds) to show the following: There is a (first-order definable) class $\mathcal{C}$ of finite structures such that for every $k \geq 0$ there is a

sentence φ of existential quantifier depth $k+1$ that is not equivalent over $\mathcal{C}$ to any monadic second-order sentence (no matter if existential or not) of set quantifier depth k.

The contribution of this chapter is that the above statement is not true for the class of grids instead of graphs, but that—to the contrary—we have the following analogue of Fact 6.1 for the class of grids instead of words.

Fact 6.4 *For the class of grids, the quantifier depth hierarchy within the existential fragment of monadic second-order logic collapses to its first level, more precisely: For every Σ_1-formula φ over τ_{Grids} there is a formula of existential quantifier depth 1 which is equivalent to φ over the class of grids.*

The proof of Fact 6.4 will be given in Section 6.2.

The following table gives an overview over two hierarchies relevant for this thesis (quantifier alternation and existential quantifier depth) in dependence on the underlying class of structures, here words, grids, and graphs.

Here, "collapse" always means "collapses to the first level".

	quantifier alternation	existential quantifier depth
words	collapse	collapse (see [Tho82])
grids	strict (see [Sch97a])	collapse (see [Mat98c])
graphs	strict (see [MT97])	strict (see [Ott95])

6.2 One Quantifier Will Do in Σ_1 over Grids

This section is concerned with the content of [Mat98c]. We wish to prove the analogue of Fact 6.1 for pictures instead of words.

It is instructive to have a look at how the proof of that Fact 6.1 can be formulated. Firstly, we have the following characterization of Σ_1-definable (i.e., regular) word languages:

Remark 6.5 Every word language definable in existential monadic second-order logic is the image of some local word language by an alphabet projection. □

(See Subsection 2.3.1 for the definition of local word languages.) Note that the above remark also holds if the word "existential" is removed.

By Remark 6.5 it suffices to show the following in order to have Fact 6.1.

Theorem 6.6 *Let L be a projection of a local word language, $\varepsilon \notin L$. Then there exists a first-order formula $\varphi(X)$ such that $L = Mod(\exists X \varphi(X))$.*

The basic idea for the proof of this theorem is the following. Let $L \subseteq \Gamma^+$ be the projection of the local word language M over alphabet Ω via the projection mapping $\alpha : \Omega \to \Gamma$. Say Ω is of size 2^n. We will explain how to construct a formula φ such that $\exists X \varphi(X)$ defines L. Let w be a word over Γ. The word w is in L iff M contains an α-pre-image of w. For convenience, let us assume that the length of w is a multiple of n. Then w can be split into sections of length n. An α-pre-image of *the first letter* of such a section can be encoded by a subset X of positions from this section. This way, a subset X of w-positions encodes the α-pre-images of every first letter of every section. A first-order formula $\varphi(X)$ can check whether this "partial" pre-image can be completed to an α-pre-image of w in M because there are only finitely many words of length n over Ω and thus only finitely many possibilities to complete the pre-image of each section.

Some modifications are necessary to capture also the case of words whose length is not a multiple of n.

We will adapt the proof sketched above to picture languages, i.e., we wish to show the following.

Theorem 6.7 *Every Σ_1-sentence over pictures is equivalent to a sentence of the form $\exists X \varphi(X)$ where φ is first-order.*

However, the proof is a little more complicated than in the word case because now we have to split a picture into sufficiently large (square) blocks rather than into sections. It will not suffice to encode only the α-pre-image of one single position of such a block, but instead we have to encode the α-pre-image of all four border rows and columns. But still these borders are small compared to the size of the inside of the block, thus such an encoding is possible. This proof will be carried out in detail in the following three subsections.

In Subsection 6.2.4, we will conclude that the above theorem holds also for formulas with free variables, thus proving Fact 6.4

Recall that in the previous chapters, in particular in Subsection 3.1.1, we showed that there are formulas that witness the strictness of the alternation hierarchy over grids and whose leading quantifier block has length 1. In Subsection 6.2.5 we will combine this "compression" with the result of this chapter in order to build formulas of in some sense minimal complexity that witness the strictness of the alternation hierarchy.

6.2.1 Picture Models—A Different Way to Colour Grids

In the previous chapters, in particular in Chapter 3 and 4, we considered "t-bit grids", i.e., grids expanded to the signature $\{S_1, S_2, X_1, \dots, X_t\}$ with unary predicates $X_1, \dots, X_t$. Such a t-bit grid may be associated to every non-empty picture over $\{0,1\}^t$.

In this section, it will be convenient to consider a different way to introduce "colours" into grids.

Let Γ be a finite alphabet. Consider the signature $\sigma_\Gamma = \{S_1, S_2, (Q_a)_{a\in\Gamma}\}$, where the Q_a are unary predicates. To every non-empty $m\times n$-picture P over Γ we associate the *picture model*

$$\underline{P} := ([m]\times[n], S_1^P, S_2^P, (Q_a^P)_{a\in\Gamma}),$$

where the two successor relations S_1^P and S_2^P are as before, i.e., $S_1^P = \{((i,j),(i+1,j)) \mid (i,j) \in [m-1]\times[n]\}$ and $S_2^P = \{((i,j),(i,j+1)) \mid (i,j) \in [m]\times[n-1]\}$, and $Q_a^P = P^{-1}(a)$ is the set of all positions that carry the letter $a \in \Gamma$.

If $P \in \Gamma^{+,+}$, $Q \in \Omega^{+,+}$ are pictures of the same size over two alphabets Γ and Ω, then we denote by $P \otimes Q$ the picture over $\Gamma\times\Omega$ for which $(P\otimes Q)\langle x\rangle = (P\langle x\rangle, Q\langle x\rangle)$ for every $x \in \mathrm{dom}P$.

In this section, we will consider monadic second-order formulas in the signature σ_Γ. Thus $Q_a(x)$ is a formula that asserts that the position x carries the letter a. This is, by the way, the signature used in [GRST96]. It is very easy to see that every first-order formula φ over signature σ_Γ can be translated into a formula φ' over signature τ_{Grids} and vice versa, such that φ and φ' are "equivalent" in the sense of Remark 6.22, so for the notion of "Σ_1-definability" etc., it does not matter to which signature we refer.

The *word model* $\underline{w}$ associated to a non-empty word[1] $w \in \Gamma^+$ is the structure over signature $\varrho_\Gamma := \{S, (Q_a)_{a\in\Gamma}\}$, where $\mathrm{dom}\underline{w} = \{1, \dots, |w|\}$ is the set of positions of w, and S^w is the successor relation on $\{1, \dots, |w|\}$, and Q_a^w is the set of those positions of w that carry an a.

6.2.2 Domino-Local Picture Languages

As in Chapter 4, it is useful to base the argument against formulas on a characterization of the expressiveness of Σ_1-formulas by projections of local picture languages. However, here it is—for technical reasons—a bit more

[1] Here we have a minor conflict with our agreement to identify words with pictures of height ≤ 1.

convenient to refer to a slightly different notion of locality, namely "domino-locality".

