Formal Language Theory of Logic Fragments

Von der Fakultät für Informatik, Elektrotechnik und Informationstechnik der Universität Stuttgart zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.) genehmigte Abhandlung

Vorgelegt von

Alexander Lauser

aus Kirchheim unter Teck

Hauptberichter: Prof. Dr. Volker Diekert

Mitberichter: Prof. Dr. Christian Glaßer

Tag der mündlichen Prüfung: 28. August 2014

Institut für Formale Methoden der Informatik

2014

Contents

List of Symbols

$:=$	Syntactic equality of formulae; *p. 42*
$\emptyset$	Empty set
$\ll$, $\gg$	Stricter orders; $i \ll j$ if and only if $j \gg i$ if and only if $i \leqslant j - 2$; *p. 31*
$[u]$	Encoding of factors over $\Lambda \cup \{\square\}$ into a single label; *p. 84*
$\square$	New blank symbol; $\square \notin \Lambda$; *p. 84*
A, B	Typical names for alphabets $A, B \subseteq \Lambda$ (usually finite); *pp. 31, 41*
A^*	Set of finite words over the alphabet A; *p. 31*
A^+	Set of non-empty finite words over the alphabet A; *p. 31*
$\mathbf{A}$	Variety of aperiodic monoids satisfying $x^{\omega+1} = x^{\omega}$; *p. 37*
$X \to Y$	Declares a total function from X into Y
$X \to_p Y$	Declares a *partial* function from X into Y
$(a, J)^{-1}\varphi$	Left residual construction for φ by the extended letter (a, J); *p. 59*
$\mathrm{alph}(u)$	Set of letters occurring in u; *p. 31*
$\mathrm{alph}_k(u)$	Set of factors of length k occurring in u; *p. 31*
$A_{(N)}$	Sliding window alphabet of radius N; *p. 84*
$\mathbb{B}\Sigma_m$	Boolean closure of Σ_m-formulae; *p. 49*
$\mathbb{B}\Sigma^2_m$	Boolean closure of Σ^2_m-formulae; *p. 123*
$\mathbb{B}\Sigma^+_1$	Boolean closure of purely existential formulae without any negation; *p. 81*
$\mathcal{C}_{all}$	Category of all homomorphisms; *p. 54*
$\mathcal{C}_{lm}$	Category of length-multiplying homomorphisms; *p. 54*
$\mathcal{C}_{lp}$	Category of length-preserving homomorphisms; *p. 54*
$\mathcal{C}_{lr}$	Category of length-reducing homomorphisms; *p. 54*
$\mathcal{C}_{ne}$	Category of non-erasing homomorphisms; *p. 54*
$\mathbf{D}$	Variety of definite semigroups satisfying $yx^{\omega} = x^{\omega}$; *p. 37*
$\mathbf{DA}$	Variety of monoids satisfying $(xy)^{\omega} = (xy)^{\omega}x(xy)^{\omega}$; *p. 37*

δ	A typical name for an offset function used in various constructions
$\delta[x/d]$	Derived offset function, mapping x to d and $y \neq x$ to $\delta(y)$
$E(S)$	Set of all idempotents of S; *p. 33*
empty	Constant formula that is true only on the empty word; *p. 41*
ε	Empty word; *p. 31*
$\exists x\ \varphi$	Existential first-order quantifier; *p. 42*
$\exists x \in \mathcal{C} \colon \varphi$	Generalized existential first-order quantifier with constraint $\mathcal{C}$; *p. 87*
$\forall x\ \varphi$	Universal first-order quantifier; *p. 42*
$\forall x \in \mathcal{C} \colon \varphi$	Generalized universal first-order quantifier with constraint $\mathcal{C}$; *p. 87*
$\exists X\ \varphi$	Existential second-order quantifier; *p. 42*
$\forall X\ \varphi$	Universal second-order quantifier; *p. 42*
$\exists^{r \bmod q} x\ \varphi$	Modular counting quantifier; true if r (modulo q) positions satisfy φ; *p. 42*
$\exists^{r \bmod q} x \in \mathcal{C} \colon \varphi$	Generalized modular counting quantifier first-order quantifier; *p. 87*
$\exists! x\ \varphi$	Uniqueness quantification; true if precisely one position satisfies φ; *p. 42*
$\mathcal{F}$	Typical name for a logic prefragment or logic fragment; *p. 45*
$\mathcal{F}[\mathcal{N}]$	Restriction to formulae that only use numerical predicates in $\mathcal{N}$; *p. 47*
$\widehat{\mathcal{F}}$	Distance-stable extension of the fragment $\mathcal{F}$; *p. 87*
$\mathfrak{F}$	Typical name for a collection of appropriate fragments
$\varphi \leqslant_{\mathcal{F}} \psi$	Syntactic preorder of fragment $\mathcal{F}$; substituting φ for ψ respects $\mathcal{F}$; *p. 46*
FO	First-order logic without modular counting quantifiers; *p. 47*
FO_m	Full level m of the FO alternation hierarchy; next half level is Σ_{m+1}; *p. 48*
FO^2_m	Full level m of the FO^2-alternation hierarchy; next half level is Σ^2_{m+1}; *p. 49*
$\mathrm{FO}_{m,n}$	Fragment of FO_m with quantifier depth at most n; *p. 48*
$\mathrm{FO}^2_{m,n}$	Fragment of FO^2_m with quantifier depth at most n; *p. 49*
FO+Mod	First-order logic with modular counting quantifiers; *p. 47*
$\mathrm{FV}(\varphi)$	Free variables of the formula φ; *p. 42*
$h_\delta(u)$	Homomorphism respecting the free-variable offset function δ; *p. 69*
$h^{-1}(L)$	Inverse image of L under h; *p. 32*

h_L	Syntactic homomorphism of the language L; *p. 35*
$h^{-1}(\varphi)$	Inverse image construction for φ under homomorphism h; *p. 68*
$\binom{I}{r}$	Set of all subsets of I with r elements; *p. 91*
$[i; j]$	Closed interval of integers $\{i, \ldots, j\}$; *p. 31*
$(i; j]$, $[i; j)$	Half-open intervals of integers $\{i + 1, \ldots, j\}$ and $\{i, \ldots, j - 1\}$; *p. 31*
$(i; j)$	Open interval of integers $\{i + 1, \ldots, j - 1\}$; *p. 31*
$x \mathrel{\mathcal{J}} y$	Green's equivalence relation $\mathcal{J}$ defined by $S^1xS^1 = S^1yS^1$; *p. 34*
$x \leqslant_{\mathcal{J}} y$	Order on $\mathcal{J}$-classes defined by $S^1xS^1 \subseteq S^1yS^1$; *p. 34*
$\mathcal{L}_A$	Set of languages over A in the language family $\mathcal{L}$; *p. 32*
$\mathcal{L}_A(\varphi)$	Language over A defined by φ; *p. 42*
$L(\varphi)$	Language over implicitly understood alphabet defined by φ; *p. 42*
$\mathcal{L}(\mathcal{F})$	Family of languages defined by φ; *p. 42*
$x \mathrel{\mathcal{L}} y$	Green's equivalence relation $\mathcal{L}$ defined by $S^1x = S^1y$; *p. 34*
$x \leqslant_{\mathcal{L}} y$	Order on $\mathcal{L}$-classes defined by $S^1x \subseteq S^1y$; *p. 34*
$\lambda(x) \in B$	Label predicate; true if x has a label in B; *p. 41*
$\lambda(x) = a$	Label predicate; true if x has label a; *p. 42*
Λ	Infinite set of labels in formulae; *p. 41*
$\Lambda \times 2^V$	Extended alphabet over $V \subseteq \mathbb{V}$; *p. 43*
$u^{-1}Lv^{-1}$	Residual of the language L; *p. 32*
$u^{-1}L$	Left residual of the language L; *p. 32*
Lv^{-1}	Right residual of the language L; *p. 32*
$\mathbf{LV}$	Variety of all semigroups S such that $eSe \in \mathbf{V}$ for all idempotents e; *p. 37*
$\mathbf{LDA}$	Variety of semigroups S such that $eSe \in \mathbf{DA}$ for all idempotents e; *p. 37*
$\max(x)$	Maximum predicate; *p. 41*
$\min(x)$	Minimum predicate; *p. 41*
MSO	Monadic second-order logic without modular counting quantifiers; *p. 47*
MSO+MOD	Monadic second-order logic with modular counting quantifiers; *p. 47*
$\mu(\varphi)$	Substitution of φ for $\circ$ in the context μ; *p. 45*

μ, ν	Typical names for contexts, *i.e.*, formula with unique occurrence of $\circ$; *p. 45*
$\circ$	0-ary placeholder in contexts; read as "hole"; *p. 45*
$\mathcal{N}$	Typical name for a relational signature; *p. 47*
$\mathbb{N}$	Set of naturals; $0 \in \mathbb{N} \subseteq \mathbb{Z}$; *p. 31*
$N(\varphi)$	Sliding window radius for the sentence φ; *p. 95*
$N_\delta(\varphi)$	Sliding window radius for the formula φ under δ; *p. 95*
ω_S	Idempotent generating power for the finite semigroup S; *p. 33*
$\mathrm{ord}(i, j)$	The order type of integers i and j; *p. 131*
$\mathrm{ord}_S(i, j)$	The successor order type of integers i and j; *p. 139*
$\mathrm{pad}_v(u)$	Padding of u with the word v; *p. 101*
φ, ψ, ξ	Typical names for formulae
$\varphi(x_1, \ldots, x_k, X_1, \ldots, X_\ell)$	At most $x_1, \ldots, x_k, X_1, \ldots, X_\ell$ are free variables of φ; *p. 43*
$\neg\varphi$	Negation; true if φ is false; *p. 42*
$\varphi \wedge \psi$	Disjunction; true if φ and ψ are true; *p. 42*
$\varphi \vee \psi$	Disjunction; true if φ or ψ is true; *p. 42*
$\bigwedge_{i \in I} \varphi_i$	Finite conjunction; true if all of the φ_i are true; *p. 42*
$\bigvee_{i \in I} \varphi_i$	Finite disjunction; true if at least one of the φ_i is true; *p. 42*
$[\![\varphi]\!]_V$	Formal semantics of φ; all structures over V that satisfy φ; *p. 43*
$[\![\varphi]\!]_{A,V}$	Formal semantics over the alphabet A of φ; *p. 43*
$\langle\varphi\rangle_{p,\delta}$	Minimum elimination construction; *p. 109*
Π_m	Negations of Σ_m-formulae; *p. 49*
Π^2_m	Negations of Σ^2_m-formulae; *p. 123*
$\psi \leqslant \varphi$	ψ is a subformula of φ; *p. 117*
$\psi < \varphi$	ψ is a proper subformula of φ; *p. 117*
$\mathrm{qd}(\varphi)$	Quantifier depth of φ; *p. 47*
$\mathcal{Q}^k_n[\mathcal{N}]$	Fragment of formulae with at most k variables and quantifier depth of at most n that only use quantifiers in $\mathcal{Q}$ and numerical predicates in $\mathcal{N}$; *p. 57*
$x \mathrel{\mathcal{R}} y$	Green's equivalence relation $\mathcal{R}$ defined by $xS^1 = yS^1$; *p. 34*

$x \leqslant_{\mathcal{R}} y$	Order on $\mathcal{R}$-classes defined by $xS^1 \subseteq yS^1$; *p. 34*
$R_{m,n}$	Rankers with m alternations and depth n; *p. 130*
$R^{\mathrm{X}}_{m,n}$	Rankers with m alternations and depth n that end on an X-modality; *p. 130*
$R^{\mathrm{Y}}_{m,n}$	Rankers with m alternations and depth n that end on a Y-modality; *p. 130*
$\widetilde{R}_{m,n}$	Look-around rankers with m alternations and depth n; *p. 137*
$\widetilde{R}^{\mathrm{X}}_{m,n}$	Look-around rankers in $\widetilde{R}_{m,n}$ ending on an X-modality; *p. 137*
$\widetilde{R}^{\mathrm{Y}}_{m,n}$	Look-around rankers in $\widetilde{R}_{m,n}$ ending on a Y-modality; *p. 137*
S^1	Smallest monoid containing the semigroup S; *p. 33*
$\mathrm{sd}_{\varphi}(x, y)$	Successor distance within φ between x and y; *p. 88*
$\sigma_N(u)$	Sliding window of radius N applied to u; *p. 84*
$\sigma_{N,\delta}(\varphi)$	Sliding window construction of radius N for φ under δ; *p. 94*
$\sigma_{\delta}(\varphi)$	Sliding window construction for φ under δ; *p. 90*
Σ_m	Half level m of the FO-alternation hierarchy; next full level is FO_m; *p. 48*
Σ^2_m	Half level m of the FO^2-alternation hierarchy; next full level is FO^2_m; *p. 49*
$\Sigma_{m,n}$	Fragment of Σ_m with quantifier depth at most n; *p. 48*
$\Sigma^2_{m,n}$	Fragment of Σ^2_m with quantifier depth at most n; *p. 49*
$\mathrm{suc}(x, y)$	Successor predicate; *p. 41*
$\equiv_L$	Syntactic congruence of the language L; *p. 35*
$\leqslant_L$	Syntactic preorder of the language L; *p. 35*
M_L	Syntactic monoid of the language L; *p. 35*
S_L	Syntactic semigroup of the language L; *p. 35*
$\top$	Constant formula that is always true; *p. 41*
$\bot$	Constant formula that is always false; *p. 41*
TL^+_m	Temporal logic formulae with negation nesting depth of at most m; *p. 129*
$\mathrm{TL}^+_{m,n}$	TL^+_m-formulae with modality nesting depth at most n; *p. 129*
TL_m	Boolean closure of TL^+_m; *p. 129*
$\mathrm{TL}_{m,n}$	Boolean closure of $\mathrm{TL}^+_{m,n}$; *p. 129*
$\lvert u\rvert$	Length of the finite word u; *p. 31*

$u + \delta$	Structure obtained by adding the offsets in δ; *p. 89*
$u[i]$	i^{th} letter of the word u; *p. 31*
$u[i;j]$	Factor of u induced by positions in $[i;j]$; *p. 31*
$u[i;j;k]$	Context given by $(u[i;j-1], u[j], u[j+1;k])$; *p. 143*
$u[1;r)$	Factor $u[1;r(u)-1]$ of u defined by ranker r; *p. 169*
$u(r;s)$	Factor $u[r(u)+1;s(u)-1]$ of u defined by rankers r and s; *p. 169*
$u[x/i]$	Structure with x interpreted by the position i; *p. 43*
$u[X/I]$	Structure with X interpreted by the set of positions I; *p. 43*
$u \cong_{m,n} v$	Equivalence relation given by $u \lesssim_{m,n} v$ and $v \lesssim_{m,n} u$; *p. 173*
$u \approx^{\mathrm{FO2}}_{m,n} v$	Equivalence relation given by $u \preccurlyeq^{\mathrm{FO2}}_{m,n} v$ and $v \preccurlyeq^{\mathrm{FO2}}_{m,n} u$; *p. 173*
$u \equiv^{R}_{m,n} v$	Equivalence relation given by $u \leqslant^{R}_{m,n} v$ and $v \leqslant^{R}_{m,n} u$; *p. 131*
$u \approx^{R}_{m,n} v$	Equivalence relation given by $u \preccurlyeq^{R}_{m,n} v$ and $v \preccurlyeq^{R}_{m,n} u$; *p. 139*
$u \leqslant^{\mathrm{FO2}}_{m,n} v$	Preorder induced by $\Sigma^2_{m,n}[<]$; *p. 131*
$u \preccurlyeq^{\mathrm{FO2}}_{m,n} v$	Preorder induced by $\Sigma^2_{m,n}[<,\mathrm{suc}]$; *p. 139*
$u \lesssim_{m,n} v$	Preorder induced by $\Sigma^2_{m,n}[<,\mathrm{suc},\min,\max]$ if $m = 1$ and $\Sigma^2_{m,n}[<,\mathrm{suc}]$ if $m \geqslant 2$; *p. 175*
$u \leqslant^{R}_{m,n} v$	Preorder induced by rankers in $R_{m,n}$; *p. 130*
$u \preccurlyeq^{R}_{m,n} v$	Preorder induced by rankers in $\widetilde{R}_{m,n}$; *p. 138*
$u \leqslant^{\mathrm{TL}}_{m,n} v$	Preorder induced by $\mathrm{TL}^{+}_{m,n}[\mathrm{XF},\mathrm{YP}]$; *p. 131*
$u \preccurlyeq^{\mathrm{TL}}_{m,n} v$	Preorder induced by $\mathrm{TL}^{+}_{m,n}[\mathrm{XXF},\mathrm{YYP},\mathrm{X},\mathrm{Y}]$; *p. 139*
$(u,x') \leqslant^{R}_{m,n} (v,x)$	Extension of $u \leqslant^{R}_{m,n} v$ to include free variable; *p. 133*
$(u,x') \preccurlyeq^{R}_{m,n} (v,x)$	Extension of $u \preccurlyeq^{R}_{m,n} v$ to include free variable; *p. 144*
$u \models \varphi$	Models relation; structure u satisfies φ; *p. 42*
$u, i_j, I_k \models \varphi(x_j, X_k)$	Explicit specification of the interpretation of free variables; *p. 43*
U_m, V_m	Specific omega-terms for $\mathrm{FO}^2[<,\mathrm{suc},\min,\max]$-alternation; *p. 152*
U'_m, V'_m	Specific omega-terms for $\mathrm{FO}^2[<]$-alternation; *p. 154*
$\mathcal{U}_V$	Universe of structures over $V \subseteq \mathbb{V}$; *p. 43*
$\mathbb{V}$	Set of all variables for formulae; *p. 41*

$\mathbb{V}_1$	Set of first-order variables for formulae; *p. 41*
$\mathbb{V}_2$	Set of second-order variables for formulae; *p. 41*
$\mathbf{V_M}$	Class of monoids in a semigroup variety $\mathbf{V}$; *p. 106*
V_n	Derived set of variables of V; *p. 69*
$\mathcal{W}_{\varphi,\delta}$	Admissible structures for $\sigma_\delta(\varphi)$; *p. 89*
$\mathcal{W}_{\varphi,p,\delta}$	Admissible structures for $\langle\varphi\rangle_{p,\delta}$; *p. 111*
$x(u)$, $X(u)$	Interpretation of x and X in the structure u; *p. 43*
x, y, z, x_i	Typical names for first-order variables $\mathbb{V}_1$; *p. 41*
X, Y, Z, X_i	Typical names for second-order variables $\mathbb{V}_2$; *p. 41*
$x \equiv r \pmod q$	Modular predicate; *p. 41*
Ξ	Infinite set of variables for omega-terms; *p. 36*
x_{max}	New variable that is always the maximal position; *p. 107*
x_{min}	New variable that is always the minimal position; *p. 107*
x^ω	Idempotent generated by x; *p. 33*
$\mathrm{X}, \mathrm{F}, \mathrm{XF}, \mathrm{XXF}$	Temporal logic future modalities; *p. 128*
$\mathrm{Y}, \mathrm{P}, \mathrm{YP}, \mathrm{YYP}$	Temporal logic past modalities; *p. 128*
$X \setminus Y$	Set difference; $X \setminus Y = \{x \in X \mid x \notin Y\}$
$X \,\Delta\, Y$	Symmetric set difference; $X \,\Delta\, Y = (X \setminus Y) \cup (Y \setminus X)$
X_a	Ranker X-modality; *p. 126*
Y_a	Ranker Y-modality; *p. 126*
Z_a	Either of X_a or Y_a; *p. 126*
X_w, XX_w	Look-around ranker X-modalities with n-context w; *p. 127*
Y_w, YY_w	Look-around ranker Y-modalities with n-context w; *p. 127*
Z_w	Either of X_w, Y_w, XX_w, or YY_w; *p. 127*
$\mathfrak{z}$	A generic first-order or second-order variable; *p. 117*
$\mathbb{Z}$	Set of integers; *p. 31*

Abstract

The purpose of this thesis is to deepen the insight into the complex interactions between logic and computers. Its focus is on regular languages of finite words for which it provides logic tools to reveal new features of the rich structure of sub-regular languages.

Part I develops an abstract theory of logic fragments. A formal definition for a set of formulae to constitute a *logic fragment* is given in terms axiomatic closure properties on syntax level. The axioms of a fragment are the following:

- Any subformula may be replaced by “true”, by “false”, or by a label predicate within a logic fragment, provided that the set of free variables is not altered.
- Any branch of a disjunction or a conjunction may be pruned within a logic fragment.
- Conversely, formulae occurring in the same context may be recombined in this context within the logic fragment by disjunction or by conjunction.
- First-order quantifiers may be omitted without leaving the logic fragment, provided that the bound variable does not appear as a free variable in the scope of the quantifier — which is thus independent of the bound variable.

Logic fragments have been extensively studied in the literature before, but there the term merely signified a set of formulae obtained by restricting certain resources. The logic framework of this thesis includes the commonly used logic constructs. Moreover, the axioms of logic fragments are few and natural, so that virtually all informal logic fragments occurring in the literature are also fragments in the formal sense of this thesis.

It is shown that this formal notion of a logic fragment leads to natural closure properties of the family of languages defined by its formulae, namely closure under (positive) Boolean operations, residuals, and inverse images of homomorphisms in $\mathcal{C}$ for a particular natural homomorphism family $\mathcal{C}$. Depending on the resources used by the logic fragment, inverses of only a subfamily $\mathcal{C}$ of homomorphisms may be possible; for example, only inverses of non-erasing homomorphisms are possible in general if the successor predicate is available.

Combining these closure properties leads to the fundamental Theorem 5.7 on logic fragments, characterizing logic fragments in terms of $\mathcal{C}$-varieties, which are a generalization of Eilenberg’s $+$-varieties and $*$-varieties. This enables to tackle decidability of definability in a logic fragment using the algebraic approach in terms of the syntactic homomorphism.

Depending on the constructs used by the logic fragment and on the desired semantic closure property, additional axioms may be necessary. Such application-specific axioms are formalized by stability notions. For closure under inverse homomorphisms, for example, the fragment needs to be *order-stable*, which states that non-strict and strict order predicates are syntactically interchangeable with respect to the fragment. More interestingly, *suc-stability* is necessary for closure under residuals; it states that the successor predicate may be replaced in any context by an equality, minimum, or maximum predicate. This gives a purely syntactic reason why the successor predicate often entails minimum and maximum predicates, especially for small logic fragments.

The introduction of the $\mathcal{C}$-variety framework by Straubing was actually motivated by logic, with the main result that a multitude of logic fragments is characterized by $\mathcal{C}$-varieties. In addition to the usual logic constructs, he also included modular counting quantifiers, which make traditional proof techniques such as Ehrenfeucht-Fraïssé-games hard to apply. In fact, he only gave an outline of the argument and to date a promised formal proof did not appear. The techniques developed in this thesis yield Theorem 5.13 which generalizes Straubing's result.

After this abstract theory of logic fragments, two concrete examples of fragments are considered. The first example is the fragment of acyclic formulae in $\Sigma_2[<]$ where order comparisons between variables are available only if no cycles are generated in the so-called comparison graph. Theorem 6.1 shows that the acyclic fragment of $\Sigma_2[<]$ coincides with the two-variable first-order fragment $\mathrm{FO}^2[<]$.

The second example fragment is $\mathbb{B}\Sigma_1^+[\leqslant]$ consisting of Boolean combinations of negation-free, purely existential first-order formulae that are only allowed to use *non-strict* order comparison. Note that neither the strict order predicate nor negations of atomic formulae are allowed. Theorem 7.1 shows that $\mathbb{B}\Sigma_1^+[\leqslant]$ is expressively complete for stutter-invariant $\mathbb{B}\Sigma_1[<]$-definable languages. In particular, the restriction to the non-strict order predicate is essential.

After the interlude with these two example fragments, the power of the abstract theory of logic fragments is further exemplified. The first application of this approach studies the expressive power of the successor predicate. It is shown in Theorem 8.6 that for many logic fragments the successor predicate does not provide expressive power beyond the ability to sample the labels of surrounding positions. This is well-known for certain concrete logic fragments and is classically obtained either indirectly by way of non-logic characterizations, or using *ad hoc* methods that are tailor-made for the specific logic fragment at hand. In contrast to these classical methods, this thesis applies a direct approach on the syntax level of formulae and provides an abstract, purely logical framework for such results.

Based on this framework, Theorem 8.23 shows that many fragments do not gain expressive power by the availability of the successor predicate in the presence of a neutral letter. Here, a letter is neutral if membership in the language is invariant under inserting or deleting the letter in any context. Building on this result, Proposition 8.26 provides a means to reduce the recognizing power of monoids in a variety $\mathbf{V}$ in terms of logic fragments to the recognizing power of semigroups in $\mathbf{V}$ in terms of logic fragments.

As a second application of the axiomatic approach to logic fragments, the expressive power of the predicates for minimum and maximum is analyzed. Theorem 9.2 shows that the minimum predicate is dispensable whenever enough information about the prefix of the word is known. By left-right symmetry, the maximum predicate only provides information about the suffix.

It is plausible that it is almost never worthwhile to quantify the same variable twice in succession. The last application of the axiomatic approach justifies this intuition on a very abstract level: Proposition 10.3 shows that any formula can be converted into an equivalent requantification-free form in a way that respects most logic fragments.

Part II turns to the quantifier alternation hierarchy within two-variable first-order logic FO^2, consisting of full levels FO^2_m with at most $m-1$ quantifier alternations, and half levels Σ^2_m with at most $m-1$ quantifier alternations that start with existential quantifiers. An important combinatorial tool in the context of FO^2 are *rankers*. Such rankers were introduced by Weis and Immerman to characterize the expressive power of $\mathrm{FO}^2_m[<]$ and $\mathrm{FO}^2_m[<, \mathrm{suc}]$ combinatorially. A ranker is a sequence of single-step instructions to navigate on a word; it is either undefined or identifies a unique position of the word. The precise mode of operation of the instructions depends on the signature of the logic, but in any case, the position to the left or to the right is sought that is nearest to the current position and that satisfies a specified alphabetic condition. For alternation in $\mathrm{FO}^2[<]$ ranker instructions seek the nearest position labeled by a specified letter, whereas for alternation in $\mathrm{FO}^2[<, \mathrm{suc}]$ they seek the nearest occurrence of a specified factor.

Chapter 11 contributes a ranker characterization for all half levels $\Sigma^2_m[<]$ and $\Sigma^2_m[<, \mathrm{suc}]$ in Theorem 11.3 and Theorem 11.16, from which ranker descriptions for $\mathrm{FO}^2_m[<]$ and $\mathrm{FO}^2_m[<, \mathrm{suc}]$ immediately follow. This yields Weis and Immerman's result for $\mathrm{FO}^2_m[<]$ as a corollary; and for $\mathrm{FO}^2_m[<, \mathrm{suc}]$ it rectifies an incorrect assertion of theirs. As a convenient intermediate step, temporal logic fragments capturing two-variable quantifier alternation are used, generalizing a result by Etessami, Vardi, and Wilke.

Chapter 12 contributes effective algebraic characterizations for all levels of the alternation hierarchy over any signature between $[<, \mathrm{suc}]$ and $[<, \mathrm{suc}, \min, \max]$ as well as over $[<]$. For the levels FO^2_1 and Σ^2_1 these characterizations are due to Simon, Knast, Pin, and Glaßer and Schmitz, depending on the signature. New and complete proofs are given for these classical results.

The characterizations in particular show that definability in any of the following fragments is decidable:

- $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$,
- $\Sigma^2_m[<, \mathrm{suc}, \min]$,
- $\Sigma^2_m[<, \mathrm{suc}, \max]$,
- $\Sigma^2_m[<, \mathrm{suc}]$,
- $\Sigma^2_m[<]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \min]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \max]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}]$,
- $\mathrm{FO}^2_m[<]$.

Decidability of the full levels $\mathrm{FO}^2_m[<]$ was already known by a result due to Kufleitner and Weil as well as independently by a result by Krebs and Straubing. Both these decidability results are obtained by means of different characterizations.

Zusammenfassung

Diese Arbeit behandelt das Zusammenspiel von Logik und Automatentheorie. Ziel ist es, mit Hilfe von Logik einen Beitrag zum Verständnis der reichhaltigen Struktur subregulärer Sprachen endlicher Wörter zu leisten.

Im ersten Teil wird eine abstrakte Theorie der Logikfragmente entwickelt. Es werden axiomatische Bedingungen angegeben, die eine Formelmenge auf Syntaxebene erfüllen muss, um ein *Logikfragment* zu bilden. Die syntaktischen Abschlusseigenschaften eines Logikfragments sind wie folgt:

- Innerhalb eines Logikfragments kann jede Teilformel durch „wahr“, durch „falsch“ oder durch ein Beschriftungsprädikat ersetzt werden, sofern die Menge der freien Variablen unverändert bleibt.
- Beliebige Zweige von Disjunktionen und Konjunktionen können innerhalb des Logikfragments gestrichen werden.
- Umgekehrt dürfen Formeln, die im selben Kontext vorkommen, innerhalb des Logikfragments in diesem Kontext durch Disjunktion und Konjunktion verknüpft werden.
- Quantoren erster Stufe dürfen getilgt werden, sofern die dadurch gebundene Variable nicht frei in der quantifizierten Formel vorkommt, diese also unabhängig von der gebundenen Variable ist.

Der Begriff eines Logikfragments wird in der einschlägigen Literatur bereits häufig verwendet, bezeichnete bisher jedoch lediglich informell jedwede durch Ressourcenbeschränkungen entstandene Formelmenge. Der formale Rahmen von Formeln ist dabei sehr allgemein gehalten und umfasst die üblicherweise in der Literatur vorkommenden Konstrukte. Ferner sind die Axiome sehr natürlich, was dazu führt, dass nahezu alle in der Literatur vorkommenden informellen Logikfragmente auch Logikfragmente im formalen Sinne sind.

Basierend auf diesen Axiomen wird anschließend gezeigt, dass die syntaktischen Abschlusseigenschaften von Logikfragmenten auf natürliche semantische Abschlusseigenschaften der definierten Sprachklasse führen. Konkret wird gezeigt, dass Logikfragmente Sprachklassen definieren, die abgeschlossen sind unter (positiven) booleschen Verknüpfungen, Residuen, sowie Urbildern von Homomorphismen in natürlichen Homomorphismenklassen $\mathcal{C}$. Der Parameter $\mathcal{C}$ spiegelt die Tatsache wider, dass die von einem Logikfragment definierte Sprachklasse je nach verwendeten Ressourcen im Allgemeinen nicht unter Urbildern von beliebigen Homomorphismen abgeschlossen ist. So ist im Allgemeinen nur ein Abschluss unter inversen nicht-löschenden Homomorphismen möglich, sofern dem Logikfragment ein Nachfolgerprädikat zur Verfügung steht.

Diese Abschlusseigenschaften führen in Theorem 5.7 auf eine Charakterisierung der Ausdrucksstärke von Logikfragmenten mittels sogenannter $\mathcal{C}$-Varietäten, welche eine Verallgemeinerung von Eilenbergs $+$-Varietäten und $*$-Varietäten darstellen. Dies eröffnet den algebraischen Ansatz, um mittels des syntaktischen Homomorphismus nachzuweisen, dass Definierbarkeit in einem betrachteten Logikfragment entscheidbar ist.

Die verfügbaren Ressourcen des Logikfragments sowie die gewünschten Abschlusseigenschaften erfordern teils weitere Axiome. Solche anwendungsspezifische Axiome werden durch Stabilitätsbegriffe abstrahiert. Für den Abschluss unter inversen Homomorphismen ist zum Beispiel die Ordnungsstabilität (engl. *order-stability*) notwendig, die besagt, dass strikte und nicht-strikte Ordnungsprädikate im Logikfragment auf Syntaxebene austauschbar sind. Interessanter ist, dass die Nachfolgerstabilität (engl. *suc-stability*) für den Residuenabschluss von Logikfragmenten notwendig ist. Diese besagt, dass ein Nachfolgerprädikat innerhalb des Logikfragments stets durch ein Gleichheits-, Minimum- oder Maximumprädikat ersetzt werden darf. Dies begründet auf rein syntaktischer Ebene, warum ein Nachfolgerprädikat meist auch Minimum- und Maximumprädikate nach sich zieht, speziell bei ausdrucksschwachen Logikfragmenten.

Die Einführung von $\mathcal{C}$-Varietäten durch Straubing war seinerzeit durch Logik motiviert, mit dem Hauptergebnis, dass eine Vielzahl von Logikfragmenten $\mathcal{C}$-Varietäten definieren. In diesem Rahmen waren modular zählende Quantoren enthalten, für welche übliche Techniken wie Ehrenfeucht-Fraïssé-Spiele nur mit Schwierigkeiten anwendbar sind. Tatsächlich umriss Straubing die Argumentation auch nur und ein angekündigter formaler Beweis ist bisher nicht erschienen. Mit Theorem 5.13 wird eine Verallgemeinerung dieses Ergebnisses von Straubing bewiesen.

Anschließend werden beispielhaft zwei bislang nicht untersuchte Logikfragmente betrachtet. Das erste Beispiel ist das Fragment der azyklischen Formeln in $\Sigma_2[<]$, in dem Ordnungsvergleiche zwischen Variablen nur erlaubt sind, sofern keine Zyklen im sogenannten Ordnungsvergleichsgraphen (engl. *comparison graph*) entstehen. In Theorem 6.1 wird gezeigt, dass dieses azyklische Fragment von $\Sigma_2[<]$ semantisch mit dem Logikfragment $\mathrm{FO}^2[<]$ übereinstimmt.

Das zweite betrachtete Beispielfragment ist $\mathbb{B}\Sigma_1^+[\leqslant]$, welches aus booleschen Kombinationen von rein existentiellen Formeln der Logik erster Stufe ohne jegliche Negationen besteht, die nur das nichtstrenge Ordungsprädikat benutzen. Insbesondere sind weder das Prädikat für die strenge Ordnung noch Negationen von atomaren Formeln verfügbar. Wie Theorem 7.1 zeigt, ist die Ausdrucksstärke von $\mathbb{B}\Sigma_1^+[\leqslant]$ vollständig für die stotter-invarianten (engl. *stutter-invariant*) $\mathbb{B}\Sigma_1[<]$-definierbaren Sprachen. Dies zeigt insbesondere, dass die Einschränkung auf das nichtstrenge Ordungsprädikat wesentlich ist. Die recht unkonventionellen Einschränkungen, durch welche diese beiden Beispiele entstehen, passen gut in den in dieser Arbeit eingeführten formalen Rahmen der Logikfragmente, mögen jedoch gleichzeitig der Grund sein, weshalb diese trotz ihrer natürlichen Charakterisierungen bislang nicht in der Literatur behandelt wurden.

Nach der Betrachtung dieser beiden konkreten Logikfragmente wird die axiomatische Theorie der Logikfragmente angewendet, um folgende Ergebnisse zu erhalten: Zunächst wird der Einfluss des Nachfolgerprädikats auf die Ausdrucksstärke von Logikfragmenten untersucht. In Theorem 8.6 wird gezeigt, dass die Verfügbarkeit des Nachfolgerprädikats für viele Logikfragmente keine zusätzliche Ausdrucksstärke bewirkt, außer dass die Beschriftung umliegender Positionen abgefragt werden kann. Für viele konkrete Logikfragmente ist dies eine bekannte Tatsache, was klassischerweise entweder indirekt, d. h. durch eine Charakterisierung außerhalb der Logik, oder *ad hoc*, d. h. durch eine auf das vorliegende Logikfragment zugeschnittene Methode, erwirkt wurde. Im Gegensatz hierzu verfolgt diese Arbeit einen Ansatz, der direkt auf Syntaxebene von Formeln ansetzt und damit einen rein formallogischen Rahmen für derartige Resultate liefert.

Auf ebendiesem Rahmen aufbauend, zeigt Theorem 8.23, dass die Verfügbarkeit des Nachfolgerprädikats für viele Logikfragmente keine zusätzliche Ausdrucksstärke bewirkt, sofern ein neutraler Buchstabe vorhanden ist. Dabei ist ein Buchstabe neutral, falls die Zugehörigkeit zur betrachteten Sprache durch Einfügen und Löschen dieses Buchstabens an beliebiger Stelle unverändert bleibt. Auf diesem Resultat fußend, zeigt Proposition 8.26, wie eine Logikbeschreibung der Halbgruppen in einer Varietät $\mathbf{V}$ in eine Logikbeschreibung der Monoide in $\mathbf{V}$ umgewandelt werden kann.

Die zweite Anwendung des axiomatischen Ansatzes von Logikfragmenten untersucht den Einfluss der Prädikate für Minimum und Maximum auf die Ausdrucksstärke eines Fragments. Wie Theorem 9.2 zeigt, ist das Minimumsprädikat häufig überflüssig, sofern genügend Information über das Präfix vorhanden ist. Dies mag intuitiv erscheinen, ist jedoch nicht *a priori* klar, da das Minimumsprädikat in der Formel jederzeit benutzt werden kann, nicht nur auf äußerster Ebene. Man beachte, dass die Anzahl der Variablen beschränkt sein kann und ein Speichern der minimalen Position in einer neuen Variablen somit im Allgemeinen nicht zulässig ist. Durch Links-Rechts-Symmetrie liefert das Maximumsprädikat entsprechend nur Informationen über das Suffix.

Es ist plausibel, dass eine direkt aufeinanderfolgende Quantifizierung ein und derselben Variablen nahezu nie lohnenswert ist. Die letzte Anwendung des axiomatischen Ansatzes rechtfertigt diese Intuition auf einer sehr abstrakten Ebene: Proposition 10.3 zeigt, dass jede Formel in eine äquivalente Formel umgeformt werden kann, die keine direkt aufeinanderfolgende Quantifizierung ein und derselben Variablen besitzt. Darüber hinaus kann diese Umformung derart vorgenommen werden, dass die meisten Logikfragmente respektiert werden.

Der zweite Teil der vorliegenden Arbeit ist der Untersuchung der Quantorenalternierungshierarchie innerhalb des Logikfragments FO^2 erster Stufe mit zwei Variablen gewidmet. Diese besteht aus den Stufen FO^2_m, in denen höchstens $m-1$ Wechsel zwischen Existenz- und Allquantoren erlaubt sind, sowie den Zwischenstufen Σ^2_m, die darüber hinaus Existenzquantoren auf äußerster Ebene fordern.

Ein wichtiges Hilfsmittel zur Untersuchung von FO^2 sind sogenannte *Ranker*. Diese wurden von Weis und Immerman eingeführt, um eine kombinatorische Beschreibung der Ausdrucksstärke der Stufen $\mathrm{FO}^2_m[<]$ und $\mathrm{FO}^2_m[<, \mathrm{suc}]$ zu erhalten. Ein Ranker ist dabei ein einfaches Programm, um auf Wörtern zu navigieren. Für ein gegebenes Wort ist dieser entweder undefiniert oder er spezifiziert eine eindeutige Position. Ein solches Programm besteht aus einer Abfolge von Einzelschrittanweisungen, die jeweils die nächste Position links bzw. rechts der aktuellen Position anspringen, die eine spezifizierte alphabetische Bedingung erfüllt. Wie diese alphabetischen Bedingungen aussehen können, wird durch die Signatur des betrachteten Logikfragments (entweder $[<]$ oder $[<, \mathrm{suc}]$) bestimmt. Im Logikfragment $\mathrm{FO}^2[<]$ kann ein einzelner Buchstabe spezifiziert werden, wohingegen im Logikfragment $\mathrm{FO}^2[<, \mathrm{suc}]$ nach kurzen Faktoren gesucht werden kann.

Kapitel 11 gibt in Theorem 11.3 und Theorem 11.16 Ranker-Beschreibungen der Ausdrucksstärke aller Zwischenstufen $\Sigma^2_m[<]$ und $\Sigma^2_m[<, \mathrm{suc}]$ an. Ranker-Beschreibungen für die Stufen $\mathrm{FO}^2_m[<]$ und $\mathrm{FO}^2_m[<, \mathrm{suc}]$ folgen unmittelbar hieraus. Für $\mathrm{FO}^2_m[<]$ folgt daraus das Resultat von Weis und Immerman. Im Falle von $\mathrm{FO}^2_m[<, \mathrm{suc}]$ berichtigt dies eine inkorrekte Behauptung von Weis und Immerman. Als geeigneter Zwischenschritt werden Fragmente einer Temporallogik für die Quantorenalternierungshierarchie angegeben. Dies verallgemeinert ein Ergebnis von Etessami, Vardi und Wilke.

Kapitel 12 stellt effektive algebraische Beschreibungen aller Stufen der Alternierungshierarchien über allen Signaturen zwischen $[<, \mathrm{suc}]$ und $[<, \mathrm{suc}, \min, \max]$ sowie über $[<]$ vor. Für die untersten Stufen FO^2_1 und Σ^2_1 sind diese bereits durch Ergebnisse von Simon, Knast, Pin sowie Glaßer und Schmitz bekannt. Für diese klassischen Resultate werden neue Beweise angegeben.

Aus den algebraischen Charakterisierungen in Kapitel 12 folgt, dass die Definierbarkeit in jedem der folgenden Logikfragmente entscheidbar ist:

- $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$,
- $\Sigma^2_m[<, \mathrm{suc}, \min]$,
- $\Sigma^2_m[<, \mathrm{suc}, \max]$,
- $\Sigma^2_m[<, \mathrm{suc}]$,
- $\Sigma^2_m[<]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \min]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \max]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}]$,
- $\mathrm{FO}^2_m[<]$.

Die Entscheidbarkeit der Stufen $\mathrm{FO}^2_m[<]$ ist bereits durch Ergebnisse von Kufleitner und Weil sowie Krebs und Straubing bekannt.

1. Introduction

Background. Logic has played a fundamental role right from the very beginning of computer science. Boldly speaking, one might even venture to say that back in the early 1930s, logic was a driving force towards computer science as we know it today. In fact Hilbert's *Entscheidungsproblem*, which asks whether there is an algorithm that can decide the truth of a given formula in a certain logic formalism, led to the mathematical foundations of computer science and precise definitions of formerly vague concepts such as *algorithm* or *decidability*: Turing introduced his famous machines in his seminal paper from 1936 and described the computation of Turing machines by a formula, thus reducing the halting problem of Turing machines to the *Entscheidungsproblem* [Tur36]; in the same year Church independently introduced his lambda-calculus to model computing devices and, using this formalism, also showed the undecidability of the *Entscheidungsproblem* [Chu36].

Later, logic emerged quite unexpectedly in the context of computational complexity theory: In 1974, Fagin showed that the complexity class NP (that is, problems solvable in polynomial time by a nondeterministic Turing machine) coincides with the family of problems expressible in existential second-order logic; *cf.* [Fag74]. This started the investigation of what is known as the *descriptive complexity* of computational problems, which aims to classify computational problems in terms of the logic resources necessary to describe the problem. It turned out that virtually all complexity classes had natural logic counterparts, including famous classes such as LOGSPACE, NLOGSPACE, P, PSPACE, EXPTIME, and the recursively enumerable sets. Refer to Immerman's monograph for references, details, and an extensive survey concerning results in descriptive complexity, including Fagin's Theorem [Imm99].

Apart from such applications in theoretical computer science, logic plays a key role in many practical applications such as databases or the algorithmic verification of systems. For instance, a relational database is a logical structure that can be accessed by queries in some *query language* like SQL. Such query languages can be seen as an engineers approach to implement logic formalisms, often variants of first-order logic [Cod70; Cod90].

A frequently occurring problem of algorithmic verification is *model checking*, where a given implementation of a system like a finite state machine is to be checked algorithmically for a given property, which is specified in some logic formalism. As an example of model checking consider an elevator system. A typical desirable property could be that every transportation request (event a) is eventually followed by a grant (event b), which can be formalized by the first-order formula $\forall x \, \big(\lambda(x) = a \to \exists y > x \colon \lambda(y) = b\big)$. The Greek letter lambda is for *label*; this means for every point in time x that is labeled by event a, the formula postulates a later point in time y that is labeled by event b.[1] The elevator system may be implemented, for example, by a transition system. A property means that every possible run of the transition system satisfies the formula. So if the

[1]Precise definitions can be found in Section 3.1.

system is given by the transition system $\mathcal{A}$ and the property is given by the formula φ, the problem boils down to checking the formal language inclusion $L(\mathcal{A}) \subseteq L(\varphi)$. For more details about this automata theoretic approach to model checking refer to [Var07].

This thesis takes this language point of view on logic, associating each formula with the formal language it defines, and studies logic under formal language theoretic aspects, with a focus on regular languages of finite words. The setting in this area was given independently by Büchi, Elgot, and Trakhtenbrot in the early 1960s, who showed that monadic second-order logic (MSO) is expressively complete for regular languages [Büc60; Elg61; Tra61]; *i.e.*, that every regular language is defined by some sentence in MSO and, conversely, every language definable in MSO is regular. The attribute *monadic* refers to the fact that second-order quantifiers range only over monadic predicates (*i.e.*, sets of positions) and not over predicates of higher arity (like sets of pairs of positions). Despite its expressiveness, monadic second-order logic has several drawbacks: For one thing, second-order constructs are not particularly easy to grasp for the human mind, thus forfeiting much of the naturalness in stating properties. For another, typical decision problems such as the above inclusion problem have huge computational complexity for MSO-formulae [Sto74]. It may therefore be desirable to describe properties in a more tractable fragment of MSO in order get algorithms with a better complexity.

Logic fragments are also interesting from the point of view of descriptive complexity: Classifying regular languages by logic fragments is an intriguing way to shed light on the rich structure of sub-regular languages. Such logic fragments are classically obtained by restricting certain resources such as

- the atomic predicates,
- the set of quantifiers,
- the quantifier depth,
- the number of alternations between different quantifier types, and
- the number of available variables.

The most common predicates are equality $x = y$, order comparison $x < y$, successor $\mathrm{suc}(x, y)$ which is true if $x + 1 = y$, minimum and maximum $\min(x)$ and $\max(x)$, and modular predicates $x \equiv r \pmod q$. Apart from first-order and second-order quantifiers, a common quantifier found in the literature is the modular counting quantifier of the form $\exists^{r \bmod q}\varphi$, which is true if there are r positions (modulo q) that make φ true.

One might wonder at this point why, for example, a successor predicate $\mathrm{suc}(x, y)$ is included when it can readily be expressed by $(x < y) \wedge \neg\exists z\ (x < z \wedge z < y)$. Although these formulae are semantically equivalent, the purpose of logic fragments is to syntactically restrict resources of formulae, and the argumentation may violate all of the classical restrictions listed above: First, the order predicate and existential quantifiers might be unavailable; second, the quantifier depth is increased; third, the negation increases the alternation depth; and fourth, a new variable name z is used. This shows that predicates that might be superfluous for more expressive logic fragments can be crucial for smaller fragments. In fact, in Büchi's sequential second-order calculus only the successor predicate is available [Büc60]. In the full second-order setting, the order predicate is expressible using only the successor predicate, but first-order logic drops in expressive power if only the successor predicate is allowed [Tho82].

Considering some concrete logic fragment $\mathcal{F}$, one can of course reconsider the typical decision problems such as satisfiability restricted to $\mathcal{F}$, which may well drop in computational complexity. Moreover, deciding definability in $\mathcal{F}$ naturally arises: On input of a property presented for instance by an MSO-formula or a finite automaton, decide whether this property is expressible in $\mathcal{F}$.

A cornerstone in this direction is the effective characterization of the first-order fragment of MSO by combining results due to McNaughton and Papert, and to Schützenberger: In 1971, McNaughton and Papert showed that first-order logic (FO) over the relational signature $[<]$ defines exactly the star-free languages; *i.e.*, those languages obtained by regular expressions generalized by a complement operator but without the Kleene star [MP71]. Already in 1965, Schützenberger had characterized star-free languages in terms of finite aperiodic semigroups; *i.e.*, finite semigroups containing only trivial subgroups [Sch65]. A decision procedure is thus as follows: First compute the syntactic semigroup of the input language.[2] Then check whether the syntactic semigroup is aperiodic, which can be done effectively. Modern proofs of the results of McNaughton and Papert, and Schützenberger can be found in the survey [DG08].

The theorem of McNaughton and Papert on the first-order fragment of MSO initiated the research into other fragments of MSO. Subsequently, the literature on the subject evolved quite extensively, including, *inter alia*, work done by Thomas [Tho82] on quantifier alternation in first-order logic; by Pin [Pin05] on existential first-order formulae with only the successor predicate and the Boolean closure thereof; by Etessami, Vardi, and Wilke [EVW02] and Thérien and Wilke [TW98] on first-order formulae with only two variables; by Barrington, Compton, Straubing, and Thérien [BCST92], Straubing, Thérien, and Thomas [STT95], Straubing [Str01], Chaubard, Pin, and Straubing [CPS06], Kufleitner and Walter [KW13], and Dartois and Paperman [DP13; DP14], all of whom incorporate modular arithmetic resources such as modular predicates or modular counting quantifiers. For an overview and more details refer to the surveys [TT07; DGK08; TT02] or Straubing's monograph [Str94].

Apart from trying to understand the expressive power of logic fragments (*e.g.* by giving alternative descriptions of the languages defined by the formulae in a logic fragment), many papers are devoted to decidability of definability in a certain logic fragment. Most of the decidability results were obtained much like for first-order logic, *i.e.*, by establishing an effective algebraic criterion to be checked on the syntactic semigroup. And indeed, it is not by accident that algebra and the syntactic semigroup enter the scene at this point. In 1976, Eilenberg showed that *varieties* (families of languages with very natural closure properties) correspond to *varieties of finite semigroups* (classes of finite semigroups with natural closure properties) [Eil76]; see also [Pin86].[3] Reiterman showed six years later that membership of a finite semigroup in a varieties of finite semigroups only depends on the validity of a finite set of formal identities [Rei82].

The closure properties required for a variety involve Boolean combinations (*i.e.*, union, intersection, and complementation), taking residuals (*i.e.*, languages of the form $u^{-1}Lv^{-1}$

[2] The syntactic semigroup of a regular language is its unique minimal recognizer. It can be thought of as the algebraic analogue of the minimal deterministic automaton and is computable from any reasonable presentation of regular languages, *cf.* Section 2.3 and [Pin86; Pin97].

[3] Varieties of finite semigroups are also known as *pseudovarieties*, where the prefix "pseudo" is used to distinguish them from varieties of infinite semigroups in Birkhoff's sense, *cf.* [Bir35].

containing all words w with $uwv \in L$), and inverse images of homomorphisms. Relaxing these requirements, Eilenberg's correspondence was extended in several directions: Pin introduced *positive varieties*, which do not require closure under complementation [Pin95]; Straubing introduced $\mathcal{C}$-varieties, where closure is only required under inverses of homomorphisms in $\mathcal{C}$ [Str02]; Kunc and later Pin and Straubing proved an equational theory, extending Reiterman's theorem to $\mathcal{C}$-varieties and their positive variants [Kun03; PS05]. A survey on the various variants of these equational theories can be found in [Pin12].

All these results aim for the same direction: Determining membership of a language in a family by checking finitely many identities for validity in a syntactic object. It might seem that this yields a meta-theorem, showing decidability of the membership problem for all varieties: Compute the syntactic semigroup and check whether the identities are valid in it. Let us hasten to stress that this meta-theorem is *not effective* and decidability does not follow *per se*. The issue is that identities may get so complex that checking them may be undecidable. However, it does provide strong ground to try the mentioned algebraic approach. In concrete cases it may well be possible to give an effective decidable criterion on the syntactic algebraic object. Consider again the example of first-order definable languages: Their algebraic counterparts are finite aperiodic semigroups defined by the identity $x^{\omega+1} = x^{\omega}$, where x^{ω} denotes the idempotent power generated by x. This identity can clearly be checked for a given finite semigroup.

Coming back to logic, definability in many fragments is open to this algebraic scheme, because the underlying logic fragment indeed has sufficient closure properties on the language level. To establish such closure properties, traditionally *ad hoc* techniques tailor-made for the concrete logic fragment at hand were applied, the most common being so-called Ehrenfeucht-Fraïssé-games; *cf.* [Str94]. The farthest reaching result so far in this direction is probably Straubing's result on $\mathcal{C}$-varieties [Str02, Theorem 3]. It gives $\mathcal{C}$-variety characterizations for a huge collection of logic fragments determined by the traditional resource restrictions in terms of predicates, quantifiers, quantifier depth, and number of variables. Unfortunately, the paper contains only a rough outline of the argument in the spirit of Ehrenfeucht-Fraïssé-games, but no formal proof. Straubing also includes modular counting quantifiers and he admits that a game-based reasoning is difficult to apply in such a setting. Indeed, a promised full proof never appeared. Moreover, Straubing's result does not include one of the traditional resources listed above, namely quantifier alternation.

Contributions of this thesis. Part I of this thesis tackles these issues by introducing an abstract formal language theory of logic fragments. In Definition 4.1 we define a formal notion of a logic *fragment* based on an axiomatic approach that requires natural syntactic closure properties. For example, any subformula can be replaced by "true" or by "false" within a fragment. The axioms being few and natural, virtually all "fragments" of the literature on the subject are indeed fragments in the formal sense. Apart from the naturalness of the axioms, this extensive coverage of *ad hoc* logic fragments is the principal design criterion for the notion of a fragment.

We shall see that the closure properties on the syntax level translate to semantic closure under residuals and inverses of homomorphisms in $\mathcal{C}$, where $\mathcal{C}$ is a certain homomorphism family that inherently depends on the constructs used by the logic fragment. This leads to $\mathcal{C}$-variety characterizations of logic fragments in Theorem 5.7, vastly extending the result of Straubing. Instead of using classical Ehrenfeucht-Fraïssé-games, we give direct syntactic constructions by an induction on the structure of the formulae, for which the axioms of a logic fragment allow to maintain a very strong induction invariant.

Straubing's result on $\mathcal{C}$-varieties is formally not a direct corollary to Theorem 5.7 due to some technical issues. However, the techniques employed to prove this theorem also allow to prove Theorem 5.13, thereby contributing the first formal proof for Straubing's theorem. Note that Ehrenfeucht-Fraïssé-games were only recently adapted to match our formal notion of logic fragments [HK14].

After this abstract theory of logic fragments, we exemplify the universality of our theory with two concrete, but rather unorthodox fragments, namely the acyclic fragment of $\Sigma_2[<]$ and $\mathbb{B}\Sigma_1^+[\leqslant]$. In acyclic formulae order comparisons between positions are only allowed if no cycles are introduced in the so-called comparison graph. For example, if the atomic formula $x < y$ is available, then $y < x$ is unavailable. The fragment $\mathbb{B}\Sigma_1^+[\leqslant]$ contains Boolean combinations of negation-free, purely existential first-order formulae that are only allowed to use the *non-strict* order predicate. Theorem 5.7 shows that both examples define varieties. Moreover, they define well-known language families: Theorem 6.1 shows that the acyclic fragment of $\Sigma_2[<]$ coincides with $\mathrm{FO}^2[<]$; *i.e.*, first-order formulae using only two variables and the order predicate. Theorem 7.1 characterizes the fragment $\mathbb{B}\Sigma_1^+[\leqslant]$ as expressively complete for stutter-invariant $\mathbb{B}\Sigma_1^+[<]$-definable languages. Stutter-invariant means that duplicating any single letter or deleting any one of two consecutive duplicate letters makes no difference with respect to membership in the language. Note that in this case the difference between $\leqslant$ and $<$ determines whether the defined language family is stutter-invariant or not. For both $\mathbb{B}\Sigma_1^+[\leqslant]$ and the acyclic fragment of $\Sigma_2[<]$, Ehrenfeucht-Fraïssé-games are difficult to apply. The abstract framework of logic fragments along with the aforementioned example applications to $\mathbb{B}\Sigma_1^+[\leqslant]$ and to the acyclic fragment of $\Sigma_2[<]$ is published in [KL12b].

After characterizing these two concrete fragments, applications exemplifying the power of the axiomatic approach to logic fragments are considered. As a first application, the expressive power of the successor predicate is studied. Intuitively, the availability of the successor predicate in a logic fragment should, under certain assumptions, amount to the ability to query the label of surrounding positions. This intuition is formalized by a sliding window of fixed diameter that is "dragged" over the word, *i.e.*, a window is centered at each position and the letter at this position is replaced by the factor visible through the window. Using these meta-letters, the formula has access to all labels in a finite environment of positions. A formula is a sliding window formula for a given formula if it is equivalent to the original formula, provided that the sliding window is applied before the evaluation. Theorem 8.6 gives such sliding window formulae that respect fragments obeying certain additional assumptions. This theorem is actually not fully satisfactory because it requires generalized quantifiers that are able specify a minimum distance between variables. However, in many cases such as the first-order quantifier alternation hierarchy, these generalized quantifiers can be eliminated within the fragment. This elimination is illustrated in Proposition 8.17. The sliding window approach is classical, though in the literature it is mostly obtained by *ad hoc* arguments such as Ehrenfeucht-Fraïssé-games. In contrast, this thesis provides a direct syntactic solution on fragment level.

Using sliding window formulae, Theorem 8.23 shows that the successor predicate does not provide any additional expressive power in the presence of a neutral letter, *i.e.*, a letter that can be inserted and deleted anywhere without changing membership in the language. This has consequences for the algebraic counterparts of logic fragments. Proposition 8.26 shows how neutral letters can be used to obtain a defining logic fragment for the monoids in a semigroup variety $\mathbf{V}$ from the corresponding logic fragment for $\mathbf{V}$.

The second application of the theory of logic fragments analyzes the expressive power of the minimum and the maximum predicates. Theorem 9.2 shows that the minimum predicate can be dispensed with if a sufficiently long prefix of the word is known. Similarly, the maximum predicate is superfluous if the suffix is specified. Even though this may be intuitive, it is not *a priori* clear as the minimum predicate is available at any stage of the evaluation of the formula and not just on the outermost level. Note that the number of variables may be restricted and it is therefore not possible in general to store the minimal position in a new variable.

The last application gives a very general justification to the intuition that it is not worthwhile to quantify the same variable twice in succession. Using the abstract theory of logic fragments, Proposition 10.3 shows that any formula can be equivalently converted into a normalized form without such requantifications in a way that is compatible with most logic fragments.

Part II of this thesis turns to a famous logic fragment, namely two-variable first-order logic FO^2. More precisely, we study several variants of the *quantifier alternation hierarchy* within FO^2 from a combinatorial and from an algebraic point of view. In this two-variable setting the usual definition of quantifier alternation in terms of the prenex normal form does not work, since new variables are necessary to write a formula in prenex form. We take a direct approach and analyze the parse tree of the formula, *i.e.*, the canonical tree structure obtained from the inductive composition rules of formulae. The quantifier alternation hierarchy of first-order logic consists of the full levels FO_m and the half levels Σ_m. The full level FO_m is the first-order fragment with at most $m-1$ alternations between different quantifier types on any path in the parse tree. In particular, formulae in FO_m may start with universal quantification. The half level Σ_m is the fragment of FO_m comprising all formulae that start with existential quantifiers. The quantifier alternation hierarchy in FO^2 is the restriction of the first-order alternation hierarchy to two-variable formulae; *i.e.*, the half levels are given by $\Sigma^2_m = \Sigma_m \cap \mathrm{FO}^2$ and the full levels by $\mathrm{FO}^2_m = \mathrm{FO}_m \cap \mathrm{FO}^2$. Notice that alternation based on second-order quantifiers is not interesting by a result Thomas, which shows that every MSO-formula can be reformulated to use only one existential second-order quantifier [Tho82].

The restriction to two-variables in first-order logic is natural because, as a consequence of Kamp's theorem [Kam68], three variables already suffice to express all first-order properties, whereas the expressive power of the two-variable fragment drops. Moreover, quantifier alternation is an interesting complexity measure of logic formulae. For instance Thomas, and Pin and Weil showed that the alternation hierarchy in FO coincides with the famous dot-depth hierarchy of star-free languages [Tho82; PW02]. This dot-depth hierarchy is based on the number of alternations between concatenation and Boolean combinations in regular expressions that instead of a Kleene star may use a complement operation. It is one of the great open challenges of automata theory to prove or disprove the following conjecture: *Every level of the dot-depth hierarchy is decidable.* Rephrasing the conjecture in terms of logic it thus becomes: *All levels* FO_m *and* Σ_m *are decidable.*

There are actually two common variants of the dot-depth hierarchy. Both coincide with quantifier alternation in FO, albeit over different relational signatures: Brzozowski's dot-depth hierarchy corresponds to the signature $[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$, whereas the Straubing-Thérien hierarchy corresponds to $[<]$. The former was introduced by Cohen and Brzozowski [CB71], the latter by Straubing and Thérien [Str81; Thé81]. Both hierarchies are strict and exhaust the star-free languages [BK78].

After four decades of research, little progress has been achieved on proving or refuting this conjecture. What is known is that decidability in one variant reduces to decidability in the other [Str85; PW02]. But only the very first full level is known to be decidable by results due to Simon (for $\mathrm{FO}_1[<]$) and to Knast (for $\mathrm{FO}_1[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$); *cf.* [Sim75; Kna83]. Apart from that, only for the levels Σ_1 and Σ_2 decidability is known [Pin95; PW97; Arf91; GS08]; see also [PW02] and [KKL11] for alternative proofs. To date, decidability of all other levels remains an open problem.[4]

Knowledge gained on the alternation hierarchy in FO^2 might possibly help in better understanding the alternation hierarchy in FO. The study of the two-variable alternation hierarchy was started by Weis and Immerman, who provided a combinatorial description in terms of *rankers* of the full levels FO^2_m [WI09]. Rankers are simple programs to navigate on a word and are composed of single-step instructions. These instructions scan for the nearest position satisfying a specified alphabetic condition either to the left or to the right and move the cursor to that position. If all the instructions of a ranker can be followed on a word, then the ranker identifies a unique position of the word. Otherwise its position on the word is undefined. The precise mode of operation of the instructions depends on the predicates available to the FO^2-formulae. For the signature $[<]$ the nearest occurrence of a single letter is sought, whereas for the signature $[<, \mathrm{suc}]$ the nearest occurrence of a factor is sought. Weis and Immerman showed that the expressive power of the full levels $\mathrm{FO}^2_m[<]$ and $\mathrm{FO}^2_m[<, \mathrm{suc}]$ is completely determined by the ability to specify the relative order between the end positions of certain pairs of rankers. The relevant pairs of rankers are obtained by restricting the number of alternations between left-searches and right-searches.

Chapter 11 refines Weis and Immerman's result, contributing a ranker description of the half levels $\Sigma^2_m[<]$ and $\Sigma^2_m[<, \mathrm{suc}]$ in Theorem 11.3 and Theorem 11.16. When the half level Σ^2_m is compared to the next full level FO^2_m in terms of rankers, it loses the ability to express certain order relations between some ranker pairs. Both theorems are original contributions of this thesis. Ranker characterizations of the full levels follow as an immediate corollary to these descriptions: Corollary 11.7 gives the ranker description of FO^2_m due to Weis and Immerman, and Corollary 11.20 rectifies the erroneous characterization of Weis and Immerman for $\mathrm{FO}^2_m[<, \mathrm{suc}]$, which is disproved by Example 11.21.

[4] A paper that was published after this thesis was submitted shows that FO_2 and Σ_3 are also decidable [PZ14].

As a convenient intermediate result towards our ranker characterizations, we provide descriptions of the expressive power of the FO^2 alternation hierarchy in terms of temporal logic. This in particular subsumes results due to Etessami, Vardi, and Wilke, characterizing full two-variable first-order logic $\mathrm{FO}^2[<]$ by the temporal logic fragment $\mathrm{TL}[\mathrm{XF}, \mathrm{YP}]$ and $\mathrm{FO}^2[<, \mathrm{suc}]$ by $\mathrm{TL}[\mathrm{X}, \mathrm{XF}, \mathrm{Y}, \mathrm{YP}]$, see [EVW02]. This concludes the combinatorial study of the two-variable first-order alternation hierarchy, and we then consider its algebraic aspects.

Chapter 12 concludes this thesis and turns to the question whether definability within the levels of the alternation hierarchy is decidable. We answer this question in the affirmative for all half levels and for all full levels over $[<]$ and over any signature between $[<, \mathrm{suc}]$ and $[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$.

Following the algebraic scheme mentioned earlier, all these decidability result are obtained by effective conditions on the syntactic semigroup. Specifically, definability is characterized by simple identities that can be verified effectively on any given semigroup. Theorem 12.1 and Theorem 12.2 provide these identities for $\Sigma^2_m[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$ and $\mathrm{FO}^2_m[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$.

A versatile proof technique is developed to obtain a formula from the algebraic object, typically considered the hard part of establishing such algebraic characterizations. Its main ingredients are a string rewriting system to obtain a quotient semigroup, relativization methods for the logic to pass assumptions on to the quotient, and a way to control the rewriting steps so that they can be lifted back to the original semigroup. A similar technique was also applied to obtain characterizations for $\mathrm{FO}_1[<, \mathrm{suc}]$-definable languages of infinite words [KL11b], for the join of so-called $\mathcal{R}$-trivial monoids and $\mathcal{L}$-trivial monoids [KL12d], and for the joins of the Trotter-Weil hierarchy [KL12c]. See Kufleitner's habilitation thesis for a treatise on the Trotter-Weil hierarchy [Kuf13]. This proof technique also yields a comparatively short proof of Knast's result on dot-depth one, given in Section 12.2.3. All these instances rely on closure under complement, however. Section 12.3 refines the technique to work in the case of the half levels as well. An analogous refinement in the less technical case without the successor predicate will be presented in the forthcoming paper [FKL14]. For the converse direction, *i.e.*, that the logic satisfies the identity, the ranker characterizations of Chapter 11 are helpful. However, note that they do not yield decidability *per se*.

It turns out in Proposition 12.3 that for the other signatures that include the successor predicate, the presence or absence of minimum or maximum predicates only makes a difference for the lowest levels Σ^2_1 and FO^2_1. Theorems 12.4 to 12.9 give effective algebraic characterizations of these levels over the various signatures with the successor predicate.

Theorem 12.11 and Theorems 12.12 finally establish a similar algebraic characterization for the alternation hierarchy over the signature $[<]$ in terms of the syntactic monoid. These results are obtained by a reduction of the signature $[<]$ (involving monoids) to the signature $[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$ (involving semigroups) using the neutral letter approach developed in Part I of the thesis. To use this property, we have to provide a sliding window formula in $\Sigma^2_m[<]$ for any formula in $\Sigma^2_m[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$.

Not all the decidability consequences obtained by these results are new. For the levels FO^2_1 and Σ^2_1 over the relational signatures $[<]$ and $[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$ decidability is well-known in the literature [Sim75; Pin95; Kna83; PW97]. Only recently, Kufleitner and Weil associated the full levels $\mathrm{FO}^2_m[<]$ with decidable levels of the Trotter-Weil hierarchy [KW12]. Independently and at about the same time, characterizing identities for the full levels $\mathrm{FO}^2_m[<]$ were established by Krebs and Straubing [KS12b]. All other results are contributions of this thesis — in particular those for the half levels. The results for Σ^2_1 and FO^2_1 over the signatures $[<, \mathrm{suc}, \min, \max]$, $[<, \mathrm{suc}, \min]$, $[<, \mathrm{suc}, \max]$, and $[<, \mathrm{suc}]$ are published in [KL12a], and slightly more complicated identities for the full levels $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$ are published in [KL13]. The identities for the half levels $\Sigma^2_m[<]$ along with a direct proof that establishes them are published in [FKL14].

Organization of this thesis. To increase the visibility of the contributions among the technical details, the main ideas and results are stated at the beginning of each chapter, deferring proofs to later parts of the chapter.

Chapter 2 gives the most basic definitions and general terminology used throughout this thesis. Concepts with only local appearances are introduced at the respective locations. To increase readability, a list of mathematical symbols can be found on page 5 *et seq.*, right after the table of contents, and an index of terminology and concepts can be found on page 203 *et seq.* at the end of this thesis. The contributions of the thesis are divided into two parts.

Part I defines monadic second-order logic and logic fragments in Chapters 3 and 4, and afterwards develops the theory of logic fragments in Chapter 5. Chapters 6 and 7 exemplify the subtlety of the abstract notion of a logic fragment with two rather unorthodox fragments that have natural characterizations. Chapters 8, 9, and 10 contain applications of the abstract theory of logic fragments analyzing, respectively, the successor predicate, the minimum and maximum predicates, and a requantification-free normal form.

Part II investigates the quantifier alternation hierarchy within two-variable first-order logic over several signatures. Chapter 11 gives combinatorial descriptions of each level in this hierarchy in terms of rankers. Finally, Chapter 12 shows decidability of all levels in this hierarchy.

2. Preliminaries and Foundations

This chapter defines terminology for the most fundamental mathematical concepts found throughout this thesis.

Section 2.1 fixes common mathematical notation for numbers and especially intervals of numbers. Section 2.2 introduces formal language theoretic concepts such as words, languages, family of languages, and regular languages. Section 2.3 establishes the link between algebra and formal language theory. Concepts introduced include Green's relations in finite semigroup theory, recognition by semigroups and monoids, syntactic semigroups and monoids, identities of omega-terms, and varieties.

Readers familiar with these concepts may choose to skip this chapter, but should remember it when stumbling upon unknown notation. Section 2.3 is relevant mostly for Chapter 12 in Part II. Its concepts have only sporadic occurrences in Part I, which are marked as such and refer to this section. It might be skipped for the moment and read on demand, or just skimmed through by the more knowledgeable reader.

2.1. Numbers and Intervals

The set of all integers is $\mathbb{Z}$. The subset of natural numbers $\mathbb{N}$ consists of all non-negative integers including 0. An integer is positive if it is a natural number except 0. Apart from the usual non-strict and strict order $i \leqslant j$ and $i < j$ for integers $i, j \in \mathbb{Z}$ with their usual meaning, $i \ll j$ is an even stricter version that is true if $i \leqslant j - 2$. The closed interval $\{k \in \mathbb{Z} \mid i \leqslant k \leqslant j\}$ is denoted by $[i; j]$. The open interval excluding i and j is denoted by $(i; j)$ and as usual, the half-open variants are denoted by $[i; j)$ and $(i; j]$.

2.2. Words and Languages

The set of all finite words over the alphabet A is A^*. It is the free monoid under concatenation generated by A. Its neutral element is the *empty word* ε. The subsemigroup A^+ is the set $A^* \setminus \{\varepsilon\}$ of finite non-empty words. Let $u \in A^*$ be a word and let $a_i \in A$ be such that $u = a_1 \cdots a_n$. The *length* of u is $|u| = n$. A word w is a *factor* of u if there exist $p, q \in A^*$ such that $u = pwq$; it is a *prefix* if p is empty; and it is a *suffix* if q is empty. Viewing a word as a labeled linear order, the set of positions of u is $\{1, \ldots, |u|\}$. Let $u[i] = a_i$ for a position i, that is, $u[i]$ denotes the i^{th} letter of u. Extending this notation to $u[i; j]$ for $i, j \in \mathbb{Z}$, let $u[i; j]$ be the factor of u induced by the interval $[i; j] \cap [1; |u|]$. All non-position integers in $[i; j]$ are disregarded and $u[i; j] = \varepsilon$ whenever $[i; j] \cap [1; |u|] = \emptyset$. A word w is a *scattered subword* of u if there are positions $i_1 < \cdots < i_{|w|}$ of u such that $w[j] = u[i_j]$ for all $j \in \{1, \ldots, |w|\}$.

The *alphabet* $\mathrm{alph}(u)$ of u is the set $\{u[1], \ldots, u[|u|]\}$ of letters occurring in u. The alphabet is also known as the *content* of a word. Extending this notation, the *factor alphabet* $\mathrm{alph}_k(u)$ is the set $\{w \in A^k \mid u = pwq \text{ for some } p, q \in A^*\}$ of factors of length k occurring in u.

Any subset of A^* is a *language* of A^*. A *residual*[1] of a language $L \subseteq A^*$ is a language of the form $u^{-1}Lv^{-1} = \{w \in A^* \mid uwv \in L\}$ for some $u, v \in A^*$. It is a *left residual*, denoted by $u^{-1}L$, if $v = \varepsilon$, and a *right residual*, denoted by Lv^{-1}, if $u = \varepsilon$. For a function $h\colon B^* \to A^*$ the *inverse image* of L under h is the language $h^{-1}(L) = \{w \in B^* \mid h(w) \in L\}$ of B^*.

Regular languages constitute one of the most fundamental language families in formal language theory. A language is regular if it can be described by a *regular expression*; *i.e.*, it can be built from $\emptyset$, $\{\varepsilon\}$, and $\{a\}$ for $a \in A$ by means union $K \cup L$, concatenation $K \cdot L = \{uv \mid u \in K,\, v \in L\}$, and Kleene star $L^* = \{u_1 \cdots u_n \mid u_i \in L,\, n \geqslant 0\}$.

Regular languages were introduced and shown to be equivalent to deterministic finite automata by Kleene in his study of nerve nets [Kle56]. As deterministic automata are closed under complementation, so are regular languages. Consequently regular expressions extended by all Boolean combinations, *i.e.*, incorporating operations for complementation $A^* \setminus L$ and intersection $K \cap L$ in addition to union, also describe only regular languages.

On the other hand *extended regular expressions* are more succinct,[2] and moreover, they make a difference when considering restrictions of regular expressions. For instance, a language is *star-free* if it can be described by an *extended* regular expression without Kleene star. A similar definition with ordinary regular expressions is not interesting: Regular expressions without Kleene operation describe only finite languages.

Note that a language may well be star-free even if it is presented by a regular expression involving a Kleene star. Consider, for example, a language of the form $A_1^* a_1 \cdots A_n^* a_n A_{n+1}^*$ for some $A_i \subseteq A$ and $a_i \in A$ (a so-called monomial). Such languages are star-free, because $B^* = A^* \setminus \bigcup_{b \notin B} A^* b A^*$ for any $B \subseteq A$ and A^* is obtained as the complement of $\emptyset$. A deep result of Schützenberger characterizes star-free languages effectively [Sch65].

Varieties of languages. Varieties of languages introduced in this section are a vital concept for the classification of regular languages. In this context there is a difference whether or not the empty word may be present in languages. This distinction is substantial in certain contexts and leads to two related yet different branches when considering language families.

A *family of $*$-languages* $\mathcal{L}$ maps each finite alphabet A to a set $\mathcal{L}_A$ of languages of A^*. A family of $*$-languages $\mathcal{L}$ is closed under

- *positive Boolean combinations* if $\mathcal{L}_A$ is closed under finite intersections and finite unions; *i.e.*, $\emptyset, A^* \in \mathcal{L}_A$ and $L, K \in \mathcal{L}_A$ implies $(L \cup K), (L \cap K) \in \mathcal{L}_A$ for each A,
- *complementation* if $L \in \mathcal{L}_A$ implies $A^* \setminus L \in \mathcal{L}_A$ for each A,
- *Boolean combinations* if it is closed under complementation and positive Boolean combinations,
- *residuals* if $L \in \mathcal{L}_A$ implies $u^{-1}Lv^{-1} \in \mathcal{L}_A$ for each finite alphabet A and all $u, v \in A^*$,
- *inverses* of h for a map $h\colon B^* \to A^*$ if $L \in \mathcal{L}_A$ implies $h^{-1}(L) \in \mathcal{L}_B$.

[1] Residuals are also called *quotients* by other authors.

[2] More precisely, extended regular expressions are non-elementary more succinct [Sto74]. Even avoiding a single complementation in a regular expression of size n makes a size of 2^{2^n} for a describing regular expression necessary in the worst-case; *cf.* [GN08; GH08].

A *(positive) $*$-variety* is a family of $*$-languages closed under (positive) Boolean combinations, residuals, and inverses of each homomorphism $h\colon B^* \to A^*$ between finitely generated free monoids.

A *family of $+$-languages* $\mathcal{L}$ is a map that assigns each finite alphabet A with a set $\mathcal{L}_A$ of languages of A^+. All of the above closure properties are defined analogously with free monoids replaced by free semigroups. In particular, complements $A^+ \setminus L$ are with respect to A^+, residuals $u^{-1}Lv^{-1} = \{w \in A^+ \mid uwv \in L\}$ are restricted to non-empty words, and inverse closure is declared only under maps $h\colon B^+ \to A^+$ between free finitely generated semigroups.

A *(positive) $+$-variety* is a family of $+$-languages closed under (positive) Boolean combinations, residuals, and inverse homomorphisms $B^+ \to A^+$ between finitely generated free monoids.

Notice that there is a huge difference in the homomorphisms involved: Homomorphisms for $*$-varieties may delete certain letters by mapping them to the empty word, which is not possible for $+$-varieties. This difference will become crucial in Chapter 5.

2.3. Algebra and Recognition

A word in advance: The algebraic view on languages given in this section is often puzzling when seen for the first time. It is further complicated by the fact that it is necessary to consider semigroups and monoids as well as ordered variants thereof. The difference in these concepts, small as it might appear, is crucial in the later application. Algebraic language theory is a vast field with many intriguing insights. This section is barely able to scratch its surface, covering only the concepts necessary for this work. For further reading a standard textbook on finite semigroup theory such as [Pin86; Pin97] may be consulted.

A *semigroup* is a set S with an associative multiplication declared within S. The multiplication symbol is usually omitted, and we write xy instead of $x \cdot y$. It is customary to identify a semigroup with its constituting set whenever the multiplication is understood from context. A *monoid* is a semigroup that possesses a *neutral element*, usually denoted by 1. Let S^1 be the monoid obtained by adjoining a neutral element to S if necessary; if S is a monoid, then $S^1 = S$. An element $e \in S$ of a semigroup S is *idempotent* if $e^2 = e$. The set of all idempotents of S is $E(S)$. In a finite semigroup S there is always a positive integer ω such that x^ω is idempotent for all $x \in S$. Whenever it is necessary to make the semigroup explicit, the integer ω is henceforth denoted by ω_S. An *ideal* of S is a subset $I \subseteq S$ such that $S^1IS^1 \subseteq I$. Similarly, a *right ideal* (*left ideal*) of S is a subset $I \subseteq S$ such that $IS^1 \subseteq I$ (respectively, $S^1I \subseteq I$).

Ordered semigroups. An *ordered semigroup* $(S, \leqslant)$ is a semigroup S equipped with a compatible partial order $\leqslant$ on its elements; *i.e.*, the partial order $\leqslant$ satisfies $x \leqslant y$ and $s \leqslant t$ implies $xs \leqslant yt$. A *$\leqslant$-order ideal* of S is a subset $P \subseteq S$ that is downward closed for $\leqslant$; that is, $x \leqslant y$ and $y \in P$ implies $x \in P$. The order of an ordered semigroup is mostly understood implicitly. Note that any non-ordered semigroup can be ordered by equality.

Cayley graph. The *Cayley graph* of a semigroup S is an edge-labeled directed graph with vertex set S and with an z-labeled edge $x \xrightarrow{z} y$ from x to y if $z \in S^1$ is such that $x \cdot z = y$ or $z \cdot x = y$. The *right Cayley graph* and *left Cayley graph* of S are the subgraphs of the Cayley graph with vertex set S where in the former an edge $x \xrightarrow{z} y$ exists if $x \cdot z = y$ and in the latter $x \xrightarrow{z} y$ exists if $z \cdot x = y$.

Green's relations. A fundamental concept in the structure theory of finite semigroups are *Green's relations* on a semigroup S, which are defined as follows:

- Let $x \mathrel{\mathcal{R}} y$ if $xS^1 = yS^1$.
- Let $x \mathrel{\mathcal{L}} y$ if $S^1x = S^1y$.
- Let $x \mathrel{\mathcal{J}} y$ if $S^1xS^1 = S^1yS^1$.

Intuitively, these relations postulate inverse-like elements in a certain context. For example, $x \mathrel{\mathcal{R}} y$ means that there are local right inverse elements $\bar{x}$ and $\bar{y}$ in S^1 such that $x\bar{x} = y$ and $y\bar{y} = x$. Another way to think of it is that x and y are in the same strongly connected component of the right Cayley graph of the semigroup; *i.e.*, there is a path from x to y and back to x. Analogously the relation $\mathcal{L}$ corresponds to strongly connected components in the left Cayley graph, and $\mathcal{J}$ corresponds to strongly connected components in the full Cayley graph.

There are natural orders on the equivalence classes of Green's equivalence relations, which are defined as follows:

- Let $x \leqslant_{\mathcal{R}} y$ if $x \in yS^1$.
- Let $x \leqslant_{\mathcal{L}} y$ if $x \in S^1y$.
- Let $x \leqslant_{\mathcal{J}} y$ if $x \in S^1yS^1$.

In terms of the right Cayley graph, the relation $x \leqslant_{\mathcal{R}} y$, for example, requires that there be a path from y to x; there need not be a path back from x to y, however. For $\mathcal{G} \in \{\mathcal{R}, \mathcal{L}, \mathcal{J}\}$ we have $x \mathrel{\mathcal{G}} y$ if and only if $x \leqslant_{\mathcal{G}} y$ and $y \leqslant_{\mathcal{G}} x$. The strict versions of Green's orders are defined as usual by setting $x <_{\mathcal{G}} y$ if $x \leqslant_{\mathcal{G}} y$ and $y \nleqslant_{\mathcal{G}} x$.

Homomorphisms and algebraic divisors. An *algebra* in the context of this thesis means an arbitrary set with an underlying semigroup or monoid structure, possibly ordered. A *homomorphism* between two algebras is a structure-preserving mapping. Specifically, a *semigroup homomorphism* between semigroups S and T is a mapping $h\colon S \to T$ such that $h(xy) = h(x)h(y)$. A *monoid homomorphism* between monoids M and N is a mapping $h\colon M \to N$ such that $h(xy) = h(x)h(y)$ and $h(1) = 1$. In the presence of an order on the underlying algebras, a homomorphism additionally has to preserve the order; *i.e.*, $x \leqslant y$ implies $h(x) \leqslant h(y)$.

An algebra T is a *subalgebra*[3] of S if there exists an homomorphic inclusion map $T \hookrightarrow S$. It is a *quotient* of S if there exists a homomorphic surjection $S \twoheadrightarrow T$. It is a *divisor* of S if it is a quotient of a subalgebra.

Note that these definitions depend on the structure of the underlying algebra and its associated homomorphisms. A subsemigroup, *e.g.*, is a subset of the semigroup that is closed under multiplication; a submonoid additionally contains the neutral element.

[3]That is to say, an (ordered) *subsemigroup* or an (ordered) *submonoid*, depending on the concrete structure considered.

Algebras as recognizers. The reason for considering semigroups and monoids as well as their ordered variants is that each of these algebraic structures has different recognition capabilities when it comes to languages.

A semigroup S *recognizes* a language $L \subseteq A^+$ if there exists a semigroup homomorphism $h\colon A^+ \to S$ such that $L = h^{-1}(P)$ for some *recognizing set* $P \subseteq S$. In other words, $u \in L$ if and only if $h(u) \in P$. For languages of A^* that a priori may include the empty word, monoids are the suitable recognizers. A monoid M *recognizes* a language $L \subseteq A^*$ if there exists a monoid homomorphism $h\colon A^* \to M$ such that $L = h^{-1}(P)$ for some subset $P \subseteq M$. A way to think about a recognizing homomorphism is as a reduction of the membership problem in L: To decide membership of $u = a_1 \cdots a_n$ in L, compute $h(u) = h(a_1) \cdots h(a_n)$ and check for membership in P. For finite semigroups or monoids this is always effective.

Note that if a language is recognized by a semigroup, then its complement is recognized by the same semigroup. For families of languages that are not closed under complementation this is not desirable. Recognition by ordered semigroups and ordered monoids solves this issue. An ordered semigroup $(S, \leqslant)$ *recognizes* a language $L \subseteq A^+$ if there exists a semigroup homomorphism $h\colon A^+ \to S$ such that $L = h^{-1}(P)$ for some $\leqslant$-order ideal P of S. Similarly, an ordered monoid $(M, \leqslant)$ *recognizes* a language $L \subseteq A^*$ if there exists a monoid homomorphism $h\colon A^* \to M$ such that $L = h^{-1}(P)$ for some $\leqslant$-order ideal P of M.

Syntactic recognizers. A canonical way to obtain quotients of algebras is by means of congruences. A *congruence* $\cong$ on an algebra M is a compatible equivalence relation; *i.e.*, an equivalence relation such that $x \cong y$ and $s \cong t$ implies $xs \cong yt$. The structure of M induces the same algebraic structure on the set equivalence classes $M/{\cong}$.

A classical example of such a relation is the *syntactic congruence* $\equiv_L$ of a language $L \subseteq A^*$ defined by $u \equiv_L v$ whenever $puq \in L$ if and only if $pvq \in L$ for all $p, q \in A^*$. The relation $u \equiv_L v$ means that L cannot distinguish the words u and v in any context. The *syntactic monoid* of L, denoted by M_L, is the quotient $A^*/{\equiv_L}$. The *syntactic (monoid) homomorphism* of L, denoted by $h_L\colon A^* \to M_L$, is the canonical homomorphism, mapping u to its $\equiv_L$-class.

The *syntactic preorder* of L is defined by $u \leqslant_L v$ if $pvq \in L$ implies $puq \in L$ for all $p, q \in A^*$. The syntactic preorder canonically induces a partial order on the equivalence classes of the syntactic congruence, and the syntactic monoid becomes an ordered monoid.

The *syntactic semigroup* of L, denoted by S_L, is the subsemigroup $h(A^+)$ of M_L. It inherits the order of the syntactic monoid. The *syntactic (semigroup) homomorphism* of L is denoted by $h_L\colon A^+ \to S_L$ and defined canonically. The syntactic semigroup is largely the same as the syntactic monoid, but it has no neutral element if only the empty word maps to the neutral element. Consequently M_L and S_L^1 are isomorphic.