Definition 6.8 Let P be a picture, and Δ_1 (and Δ_2) be a set of pictures of size $(2,1)$ (or $(1,2)$, respectively) over the same alphabet. Δ_1 (or Δ_2) *tiles* P iff all subblocks of P of size $(2,1)$ (or $(1,2)$) are in Δ_1 (or Δ_2, respectively). We call (Δ_1, Δ_2) a *(local) domino tiling system.*

Let P be a picture over Γ. Recall that $\hat{P}$ denotes the picture that results from P by surrounding it with the fresh boundary symbol #.

A picture language L over Γ is *domino-local* iff there exists a domino tiling system (Δ_1, Δ_2) such that L is the set of non-empty pictures P for which both Δ_1 and Δ_2 tile $\hat{P}$. In that case we write $\mathcal{L}((\Delta_1, \Delta_2)) := L$. □

Domino-local picture languages are called "hv-local" in [GR96]. We have indeed the analogue to Remark 6.5 because every local picture language is a projection of some domino-local picture language.

Theorem 6.9 *([LS94, Mat95, GR96]) Every Σ_1-definable picture language is a projection of some domino-local picture language.*

Theorem 6.9 adds another characterization of recognizable picture languages to Theorem 4.4.

6.2.3 Compression of Existential Quantifier Block

By Theorem 6.9 it suffices to show the following in order to obtain Theorem 6.7:

Theorem 6.10 *Let L be a projection of a domino-local picture language. Then there exists some first-order formula $\varphi(X)$ with $L = Mod(\exists X\, \varphi(X))$.*

The proof will be done in two steps—firstly for pictures of bounded height and then in the general case.

Pictures of Bounded Height

The first step for the proof of Theorem 6.10 is to consider a domino-local picture language restricted to pictures of a fixed height. In this case, the compression of existential quantifier prefixes of Σ_1-formulas over picture models can easily be reduced to the word model case.

Theorem 6.11 *Let L be the image of a domino-local picture language M over Ω and $m \geq 1$ a fixed height. Then there exists a first-order formula $\varphi(X)$ (in signature σ_Ω) such that $Mod(\exists X\, \varphi(X))$ is the set of pictures in L that have height m.*

Proof Assume that L is the image of M over alphabet Ω under a projection mapping π defined on Ω.

Consider the height-m fragment $L_0 = L[m]$ over alphabet $\Omega^{m,1}$. Then L_0 is a projection of the local word language $M[m]$ via the canonical extension of π to columns. By Theorem 6.1 there exists a first-order formula $\varphi_0(X)$ over signature $\varrho_{\Omega^{m,1}}$ such that $L_0 = Mod(\exists X\, \varphi_0(X))$.

Let φ' result from φ_0 as follows: Replace atomic formulas $Q_{\overline{a}}x$ (where $\overline{a} = \begin{pmatrix} a_0 \\ \vdots \\ a_{m-1} \end{pmatrix} \in \Omega^{m,1}$ and x is some first-order variable) by

$$\exists x_0 \ldots \exists x_{m-1} \left(x = x_0 \wedge \bigwedge_{i=0}^{m-2} S_1 x_i x_{i+1} \wedge \bigwedge_{i=0}^{m-1} Q_{a_i} x_i \right),$$

(where $x_0, \ldots x_{m-1}$ are fresh variables,) and replace first-order quantifiers $\exists x(\psi)$ by $\exists x(top(x) \wedge \psi)$ (i.e., first-order quantifiers are relativized by the first-order formula $top(x)$ asserting that position x is in the top-most row.)

For a picture P of height m and some set $U \subseteq \mathrm{dom}P$, we have $\underline{P} \models \varphi'[U]$ iff $\underline{column\text{-}word(P)} \models \varphi_0[\{j \mid (0,j) \in U\}]$. Thus $\underline{P} \models \exists X \varphi'$ iff $\underline{column\text{-}word(P)} \models \exists X \varphi_0$ for every non-empty picture P of height m.

The desired formula φ is obtained by building the conjunction of φ' with a first-order formula that holds exactly for pictures of height m. □

Pictures of Unbounded Height

The second step for the proof of Theorem 6.10 is to consider projections of domino-local picture languages without the restriction of a fixed height. We will sketch the main construction.

The aim is to construct a special Σ_1-formula whose models are exactly the pictures of some projection of a given domino-local picture language M under an alphabet projection π. Let (Δ_1, Δ_2) be a domino tiling system with $\mathcal{L}((\Delta_1, \Delta_2)) = M$.

Without the additional limitation to one single monadic quantifier, such a formula may, informally speaking, work as follows: (1) Guess a colouring of the input picture, where the colours encode the letters of the alphabet of M, and then (2) check that the local restrictions are fulfilled, i.e., the picture obtained this way is tiled by Δ_1 and Δ_2.

Let us call the colouring guessed in step (1) "polychrome". This polychrome colouring has potentially as many colours as there are letters in the alphabet of M. That is why a formula that is restricted to one single monadic quantifier is not able to guess the polychrome colouring in step (1) but only a $\{0,1\}$-colouring. However, if we partition the picture into sufficiently large blocks, then it is possible to guess the polychrome colouring *of the border* of each block and store this in a $\{0,1\}$-colouring of the whole block. Then a finite disjunction may check if there really is a polychrome colouring of the block with that border such that Δ_1 and Δ_2 tile the inside of it. What remains to be done by the formula is to check whether the guessed $\{0,1\}$-colouring of the borders of neighboured blocks fit to each other in the sense that the (2×1) or (1×2)-subblocks (of "the" polychromely coloured picture) along the edges of blocks are in Δ_1 or Δ_2, respectively. This can be checked (in the actual $\{0,1\}$-coloured picture) by a first-order formula because the information about the polychrome colouring of the border of the block is coded in the $\{0,1\}$-colouring of its inside.

Some technicalities are necessary to choose the $\{0,1\}$-colouring in such a way that a first-order formula can determine the borders of blocks.

We prepare the proof with some definitions.

Definition 6.12 Let $P \in \Gamma^{+,+}$ be a picture of size (m,n). Then

$$\begin{aligned}
\mathit{left}(P) &= (P\langle 0,0\rangle \ldots P\langle m-1,0\rangle)^\top,\\
\mathit{right}(P) &= (P\langle 0,n-1\rangle \ldots P\langle m-1,n-1\rangle)^\top,\\
\mathit{top}(P) &= (P\langle 0,0\rangle \ldots P\langle 0,n-1\rangle),\\
\mathit{bottom}(P) &= (P\langle m-1,0\rangle \ldots P\langle m-1,n-1\rangle),\\
\mathit{border}(P) &= (\mathit{left}(P), \mathit{right}(P), \mathit{top}(P), \mathit{bottom}(P)).
\end{aligned}$$

Here, $.^\top$ denotes transposition. □

The next definition will help to define the partition of a picture into blocks.

Let $m \geq d \geq 1$. Choose $n \geq 0$ and $r < d$ such that $m = (n+1)d + r$. The tuple $(0, d, 2d, \ldots, nd, m)$ is called the *d-step sequence in m*. Note that if $(i_0, \ldots, i_{n+1})$ is the d-step sequence in m, then $d \leq i_{n+1} - i_n < 2d$.