It is well-known that the syntactic monoid is a divisor of every recognizing monoid. It is the minimal recognizer, which is unique up to isomorphism. It is the algebraic analogon of the minimal automaton of a language. Another classical fact is that the syntactic monoid M_L is finite if and only if L is regular. Indeed, it is the so-called transition monoid of the minimal automaton and as such computable from any reasonable presentation of a regular language, including but not limited to, presentations in terms of regular expressions, deterministic or nondeterministic finite automata, monoids, or monadic second-order formulae. Confer Pin's textbook [Pin86] for these classical results. As this thesis is concerned with regular languages, most semigroups and monoids considered here are finite, with the exception of their free variants.

Varieties of semigroups and varieties of monoids. A *variety of finite algebras* is a class of finite algebras that is closed under taking divisors and finite direct products. The notion of a divisor depends on the concrete algebraic structure considered, leading to different notions of varieties. In the context of this thesis this specifically defines the concretization of a *variety of (ordered) semigroups* as well as of a *variety of (ordered) monoids*.[4]

The reason why varieties of finite algebras are interesting for the purpose of this thesis is because they are closely tied up with varieties of languages, which makes them an important tool in the classification of regular languages. Indeed, for every variety of languages $\mathcal{V}$ there exists a variety of finite algebras $\mathbf{V}$ such that a language belongs to $\mathcal{V}$ if and only if its syntactic algebra belongs to $\mathbf{V}$. Of course, the precise form of the correspondence depends on the structure of the algebra: Semigroups correspond to $+$-languages, monoids correspond to $*$-languages, and varieties of ordered algebras correspond to positive varieties of languages.[5]

Membership of a language in $\mathcal{V}$ thus amounts to the same as membership of its syntactic algebra in $\mathbf{V}$. As the latter often boils down to a verification of some simple formal identity within the syntactic algebra, this correspondence between $\mathcal{V}$ and $\mathbf{V}$ often becomes the key to many decidability results about the membership problem for varieties of languages. This is true in particular for the decidability results covered in Chapter 12.

With respect to the necessary computations to decide membership of a finite algebra in a variety, it is well-known that actually *every* such variety is defined by a finite set of formal identities [Rei82; Kun03; PS05]; see also [Wei02; Alm94]. The problem with these identities is that their formulation is in a so-called profinite extension, which can become extremely complicated. Indeed, there is a finite set of identities whose verification on a given input semigroup is undecidable [ABR92]. For the purpose of this thesis, the full formalism of such identities is not necessary. We restrict ourselves to identities of a simple form, namely identities of omega-terms, that can always be verified.

Identities of omega-terms. Let Ξ be an infinite set of variables. An *omega-term* over Ξ is a word over Ξ with an additional operator for a formal ω-power. More precisely, omega-terms are given by the following inductive definition: Every x in Ξ is an omega-term. If U and V are omega-terms, then UV and U^ω are also omega-terms. Note that ω is just a formal symbol in this context, not some particular power. An *identity of omega-terms* is a formal equality $U = V$ or a formal inequality of the

[4]Some authors prefer to write *pseudovariety* instead of *variety* to distinguish them from varieties of infinite algebras in Birkhoff's sense, *cf.* [Bir35]. This is not necessary in the context of this thesis.

[5]The connection between varieties of languages and varieties of finite algebras is actually even tighter than described here. This is commonly called Eilenberg correspondence; *cf.* [Pin97]. We do not elaborate on this, as it is not used in this thesis.

form $U \leqslant V$. For an omega-term U and a positive integer n let $U^{(n)}$ denote the word over Ξ that is obtained from U by replacing every occurrence of the formal symbol ω by the concrete power n.

A semigroup S *satisfies* the identity $U = V$ if it is finite and $h(U^{(\omega_S)}) = h(V^{(\omega_S)})$ holds for every homomorphism $h\colon \Xi^+ \to S$, where ω_S is the idempotent generating power from above. An ordered semigroup S satisfies the identity $U \leqslant V$ if it is finite and $h(U^{(\omega_S)}) \leqslant h(V^{(\omega_S)})$ holds for every homomorphism $h\colon \Xi^+ \to S$. This means that ω-powers are interpreted as the idempotent generated by their base, and that the (in)equality has to be true no matter what values are inserted for the variables. For example, a semigroup is commutative if and only if it satisfies the identity $xy = yx$.

For monoids, we also allow the empty word over Ξ as an omega-term, which is denoted by 1. A monoid M *satisfies* $U = V$ if it is finite and $h(U^{(\omega_M)}) = h(V^{(\omega_M)})$ for each homomorphism $h\colon \Xi^* \to M$. An ordered monoid M satisfies $U \leqslant V$ if it is finite and $h(U^{(\omega_M)}) \leqslant h(V^{(\omega_M)})$ for each homomorphism $h\colon \Xi^* \to M$.

Selected varieties. The most important varieties for the purpose of this thesis are the following:

- $\mathbf{A}$ consists of all *aperiodic monoids* satisfying $x^{\omega+1} = x^{\omega}$.
- $\mathbf{DA}$ consists of all monoids satisfying $(xy)^{\omega} = (xy)^{\omega}x(xy)^{\omega}$.
- $\mathbf{LDA}$ consists of all semigroups S such that $eSe \in \mathbf{DA}$ for all idempotents e.
- $\mathbf{D}$ consists of all *definite semigroups* satisfying $yx^{\omega} = x^{\omega}$.

The set eSe occurring in the definition of $\mathbf{LDA}$ is the *local monoid* of S at e. It is a subsemigroup of S that forms a monoid with e as neutral element. More generally, for each variety of monoids $\mathbf{V}$ let $\mathbf{LV}$ be the variety of all semigroups S that are *locally* $\boldsymbol{V}$; that is, $S \in \mathbf{LV}$ if $eSe \in \mathbf{V}$ for each idempotent e.

In terms of identities, the variety $\mathbf{LDA}$ comprises all semigroups that satisfy the identity $(exeye)^{\omega} = (exeye)^{\omega}x(exeye)^{\omega}$, where $e = z^{\omega}$.

Note that if $M \in \mathbf{DA}$, then

$$(xy)^{\omega} = (xy)^{\omega-1}x(yx)^{\omega}y = (xy)^{\omega-1}x(yx)^{\omega}y(yx)^{\omega}y = (xy)^{2\omega}y(xy)^{\omega} = (xy)^{\omega}y(xy)^{\omega}$$

for all $x, y \in M$. The second equality uses the $\mathbf{DA}$-identity, all other equalities follow from associativity and idempotence. This in particular shows that the varieties $\mathbf{DA}$ and $\mathbf{LDA}$ are left-right symmetric, despite their *a priori* asymmetric definition.

Part I.

Logic Fragments

3. Monadic Second-Order Logic over Words

That that is is that that is is not not.

— Unknown origin

This chapter defines *monadic second-order logic* interpreted over finite words. The logic framework introduced in this chapter is used throughout this thesis. It in particular sets the stage for the central notion of a logic *fragment* as defined in the next chapter. In the following, syntax and semantics of formulae are defined, and notation to speak about truth values and about formulae with free variables is fixed. The focus is on providing a profound intuition about logic formulae, not to give all formal details of the definitions. A rigorous mathematical definition can be found in Appendix A on page 209.

3.1. Formulae and Formal Languages

In the context of logic, words are viewed as labeled linear orders, where positions are indexed by consecutive natural numbers, starting with 1 as the minimal position. Labels come from an infinite universe of letters Λ. Typical symbols to denote labels are a, b, c. An *alphabet* A is henceforth any subset of Λ.[1] The set of variables is $\mathbb{V} = \mathbb{V}_1 \dot\cup \mathbb{V}_2$, where $\mathbb{V}_1$ is an infinite set of first-order variables, and $\mathbb{V}_2$ is a disjoint infinite set of second-order variables. First-order variables (generally denoted by lowercase letters like $x, y, x_i \in \mathbb{V}_1$) range over positions of the word, whereas second-order variables (denoted by uppercase letters like $X, Y, X_i \in \mathbb{V}_2$) range over subsets of positions. Atomic formulae are

- the constants $\top$ for true and $\perp$ for false,
- the 0-ary predicate "empty", which is true only for the empty word,
- the label predicate $\lambda(x) \in B$ for $B \subseteq \Lambda$, which is true if x has a label in B, and
- the second-order predicate $x \in X$, which is true if x is contained in X

as well as the following numerical predicates:

- the equality predicate $x = y$,
- the strict and non-strict order predicates $x < y$ and $x \leqslant y$,
- the successor predicate $\mathrm{suc}(x, y)$, which is true if $x + 1 = y$,
- the minimum and maximum predicate $\min(x)$ and $\max(x)$, which are true if x is the first or the last position, respectively, and
- the modular predicate $x \equiv r \pmod q$, where $r, q \in \mathbb{Z}$, which is true if the position of x is congruent to r modulo q.

[1] Readers familiar with formal languages may wonder why we use an "infinite alphabet" of labels Λ. Anticipating the application to language theory, the idea is that we want to have a uniform logic framework *for all* finite alphabets A. If only finitely many labels were allowed, we would be forced to re-define formulae for every alphabet, leading to clumsy notation. See also Remark 3.1 for a discussion about labels used by a formula.

Formulae φ and ψ can be composed by the Boolean connectives

- negation $\neg\varphi$,
- disjunction $\varphi \vee \psi$ and conjunction $\varphi \wedge \psi$,

by the common first-order and second-order quantifiers, namely

- existential and universal first-order quantification $\exists x\, \varphi$ and $\forall x\, \varphi$, and
- existential and universal monadic second-order quantification $\exists X\, \varphi$ and $\forall X\, \varphi$

with their usual meanings as well as by

- modular counting quantification $\exists^{r \bmod q} x\, \varphi$, where $r, q \in \mathbb{Z}$, which is true if there are r positions (modulo q) for x which make φ true.

The parameter r in the modular predicate $x \equiv r \pmod q$ and in the modular counting quantifier $\exists^{r \bmod q}$ is the *remainder* and q is the *modulus*.

For convenience we let $\lambda(x) = a$ for $a \in A$ be an abbreviation for $\lambda(x) \in \{a\}$, thus associating it with its usual meaning that x is an a-position. Parentheses may be used to disambiguate or to increase readability. As usual for a finite index set I we let $\bigvee_{i \in I} \varphi_i$ and $\bigwedge_{i \in I} \varphi_i$ be, respectively, the disjunction and conjunction of all formulae φ_i. For the empty index set this disjunction is $\bot$, whereas the conjunction is $\top$. We use the symbol $:=$ to denote syntactic equality of formulae to distinguish it from the equality sign occurring in atomic formulae.

The above list conveyed the meaning of the respective formulae right away, albeit on a very informal level. Remember that a thorough mathematical definition is given in Appendix A. Note that we also allow a modulus $q = 0$, in which case the modular predicate $x \equiv r \pmod q$ degenerates to the equality $x = r$, and the counting quantifier $\exists^{r \bmod q} x\, \varphi$ counts the exact number of positions making φ true. In particular, $\exists^{1 \bmod 0} x\, \varphi$ is the uniqueness quantification, often denoted by $\exists! x\, \varphi$, requiring that there be one and only one position making φ true.

Tracing the inductive rules to build formulae canonically associates a rooted ordered tree of degree 2 with each formula, the so-called *parse tree* of the formula. Atomic formula correspond to leafs, disjunction and conjunction are inner nodes with two children, and negations and quantifiers are inner nodes with one child. Each node of the tree is labeled by the respective construct it corresponds to.

The set $\mathrm{FV}(\varphi)$ of *free variables* of φ is defined as usual. A *sentence* is a formula without free variables. Let $u \models \varphi$ if φ is a sentence that is true on the word $u \in \Lambda^*$. The symbol $u \models \varphi$ may be read as "*u models* φ" or "φ is *true* on u" or "*u satisfies* φ".

Logic and families of formal languages. Depending on the alphabet, a sentence φ may define different languages. For a finite alphabet $A \subseteq \Lambda$ and a sentence φ, the *language defined* by φ over A is the set $\mathcal{L}_A(\varphi) = \{u \in A^* \mid u \models \varphi\}$. If the alphabet is clear from context (*e.g.*, if it is fixed), we also write $L(\varphi)$ instead of $\mathcal{L}_A(\varphi)$.

Let $\mathcal{F}$ be a set of formulae. An $\mathcal{F}$*-formula* is a formula in $\mathcal{F}$. Denote by $\mathcal{L}(\mathcal{F})$ the *family of languages defined* by $\mathcal{F}$. It maps each finite alphabet $A \subseteq \Lambda$ to the set $\mathcal{L}_A(\mathcal{F})$ consisting of all languages $\mathcal{L}_A(\varphi)$ defined by some $\mathcal{F}$-formula φ.

A language $L \subseteq A^*$ is *definable in* $\mathcal{F}$ if $L \in \mathcal{L}_A(\mathcal{F})$. Let $\mathcal{G}$ be any family of languages. The family of languages *defined by* $\mathcal{F}$ *over* $\mathcal{G}$ maps A to $\mathcal{L}_A(\mathcal{F}) \cap \mathcal{G}_A$. The family of languages *defined by* $\mathcal{F}$ *over non-empty words* maps A to $\mathcal{L}_A(\mathcal{F}) \cap A^+$.

Remark 3.1 (On label symbols used by a formula)
According to the definition, the atom $\lambda(x) \in B$ may involve infinite sets of labels. As soon as a finite alphabet is fixed, we may assume all sets of labels to be finite. Suppose some finite alphabet A is fixed over which structures are built. Replacing $\lambda(x) \in B$ by $\lambda(x) \in (A \cap B)$ yields a formula which is equivalent for all structures over A.

Observe that the alphabet A may well be different from the set of labels used in a formula φ. For example, consider the formula $\exists x\ \lambda(x) = a$ requiring that there be an a-position. If $a \notin A$, then $\mathcal{L}_A(\varphi) = \emptyset$. Otherwise φ defines the language $A^* a A^*$.

This may seem odd at first glance. However, it allows to deal with languages over different alphabets in a uniform way, in particular when considering closure under inverse homomorphisms in Section 5.5. $\diamond$

3.2. Formulae with Free Variables

We introduce notation for formulae with free variables which will come in handy when phrasing results about logic fragments.

As usual, we write $\varphi(x_1, \dots, x_k, X_1, \dots, X_\ell)$ to indicate that no variables other than mentioned appear freely in φ; that is, $\mathrm{FV}(\varphi) \subseteq \{x_1, \dots, x_k, X_1, \dots, X_\ell\}$. To ascertain truth in the presence of free variables, these have to be assigned an *interpretation*. This interpretation is encoded into an additional second component, leading to the *extended alphabet* $\Lambda \times 2^V$, where V contains all free variables of the formula. The first component specifies the underlying word over Λ. The *universe of structures* with free variables V is denoted by $\mathcal{U}_V$. It consists of all words over the extended alphabet $\Lambda \times 2^V$ with a well-defined interpretation for all first-order variables. Specifically, for a word u in $\mathcal{U}_V$ we require that for each first-order variable $x \in V$ there is a unique position i containing x in the second component of its label. The interpretation of x, denoted by $x(u)$, is this unique index i. For second-order variables we do not impose any restriction whatsoever, and the interpretation of X, denoted by $X(u)$, is the (possibly empty) set of all positions i containing X in the second component of their label.

Let $u \in \mathcal{U}_V$. Every formula φ whose free variables are contained in V can be assigned a well-defined truth value on u as usual. The *formal semantics* $[\![\varphi]\!]_V$ of φ (with respect to V) is the set of all words $u \in \mathcal{U}_V$ on which φ is true. Note that for a sentence φ we have $a_1 \cdots a_n \models \varphi$ if and only if $(a_1, \emptyset) \cdots (a_n, \emptyset) \in [\![\varphi]\!]_\emptyset$. We are also interested in the behavior of formulae on words over a concrete alphabet $A \subseteq \Lambda$. Let $[\![\varphi]\!]_{A,V}$ be the *formal semantics over* A, containing all models in $[\![\varphi]\!]_V$ whose first components only contain labels in A, *i.e.*; we let $[\![\varphi]\!]_{A,V} = [\![\varphi]\!]_V \cap \left(A \times 2^V\right)^*$.

The following substitution notations $u[X/I]$ and $u[x/i]$ are useful to manipulate the interpretation of X and x; the underlying word in the first component and the interpretation of all other variables remain unchanged. For $X \in \mathbb{V}$ the interpretation of X on the word $u[X/I]$ is the set of all positions of u contained in I. For a first-order variable x we write $u[x/i]$ instead of $u[x/\{i\}]$; *i.e.*, the variable x is mapped to the position i. Note that if i is not a position of u, then x is contained in no second component.

We extend the truth value relation $u \models \varphi$ to words u over the extended alphabet $\Lambda \times V$ by letting $u \models \varphi$ if $u \in [\![\varphi]\!]_V$. It is sometimes convenient to give the interpretation explicitly. For $u \in \Lambda^*$ we write $u, i_1, \dots, i_k, I_1, \dots, I_\ell \models \varphi(x_1, \dots, x_k, X_1, \dots, X_\ell)$ to signify that φ is true on u with x_j interpreted by i_j and X_j interpreted by I_j. More precisely, if $V = \{x_1, \dots, x_k, X_1, \dots, X_\ell\}$, then this truth value is defined by

$$(u[1], \emptyset) \cdots (u[|u|], \emptyset)[x_1/i_1, \dots, x_k/i_k, X_1/I_1, \dots, X_\ell/I_\ell] \in [\![\varphi]\!]_V.$$

4. Logic Fragments

This chapter defines the central notion of this first part, namely *fragments*. Fragments are subsets of formulae with natural closure properties on the syntactic level. Chapter 5 will show how these syntactic closure properties lead to natural semantic closure properties. Section 4.1 and Section 4.2 define important and widespread instances of fragments like first-order logic or the quantifier alternation hierarchy.

To define fragments, we first introduce contexts which allow to conveniently formalize subformulae. A *context*, typically denoted by the symbol μ, is a formula with a unique occurrence of an additional 0-ary predicate $\circ$. The symbol $\circ$ is to be read as "hole" and is a placeholder for a formula. Let $\mu(\varphi)$ be the result of substituting φ for the unique occurrence of $\circ$ in μ. A formula ψ is a *subformula* of φ if there exists a context μ such that $\varphi = \mu(\psi)$. In terms of the parse tree this means that ψ is a subtree of φ as depicted in Figure 4.1.

Definition 4.1 (Fragment)
A prefragment $\mathcal{F}$ *is a non-empty set of formulae that for all contexts μ, all formulae φ and ψ, and all first-order variables $x \in \mathbb{V}_1$ complies with the following axioms:*

1. *If $\mu(\varphi) \in \mathcal{F}$, then $\mu(\top) \in \mathcal{F}$ and $\mu(\bot) \in \mathcal{F}$.*
2. *$\mu(\varphi \vee \psi) \in \mathcal{F}$ if and only if $\mu(\varphi) \in \mathcal{F}$ and $\mu(\psi) \in \mathcal{F}$.*
3. *$\mu(\varphi \wedge \psi) \in \mathcal{F}$ if and only if $\mu(\varphi) \in \mathcal{F}$ and $\mu(\psi) \in \mathcal{F}$.*
4. *If $x \notin \mathrm{FV}(\varphi)$ and either $\mu(\exists x\ \varphi) \in \mathcal{F}$ or $\mu(\forall x\ \varphi) \in \mathcal{F}$, then $\mu(\varphi) \in \mathcal{F}$.*

A prefragment $\mathcal{F}$ is a fragment *if it complies with the additional axiom*

5. *If $\mu(\varphi) \in \mathcal{F}$ and $\mathrm{FV}(\mu(\lambda(x) \in B)) \subseteq \mathrm{FV}(\mu(\varphi))$, then $\mu(\lambda(x) \in B) \in \mathcal{F}$*

for every context μ, every formula φ, every $B \subseteq \Lambda$, and every variable $x \in \mathbb{V}_1$. A set of formulae $\mathcal{F}$ is closed under negation *if $\varphi \in \mathcal{F}$ implies $\neg\varphi \in \mathcal{F}$.*

Axiom (1) means that a subformula may always be replaced by $\top$ or by $\bot$ without leaving the prefragment. The implication from left to right in axioms (2) and (3) allow

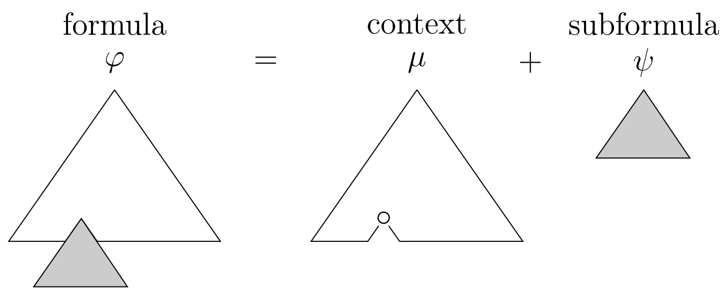

Figure 4.1.: Illustration of ψ being a subformula of φ: The formula φ can be obtained by substituting the formula ψ for the unique occurrence of the placeholder $\circ$ in some context μ.

to prune one of the branches of disjunctions and conjunctions within a prefragment. The converse implication allows to recombine formulae in a context whenever both formulae already appear in this context. Axiom (4) states that whenever a formula is independent of the interpretation of x, we may as well delete a preceding quantification $\exists x$ or $\forall x$ without leaving the prefragment. Within a fragment all these properties also hold. In addition, axiom (5) ensures that a subformula can always be replaced by a label atom — provided this operation does not introduce additional free variables. Note that all axioms ensure that replacing a formula does not introduce new free variables.

It is a recurrent theme of the abstract theory of logic fragments that some formula φ can replace another formula ψ in any context in a way that respects $\mathcal{F}$. We introduce a notation for this in the following. Let $\mathcal{F}$ be a set of formulae. The *syntactic preorder* of $\mathcal{F}$, denoted by $\leqslant_{\mathcal{F}}$, is defined by $\varphi \leqslant_{\mathcal{F}} \psi$ if φ and ψ are formulae and $\mu(\psi) \in \mathcal{F}$ implies $\mu(\varphi) \in \mathcal{F}$ for every context μ.

One can intuitively think of $\varphi \leqslant_{\mathcal{F}} \psi$ as the syntactic structure of φ being not more "complex" with respect to $\mathcal{F}$ than that of ψ. This relation allows to rephrase some of the axioms of a fragment. The following conditions necessarily hold if $\mathcal{F}$ is a prefragment:

- $\top \leqslant_{\mathcal{F}} \varphi$ and $\perp \leqslant_{\mathcal{F}} \varphi$,
- $\varphi \leqslant_{\mathcal{F}} (\varphi \vee \psi)$ and $\psi \leqslant_{\mathcal{F}} (\varphi \vee \psi)$,
- $\varphi \leqslant_{\mathcal{F}} (\varphi \wedge \psi)$ and $\psi \leqslant_{\mathcal{F}} (\varphi \wedge \psi)$, and
- $\varphi \leqslant_{\mathcal{F}} (\exists x\ \varphi)$ and $\varphi \leqslant_{\mathcal{F}} (\forall x\ \varphi)$ if $x \notin \mathrm{FV}(\varphi)$

for all formulae φ, ψ. Let us hasten to add that these conditions are *not sufficient* for a prefragment. Indeed, the syntactic preorder does not allow to formalize that $\mu(\varphi) \in \mathcal{F}$ and $\mu(\psi) \in \mathcal{F}$ implies $\mu(\varphi \wedge \psi) \in \mathcal{F}$ and $\mu(\varphi \vee \psi) \in \mathcal{F}$. This is because $\leqslant_{\mathcal{F}}$ inherently makes a statement about *all* contexts, but the recombination requires that φ and ψ appear in the *fixed* context μ. On the other hand, if $\varphi_1 \leqslant_{\mathcal{F}} \varphi$ and $\varphi_2 \leqslant_{\mathcal{F}} \varphi$, then also $(\varphi_1 \vee \varphi_2) \leqslant_{\mathcal{F}} \varphi$ and $(\varphi_1 \wedge \varphi_2) \leqslant_{\mathcal{F}} \varphi$. This fact shall tacitly be used a lot in later chapters.

For a fragment $\mathcal{F}$ we furthermore have

$$(\lambda(x) \in B) \leqslant_{\mathcal{F}} \varphi$$

for all $x \in \mathbb{V}_1$, all $B \subseteq \Lambda$, and all formulae φ with $\mathrm{FV}(\mu(\lambda(x) \in B)) \subseteq \mathrm{FV}(\mu(\varphi))$. In other words, whenever no additional free variable is introduced, any subformula may be replaced by a label predicate.

We immediately get closure under disjunction and conjunction for prefragments on the syntax level. On the semantic level this means that the language families defined by prefragments are closed under finite union and finite intersection. The following lemma records this for later reference.

Lemma 4.2
Let $\mathcal{F}$ be a prefragment. If $\varphi \in \mathcal{F}$ and $\psi \in \mathcal{F}$, then $(\varphi \vee \psi) \in \mathcal{F}$ and $(\varphi \wedge \psi) \in \mathcal{F}$.

Proof. This is a straightforward application of axioms (2) and (3) in Definition 4.1 using the context $\circ$. Note that $\circ(\varphi) = \varphi$ and $\circ(\psi) = \psi$. $\square$

4.1. Selected Logic Fragments

We write "FO" for the first-order quantifiers, "MSO" for the first-order and monadic second-order quantifiers, and "MOD" for the modular counting quantifiers. This yields the following fragments:

- MSO+MOD is the set of all formulae.
- FO+MOD is the set of all first-order formulae including quantifiers $\exists^{r \bmod q}$.
- MSO is the set of all formulae without quantifiers $\exists^{r \bmod q}$.
- FO is the set of all first-order formulae without quantifiers $\exists^{r \bmod q}$.

A common way to get new fragments is to restrict the allowed predicates. Any subset $\mathcal{N}$ of the set of numerical predicates $\{<, \leqslant, =, \mathrm{mod}, \mathrm{suc}, \min, \max, \mathrm{empty}\}$ is called a *relational signature* or just *signature*. Here, "mod" stands for the atomic formulae $x \equiv r \pmod q$. A set of formulae $\mathcal{F}$ is said to use some specific predicate (like the successor predicate) if there is a formula in $\mathcal{F}$ which uses this predicate.

For a set of formulae $\mathcal{F}$ and a signature $\mathcal{N}$, we denote by $\mathcal{F}[\mathcal{N}]$ the restriction to formulae in $\mathcal{F}$ that, apart from $\top$, $\bot$, labels, and atomic formulae of the form $x \in X$, only use predicates in $\mathcal{N}$. This notation is refined as follows. For a collection $\mathcal{N}$ of atomic formulae, let $\mathcal{F}[\mathcal{N}]$ be the set of formulae in $\mathcal{F}$ which, apart from $\top$, $\bot$, labels, and atomic formulae of the form $x \in X$, only uses atomic formulae in $\mathcal{N}$.

The latter restriction with respect to a fixed set of atomic formulae (instead of to a fixed set of predicates, which can be instantiated arbitrarily), is not yet common in the literature. We feel, however, that interesting sets of formulae might be obtained in this way, for which simple predicate restrictions are too coarse. What is more, it fits well in the fine-grained theory of logic fragments developed in this thesis. For example, FO[<] consists of all first-order formulae only using atomic formulae of the form $\top$, $\bot$, $\lambda(x) \in B$, and $x < y$ for arbitrary $x, y \in \mathbb{V}_1$. In contrast, for fixed $x, y \in \mathbb{V}_1$, the formulae in FO$[x < y]$ are only allowed to use atomic formulae $\top$, $\bot$, $\lambda(x) \in B$, and $x < y$; no comparison between any other first-order variables is allowed.

4.2. Quantifier Depth, Quantifier Alternation, and Number of Variables

A traditional way to obtain fragments, particularly for first-order formulae, is to restrict the quantification resources. For example, one can allow only r different variables, a quantifier depth of at most n, or at most $m-1$ quantifier alternations. We render these notions more precisely in the following.

Let FO^r be the fragment of FO which uses at most r variables. It is known that the fragment FO^3 with three variable names is already expressively complete for full first-order logic FO [Kam68], whereas FO^2 is strictly less expressive than FO and has a huge number of interesting characterizations, *cf.* [DGK08] for an overview. The *quantifier depth* $\mathrm{qd}(\varphi)$ of a formula φ is the maximal number of nested quantifiers on any root-leaf-path in the parse tree of φ. More precisely let $\mathrm{qd}(\varphi) = 0$ if φ is atomic and recursively let $\mathrm{qd}(\neg\varphi) = \mathrm{qd}(\varphi)$ and $\mathrm{qd}(\varphi \vee \psi) = \mathrm{qd}(\varphi \wedge \psi) = \max\{\mathrm{qd}(\varphi),\ \mathrm{qd}(\psi)\}$ as well as $\mathrm{qd}(\mathrm{Q}\ \varphi) = 1 + \mathrm{qd}(\varphi)$, where $\mathrm{Q} \in \{\exists x,\ \forall x,\ \exists X,\ \forall X,\ \exists^{r \bmod q} x\}$.

Next, we introduce the *quantifier alternation hierarchy* within first-order logic FO. It consists of *half levels*, denoted by Σ_m for $m \geqslant 0$, and *full levels*, denoted by FO_m for $m \geqslant 0$. These levels are obtained in the next definition by disregarding universal quantifiers and counting the number of nested negations ranging over quantifiers instead.

Definition 4.3 (Quantifier alternation hierarchy)
Let FO_0 *and* Σ_0 *consist of Boolean combinations of atomic formulae. For* $m \geqslant 1$ *let* Σ_m *be the smallest set of formulae such that*

- $\{\varphi,\ \neg\varphi\} \subseteq \Sigma_m$ *for all* $\varphi \in \Sigma_{m-1}$.
- $\{\exists x\ \varphi,\ \varphi \vee \psi,\ \varphi \wedge \psi\} \subseteq \Sigma_m$ *for all* $\varphi, \psi \in \Sigma_m$.

For $m \geqslant 1$ *let* FO_m *be the smallest set of formulae such that*

- $\{\varphi,\ \neg\varphi\} \subseteq \mathrm{FO}_m$ *for all* $\varphi \in \Sigma_m$.
- $\{\varphi \vee \psi,\ \varphi \wedge \psi\} \subseteq \mathrm{FO}_m$ *for all* $\varphi, \psi \in \mathrm{FO}_m$.

Let $m, n \geqslant 0$. *The fragments* $\Sigma_{m,n}$ *and* $\mathrm{FO}_{m,n}$ *comprise all formulae in* Σ_m *and* FO_m, *respectively, of quantifier depth at most* n.

A convenient convention is to call FO_m the m^{th} *level* and Σ_m the $(m - 1/2)^{th}$ *level* of the alternation hierarchy. Note that this ensures the expressive power to increase with increasing level, and that this is consistent with Σ_1 being the first half level. The structure of the quantifier alternation hierarchy is depicted in Figure 4.2.

There are other ways to define quantifier alternation that incorporate universal quantifiers, but it is often more convenient to handle only existential quantification and negations. The above definition using negation nesting faithfully captures quantifier alternation because of the duality of existential and universal quantifiers in terms of the equivalences $\neg\exists x\ \varphi \equiv \forall x\ \neg\varphi$ and $\neg\forall x\ \varphi \equiv \exists x\ \neg\varphi$. We can thus think of a negation as switching to the other kind of quantifier. More precisely, consider a formula φ and let φ' be the formula that arises from φ by using the above duality and De Morgan's laws[1] to move negations inwards so that no quantifier is in the range of a negation. The formula φ is in FO_m if the formula φ' has the property that on each root-leaf-path of its parse tree there are at most $m-1$ quantifier alternations or equivalently, at most m blocks of quantifiers. The fragment Σ_m consists of those formulae φ in FO_m such that φ' starts with an existential quantifier or have at most $m-1$ blocks of quantifiers.

The probably most common way to define quantifier alternation in the literature is to rely on the *prenex normal form*. A formula is in prenex normal form if it can be decomposed into a quantifying prefix that contains only quantifiers and a quantifier-free suffix. Formulae in FO_m are then those in prenex normal form with at most m blocks of quantifiers in the quantification part. Disregarding the quantifier depth, this defines the same hierarchy with respect to expressive power. This can be seen using the usual procedure to obtain an equivalent prenex normal form; *i.e.*, introduce new variables and move quantifiers outwards over disjunctions and conjunctions in a suitable order.

Relying on the prenex normal form is too coarse for our purposes for at least two reasons. It does not respect variable restrictions, so quantifier alternation within FO^2 would have no meaning in this setting. What is more, the transition to the prenex normal form utterly destroys the quantifier depth. Indeed, the quantifier depth of an

[1] De Morgan's laws are the equivalences $\neg(\varphi \vee \psi) \equiv \neg\varphi \wedge \neg\psi$ and $\neg(\varphi \wedge \psi) \equiv \neg\varphi \vee \neg\psi$.

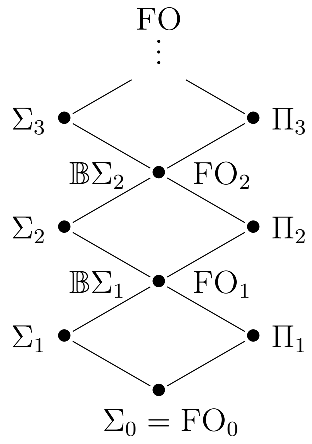

Figure 4.2.: The zigzag lattice of the first-order alternation hierarchy: $\mathbb{B}\Sigma_m$ is the Boolean closure of Σ_m and Π_m consists of negations of Σ_m-formulae.

equivalent formula in prenex normal form is the *total* number of quantifiers occurring in the original formula, which may be vast, even for formulae with moderate quantifier depth. Indeed, on a purely syntactic level, even formulae with quantifier depth 1 have an unbounded total number of quantifiers. Notice that a hierarchy based alternation of second-order quantifiers collapses to its first half level by a result of Thomas [Tho82, Theorem 5.2]. More precisely, Thomas showed that every MSO-formula has an equivalent MSO-formula with a total of only one existential second-order quantifier. Therefore, not even the depth of second-order quantifiers yields a strict hierarchy.

Part II of this thesis focuses on the two-variable first-order alternation hierarchy Σ^2_m and FO^2_m as well as $\Sigma^2_{m,n}$ and $\mathrm{FO}^2_{m,n}$ combining two-variable restriction, quantifier depth, and alternation depth. More generally, for $r \geqslant 0$ we let

- $\Sigma^r_m = \Sigma_m \cap \mathrm{FO}^r$,
- $\Sigma^r_{m,n} = \Sigma_{m,n} \cap \mathrm{FO}^r$,
- $\mathrm{FO}^r_m = \mathrm{FO}_m \cap \mathrm{FO}^r$,
- $\mathrm{FO}^r_{m,n} = \mathrm{FO}_{m,n} \cap \mathrm{FO}^r$.

The intersections are truly meant to be on the syntactic level and not on the semantic level. Consider the following to make this point clear: A classical result by Thérien and Wilke about FO^2 is often written in the literature as "$\mathrm{FO}^2 = \Sigma_2 \cap \Pi_2$", where Π_2 consists of negations of Σ_2-formulae [TW98]. What it signifies is not that FO^2-formulae are those which are both in Σ_2 and in Π_2 (which is not often the case), but that a language is *definable* in FO^2 if and only if it is *definable* in both Σ_2 and in Π_2.

A formula is thus in $\Sigma^r_{m,n}$, for example, if it is in Σ_m, has a quantifier depth of at most n, and uses at most r variables. These definitions make sense in our setting, whereas this would not be the case when relying on the prenex normal form. Let us stress that prenex normal forms could be used indirectly: A formula is in Σ_m if and only if it is equivalent to a formula in prenex normal form with at most m blocks of quantifiers.

These fragments may be further refined by restricting the relational signature of the fragment as mentioned above. For example, $\mathrm{FO}^r_{m,n}[<, \mathrm{suc}]$ is the fragment of formulae in $\mathrm{FO}^r_{m,n}$ that only use the numerical predicates $<$ and suc.

5. Fragments and $\mathcal{C}$-Varieties of Formal Languages

This chapter proves closure properties of the language families defined by logic fragments. More precisely, syntactic closure conditions of fragments are given such that the defined language families are closed under residuals (Proposition 5.3) and under the inverse image of certain homomorphisms (Proposition 5.6). Combining closure under residuals and inverse homomorphisms yields the main theorem of fragments, characterizing languages defined by fragments in terms of $\mathcal{C}$-varieties (Theorem 5.7). The results of this chapter are published in [KL12b].

5.1. Closure under Residuals

Depending on the logic resources used, further syntactic closure axioms are necessary in addition to the axioms of fragments. Such amendments are encapsulated in so-called stability conditions. The following additional assumptions are necessary when the successor, minimum, or maximum predicate is used.

Definition 5.1 (suc-stability)
A prefragment $\mathcal{F}$ is suc-stable *if for all $x, y \in \mathbb{V}_1$ all of the following hold:*

1. $(x = y) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$.
2. $\mathrm{empty} \leqslant_{\mathcal{F}} \max(x) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$.
3. $\mathrm{empty} \leqslant_{\mathcal{F}} \min(y) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$.

This means that a successor predicate $\mathrm{suc}(x, y)$ may always be replaced by equality or by $\max(x)$ or $\min(y)$, both of which may be replaced by the predicate empty.

For the residual of formulae using a modular predicate or a modular counting quantifier, the next definition gives the additional assumptions.

Definition 5.2 (remainder-stability; mod-stability)
A prefragment $\mathcal{F}$ is remainder-stable *if for all formulae φ, all $x \in \mathbb{V}_1$, and all $q, r, s \in \mathbb{Z}$:*

1. $\big(x \equiv s \pmod q\big) \leqslant_{\mathcal{F}} \big(x \equiv r \pmod q\big)$.
2. $(\exists^{s \bmod q} x\ \varphi) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$.

A prefragment $\mathcal{F}$ is mod-stable *if it is remainder-stable and for all φ, x, q, and r:*

3. $\varphi \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$ *whenever* $x \notin \mathrm{FV}(\varphi)$.
4. $\neg\varphi \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$ *whenever* $x \notin \mathrm{FV}(\varphi)$.

Axioms (1) and (2) state that the remainder parameter r can be altered for modular predicates and modular quantifiers. Axiom (3) extends the similar axiom of prefragments to modular quantifiers. That axiom (4) should introduce a negation might not seem natural at first glance. However, it is not too surprising considering the fact that a prefragment complying with axiom (2) is already able to negate modular quantifiers on the semantic level. Indeed, $\neg\exists^{r \bmod q} x\ \varphi$ for $q \neq 0$ is equivalent to the disjunction of $\exists^{s \bmod q} x\ \varphi$, where s ranges over $\{0, \ldots, q-1\} \setminus \{r \bmod q\}$.

Proposition 5.3
Let $\mathcal{F}$ be a suc-stable and mod-stable prefragment. The family of languages defined by $\mathcal{F}$ is closed under residuals.

Remember that a residual of L is a language of the form $u^{-1}Lv^{-1} = \{w \mid uwv \in L\}$. Residuals are of particular interest in algebraic contexts: If $\mathcal{F}$ defines a language family that is closed under residuals, then the equivalence $u \cong_{\mathcal{F}} v$ given by $u \models \varphi \Leftrightarrow v \models \varphi$ for all $\varphi \in \mathcal{F}$ is actually a congruence, and the quotient $A^*/{\cong}$ is naturally equipped with a monoid structure.

Note that a fragment is vacuously suc-stable if it uses neither of the predicates suc, max, or min. Similarly, $\mathcal{F}$ is vacuously mod-stable if it uses neither modular counting quantifiers nor the modular predicate.

The complete formal proof of the proposition employs a rather laborious construction and can be found in Section 5.4. Here we present only its gist. It suffices to consider left residuals $a^{-1}L$ and right residuals La^{-1} by a letter a. Let us concentrate on left residuals, for which truth of a formula φ on aw is to be determined by a derived formula $a^{-1}\varphi$ on w; that is, $a^{-1}\varphi$ is to be constructed such that $aw \models \varphi$ if and only if $w \models a^{-1}\varphi$. In the evaluation of $a^{-1}\varphi$, the first position of aw is not present but thought of as a *virtual* position. Although the virtual position is not actually present, it must be possible for quantifiers to place variables on it, which is simulated by syntactic bookkeeping methods.

The construction of $a^{-1}\varphi$ is done by induction on the structure of φ, in which the stability conditions allow to maintain the syntactic invariant $a^{-1}\varphi \leqslant_{\mathcal{F}} \varphi$. This is much stronger than necessary for the proof of the proposition and immediately yields that $\varphi \in \mathcal{F}$ implies $a^{-1}\varphi \in \mathcal{F}$. This shows closure of the language family defined by $\mathcal{F}$ under residuals.

5.2. Closure under Inverses of Homomorphisms

The following additional axioms are necessary in case order predicates or second-order quantifiers are available.

Definition 5.4 (order-stability)
A prefragment $\mathcal{F}$ is order-stable *if for all $x, y \in \mathbb{V}_1$:*

1. $(x < y) \leqslant_{\mathcal{F}} (x \leqslant y)$.

2. $(x \leqslant y) \leqslant_{\mathcal{F}} (x < y)$.

Definition 5.5 (MSO-stability)
A prefragment $\mathcal{F}$ is MSO-stable *if for all formulae φ, all $x \in \mathbb{V}_1$, and all $X, Y \in \mathbb{V}_2$:*

1. $(x \in Y) \leqslant_{\mathcal{F}} (x \in X)$.

2. $\exists Y\, \exists X\, \varphi \leqslant_{\mathcal{F}} \exists X\, \varphi$.

3. $\forall Y\, \forall X\, \varphi \leqslant_{\mathcal{F}} \forall X\, \varphi$.

A homomorphism $h\colon B^* \to A^*$ between finitely generated free monoids is said to be

- *non-erasing* if $|h(b)| \geqslant 1$ for all $b \in B$,
- *length-reducing* if $|h(b)| \leqslant 1$ for all $b \in B$,
- *length-preserving* if $|h(b)| = 1$ for all $b \in B$, and
- *length-multiplying* if there exists $m \in \mathbb{N} \setminus \{0\}$ such that $|h(b)| = m$ for all $b \in B$.

Note that the homomorphism h is completely determined by the images of the letters by the universal property of free monoids.

Let $\mathcal{C}$ be a family of homomorphisms between finitely generated free monoids. The members of $\mathcal{C}$ are called *$\mathcal{C}$-homomorphisms*. The next proposition gives conditions for a fragments to be closed under inverses of $\mathcal{C}$-homomorphisms for natural homomorphism families $\mathcal{C}$.

Proposition 5.6
Let $\mathcal{F}$ be a fragment and let $\mathcal{C}$ be a family of homomorphisms between finitely generated free monoids. Suppose that all of the following conditions hold:

1. *If $\mathcal{F}$ uses the predicate $\leqslant$ or $<$, then $\mathcal{F}$ is order-stable or all $\mathcal{C}$-homomorphisms are length-reducing.*
2. *If $\mathcal{F}$ uses the predicate* suc, min, max, *or* empty, *then $(x = y) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$ for all $x, y \in \mathbb{V}_1$ and all $\mathcal{C}$-homomorphisms are non-erasing.*
3. *If $\mathcal{F}$ uses a modular predicate, then all $\mathcal{C}$-homomorphisms are length-multiplying and either $\mathcal{F}$ is mod-stable or all $\mathcal{C}$-homomorphisms are length-preserving.*
4. *If $\mathcal{F}$ uses a modular quantifier, then $\mathcal{F}$ is mod-stable or all $\mathcal{C}$-homomorphisms are length-reducing.*
5. *If $\mathcal{F}$ uses a second-order quantifier, then $\mathcal{F}$ is MSO-stable or all $\mathcal{C}$-homomorphisms are length-reducing.*

Under these assumptions the family of languages defined by $\mathcal{F}$ is closed under inverses of $\mathcal{C}$-homomorphisms.

Using more logic constructs in a fragment typically necessitates more syntactic stability conditions or leads to closure under fewer inverse homomorphisms. This trade-off between syntactic closure properties and inverse homomorphisms is formalized by the series of implications in Proposition 5.6, each implication covering certain logic resources. Proposition 5.6 in particular implies that every fragment is closed under length-preserving homomorphisms.

At this point we only sketch the idea for showing Proposition 5.6 as the full formal proof is rather lengthy and technically involved; it can be found in Section 5.5.

Consider a formula φ and a homomorphism $h\colon B^* \to A^*$. The goal is to determine truth of φ on the word $h(w)$ by a derived formula, denoted by $h^{-1}(\varphi)$, that is evaluated on w; that is, $h^{-1}(\varphi)$ is to be constructed such that $h(w) \models \varphi$ if and only if $w \models h^{-1}(\varphi)$. Each position of $h(w)$ is encoded by its originating position j in w together with an offset $d \in \{1, \dots, |h(w[j])|\}$ to specify the sub-position within the image of the letter $w[j]$. The offset is thus bounded by the maximal length $|h(b)|$ for letters $b \in B$. One such offset suffices for defining the position of a first-order variable. This offset can be stored syntactically. For second-order variables this is not possible because every occurrence of the variable in $h(w)$ has its own, independent offset. The solution is to distribute a second-order variable X in $h(w)$ over multiple second-order variables X_d in w, according to the offset d.

The definition of $h^{-1}(\varphi)$ is by induction on the structure of φ. Using the stability conditions allows to obtain $h^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi$ as a syntactic invariant, which in particular yields that $\varphi \in \mathcal{F}$ implies $h^{-1}(\varphi) \in \mathcal{F}$. Therefore, the language family defined by $\mathcal{F}$ is closed under inverses of h.

5.3. $\mathcal{C}$-Varieties

This section characterizes fragments in terms of so-called $\mathcal{C}$-varieties of languages. A *category* $\mathcal{C}$ of homomorphisms is a family of homomorphisms between finitely generated free monoids that is closed under composition and contains all length-preserving homomorphisms. Examples of categories include:

- The family $\mathcal{C}_{all}$ of all homomorphisms between finitely generated free monoids,
- The family $\mathcal{C}_{ne}$ of non-erasing homomorphisms,
- The family $\mathcal{C}_{lm}$ of length-multiplying homomorphisms,
- The family $\mathcal{C}_{lr}$ of length-reducing homomorphisms, and
- The family $\mathcal{C}_{lp}$ of length-preserving homomorphisms.

A *positive $\mathcal{C}$-variety* for a category $\mathcal{C}$ is a family of $*$-languages that is closed under positive Boolean combinations, residuals, and inverse images of $\mathcal{C}$-homomorphisms. A *$\mathcal{C}$-variety* is a positive $\mathcal{C}$-variety that is closed under complementation.

The notion of $\mathcal{C}$-varieties was contrived by Straubing as a generalization of Eilenberg's varieties as introduced in Section 2.2 [Str02]. For varieties there is a close connection between languages and algebraic recognizers like monoids and semigroups. In some cases monoids are better suited than semigroups, while in other cases it is the other way round. This leads to two parallel theories of Eilenberg-varieties, namely $*$-varieties in the monoid setting and $+$-varieties in the semigroup setting. The framework of $\mathcal{C}$-varieties unifies this by considering homomorphisms as algebraic recognizers for which $*$-varieties correspond to $\mathcal{C}_{all}$-varieties and $+$-varieties basically correspond to $\mathcal{C}_{ne}$-varieties. Many desirable properties of varieties are also valid more generally for $\mathcal{C}$-varieties; in particular membership of a language in a $\mathcal{C}$-variety only depends on whether a certain finite set of identities is satisfied by the syntactic homomorphism. For further details about the equational theory of $\mathcal{C}$-varieties refer to [Pin12].

The following fundamental theorem on fragments states that in many cases a fragment defines a $\mathcal{C}$-variety for some natural category $\mathcal{C}$.

Theorem 5.7
Let $\mathcal{F}$ be a suc-stable and mod-stable fragment. Let $\mathcal{C}$ be a category of homomorphisms between finitely generated free monoids. Suppose all of the following conditions hold:

1. *If $\mathcal{F}$ uses the predicate $\leqslant$ or $<$, then $\mathcal{F}$ is order-stable or all $\mathcal{C}$-homomorphisms are length-reducing.*
2. *If $\mathcal{F}$ uses the predicate* suc, min, max, *or* empty, *then all $\mathcal{C}$-homomorphisms are non-erasing.*
3. *If $\mathcal{F}$ uses a modular predicate, then all $\mathcal{C}$-homomorphisms are length-multiplying.*
4. *If $\mathcal{F}$ uses a second-order quantifier, then $\mathcal{F}$ is MSO-stable or all $\mathcal{C}$-homomorphisms are length-reducing.*

Under these assumptions the family of languages defined by $\mathcal{F}$ is a positive $\mathcal{C}$-variety.

Proof. We have to show that $\mathcal{L}(\mathcal{F})$ is closed under union, intersection, residuals, and inverse $\mathcal{C}$-homomorphisms. Lemma 4.2 immediately yields closure under union and intersection. It remains to show that $\mathcal{L}(\mathcal{F})$ is closed under residuals and inverses of $\mathcal{C}$-homomorphisms. Closure under residuals is given by Proposition 5.3 and closure under inverse $\mathcal{C}$-homomorphisms is shown in Proposition 5.6. $\square$

As for closure under inverse homomorphisms in Proposition 5.6, there is a trade-off between the closure properties of the language family and the fragment: Either we need to weaken the closure conditions on the family of languages, or we need to strengthen the stability conditions of the fragment. Theorem 5.7 formalizes this as a series of implications. The next remark examines the necessity of the premises in Theorem 5.7.

Remark 5.8
We start by arguing for the necessity of suc-stability. Consider the fragment $\mathcal{G}$ generated by $\exists x \exists y\ (\mathrm{suc}(x, y) \wedge \lambda(x) = a \wedge \lambda(y) = b)$. Over the alphabet A this sentence defines A^*abA^*. It is not hard to see, that every language definable in $\mathcal{G}$ is an ideal of A^*; *i.e.*, $L \in \mathcal{L}_A(\mathcal{G})$ implies $L = A^*LA^*$. (We will actually prove in Proposition 12.29 more generally that all $\Sigma_1^2[<, \mathrm{suc}]$-languages are ideals. Note that $\mathcal{G} \subseteq \Sigma_1^2[<, \mathrm{suc}]$.) However, the languages $a^{-1}L = L \cup bA^*$ and $Lb^{-1} = L \cup A^*a$ are not ideals and consequently cannot be definable in $\mathcal{G}$. This argument heavily builds on the unavailability of the minimum and maximum predicates because $\Sigma_1^2[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$ can define non-ideal languages. This shows that suc-stability has to introduce the minimum and maximum predicates.

Let now $\mathcal{G}$ be the fragment generated by $\exists x\ (\lambda(x) = a \wedge \max(x))$. The language L defined by this sentence is A^*a and $a^{-1}L = \{\varepsilon\} \cup L$. Obviously $a^{-1}L \neq A^*$, but all sentences in $\mathcal{G}$ that are satisfied by the empty word are equivalent to $\top$. This shows that suc-stability has to introduce the predicate empty.

Let us discuss why mod-stability is required. That we must be able to alter the remainder parameters of the modular predicate and the modular quantifier to get closure under residual is not hard to verify by considering the fragments generated by the sentence $\exists x\ (\lambda(x) = a \wedge x \equiv 0 \pmod 2)$ and by the sentence $\exists^{0 \bmod 2} x\colon \lambda(x) = a$.

We next argue that order-stability in condition (1) is required. Consider the fragment $\mathbb{B}\Sigma_1^+[\leqslant]$ of Boolean combinations of existential first-order formulae using only non-strict order without any negations — not even negations over atomic formulae. This fragment is not order-stable, and indeed, it does not define a positive $\mathcal{C}_{all}$-variety because it is not closed under inverses of arbitrary homomorphisms. This is shown by the following example. In anticipation of Theorem 7.1, we claim that every language L that is definable in this fragment is *stutter-invariant*; *i.e.*, $paq \in L$ if and only if $paaq \in L$ for all letters a and all words p and q. In other words, in every context a single letter may be duplicated or several copies of a letter may be reduced to just one copy. Consider the language $L = A^*aA^*bA^*aA^*bA^*$ over the two-letter alphabet $A = \{a, b\}$. Observe that if $x \leqslant y$ with x being an a-position and y being a b-position, then $x < y$. With this it is straightforward to verify that L is $\Sigma_1^+[\leqslant]$-definable. Letting $h\colon \{c\}^* \to A^*$ be the homomorphism given by $c \mapsto ab$, we have $h(c) \notin L$ but $h(cc) \in L$. This shows that $h^{-1}(L)$ is not stutter-invariant and consequently $h^{-1}(L)$ cannot be $\mathbb{B}\Sigma_1^+[\leqslant]$-definable.

We come to (2) and reason why only non-erasing homomorphisms are allowed when at least one of the predicates suc, min, max, or empty is used. Consider the languages A^*aaA^* and aA^* over the alphabet $A = \{a, b, c\}$ and consider the erasing homomorphism $h\colon A^* \to A^*$ given by $a \mapsto a$, $b \mapsto b$, and $c \mapsto \varepsilon$. Both languages are definable in $\Sigma_1[<, \mathrm{suc}, \mathrm{min}]$, but neither of the inverse images $h^{-1}(A^*aaA^*)$ and $h^{-1}(aA^*)$ is definable in $\Sigma_1[<, \mathrm{suc}, \mathrm{min}]$. A symmetric argument applies for the maximum predicate. For the predicate empty note that the language $\{\varepsilon\}$ is definable in $\Sigma_1[<, \mathrm{empty}]$ but not in $\Sigma_1[<]$.

We now consider condition (3) and justify the restriction to length-multiplying homomorphisms. The fragment $\mathrm{FO}[<, \mathrm{mod}]$ defines a $\mathcal{C}_{lm}$-variety by Theorem 5.7, but not a $\mathcal{C}_{all}$-variety. To see this, consider the language $L = \{u \in \{a\}^* \mid |u| \equiv 0 \pmod 2\}$ and the homomorphism $h\colon \{a,b\}^* \to \{a\}^*$ defined by $a \mapsto a$ and $b \mapsto aa$. The language is definable in $\mathrm{FO}[<, \mathrm{mod}]$, but $h^{-1}(L)$ is not: $h^{-1}(L)$ consists of all words with an even number of occurrences of the letter a and in particular, $(ab)^n \in h^{-1}(L)$ if and only if $(ab)^{n+1} \notin h^{-1}(L)$ for all n. On the other hand, one can verify that for sufficiently large n, a fixed sentence in $\mathrm{FO}[<, \mathrm{mod}]$ cannot distinguish the word $(ab)^n$ from $(ab)^{n+1}$. This shows that no such sentence is able to define the inverse image $h^{-1}(L)$. $\diamond$

Let us state some consequences of Theorem 5.7. We first record corollaries for fragments without modular resources: More precisely, we consider fragments of MSO and of FO both with and without successor predicates. After that, we turn to specific fragments with modular resources in Theorem 5.13, generalizing a result by Straubing [Str02]. The following is an immediate consequence of Theorem 5.7.

Corollary 5.9
Let $\mathcal{F} \subseteq \mathrm{MSO}[<, \leqslant, =]$ be a fragment that is MSO-stable and order-stable. The family of languages defined by $\mathcal{F}$ is a positive $*$-variety. $\square$

Remember that a positive $*$-variety is a positive $\mathcal{C}_{all}$-variety and that a positive $+$-variety is a positive $\mathcal{C}_{ne}$-variety of languages of non-empty words.

There is a similar result for formulae using the successor predicate. The natural setting in this case is non-empty words, thus leading to $+$-varieties of languages instead of $*$-varieties. In this case the empty-predicate is of course meaningless and consequently, it is omitted from the relational signature.

Corollary 5.10
Let $\mathcal{F} \subseteq \mathrm{MSO}[<, \leqslant, =, \mathrm{suc}, \min, \max]$ be an MSO-stable and order-stable fragment. Suppose $\min(y) \leqslant_{\mathcal{F}} \mathrm{suc}(x,y)$ and $\max(x) \leqslant_{\mathcal{F}} \mathrm{suc}(x,y)$ as well as $(x = y) \leqslant_{\mathcal{F}} \mathrm{suc}(x,y)$ for all first-order variables x and y. The family of languages defined by $\mathcal{F}$ over non-empty words is a positive $+$-variety.

Proof. Let $\mathcal{G}$ be the smallest fragment that contains $\mathcal{F}$ and satisfies $(\mathrm{empty} \leqslant_{\mathcal{G}} \min(x))$ as well as $(\mathrm{empty} \leqslant_{\mathcal{G}} \max(x))$ for all first-order variables x. The fragment $\mathcal{G}$ is order-stable, suc-stable, and MSO-stable. By Theorem 5.7, the family of languages defined by $\mathcal{G}$ is a positive $\mathcal{C}_{ne}$-variety. If in a formula $\varphi \in \mathcal{G}$ all empty-predicates are replaced by $\perp$ to obtain φ', then φ' is in $\mathcal{F}$ by construction of $\mathcal{G}$. Moreover, φ and φ' are equivalent over non-empty words, *i.e.*, $\mathcal{L}_A(\varphi) \cap A^+ = \mathcal{L}_A(\varphi') \cap A^+$. This shows that the language family defined by $\mathcal{F}$ over non-empty words is a positive $+$-variety. $\square$

As every first-order fragment is vacuously MSO-stable, we immediately get the following corollaries.

Corollary 5.11
Let $\mathcal{F} \subseteq \mathrm{FO}[<, \leqslant, =]$ be an order-stable fragment. The family of language defined by $\mathcal{F}$ is a positive $*$-variety. $\square$

Corollary 5.12
Let $\mathcal{F} \subseteq \mathrm{FO}[<, \leqslant, =, \mathrm{suc}, \min, \max]$ be an order-stable fragment. Suppose we have $\min(y) \leqslant_{\mathcal{F}} \mathrm{suc}(x,y)$ and $\max(x) \leqslant_{\mathcal{F}} \mathrm{suc}(x,y)$ and $(x = y) \leqslant_{\mathcal{F}} \mathrm{suc}(x,y)$ for all first-order variables x and y. The family of languages defined by $\mathcal{F}$ over non-empty words is a positive $+$-variety. □

Our techniques also allow to obtain a generalization of the main result of a paper by Straubing [Str02, Theorem 3], characterizing various combinations of classical restrictions in terms of admissible quantifiers, available predicates, quantifier depth, and available variables. As quantifiers we consider arbitrary combinations of the first-order quantifiers and the modular quantifier. As numerical predicates we allow arbitrary combinations of equality, order, successor, modular predicates, and an additional modular length predicate. The notation in the following is mostly adopted from Straubing.

Consider a signature

$$\begin{aligned}\mathcal{N} \subseteq \{&x = y,\ x < y,\ \mathrm{suc}(x,y),\ x \equiv r \pmod q,\\ &\qquad\qquad \min(x),\ \max(x),\ \mathrm{len} \equiv r \pmod q \mid x, y \in \mathbb{V}_1,\ r, q \in \mathbb{Z}\},\end{aligned}$$

and let $\mathcal{Q} \subseteq \{\exists,\ \forall,\ \exists^{r \bmod q} \mid r, q \in \mathbb{Z}\}$ be a set of admissible quantifiers. The additional length predicate len is included for compatibility with Straubing's setting. The atomic formula $\mathrm{len} \equiv r \pmod q$ is true if the structure has length r (modulo q); it can be regarded as an abbreviation of $\exists^{r \bmod q} \top$. For $n, k \geqslant 0$ let

$$\mathcal{Q}_n^k[\mathcal{N}] \tag{5.1}$$

be the fragment of FO+MOD-formulae with quantifier depth at most n that use at most k variables, only quantifiers in $\mathcal{Q}$, and only numerical predicates in $\mathcal{N}$. The following generalizes Theorem 3 in [Str02].

Theorem 5.13
Let $n, k \geqslant 0$. If $\mathcal{Q}_n^k[\mathcal{N}]$ is remainder-stable, then it defines a $\mathcal{C}$-variety, where

1. *$\mathcal{C} = \mathcal{C}_{lm}$ if $\mathcal{N}$ contains a modular predicate,*
2. *$\mathcal{C} = \mathcal{C}_{ne}$ if $\mathcal{N}$ contains a successor predicate but no modular predicate, and*
3. *$\mathcal{C} = \mathcal{C}_{all}$ in all other cases.*

We come to the proof of this theorem in a short while. Note that by reading $\mathrm{len} \equiv r \pmod q$ as a macro for $\exists^{r \bmod q} \top$, the fragment $\mathcal{Q}_n^k[\mathcal{N}]$ is remainder-stable if and only if for all $r, s \in \mathbb{Z}$

- $\exists^{s \bmod q} \in \mathcal{Q}$ whenever $\exists^{r \bmod q} \in \mathcal{Q}$,
- $(x \equiv s \pmod q) \in \mathcal{N}$ whenever $(x \equiv r \pmod q) \in \mathcal{N}$, and
- $(\mathrm{len} \equiv s \pmod q) \in \mathcal{N}$ whenever $(\mathrm{len} \equiv r \pmod q) \in \mathcal{N}$.

Of course, neither $\mathcal{Q}$ nor $\mathcal{N}$ are obliged to contain modular resources, but for an available modulus, all remainder parameters are available.

Apart from all combinations of admissible quantifiers $\mathcal{Q}$ and predicates $\mathcal{N}$ considered by Straubing, Theorem 5.13 allows further combinations. For one thing, Straubing allows only a single fixed modulus for the modular predicate and another single fixed modulus for the modular quantifiers; in contrast, we allow an arbitrary set of available moduli for both the modular predicate as well as for the modular quantifiers. For another, Straubing considers only the five signatures $\{=\}$ with only equality, $\{=, \text{suc}\}$ with equality and successor, $\{<\}$ with only order, $\{<, \text{suc}\}$ with order and successor, and the signature $\{x < y,\ x \equiv r \pmod{q},\ \text{len} \equiv r \pmod{q} \mid x, y \in \mathbb{V}_1,\ r \in \mathbb{Z}\}$ with order and modular predicates of fixed modulus q. Theorem 5.13 additionally allows, *e.g.*, combinations of the modular predicate and the successor predicate.

A novelty of Straubing's result was that it includes modular resources. As Straubing pointed out himself, classical arguments based on Ehrenfeucht-Fraïssé-games are "difficult to adapt to modular quantifiers". Indeed, the announced full paper of [Str02] never appeared, and Theorem 5.13 provides the first full formal proof of the result.

Remark 5.14 (Ehrenfeucht-Fraïssé-games)
Ehrenfeucht-Fraïssé-games are a proof technique for certain logic fragments in which two players play a game on two words. We do not use them in this thesis and thus only sketch the concept: The first player tries to distinguish the two words, whereas her opponent tries to conceal differences between them. The game is played in rounds; the first player places a pebble on her current word, and the second player has to answer in an appropriate way. The precise rules and winning situations are determined by the fragment considered. Roughly speaking, placing pebbles corresponds to an existential quantifier, interchanging the words corresponds to a negation, and the number of rounds corresponds to the quantifier depth.

This game-based approach is intuitive, but has major drawbacks: First and foremost, the rules heavily depend on the fragment considered, so every fragment requires a new game. Second, it is not always clear how to translate logic resources into game rules; this is what Straubing refers to when calling Ehrenfeucht-Fraïssé-games difficult to adapt. Third and last, Ehrenfeucht-Fraïssé-games are not a good way to write proofs formally: Winning strategies are complex objects and they can be difficult to specify and to reason about in a precise way. More details about Ehrenfeucht-Fraïssé-games can be found in the textbooks [Str94; EF95] or in the article [HK14], which extends Ehrenfeucht-Fraïssé-games to match our notion of logic fragments in the more general setting of omega-terms. $\diamond$

Proof of Theorem 5.13. As the signature may contain the successor predicate but the minimum and maximum predicates may be unavailable, $\mathcal{Q}^k_n[\mathcal{N}]$ cannot be suc-stable in general and Theorem 5.7 does not apply. Similarly, there is no non-strict order predicate and the fragment cannot be order-stable in general. We thus have to show the necessary closure properties explicitly. Closure under union, intersection, and complementation is given by Lemma 4.2 and by closure under negation. Proposition 5.6 is applicable for closure under inverses of homomorphisms in $\mathcal{C}$. Note that the nonavailability of the non-strict order predicate is easily remedied by replacing $x \leqslant y$ by $\neg(y < x)$. This is possible in all contexts without violating any of the syntactic restrictions of $\mathcal{Q}^k_n[\mathcal{N}]$.

Missing for a $\mathcal{C}$-variety is closure under residuals. The proof of this is deferred and can be found on page 67. The reason is that it needs to use internal details of the construction employed to prove Proposition 5.3, which is given in the next section. We presently only hint briefly at the solution to the issue.

We follow a similar approach as for the non-strict order when the minimum and maximum predicates are unavailable, and we read for instance the minimum predicate as a macro for "there is no predecessor position", which depending on the available logic resources can be implemented accordingly. The implicit quantifier alternation introduced by the negation is no issue here, as $\mathcal{Q}_n^k[\mathcal{N}]$ does not restrict alternation. But we do consider quantifier depth and the substitution may *a priori* violate this restriction. Our construction for residuals will guarantee, however, that we may indeed replace minimum and maximum predicates by the above macro: Minimum and maximum predicates generated by our construction will always originate in successor predicates; moreover, minimum and maximum predicates are only generated if one level of quantification has already been saved. A further complication is that suc-stability introduces the empty-predicate, which can also be eliminated in this setting. □

5.4. Proving Closure under Residuals

In order to prove Proposition 5.3, we give a construction of a formula defining the residual. We concentrate on left residuals by letters. Right residuals are mostly left-right symmetric, and we only make those statements explicit that are not just straightforward symmetric versions of the statements for left residuals. Residuals by words are obtained as repeated residuals by letters.

In the later application we are mostly interested in sentences, but intermediately we also have to handle free variables in our construction. This is formalized by a more general construction for the residuals over the extended alphabet $\Lambda \times 2^V$, which is used to encode the interpretation of free variables. More concretely, for a formula φ and an extended letter (a, J), we want to ascertain truth of φ over the word $(a, J)w$ by evaluating another formula termed $(a, J)^{-1}\varphi$ over w, where w is a word over the extended alphabet. The intuition of the construction of the formula $(a, J)^{-1}\varphi$ is as follows: We imagine that there is a "virtual" (a, J)-position just in front of w, as illustrated in Figure 5.1. This virtual position is dealt with syntactically, using the set J to keep track of the variables that were notionally placed onto it. The closure properties of prefragments then yield that the syntactic structure of $(a, J)^{-1}\varphi$ is not too complicated.

The definition of $(a, J)^{-1}\varphi$ is by induction on the structure of formulae. Keeping the construction extensible, this leads to a relatively large number of lemmas due to different premises for different logic constructs: Four for the atomic formulae, one for Boolean connectives, one for first-order and second-order quantification, respectively, and one for modular quantification. However, all those lemmas are of a very similar

a	b_1	b_2	$\cdots$	b_n
$\{x,z,X\}$	$\{y\}$	$\emptyset$	$\cdots$	$\{X\}$

Figure 5.1.: Our mental picture of the situation for $(a, J)^{-1}\varphi$. The a-position in front of the actual structure $b_1 \cdots b_n$ is only virtual and drawn in gray. The set J under the letter, which is $\{x, z, X\}$ in this example, keeps track of the variables placed upon the virtual position. In this example, x, z, and X are on the virtual position, y is on the position of b_1, and X also contains the maximal position.

structure: Starting with a formula φ, they postulate a formula denoted $(a,J)^{-1}\varphi$ that satisfies

1. $(a,J)^{-1}\varphi \leqslant_{\mathcal{F}} \varphi$ for all appropriate prefragments $\mathcal{F}$, and
2. $[\![(a,J)^{-1}\varphi]\!]_{V'} = (a,J)^{-1}\,[\![\varphi]\!]_V$

for all sets of variables V with $J \cup \mathrm{FV}(\varphi) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$.

Property (1) is syntactic in nature; most notably it yields that $(a,J)^{-1}\varphi$ is in $\mathcal{F}$ whenever φ is in $\mathcal{F}$. An *appropriate prefragment* is a prefragment which, depending on the predicates occurring in φ, potentially is to have some additional closure properties. In the lemmas below, $\mathfrak{F}$ will denote a collection of appropriate prefragments.

Property (2) is semantic correctness, *i.e.*, $(a,J)^{-1}\varphi$ actually defines the left residual of φ. Note that for $(a,J)^{-1}\,[\![\varphi]\!]_V$ to be meaningful, V has to contain J apart from all free variables of φ. This property also implies that no first-order variable in J can appear freely in $(a,J)^{-1}\varphi$ because in particular $[\![(a,J)^{-1}\varphi]\!]_{V'}$ is defined.

Note that we have one formula for all appropriate prefragments, which is stronger than actually needed for the closure under left residuals of a fixed prefragment.

The atomic formulae. The following lemmas give formulae for the left residual of languages defined by one of the atomic formulae. Lemma 5.15 deals with the formulae $\top$, $\bot$, empty, $\min(x)$, $\lambda(x) \in B$, $x = y$, $x < y$, $x \leqslant y$, and $x \in X$ for which the closure properties of a prefragment suffice. Lemma 5.16 and Lemma 5.17 are for $\max(x)$ and for the successor predicate $\mathrm{suc}(x,y)$, respectively. For $\max(x)$ the prefragment must allow to replace $\max(x)$ by the empty-predicate. The construction for $\mathrm{suc}(x,y)$ relies on replacing $\mathrm{suc}(x,y)$ by $\min(y)$. Hence both lemmas restrict the collection of appropriate prefragments. Lemma 5.18 finally gives the construction for the modular predicate $x \equiv r \pmod{q}$ where we have to be able to alter the remainder parameter r.

Lemma 5.15
Let φ be an atomic formula of the form $\top$, $\bot$, $\lambda(x) \in B$, $x = y$, $x < y$, $x \leqslant y$, empty*, $\min(x)$, or $x \in X$. For every $a \in \Lambda$ and every set of variables J there exists a formula $(a,J)^{-1}\varphi$ that satisfies*

$$(a,J)^{-1}\varphi \leqslant_{\mathcal{F}} \varphi \quad \textit{for all prefragments } \mathcal{F} \textit{ and}$$
$$[\![(a,J)^{-1}\varphi]\!]_{V'} = (a,J)^{-1}\,[\![\varphi]\!]_V$$

where V is a set of variables with $J \cup \mathrm{FV}(\varphi) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. Let $(a,J)^{-1}\top := \top$ and $(a,J)^{-1}\bot := (a,J)^{-1}\,\text{empty} := \bot$.

In the following, w denotes some word over $\Lambda \times 2^{V'}$. Note that there exists no w such that $(a,J)w$ is empty which establishes the correctness of $(a,J)^{-1}\,\text{empty}$.

The label predicate is given by

$$(a,J)^{-1}\big(\lambda(x) \in B\big) \;:=\; \begin{cases} \lambda(x) \in B & \text{if } x \notin J, \\ \top & \text{if } x \in J \text{ and } a \in B, \\ \bot & \text{else.} \end{cases}$$

If $x \in J$, then $x((a,J)w) = 1$. Consequently $(a,J)w$ has a label in B at position of x if and only if $a \in B$; for $x \notin J$ we have $((a,J)w)[x((a,J)w)] = w[x(w)]$. This shows that the construction is correct.

For equality and the order predicates let

$$(a,J)^{-1}(x=y) \;:=\; \begin{cases} x=y & \text{if } x\notin J \text{ and } y\notin J,\\ \top & \text{if } x\in J \text{ and } y\in J,\\ \bot & \text{else,}\end{cases}$$

$$(a,J)^{-1}(x<y) \;:=\; \begin{cases} x<y & \text{if } x\notin J \text{ and } y\notin J,\\ \top & \text{if } x\in J \text{ and } y\notin J,\\ \bot & \text{else,}\end{cases}$$

$$(a,J)^{-1}(x\leqslant y) \;:=\; \begin{cases} x\leqslant y & \text{if } x\notin J \text{ and } y\notin J,\\ \top & \text{if } x\in J,\\ \bot & \text{else.}\end{cases}$$

We have $x((a,J)w) = y((a,J)w)$ if and only if either $x,y\in J$ or $x,y\notin J$ and $x(w)=y(w)$. We have $x((a,J)w) < y((a,J)w)$ if and only if either $x\in J$ and $y\notin J$ or $x,y\notin J$ and $x(w)<y(w)$. For the non-strict order we have $x((a,J)w)\leqslant y((a,J)w)$ if and only if either $x\in J$ or $x,y\notin J$ and $x(w)\leqslant y(w)$. This shows the correctness of the constructions.

The formula $\min(x)$ is true over $(a,J)w$ if and only if $x\in J$. Thus

$$(a,J)^{-1}\big(\min(x)\big) \;:=\; \begin{cases} \top & \text{if } x\in J,\\ \bot & \text{else.}\end{cases}$$

Finally for the second-order predicate $x\in X$ let

$$(a,J)^{-1}\big(x\in X\big) \;:=\; \begin{cases} x\in X & \text{if } x\notin J,\\ \top & \text{if } x\in J \text{ and } X\in J,\\ \bot & \text{else.}\end{cases}$$

If $x\in J$, then $x\in X((a,J)w)$ if and only if $X\in J$. For $x\notin J$ we have $x\in X((a,J)w)$ if and only if $x\in X(w)$.

It is easy to see that all these formulae satisfy the syntactic property $(a,J)^{-1}\varphi\leqslant_{\mathcal{F}}\varphi$ for all prefragments $\mathcal{F}$. □

There is a good reason why the maximum predicate $\max(x)$ does not appear in this lemma. Indeed, if $x\in J$, in other words x is on the virtual first position of $(a,J)w$, then the only possible way for x to be on the maximal position is when w is empty. This is the first time when we have to restrict the appropriate prefragments $\mathfrak{F}$.

Lemma 5.16
Consider an atomic formula of the form $\max(x)$ *for some* $x\in\mathbb{V}_1$. *Let* $\mathfrak{F}$ *be a collection of prefragments* $\mathcal{F}$ *such that* empty $\leqslant_{\mathcal{F}}\max(x)$. *For every* $a\in\Lambda$ *and every set of variables* J *there exists a formula* $(a,J)^{-1}\big(\max(x)\big)$ *that satisfies*

$$(a,J)^{-1}\big(\max(x)\big)\leqslant_{\mathcal{F}}\max(x) \quad \textit{for all prefragments } \mathcal{F}\in\mathfrak{F} \textit{ and}$$
$$[\![(a,J)^{-1}\big(\max(x)\big)]\!]_{V'} = (a,J)^{-1}[\![\max(x)]\!]_V$$

where V *is a set of variables with* $J\cup\{x\}\subseteq V$ *and* $V' = V\setminus(J\cap\mathbb{V}_1)$.

Proof. Let

$$(a, J)^{-1}\big(\max(x)\big) \;:=\; \begin{cases} \max(x) & \text{if } x \notin J, \\ \text{empty} & \text{else.} \end{cases}$$

Let w be a word over $\Lambda \times 2^{V'}$. If $x \in J$, then $x((a, J)w) = 1$ and $1 = |(a, J)w|$ if and only if $|w| = 0$. For $x \notin J$ we have $x((a, J)w) = x(w) + 1$.

By assumption, we have the syntactic property $(a, J)^{-1}\big(\max(x)\big) \leqslant_{\mathcal{F}} \max(x)$ for all prefragments $\mathcal{F} \in \mathfrak{F}$. $\square$

There is a special case for left residual of the successor predicate. If x is on the virtual position, then $\mathrm{suc}(x, y)$ is true if and only if y is on the first real position. Again, we have to restrict appropriate prefragments.

Lemma 5.17
Consider an atomic formula of the form $\mathrm{suc}(x, y)$ *for some* $x, y \in \mathbb{V}_1$. *Let* $\mathfrak{F}$ *be a collection of prefragments* $\mathcal{F}$ *such that* $\min(y) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$. *For every* $a \in \Lambda$ *and every set of variables* J *there exists a formula* $(a, J)^{-1}\big(\mathrm{suc}(x, y)\big)$ *that satisfies*

$(a, J)^{-1}\big(\mathrm{suc}(x, y)\big) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$ *for all prefragments* $\mathcal{F} \in \mathfrak{F}$ *and*
$[\![(a, J)^{-1}\big(\mathrm{suc}(x, y)\big)]\!]_{V'} = (a, J)^{-1}[\![\mathrm{suc}(x, y)]\!]_V$

where V *is a set of variables with* $J \cup \{x, y\} \subseteq V$ *and* $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. Let

$$(a, J)^{-1}\big(\mathrm{suc}(x, y)\big) \;:=\; \begin{cases} \mathrm{suc}(x, y) & \text{if } x \notin J \text{ and } y \notin J, \\ \min(y) & \text{if } x \in J \text{ and } y \notin J, \\ \bot & \text{else.} \end{cases}$$

Consider a word w over $\Lambda \times 2^{V'}$. First suppose $x \in J$. We thus have $x((a, J)w) + 1 = y((a, J)w)$ if and only if $y((a, J)w) = 2$ if and only if $y \notin J$ and $y(w) = 1$. Suppose $x \notin J$. Then $x((a, J)w) + 1 = y((a, J)w)$ if and only if $y \notin J$ and $x(w) + 1 = y(w)$. This shows semantic correctness.

By assumption the syntactic property $(a, J)^{-1}\big(\mathrm{suc}(x, y)\big) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$ for all prefragments $\mathcal{F} \in \mathfrak{F}$ is satisfied by this formula. $\square$

For the modular predicates, the only difference between $(a, J)w$ and w is an offset of one position. This means, we need to be able to change the remainder parameter.

Lemma 5.18
Consider an atomic formula of the form $x \equiv r \pmod{q}$ *for some* $x \in \mathbb{V}_1$ *and some* $q, r \in \mathbb{Z}$. *Let* $\mathfrak{F}$ *be a collection of prefragments* $\mathcal{F}$ *which for all integers* $s \in \mathbb{Z}$ *satisfy* $\big(x \equiv s \pmod{q}\big) \leqslant_{\mathcal{F}} \big(x \equiv r \pmod{q}\big)$. *For every* $a \in \Lambda$ *and every set of variables* J *there exists a formula* $(a, J)^{-1}\big(x \equiv r \pmod{q}\big)$ *such that*

$(a, J)^{-1}\big(x \equiv r \pmod{q}\big) \leqslant_{\mathcal{F}} \big(x \equiv r \pmod{q}\big)$ *for all prefragments* $\mathcal{F} \in \mathfrak{F}$ *and*
$[\![(a, J)^{-1}\big(x \equiv r \pmod{q}\big)]\!]_{V'} = (a, J)^{-1}\,[\![\big(x \equiv r \pmod{q}\big)]\!]_V$

where V *is a set of variables with* $J \cup \{x\} \subseteq V$ *and* $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. Let

$$(a,J)^{-1}\big(x \equiv r \pmod q\big) \;:=\; \begin{cases} x \equiv r-1 \pmod q & \text{if } x \notin J, \\ \top & \text{if } x \in J \text{ and } r \equiv 1 \pmod q, \\ \bot & \text{else.} \end{cases}$$

Let w be a word over $\Lambda \times 2^{V'}$. If $x \in J$, then $x((a,J)w) = 1$. If $x \notin J$, then $x((a,J)w) = 1 + x(w)$. Therefore, $x((a,J)w) \equiv r \pmod q$ if and only if $x(w) \equiv r-1 \pmod q$.

By assumption, this formula satisfies the syntactic property for all prefragments in $\mathfrak{F}$. □

Boolean connectives and quantifiers. The next lemmas lift the construction to formulae composed by Boolean combinations and quantifiers. Suppose we constructed formulae for the left residual, and suppose $\mathfrak{F}$ contains the appropriate prefragments. Under these premises, the following lemmas give formulae for the left residual of Boolean combinations (Lemma 5.19) as well as for first-order quantification and second-order quantification (Lemma 5.20 and Lemma 5.21, respectively) for which the prefragments in $\mathfrak{F}$ are still appropriate. For the left residual of the modular counting quantifier there is also a construction (Lemma 5.22), for which we have to further restrict the appropriate prefragments.

Lemma 5.19

Let ψ be a formula of the form $\neg\varphi_1$ or $\varphi_1 \vee \varphi_2$ or $\varphi_1 \wedge \varphi_2$. Let $\mathfrak{F}$ be a collection of prefragments. Suppose for every $a \in \Lambda$ and every set of variables J there exists a formula $(a,J)^{-1}\varphi_i$ for $i \in \{1,2\}$ that satisfy

$$(a,J)^{-1}\varphi_i \leqslant_{\mathcal{F}} \varphi_i \quad \textit{for all prefragments } \mathcal{F} \in \mathfrak{F} \textit{ and}$$
$$[\![(a,J)^{-1}\varphi_i]\!]_{V'} = (a,J)^{-1}\,[\![\varphi_i]\!]_V$$

where V is a set of variables $J \cup \mathrm{FV}(\varphi_i) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$. For every $a \in \Lambda$ and every set of variables J there exists a formula $(a,J)^{-1}\psi$ that satisfies

$$(a,J)^{-1}\psi \leqslant_{\mathcal{F}} \psi \quad \textit{for all prefragments } \mathcal{F} \in \mathfrak{F} \textit{ and}$$
$$[\![(a,J)^{-1}\psi]\!]_{V'} = (a,J)^{-1}\,[\![\psi]\!]_V$$

where V is a set of variables with $J \cup \mathrm{FV}(\psi) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. The constructions for disjunction and conjunction are straightforward:

$$(a,J)^{-1}(\varphi_1 \vee \varphi_2) \;:=\; \big((a,J)^{-1}\varphi_1\big) \vee \big((a,J)^{-1}\varphi_2\big),$$
$$(a,J)^{-1}(\varphi_1 \wedge \varphi_2) \;:=\; \big((a,J)^{-1}\varphi_1\big) \wedge \big((a,J)^{-1}\varphi_2\big).$$

Let us see why the syntactic property holds. Let μ be a context, let $\mathcal{F} \in \mathfrak{F}$, and suppose $\mu(\varphi_1 \vee \varphi_2) \in \mathcal{F}$. Since $\mathcal{F}$ is a prefragment, we see $\mu(\varphi_i) \in \mathcal{F}$ (for $i \in \{1,2\}$). By assumption $\mu((a,J)^{-1}\varphi_i) \in \mathcal{F}$ and finally $\mu\big((a,J)^{-1}\varphi_i \vee (a,J)^{-1}\varphi_i\big) \in \mathcal{F}$. This shows $(a,J)^{-1}(\varphi_1 \vee \varphi_2) \leqslant_{\mathcal{F}} (\varphi_1 \vee \varphi_2)$. By a dual argument we see $(a,J)^{-1}(\varphi_1 \wedge \varphi_2) \leqslant_{\mathcal{F}} (\varphi_1 \wedge \varphi_2)$.

Negation is also straightforward. Let

$$(a,J)^{-1}(\neg\varphi_1) \;:=\; \neg\big((a,J)^{-1}\varphi_1\big).$$

Suppose $\mu(\neg\varphi_1) \in \mathcal{F}$ for some context μ. This means $\mu'(\varphi_1) \in \mathcal{F}$ for the context $\mu' = \mu(\neg\circ)$. By assumption $\mu'\big((a,J)^{-1}\varphi_1\big) \in \mathcal{F}$. Therefore, $\mu\big(\neg(a,J)^{-1}\varphi_1\big) \in \mathcal{F}$. This shows $(a,J)^{-1}(\neg\varphi_1) \leqslant_{\mathcal{F}} \neg\varphi_1$. □

Lemma 5.20

Consider a formula of the form $\exists x\ \varphi$ or $\forall x\ \varphi$ for some $x \in \mathbb{V}_1$. Let $\mathfrak{F}$ be a collection of prefragments. Suppose for every $a \in \Lambda$ and every set of variables J there exists a formula $(a,J)^{-1}\varphi$ that satisfies

$$(a,J)^{-1}\varphi \leqslant_{\mathcal{F}} \varphi \ \text{ for all prefragments } \mathcal{F} \in \mathfrak{F} \text{ and}$$
$$[\![(a,J)^{-1}\varphi]\!]_{V'} = (a,J)^{-1}\,[\![\varphi]\!]_V$$

where V is a set of variables with $J \cup \mathrm{FV}(\varphi) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$. For every $a \in \Lambda$ and every set of variables J there exist formulae $(a,J)^{-1}\big(\exists x\ \varphi\big)$ and $(a,J)^{-1}\big(\forall x\ \varphi\big)$ that for each prefragment $\mathcal{F} \in \mathfrak{F}$ satisfy

$$(a,J)^{-1}\big(\exists x\ \varphi\big) \leqslant_{\mathcal{F}} \exists x\ \varphi, \qquad [\![(a,J)^{-1}\big(\exists x\ \varphi\big)]\!]_{V'} = (a,J)^{-1}\,[\![\exists x\ \varphi]\!]_V,$$
$$(a,J)^{-1}\big(\forall x\ \varphi\big) \leqslant_{\mathcal{F}} \forall x\ \varphi, \qquad [\![(a,J)^{-1}\big(\forall x\ \varphi\big)]\!]_{V'} = (a,J)^{-1}\,[\![\forall x\ \varphi]\!]_V,$$

where V is a set of variables with $J \cup (\mathrm{FV}(\varphi) \setminus \{x\}) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. Let $\varphi_1 := (a, J \cup \{x\})^{-1}\varphi$, let $\varphi_2 := (a, J \setminus \{x\})^{-1}\varphi$, and let

$$(a,J)^{-1}\big(\exists x\ \varphi\big) \ := \ \varphi_1 \vee \exists x\ \varphi_2,$$
$$(a,J)^{-1}\big(\forall x\ \varphi\big) \ := \ \varphi_1 \wedge \forall x\ \varphi_2.$$

Let us see semantic correctness of the existential quantifier. Universal quantification is similar. Let w be a word over $\Lambda \times 2^{V'}$ and let $w' = (a,J)w$. We distinguish cases of i that make $w'[x/i] \in [\![\varphi]\!]_{V\cup\{x\}}$ true.