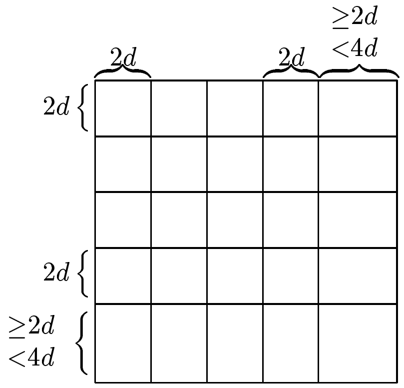


Figure 6.1: Partition into $2d$-blocks

Let P be a picture of size (m,n) (with $m,n \geq d$) over any alphabet. Let i,i' (and j,j', respectively) be consecutive components in the d-step sequence in m (in n, respectively). The *d-block of P at position* (i,j) is the subblock $\mathit{Block}_d(P,i,j) := P([i,i'-1]\times[j,j'-1]))$ of P. See Figure 6.1 for an illustration.

Definition 6.13 Let P, Q be non-empty pictures over Γ and Θ, respectively, such that P and Q have the same size. Let (Δ_1, Δ_2) be a domino tiling system over $\Gamma \times \Theta$.

Q is a *run on P wrt.* (Δ_1,Δ_2) iff (Δ_1,Δ_2) tiles $\widehat{P \otimes Q}$. □

It will be useful to reformulate the definition of a run in the presence of some partitioning of a picture into $2d$-blocks.

Remark 6.14 Let (Δ_1,Δ_2) be a domino tiling system over $\Omega = \Gamma \times \Theta$ for two alphabets Γ, Θ. Let P be a picture of size (m,n) over alphabet Γ with $m,n \geq 2d$. Let $i_{\max}$ and $j_{\max}$ be the last components but one of the $2d$-step sequences of m and n, respectively.

A picture $Q \in \Theta^{m,n}$ is a run on P iff

1. For all $i < m$ and $j < n$ in the $2d$-step sequence in m and n, respectively, both Δ_1 and Δ_2 tile $\mathit{Block}_{2d}(P \otimes Q, i, j)$.

2. For all $i, i' < m$ (and $j, j' < n$, respectively) in the $2d$-step sequence of m (and n, respectively) we have:

(a) if i, i' are consecutive in the $2d$-step sequence in m, then Δ_1 tiles $\mathit{bottom}(\mathit{Block}_{2d}(P \otimes Q), i, j) \ominus \mathit{top}(\mathit{Block}_{2d}(P \otimes Q), i', j)$;

(b) if j, j' are consecutive in the $2d$-step sequence in n, then Δ_2 tiles $\mathit{right}(\mathit{Block}_{2d}(P \otimes Q), i, j) \mathbin{\Phi} \mathit{left}(\mathit{Block}_{2d}(P \otimes Q), i, j')$;

(c) Δ_1 tiles $\# \ldots \# \ominus \mathit{top}(\mathit{Block}_{2d}(P \otimes Q), 0, j)$;

(d) Δ_1 tiles $\mathit{bottom}(\mathit{Block}_{2d}(P \otimes Q), i_{\max}, j) \ominus \# \ldots \#$;

(e) Δ_2 tiles $(\# \ldots \#)^\top \mathbin{\Phi} \mathit{left}(\mathit{Block}_{2d}(P \otimes Q), i, 0)$;

(f) Δ_2 tiles $\mathit{right}(\mathit{Block}_{2d}(P \otimes Q), i, j_{\max}) \mathbin{\Phi} (\# \ldots \#)^\top$.

□

As explained in the introduction of this section, a Σ_1-formula with only one quantifier cannot assert the existence of a run that easily because no quantification over all possible Θ-colourings is possible. However, if we have a partition into $2d$-blocks (as used in the above remark) then it suffices to quantify over all possible Θ-colourings *of borders* of blocks. This can be mimicked by quantifying over $\{0,1\}$-colourings that contain the "essential" information of each block.

The latter concept is put precisely in the following definition.

Definition 6.15 Let $d \geq 1$ an Ω be an alphabet. A mapping *ess* from $\bigcup\{\Omega^{i,j} \mid i,j \in [2d, 4d-1]\}$ into $\{0,1\}^{*,*}$ is called *border respecting compression of* $2d$*-blocks (over* Ω*)* iff for every B in the domain of *ess*,

- $\mathrm{size}(B) = \mathrm{size}(\mathit{ess}(B))$ and
- if $\mathit{ess}(B) = \mathit{ess}(B')$, then $\mathit{border}(B) = \mathit{border}(B')$.

□

Later we show that if d is sufficiently large, then there is a border respecting compression of $2d$-blocks.

In Lemma 6.17 we state that for a given border respecting compression, the existence of a run is equivalent to the existence of a $\{0,1\}$-colouring that is the block-wise image of a run. For this aim we need the following definition.

Definition 6.16 Let Γ, Ω, Θ, Δ_1, Δ_2, P, m, n, d, $i_{\max}$, and $j_{\max}$ be as in Remark 6.14. Let *ess* be a border respecting compression of $2d$-blocks over Ω.

A picture $\check{Q} \in \{0,1\}^{m,n}$ is a *run trace on* P *wrt.* (Δ_1, Δ_2) *and ess* iff

1. For all $i < m$ and $j < n$ in the $2d$-step sequence of m and n, respectively, there is a $Q \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, i, j))$ such that Δ_1 and Δ_2 tile $\mathit{Block}_{2d}(P, i, j) \otimes Q$.

2. For all $i, i' < m$ (and $j, j' < n$, respectively) in the $2d$-step sequence of m (and n, respectively), we have:

   (a) if i, i' are consecutive in the $2d$-step sequence of m, then there are $Q \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, i, j))$ and $Q' \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, i', j))$ such that Δ_1 tiles
   $$\mathit{bottom}(\mathit{Block}_{2d}(P, i, j) \otimes Q) \ominus \mathit{top}(\mathit{Block}_{2d}(P, i', j) \otimes Q');$$

   (b) if j, j' are consecutive in the $2d$-step sequence of n, then there are $Q \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, i, j))$ and $Q' \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, i, j'))$ such that Δ_2 tiles
   $$\mathit{right}(\mathit{Block}_{2d}(P, i, j) \otimes Q) \mathbin{\Phi} \mathit{left}(\mathit{Block}_{2d}(P, i, j') \otimes Q');$$

   (c) there is a $Q \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, 0, j))$ such that Δ_1 tiles $\# \ldots \# \ominus \mathit{top}(\mathit{Block}_{2d}(P, 0, j) \otimes Q)$;

   (d) there is a $Q \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, i_{\max}, j))$ such that Δ_1 tiles $\mathit{bottom}(\mathit{Block}_{2d}(P, i_{\max}, j) \otimes Q) \ominus \# \ldots \#$;

   (e) there is a $Q \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, i, 0))$ such that Δ_2 tiles $(\# \ldots \#)^\top \mathbin{\Phi} \mathit{left}(\mathit{Block}_{2d}(P, i, 0) \otimes Q)$;

   (f) there is a $Q \in \mathit{ess}^{-1}(\mathit{Block}_{2d}(\check{Q}, i, j_{\max}))$ such that Δ_2 tiles $\mathit{right}(\mathit{Block}_{2d}(P, i, j_{\max}) \otimes Q) \mathbin{\Phi} (\# \ldots \#)^\top$.

□

The next lemma states that the existence of a run trace is equivalent to the existence of a run. This is useful because the existence of a run trace can be asserted by a Σ_1-formula with only one quantifier (see Lemma 6.18).