First consider the case $i = 1$ and let $w'[x/1] = (a, J \cup \{x\})w''$. Note that w'' is w with x removed from all second components. We have $w'[x/1] \in [\![\varphi]\!]_{V\cup\{x\}}$ if and only if $w'' \in [\![(a, J \cup \{x\})^{-1}\varphi]\!]_{V'\setminus\{x\}}$ by assumption. The latter in turn is equivalent to $w \in [\![\varphi_1]\!]_{V'}$. This is because x is not a free variable of φ_1 and consequently the truth value does not depend on the interpretation of x.

For $i \geqslant 2$ let $w'[x/i] = (a, J \setminus \{x\})w''$. Note that $w'' = w[x/i-1]$. By assumption $w'[x/i] \in [\![\varphi]\!]_{V\cup\{x\}}$ if and only if $w'' \in [\![(a, J \setminus \{x\})^{-1}\varphi]\!]_{V'\setminus\{x\}}$.

It remains to show the syntactic property. Let $\mathcal{F}$ be a prefragment in $\mathfrak{F}$. We have $\varphi_1 \leqslant_{\mathcal{F}} (\exists x\ \varphi_1) \leqslant_{\mathcal{F}} (\exists x\ \varphi)$ by assumption on φ as $x \notin \mathrm{FV}(\varphi_1)$. Together with $\exists x\ \varphi_2 \leqslant_{\mathcal{F}} \exists x\ \varphi$ this yields $(a,J)^{-1}(\exists x\ \varphi) \leqslant_{\mathcal{F}} (\exists x\ \varphi)$. The argument for the universal quantifier is analogous. □

Lemma 5.21

Consider a formula of the form $\exists X\ \varphi$ and $\forall X\ \varphi$ for some $X \in \mathbb{V}_2$. Let $\mathfrak{F}$ be a collection of prefragments. Suppose for every $a \in \Lambda$ and every set of variables J there exists a formula $(a,J)^{-1}\varphi$ that satisfies

$$(a,J)^{-1}\varphi \leqslant_{\mathcal{F}} \varphi \ \text{ for all prefragments } \mathcal{F} \in \mathfrak{F} \text{ and}$$
$$[\![(a,J)^{-1}\varphi]\!]_{V'} = (a,J)^{-1}[\![\varphi]\!]_V$$

where V is a set of variables with $J \cup \mathrm{FV}(\varphi) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$. For every $a \in \Lambda$ and every set of variables J there exist formulae $(a, J)^{-1}(\exists X\, \varphi)$ and $(a, J)^{-1}(\forall X\, \varphi)$ that for each prefragment $\mathcal{F} \in \mathfrak{F}$ satisfy

$$(a, J)^{-1}(\exists X\, \varphi) \leqslant_{\mathcal{F}} \exists X\, \varphi, \qquad [\![(a, J)^{-1}(\exists X\, \varphi)]\!]_{V'} = (a, J)^{-1}[\![\exists X\, \varphi]\!]_V,$$
$$(a, J)^{-1}(\forall X\, \varphi) \leqslant_{\mathcal{F}} \forall X\, \varphi, \qquad [\![(a, J)^{-1}(\forall X\, \varphi)]\!]_{V'} = (a, J)^{-1}[\![\forall X\, \varphi]\!]_V,$$

where V is a set of variables with $J \cup (\mathrm{FV}(\varphi) \setminus \{X\}) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. Let $\varphi_1 := (a, J \cup \{X\})^{-1}\varphi$ and $\varphi_2 := (a, J \setminus \{X\})^{-1}\varphi$

$$(a, J)^{-1}(\exists X\, \varphi) \;:=\; \exists X\, (\varphi_1 \vee \varphi_2),$$
$$(a, J)^{-1}(\forall X\, \varphi) \;:=\; \forall X\, (\varphi_1 \wedge \varphi_2).$$

Let us argue for the correctness of the construction for existential quantifiers. Universal quantifiers are similar. Let $w \in \left(\Lambda \times 2^{V'}\right)^*$ and let $w' = (a, J)w$. Consider a set $I \subseteq \{1, \ldots, |w'|\}$ and let $K = \{i - 1 \mid i \in I,\ i \geqslant 2\}$. Suppose $1 \in I$. We have that $w'[X/I] \in [\![\varphi]\!]_{V \cup \{X\}}$ if and only if $(a, J \cup \{X\})(w[X/K]) \in [\![\varphi]\!]_{V \cup \{X\}}$. The latter is equivalent to $w[X/K] \in [\![(a, J \cup \{X\})^{-1}\varphi]\!]_{V' \cup \{X\}}$ by assumption.

A similar argument shows that if $1 \notin I$, then $w'[X/I] \in [\![\varphi]\!]_{V \cup \{X\}}$ is equivalent to $w[X/K] \in [\![(a, J \setminus \{X\})^{-1}\varphi]\!]_{V' \cup \{X\}}$.

It remains to show the syntactic property. Let $\mathcal{F}$ be a prefragment in $\mathfrak{F}$. The assumption on φ yields $\varphi_1 \leqslant_{\mathcal{F}} \varphi$ and $\varphi_2 \leqslant_{\mathcal{F}} \varphi$. Thus $\varphi_1 \vee \varphi_2 \leqslant_{\mathcal{F}} \varphi$ and finally $(a, J)^{-1}\, \exists X\, \varphi \leqslant_{\mathcal{F}} \exists X\, \varphi$. The argument for the universal quantifier is analogous. □

The following lemma lifts the residual construction for the modular counting quantifier. The inner formula of the counting quantifier may or may not hold on the virtual position, and we have to adapt the remainder parameter accordingly.

Lemma 5.22
Consider a formula of the form $\exists^{r \bmod q} x\, \varphi$ for some $x \in \mathbb{V}_1$ and some $r, q \in \mathbb{Z}$. Let $\mathfrak{F}$ be a collection of prefragments $\mathcal{F}$ such that $(\exists^{s \bmod q} x\, \varphi) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\, \varphi)$ for all $s \in \mathbb{Z}$. Suppose that $\psi \leqslant_{\mathcal{F}} \exists^{r \bmod q} x\, \psi$ and $\neg\psi \leqslant_{\mathcal{F}} \exists^{r \bmod q} x\, \psi$ for all $\mathcal{F} \in \mathfrak{F}$ and all formulae ψ with $x \notin \mathrm{FV}(\psi)$. Suppose that for every $a \in \Lambda$ and every set of variables J there exists a formula $(a, J)^{-1}\varphi$ that satisfies

$$(a, J)^{-1}\varphi \leqslant_{\mathcal{F}} \varphi \quad \textit{for all prefragments } \mathcal{F} \in \mathfrak{F} \textit{ and}$$
$$[\![(a, J)^{-1}\varphi]\!]_{V'} = (a, J)^{-1}\, [\![\varphi]\!]_V$$

where V is a set of variables with $J \cup \mathrm{FV}(\varphi) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$. For every $a \in \Lambda$ and every set of variables J there exists a formula $(a, J)^{-1}(\exists^{r \bmod q} x\, \varphi)$ with

$$(a, J)^{-1}(\exists^{r \bmod q} x\, \varphi) \leqslant_{\mathcal{F}} \exists^{r \bmod q} x\, \varphi \quad \textit{for all prefragments } \mathcal{F} \in \mathfrak{F} \textit{ and}$$
$$[\![(a, J)^{-1}(\exists^{r \bmod q} x\, \varphi)]\!]_{V'} = (a, J)^{-1}\, [\![\exists^{r \bmod q} x\, \varphi]\!]_V$$

where V is a set of variables with $J \cup \mathrm{FV}(\exists^{r \bmod q} x\, \varphi) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. Let $\varphi_1 := (a, J \cup \{x\})^{-1}\varphi$ and $\varphi_2 := (a, J \setminus \{x\})^{-1}\varphi$ be the formulae from the premise. Let

$$(a, J)^{-1}(\exists^{r \bmod q} x\, \varphi) \;:=\; \left(\varphi_1 \wedge \exists^{r-1 \bmod q} x\, \varphi_2\right) \vee \left(\neg\varphi_1 \wedge \exists^{r \bmod q} x\, \varphi_2\right).$$

Suppose we are given a model w. The formula realizes a straightforward case distinction: Either setting x to the first position of $(a, J)w$ makes φ true, or it does not. In the former case the number of positions in the remaining factor w that make φ true has to be $r-1$ (modulo q), whereas in the latter case it has to be r (modulo q).

It remains to show the syntactic property. Given the closure properties of prefragments $\mathcal{F} \in \mathfrak{F}$, one can see that $(a,J)^{-1}(\exists^{r \bmod q} x\ \varphi) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$ is inherited from φ_1 and φ_2 by a similar argument as for existential quantifiers. Notice $x \notin \mathrm{FV}(\varphi_1)$. □

The preceding lemma in particular applies if $\mathfrak{F}$ contains only mod-stable prefragments.

Right Residuals. There are of course left-right dual statements providing formulae $\varphi(a,J)^{-1}$ for the right residual. We shall only make those explicit where some attention has to be paid for the premises.

Lemma 5.23
Consider an atomic formula of the form $\mathrm{suc}(x,y)$ *for some* $x, y \in \mathbb{V}_1$. *Let* $\mathfrak{F}$ *be a collection of prefragments* $\mathcal{F}$ *such that* $\max(x) \leqslant_{\mathcal{F}} \mathrm{suc}(x,y)$. *For every* $a \in \Lambda$ *and every set of variables* J *there exists a formula* $\big(\mathrm{suc}(x,y)\big)(a,J)^{-1}$ *that satisfies*

$$\big(\mathrm{suc}(x,y)\big)(a,J)^{-1} \leqslant_{\mathcal{F}} \mathrm{suc}(x,y) \quad \textit{for all prefragments } \mathcal{F} \in \mathfrak{F} \textit{ and}$$
$$[\![\big(\mathrm{suc}(x,y)\big)(a,J)^{-1}]\!]_{V'} = [\![\mathrm{suc}(x,y)]\!]_V\ (a,J)^{-1}$$

where V *is a set of variables with* $J \cup \{x,y\} \subseteq V$ *and* $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. Let

$$\big(\mathrm{suc}(x,y)\big)(a,J)^{-1} \ := \begin{cases} \mathrm{suc}(x,y) & \text{if } x \notin J \text{ and } y \notin J,\\ \max(x) & \text{if } x \notin J \text{ and } y \in J,\\ \bot & \text{else.} \end{cases}$$

This formula obviously satisfies the syntactic property. For the semantic property let w be a word over $\Lambda \times 2^{V'}$. If $y \in J$, then $y(w(a,J)) = |w(a,J)|$. Hence $x(w(a,J)) + 1 = y(w(a,J))$ if and only if $x(w) = |w|$. If $y \notin J$, then $x((a,J)w) + 1 = y((a,J)w)$ if and only if $x \notin J$ and $x(w) + 1 = y(w)$. □

Lemma 5.24
Consider an atomic formula of the form $\min(x)$ *for some* $x \in \mathbb{V}_1$. *Let* $\mathfrak{F}$ *be a collection of prefragments* $\mathcal{F}$ *such that* $\mathrm{empty} \leqslant_{\mathcal{F}} \min(x)$. *For every* $a \in \Lambda$ *and every set of variables* J *there exists a formula* $\big(\min(x)\big)(a,J)^{-1}$ *that satisfies*

$$\big(\min(x)\big)(a,J)^{-1} \leqslant_{\mathcal{F}} \min(x) \quad \textit{for all prefragments } \mathcal{F} \in \mathfrak{F} \textit{ and}$$
$$[\![\big(\min(x)\big)(a,J)^{-1}]\!]_{V'} = [\![\min(x)]\!]_V (a,J)^{-1}$$

where V *is a set of variables with* $J \cup \{x\} \subseteq V$ *and* $V' = V \setminus (J \cap \mathbb{V}_1)$.

Proof. Let

$$\big(\min(x)\big)(a,J)^{-1} \ := \begin{cases} \min(x) & \text{if } x \notin J,\\ \mathrm{empty} & \text{else.} \end{cases}$$

This formula obviously satisfies the syntactic property. To see semantic correctness let w be a word over $\Lambda \times 2^{V'}$. Suppose $x \in J$. In this case $x(w(a,J)) = |w(a,J)|$ and consequently $x(w(a,J)) = 1$ if and only if $|w| = 0$. If $x \notin J$, then $x(w(a,J)) = x(w)$. □

We are now ready to prove Proposition 5.3.

Proof of Proposition 5.3. We show closure under left residuals by letters. Right residuals are similar and residuals by words are obtained as iterated residuals by letters. We proceed by induction on the structure of φ and show that for every $a \in \Lambda$ and every set of variables J there exists a formula $(a,J)^{-1}\varphi$ that satisfies $(a,J)^{-1}\varphi \leqslant_{\mathcal{F}} \varphi$ and $[\![(a,J)^{-1}\varphi]\!]_{V'} = (a,J)^{-1}\,[\![\varphi]\!]_V$ where V is a set of variables with $J \cup \mathrm{FV}(\varphi) \subseteq V$ and $V' = V \setminus (J \cap \mathbb{V}_1)$.

For the atomic modalities $\top$, $\bot$, $\lambda(x) \in B$, $x = y$, $x < y$, $x \leqslant y$, empty, $\min(x)$, and $x \in X$ this is Lemma 5.15. Let $\mathfrak{F} = \{\mathcal{F}\}$. The predicates $\mathrm{suc}(x,y)$, $\max(x)$, and $x \equiv r \pmod q$ are Lemma 5.16, Lemma 5.17, and Lemma 5.18, respectively. Note that in all cases $\mathfrak{F}$ meets the requirements of the respective lemmas as $\mathcal{F}$ is suc-stable as well as mod-stable. For Boolean connectives, first-order quantification, second-order quantification, and modular quantification, the claim follows by induction and Lemma 5.19, Lemma 5.20, Lemma 5.21, and Lemma 5.22, respectively.

Now, using this and setting $a^{-1}\varphi$ to be $(a,\emptyset)^{-1}\varphi$ yields $a^{-1}\varphi \leqslant_{\mathcal{F}} \varphi$ and $a^{-1}\mathcal{L}_A(\varphi) = \mathcal{L}_A(a^{-1}\varphi)$ for every finite alphabet $A \subseteq \Lambda$. In particular, if $\varphi \in \mathcal{F}$, then $a^{-1}\varphi \in \mathcal{F}$ and $L \in \mathcal{L}_A(\mathcal{F})$ implies $a^{-1}L \in \mathcal{L}_A(\mathcal{F})$. Closure of $\mathcal{L}_A(\mathcal{F})$ under right residuals follows symmetrically. This concludes the proof of Proposition 5.3. $\square$

Proof of closure under residuals in Theorem 5.13. We turn to closure under residuals of the languages defined by $\mathcal{Q}_n^k[\mathcal{N}]$ in Theorem 5.13. We concentrate on left residuals, as right residuals follow by left-right symmetry. Be remembered that this fragment is not suc-stable in general, and there are quite some technical issues to be solved to deal with that. We may assume that both n and r are at least 2 and that $\mathrm{suc} \in \mathcal{N}$. Indeed, if $n \leqslant 1$ or $r \leqslant 1$, all binary predicates are rendered useless; and if no successor is present, then $\mathcal{Q}_n^k[\mathcal{N}]$ is vacuously suc-stable.

Let $\mathcal{G}$ be the smallest suc-stable fragment containing $\mathcal{Q}_n^k[\mathcal{N}]$. It is also mod-stable as $\mathcal{Q}_n^k[\mathcal{N}]$ is. Consider the formulae $a^{-1}\varphi$ as above. Starting with a sentence $\varphi \in \mathcal{Q}_n^k[\mathcal{N}]$ yields a sentence $a^{-1}\varphi \in \mathcal{G}$ defining the left residual $a^{-1}L(\varphi)$. In the following we construct an equivalent formula in $\mathcal{Q}_n^k[\mathcal{N}]$. For this we have to eliminate the additional predicates min, max, and empty that are possibly introduced by suc-stability. Note that interpreting the length predicate $\mathrm{len} \equiv r \pmod q$ as a macro for $\exists^{r \bmod q}\, \top$, the construction only changes the remainder parameter r.

By the axioms of suc-stability, a newly introduced predicate min or max always has its origins in a successor predicate. Analyzing the proof of Lemma 5.17, we see that min is only introduced if one of the variables is placed upon the virtual position. (Similarly, a predicate max is only introduced for *right* residuals when one of the variables is placed on the virtual position.) As we start with sentences, first-order variables must have been placed by a first-order quantifier (Lemma 5.20) or by a modular quantifier (Lemma 5.22). In both cases, a variable is placed upon the virtual position (*i.e.*, taking $J \cup \{x\}$ as the decorating set) only if we save the quantifier on this branch of the parse tree. Depending on the quantifiers available, we can thus replace $\min(x)$ by $\neg\exists y\ \mathrm{suc}(y,x)$, by $\forall y\ \neg\,\mathrm{suc}(y,x)$, or by $\exists^{0 \bmod q}\ \mathrm{suc}(y,x)$. Here, y is another available variable. Remember that we may assume $r \geqslant 2$ as otherwise the successor predicate is useless, anyway. For the third case note that there is at most one predecessor. Therefore, $\exists^{0 \bmod q}\ \mathrm{suc}(y,x)$ is true if and only if there exits no predecessor of x — provided $|q| \neq 1$.

Modular quantifiers with modulus ± 1 are always true and therefore not worthwhile. The maximum predicate can be replaced similarly.

Let us turn to the empty-predicate. As soon as first-order quantifiers are available, (non-)emptiness can be stated in the obvious way. Some care has to be taken when only modular quantifiers available, though. In this case consider the formula

$$\xi_{\bar{\varepsilon}} \;:=\; \exists^{1 \bmod q} x \; (\exists^{0 \bmod q} y \;\, \mathrm{suc}(y, x)),$$

where x and y are two distinct available variables, and $q \neq \pm 1$. A word is non-empty if and only if there exists a first position (*i.e.*, a position without predecessor). Hence $\xi_{\bar{\varepsilon}}$ is true if and only if the model at hand is non-empty. (In case of first-order quantifiers available, we could take $\xi_{\bar{\varepsilon}} := \exists x \; \top$ or $\xi_{\bar{\varepsilon}} := \neg \forall x \; \bot$.)

Suppose ψ is the formula which was obtained from $a^{-1}\varphi$ by eliminating min and max using the above procedure. Let

$$\psi' \;:=\; (\xi_{\bar{\varepsilon}} \wedge \psi[\mathrm{empty}/\bot]) \vee (\neg\xi_{\bar{\varepsilon}} \wedge \psi[\mathrm{empty}/\top]),$$

where in $\psi[\mathrm{empty}/\bot]$ and $\psi[\mathrm{empty}/\top]$ all occurrences of the empty-predicate are replaced by $\bot$ and $\top$, respectively. The formula ψ' is equivalent to ψ and thus also to $a^{-1}\varphi$. Moreover $\psi' \in \mathcal{Q}_n^k[\mathcal{N}]$, which concludes the proof. $\square$

5.5. Proving Closure under Inverse Homomorphisms

In order to proof Proposition 5.6, we give a construction for the formula defining the inverse homomorphic image. More specifically, let $A, B \subseteq \Lambda$ be finite alphabets. For a homomorphism $h\colon B^* \to A^*$ and a formula φ we construct a formula $h^{-1}(\varphi)$ that, interpreted over w, has the same truth value as φ interpreted over $h(w)$. In addition, the syntactic structure of $h^{-1}(\varphi)$ is not to complicated.

The later application will be mainly sentences, but we need to handle formulae with free variables intermediately. A position of $h(w)$ is encoded by its origin position in w and an offset. Let us state this formally. Let $h\colon B^* \to A^*$ be a homomorphism, let $w = b_1 \cdots b_m$ with $b_i \in B$, and let $h(w) = a_1 \cdots a_n$ with $a_i \in A$. For every position $i \in \{1, \ldots, n\}$ of $h(w)$ there exist unique integers $j \in \{1, \ldots, m\}$ and $d \in \{1, \ldots, |h(b_j)|\}$ such that $|a_1 \cdots a_i| = |h(b_1 \cdots b_{j-1})| + d$. The position j of w is the *origin* of i and identifies the letter whose image generated i. The integer d is the *offset* within the image of b_j. We call (j, d) the *h-coordinates* in w of the position i of $h(w)$. Figure 5.2 illustrates this principle. Note that since B is finite, $\max\{b \in B \mid |h(b)|\}$ is a well-defined upper bound for d.

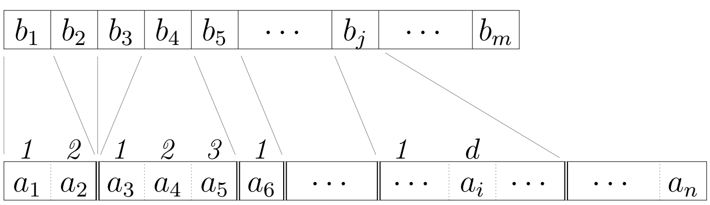

Figure 5.2.: Illustration of the h-coordinates. In the depicted example $h(b_1) = a_1 a_2$, $h(b_2) = h(b_3) = \varepsilon$, $h(b_4) = a_3 a_4 a_5$, and $h(b_5) = a_6$. Position 5, for example, has h-coordinates $(4, 3)$. Position i has h-coordinates (j, d).

Consider a position i of $h(w)$ with h-coordinates (j, d). Suppose a first-order variable x is notionally to be placed in $h(w)$ on position i. Note that that formula is interpreted over w, so placing the variable in $h(w)$ hast to be simulated. This is realized by placing x in w on its origin position j and storing the offset syntactically in a partial mapping $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ with finite domain. The index p indicates that δ is a partial function; *i.e.*, need not be defined for all $\mathbb{V}_1$. As second-order variables can be placed on many positions, the offset cannot be stored in the same way. We distribute the interpretation of X on $h(w)$ over several new variables X_d according to their offset. Specifically, placing X in $h(w)$ on position i is realized in w by placing X_d on position j.

To formalize this, we first introduce for every set of variables V a derived set of variables V_n with the same first-order variables as V such that for every second-order variable X there are n distinct variables $X_1, \ldots, X_n$. Together with an offset function $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ as mentioned above, the encoding we just described is realized by the homomorphism $h_\delta\colon \left(B \times 2^{V_n}\right)^* \to \left(A \times 2^V\right)^*$ of the following definition.

Definition 5.25 (V_n; h_δ)
For every V and every $n \geqslant 0$ let V_n be an arbitrary but fixed minimal set of variables with $V_n \cap \mathbb{V}_1 = V \cap \mathbb{V}_1$ such that for every second-order variable $X \in V$ we have $X \in V_n$ and there exist distinct second-order variables $X = X_1, \ldots, X_n \in V_n$ such that $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_n\}$ are disjoint for $X, Y \in V$ with $X \neq Y$.

Let $h\colon B^* \to A^*$ be a homomorphism, let $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$, and let $n \geqslant 0$. Let the homomorphism $h_\delta\colon \left(B \times 2^{V_n}\right)^* \to \left(A \times 2^V\right)^*$ be given by $h(b, J) = (a_1, J_1) \cdots (a_\ell, J_\ell)$ if $h(b) = a_1 \cdots a_\ell$ and for all first-order variables x and all second-order variables X:

1. *$x \in J_d$ if and only if $x \in J$ and $\delta(x) = d$, and*
2. *$X \in J_d$ if and only if $X_d \in J$.*

To avoid an all to tedious notation, the parameter n is understood implicitly in h_δ. We now come to the actual construction and for every formula φ and every appropriate homomorphism h_δ, we give a formula $h_\delta^{-1}(\varphi)$ such that

1. $h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi$ for all appropriate fragments $\mathcal{F}$, and
2. $[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V})$,

where V is a set of variables containing all free variables of φ. What *appropriate* homomorphisms and what *appropriate* fragments are depends on the resources used by the formula.

The first property in particular yields $h_\delta^{-1}(\varphi) \in \mathcal{F}$ whenever $\varphi \in \mathcal{F}$. It is stronger, however, than merely closure of a fragment under inverse homomorphisms because we get *one* formula which works for *all* appropriate fragments. The second property is semantic correctness, *i.e.*, $h_\delta^{-1}(\varphi)$ indeed defines the inverse image of h_δ. Note that for first-order formulae we may always choose $V \subseteq \mathbb{V}_1$ for which $V_n = V$.

The construction of $h_\delta^{-1}(\varphi)$ is by induction on the structure of φ. In the following, we give the construction for each of the logical constituents. Different constituents require different assumptions and admit different homomorphisms. As for the residual construction in the previous section, formulating things in a general way leads to quite a lot of lemmas. These lemmas can be seen as a toolbox for the closure under inverse homomorphisms of which one needs to consider only those lemmas that are relevant in a given situation. These are then connected by an easy induction. For example, for a first-order fragment without modular quantifiers using only the order predicates, it suffices to consider Lemma 5.26 for the formulae true, false and label, Lemma 5.27 for the order predicates, Lemma 5.31 for the Boolean connectives, and Lemma 5.32 for first-order quantification.

The atomic formulae. We start with atomic formulae. Lemma 5.26 deals with $\top$, $\bot$, $\lambda(x) \in C$, and $x = y$; Lemma 5.27 is for the orders predicates $x < y$ and $x \leqslant y$; Lemma 5.28 is for the successor predicate $\mathrm{suc}(x, y)$ and its entourage $\min(x)$, $\max(x)$, and empty; Lemma 5.29 is for $x \equiv r \pmod q$; and Lemma 5.30 is for the second-order predicate $x \in X$.

Lemma 5.26
Let φ be an atomic formula of the form $\top$, $\bot$, $\lambda(x) \in C$, or $x = y$. Let V be a set of variables with $\mathrm{FV}(\varphi) \subseteq V$. For each homomorphism $h\colon B^* \to A^*$ and each $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\varphi)$ such that for $n = \max\{|h(b)| \mid b \in B\}$ the following hold:

$$h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi \quad \text{for all fragments } \mathcal{F} \text{ and}$$
$$[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V}).$$

Proof. Let $\lambda_w(x)$ for a first-order variable x be an abbreviation for $w[x(w)]$. We have $h_\delta(w) \in [\![\top]\!]_{A,V}$ if and only if $\delta(x)$ is defined and if $1 \leqslant \delta(x) \leqslant |h_\delta(\lambda_w(x))|$ for all first-order variables $x \in V$. Therefore, if $\delta(x)$ is undefined for some $x \in V \cap \mathbb{V}_1$, then we let $h_\delta^{-1}(\top) = \bot$. Else we let

$$h_\delta^{-1}(\top) \;:=\; \bigwedge_{x \in V \cap \mathbb{V}_1} \lambda(x) \in B_{\delta(x)},$$

where $B_d = \{b \in B \mid 1 \leqslant d \leqslant |h(b)|\}$ is the set of letters $b \in B$ such that d is a position of $h(b)$. By the above considerations we see $[\![h_\delta^{-1}(\top)]\!]_{B,V_n} = h_\delta^{-1}([\![\top]\!]_{A,V})$.

Note that for length-multiplying h we could also set $h_\delta^{-1}(\top) = \top$ if $\delta(x)$ is defined and in $\{1, \dots, n\}$, and $h_\delta^{-1}(\top) = \bot$ otherwise. In this case we can relax the assumptions to prefragments and still get the syntactic property.

Consider now a position i of $h(w)$ with h-coordinates (j, d). Position i being a c-position of $h(w)$ is equivalent to d being a c-position of $w[j]$. Let $\widehat{C}$ be the set $\{b \in B \mid (h(b))[\delta(x)] \in C\}$ of letters $b \in B$ such that $\delta(x)$ is a c-position of $h(b)$ for some $c \in C$. Let i' be a position of $h(w)$ with h-coordinates (j', d'). Then $i = i'$ if and only if $j = j'$ and $d = d'$. We therefore let

$$\begin{aligned}
h_\delta^{-1}(\bot) \;&:=\; \bot \\
h_\delta^{-1}(\lambda(x) \in C) \;&:=\; h_\delta^{-1}(\top) \wedge \bigl(\lambda(x) \in \widehat{C}\bigr), \\
h_\delta^{-1}(x = y) \;&:=\; \begin{cases} h_\delta^{-1}(\top) \wedge (x = y) & \text{if } \delta(x) = \delta(y), \\ \bot & \text{else.} \end{cases}
\end{aligned}$$

It is easy to see that these formulae satisfy the syntactic property $h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi$ for all fragments $\mathcal{F}$. $\square$

Note that the formula $h_\delta^{-1}(\top)$ ensures that the homomorphic image is admissible in the sense that all first-order variables in V have a well-defined interpretation. All formulae have to ensure this, and we shall thus use this formula a lot in the following.

Lemma 5.27
Let φ be an atomic formula of the form $x < y$ or $x \leqslant y$ for some $x, y \in \mathbb{V}_1$. Let V be a set of variables with $\mathrm{FV}(\varphi) \subseteq V$. For every homomorphism $h\colon B^* \to A^*$ and every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\varphi)$ such that for $n = \max\{|h(b)| \mid b \in B\}$ the following hold:

$$h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi \text{ for all order-stable fragments } \mathcal{F} \text{ and}$$
$$[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V}).$$

Moreover, if h is length-reducing, then $h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi$ for all fragments $\mathcal{F}$.

Proof. If h is not length-reducing let

$$h_\delta^{-1}(x < y) \ := \ h_\delta^{-1}(\top) \wedge \begin{cases} x \leqslant y & \text{if } \delta(x) < \delta(y), \\ x < y & \text{else,} \end{cases}$$
$$h_\delta^{-1}(x \leqslant y) \ := \ h_\delta^{-1}(\top) \wedge \begin{cases} x \leqslant y & \text{if } \delta(x) \leqslant \delta(y), \\ x < y & \text{else.} \end{cases}$$

Here, $h_\delta^{-1}(\top)$ is the formula from Lemma 5.26 for the set V. For length-reducing h let $h_\delta^{-1}(x < y) := (h_\delta^{-1}(\top) \wedge x < y)$ and $h_\delta^{-1}(x \leqslant y) := (h_\delta^{-1}(\top) \wedge x \leqslant y)$. It is not hard to verify that these formulae satisfy $h_\delta^{-1}(x < y) \leqslant_{\mathcal{F}} (x < y)$ and $h_\delta^{-1}(x \leqslant y) \leqslant_{\mathcal{F}} (x \leqslant y)$ for all order-stable fragments $\mathcal{F}$.

For correctness consider positions i and i' of $h(w)$ with h-coordinates (j, d) and (j', d'), respectively. We have $i < i'$ if and only if $j < j'$ or if $j \leqslant j'$ and $d < d'$. And we have $i \leqslant i'$ if and only if $j < j'$ or if $j \leqslant j'$ and $d \leqslant d'$. If h is length-reducing, then $d = d' = 1$ and hence obviously $i < i'$ if and only if $j < j'$; and $i \leqslant i'$ if and only if $j \leqslant j'$. ☐

Lemma 5.28
Let φ be an atomic formula of the form $\mathrm{suc}(x, y)$, $\min(x)$, $\max(x)$, or empty. Let $\mathfrak{F}$ be a collection of fragments $\mathcal{F}$ with $(x = y) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$ for all $x, y \in \mathbb{V}_1$. Let V be a set of variables with $\mathrm{FV}(\varphi) \subseteq V$. For every non-erasing homomorphism $h\colon B^* \to A^*$ and every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\varphi)$ such that for $n = \max\{|h(b)| \mid b \in B\}$ the following hold:

$$h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi \text{ for all fragments } \mathcal{F} \in \mathfrak{F} \text{ and}$$
$$[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V}).$$

Proof. If h is non-erasing, then clearly $h(w)$ is empty only if w is empty. Consider a position i of $h(w)$ with h-coordinates (j, d). Then $i = 1$ is equivalent to $j = d = 1$; note that for erasing homomorphisms, in contrast, we may well have $i = 1$ but nonetheless $j > 1$. We therefore let

$$h_\delta^{-1}(\text{empty}) \ := \ \text{empty},$$
$$h_\delta^{-1}(\min(x)) \ := \ \begin{cases} h_\delta^{-1}(\top) \wedge \min(x) & \text{if } \delta(x) = 1, \\ \perp & \text{else.} \end{cases}$$

For the max-predicate we observe that $i = |h(w)|$ if and only if $j = |w|$ and $d = |h(b)|$ where $b = w[j]$. For erasing homomorphisms this would not necessarily be true.

We turn to the successor predicate. Consider positions i and i' of $h(w)$ with h-coordinates (j, d) and (j', d'), respectively. Suppose $i + 1 = i'$. There are two cases how this may happen. If $d' = 1$, then necessarily $j + 1 = j'$ and $d = |h(w[j])|$. Otherwise if $d' > 1$, then $j = j'$ and $d + 1 = d'$. (Again, for erasing homomorphisms this would not necessarily be valid.)

Let $C = \{b \in B \mid \delta(x) = |h(b)|\}$ be the set of labels such that $\delta(x)$ is the maximum position in the image under h. With this let

$$h_\delta^{-1}(\max(x)) \;:=\; h_\delta^{-1}(\top) \wedge \max(x) \wedge \lambda(x) \in C,$$

$$h_\delta^{-1}(\mathrm{suc}(x,y)) \;:=\; \begin{cases} h_\delta^{-1}(\top) \wedge \mathrm{suc}(x,y) \wedge \lambda(x) \in C & \text{if } \delta(y) = 1, \\ h_\delta^{-1}(\top) \wedge x = y & \text{else if } \delta(x) + 1 = \delta(y), \\ \bot & \text{else.} \end{cases}$$

It is easy to see that these formulae satisfy the syntactic property $h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi$ for all fragments $\mathcal{F}$. □

As for residuals, we also need to be able to change the remainder parameter for the modular predicate. Moreover, we only get closure under inverses of length-multiplying homomorphisms.

Lemma 5.29

Let φ be an atomic formula of the form $x \equiv r \pmod q$ for some $x \in \mathbb{V}_1$ and $r, q \in \mathbb{Z}$. Let $\mathfrak{F}$ be a collection of fragments $\mathcal{F}$ such that $\big(x \equiv s \pmod q\big) \leqslant_{\mathcal{F}} \big(x \equiv r \pmod q\big)$ for all $s \in \mathbb{Z}$. Let V be a set of variables containing x. For every length-multiplying homomorphism $h\colon B^* \to A^*$ and every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\varphi)$ such that for $n = \max\{|h(b)| \mid b \in B\}$ the following hold:

$$h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi \quad \textit{for all fragments } \mathcal{F} \in \mathfrak{F} \textit{ and}$$
$$[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V}).$$

Moreover, if h is length-preserving, then $h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi$ for all fragments $\mathcal{F}$.

Proof. Let h be length-multiplying, and let $n = |h(b)|$ for some $b \in B$. If $n = 0$ or if $\delta(y)$ is undefined for some $y \in V \cap \mathbb{V}_1$, then $h_\delta^{-1}(x \equiv r \pmod q) = \bot$. Let i be a position of $h(w)$ with h-coordinates (j, d). As h is length-multiplying, we have $i = n(j-1) + d$. Let $t = \gcd(q, n)$ be the greatest common divisor of q and n. Let $q = pt$, $n = \ell t$, and $r' = r + n - d$. With this $i \equiv r \pmod q$ if and only if $nj \equiv r' \pmod q$. Hence, if t is a divisor of r', then $nj \equiv r' \pmod q$ is equivalent to $\ell j \equiv r'/t \pmod p$. Since $\gcd(\ell, p) = 1$ there exists an integer ℓ^{-1} such that $\ell^{-1}\ell \equiv 1 \pmod p$ and $\ell j \equiv r'/t \pmod p$ if and only if $j \equiv \ell^{-1} r'/t \pmod p$. Now the latter is equivalent to the existence of $0 \leqslant k < t$ such that $j \equiv \ell^{-1} r'/t + kp \pmod q$.

These considerations lead to the following formula. If $r + n - \delta(x) \not\equiv 0 \pmod t$, then let $h_\delta^{-1}(x \equiv r \pmod q) = \bot$. Let otherwise $s = \ell^{-1}(r + n - \delta(x))/t$ and

$$h_\delta^{-1}(x \equiv r \pmod q) \;:=\; h_\delta^{-1}(\top) \wedge \bigvee_{k=0}^{t-1} x \equiv s + kp \pmod q.$$

The syntactic property $h_\delta^{-1}(x \equiv r \pmod q) \leqslant_{\mathcal{F}} \big(x \equiv r \pmod q\big)$ for all fragments $\mathcal{F} \in \mathfrak{F}$ is easy to verify.

Note that for length-preserving h, $h_\delta^{-1}(x \equiv r \pmod q) = h_\delta^{-1}(\top) \wedge x \equiv r \pmod q$ because $n = 1$, $t = 1$, and $s = r$. In this case $h_\delta^{-1}\big(x \equiv r \pmod q\big) \leqslant_{\mathcal{F}} \big(x \equiv r \pmod q\big)$ for all fragments $\mathcal{F}$. □

Lemma 5.30
Consider an atomic formula of the form $x \in X$ for some $x \in \mathbb{V}_1$ and $X \in \mathbb{V}_2$. Let V be a set of variables with $\mathrm{FV}(x \in X) \subseteq V$. For every homomorphism $h\colon B^* \to A^*$ and every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(x \in X)$ such that for $n \geqslant \max\{|h(b)| \mid b \in B\}$ the following hold:

$$h_\delta^{-1}(x \in X) \leqslant_{\mathcal{F}} (x \in X) \text{ for all MSO-stable fragments } \mathcal{F} \text{a nd}$$
$$[\![h_\delta^{-1}(x \in X)]\!]_{B,V_n} = h_\delta^{-1}([\![x \in X]\!]_{A,V}).$$

Moreover, if h is length-reducing, then $h_\delta^{-1}(x \in X) \leqslant_{\mathcal{F}} (x \in X)$ for all fragments $\mathcal{F}$.

Proof. Let i be a position of $h(w)$ with h-coordinates (j, d). By definition of h_δ, we see that i is an (a, J)-position of $h_\delta(w)$ for some $J \subseteq V$ with $X \in J$ if and only if j is a (b, J')-position of w for some $J' \subseteq V_n$ with $X_d \in J'$. Now, if $\delta(x) \notin \{1, \ldots, n\}$, then $h_\delta^{-1}(x \in X) = \bot$; otherwise let

$$h_\delta^{-1}(x \in X) \ := \ h_\delta^{-1}(\top) \wedge (x \in X_{\delta(x)}).$$

This formula is easily seen to satisfy the syntactic property $h_\delta^{-1}(x \in X) \leqslant_{\mathcal{F}} (x \in X)$ for all MSO-stable fragments $\mathcal{F}$. If h is length-reducing, then $h_\delta^{-1}(x \in X) \leqslant_{\mathcal{F}} (x \in X)$ for all fragments $\mathcal{F}$; notice that $X_1 = X$ by definition of V_n. □

Boolean connectives and quantifiers. Next, we give the following lifting lemmas: If for some formulae the inverse homomorphic images are definable, then so are the inverse homomorphic images of their Boolean combinations (Lemma 5.31), their first-order and second-order quantification (Lemma 5.32 and Lemma 5.33, respectively), and their modular counting quantification (Lemma 5.34). Moreover, the construction respects every family of homomorphisms and every collection of fragments (in the case of second-order quantification every collection of MSO-stable fragments; and for the modular counting quantifier every collection of mod-stable fragments).

Lemma 5.31
Let ψ be a formula of the form $\neg\varphi_1$ or $\varphi_1 \vee \varphi_2$ or $\varphi_1 \wedge \varphi_2$. Let $\mathfrak{F}$ be a collection of fragments, let $h\colon B^* \to A^*$ be a homomorphism, and let $n = \max\{|h(b)| \mid b \in B\}$. Suppose that for every set of variables V containing $\mathrm{FV}(\varphi_i)$, and every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exist formulae $h_\delta^{-1}(\varphi_i)$ for $i \in \{1, 2\}$ such that the following hold:

$$h_\delta^{-1}(\varphi_i) \leqslant_{\mathcal{F}} \varphi_i \text{ for all fragments } \mathcal{F} \in \mathfrak{F} \text{ and}$$
$$[\![h_\delta^{-1}(\varphi_i)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi_i]\!]_{A,V}).$$

Let V be a set of variables containing $\mathrm{FV}(\psi)$. For every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\psi)$ such that the following hold:

$$h_\delta^{-1}(\psi) \leqslant_{\mathcal{F}} \psi \text{ for all fragments } \mathcal{F} \in \mathfrak{F} \text{ and}$$
$$[\![h_\delta^{-1}(\psi)]\!]_{B,V_n} = h_\delta^{-1}([\![\psi]\!]_{A,V}).$$

Proof. The formulae for disjunction and conjunction are straightforward:

$$h_\delta^{-1}(\varphi_1 \vee \varphi_2) \ := \ h_\delta^{-1}(\varphi_1) \vee h_\delta^{-1}(\varphi_2),$$
$$h_\delta^{-1}(\varphi_1 \wedge \varphi_2) \ := \ h_\delta^{-1}(\varphi_1) \wedge h_\delta^{-1}(\varphi_2).$$

For negation we have to be more careful. For a word $w \in \mathcal{U}_{V_n}$ over the extended alphabet $B \times V_n$ the reason why $h_\delta(w)$ may not be in $[\![\varphi_1]\!]_{A,V}$ is as follows: The word

$h_\delta(w)$ may not permit to interpret all first-order variables in V. This happens whenever δ does not specify a valid offset. This means that $[\![\neg h_\delta^{-1}(\varphi_1)]\!]_{B,V_n}$ is in general not a subset of $[\![\top]\!]_{B,V_n}$. These spurious words are removed by an additional conjunctive term $h_\delta^{-1}(\top)$. We hence let $h_\delta^{-1}(\neg\varphi_1) := h_\delta^{-1}(\top) \wedge \neg h_\delta^{-1}(\varphi_1)$.