Lemma 6.17 *Let* $P \in \Gamma^{m,n}$ *and* $m, n \geq 2d$*, and let ess be a border respecting compression of* $2d$*-blocks. There is a run on* P *iff there is a run trace on* P *wrt.* (Δ_1, Δ_2) *and ess.*

Proof Let Q be a run on P. Let $\check{Q}$ be the picture of size (m,n) over $\{0,1\}$ defined by

$$Block_{2d}(\check{Q},i,j) = ess(Block_{2d}(Q,i,j)) \tag{6.1}$$

for all $i < m$ and all $j < n$ in the $2d$-step sequences of m and n, respectively. Then it is straightforward to check that $\check{Q}$ is a run trace on P wrt. (Δ_1, Δ_2).

Conversely, let $\check{Q}$ be a run trace on P wrt. (Δ_1, Δ_2). By Property 1 of Definition 6.16, there is a $Q \in \Theta^{m,n}$ such that (6.1) holds for all $i < m$ and $j < n$ in the $2d$-step sequences of m and n, respectively. We have to show that $P \otimes Q$ fulfills Properties 2a to 2f of Remark 6.14.

For Property 2a, let $i, i' < m$ be consecutive components of the $2d$-step sequence of m, and let $j < n$ be in the $2d$-step-sequence of n. By Property 2a in Definition 6.16, there are $R, R' \in \Theta^{+,+}$ such that $ess(R) = Block_{2d}(\check{Q},i,j)$ and $ess(R') = Block_{2d}(\check{Q},i',j)$ and Δ_1 tiles

$$bottom(Block_{2d}(P,i,j) \otimes R) \ominus top(Block_{2d}(P,i',j) \otimes R').$$

Since $ess(Block_{2d}(Q,i,j)) = Block_{2d}(\check{Q},i,j) = ess(R)$, we have

$$border(Block_{2d}(Q,i,j)) = border(R).$$

In particular, $bottom(Block_{2d}(Q,i,j)) = bottom(R)$. Analogously, one shows $top(Block_{2d}(Q,i',j)) = top(R')$. Thus

$$\begin{aligned} & bottom(Block_{2d}(P \otimes Q),i,j) \ominus top(Block_{2d}(P \otimes Q),i',j) \\ = \; & bottom(Block_{2d}(P,i,j) \otimes R) \ominus top(Block_{2d}(P,i',j) \otimes R'), \end{aligned}$$

which is tiled by Δ_1 by choice of R and R'. This completes the proof of Property 2a of Remark 6.14. The proofs of Properties 2b to 2f are analogous, so by Remark 6.14, Q is a run on P. □

If $X \subseteq \mathrm{dom}P$ for a non-empty picture P, then the *characteristic picture of X* is the picture Q over $\{0,1\}$ of the same size with $Q\langle x\rangle = 1$ iff $x \in X$ for every $x \in \mathrm{dom}P$.

Lemma 6.18 *Let (Δ_1, Δ_2) be a domino tiling system over Ω. Let ess be a border respecting compression of $2d$-blocks with the following property: for all pictures P in the image of ess:*

1. $\forall i,j : P\langle i,j\rangle = P\langle i+1,j\rangle = 1 \Leftrightarrow j = 0$,
2. $\forall i,j : P\langle i,j\rangle = P\langle i,j+1\rangle = 1 \Leftrightarrow i = 0$.

There is a first-order formula run-trace(X) that asserts for a picture P with $\mathrm{size}(P) \geq (2d,2d)$ *and some $X \subseteq \mathrm{dom}P$ that the characteristic picture of X is a run trace on P wrt. (Δ_1,Δ_2) and ess.*

Proof By the assumption on *ess* there is a first-order formula *seq-cols*(X) that asserts for a picture of length $n \geq 2d$ and a subset X of its domain that (1) whenever a column contains two vertically neighboured positions in X, then the complete column is in X and (2) the column numbers of these columns that are subsets of X are exactly the components $< n$ in the $2d$-step-sequence in n.

A similar formula *seq-rows*(X) asserts the corresponding fact for rows, horizontal neighbours, and $m \geq 2d$.

We claim that there are first-order formulas *inner-tile*(X), *vert-tile*(X), *horiz-tile*(X), *top-tile*(X), *bottom-tile*(X), *left-tile*(X), and *right-tile*(X), respectively, that assert that the characteristic picture of X fulfills Conditions 1, 2a to 2f, respectively, of Defintion 6.16, provided that X fulfills *seq-cols*$(X) \wedge$ *seq-rows*(X).

Then we are finished because the formula *run-trace*(X) = *seq-cols*$(X) \wedge$ *seq-rows*$(X) \wedge$ *inner-tile*$(X) \wedge$ *vert-tile*$(X) \wedge$ *horiz-tile*$(X) \wedge$ *top-tile*$(X) \wedge$ *bottom-tile*$(X) \wedge$ *left-tile*$(X) \wedge$ *right-tile*(X) asserts the desired property by definition of run traces.

We need some more auxiliary formulas. There is a first-order formula *tl-corner*(X,x) (or *br-corner*(X,x), respectively) that asserts for a position $x = (i,j)$ in a picture of size (m,n), where $m,n \geq 2d$, that i (or $i+1$, respectively) is in the $2d$-step-sequence of m, and that j (or $j+1$, respectively) is in the $2d$-step-sequence of n, provided that X fulfills the formula *seq-cols*$(X) \wedge$ *seq-rows*(X).

Choose a first-order *vert-dist*$_k(x,y)$ that asserts that $row(x)+k = row(y)$, i.e., that x and y are in the same column with vertical distance k.

Besides, for every $B \in \{0,1\}^{p,q}$ with $p,q \in [2d, 4d-1]$, there is a first-order formula X-*block*$_B(X,x,y)$ that asserts that the subblock of the characteristic picture of X that has top-left corner x and bottom-right corner y equals B.

Similarly, for every $A \in \Gamma^{p,q}$ with $p,q \in [2d,4d-1]$, there is a formula P-*block*$_A(x,y)$ that asserts that the subblock of P with top-left corner x and bottom-right corner y equals A.

Let

$$\begin{aligned} \textit{inner-tile}(X) = \; & \forall x \Big(\textit{tl-corner}(X,x) \rightarrow \exists y \Big(\textit{br-corner}(X,y) \wedge \\ & \textstyle\bigvee_{p,q} \bigvee_{B,A} (\textit{X-block}_B(X,x,y) \wedge \textit{P-block}_A(x,y)) \Big) \Big), \end{aligned}$$

where the disjunctions range over all $p,q \in [2d, 4d-1]$ and over all $B \in \{0,1\}^{p,q}$, $A \in \Gamma^{p,q}$ for which there is a $Q \in \textit{ess}^{-1}(B)$ such that Δ_1 and Δ_2 tile $A \otimes Q$. Let

$$\begin{aligned} & \textit{vert-tile}(X) = \\ & \quad \forall x \forall x' \Big(\textit{tl-corner}(X,x) \wedge \textit{tl-corner}(X,x') \wedge \textit{vert-dist}_{2d}(x,x') \rightarrow \\ & \quad\quad \exists y \exists y' \big(\textit{br-corner}(y) \wedge \textit{br-corner}(y') \wedge \\ & \quad\quad\quad \textstyle\bigvee_{p,p',q} \bigvee_{B,B',A,A'} \; (\textit{X-block}_B(X,x,y) \wedge \textit{X-block}_{B'}(X,x',y') \wedge \\ & \quad\quad\quad\quad \textit{P-block}_A(x,y) \wedge \textit{P-block}_{A'}(x',y'))) \Big), \end{aligned}$$

where the disjunctions range over all $p,p',q \in [2d, 4d-1]$ and all $B \in \{0,1\}^{p,q}$, $B' \in \{0,1\}^{p',q}$, $A \in \Gamma^{p,q}$, and $A \in \Gamma^{p',q}$ such that there are $Q \in \textit{ess}^{-1}(B)$ and $Q' \in \textit{ess}^{-1}(B')$ such that Δ_1 tiles $\textit{bottom}(A \otimes Q) \ominus \textit{top}(A' \otimes Q')$.

The formula $\textit{horiz-tile}(X)$ can be defined very similarly.

$$\begin{aligned} \textit{top-tile}(X) = \; & \forall x \Big(\textit{top}(x) \wedge \textit{tl-corner}(X,x) \rightarrow \\ & \exists y \big(\textit{br-corner}(X,y) \wedge \\ & \textstyle\bigvee_{p,q} \bigvee_{B,A} (\textit{X-block}_B(X,x,y) \wedge \textit{P-block}_A(x,y))) \Big), \end{aligned}$$

where the disjunctions range over all $p,q \in [2d, 4d-1]$ and all $B \in \{0,1\}^{p,q}$ and all $A \in \Gamma^{p,q}$ such that there is $Q \in \textit{ess}^{-1}(B)$ such that Δ_1 tiles $\textit{top}(A \otimes Q) \ominus (\#)^p$.

The formulas $\textit{bottom-tile}(X)$, $\textit{left-tile}(X)$, and $\textit{right-tile}(X)$ can be defined similarly. This completes the proof. $\square$

Now we are ready to prove Theorem 6.10, which states that every projection of a domino-local picture language is defined by some formula of the form $\exists X \varphi(X)$ with first-order $\varphi(X)$.

Proof (of Theorem 6.10) There is a domino-local picture language M over alphabet $\Omega = \Gamma \times \Theta$ such that L is over alphabet Γ, and L is the image of M under the alphabet projection $\pi : \Omega \rightarrow \Gamma$, $(a,b) \mapsto a$. Let (Δ_1, Δ_2) be a domino tiling system with $\mathcal{L}((\Delta_1, \Delta_2)) = M$.

Choose d sufficiently large such that there exists an injective mapping

$$f : \bigcup_{m,n \in [2d, 4d-1]} (\Omega^{m,1} \times \Omega^{m,1} \times \Omega^{1,n} \times \Omega^{1,n}) \rightarrow \{0,1\}^{d \times d}.$$

(This is possible because the cardinality of the left side is $2^{\mathcal{O}(d)}$, whereas the cardinality of the right side is $2^{\mathcal{O}(d^2)}$.)

For a block B over Ω whose width and height are $\geq 2d$ and $< 4d$, let $\textit{ess}(B)$ be the picture of same size as B such that

$$\textit{ess}(B)(i,j) = \begin{cases} 1 & \text{if } i = 0 \vee j = 0, \\ f(\textit{border}(B))(\frac{1}{2}(i+1), \frac{1}{2}(j+1)) & \text{if } i,j < 2d \text{ are odd}, \\ 0 & \text{else}. \end{cases}$$

Then $\textit{ess}$ is a border respecting compression of $2d$-blocks and has the additional property of Lemma 6.18.

By Lemma 6.18 there is a formula $\textit{run-trace}(X)$ such that for every picture P over Γ with $\underline{\overline{P}}, |P| \geq 2d$ we have $P \in L$ iff $P \models \exists X \, \textit{run-trace}(X)$. Let ψ be the conjunction of $\textit{run-trace}(X)$ with a first-order sentence that asserts for a picture P that $\underline{\overline{P}}, |P| \geq 2d$.

By Theorem 6.11 there is, for every $m < 2d$, a formula $\varrho_m(X)$ such that $\textit{Mod}(\exists X \varphi(X))$ is the set of pictures in L that have height m. By duality, there is, for every $m < 2d$, a formula $\gamma_m(X)$ such that $\textit{Mod}(\exists X \varphi(X))$ is the set of pictures in L that have width m.

Choose $\varphi(X) = \psi(X) \vee \bigvee_{m<2d} (\varrho_m(X) \vee \gamma_m(X))$. Then φ is first-order, and we have $L = \textit{Mod}(\exists X \psi \vee \bigvee_{m<2d} (\exists X \varrho_m \vee \exists X \gamma_m)) = \textit{Mod}(\exists X \varphi)$. This completes the proof of Theorem 6.10. $\square$

As mentioned above, the main result of this Chapter, Theorem 6.7, which states that every Σ_1-sentence over pictures is equivalent to a sentence of the form $\exists X \varphi(X)$ where φ is first-order, is an immediate consequence of Theorems 6.9 and 6.10.

6.2.4 Compression in the Presence of Free Variables

We wish to show that Theorem 6.7 remains true if "sentence" is replaced by "formula", i.e., that the "compression" of an existential quantifier block also works for formulas with free variables. At the same time we will show that the two different types of signatures we used for "coloured" grids, namely $\tau_{\textit{Grids}_t} = \{S_1, S_2, X_1, \ldots, X_t\}$ and $\sigma_\Gamma = \{S_1, S_2, (Q_a)_{a \in \Gamma}\}$, do not differ wrt. to the induced notion of first-order definability of picture languages.

Let Γ be a fixed alphabet.

Definition 6.19 Let $\overline{X}$ and $\overline{Y}$ be the variable tuples $(X_1, \ldots, X_t)$ and $(Y_1, \ldots, Y_r)$, respectively. A formula $\varphi(\overline{X}, \overline{Y})$ over σ_Γ is called $\overline{Y}$*-equivalent* to a formula $\varphi'(\overline{Y})$ over $\sigma_{\Gamma \times \{0,1\}^t}$ iff for every $P \in \Gamma^{+,+}$ and all tuples $\overline{U} = (U_1, \ldots, U_t)$ and $\overline{V} = (V_1, \ldots, V_r)$ of subsets of dom$P$,

$$\underline{P} \models \varphi[\overline{U}, \overline{V}] \quad \text{iff} \quad \underline{P \otimes Q} \models \varphi'[\overline{V}]$$

for the picture Q over alphabet $\{0,1\}^t$ given by

$$\forall x \in \mathrm{dom}P\ \forall s \leq t: \quad Q\langle x\rangle(i) = \begin{cases} 1 & \text{if } x \in U_s, \\ 0 & \text{else.} \end{cases}$$ □

Lemma 6.20 *Let $\overline{X}$, $\overline{Y}$ as above. For every first-order formula $\varphi(\overline{X}, \overline{Y})$ over σ_Γ there is a $\overline{Y}$-equivalent first-order formula $\varphi'(\overline{Y})$ over $\sigma_{\Gamma \times \{0,1\}^t}$ and vice versa.*

Proof We may assume that $X_1, \ldots, X_t$ are not bound in φ. In order to transform φ into φ', one may replace every atomic formula $Q_a(x)$ by $\bigvee_{b \in \{0,1\}^t} Q_{(a,b)}(x)$ and every atomic formula $X_i(x)$ (where $i \leq t$) by $\bigvee_{(a,b)} Q_{(a,b)}(x)$, where the disjunction ranges over those $(a, b) \in \Gamma \times \{0,1\}^t$ for which $b(i) = 1$.