It is easy to verify that $h_\delta^{-1}(\psi)$ inherits the syntactic property of $h_\delta^{-1}(\varphi_i)$, *i.e.*, we have $h_\delta^{-1}(\psi) \leqslant_{\mathcal{F}} \psi$ for all $\mathcal{F} \in \mathfrak{F}$. Note that $\neg h_\delta^{-1}(\varphi_1) \leqslant_{\mathcal{F}} \neg\varphi_1$. □

Lemma 5.32
Consider a formula of the form $\exists x\, \varphi$ or $\forall x\, \varphi$ for some $x \in \mathbb{V}_1$. Let $\mathfrak{F}$ be a collection of fragments, let $h\colon B^* \to A^*$ be a homomorphism, and let $n = \max\{|h(b)| \mid b \in B\}$. Suppose that for every set of variables V containing $\mathrm{FV}(\varphi)$, and every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\varphi)$ such that the following hold:

$$h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi \quad \text{for all fragments } \mathcal{F} \in \mathfrak{F} \text{ and}$$
$$[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V}).$$

Let V be a set of variables containing $\mathrm{FV}(\varphi) \setminus \{x\}$. For every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exist formulae $h_\delta^{-1}(\exists x\, \varphi)$ and $h_\delta^{-1}(\forall x\, \varphi)$ such that for all $\mathcal{F} \in \mathfrak{F}$ the following hold:

$$h_\delta^{-1}(\exists x\, \varphi) \leqslant_{\mathcal{F}} \exists x\, \varphi, \qquad [\![h_\delta^{-1}(\exists x\, \varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\exists x\, \varphi]\!]_{A,V}),$$
$$h_\delta^{-1}(\forall x\, \varphi) \leqslant_{\mathcal{F}} \forall x\, \varphi, \qquad [\![h_\delta^{-1}(\forall x\, \varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\forall x\, \varphi]\!]_{A,V}).$$

Proof. Let i be a position of $h_\delta(w)$ with h-coordinates (j, d). For this position we have $h_\delta(w)[x/i] = h_{\delta[x/d]}(w[x/j])$, where $\delta[x/d]$ is given by $x \mapsto d$ and $y \mapsto \delta(y)$ if $y \neq x$. For the existential quantifier this leads to

$$h_\delta^{-1}(\exists x\, \varphi) \;:=\; \exists x \bigvee_{1 \leqslant d \leqslant n} h_{\delta[x/d]}^{-1}(\varphi),$$

where $h_{\delta[x/d]}^{-1}(\varphi)$ is the formula from the premise for the mapping $h_{\delta[x/d]}$ with respect to the set of variables $V \cup \{x\}$. Note that the h-coordinates (j, d) of every position of $h(w)$ satisfies $1 \leqslant d \leqslant n$ by choice of n. Duality of existential and universal quantifiers together with the rule for negation yields $h_\delta^{-1}(\forall x\, \varphi) := \forall x\, \neg h_{\delta[x/d]}^{-1}(\top) \vee h_{\delta[x/d]}^{-1}(\varphi)\big)$, where the negation of $h_{\delta[x/d]}^{-1}(\top)$ can be incorporated on the atomic level.

The syntactic properties are straightforward to verify. □

Lemma 5.33
Consider a formula of the form $\exists X\, \varphi$ or $\forall X\, \varphi$ for some $X \in \mathbb{V}_2$. Let $\mathfrak{F}$ be a collection of fragments, let $h\colon B^* \to A^*$ be a homomorphism, and let $n = \max\{|h(b)| \mid b \in B\}$. Suppose that for every set of variables V containing $\mathrm{FV}(\varphi)$, and every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\varphi)$ such that the following hold:

$$h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi \quad \text{for all fragments } \mathcal{F} \in \mathfrak{F} \text{ and}$$
$$[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V}).$$

Let V be a set of variables containing $\mathrm{FV}(\varphi) \setminus \{X\}$. For every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exist formulae $h_\delta^{-1}(\exists X\, \varphi)$ and $h_\delta^{-1}(\forall X\, \varphi)$ such that for all MSO-stable fragments $\mathcal{F} \in \mathfrak{F}$ the following hold:

$$h_\delta^{-1}(\exists X\, \varphi) \leqslant_{\mathcal{F}} \exists X\, \varphi, \qquad [\![h_\delta^{-1}(\exists X\, \varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\exists X\, \varphi]\!]_{A,V}),$$
$$h_\delta^{-1}(\forall X\, \varphi) \leqslant_{\mathcal{F}} \forall X\, \varphi, \qquad [\![h_\delta^{-1}(\forall X\, \varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\forall X\, \varphi]\!]_{A,V}).$$

Moreover, if h is length-reducing, then $h_\delta^{-1}(\exists X\,\varphi) \leqslant_{\mathcal{F}} \exists X\,\varphi$ and $h_\delta^{-1}(\forall X\,\varphi) \leqslant_{\mathcal{F}} \forall X\,\varphi$ for all fragments $\mathcal{F} \in \mathfrak{F}$.

Proof. Let i be a position of $h_\delta(w)$ with h-coordinates (j,d). By definition of h_δ, the position i contains X in its label if and only if position j of w contains X_d. This leads to

$$\begin{aligned} h_\delta^{-1}(\exists X\,\varphi) &:= \exists X_1 \cdots \exists X_N\; h_\delta^{-1}(\varphi),\\ h_\delta^{-1}(\forall X\,\varphi) &:= \forall X_1 \cdots \forall X_N\; h_\delta^{-1}(\varphi), \end{aligned}$$

where $N = \max\{1,n\}$ and $h_\delta^{-1}(\varphi)$ is the formula from the premise with respect to the the set of variables $V \cup \{X\}$. Note that for the h-coordinates (j,d) of every position of $h(w)$ we have $1 \leqslant d \leqslant n$ by choice of n.

The syntactic properties $h_\delta^{-1}(\exists X\;\varphi) \leqslant_{\mathcal{F}} (\exists X\;\varphi)$ and $h_\delta^{-1}(\forall X\;\varphi) \leqslant_{\mathcal{F}} (\forall X\;\varphi)$ hold for $\mathcal{F} \in \mathfrak{F}$ whenever either $\mathcal{F}$ is MSO-stable or h is length-reducing. In the latter case $N = 1$ and $X_1 = X$ by definition of V_n. ☐

The following lemma lifts the inverse homomorphism construction to modular counting quantifiers. It in particular applies to mod-stable fragments, but holds in a more general setting. Specifically, we do not need the closure under negation required by mod-stability.

Lemma 5.34
Consider a formula of the form $\exists^{r \bmod q} x\;\varphi$ for some $x \in \mathbb{V}_1$ and some $r,q \in \mathbb{Z}$. Let $\mathfrak{F}$ be a collection of fragments, let $h\colon B^* \to A^*$ be a homomorphism, and let $n = \max\{|h(b)| \mid b \in B\}$. Suppose that for every set of variables V containing $\mathrm{FV}(\varphi)$, and every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\varphi)$ such that the following hold:

$$\begin{aligned} &h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi \quad \text{for all fragments } \mathcal{F} \in \mathfrak{F} \text{ and}\\ &[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V}). \end{aligned}$$

Let V be a set of variables containing $\mathrm{FV}(\exists^{r \bmod q} x\;\varphi)$. For every $\delta\colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\exists^{r \bmod q} x\;\varphi)$ such that for every fragment $\mathcal{F} \in \mathfrak{F}$ that satisfies $\left(\exists^{s \bmod q} x\;\varphi\right) \leqslant_{\mathcal{F}} \left(\exists^{r \bmod q} x\;\varphi\right)$ (for all $s \in \mathbb{Z}$) the following hold:

$$\begin{aligned} &h_\delta^{-1}(\exists^{r \bmod q} x\;\varphi) \leqslant_{\mathcal{F}} \exists^{r \bmod q} x\;\varphi \quad \text{and}\\ &[\![h_\delta^{-1}(\exists^{r \bmod q} x\;\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\exists^{r \bmod q} x\;\varphi]\!]_{A,V}). \end{aligned}$$

Moreover, if h is length-reducing, then $h_\delta^{-1}(\exists^{r \bmod q} x\;\varphi) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\;\varphi)$ for all $\mathcal{F} \in \mathfrak{F}$.

Proof. For $d \in \mathbb{N}$ we denote by $\delta[x/d]\colon \mathbb{V}_1 \to_p \mathbb{N}$ the mapping $x \mapsto d$ and $y \mapsto \delta(y)$ if $y \neq x$. Let $h_{\delta[x/d]}^{-1}(\varphi)$ be the formula from the premise for the set of variables $V \cup \{x\}$. Let

$$h_\delta^{-1}(\exists^{r \bmod q} x\;\varphi) \;:=\; \bigvee_{s \in S} \bigwedge_{d=1}^{n} \exists^{s(d) \bmod q}\, x\colon\; h_{\delta[x/d]}^{-1}(\varphi),$$

where S is the set of functions $s\colon \{1,\dots,n\} \to \{0,\dots,q-1\}$ such that $\sum_{d=1}^{n} s(d) \equiv r \pmod q$. We say that a position is a *φ-position* if φ holds when x is interpreted by this position. The idea is that the φ-positions of $h_\delta(w)$ are partitioned; for every $d \in \{1,\dots,n\}$ the number of φ-positions of $h_\delta(w)$ originating from a position in w with offset d is counted separately (modulo q). The total sum of these counts then has to be r (modulo q). Note that S is finite and that for the h-coordinates (j,d) of every position of $h(w)$ we have $1 \leqslant d \leqslant n$ by choice of n. Hence, every φ-position of $h_\delta(w)$ is counted in precisely one of the terms of the conjunction. The syntactic property $h_\delta^{-1}(\exists^{r \bmod q} x\ \varphi) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$ for all $\mathcal{F} \in \mathfrak{F}$ with $\left(\exists^{s \bmod q} x\ \varphi\right) \leqslant_{\mathcal{F}} \left(\exists^{r \bmod q} x\ \varphi\right)$ is easily verified.

Suppose h is length-reducing, *i.e.*, $n \leqslant 1$. Consider first the case $n = 0$. If $r \equiv 0 \pmod q$, then $h_\delta^{-1}(\exists^{r \bmod q} x\ \varphi) = \top$; else $h_\delta^{-1}(\exists^{r \bmod q} x\ \varphi) = \bot$. If $n = 1$, then S contains only the function s with $s(1) = (r \bmod q)$ and we redefine $h_\delta^{-1}(\exists^{r \bmod q} x\ \varphi) = \exists^{r \bmod q} x\ h_{\delta[x/1]}^{-1}(\varphi)$. In both cases the formulae satisfy $h_\delta^{-1}(\exists^{r \bmod q} x\ \varphi) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$ for all $\mathcal{F} \in \mathfrak{F}$. $\square$

We are now ready to prove Proposition 5.6.

Proof of Proposition 5.6. An induction on the structure of φ shows that for all sets of variables V containing $\mathrm{FV}(\varphi)$, for all $\mathcal{C}$-homomorphisms $h \colon B^* \to A^*$, and for all $\delta \colon \mathbb{V}_1 \to_p \mathbb{N}$ there exists a formula $h_\delta^{-1}(\varphi)$ that satisfies $h_\delta^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi$ and $[\![h_\delta^{-1}(\varphi)]\!]_{B,V_n} = h_\delta^{-1}([\![\varphi]\!]_{A,V})$, where $n = \max\{|h(b)| \mid b \in B\}$. For the atomic modalities $\top$, $\bot$, $\lambda(x) \in C$, and $x = y$ this is Lemma 5.26; for $x < y$ and $x \leqslant y$ it is Lemma 5.27; for $\mathrm{suc}(x, y)$, $\min(x)$, $\max(x)$, and empty it is Lemma 5.28; for the modular predicate $x \equiv r \pmod q$ it is Lemma 5.29; and for $x \in X$ it is Lemma 5.30. Note that in all cases the lemmas do apply by the assumptions on $\mathcal{C}$.

For Boolean connectives, first-order quantification, second-order quantification, and modular quantification, this follows by induction and Lemma 5.31, Lemma 5.32, Lemma 5.33, and Lemma 5.34, respectively (with $\mathfrak{F} = \{\mathcal{F}\}$).

Using this, closure under inverse homomorphic images follows readily: Suppose φ is a sentences and let $h^{-1}(\varphi) = h_\delta^{-1}(\varphi)$ for an arbitrary δ. We have $h^{-1}(\varphi) \leqslant_{\mathcal{F}} \varphi$ and $h^{-1}(\mathcal{L}_A(\varphi)) = \mathcal{L}_B(h^{-1}(\varphi))$. In particular $\varphi \in \mathcal{F}$ implies $h^{-1}(\varphi) \in \mathcal{F}$, that is, if $L \in \mathcal{L}_A(\mathcal{F})$, then $h^{-1}(L) \in \mathcal{L}_B(\mathcal{F})$. $\square$

6. Two-Variable Logic $\mathrm{FO}^2[<]$ and Acyclic $\Sigma_2[<]$-Formulae

In this chapter we consider a rather unorthodox restriction of first-order formulae, namely the fragment of $\Sigma_2[<]$ that allows to query information about the relative order of positions only in an asymmetric way. For example, such formulae may contain an atomic query $x < y$, but are then not allowed to use $y < x$. This restriction can be captured by purely syntactic means, restricting a certain induced graph (called comparison graph) to have no cycles. The goal is to show that such acyclic $\Sigma_2[<]$-formulae are expressively complete for the two-variable fragment $\mathrm{FO}^2[<]$ of first-order logic. Remember that Σ_2 is the second half level of the first-order quantifier alternation hierarchy as introduced in Section 4.2.

The *comparison graph* of a formula φ is the directed graph with the set of all variables used in φ as vertices and edges E_φ defined by

- $E_\varphi = \{(x, y)\}$ if φ is one of $x < y$, $x \leqslant y$, $x = y$, or $y = x$,
- $E_\varphi = \emptyset$ if φ is any other atomic formula,

and the inductive rules $E_{\varphi \vee \psi} = E_{\varphi \wedge \psi} = E_\varphi \cup E_\psi$ and $E_{\neg\varphi} = \{(x, y) \mid (y, x) \in E_\varphi\}$ and $E_{\exists x\, \varphi} = E_\varphi$. Without negations, an edge from x to y thus means that somewhere a comparison is used that contains information on whether $x \leqslant y$. Negations reverse the direction of the order and thus the edges of the comparison graph.

Consider a fixed strict linear order $x_1 \prec x_2 \prec x_3 \prec \cdots$ of all first-order variables. We say that the comparison graph of φ is *acyclic* if $(x, y) \in E_\varphi$ implies $x \prec y$. In other words, the given ordering of the first-order variables determines the permitted usage of order comparisons. Note that equality cannot be used with an acyclic comparison graph at all. Also note that an acyclic comparison graph cannot have a cycle; however, not every formula whose comparison graph is cycle-free is acyclic. This is just an issue of naming the variables, however, and renaming the variables, we can find an equivalent acyclic formula for cycle-free formulae.

Formulae in $\Sigma_2[<]$ with cycle-free comparison graph do not form a fragment. Indeed, the disjunction of two cycle-free formulae may no longer be cycle free — just consider $x < y$ and $y < x$, for example. By fixing the order of the variables, $\Sigma_2[<]$-formulae with an acyclic comparison graph do form a fragment. Reading $x \leqslant y$ as a shorthand for $\neg(y < x)$, this acyclic fragment of $\Sigma_2[<]$ is order-stable, and Corollary 5.11 implies that it defines a $*$-variety. We now show that it actually coincides with a very famous fragment on language level.

Theorem 6.1
A language is definable in $\mathrm{FO}^2[<]$ *if and only if it is definable by a formula in* $\Sigma_2[<]$ *with an acyclic comparison graph.*

Proof. A *monomial* (of degree n) is a language of the form $P = A_1^* a_1 \cdots A_n^* a_n A_{n+1}^*$; it is *unambiguous* if each $w \in P$ has a unique factorization $w = w_1 a_1 \cdots w_n a_n w_{n+1}$ with

$w_i \in A_i^*$. It is well-known that every $\mathrm{FO}^2[<]$-definable language is a finite union of unambiguous monomials, see [TT02; DGK08].

For an unambiguous monomial P there exists $i \in \{1, \ldots, n\}$ such that $a_i \notin A_1 \cap A_{n+1}$. (Assuming the contrary, $(a_1 \cdots a_n)^2$ would admit two different factorizations.) By symmetry, we may suppose $a \notin A_1$ for some $a = a_i$. Making the first a-occurrence explicit, we see that P is a finite union of languages of the form $Q_1 a Q_2$, where Q_1 and Q_2 are unambiguous monomials with a smaller degree than P. By induction, Q_j is defined by some $\varphi_j \in \Sigma_2[<]$ with acyclic comparison graph. We may assume that the variables used by φ_1 and φ_2 are disjoint.

The next step is to relativize φ_1 to the factor to the left of the first a-position. Let

$$\varphi_1' \ := \ \exists x_1 \exists x_2 \left(\varphi_1'' \wedge \bigwedge_{j \in \{1,2\}} \lambda(x_j) = a \wedge \neg \exists y \ (y < x_j \wedge \lambda(y) = a)\right),$$

where x_1, x_2, y are new variables. Before turning to the formula φ_1'' note that x_1 and x_2 both specify the first a-position, thus ensuring $x_1 = x_2$ without actually using equality. This will be important for acyclicity.

The construction of φ_1'' is by induction on the formula. Let $\psi'' := \psi$ for atomic ψ, $(\psi_1 \vee \psi_2)'' := \psi_1'' \vee \psi_2''$ and $(\psi_1 \wedge \psi_2)'' := \psi_1'' \wedge \psi_2''$, and $(\exists z\ \psi)'' := \exists z\ (z < x_1 \wedge \psi'')$. For negations we have to be careful lest we should introduce a cycle. The variable x_1 is used to restrict quantification to positions smaller than the first a-position. By construction, all edges involving x_1 so far have x_1 as a target (with the sole exception of the edge (x_1, y), of course). After negations, x_1 will be the source, and we thus have to switch to the second variable x_2 to ensure that no cycles are introduced. Let $(\neg\psi)'' := \neg(\psi''[x_1 \leftrightarrow x_2])$, where $\psi''[x_1 \leftrightarrow x_2]$ denotes the formula obtained from ψ'' by interchanging x_1 and x_2.

The formula φ_1' is true if and only if there is an a-position and φ_1 holds on the factor before the first a-position. One can verify that φ_1' is cycle-free since φ_1 is cycle-free. A similar construction yields a cycle-free formula φ_2''', evaluating φ_2 on the factor beyond the first a-position. The formula $\varphi_1' \wedge \varphi_2'''$ defines $Q_1 a Q_2$. This shows that P is a disjunction of cycle-free Σ_2-formulae. Renaming variables yields an acyclic Σ_2-formula for P.

We come to the converse implication. It is well-known that a language is definable in $\mathrm{FO}^2[<]$ if and only if its syntactic monoid is in **DA**, see for instance [DGK08]. A finite monoid M is in **DA** if and only if it is aperiodic and there exists n such that $(xy)^n x (xy)^n = (xy)^{3n}$ for all $x, y \in M$. (See Section 2.3 for a definition of these algebraic terms.) It is also well-known that full first-order logic $\mathrm{FO}[<]$ corresponds to aperiodic finite monoids [DG08], so in particular every $\Sigma_2[<]$-definable language has an aperiodic syntactic monoid. It remains to show the identity mentioned earlier.

The implication $p(uv)^{3n}q \models \varphi \Rightarrow p(uv)^n u (uv)^n q \models \varphi$ holds for all $\varphi \in \Sigma_2[<]$, even without acyclicity condition [TT02; DGK08] and it suffices to show the converse implication for some sufficiently large n. Negations reverse edges, which makes our arguments unnecessarily complicated. The first step is therefore to eliminate negations. Formulae in prenex normal form are normalized to $\Sigma_2[<, \leqslant]$-formulae with no negation at all in its propositional matrix. For this we use De Morgan's laws to move negations inwards over disjunction and conjunction and the equivalence $\neg(x < y) \equiv (y \leqslant x)$ to eliminate negations. Note that acyclicity of the comparison graph is preserved and that $\leqslant$ is allowed henceforth in the syntax.

Let us start with an informal sketch of the idea, best explained in terms of a game. (Readers familiar with Ehrenfeucht-Fraïssé-games will recognize a strong reminiscence of those, yet it is not an Ehrenfeucht-Fraïssé-game in the usual sense due to an unsymmetrical winning condition; see Remark 5.14 for a brief explanation of Ehrenfeucht-Fraïssé-games.) The game is played between two players, called *Spoiler* and *Duplicator*, on the words $p(uv)^n u(uv)^n q$ and $p(uv)^{3n} q$. At the beginning, both players agree on a number ℓ. Spoiler starts on $p(uv)^n u(uv)^n q$, Duplicator starts on $p(uv)^{3n} q$. The goal of Spoiler is to reveal a difference between the two words, whereas Duplicator tries to conceal any difference. The game is played in two rounds according to the following rules.

In the first round, Spoiler places k pebbles, $k \leqslant \ell$, with labels $x_1, \ldots, x_k$ on positions of the word $p(uv)^n u(uv)^n q$; Duplicator answers by placing pebbles $x_1, \ldots, x_k$ on $p(uv)^{3n} q$. Then the game boards are switched and Spoiler places pebbles $x_{k+1}, \ldots, x_\ell$ on $p(uv)^{3n} q$ in the second round; Duplicator places pebbles $x_{k+1}, \ldots, x_\ell$ on $p(uv)^n u(uv)^n q$. Duplicator wins the game if all pebbles have the same label on both words, and if $x_i < x_j$ (respectively, $x_i \leqslant x_j$) on $p(uv)^n u(uv)^n q$ implies $x_i < x_j$ (respectively, $x_i \leqslant x_j$) on $p(uv)^{3n} q$ for every pair of pebbles with $x_i \prec x_j$; otherwise Spoiler wins. Switching the game board corresponds to a negation in the logic, which is why Spoiler may only switch once in the game for Σ_2.

To show the desired implication, we take Duplicator's part and describe her winning strategy. After Spoiler placed his pebbles in the first round, Duplicator places all pebbles outside the central u on the respective position in the prefix $p(uv)^n$ or suffix $(uv)^n q$ of her word. She places the remaining pebbles so as to make as many atomic formulae true on $p(uv)^{3n} q$ as possible. This greedy approach needs acyclicity.

Having chosen n large enough, we know that on $p(uv)^n u(uv)^n q$, there are large "gaps" with no pebbles left and right of the central factor u. This allows Duplicator to win the second round: No matter how Spoiler places his pebbles on $p(uv)^{3n} q$, Duplicator can use the gaps on $p(uv)^n u(uv)^n q$ and acyclicity to obtain a situation where as many atomic formulae as possible are false on $p(uv)^n u(uv)^n q$. The matrix ψ of the original formula is a positive Boolean combinations of its atoms. Thus, if ψ is true on $p(uv)^n u(uv)^n q$, then it is true on $p(uv)^{3n} q$ by monotonicity. The following formalizes this idea.

Let $L \subseteq A^*$ be defined by the $\Sigma_2[<]$-sentence φ with acyclic comparison graph. Using the procedure described above, we may suppose $\varphi \in \Sigma_2[<, \leqslant]$ to be in prenex form $\exists x_1 \cdots \exists x_k \neg \exists x_{k+1} \cdots \exists x_\ell \; \psi$ for some negation-free propositional ψ. We shall show that $p(uv)^n u(uv)^n q \models \varphi$ implies $p(uv)^{3n} q \models \varphi$ for all $n \geqslant \ell(\ell+1)$ and all $u, v, p, q \in A^*$.

Suppose $p(uv)^n u(uv)^n q \models \varphi$. Choose positions $z_1, \ldots, z_k$ of $p(uv)^n u(uv)^n q$ such that there do not exist positions $z_{k+1}, \ldots, z_\ell$ such that the formula ψ holds on $p(uv)^n u(uv)^n q$ with x_i interpreted by z_i.

By choice of n there is a factorization $p(uv)^n u(uv)^n q = p'(uv)^\ell w(uv)^\ell q'$ such that

- $p'(uv)^\ell$ is a prefix of $p(uv)^n$,
- $(uv)^\ell q'$ is a suffix of $(uv)^n q$, and
- none of $z_1, \ldots, z_k$ is in one of the factors $(uv)^\ell$ of this factorization (as $k \leqslant \ell$).

The last condition more specifically means $z_j \in I_1 \cup I_2 \cup I_3$ for all $1 \leqslant j \leqslant k$, where

- $I_1 = \left\{ z \in \mathbb{N} \;\middle|\; 1 \leqslant z \leqslant |p'| \right\}$,
- $I_2 = \left\{ z \in \mathbb{N} \;\middle|\; |p'(uv)^\ell| < z \leqslant |p'(uv)^\ell w| \right\}$, and
- $I_3 = \left\{ z \in \mathbb{N} \;\middle|\; |p'(uv)^\ell w(uv)^\ell| < z \leqslant \left|p'(uv)^\ell w(uv)^\ell q'\right| \right\}$.

We can factorize $p(uv)^{3n}q$ as $p'(uv)^{\ell}w'(uv)^{\ell}q'$ for some w'. Let I_1', I_2', and I_3' be defined analogously to I_1, I_2, and I_3 with w replaced by w'. Let $\pi : \mathbb{N} \to \mathbb{N}$ be an order-preserving injection, mapping I_j to I_j'. The factors $(uv)^{\ell}$ left and right of the center are informally called the left and the right *gap*.

We construct an interpretation of the variables $x_1, \ldots, x_k$ on $p(uv)^{3n}q$ as follows. For $z_j \in I_1 \cup I_3$ we set $z_j' = \pi(z_j)$. For positions $z_j \in I_2$ let $z_j' \in I_2'$ be such that z_j' and z_j have the same label and such that $z_j' < z_{j'}'$ if and only if $x_j \prec x_{j'}$. Note that this is possible as w' contains a factor $(uv)^{\ell}$.

Assume that we can choose positions $z_{k+1}', \ldots, z_{\ell}'$ of $p(uv)^{3n}q$ such that ψ is false on $p(uv)^{3n}q$ with x_i interpreted by z_i'. We want to show that this is impossible, thus implying $p(uv)^{3n}q \models \varphi$. For this we construct an interpretation $z_{k+1}, \ldots, z_{\ell}$ of $x_{k+1}, \ldots, x_{\ell}$ such that ψ is false on $p(uv)^n u(uv)^n q$ with x_i interpreted by z_i, contradicting the choice of $z_1, \ldots, z_k$.

For $z_j' \in I_1' \cup I_3'$ let $z_j = \pi^{-1}(z_j')$. For $z_j' \notin I_1' \cup I_3'$ we have to be more careful. Let

- $L(x_j) = \{x_i \mid i \in \{1, \ldots, k\},\, z_j' \geqslant z_i' \in I_2'\}$ contain the variables in I_2' left of z_j',
- $R(x_j) = \{x_i \mid i \in \{1, \ldots, k\},\, z_j' \leqslant z_i' \in I_2'\}$ contain the variables in I_2' right of z_j'.

At least one of the following cases applies:

1. $x_i \prec x_j$ for all $x_i \in L(x_j)$, or
2. $x_j \prec x_i$ for all $x_i \in R(x_j)$.

Indeed, assuming contrary there are $x_i \in L(x_j)$ and $x_{i'} \in R(x_j)$ such that $x_{i'} \prec x_j \prec x_i$. By definition $z_i' \leqslant z_{j'}'$. Thus $x_i \preccurlyeq x_{j'}$ by construction, contradicting $x_{i'} \prec x_i$.

Let L consist of all x_j with $j \in \{k+1, \ldots, \ell\}$ for which case (1) applies. Let R consist of all x_j with $j \in \{k+1, \ldots, \ell\}$ for which case (2) applies. Using this classification of variables, allows to finally declare $z_{k+1}, \ldots, z_{\ell}$. The idea is that placing variables in L within the left gap only makes atomic formulae false but never true, and similarly placing variables in R within the right gap only makes more atomic formulae false.

Formally, the positions z_j with $x_j \in L$ are set label-respecting in the range between $|p'| + 1$ and $|p'(uv)^{\ell}|$ such that $z_j < z_{j'}$ for all $x_j, x_{j'} \in L$ with $x_{j'} \prec x_j$. The remaining positions z_j with $x_j \in R \setminus L$ are similarly set label-respecting in the range between $|p'(uv)^{\ell}w| + 1$ and $|p'(uv)^{\ell}w(uv)^{\ell}|$ such that $z_j < z_{j'}$ for all $x_j, x_{j'} \in R \setminus L$ with $x_{j'} \prec x_j$. (For the variables in $L \cap R$ it does not really matter whether they are placed within the left gap (as we did) or the right gap; it is important to do it consistently for all variables in $L \cap R$, though.)

Under this interpretation, every atomic formula that is true on $p(uv)^n u(uv)^n q$ is also true on $p(uv)^{3n}q$. Since ψ does not contain negations, it is monotonic in its atoms. Therefore, ψ is false on $p(uv)^n u(uv)^n q$ as it is false on $p(uv)^{3n}q$, a contradiction. $\square$

7. Stutter-Invariant Piecewise Testable Languages

This chapter investigates a second example of an unorthodox fragment, namely the Boolean closure of purely existential formulae without any negation, denoted by $\mathbb{B}\Sigma_1^+$. In contrast to FO_1, this fragment disallows negations even over atomic formulae. More specifically, we are interested in the fragment $\mathbb{B}\Sigma_1^+[\leqslant]$ with non-strict order as only numerical predicate.

The usual trick of reading $x < y$ as a shorthand for $\neg(y \leqslant x)$ does not work within this fragment. We shall see in this chapter that this is not just for want of a creative replacement of the strict order but that $\mathbb{B}\Sigma_1^+[\leqslant]$ is truly less expressive than $\mathbb{B}\Sigma_1^+[<]$. Indeed, $\mathbb{B}\Sigma_1^+[\leqslant]$ can define only languages that are stutter-invariant, while it is not hard to see that there are languages definable in $\mathbb{B}\Sigma_1^+[<]$ that are not stutter-invariant.

A language $L \subseteq A^*$ is *piecewise testable* if it is a finite Boolean combination of languages of the form $A^*a_1 \cdots A^*a_nA^*$ for $a_i \in A$. It is *stutter-invariant* if $paq \in L$ if and only if $paaq \in L$ for all words $p, q \in A^*$ and all letters $a \in A$.

A seminal result due to Simon shows that a regular language is piecewise testable if and only if its syntactic monoid is $\mathcal{J}$-trivial [Sim75].[1] Moreover, a regular language is stutter-invariant if and only if every letter has an idempotent image under the syntactic homomorphism. Both conditions can be checked effectively on the syntactic homomorphism and consequently it is decidable whether a regular language is piecewise testable and stutter-invariant.

Klíma and Polák developed a general framework for considering the stutter-invariant fragment of a variety [KP08]. Among others, they also consider the variety of piecewise testable languages. Stutter-invariant piecewise testable languages also arise in the study of so-called quantum automata, *cf.* [CB12].

The following theorem shows that this decision procedure for stutter-invariant piecewise testable languages transfers to definability in the fragment $\mathbb{B}\Sigma_1^+[\leqslant]$.

Theorem 7.1
Let $L \subseteq A^*$ be a language. The following are equivalent:

1. *L is definable in $\mathbb{B}\Sigma_1^+[\leqslant]$.*
2. *L is piecewise testable and stutter-invariant.*
3. *L is a Boolean combination of languages $A^*a_1 \cdots A^*a_nA^*$ with $a_i \neq a_{i+1}$ for all i.*

Proof. (1) $\Rightarrow$ (2): Any $\mathbb{B}\Sigma_1^+[\leqslant]$-definable language is $\mathrm{FO}_1[<]$-definable. The fragment $\mathrm{FO}_1[<]$ is well-known to be expressively complete for piecewise testable languages, see *e.g.* [DGK08]. It remains to show that L is stutter-invariant.

Consider a sentence φ in the negation-free fragment $\Sigma_1^+[\leqslant]$. Without restriction suppose φ is of the form $\exists x_1 \cdots \exists x_n\ \psi$ with a propositional formula ψ. Let $u = paq$ and let $u' = paaq$ for some $p, q \in A^*$ and $a \in A$. We shall show $u \models \varphi$ if and only if $u' \models \varphi$.

[1] $\mathcal{J}$-trivial means that Green's $\mathcal{J}$-relation is the identity. Refer to Section 2.3 for details about the algebraic terms such as syntactic monoid, the relation $\mathcal{J}$, and idempotent elements. For the present considerations, their precise meaning is not all too important, however.

First suppose $u \models \varphi$. Let $k_1, \ldots, k_n$ be such that $u, k_1, \ldots, k_n \models \psi(x_1, \ldots, x_n)$. Let $k'_i = k_i$ if $k_i \leqslant |pa|$ and $k'_i = k_i + 1$ otherwise. Each position k_i of u has the same label as the position k'_i of u'. Moreover, $u, k_i, k_j \models (x_i \leqslant x_j)$ if and only if $u', k'_i, k'_j \models (x_i \leqslant x_j)$. Therefore, $u', k'_1, \ldots, k'_n \models \psi(x_1, \ldots, x_n)$ and thus $u \models \varphi$.

Next suppose $u' \models \varphi$ and let $k'_1, \ldots, k'_n$ be such that $u', k'_1, \ldots, k'_n \models \psi(x_1, \ldots, x_n)$. Let $k_i = k'_i$ if $k'_i \leqslant |pa|$ and let $k_i = k'_i - 1$ otherwise. Again, the label of the position k_i of u and that of the position k'_i of u' is the same. If $u', k'_i, k'_j \models (x_i \leqslant x_j)$, then $u, k_i, k_j \models (x_i \leqslant x_j)$. Observe, however, that the converse need not be true; *i.e.*, it is well possible that $u', k'_i, k'_j \models (x_i \leqslant x_j)$ is not true but $u, k_i, k_j \models (x_i \leqslant x_j)$ is. (Namely for $k'_i = |paa|$ and $k'_j = |pa|$.) Then again this is not a problem. We just showed that every atom of ψ that is true on u' is also true on u. Being negation-free, ψ is monotonic in its atoms and thus true on u. To be more precise $u, k_1, \ldots, k_n \models \psi(x_1, \ldots, x_n)$, which shows $u \models \varphi$. This shows that every $\Sigma_1^+[\leqslant]$-definable language is stutter-invariant. The claim follows because stutter-invariant languages are closed under Boolean operations.

(2) $\Rightarrow$ (3): Let $L \subseteq A^*$ be piecewise testable and stutter-invariant. The intersection of two languages of the form $A^*a_1 \cdots A^*a_nA^*$ is itself of the same form. We can hence write $L = \bigcup_{i=1}^{s} \big(P_i \setminus \bigcup_{j=1}^{t} Q_{i,j}\big)$ where P_i and $Q_{i,j}$ are languages of the form $A^*a_1 \cdots A^*a_nA^*$.

We introduce a contraction operator $\mathrm{ct}(P)$ as follows. Let $\mathrm{ct}(P) = A^*a_1 \cdots A^*a_nA^*$ for a language P presented by the expression $P = (A^*a_1)^{e_1} \cdots (A^*a_n)^{e_n}A^*$, with positive integers e_i and letters $a_i \neq a_{i+1}$. In other words, any two successive a_i that are equal are contracted. Note that $P \subseteq \mathrm{ct}(P)$ and that $\mathrm{ct}(P)$ is stutter-invariant. We extend this contraction to L by setting $\mathrm{ct}(L) = \bigcup_i \big(\mathrm{ct}(P_i) \setminus \bigcup_j \mathrm{ct}(Q_{i,j})\big)$. We shall show $L = \mathrm{ct}(L)$. Note that the language $\mathrm{ct}(L)$ is stutter-invariant.

We start with the inclusion $\mathrm{ct}(L) \subseteq L$. Suppose $u \in \mathrm{ct}(P_i)$ and $u \notin \bigcup_j \mathrm{ct}(Q_{i,j})$ for some i. Let $\mathrm{ct}(P_i) = A^*a_1 \cdots A^*a_nA^*$ and let $u = v_0a_1 \cdots v_{n-1}a_nv_n$. By definition of contraction, there exists $u' = v_0a_1^{e_1} \cdots v_{n-1}a_n^{e_n}v_n$ for some e_j such that $u' \in P_i$. As the languages $\mathrm{ct}(Q_{i,j})$ are stutter-invariant, $u' \notin \mathrm{ct}(Q_{i,j}) \supseteq Q_{i,j}$ for all j. This shows $u' \in P_i \setminus \bigcup_j Q_{i,j} \subseteq L$. By stutter-invariance we conclude $u \in L$.

Next $L \subseteq \mathrm{ct}(L)$ is shown. Under the assumption that there be u with $u \in L \setminus \mathrm{ct}(L)$ we construct an infinite sequence of words $u = u_0, u_1, \ldots$ and an infinite sequence of sets $\emptyset = I_0 \subsetneq I_1 \subsetneq \cdots \subseteq \{1, \ldots, s\}$ with the following properties for all k:

1. $u_k \in L \setminus \mathrm{ct}(L)$, and
2. $u_k \in \bigcup_j Q_{i,j}$ for all $i \in I_k$.

Clearly such a strictly increasing chain of subsets of a fixed finite set cannot be. Therefore, the assumption must be false.

The construction of the sequences is inductive in nature. For $k = 0$, where $u_0 = u$ and $I_0 = \emptyset$, property (1) holds by assumption and (2) is vacuously true.

Suppose $k \geqslant 1$. Since $u_{k-1} \in L$ there exists $i \notin I_{k-1}$ such that $u_{k-1} \in P_i \subseteq \mathrm{ct}(P_i)$. Since $u_{k-1} \notin \mathrm{ct}(L)$, we necessarily have $u_{k-1} \in \mathrm{ct}(Q_{i,j})$ for some j. Suppose that $\mathrm{ct}(Q_{i,j}) = A^*a_1 \cdots A^*a_nA^*$ and $u_{k-1} = v_0a_1 \cdots v_{n-1}a_nv_n$. By definition of contraction $u_k = v_0a_1^{e_1} \cdots v_{n-1}a_n^{e_n}v_n \in Q_{i,j}$ for some e_j. Stutter-invariance of the languages L and $\mathrm{ct}(L)$ establishes (1). Let $I_k = I_{k-1} \cup \{i\}$. We have $u_k \in \bigcup_j Q_{\ell,j}$ for all $\ell \in I_{k-1}$ because u_{k-1} is a scattered subword of u_k. This shows (2).

(3) $\Rightarrow$ (1): The language $A^*a_1 \cdots A^*a_nA^*$ with $a_i \neq a_{i+1}$ for all i is defined by the sentence $\exists x_1 \cdots \exists x_n \ \bigwedge_{1 \leqslant i \leqslant n} \ \lambda(x_i) = a_i \wedge \ \bigwedge_{1 \leqslant i < n} \ x_i \leqslant x_{i+1}$, which is a sentence in $\mathbb{B}\Sigma_1^+[\leqslant]$. Note that $x_i \leqslant x_{i+1}$ implies $x_i < x_{i+1}$, as $a_i \neq a_{i+1}$. $\square$

8. On the Expressive Power of the Successor Predicate

The purpose of this chapter is to study the influence of the successor predicate on the expressive power of fragments of FO+MOD. More specifically, how does the expressive power of a fragment $\mathcal{F}[\mathcal{N}, \mathrm{suc}, \min, \max]$ for some signature $\mathcal{N}$ relate to that of $\mathcal{F}[\mathcal{N}]$ without predicates for successor, minimum, or maximum? Obviously, the successor predicate does not provide additional expressive power for powerful fragments such as full first-order logic: x is the predecessor of y if and only if x is smaller than y and no position is strictly in between x and y, that is, $\mathrm{suc}(x, y)$ is equivalent to $x < y \wedge \forall z\ (z \leqslant x \vee y \leqslant z)$. This reasoning does not work for general fragments $\mathcal{F}$, however: First, several order comparisons are introduced, although the order predicate may be unavailable in $\mathcal{F}$; second, a new quantifier is needed, potentially violating restrictions of the quantifier depth and the quantifier alternation; and third, a new variable z is necessary.

This chapter strives to give a finer-grained answer to the above question. Section 8.1 pursues a *sliding window approach* to capture the expressive power of an additional successor predicate. The basic idea is that equipping formulae with the ability to sample the surrounding labels should more or less capture what the fragment gains from using the successor predicate. This look-around is realized by a sliding window of fixed diameter that is dragged over the word: The window is centered on each position of the word, and the original letter of that position is replaced by the factor visible through the window. This sliding window approach is well-known in the literature [Str85; BP91; Kuf13]. For want of a formal notion of fragments, classically *ad hoc* methods were applied that are tailor-made for the specific fragment considered.

In contrast, Section 8.1 implements the sliding window approach more universally on fragment level. The solution to the sliding window approach, given in Theorem 8.6, is not completely satisfactory, though, because it necessitates more general quantifiers. It is nonetheless quite far-reaching as the generalized quantifiers can often be expressed within the fragment itself. For this we have to leave the abstract level and start to tailor the approach to concrete fragments. Proposition 8.17 exemplifies this, subsuming in particular the first-order quantifier alternation hierarchy.

Building on the sliding window approach, Section 8.2 studies the power of the successor predicate in the presence of a neutral letter, *i.e.*, a letter that can be inserted or deleted anywhere without changing membership in the language. The main result of this section is that if the sliding window approach works for the fragment, then all languages with a neutral letter that are definable in $\mathcal{F}[\mathcal{N}, \mathrm{suc}, \min, \max, \mathrm{mod}]$ are already $\mathcal{F}[\mathcal{N}]$-definable. An algebraic corollary can be drawn from this logic-related result. If a logic fragment $\mathcal{F}[\mathcal{N}, \mathrm{suc}, \min, \max]$ admits successor-free sliding window formulae and characterizes a variety of semigroups $\mathbf{V}$, then languages recognized by monoids in $\mathbf{V}$ are definable in $\mathcal{F}[\mathcal{N}]$. This algebraic corollary to the neutral letter approach is used later in Chapter 12, where the algebraic description quantifier alternation hierarchy within $\mathrm{FO}^2[<]$ is reduced to that of the alternation hierarchy in $\mathrm{FO}^2[<, \mathrm{suc}, \min, \max]$.

8.1. The Successor Predicate and Sliding Windows

There are two formal issues to overcome when implementing the above mentioned sliding window. First, the sliding window may protrude the word near the borders (in which case we pad with a new blank symbol), and second, our logic framework does not allow factors as labels (which is remedied by encoding factors into single labels). The following formalizes this.

Let $\square$ be an arbitrary but fixed *blank symbol* with $\square \notin \Lambda$, and let $[\cdot]$ be an arbitrary but fixed encoding, mapping any word u over the alphabet $\Lambda \cup \{\square\}$ injectively to $[u] \in \Lambda$. Note that such an encoding exists because Λ is infinite.

Definition 8.1 (Sliding window)
For any alphabet $A \subseteq \Lambda$ and any integer $N \geqslant 0$, let $A_{(N)}$ be the sliding window alphabet *of radius N consisting of all encodings $[u]$ of words $u \in (A \cup \{\square\})^{2N+1}$ of length $(2N+1)$.*

The sliding window *of radius N is the length-preserving map $\sigma_N \colon \Lambda^* \to \Lambda^*_{(N)}$ with $a_1 \cdots a_n \mapsto b_1 \cdots b_n$ for $n \geqslant 0$ and $a_i \in \Lambda$, where for all $i \in \{1, \dots, n\}$ the i^{th} letter is $b_i = [a_{i-N} \cdots a_i \cdots a_{i+N}]$ with $a_j = \square$ for $j \notin \{1, \dots, n\}$.*

The lower case Greek letter σ for "*s*" is a mnemonic for "*s*liding window". Any word u can be recovered from $\sigma_N(u)$ by projecting to the letter in the center of each $\Lambda_{(N)}$-letter, *i.e.*, $u = \pi(\sigma_N(u))$ for the homomorphism $\pi \colon \Lambda^*_{(N)} \to \Lambda^*$ given by $[a_{-N} \cdots a_0 \cdots a_N] \mapsto a_0$.

Note that for $N \geqslant 1$ the mapping σ_N is not a homomorphism, and not all words in $\Lambda^*_{(N)}$ are obtained by applying the sliding window σ_N to words in Λ^*.

A word $u \in \Lambda^*_{(N)}$ is said to be *well-formed* with respect to σ_N if $u \in \sigma_N(\Lambda^*)$, *i.e.*, if it is the image under σ_N of some word in Λ^*. Well-formed words with respect to σ_1 include ε, $[\square a \square]$, and $[\square ab][abb][bb\square]$, for example, being the image of ε, a, and abb, respectively. The word $[abb][bb\square]$ is not well-formed as the first letter $[\square ab]$ is missing; $[\square ab][bb\square]$ is not well-formed because the overlap does not match.

The generic scheme of the sliding window approach for a fragment $\mathcal{F}$ is to show the following sliding window property.

Definition 8.2 (Sliding window property)
A fragment $\mathcal{F}$ has the sliding window property *if for every language $L \subseteq A^*$ the following conditions are equivalent:*

1. *L is definable in $\mathcal{F}$.*
2. *There exists an integer $N \geqslant 0$ and a language $K \subseteq A^*_{(N)}$ definable in $\mathcal{F}$ by a sentence that does not use successor, minimum, or maximum predicates with $L = \sigma_N^{-1}(K)$.*

The condition $L = \sigma_N^{-1}(K)$ in property (2) in particular requires that the sentence defining K need to be consistent only on well-formed words; *i.e.*, its behavior on other words does not matter. This motivates the following notion: A formula ψ is a *sliding window formula* of radius N for φ if $u \models \varphi$ is equivalent to $\sigma_N(u) \models \psi$ for all $u \in \Lambda^*$.

The goal of this section, stated in this terminology, is to prove the sliding window property with as few assumptions on the fragment as possible. The relevance of the sliding window property is twofold. First, from a model-theoretic point of view, it is natural *per se* to describe a fragment with successor predicates in terms of its counterpart without successor. A more practical reason is that the sliding window property immediately yields an algebraic transfer theorem in terms of so-called semidirect products with definite monoids: Suppose $\mathcal{F}[\mathcal{N}, \mathrm{suc}, \min, \max]$ has the sliding window property. If the language family $\mathcal{L}(\mathcal{F}[\mathcal{N}])$ corresponds to the variety $\mathbf{V}$, then the language family $\mathcal{L}(\mathcal{F}[\mathcal{N}, \mathrm{suc}, \min, \max])$ corresponds to $\mathbf{V} * \mathbf{D}$.[1] This is particularly interesting when the variety $\mathbf{V}$ is known to be local, *i.e.*, if $\mathbf{V} * \mathbf{D} = \mathbf{LV}$. Deciding definability in $\mathcal{L}(\mathcal{F}[\mathcal{N}, \mathrm{suc}, \min, \max])$ in such a case reduces to deciding definability in $\mathcal{L}(\mathcal{F}[\mathcal{N}])$ by testing each local monoid of the syntactic semigroup for membership in $\mathbf{V}$.

Let us note that the sliding window approach does not work well for MSO-fragments because adjoining the successor predicate can lead to a huge gap in the expressive power. The fragment MSO[=], for example, is characterized by the ability to count the number of occurrences of letters up to a certain threshold. In contrast to this very small subfamily of regular languages, every regular language is definable in MSO[=, suc] with the successor predicate available *cf.* [Büc60].

We start the proof of the sliding window property with the easy direction from condition (2) to (1), using the successor predicate to eliminate sliding window letters.

A fragment $\mathcal{F}$ is *factor-stable* if the following conditions hold for all contexts μ, all $x, y \in \mathbb{V}_1$, all $q, r \in \mathbb{Z}$, and all formulae φ:

1. $\mu(\varphi) \in \mathcal{F}$ and $\mathrm{FV}\big(\mu(\mathrm{suc}(x, y))\big) \subseteq \mathrm{FV}(\mu(\varphi))$ implies $\mu(\mathrm{suc}(x, y)) \in \mathcal{F}$.
2. $\mu(\mathrm{suc}(x, y)) \in \mathcal{F}$ implies $\{\mu(\min(y)), \mu(\max(x))\} \subseteq \mathcal{F}$.
3. $\mu(\exists x\ \varphi) \in \mathcal{F}$ implies there exists $z \neq x$ with $\mu(\exists x \exists z\ \varphi) \in \mathcal{F}$.
4. $\mu(\forall x\ \varphi) \in \mathcal{F}$ implies there exists $z \neq x$ with $\mu(\forall x \forall z\ \varphi) \in \mathcal{F}$.
5. $\mu(\exists^{r \bmod q} x\ \varphi) \in \mathcal{F}$ implies there exists $z \neq x$ with $\mu(\exists^{r \bmod q} x\, \exists^{1 \bmod q} z\ \varphi) \in \mathcal{F}$.

Note that factor-stability is incompatible with quantifier depth, but respects quantifier alternations.

Factor-stability suffices to sample the environment at quantification time, leading to the following proposition.

Proposition 8.3
For all sentences $\psi \in$ FO+MOD, all finite alphabets A, and all integers $N \geqslant 0$ there exists a sentence φ such that $\mathcal{L}_A(\varphi) = \sigma_N^{-1}(\mathcal{L}_{A_{(N)}}(\psi))$ and $\varphi \leqslant_{\mathcal{F}} \psi$ for every factor-stable fragment $\mathcal{F}$. Moreover, φ can be chosen such that $\mathrm{qd}(\varphi) \leqslant \mathrm{qd}(\psi) + N$.

Proof. The idea is to query the factor around every newly quantified variable at quantification time. Formally, we are going to construct a formula $\langle\psi\rangle_f$ for ψ and $f \colon V \to A_{(N)}$ with $\mathrm{FV}(\psi) \subseteq V$. This formula is true on a word if and only if ψ is true on the word with the sliding window applied, provided $f(x)$ correctly specifies the factor around the position x. That $f(x)$ is consistent in this sense is ensured at quantification time of x. The proposition follows by setting φ to the sentence $\langle\psi\rangle_\emptyset$.

The construction of $\langle\psi\rangle_f$ is by induction on the structure of ψ. For the label predicate let

$$\langle\lambda(x) \in B\rangle_f \ := \begin{cases} \top & \text{if } f(x) \in B, \\ \bot & \text{else.} \end{cases}$$

[1] This classical result is due to Straubing [Str85] who used different terminology, however. A proof in the terminology of this thesis is given in [Kuf13, Theorem 3.36], where also the necessary definitions can be found. As we do not use semidirect products beyond this motivation, we do not even define them. All other algebraic terms are introduced in Section 2.3.

For all other atomic formulae ψ let $\langle\psi\rangle_f := \psi$. Boolean connectives are straightforward: $\langle\varphi \vee \psi\rangle_f := \langle\varphi\rangle_f \vee \langle\psi\rangle_f$ and $\langle\varphi \wedge \psi\rangle_f := \langle\varphi\rangle_f \wedge \langle\psi\rangle_f$ as well as $\langle\neg\psi\rangle_f := \neg\langle\psi\rangle_f$.

For existential quantification let

$$\langle\exists x\ \psi\rangle_f \ := \ \exists x \bigvee_{\substack{p,q\in(A\cup\{\Box\})^N,\\ a\in A}} \big(\lambda(x) = (p,a,q)\big) \wedge \langle\psi\rangle_{f[x/[paq]]},$$

where $f[x/[paq]]$ is the function, that maps x to $[paq]$ and each $y \neq x$ to $f(y)$. The generalized label $\lambda(x) = (p,a,q)$ requires that the positions surrounding x be labeled by the factor paq with x being on the central a. It is the Σ_1^2-formula

$$\big(\lambda(x) = (p,a,q)\big) \ := \ \big(\lambda(x) = \overleftarrow{pa}\big) \wedge \big(\lambda(x) = \overrightarrow{aq}\big) \tag{8.1}$$

of quantifier depth at most N, where

$$\lambda(x) = \overleftarrow{pa} \ := \ \begin{cases} (\lambda(x) = a) & \text{if } p = \varepsilon,\\ (\lambda(x) = a) \wedge \min(x) & \text{if } p \in \{\Box\}^+,\\ (\lambda(x) = a) \wedge \exists y\ \big(\mathrm{suc}(y,x) \wedge \lambda(y) = \overleftarrow{p}\big) & \text{otherwise,}\end{cases}$$

$$\lambda(x) = \overrightarrow{aq} \ := \ \begin{cases} (\lambda(x) = a) & \text{if } q = \varepsilon,\\ (\lambda(x) = a) \wedge \max(x) & \text{if } q \in \{\Box\}^+,\\ (\lambda(x) = a) \wedge \exists y\ \big(\mathrm{suc}(x,y) \wedge \lambda(y) = \overrightarrow{q}\big) & \text{otherwise.}\end{cases}$$

The variable y occurring in these formulae is the variable from axiom (3) of factor-stability. The formula $\lambda(x) = \overleftarrow{pa}$ is true if the factor pa ends at position x, and $\lambda(x) = \overrightarrow{aq}$ is true if aq starts at position x. Note that the minimum and maximum predicates can be inserted because $\min(x) \leqslant_{\mathcal{F}} \mathrm{suc}(x,x)$ and $\max(x) \leqslant_{\mathcal{F}} \mathrm{suc}(x,x)$ in every factor-stable fragment $\mathcal{F}$. Universal quantification follows by duality.

Consider the modular quantification $\exists^{r \bmod q} x\ \psi$. We let $\langle\exists^{r \bmod 1} x\ \psi\rangle_f := \top$ and we may assume $q \neq 1$. Similar to ordinary existential quantification let

$$\langle\exists^{r \bmod q} x\ \psi\rangle_f \ := \ \exists^{r \bmod q} x \bigvee_{\substack{p,q\in(A\cup\{\Box\})^N,\\ a\in A}} \big(\lambda(x) = (p,a,q)\big) \wedge \langle\psi\rangle_{f[x/[paq]]}.$$

The formula $\lambda(x) = (p,a,q)$ in this formula has the same semantics as above, but uses modular quantifiers. Specifically, the quantifier $\exists y$ in the formulae $\lambda(x) = \overleftarrow{pa}$ and $\lambda(x) = \overrightarrow{aq}$ is replaced by the quantifier $\exists^{1 \bmod q} y$. Note that this makes no semantic difference because there is at most one successor position of x. $\Box$

This immediately yields the following strong version for one half of the sliding window property as a corollary. Note that the sentence defining $\sigma_N^{-1}(L)$ in the assumption of the corollary is allowed use the successor predicate.

Corollary 8.4
Let $\mathcal{F} \subseteq$ FO+MOD be a factor-stable fragment. If $K \subseteq A^*_{(N)}$ is definable in $\mathcal{F}$ for some integer N, then $\sigma_N^{-1}(K)$ is also definable in $\mathcal{F}$. $\Box$

We still have to show the more involved implication of the sliding window property from condition (1) to (2), which eliminates the successor predicate in the sliding window formula. For this we have to provide for every sentence $\varphi \in \mathcal{F}$ a sliding window formula $\psi \in \mathcal{F}$ that does not use suc, min, or max. We provide a fairly general solution to this, but we have to relax the sliding window property. The sliding window formula will not be in the same fragment $\mathcal{F}$ as the original formula, but in a extension thereof, denoted by $\widehat{\mathcal{F}}$. This extension, defined in Definition 8.5, can require that certain distance constraints between variables are met.

An *extended formula* is a formula that, apart from the usual logic constructs of ordinary formulae given in Section 3.1, is additionally allowed to use

- generalized existential quantifiers $\exists x \in \mathcal{C} \colon \varphi$,
- generalized universal quantifiers $\forall x \in \mathcal{C} \colon \varphi$, and
- generalized modular counting quantifiers $\exists^{r \bmod q} x \in \mathcal{C} \colon \varphi$,

where $r, q \in \mathbb{Z}$ are integers, φ is an extended formula, and $\mathcal{C}$ is a *quantifier constraint* (or just *constraint* for short) of the form $\{y\}$ for some first-order variable $y \neq x$ or of the form $\{i \mid |y - i| > d(y),\ y \in V\}$ for some finite subset V of first-order variables with $x \notin V$ and some $d \colon V \to \mathbb{N}$. The values $d(x)$ are called *distance parameters*.

The set of free variables of a constraint $\mathcal{C} = \{y\}$ is $\mathrm{FV}(\mathcal{C}) = \{y\}$ and that of a constraint $\mathcal{C} = \{i \mid |y - i| > d(y),\ y \in V\}$ is $\mathrm{FV}(\mathcal{C}) = V$. The set of free variables of extended formulae is defined as usual with $\mathrm{FV}(\mathrm{Q}x \in \mathcal{C} \colon \varphi) = \mathrm{FV}(\mathrm{Q}x\ \varphi) \cup \mathrm{FV}(\mathcal{C})$.

Up until now these quantifiers are just syntactical symbols, albeit with rather suggestive names. Their semantics is as follows. Consider a structure $u \in \mathcal{U}_{\mathbb{V}}$. A position i of u *meets* the constraint $\{i \mid |y - i| > d(y),\ y \in V\}$ if $|y(u) - i| > d(y)$ for all $y \in V$, and it meets the constraint $\{y\}$ if $i = y(u)$. Let $u \models \exists x \in \mathcal{C} \colon \varphi$ if there exists a position i meeting $\mathcal{C}$ such that $u[x/i] \models \varphi$. Let $u \models \forall x \in \mathcal{C} \colon \varphi$ if $u \not\models \exists x \in \mathcal{C} \colon \neg\varphi$, i.e., all positions meeting $\mathcal{C}$ make φ true. Let $u \models \exists^{r \bmod q} x \in \mathcal{C} \colon \varphi$ if the cardinality of the set of positions i such that i meets $\mathcal{C}$ and $u[x/i] \models \varphi$ is congruent r modulo q.

An *extended context* is an extended formula with a unique occurrence of a 0-ary atomic formula $\circ$ as a placeholder. For a set $\mathcal{F}$ of extended formulae, let $\varphi \leqslant_{\mathcal{F}} \psi$ if $\mu(\psi) \in \mathcal{F}$ implies $\mu(\varphi) \in \mathcal{F}$ for all extended contexts μ.

Definition 8.5 (distance-stability)
A set $\mathcal{F}$ of extended formulae is distance-stable *if for all extended formulae φ, all $x \in \mathbb{V}_1$, all $q, r, s \in \mathbb{Z}$, and all quantifier constraints $\mathcal{C}$ with* $\mathrm{FV}(\mathcal{C}) \subseteq \mathrm{FV}(\varphi) \setminus \{x\}$:

1. $\big(\exists x \in \mathcal{C} \colon \varphi\big) \leqslant_{\mathcal{F}} \big(\exists x\ \varphi\big)$.
2. $\big(\forall x \in \mathcal{C} \colon \varphi\big) \leqslant_{\mathcal{F}} \big(\forall x\ \varphi\big)$.
3. $\big(\exists^{s \bmod q} x \in \mathcal{C} \colon \varphi\big) \leqslant_{\mathcal{F}} \big(\exists^{r \bmod q} x\ \varphi\big)$.
4. $\big(x \equiv s \pmod{q}\big) \leqslant_{\mathcal{F}} \big(x \equiv r \pmod{q}\big)$.

The distance-stable extension $\widehat{\mathcal{F}}$ *of a fragment $\mathcal{F}$ is the smallest distance-stable set of extended formulae that contains $\mathcal{F}$.*

Any ordinary quantifier can thus be replaced by its generalized variant within the distance-stable extension, and the remainder parameters can be altered.

Theorem 8.6 (Existence of sliding window formulae)
For each sentence $\varphi \in$ FO+MOD there exist a natural number $N \leqslant 2^{\mathrm{qd}(\varphi)} + 1$ and an extended sentence ψ that uses neither successor, minimum, nor maximum predicates such that $\mathcal{L}_A(\varphi) = \sigma_N^{-1}(\mathcal{L}_{A(N)}(\psi))$ for every alphabet A, and $\psi \leqslant_{\widehat{\mathcal{F}}} \varphi$ for every fragment $\mathcal{F}$.

Moreover, ψ can be chosen such that the absolute value of every distance parameter occurring in a quantifier constraint in ψ is less than N.

The proof of this theorem will soon be given. In one aspect this theorem is much stronger than what the sliding window approach requires: For each formula there is *one* extended sliding window formula that respects all fragments. However, it is also weaker in that the sliding window formula is an extended formula. The following weak solution to the sliding window approach is an immediate corollary.

Corollary 8.7
Let $\mathcal{F} \subseteq$ FO+MOD be a fragment. If $L \subseteq A^*$ is $\mathcal{F}$-definable, then there exist an integer $N \geqslant 0$ and a language $K \subseteq A^*_{(N)}$ definable in $\widehat{\mathcal{F}}$ by a sentence that does not use successor, minimum, or maximum predicates such that $L = \sigma_N^{-1}(K)$. □

Although this is not completely satisfactory, it gives a far-reaching partial solution to the sliding window property. In concrete cases it is possible to eliminate quantifier constraints within the considered fragment, thus showing $\mathcal{F}$-definability of K in the previous corollary. Such an elimination procedure is exemplified in Proposition 8.17.

In the following Theorem 8.6 is proven. It employs a rather involved construction, so let us start with an intuition.

Motivation. Recall that the main goal is to get rid of the successor predicate by drawing the necessary information from the sliding window alphabet. Let us illustrate the main difficulty for this and how we overcome it. Consider the formula $\mathrm{suc}(x, y)$ with free variables x and y. Without successor predicates and without further assumptions on x and y it seems impossible to ascertain truth of $\mathrm{suc}(x, y)$ — at least without introducing new quantifiers and variables, therewith "destroying" the structure of the formula. The solution is to ensure in advance (that is, at quantification time) that x and y are not adjacent, thus reducing $\mathrm{suc}(x, y)$ to $\perp$. Let us elaborate on this, supposing $\mathrm{suc}(x, y)$ is part of a formula $\exists x\ (\mathrm{suc}(x, y) \vee \varphi)$, where y is already bound. This quantification is handled by a case distinction: Either there exists a satisfying x adjacent to y, or there exists a satisfying x with $|x - y| > 1$. In the first case, x is bound to the same position as y, and a syntactic bookkeeping method keeps track of the relative distance between x and y. For the second case let us develop this example further, suppose that φ is the formula $\exists z\ (\mathrm{suc}(x, z) \wedge \mathrm{suc}(z, y))$, requiring that there be precisely one position in between x and y, or in other words $y = x + 2$. As all we ensured is $|x - y| > 1$, it is difficult to ascertain truth of φ without using successor. The solution is the same as before: Handle all positions x with $|x - y| \leqslant 2$ syntactically and ensure $|x - y| > 2$. If φ had been more complicated, an even larger distance between x and y would have had to be ensured. Since only finitely many neighboring positions can be handled syntactically, the main question is then how many such positions suffice. The answer depends on φ and, broadly speaking, is the longest coherent chain of successive positions that φ can specify. A bound on this number can be estimated by a purely syntactic analysis of the structure of φ, which motivates the following definition.

Definition 8.8 (Successor distance $\mathrm{sd}_\varphi(x, y)$)
The successor distance $\mathrm{sd}_\varphi(x, y)$ *within a formula $\varphi \in$ FO+MOD between distinct first-order variables x and y is defined by the following inductive scheme:*

- *let* $\mathrm{sd}_\varphi(x, y) = 1$ *if* $\varphi \in \{\mathrm{suc}(x, y), \mathrm{suc}(y, x)\}$*;*
- *let* $\mathrm{sd}_\varphi(x, y) = 0$ *if φ is any other atomic formula.*

– *let* $\mathrm{sd}_{\neg\varphi}(x,y) = \mathrm{sd}_{\varphi}(x,y)$*;*
– *let* $\mathrm{sd}_{\varphi\vee\psi}(x,y) = \mathrm{sd}_{\varphi\wedge\psi}(x,y) = \max\{\mathrm{sd}_{\varphi}(x,y),\ \mathrm{sd}_{\psi}(x,y)\}$*;*
– *let* $\mathrm{sd}_{\mathrm{Q}z\,\varphi}(x,y) = \max\{\mathrm{sd}_{\varphi}(x,y),\ \mathrm{sd}_{\varphi}(x,z) + \mathrm{sd}_{\varphi}(z,y)\}$ *if* $z \notin \{x,y\}$*, and* *let* $\mathrm{sd}_{\mathrm{Q}z\,\varphi}(x,y) = 0$ *if otherwise* $z \in \{x,y\}$*, where* $\mathrm{Q} \in \{\exists,\ \forall,\ \exists^{r \bmod q}\}$.

The successor distance is symmetric in its arguments: $\mathrm{sd}_{\varphi}(x,y) = \mathrm{sd}_{\varphi}(y,x)$. For the quantifiers this definition is motivated by the idea that by placing z in between x and y, a chain of successive positions from x to y can be obtained from chains connecting x with z and z with y. However, if the quantifier rebinds x, say, then the original value of x is obliterated and φ does not connect the original value of x with y by successor predicates.

It is not hard to verify that $\mathrm{sd}_{\varphi}(x,y)$ is at most exponential in the quantifier depth of φ; that is, $\mathrm{sd}_{\varphi}(x,y) \leqslant 2^{\mathrm{qd}(\varphi)}$ for all $x \neq y$ and all $\varphi \in \mathrm{FO{+}M{\scriptstyle OD}}$. Note that if φ uses only two-variables, then $\mathrm{sd}_{\varphi}(x,y) \leqslant 1$ for all $x \neq y$. The following example shows that the exponential upper bound is tight in general.

Example 8.9
There are formulae $\varphi_n(x,y)$ of quantifier depth n such that $\mathrm{sd}_{\varphi_n}(x,y) = 2^n$. An example of such a formula is $\varphi_n(x,y)$ defined by $\varphi_0(x,y) := \mathrm{suc}(x,y)$ and for $n \geqslant 1$ by

$$\varphi_n(x,y) \ := \ \exists z\ \big(\varphi_{n-1}(x,z) \wedge \varphi_{n-1}(z,y)\big),$$

where y and z switch roles in $\varphi_{n-1}(x,z)$, and x and z change roles in $\varphi_{n-1}(z,y)$. This formula is true if y is 2^n positions greater than x, that is, $u \models \varphi_n(x,y)$ if and only if $x(u) + 2^n = y(u)$. ◇

The syntactic bookkeeping method mentioned above makes use of an *offset function* $\delta\colon V \to \mathbb{Z}$, where $V \subseteq \mathbb{V}_1$ contains the free variables of the formula. In the evaluation of the formula, the actual value of a variable x is notionally offset by $\delta(x)$ positions, with negative values denoting a shift to the left: For $V = \{x_1,\ldots,x_\ell\}$ and a structure $u \in \mathcal{U}_V$ let $u+\delta$ be the structure $u[x_1/x_1(u)+\delta(x_1),\ldots,x_\ell/x_\ell(u)+\delta(x_\ell)]$, offsetting the interpretation of each free variable x_i by $\delta(x_i)$ positions.

For a formula φ in $\mathrm{FO{+}M{\scriptstyle OD}}$ and an offset function δ we are going to give an extended formula $\sigma_\delta(\varphi)$ such that $u \models \sigma_\delta(\varphi)$ if and only if $u+\delta \models \varphi$ for admissible words. The following defines admissible words for our construction, *i.e.*, those words for which we can guarantee correctness.

Definition 8.10 (Admissibility $\mathcal{W}_{\varphi,\delta}$)
Let φ be a formula, and let $\delta\colon V \to \mathbb{Z}$ be an offset function with $\mathrm{FV}(\varphi) \subseteq V$. A structure $u \in \mathcal{U}_V$ is admissible *for $\sigma_\delta(\varphi)$ if $u(x)+\delta(x) \in \{1,\ldots,|u|\}$ for all $x \in V$ and for all $x,y \in V$ either $x(u) = y(u)$, or $|x(u)-y(u)| > |\delta(x)| + |\delta(y)| + \mathrm{sd}_{\varphi}(x,y)$. Let $\mathcal{W}_{\varphi,\delta}$ be the set of admissible structures for $\sigma_\delta(\varphi)$.*

Note that if u is admissible for $\sigma_\delta(\varphi)$, then all variables in V have a well-defined interpretation in $u+\delta$ by definition and consequently $u+\delta \in \mathcal{U}_V$.

At this point of the construction, we do not want to bother with the radius of the sliding window. The means to do this are generalized label predicates, allowing to sample surrounding positions. These generalized label predicates can be eliminated afterwards using a large enough sliding window. Apart from generalized quantifiers, the formula $\sigma_\delta(\varphi)$ will thus also use *offset label predicates* of the form $\lambda(x+d) \in B$ for $d \in \mathbb{Z}$ and $B \subseteq \Lambda \cup \{\square\}$. The semantics of offset label predicates is canonical: If the interpretation of x on the structure is position i, then $\lambda(x+d) \in B$ is true if and only if $i+d$ is labeled by a letter in B. Here the convention is that all non-position integers are labeled by $\square$. As usual, we write $\lambda(x+d) = b$ whenever $B = \{b\}$. Let us stress once more that offset label predicates are just a means to write things down smoothly and will be eliminated later on.

The construction. The following gives the construction of $\sigma_\delta(\varphi)$ for $\varphi \in \mathrm{FO}{+}\textsc{Mod}$ and $\delta \colon V \to \mathbb{Z}$ with $\mathrm{FV}(\varphi) \subseteq V$. The focus is on conciseness and we provide only a rough intuition and only for more involved constructs like quantifiers. The main properties of the construction, such as semantic correctness, are given later on.

The whole point of the construction is to rid the predicates for successor, minimum, and maximum. This is done as follows:

$$\begin{aligned}
\sigma_\delta(\mathrm{suc}(x,y)) &:= \begin{cases} x = y & \text{if } \delta(x)+1 = \delta(y), \\ \bot & \text{otherwise,} \end{cases} \\
\sigma_\delta(\min(x)) &:= \big(\lambda(x+\delta(x)-1) = \square\big), \\
\sigma_\delta(\max(x)) &:= \big(\lambda(x+\delta(x)+1) = \square\big).
\end{aligned}$$

For $\varphi \in \{\top, \bot, \mathrm{empty}\}$ we let $\sigma_\delta(\varphi) := \varphi$. The formula $\sigma_\delta(\lambda(x) \in B)$ for the label predicate is $\lambda(x+\delta(x)) \in B$. For the remaining atomic formulae let

$$\begin{aligned}
\sigma_\delta(x = y) &:= \begin{cases} x = y & \text{if } \delta(x) = \delta(y), \\ \bot & \text{otherwise,} \end{cases} \\
\sigma_\delta(x < y) &:= \begin{cases} x \leqslant y & \text{if } \delta(x) < \delta(y), \\ x < y & \text{otherwise,} \end{cases} \\
\sigma_\delta(x \leqslant y) &:= \begin{cases} x \leqslant y & \text{if } \delta(x) \leqslant \delta(y), \\ x < y & \text{otherwise,} \end{cases}
\end{aligned}$$

and $\sigma_\delta(x \equiv r \pmod q) := (x \equiv r - \delta(x) \pmod q)$. For the Boolean connectives let

$$\begin{aligned}
\sigma_\delta(\varphi_1 \vee \varphi_2) &:= \sigma_\delta(\varphi_1) \vee \sigma_\delta(\varphi_2), \\
\sigma_\delta(\varphi_1 \wedge \varphi_2) &:= \sigma_\delta(\varphi_1) \wedge \sigma_\delta(\varphi_2), \\
\sigma_\delta(\neg\varphi) &:= \neg\sigma_\delta(\varphi).
\end{aligned}$$

We come to the existential first-order quantifier. While the construction so far was fairly straightforward, quantifiers are more involved and we start with an intuition. The idea is there exists an x that makes φ true if and only if there exists such an x in the "vicinity" of another variable y (*i.e.*, at most $|\delta(y)| + \mathrm{sd}_\varphi(x,y)$ positions away) or there exists such an x which is "far away" from all other variables.

For technical reasons that will be become apparent later, we handle $x \notin \mathrm{FV}(\varphi)$ separately and in this case we define $\sigma_\delta(\exists x\ \varphi) := \exists x \colon \sigma_{\delta[x/0]}(\varphi)$. Suppose now $x \in \mathrm{FV}(\varphi)$. For an offset function $\delta \colon V \to \mathbb{Z}$ and $d \in \mathbb{Z}$ let $\delta[x/d]$ be the function $V \cup \{x\} \to \mathbb{Z}$ defined by $x \mapsto d$ and $y \mapsto \delta(y)$ for $y \neq x$. For a first-order variable $y \in V$ define the interval $I(y) = \{d \in \mathbb{Z} \mid |d| \leqslant |\delta(y)| + \mathrm{sd}_\varphi(x, y)\}$. Depending on φ and x, define a quantifier constraint $\mathcal{C}$ by

$$\mathcal{C} = \{i \mid |y - i| > |\delta(y)| + \mathrm{sd}_\varphi(x, y),\ y \in \mathrm{FV}(\varphi) \setminus \{x\}\}. \tag{8.2}$$

Let $\sigma_\delta(\exists x\ \varphi) := \varphi' \vee \varphi''$, where

$$\varphi' \ := \bigvee_{\substack{y \in \mathrm{FV}(\exists x\ \varphi) \\ d \in I(y)}} (\lambda(y + d) \neq \square) \wedge \exists x \in \{y\} \colon \sigma_{\delta[x/d]}(\varphi),$$

$$\varphi'' \ := \ \exists x \in \mathcal{C} \colon \sigma_{\delta[x/0]}(\varphi).$$

The formula $\lambda(y + d) \neq \square$ stands for $\lambda(y + d) \in \Lambda$ and is used for clarity because it expresses more clearly the requirement that y be labeled by anything but $\square$. The quantifier constraint $\mathcal{C}$ is just a formalization of the requirement that the newly quantified position be such that the arising structure is admissible.

For universal quantifiers let $\sigma_\delta(\forall x\ \varphi) := \forall x \colon \sigma_{\delta[x/0]}(\varphi)$ if $x \notin \mathrm{FV}(\varphi)$. If $x \in \mathrm{FV}(\varphi)$, then let $\mathcal{C}$ be given by (8.2) and let $\sigma_\delta(\forall x\ \varphi) := \bar{\varphi}' \wedge \bar{\varphi}''$, where

$$\bar{\varphi}' \ := \bigwedge_{\substack{y \in \mathrm{FV}(\forall x\ \varphi) \\ d \in I(y)}} (\lambda(y + d) = \square) \vee \forall x \in \{y\} \colon \sigma_{\delta[x/d]}(\varphi),$$

$$\bar{\varphi}'' \ := \ \forall x \in \mathcal{C} \colon \sigma_{\delta[x/0]}(\varphi).$$

The construction for the modular counting quantifier is even more involved and we again start with an intuition. We count how many positions in the "vicinity" of another free variable make φ true. Using this knowledge, we can then infer how many positions "far away" have to make φ true in order to make $\exists^{r \bmod q} x\ \varphi$ true. Observe that for a modulus $|q| = 1$ all formulae of the form $\exists^{r \bmod q} x\ \varphi$ are always true, and we may set $\sigma_\delta(\exists^{r \bmod q} x\ \varphi) := \top$. We may therefore assume $|q| \neq 1$ in the following.

If $x \notin \mathrm{FV}(\varphi)$, then let $\sigma_\delta(\exists^{r \bmod q} x\ \varphi) := \exists^{r \bmod q} x \colon \sigma_{\delta[x/0]}(\varphi)$. Suppose $x \in \mathrm{FV}(\varphi)$. For an arbitrary set I and an integer r we denote the set of all subsets of I with precisely r elements by $\binom{I}{r}$. For given d and M let $[d \in M]$ denote the Iverson bracket; it is the integer 1 in case $d \in M$ and 0 otherwise. Let $\mathfrak{R}$ be the set of all functions R from $\mathrm{FV}(\varphi) \setminus \{x\}$ into $\mathbb{N}$ with $R(y) \in \{0, \ldots, |I(y)|\}$ for all $y \in \mathrm{FV}(\varphi) \setminus \{x\}$. Note that $\mathfrak{R}$ is finite. With this terminology let

$$\sigma_\delta(\exists^{r \bmod q} x\ \varphi) \ := \bigvee_{R \in \mathfrak{R}} (\hat{\varphi}' \wedge \hat{\varphi}''),$$

where the formulae $\hat{\varphi}'$ and $\hat{\varphi}''$ depend on R and are defined as follows:

$$\hat{\varphi}' \ := \bigwedge_{y \in \mathrm{FV}(\varphi) \setminus \{x\}} \ \bigvee_{M \in \binom{I(y)}{R(y)}} \ \bigwedge_{d \in I(y)} \exists^{[d \in M] \bmod q} x \in \{y\} \colon \big((\lambda(y + d) \neq \square) \wedge \sigma_{\delta[x/d]}(\varphi)\big),$$

$$\hat{\varphi}'' \ := \ \exists^{s \bmod q} x \in \mathcal{C} \colon \sigma_{\delta[x/0]}(\varphi),$$

where $\mathcal{C}$ is given by (8.2) and where $s = r - \sum_y R(y)$ with the sum ranging over all $y \in \mathrm{FV}(\varphi) \setminus \{x\}$. This means we guess that there be $R(y)$ positions near the variable y. This guess is verified in $\hat{\varphi}'$. The verification is done by testing for existence of any distribution M of $R(y)$ offsets around the variable y. Note that there are only finitely many such distributions as we only consider offsets in $I(y)$. The last conjunction over $d \in I(y)$ ensures that the offsets in M make φ true (because then $[d \in M] = 1$), whereas all other offsets make φ false (because then $[d \in M] = 0$). Note that the relative position of y with offset d is unique and the quantifier in $\hat{\varphi}'$ thus counts exactly (as $q \notin \{-1, +1\}$). In $\hat{\varphi}''$ the positions that are far away from all other free variables are accounted for. This concludes the definition of the construction.

Properties of the construction. To formulate semantic correctness, we extend the sliding window in the presence of free variables. For $u = (a_1, J_1) \cdots (a_n, J_n) \in \mathcal{U}_V$ let $\sigma_N(u) = (b_1, J_1) \cdots (b_n, J_n)$, where $\sigma_N(a_1 \cdots a_n) = b_1 \cdots b_n$ with $b_i \in \Lambda_{(N)}$. This means that the sliding window is applied only to the underlying word, the interpretation of the free variables is left untouched.

The following lemma establishes semantic correctness of the sliding window formula $\sigma_\delta(\varphi)$ for admissible words.

Lemma 8.11
For every $\varphi \in \mathrm{FO{+}M{\small OD}}$ and every offset function $\delta \colon V \to \mathbb{Z}$ with $\mathrm{FV}(\varphi) \subseteq V$ the equivalence $u + \delta \models \varphi$ if and only if $u \models \sigma_\delta(\varphi)$ holds for all $u \in \mathcal{W}_{\varphi,\delta}$.

Proof. The proof is by structural induction. For the remainder of this proof let $u \in \mathcal{W}_{\varphi,\delta}$. Recall that $x(u + \delta) = x(u) + \delta(x)$.

Atomic formulae. The claim is trivially true for $\varphi \in \{\top, \perp, \text{empty}\}$. For the label predicate $u + \delta \models \lambda(x) \in B$ is equivalent to $u \models \lambda(x + \delta(x)) \in B$ by the semantics of the offset label predicate.

Consider a formula of the form $\mathrm{suc}(x, y)$. If $\delta(x)+1 = \delta(y)$, then $x(u+\delta)+1 = y(u+\delta)$ if and only if $x(u) = y(u)$. Suppose $x(u)+\delta(x)+1 = y(u)+\delta(y)$. The triangle inequality yields $|x(u) - y(u)| \leqslant |\delta(x)| + |\delta(y)| + 1$. Since u is admissible $x(u) = y(u)$, which yields $\delta(x) + 1 = \delta(y)$.

Consider a formula of the form $\max(x)$, and consider a position i of u and an offset d. The integer $i + d$ is a position of u if and only if $(u + \delta)[x/i] \models \big(\lambda(x + d) \in \Lambda\big)$; that is, $i + d$ is labeled by anything but $\square$. Hence $x(u + \delta)$ is the maximal position of u if and only if $x(u + \delta) + 1$ is no longer a position of u. The latter is the case if and only if $u + \delta \models \big(\lambda(x + 1) = \square\big)$, which in turn is equivalent to $u \models \big(\lambda(x + \delta(x) + 1) = \square\big)$. For minimum predicates a symmetric reasoning applies.

Consider a formula of the form $x = y$. If $\delta(x) = \delta(y)$, then clearly $x(u + \delta) = y(u + \delta)$ if and only if $x(u) = y(u)$. Assuming $x(u+\delta) = y(u+\delta)$ yields $x(u)+\delta(x) = y(u)+\delta(y)$ and thus $|x(u) - y(u)| \leqslant |\delta(x)| + |\delta(y)|$. Admissibility of u yields $x(u) = y(u)$, which implies $\delta(x) = \delta(y)$.

Of the order predicates we only consider $x < y$; the argument for $x \leqslant y$ is similar. First suppose $\delta(x) < \delta(y)$. Clearly $x(u + \delta) < y(u + \delta)$ if $x(u) \leqslant y(u)$. Conversely, if $x(u) > y(u)$, admissibility of u implies $x(u+\delta) \geqslant x(u) - |\delta(x)| > y(u) + |\delta(y)| \geqslant y(u+\delta)$. Suppose $\delta(x) \geqslant \delta(y)$. If $x(u) \geqslant y(u)$, then $x(u+\delta) = x(u)+\delta(x) \geqslant y(u)+\delta(y) = y(u+\delta)$. Therefore, $x(u + \delta) < y(u + \delta)$ only if $x(u) < y(u)$. Conversely, if $x(u) < y(u)$, then admissibility yields $x(u + \delta) \leqslant x(u) + |\delta(x)| < y(u) - |\delta(y)| \leqslant y(u + \delta)$.

Consider a predicate of the form $x \equiv r \pmod{q}$. We have $x(u + \delta) \equiv r \pmod{q}$ if and only if $x(u) + \delta(x) \equiv r \pmod{q}$ if and only if $x(u) \equiv s \pmod{q}$, where $s = r - \delta(x)$.

Boolean combinations. It is straightforward that the formulae for disjunction $\varphi_1 \vee \varphi_2$, conjunction $\varphi_1 \wedge \varphi_2$, and for negation $\neg\varphi$ inherit correctness from their subformulae. Notice that $\mathcal{W}_{\varphi_1 \vee \varphi_2,\delta} = \mathcal{W}_{\varphi_1 \wedge \varphi_2,\delta} \subseteq \mathcal{W}_{\varphi_1,\delta} \cap \mathcal{W}_{\varphi_2,\delta}$.

Quantifiers. Consider a formula $\mathrm{Q}x\ \varphi$ for some $\mathrm{Q} \in \{\exists,\ \forall,\ \exists^{r \bmod q}\}$, and let $\delta\colon V \to \mathbb{Z}$ and $u \in \mathcal{W}_{\mathrm{Q}x\,\varphi,\delta}$. We have to show $u + \delta \models \mathrm{Q}x\ \varphi$ if and only if $u \models \sigma_\delta(\mathrm{Q}x\ \varphi)$.

We first handle the special case $x \notin \mathrm{FV}(\varphi)$ uniformly for all quantifiers. For empty u the claim is trivially true. Suppose u is non-empty. If $V \neq \emptyset$, then let $i = y(u)$ for an arbitrary but fixed $y \in V$. If $V = \emptyset$, let i be an arbitrary but fixed position of u. In any case $u[x/i] \in \mathcal{W}_{\varphi,\delta[x/0]}$. Let $\langle \mathrm{Q}x\ \varphi\rangle := (\varphi \wedge \mathrm{Q}x\ \top) \vee (\neg\varphi \wedge \mathrm{Q}x\ \bot)$. If $x \notin \mathrm{FV}(\varphi)$, then $\mathrm{Q}x\ \varphi$ and $\langle \mathrm{Q}x\ \varphi\rangle$ are equivalent.

This yields $u + \delta \models \mathrm{Q}x\ \varphi$ if and only if $u[x/i] + \delta[x/0] \models \langle \mathrm{Q}x\ \varphi\rangle$ if and only if $u[x/i] \models \langle \mathrm{Q}x\colon \sigma_{\delta[x/0]}(\varphi)\rangle$ if and only if $u \models \mathrm{Q}x\colon \sigma_{\delta[x/0]}(\varphi)$. The first equivalence holds because $x \notin \mathrm{FV}(\varphi)$, the second equivalence uses $u[x/i] \in \mathcal{W}_{\varphi,\delta[x/0]}$, and the third equivalence holds because $x \notin \mathrm{FV}(\sigma_{\delta[x/0]}(\varphi))$.

Existential quantification. Consider a formula of the form $\exists x\ \varphi$ with $x \in \mathrm{FV}(\varphi)$. The idea of the construction is to exploit the fact that there exists x that makes φ true if and only if either there is such an x in the "vicinity" of another variable y (*i.e.*, at most $|\delta(y)| + \mathrm{sd}_\varphi(x, y)$ positions away) or there is x which is "far away" from all other variables. The first case is captured by the formula φ', whereas φ'' captures the second.

The following renders this idea formally. First consider φ'. Suppose $y \in \mathrm{FV}(\exists x\ \varphi)$ and let $d \in I(y) = \{d \in \mathbb{Z} \mid |d| \leqslant |\delta(y)| + \mathrm{sd}_\varphi(x, y)\}$ be such that $y(u) + d$ is a position of u. Note that $y(u) + d$ is a position of u if and only if $u \models (\lambda(y + d) \neq \square)$.

We shall show $u[x/y(u)] \in \mathcal{W}_{\varphi,\delta[x/d]}$ in a short while. Assuming this for the time being yields $u[x/y(u)] \models \sigma_{\delta[x/d]}(\varphi)$ if and only if $u[x/y(u)] + \delta[x/d] = (u + \delta)[x/y(u) + d] \models \varphi$. This shows $u \models \varphi'$ if and only if there exists a position i of u such that $(u + \delta)[x/i] \models \varphi$ and $|i - y(u)| \leqslant |\delta(y)| + \mathrm{sd}_\varphi(x, y)$ for some $y \in \mathrm{FV}(\exists x\ \varphi)$.

We now show $u[x/y(u)] \in \mathcal{W}_{\varphi,\delta[x/d]}$. For brevity let $\delta' = \delta[x/d]$ and $u' = u[x/y(u)]$. Consider variables $z_1, z_2 \in V \cup \{x\}$. We have to show that either $z_1(u') = z_2(u')$ or $|z_1(u') - z_2(u')| > |\delta'(z_1)| + |\delta'(z_2)| + \mathrm{sd}_\varphi(z_1, z_2)$.