For the converse direction, replace every atomic formula $Q_{(a,b)}(x)$ (with $(a, b) \in \Gamma \times \{0,1\}^t$) by $Q_a(x) \wedge \bigwedge_{i \in b^{-1}(1)} X_i(x) \wedge \bigwedge_{i \in b^{-1}(0)} \neg X_i(x)$. Here, $b^{-1}(j)$ is the set of all i with $b(i) = j$. □

Corollary 6.21 *Every formula $\exists \overline{Y} \varphi$ over signature σ_Γ, where $\overline{Y}$ is some variable tuple and φ is first-order, is equivalent to a formula $\exists Y \psi$ for some first-order formula ψ over σ_Γ.*

Proof Let $\overline{X} = (X_1, \ldots, X_t)$ be the tuple of free variables of $\exists \overline{Y} \varphi$. By Lemma 6.20, $\varphi(\overline{X}, \overline{Y})$ is $\overline{Y}$-equivalent to a first-order formula $\varphi'(\overline{Y})$ over $\sigma_{\Gamma \times \{0,1\}^t}$. By Theorem 6.7, the sentence $\exists \overline{Y} \varphi'(\overline{Y})$ is equivalent to some sentence $\exists Y \psi'(Y)$ for a first-order formula ψ' over $\sigma_{\Gamma \times \{0,1\}^t}$. Again using Lemma 6.20 we obtain a first-order σ_Γ-formula $\psi(\overline{X}, Y)$ that is Y-equivalent to ψ'.

Then $\exists Y \psi$ is equivalent to $\exists \overline{Y} \varphi$. □

Remark 6.22 In the same spirit as Definition 6.19 we may define when a formula $\varphi(X_1, \ldots, X_t)$ over signature τ_{Grids} (the signature used in all the previous chapters) is *equivalent* to a sentence φ' in signature $\sigma_{\{0,1\}^t}$. The mutual translation of first-order formulas as in Lemma 6.20 works then, too. □

6.2.5 Combining the Results

We may combine the results of the previous chapters with the compression result of this chapter and get the following theorem. It states that in order to build a Σ_{k+1}-sentence that exceeds Σ_k in expressiveness, we need only two quantifications over sets: one in the outermost quantifier block and one in the innermost one.

Theorem 6.23 *Let $k \geq 1$. There is a sentence of the form*

$$\exists X \neg \exists \overline{x}_1 \neg \ldots \neg \exists \overline{x}_k \exists Y \psi$$

(where ψ is first-order and $\overline{x}_1, \ldots, \overline{x}_k$ are first-order variable tuples) over the signature τ_{Grids} that is not equivalent over grids to any sentence in Σ_k.

Proof By Corollary 2.25 there is a formula $\varphi(X_1) \in \Pi^0_k(\Delta_1)$ in signature τ_{Grids} that is not equivalent over grids to any formula in Σ_k. By Remark 2.3, φ can be written in the form

$$\neg \exists \overline{x}_1 \neg \exists \overline{x}_2 \ldots \neg \exists \overline{x}_k \psi',$$

where $\psi' \in \Sigma_1$. By Corollary 6.21, this formula ψ' is equivalent (over grids) to a formula of the form $\exists Y \psi$ for some first-order ψ.

Thus the τ_{Grids}-sentence

$$\exists X_1 \neg \exists \overline{x}_1 \neg \exists \overline{x}_2 \ldots \neg \exists \overline{x}_k \exists Y \psi$$

is not equivalent to any sentence in Σ_k. □

Chapter 7

Conclusion

As the main result of this thesis we have proved Fact 1.4, i.e., the strictness of the monadic second-order quantifier alternation hierarchy. It turned out that this separation result could be strengthened essentially: "Very little" of monadic quantification is necessary to exceed a particular level of the monadic hierarchy in expressiveness in the sense that the alternation of first-order quantifiers or—in the context of pictures—even the alternation of the column concatenation and negation gives expressive power that cannot be captured by a smaller number of alternations of set quantifiers.

7.1 Monadic vs. Polynomial

Since part of the interest in the monadic second-order alternation hierarchy stems from the desire to achieve progress in complexity theoretical questions, we should review our proofs with regard to whether they provide methods for attacking at least other fragments of the second-order hierarchy.

As a striking difference to most of the other approaches to separations in monadic second-order logic, our non-definability argument (the "lower bound proof", Chapter 4) does not use Ehrenfeucht-Fraïssé games but automata-theoretic arguments. The ultimate core is the standard pumping lemma for finite automata. There is little or no hope that such techniques will carry over to, say, the binary fragment of second-order logic because a run of an automaton is a *colouring* of the input and thus inherently something monadic.

The reason why these automata-theoretic techniques work is that the strictness of the monadic hierarchy is less related to the strictness of the full second-order hierarchy than Fact 1.1 suggests. As pointed out in [AFS98], the levels of the monadic hierarchy are not sufficiently robust. In some sense our result that the dot-depth hierarchy cuts through the monadic hierarchy stresses this fact, and so does the observation that our witness sets for the strictness of the monadic hierarchy over pictures are all in FO^{TC} and thus in NLOG, i.e. very simple in complexity theoretic terms. (A closer investigation shows that they are also in DLOG.)

Comparing two Hierarchies

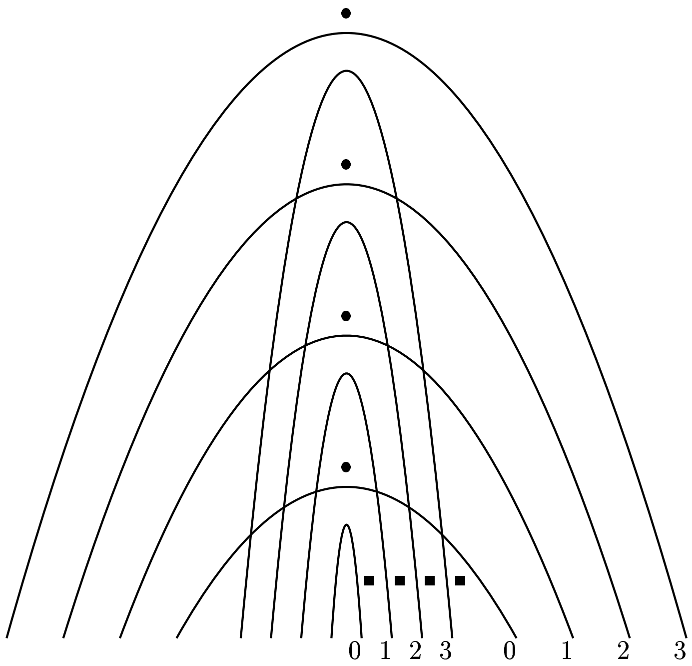
Our proof for the strictness of the monadic alternation hierarchy shows that the monadic alternation hierarchy and the polynomial hierarchy can be visualized like in Figure 7.1, provided that the polynomial hierarchy is strict. The •-witnesses, which show that there are monadic properties in level $k+1$ that are not in level k of the polynomial hierarchy, are provided by Fact 1.1 (even for the class of pictures in the version of [Sch97a]). The ■-witnesses, which are all in level 0 of the polynomial hierarchy but beyond some level of the monadic hierarchy, are provided by Corollary 2.25.

7.2 Formal Languages

The investigation of our problems motivated from finite model theory led to some new insights into problems that are located rather in the theory of formal languages. The investigation of two-dimensional languages (called "picture languages") leads, in my opinion, to the following conclusion: there is no neat "theory of regular picture languages": whenever one considers two concepts which are equivalent in the world of word languages, then there are several "straightforward" adaptations to pictures, and these are not equivalent unless for trivial reasons.

A similar picture as Figure 7.1 can be drawn for the dot-depth hierarchy as the "inner" hierarchy and the monadic alternation hierarchy over pictures as the "outer" one.

The •-witnesses are provided by our picture languages which are of dot-depth $k+1$ but not in level k of the monadic hierarchy. (The numeration of the "outer" monadic hierarchy in the figure has to be shifted by one because we have only $\mathit{dot\text{-}depth}_k \subseteq \underline{B(\Sigma_k)} \subseteq \underline{\Sigma_{k+1}}$.)



The outer curves visualize the polynomial hierarchy, whereas the inner ones represent the monadic hierarchy.
Provided that the polynomial hierarchy is strict, the monadic hierarchy and the polynomial hierarchy are as unrelated as possible in the sense that there are properties (indicated by •) that are hard for the levels of both hierarchies, and other properties (indicated by ▪) that are hard for the levels of the monadic hierarchy but in level 0 of the polynomial hierarchy (i.e., in P, they are even in NLOG).

Figure 7.1: Two Very Unrelated Hierarchies

The well-known witnesses for the strictness of the dot-depth hierarchy are word languages, i.e. sets of pictures of height 1, and thus all on level 1 of the alternation hierarchy because the alternation hierarchy over words collapses. In this figure, these are the ▪-witnesses, provided the lowest curve of the monadic hierarchy represents level 1 rather than zero.

The result that the class of starfree picture languages is not captured within any level of the monadic hierarchy stresses the above conclusion once again: the class of starfree word languages is such a small subclass of recognizable word languages, and yet starfreeness of picture languages is not captured by recognizability and not even by any higher level of the monadic hierarchy.

7.3 Some Open Questions

I conclude with some open questions, a few of which I have tried to solve myself.

Question 7.1 *Is the hierarchy* $\underline{\Sigma_0(FO^{\mathrm{TC}})} \subseteq \underline{\Sigma_1(FO^{\mathrm{TC}})} \subseteq \dots$ *over graphs strict?*

Instead of adding the unary TC-operator to the first-order kernel, as above, one might add the two partial orderings to the signature of *grids* and ask whether the hierarchy $\underline{\Sigma_0(FO^{\leq_1,\leq_2})} \subseteq \underline{\Sigma_1(FO^{\leq_1,\leq_2})} \subseteq \dots$ is strict.

Question 7.2 *Has the binary fragment of second-order logic a strict alternation hierarchy?*

Without complexity theoretic assumptions it is not even known whether the lowest level (often called "binary NP") is closed under complement nor whether it equals NP.

For preliminary study I suggest investigating the hierarchy of the fragment of second-order logic where quantification is allowed over functions or—more restrictively—permutations of the universe. The existential fragments of some of such fragments have been compared in [DLS98].

Question 7.3 *Is the alternation of the "permutation fragment" of second-order logic strict?*

I agree with [AFS98] that it is promising to study the following question.

Question 7.4 *Is the closed hierarchy strict?*

Recall that level $k+1$ of the closed hierarchy results from the boolean closure of level k by closing under existential second-order quantification as well as under (existential and universal) first-order quantification.

As a preliminary step one might consider the following hierarchy: $\Sigma_0 = FO$, and $\Sigma_{k+1} = \Sigma_1(FO(\Sigma_{k+1}))$, i.e., when passing from level k to $k+1$ one performs the closures one after the other rather than at the same time.

This hierarchy seems easier than the closed one because its levels are more similar to the first-order closures of the levels of the plain monadic hierarchy, so some hope is justified that similar arguments as in Chapter 4 will work.

All of the above hierarchies are infinite unless the polynomial hierarchy collapses.

Concerning the results of Chapter 6, there are many open questions.

Question 7.5 *For which classes of finite structures do we have the analogue of Fact 6.4, i.e., that every existential monadic second-order formula is equivalent to one that uses only one quantifier?*

We know that the class of pictures is an example where this property holds. There is hope that a general answer to this question will also give a more elegant proof of Fact 6.4, avoiding technical details that are specific for pictures.

Question 7.6 *Can the length of each quantifier block in a Σ_k-formula over grids be limited to one without loss of expressiveness?*

We know that such a limitation is possible for the innermost block, but it would be nice if this carried over to all blocks.

On the other hand, I conjecture that Otto's result Fact 6.3 carries over to the higher levels of the alternation hierarchy, i.e., that over the class of graphs, the length of the outermost quantifier block of a Σ_k-formula cannot be limited without loss of expressiveness.

Question 7.7 *Is there a class of structures for which the monadic quantifier depth hierarchy is strict but the alternation hierarchy is not?*

One candidate for such a class is the one considered in [Ott95].

Acknowledgements

First of all, I would like to thank my supervisor Prof. Wolfgang Thomas for his guidance and continuous encouragements in the whole research and critical review of the manuscript.

Grateful thanks go to Dr. Thomas Wilke, who gave valuable advice and patiently listened to my (often quite odd) explanations of (often quite wrong) preliminary results.

I owe a lot of gratitude to Nicole Schweikardt, who checked my parts in our joint paper with much more care than vice versa. Her corrections certainly improved also this thesis.

I thank my father for his tedious search for linguistic mistakes.

Bibliography

[AFS97] M. Ajtai, R. Fagin, and L.J. Stockmeyer. The closure of monadic NP. Research Report RJ 10092, IBM, Dec 1997.

[AFS98] M. Ajtai, R. Fagin, and L.J. Stockmeyer. The closure of monadic NP. In *The Thirtieth Annual ACM Symposium on Theory of Computing*, pages 309–318. SIGACT, 1998.

[Bir96] J.-C. Birget. The state complexity of $\overline{\Sigma^* \overline{L}}$. *Information Processing Letters*, 58:185–188, 1996.

[BK71] J.A. Brzozowski and R. Knast. The dot-depth hierarchy of star-free languages is infinite. *Journal of Computer and System Sciences*, 5:1–16, 1971.

[Bra42] A. Brauer. On a problem of partitions. *Amer. J. Math.*, 64:299–312, 1942.

[Bra96] J. Bradfield. The modal mu-calculus alternation hierarchy is strict. In U. Montanari and V. Sassone, editors, *CONCUR '96: Concurrency Theory. Proceedings*, volume 1119 of *Lecture Notes in Computer Science*, pages 233–246. Springer Verlag, 1996.

[Chr86] M. Chrobak. Finite automata and unary languages. *Theoretical Computer Science*, 47:149–158, 1986.

[Cos93] S.S. Cosmadakis. Logical reducibility and monadic NP. In *Proceedings of 34th Symposium on Foundations of Computer Science*, pages 52–61, 1993.

[DLS98] Arnaud Durand, Clemens Lautemann, and Thomas Schwentick. Subclasses of binary NP. *Journal of Logic and Computation*, 8:189–207, 1998.

[EF95] H.D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer-Verlag, New York, 1995.

[Fag74] R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. In R.M. Karp, editor, *Complexity of Computation*, volume 7 of *SIAM-Proceedings*, pages 43–73. AMS, 1974.

[Fag75] R. Fagin. Monadic generalized spectra. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:89–96, 1975.

[FSV95] R. Fagin, L.J. Stockmeyer, and M.Y. Vardi. On monadic NP vs. monadic co-NP. *Information and Computation*, 120:78–92, 1995.

[Gia94] D. Giammarresi. Two-dimensional languages and recognizable functions. In G. Rozenberg and A. Salomaa, editors, *Developments in Language Theory, Proceedings of the conference, Turku (Finnland) '93*, pages 290–301. world scientific, Singapore, 1994.

[GR92] D. Giammarresi and A. Restivo. Recognizable picture languages. In *Proceedings First International Colloqium on Parallel Image Processing 1991*, number 6 in International Journal Pattern Recognition and Artificial Intelligence, pages 241–256, 1992.

[GR96] D. Giammarresi and A. Restivo. Two-dimensional languages. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Language Theory*, volume III, pages 125–267. Springer-Verlag, New York, 1996.

[GRST96] D. Giammarresi, A. Restivo, S. Seibert, and W. Thomas. Monadic second-order logic and recognizability by tiling systems. *Information and Computation*, 125:32–45, 1996.

[GS96] I. Glaister and J. Shallit. A lower bound technique for the size of nondeterministic finite automata. *Information Processing Letters*, 125:32–45, 1996.

[Hog93] W. Hogdes. *Model theory*, volume 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.

[Lan09] E. Landau. *Handbuch über die Verteilung der Primzahlen*, pages 222–229. Teubner, Leipzig, 1909.

[Liu64] U.I. Liubicz. Bounds for the optimal determinization of nondeterministic autonomic automata. *Sibirskii Mathemat. Journal*, pages 337–355, 1964. (in Russian).

[LS94] M. Latteux and D. Simplot. Recognizable picture languages and domino tiling. Internal Report IT-94-264, Laboratoire d'Informatique Fondamentale de Lille, Université de Lille, France, 1994.

[Mat95] O. Matz. *Klassifizierung von Bildsprachen mit rationalen Ausdrücken, Grammatiken und Logik-Formeln.* Diploma thesis, Chrstian-Albrechts-Universität Kiel, 1995. (in German).

[Mat98a] O. Matz. First-order closure and the monadic second-order alternation hierarchy. Technical Report 9807, Institut für Informatik, Christian-Albrechts-Universität Kiel, 1998.

[Mat98b] O. Matz. On piecewise testable, starfree, and recognizable picture languages. In Maurice Nivat, editor, *Foundations of Software Science and Computation Structures*, volume 1378 of *Lecture Notes in Computer Science*, pages 203–210. Springer, 1998.

[Mat98c] O. Matz. One quantifier will do in existential monadic second-order logic over pictures. In Luboš Brim, Jozef Gruska, and Jiří Zlatuška, editors, *Mathematical Foundations of Computer Science*, volume 1450 of *Lecture Notes in Computer Science*, pages 751–759. Springer, 1998.

[MP94] J.A. Makowsky and Y.B. Pnueli. Arity vs. alternation in second order logic. In A. Nerode and Yu.V. Matiyasevich, editors, *Logical Foundations of Computer Science*, volume 813 of *Lecture Notes in Computer Science*, pages 240–252. Springer Verlag, Berlin, 1994.

[MP96] J.A. Makowsky and Y.B. Pnueli. Arity and alternation in second-order logic. *Annals of Pure and Applied Logic*, 78(1-3):189–202, 1996.

[MST99] O. Matz, N. Schweikardt, and W. Thomas. The monadic quantifier alternation hierarchy over grids and graphs. *Information and Computation, LICS '97 special issue*, 1999. to appear.

[MT97] O. Matz and W. Thomas. The monadic quantifier alternation hierarchy over graphs is infinite. In *Twelfth Annual IEEE Symposium on Logic in Computer Science*, pages 236–244, Warsaw, Poland, 1997. IEEE.

[Ott95] M. Otto. Note on the number of monadic quantifiers in monadic Σ_1^1. *Information Processing Letters*, 53:337–339, 1995.

[Pin86] J.-E. Pin. *Varieties of Formal Languages.* Oxford University Press, New York, 1986.

[Rab77] M. Rabin. *Handbook of Mathematical Logic*, volume 90 of *Studies in Logic and the Foundations of Mathematics*, chapter C.3. Elsevier, North-Holland, 1977.

[Sch97a] N. Schweikardt. The monadic quantifier alternation hierarchy over grids and pictures. In Mogens Nielson and Wolfgang Thomas, editors, *Computer Science Logic*, volume 1414 of *Lecture Notes in Computer Science*, pages 441–460. Springer, 1997.

[Sch97b] N. Schweikardt. The monadic second-order quantifier alternation hierarchy over grids and pictures. diploma thesis, Johannes Gutenberg-Universität Mainz, 1997.

[Sto74] L.J. Stockmeyer. *The complexity of decision problems in automata theory and logic.* PhD thesis, Report MAC TR-133, Project MAC, Massachusetts Institute of Technology, Department of Electrical Engineering, Cambridge, Massachusetts, 1974.

[Sza80] M. Szalay. On the maximal order in s_n and s_n^*. *Acta Arithm.*, 37:321–331, 1980.

[Tho82] W. Thomas. Classifying regular events in symbolic logic. *Journal of Computer and System Sciences*, 25:360–376, 1982.

[Wil97] Th. Wilke. Star-free picture expressions are strictly weaker than first-order logic. In Pierpaolo Degano, Roberto Gorrieri, and Alberto Marchetti-Spaccamela, editors, *Automata, Languages and Programming*, volume 1256, pages 347–357. Springer, Bologna, Italy, 1997.

Symbol Index

Index

Lebenslauf

26.03.70:	Geboren in Neumünster.
03.06.89:	Abitur am Immanuel-Kant-Gymnasium Neumünster.
07/89 – 09/90:	Zivildienst beim Malteser Hilfsdienst Neumünster.
10/90 – 10/92:	Grundstudium der Informatik mit Nebenfach Mathematik an der Christian-Albrechts-Universität Kiel.
13.10.92:	Vordiplom.
10/92 – 03/96:	Hauptstudium an der CAU Kiel.
29.03.96:	Diplom der Informatik.
05/96 – 08/98:	Wissenschaflicher Mitarbeiter im Institut für Informatik an der CAU Kiel.
08/98 – 08/99:	Wissenschaflicher Mitarbeiter in der Fachgruppe Informatik der Rheinisch-Westfälischen Technischen Hochschule Aachen.
19.03.99:	Einreichung des Promotionsgesuchs an der RWTH Aachen.
14.06.99:	Promotionsprüfung an der RWTH Aachen.