Suppose $z_1(u') \neq z_2(u')$. If x is neither of the z_j, this is immediate by admissibility of u. Note that $\delta'(z_j) = \delta(z_j)$, that $z_j(u') = z_j(u)$, and that $\mathrm{sd}_{\exists x\varphi}(z_1, z_2) \geqslant \mathrm{sd}_\varphi(z_1, z_2)$. By symmetry we may suppose $x = z_2$ for the remaining cases. For $z_1 = y$ we have $x(u') = y(u')$ and there is nothing to show. Hence suppose $z_1 \neq y$. Admissibility yields $|z_1(u') - y(u')| > |\delta'(z_1)| + |\delta'(y)| + \mathrm{sd}_{\exists x\varphi}(y, z_1)$. And since $x(u') = y(u')$, we get

$$\begin{aligned} |z_1(u') - x(u')| = |z_1(u') - y(u')| &> |\delta'(z_1)| + |\delta'(y)| + \mathrm{sd}_\varphi(y, x) + \mathrm{sd}_\varphi(x, z_1) \\ &\geqslant |\delta'(z_1)| + |\delta'(x)| + \mathrm{sd}_\varphi(x, z_1). \end{aligned}$$

This uses $\mathrm{sd}_{\exists x\varphi}(y, z_1) \geqslant \mathrm{sd}_\varphi(y, x) + \mathrm{sd}_\varphi(x, z_1)$ for the first inequality and for the second $|\delta'(x)| \leqslant |\delta'(y)| + \mathrm{sd}_\varphi(y, x)$ is used.

We now come to the second case φ''. We only give a sketch of the main arguments because the formal reasoning is similar to the case φ'. We claim that $u \models \varphi''$ if and only if there exists a position i of u such that $(u + \delta)[x/i] \models \varphi$ with

$$|i - y(u)| > |\delta(y)| + \mathrm{sd}_\varphi(x, y) \quad \text{for all } y \in \mathrm{FV}(\exists x\ \varphi). \tag{8.3}$$

Note that a position i satisfies (8.3) if and only if i meets $\mathcal{C}$ as defined by (8.2); and for such a position, the structure $u[x/i]$ is admissible for $\sigma_{\delta[x/0]}(\varphi)$. We have $u \models \varphi''$ if and only if there exists a position i meeting $\mathcal{C}$ with $u[x/i] \models \sigma_{\delta[x/0]}(\varphi)$. By induction this is equivalent to $u[x/i] + \delta[x/0] = (u + \delta)[x/i] \models \varphi$ for some position satisfying (8.3).

It is clear that this dichotomy of near positions in φ' and far positions in φ'' is exhaustive. Therefore, there exists i with $(u + \delta)[x/i] \models \varphi$ if and only if $u \models \sigma_\delta(\exists x\ \varphi)$.

Universal quantification. This follows from existential quantification using duality.

Modular quantification. A very rough intuition was already given at the construction. The formal details of the proof (such as proving that induction applies) are similar to that of existential quantification and we argue on a higher level.

Consider a formula of the form $\exists^{r \bmod q} x\ \varphi$ with $|q| \neq 1$ and $x \in \mathrm{FV}(\varphi)$. The construction first counts how many positions $R(y)$ near any variable $y \in \mathrm{FV}(\exists^{r \bmod q} x\ \varphi)$ there are that make φ. The "near" positions are those with offset in $I(y)$ around y, and in particular, there are at most $|I(y)|$ such positions. To do this counting of near positions, we first guess for each variable y their number $R(y)$. This guess is subsequently verified by $\hat{\varphi}'$ as explained in a short while.

Supposing for the time being that the $R(y)$ correctly specify the number of nearby satisfying positions, let us have a look at $\hat{\varphi}''$. Summing up over all other variables, there are a total of $\sum_y R(y)$ satisfying positions near any other variable. This means that there have to be $r - \sum_y R(y)$ satisfying positions far away from all other variables, which is precisely what by $\hat{\varphi}''$ ensures. As for existential quantification, $\mathcal{C}$ ensures that the quantification happens only over positions far away.

Let us finally have a look at how $\hat{\varphi}'$ verifies $R(y)$. For fixed y, all possible distributions M of $R(y)$ offsets over the interval $I(y)$ are considered. For a fixed M it is then ensured that of all offsets d in $I(y)$ precisely those in M yield positions around y that make φ true. Testing that the position offset from y by d is such that it makes φ true is done as follows. By the quantifier constraint $x \in \{y\}$, the quantifier actually ranges over one position only, namely $y(u)$. Together with a fixed offset d, the formula $\sigma_{\delta[x/d]}(\varphi)$ in this context thus speaks about the unique position $y(u) + d$. Therefore, there is either one or no position in this context that makes $\sigma_{\delta[x/d]}(\varphi)$ true, and the quantifier counts exactly in this case as $|q| \neq 1$. This shows that the formula

$$\exists^{1 \bmod q} x \in \{y\} \colon \big((\lambda(y + d) \neq \square) \wedge \sigma_{\delta[x/d]}(\varphi)\big)$$

is true if and only if $y(u)+d$ is a position of u that makes φ true. Changing the remainder from 1 to 0 to get the quantifier $\exists^{0 \bmod q} x \in \{y\}$ negates the meaning; *i.e.*, either $y(u)+d$ is not a position of u, or it does not make φ true. The Iverson bracket $[d \in M]$ in the formula $\sigma_\delta(\exists^{r \bmod q} x\ \varphi)$ is a means to write both conditions in a uniform way. $\square$

After semantic correctness is established, we now turn to syntactic properties. We start by showing how offset label predicates in $\sigma_\delta(\varphi)$ can be eliminated using the sliding window alphabet of a sufficiently large diameter. The following definition formalizes this elimination of offset label predicates, instantiating $\sigma_\delta(\varphi)$ for usage with the sliding window of radius N.

Definition 8.12 (Sliding window formula $\boldsymbol{\sigma_{N,\delta}(\varphi)}$)
Let $\varphi \in \mathrm{FO{+}M\small{OD}}$ and $\delta \colon V \to \mathbb{Z}$ with $\mathrm{FV}(\varphi) \subseteq V$, and let N be such that $\lambda(x+d) \in B$ being a subformula of $\sigma_\delta(\varphi)$ implies $N \geqslant |d|$. Let $\sigma_{N,\delta}(\varphi)$ be obtained from $\sigma_\delta(\varphi)$ by replacing all occurrences of offset labels $\lambda(x + d) \in B$ by the ordinary label $\lambda(x) \in B_d$, where $B_d = \{[ubv] \in \Lambda \mid u \in (\Lambda \cup \{\square\})^{N+d},\ b \in B,\ v \in (\Lambda \cup \{\square\})^{N-d}\}$.

This yields $u \models \sigma_\delta(\varphi)$ if and only if $\sigma_N(u) \models \sigma_{N,\delta}(\varphi)$ for all $u \in \mathcal{U}_V$. For all N that do not satisfy the premises, not all offset labels can be rewritten in terms of the sliding window alphabet and $\sigma_{N,\delta}(\varphi)$ is undefined. Note that the resulting formula $\sigma_{N,\delta}(\varphi)$ still uses generalized quantifiers and therefore is still an *extended* formula.

The next definition gives radii N that are sufficient for $\sigma_{N,\delta}(\varphi)$ to be defined. More precisely, it defines $N_\delta(\varphi)$ for which the subsequent lemma shows that every radius that is at least $N_\delta(\varphi) + 1$ is a valid sliding window radius for our construction.

Definition 8.13 (Sliding window radius $N(\varphi)$ and $N_\delta(\varphi)$)
The radius $N_\delta(\varphi)$ is defined inductively as follows for $\varphi, \psi \in \mathrm{FO{+}M\small{OD}}$, $\delta\colon V \to \mathbb{Z}$ with $\mathrm{FV}(\varphi) \subseteq V$, $x \in \mathbb{V}_1$, and $\mathrm{Q} \in \{\exists,\ \forall,\ \exists^{r \bmod q} \mid r, q \in \mathbb{Z}\}$:

- $N_\delta(\varphi) = \max(\{0\} \cup \{|\delta(x)| \mid x \in \mathrm{FV}(\varphi)\})$ *if φ is an atomic formula,*
- $N_\delta(\neg\varphi) = N_\delta(\varphi)$,
- $N_\delta(\varphi \vee \psi) = N_\delta(\varphi \wedge \psi) = \max\{N_\delta(\varphi),\ N_\delta(\psi)\}$,
- $N_\delta(\mathrm{Q}x\ \varphi) = \max\big(\big\{N_{\delta[x/0]}(\varphi)\big\} \cup \big\{N_{\delta[x/|\delta(y)|+\mathrm{sd}_\varphi(x,y)]}(\varphi) \ \big|\ y \in \mathrm{FV}(\mathrm{Q}x\ \varphi)\big\}\big)$.

For a sentence $\varphi \in \mathrm{FO{+}M\small{OD}}$ let the sliding window radius $N(\varphi)$ *be* $N_\emptyset(\varphi)$.

Consider some arbitrary but fixed formula φ. The radius function $N_\delta(\varphi)$ is monotonic in its offset parameter δ; that is, $\delta \leqslant \delta'$ implies $N_\delta(\varphi) \leqslant N_{\delta'}(\varphi)$. Here, offset functions are partially ordered by a component-wise comparison $\delta, \delta'\colon V \to \mathbb{Z}$; *i.e.*, $\delta \leqslant \delta'$ whenever $|\delta(y)| \leqslant |\delta'(y)|$ for all $y \in V$. This monotony property can be verified by a straightforward, monotonous structural induction. Moreover, an easy induction on the structure of φ yields that $N_\delta(\varphi) \geqslant |\delta(x)|$ for all $x \in \mathrm{FV}(\varphi)$. The next lemma establishes that $N_\delta(\varphi) + 1$ is a valid choice as a sliding window for our construction.

Lemma 8.14
If $\lambda(x + d) \in B$ is a subformula of $\sigma_\delta(\varphi)$, then $N_\delta(\varphi) + 1 \geqslant |d|$.

Proof. The proof is by induction on the structure of φ. Suppose φ is atomic. If φ is neither a label, minimum, or maximum predicate, then there is nothing to show. Suppose φ is a label of the form $\lambda(x) \in B$. In this case we have $N_\delta(\varphi) \geqslant |\delta(x)|$. This shows the claim as $\lambda(x + \delta(x)) \in B$ is the only subformula of $\sigma_\delta(\varphi)$. Suppose φ is of the form $\min(x)$ or $\max(x)$. By definition $N_\delta(\varphi) + 1 \geqslant |\delta(x)| + 1$ which shows the claim because $\lambda(x + \delta(x) \pm 1) = \square$ is the only subformula of $\sigma_\delta(\varphi)$.

Suppose φ is a disjunction of the form $\varphi_1 \vee \varphi_2$. By definition $\sigma_\delta(\varphi) := \sigma_\delta(\varphi_1) \vee \sigma_\delta(\varphi_2)$ and any label subformula $\lambda(x + d) \in B$ is a subformula of of $\sigma_\delta(\varphi_1)$ or of $\sigma_\delta(\varphi_2)$, say φ_1 for simplicity. Induction yields $N_\delta(\varphi) \geqslant N_\delta(\varphi_1) \geqslant |d| - 1$. Conjunction and negation follow similarly.

Suppose ψ is of the form $\exists x\ \varphi$, $\forall x\ \varphi$, or $\exists^{r \bmod q} x\ \varphi$. For $y \in \mathrm{FV}(\varphi) \setminus \{x\}$ let in the following $\delta_y = \delta[x/|\delta(y)| + \mathrm{sd}_\varphi(x, y)]$. First consider any label $\lambda(y + d) \neq \square$ or $\lambda(y + d) = \square$ introduced for handling the quantifier, *i.e.*, not introduced by the inductive procedure. We have $N_\delta(\psi) \geqslant |\delta(y)| + \mathrm{sd}_\varphi(x, y) \geqslant |d|$. The first inequality is implied by $N_\delta(\psi) \geqslant N_{\delta_y}(\varphi)$ due to $x \in \mathrm{FV}(\varphi)$; the second inequality is by construction of $\sigma_\delta(\psi)$.

Consider the labels introduced by the inductive procedure; *i.e.*, consider a subformula $\lambda(z + i) \in B$ introduced by a formula of the form $\sigma_{\delta[x/d]}(\varphi)$. For $d = 0$ we have that $N_\delta(\psi) \geqslant N_{\delta[x/0]}(\varphi) \geqslant |i|-1$; the first inequality is by definition of $N_\delta(\psi)$ and the second is by induction. Note that this also deals with the special case $x \notin \mathrm{FV}(\varphi)$. For $d \neq 0$ there exists $y \in \mathrm{FV}(\psi)$ such that $\delta[x/d] \leqslant \delta_y$. Monotony yields $N_{\delta[x/d]}(\varphi) \leqslant N_{\delta_y}(\varphi) \leqslant N_\delta(\psi)$, where the second inequality holds by definition of $N_\delta(\psi)$. This yields $N_\delta(\psi) \geqslant |i| - 1$ by induction, thus concluding the proof. $\square$

The additive term $+1$ in the preceding lemma is only necessary if φ uses a minimum or a maximum predicate. Indeed, analyzing the proof, we see that if φ contains neither minimum nor maximum predicates and $\lambda(x+d) \in B$ is a subformula of $\sigma_\delta(\varphi)$, then $N_\delta(\varphi) \geqslant |d|$.

Let us turn to explicit upper bounds on the sliding window radius $N(\varphi)$. In general, we obtain an exponential bound, whereas two-variable formulae admit a linear bound.

Proposition 8.15
For every sentence $\varphi \in \mathrm{FO{+}M{\small OD}}$ *with* $\mathrm{qd}(\varphi) \geqslant 1$ *the following assertions hold:*

1. $N(\varphi) \leqslant 2^{\mathrm{qd}(\varphi)-1} - 1$.
2. $N(\varphi) \leqslant \mathrm{qd}(\varphi) - 1$ *if* φ *uses only two-variables.*

Proof. Let $\sup(X) = \max(X \cup \{0\})$ for $X \subseteq \mathbb{N}$ be the *supremum* of X in the natural numbers. We have $\sup(X) = \max(X)$ for non-empty X. The only difference is that $\sup(\emptyset) = 0$, whereas $\max(\emptyset)$ is undefined.

(1): More generally $N_\delta(\varphi) + 1 \leqslant \sup\{|\delta(z)| \mid z \in \mathrm{FV}(\varphi)\} + 2^{\mathrm{qd}(\varphi)}$ for all $\delta\colon V \to \mathbb{Z}$ and all $\varphi \in \mathrm{FO{+}M{\small OD}}$ with $\mathrm{FV}(\varphi) \subseteq V$. The proof is by induction on the structure of φ. For atomic formulae this is clear, and Boolean combinations are straightforward.

Quantifiers are all handled in the same way. We consider a quantification $\mathrm{Q}x\ \varphi$ for some $\mathrm{Q} \in \{\exists,\ \forall,\ \exists^{r \bmod q}\}$. Let $\delta_y = \delta[x/|\delta(y)| + \mathrm{sd}_\varphi(x,y)]$. For $N_{\delta_y}(\varphi)$ we have

$$\begin{aligned} N_{\delta_y}(\varphi) + 1 &\leqslant \sup\{|\delta_y(z)| \mid z \in \mathrm{FV}(\varphi)\} + 2^{\mathrm{qd}(\varphi)} \\ &\leqslant \sup\{|\delta(z)| + \mathrm{sd}_\varphi(x,y) \mid z \in \mathrm{FV}(\varphi) \setminus \{x\}\} + 2^{\mathrm{qd}(\varphi)} \\ &\leqslant \sup\{|\delta(z)| \mid z \in \mathrm{FV}(\mathrm{Q}x\ \varphi)\} + 2^{\mathrm{qd}(\mathrm{Q}x\ \varphi)}. \end{aligned}$$

The first inequality is by induction. The second holds since $|\delta_y(z)| \leqslant |\delta(z)| + \mathrm{sd}_\varphi(x,y)$; also note that δ_y is independent of the value $\delta(x)$. The third uses $\mathrm{sd}_\varphi(x,y) \leqslant 2^{\mathrm{qd}(\varphi)}$ and $\mathrm{qd}(\mathrm{Q}x\ \varphi) = \mathrm{qd}(\varphi) + 1$. This bound is independent of the variable y. Moreover, by induction $N_{\delta[x/0]}(\varphi) + 1$ is at most $\sup\{|\delta[x/0](z)| \mid z \in \mathrm{FV}(\varphi)\} + 2^{\mathrm{qd}(\varphi)}$, which in turn is bounded by $\sup\{|\delta(z)| \mid z \in \mathrm{FV}(\mathrm{Q}x\ \varphi)\} + 2^{\mathrm{qd}(\mathrm{Q}x\ \varphi)}$. Both bounds together yield

$$\begin{aligned} N_\delta(\mathrm{Q}x\ \varphi) + 1 &= \max\bigl(\{N_{\delta[x/0]}(\varphi)\} \cup \{N_{\delta_y}(\varphi) \mid y \in \mathrm{FV}(\mathrm{Q}x\ \varphi)\}\bigr) + 1 \\ &\leqslant \sup\{|\delta(z)| \mid z \in \mathrm{FV}(\mathrm{Q}x\ \varphi)\} + 2^{\mathrm{qd}(\mathrm{Q}x\ \varphi)}. \end{aligned}$$

Let us now come back to the claim of the lemma and consider a sentence φ in $\mathrm{FO{+}M{\small OD}}$. Without restriction we may assume $\varphi := \mathrm{Q}x\ \psi$. Let $\{x \mapsto 0\}$ be the offset function with domain $\{x\}$ that maps x to 0. Using the above claim, we conclude $N(\mathrm{Q}x\ \psi) = N_{\{x\mapsto 0\}}(\psi) \leqslant 2^{\mathrm{qd}(\psi)} - 1 = 2^{\mathrm{qd}(\mathrm{Q}x\ \psi)-1} - 1$ as desired.

(2): This is similar to (1). We claim $N_\delta(\varphi) \leqslant \sup\{|\delta(z)| \mid z \in \mathrm{FV}(\varphi)\} + \mathrm{qd}(\varphi)$ for all $\delta\colon V \to \mathbb{Z}$ and all two-variable $\varphi \in \mathrm{FO{+}M{\small OD}}$ with $\mathrm{FV}(\varphi) \subseteq V$. We prove this by induction on the structure of the formula. Atomic formulae and Boolean combinations are straightforward. Consider a quantification $\mathrm{Q}x\ \varphi$ for some quantifier $\mathrm{Q} \in \{\exists,\ \forall,\ \exists^{r \bmod q}\}$. For $N_{\delta_y}(\varphi)$ as above we have

$$\begin{aligned} N_{\delta_y}(\varphi) &\leqslant \sup\{|\delta_y(z)| \mid z \in \mathrm{FV}(\varphi)\} + \mathrm{qd}(\varphi) \\ &\leqslant \sup\{|\delta(z)| + \mathrm{sd}_\varphi(x,y) \mid z \in \mathrm{FV}(\varphi) \setminus \{x\}\} + \mathrm{qd}(\varphi) \\ &\leqslant \sup\{|\delta(z)| \mid z \in \mathrm{FV}(\mathrm{Q}x\ \varphi)\} + \mathrm{qd}(\mathrm{Q}x\ \varphi). \end{aligned}$$

Moreover, by induction $N_{\delta[x/0]}(\varphi) \leqslant \sup\{|\delta(z)| \mid z \in \mathrm{FV}(\mathrm{Q}x\ \varphi)\} + \mathrm{qd}(\mathrm{Q}x\ \varphi)$. Together we thus have $N_\delta(\mathrm{Q}x\ \varphi) \leqslant \sup\{|\delta(z)| \mid z \in \mathrm{FV}(\mathrm{Q}x\ \varphi)\} + \mathrm{qd}(\mathrm{Q}x\ \varphi)$.

Coming back to the claim of the lemma, consider a two-variable formula φ in FO+MOD without free variables. Without restriction we may assume $\varphi := \mathrm{Q}x\ \psi$. The above claim yields $N(\mathrm{Q}x\ \psi) = N_{\{x\mapsto 0\}}(\psi) \leqslant \mathrm{qd}(\psi) = \mathrm{qd}(\mathrm{Q}x\ \psi) - 1$ as desired. $\square$

The following lemma records the most important syntactic properties of the sliding window formulae $\sigma_{N,\delta}(\varphi)$.

Lemma 8.16
For $\varphi \in$ FO+MOD, $\delta\colon V \to \mathbb{Z}$ with $\mathrm{FV}(\varphi) \subseteq V$, and $N \geqslant N_\delta(\varphi) + 1$ the following hold:

1. *$\sigma_{N,\delta}(\varphi) \leqslant_{\widehat{\mathcal{F}}} \varphi$ for all order-stable fragments $\mathcal{F}$ with $(x = y) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$ for all x, y.*
2. *$\sigma_{N,\delta}(\varphi)$ uses neither the successor, minimum, nor maximum predicate.*
3. *If $d \in \mathbb{N}$ is a distance parameter occurring in $\sigma_{N,\delta}(\varphi)$, then $d < N$.*

Proof. (1): This can be seen by a straightforward induction on the structure of φ using distance-stability, order-stability, and the axioms of fragments. Note that the conjunctions and disjunctions of $\leqslant_{\widehat{\mathcal{F}}}$-smaller formulae are themselves $\leqslant_{\widehat{\mathcal{F}}}$-smaller.

(2): This is obvious, as neither successor, minimum, nor maximum predicates are used at any point in the right hand side in the definition of $\sigma_\delta(\varphi)$ and passing to $\sigma_{N,\delta}(\varphi)$ does not introduce them either.

(3): The proof of Lemma 8.14 with minimum and maximum predicates disregarded implicitly shows $d \leqslant N_\delta(\varphi)$. Note that the construction of $\sigma_\delta(\varphi)$ ensures that every distance parameter occurs as an offset parameter of some generalized label predicate. $\square$

Proof of Theorem 8.6. Let $\varphi \in$ FO+MOD be a sentence, let $\delta\colon \emptyset \to \mathbb{Z}$ be the trivial offset function, and let $N = N_\delta(\varphi) + 1$ for which $N \leqslant 2^{\mathrm{qd}(\varphi)-1}$ by Proposition 8.15. The formula $\sigma_{N,\delta}(\varphi)$ in Definition 8.12 satisfies almost all of the properties proclaimed by Theorem 8.6. It uses neither successor, minimum, nor maximum predicates and satisfies the bound on the distance parameters by Lemma 8.16. Lemma 8.11 shows semantic correctness: $u \in \mathcal{L}_A(\varphi)$ if and only if $\sigma_N(u) \in \mathcal{L}_{A(N)}(\sigma_{N,\delta}(\varphi))$ for $u \in A^*$. Note that vacuously $u \in \mathcal{W}_{\varphi,\delta}$ and $u = u + \delta$, where for simplicity the word u is identified with the structure $(u[1], \emptyset) \cdots (u[|u|], \emptyset)) \in \mathcal{U}_\emptyset$.

Lemma 8.16 property (1) shows $\sigma_{N,\delta}(\varphi) \leqslant_{\widehat{\mathcal{F}}} \varphi$, albeit only if $\mathcal{F}$ is order-stable and satisfies $(x = y) \leqslant_{\mathcal{F}} \mathrm{suc}(x, y)$. It remains to free ourselves of these requirements for $\mathcal{F}$. This is possible since the extended quantifiers implicitly encompass knowledge about equality of variables. The reason for including the additional axioms nonetheless is that they allow to define the formulae $\sigma_{\delta'}(\mathrm{suc}(x, y))$ as well as $\sigma_{\delta'}(x < y)$ and $\sigma_{\delta'}(x \leqslant y)$ more clearly.

Analyzing the construction of $\sigma_\delta(\varphi)$, we see that an inspection of the syntactic structure allows to ascertain in any context μ whether two free variables x and y specify equal positions or not. Specifically, consider the last extended quantifier binding either x or y; *i.e.*, let $\mathrm{Q}z \in \mathcal{C}$ be the last quantifier on the path in μ from the root to the placeholder $\circ$ such that $z \in \{x, y\}$. For simplicity assume $z = x$. In this case the variables x and y specify the same position if and only if $\mathcal{C} = \{y\}$. For all other quantifier constraints, x and y are different positions.

Coming back to the construction, suppose $\sigma_{\delta'}(\mathrm{suc}(x, y))$ occurs in context μ in the formula $\sigma_\delta(\varphi)$. The equality $x = y$ used if $\delta(x) + 1 = \delta(y)$ can be replaced by $\top$ in case $\mathcal{C} = \{y\}$ and by $\perp$ otherwise. Likewise if $\sigma_{\delta'}(x < y)$ occurs in context μ, the non-strict order predicate $x \leqslant y$ for $\delta(x) < \delta(y)$ can be replaced by $\top$ if $\mathcal{C} = \{y\}$ and by $x < y$ otherwise. A similar argument applies to eliminate the usage of the strict order predicate in occurrences of $\sigma_{\delta'}(x \leqslant y)$. □

We conclude this section by example conditions that suffice to eliminate the generalized quantifiers introduced by Theorem 8.6.

A fragment $\mathcal{F}$ is *expansion-stable* if it is remainder-stable and if for all $x, y \in \mathbb{V}_1$ and all formulae φ with $y \in \mathrm{FV}(\varphi) \setminus \{x\}$ there exists $z \in \mathbb{V}_1 \setminus \{x, y\}$ such that the following properties hold:

1. $(\exists x\ (x = y)) \leqslant_{\mathcal{F}} (\exists x\ \varphi)$.
   $(\exists x\ (x < y) \vee (y < x)) \leqslant_{\mathcal{F}} (\exists x\ \varphi)$.
   $(\exists x \exists z\ (x < z \wedge z < y) \vee (y < z \wedge z < x)) \leqslant_{\mathcal{F}} (\exists x\ \varphi)$.
2. $(\forall x\ \neg(x = y)) \leqslant_{\mathcal{F}} (\forall x\ \varphi)$.
   $(\forall x\ (y \leqslant x) \wedge (x \leqslant y)) \leqslant_{\mathcal{F}} (\forall x\ \varphi)$.
   $(\forall x \forall z\ (z \leqslant x \vee y \leqslant z) \wedge (z \leqslant y \vee x \leqslant z)) \leqslant_{\mathcal{F}} (\forall x\ \varphi)$.
3. $(\exists^{r \bmod q} x\ (x = y)) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$.
   $(\exists^{r \bmod q} x\ (x < y) \vee (y < x)) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$.
   $(\exists^{r \bmod q} x\ \exists z\ (x < z \wedge z < y) \vee (y < z \wedge z < x)) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$.

Examples of expansion-stable fragments include the first-order quantifier alternation hierarchy and also the r-variable fragment of the alternation hierarchy for $r \geqslant 3$. Note that the two-variable alternation hierarchy is not expansion-stable.

Proposition 8.17
For every sentence $\varphi \in \mathrm{FO{+}M{\small OD}}$ there exists $N \leqslant 2^{\mathrm{qd}(\varphi)}+1$ and a sentence ψ that uses neither successor, minimum, nor maximum predicates such that $\mathcal{L}_A(\varphi) = \sigma_N^{-1}(\mathcal{L}_{A(N)}(\psi))$ and $\psi \leqslant_{\mathcal{F}} \varphi$ for every expansion-stable fragment $\mathcal{F}$. Moreover, $\mathrm{qd}(\psi) \leqslant 2\mathrm{qd}(\varphi) + 1$.

Proof. Let N and ψ' be given by Theorem 8.6 for φ. We eliminate one generalized quantifier in ψ' after the other until no generalized quantifiers are left.

First consider constraints of the form $\mathcal{C} = \{y\}$ and suppose $\mathcal{C}$ occurs in a quantifier with context μ. We distinguish the type of the quantifier using $\mathcal{C}$ and

- for $\psi' := \mu(\exists x \in \mathcal{C}\colon \chi)$ we let $\psi'' := \mu\big(\exists x\ ((x = y) \wedge \chi)\big)$,
- for $\psi' := \mu(\forall x \in \mathcal{C}\colon \chi)$ we let $\psi'' := \mu\big(\forall x\ (\neg(x = y) \vee \chi)\big)$, and
- for $\psi' := \mu(\exists^{r \bmod q} x \in \mathcal{C}\colon \chi)$ we let $\psi'' := \mu\big(\exists^{r \bmod q} x\ ((x = y) \wedge \chi)\big)$.

The resulting extended formula ψ'' is clearly equivalent to ψ' and uses one constraint less. The constraint elimination procedure is repeated recursively on ψ''.

Consider a constraint of the form $\mathcal{C} = \{i \mid |y - i| > d(y),\ y \in V\}$ and suppose $\mathcal{C}$ is used by a quantifier in context μ. We again distinguish the quantifier types. First consider existential quantification, *i.e.*, $\psi' := \mu(\exists x \in \mathcal{C}\colon \chi)$. We need an auxiliary construction to handle this case.

For $y \in V$ and $d \in \mathbb{N}$ we are going to construct a formula $\langle x < y - d\rangle$ with free variables x and y such that $u \models \langle x < y - d\rangle$ if and only if $x(u) < y(u) - d$ for all u; *i.e.*, there are at least d positions strictly after $x(u)$ and before $y(u)$. The construction is by induction on d. For $d = 0$ we can clearly put $\langle x < y - 0\rangle := (x < y)$. For $d \geqslant 1$ we postulate a position z in between x and y. Moreover, between x and z we require

$\lceil (d-1)/2 \rceil$ positions and between z and y we require $\lfloor (d-1)/2 \rfloor$ positions. Not counting z, this sums up to $\lceil (d-1)/2 \rceil + \lfloor (d-1)/2 \rfloor = d - 1$ between x and y. Let therefore

$$\langle x < y - d \rangle := \exists z \left(\langle x < z - \lceil (d-1)/2 \rceil \rangle \wedge \langle z < y - \lfloor (d-1)/2 \rfloor \rangle \right),$$

where $z \notin \{x, y\}$ is the variable postulated by expansion-stability. Let $\langle x > y + d \rangle$ be the dual formula of $\langle x < y - d \rangle$ with *less than* replaced by *greater than* so that $u \models \langle x > y + d \rangle$ if and only if $x(u) > y(u) + d$ for all u.

We define another convenience formula $|y - x| > d$ with free variables x and y by $(|y - x| > d) := \langle x < y - d \rangle \vee \langle x > y + d \rangle$. For this formula we have $u \models (|y - x| > d)$ if and only if $|y(u) - x(u)| > d$ for all u. Using these shortcut formulae let

$$\psi'' := \mu\Big(\exists x \,\big(\bigwedge_{y \in V} |y - x| > d(y)\big) \wedge \chi\Big).$$

By construction it is clear that ψ'' and ψ' are equivalent. Moreover, for every expansion-stable fragment $\mathcal{F}$ we have $\psi'' \leqslant_{\widehat{\mathcal{F}}} \mu(\exists x\ \chi) \leqslant_{\widehat{\mathcal{F}}} \mu(\exists x \in \mathcal{C} \colon \chi)$. The first $\leqslant_{\widehat{\mathcal{F}}}$-relation follows by construction of $|y - x| > d(y)$ using axiom (1) of expansion-stability for $\mathcal{F}$. The second $\leqslant_{\widehat{\mathcal{F}}}$-relation holds because every generalized quantifier in the distance-stable extension originates from an ordinary quantifier of the same type.

For universal quantification $\psi' := \mu(\forall x \in \mathcal{C} \colon \chi)$ a similar reasoning applies: By an analogous construction using universal instead of existential quantifiers, we can define a formula $|y - x| \leqslant d$ with the canonical meaning. With this let

$$\psi'' := \mu\Big(\forall x \,\big(\bigvee_{y \in V} |y - x| \leqslant d(y)\big) \vee \chi\Big).$$

It remains to consider modular quantifiers $\psi' := \mu(\exists^{r \bmod q} x \in \mathcal{C} \colon \chi)$. Using the formulae $|y - x| > d(y)$ from the case for existential quantifiers we let

$$\psi'' := \mu\Big(\exists^{r \bmod q} x \,\big(\bigwedge_{y \in V} |y - x| > d(y)\big) \wedge \chi\Big).$$

In any case, the formula ψ'' is equivalent to ψ' and satisfies $\psi'' \leqslant_{\widehat{\mathcal{F}}} \psi'$ for all expansion-stable fragments $\mathcal{F}$. Moreover, ψ'' has one generalized quantifier less than ψ'. Iterating the procedure thus terminates after finitely many steps, resulting in a non-extended formula ψ that is equivalent to ψ' and satisfies $\psi \leqslant_{\widehat{\mathcal{F}}} \psi'$ for all expansion-stable fragments $\mathcal{F}$.

Since we have $\psi' \leqslant_{\widehat{\mathcal{F}}} \varphi$, we get $\psi \leqslant_{\widehat{\mathcal{F}}} \varphi$ for all expansion-stable fragments $\mathcal{F}$. We have to show $\psi \leqslant_{\mathcal{F}} \varphi$. Let μ be a non-extended context such that $\mu(\varphi) \in \mathcal{F}$. Using $\psi \leqslant_{\widehat{\mathcal{F}}} \varphi$ yields that $\mu(\psi)$ is in the distance-stable extension $\widehat{\mathcal{F}}$. Since $\mu(\psi)$ is a non-extended formula, this already implies $\mu(\psi) \in \mathcal{F}$. Note that for remainder-stable fragments, passing to the distance-stable extension does not introduce new non-extended formulae. Consequently any non-extended formula in the remainder-stable extension of $\mathcal{F}$ is already in $\mathcal{F}$. This shows $\psi \leqslant_{\mathcal{F}} \varphi$ as desired.

It remains to show $\mathrm{qd}(\psi) \leqslant 2\mathrm{qd}(\varphi)$. The quantifier depth of ψ increases at most by the maximal quantifier depth of the formulae $\langle x < y - d(y) \rangle$ occurring in the above construction. We thus analyze the quantifier depth of the formulae $\langle x < y - d \rangle$. The quantifier depth $t(d)$ of $\langle x < y - d \rangle$ obeys the recursive inequality $t(1) = 1$ and $t(d) \leqslant 1 + t(\lceil (d-1)/2 \rceil) \leqslant 1 + t(d/2)$ for $d \geqslant 2$. A straightforward induction now shows $t(d) \leqslant 1 + \lfloor \log_2(d) \rfloor$. By Theorem 8.6 we can assume that if d is a distance parameter occurring in a quantifier constraint in ψ', then $d \leqslant 2^{\mathrm{qd}(\varphi)}$. With the aforementioned bound $1 + \lfloor \log_2(d) \rfloor$ on the quantifier depth of $\langle x < y - d \rangle$ this yields the claim. $\square$

This means that Corollary 8.7 can be strengthened for expansion-stable fragments.

Corollary 8.18
Let $\mathcal{F} \subseteq \mathrm{FO{+}M{\small OD}}$ be an expansion-stable fragment. If $L \subseteq A^*$ is $\mathcal{F}$-definable, then there exist an integer $N \geqslant 0$ and a language $K \subseteq A^*_{(N)}$ definable in $\mathcal{F}$ by a sentence that does not use successor, minimum, or maximum predicates such that $L = \sigma_N^{-1}(K)$. □

Combining Corollary 8.18 and Corollary 8.4 immediately yields the sliding window property 8.2 for fragments that are both expansion-stable and factor-stable. The following corollary records this for later reference.

Corollary 8.19
Every fragment $\mathcal{F} \subseteq \mathrm{FO{+}M{\small OD}}$ that is factor-stable and expansion-stable has the sliding window property; i.e., for every $L \subseteq A^*$ the following are equivalent:

1. *L is definable in $\mathcal{F}$.*
2. *There exists an integer $N \geqslant 0$ and a language $K \subseteq A^*_{(N)}$ definable in $\mathcal{F}$ by a sentence that does not use successor, minimum, or maximum predicates with $L = \sigma_N^{-1}(K)$.* □

This yields a purely logical proof of the classical result that the first-order quantifier alternation hierarchy has the sliding window property.

Corollary 8.20
For every relational signature $\mathcal{N}$ containing $\{\mathrm{suc}, \min, \max\}$ and every $r \geqslant 3$ each of the fragments $\Sigma_m[\mathcal{N}]$, $\mathrm{FO}_m[\mathcal{N}]$, $\Sigma^r_m[\mathcal{N}]$, and $\mathrm{FO}^r_m[\mathcal{N}]$ has the sliding window property.

Proof. All these fragments are obviously factor-stable and expansion-stable, and consequently Corollary 8.19 applies. Note that the predicates suc, min, and max must be available in the relational signature to obtain factor-stability. □

As another application let us turn to Straubing's fragments $\mathcal{Q}^r_n[\mathcal{N}]$ defined by (5.1) on page 57. As already mentioned, factor-stability and expansion-stability are incompatible with quantifier depth restrictions and our results do not apply to $\mathcal{Q}^r_n[\mathcal{N}]$. The respective fragments without quantifier depth restrictions given by $\mathcal{Q}^r[\mathcal{N}] = \bigcup_{n \geqslant 0} \mathcal{Q}^r_n[\mathcal{N}]$ on the other hand may well be factor-stable and expansion-stable. Corollary 8.19 lists combinations of the parameters $\mathcal{Q}$, r, and $\mathcal{N}$ for this to be the case, thus obtaining the sliding window property.

Corollary 8.21
Let $\mathcal{N}$ be a signature containing $\{x < y,\ \mathrm{suc}(x, y),\ \min(x),\ \max(x) \mid x, y \in \mathbb{V}_1\}$. For every $r \geqslant 3$ and every set $\mathcal{Q} \subseteq \{\exists,\ \forall,\ \exists^{i \bmod q} \mid i, q \in \mathbb{Z}\}$ of admissible quantifiers that does not consist solely of modular quantifiers, the fragment $\mathcal{Q}^r[\mathcal{N}]$ has the sliding window property whenever it is remainder-stable.

Proof. The fragment $\mathcal{Q}^r[\mathcal{N}]$ is easily seen to be factor-stable. Note that there are at least two variables, quantifier depth is not restricted, and suc, min, max are available in any context.

However, $\mathcal{Q}^r[\mathcal{N}]$ cannot be expansion-stable on syntax level since the equality predicate and the non-strict order predicate may be unavailable. As negations do not matter in $\mathcal{Q}^r[\mathcal{N}]$, those predicates can be expressed in terms of the order predicate, reading

$x \leqslant y$ as a macro for $\neg(y < x)$, and $x = y$ as a macro for $(x \leqslant y) \wedge (y \leqslant x)$. To see the quantifier expansion properties (1) to (3) of expansion-stability note that for any two-variables there is always a third variable which can be quantified over because $r \geqslant 3$. The only difficulty arises from modular quantifiers when no existential quantifiers are available: The axiom $(\exists^{r \bmod q} x\ \exists z\ (x < z \wedge z < y) \vee (y < z \wedge z < x)) \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$ requires that an existential quantifier may be inserted after any modular quantifier. Using negations this is again readily remedied by reading $\exists z\ \psi$ as a macro for $\neg\forall z\ \neg\psi$. Note that by assumption $\exists \notin \mathcal{Q}$ implies $\forall \in \mathcal{Q}$. $\square$

8.2. The Successor Predicate and Neutral Letters

Building on sliding window formulae, introduced in the preceding section, we now study the expressive power of the successor predicate in the presence of a neutral letter. A letter c is *neutral* for the language $L \subseteq A^*$ if $pcq \in L \Leftrightarrow pq \in L$ for all $p, q \in A^*$; *i.e.*, the letter can be inserted and deleted anywhere in any word without changing membership in the language. Let $\mathcal{N} \subseteq \{<, \text{suc}, \text{min}, \text{max}, \text{mod}\}$ be a fixed signature and suppose that $\mathcal{F}[\mathcal{N}]$ is a fragment that admits successor-free sliding window formulae in the sense of Section 8.1. The goal of this section is to show that then all $\mathcal{F}[\mathcal{N}]$-definable languages with a neutral letter are already $\mathcal{F}[<]$-definable. This result can be seen as a proof for the Crane Beach conjecture for $\mathcal{F}$ over the signature $[<, \text{suc}, \text{min}, \text{max}, \text{mod}]$.

More generally, the Crane Beach conjecture for a fragment $\mathcal{F}$ over any set of numerical predicates $\mathcal{N}$ states that in the presence of a neutral letter all available numerical predicates reduce to the order predicate. In other words, $\mathcal{F}[<]$ can define every language in $\mathcal{F}[\mathcal{N}]$ with a neutral letter. It was baptized after the location where a flawed proof for the conjecture for first-order logic over arbitrary numerical predicates was written. That this proof was erroneous is not by accident, considering that the conjecture was later refuted for first-order logic over addition and multiplication [BILST05]. Since then, proving the conjecture for special cases has received quite some attention. It has applications, for instance, in separating circuit complexity classes, see *e.g.* [BILST05; LTT06; KS12a].

In the course of proving our result we need to pad words with neutral letters. Let $\text{pad}_v \colon A^* \to A^*$ be the *padding* with $v \in A^*$ given by $\text{pad}_v(\varepsilon) = v$ and $\text{pad}_v(ua) = \text{pad}_v(u)av$ if $a \in A$. In other words, $u = a_1 a_2 \cdots a_\ell$ is mapped to $va_1va_2v \cdots a_\ell v$, *e.g.*, $\text{pad}_{cc}(abba) = ccaccbccbccacc$. Note that pad_v is not a homomorphism for non-empty v.

Definition 8.22 (pad-stability)
A fragment $\mathcal{F}$ of FO+MOD *is* pad-stable *if it is order-stable, suc-stable, mod-stable, and if* $\big(\text{len} \equiv r \pmod q\big) \leqslant_{\mathcal{F}} \big(x \equiv r \pmod q\big)$ *for all $x \in \mathbb{V}_1$ and all $r, q \in \mathbb{Z}$.*

Here, the modular length-predicate $\text{len} \equiv r \pmod q$ is a new 0-ary predicate that is true if the structure has length r (modulo q). It can be seen as a macro for $\exists^{r \bmod q} x\ \top$.

The following is the main theorem of this section.

Theorem 8.23
Let $\mathcal{F}$ be a pad-stable fragment of FO+MOD*. Let $\varphi \in \mathcal{F}$ and suppose there is a sliding window formula in $\mathcal{F}[=, <, \leqslant, \text{mod}]$ for φ. If $L(\varphi)$ has a neutral letter, then $L(\varphi)$ is definable in $\mathcal{F}[=, <, \leqslant]$.*

In particular, $L(\varphi)$ is $\mathcal{F}[<]$-definable under the assumptions of this theorem whenever equality and inequality predicates are available in any context — as they are usually considered to be. The key step towards the proof is the following lemma. It states that the sliding window over a padded word can be simulated within pad-stable fragments.

Lemma 8.24
Let $\varphi \in \mathrm{FO{+}MOD}$ be a sentence, and let $N \geqslant 1$. For all $v \in \Lambda^+$ with $|v| \geqslant 2N$ there exists a sentence $\mathrm{pad}_v^{-1}(\sigma_N^{-1}(\varphi))$ with $\mathrm{pad}_v^{-1}(\sigma_N^{-1}(\varphi)) \leqslant_{\mathcal{F}} \varphi$ for all pad-stable fragments $\mathcal{F}$ such that $u \models \mathrm{pad}_v^{-1}(\sigma_N^{-1}(\varphi))$ if and only if $\sigma_N(\mathrm{pad}_v(u)) \models \varphi$ for all $u \in \Lambda^*$.

Moreover, if v is such that $x \equiv r \pmod q$ being a subformula of φ implies that q divides $|v|+1$, then $\mathrm{pad}_v^{-1}(\sigma_N^{-1}(\varphi))$ does not use the modular predicate.

Proof. On our way, we also have to handle free variables. We thus start by reformulating the statement so as to encompass free variables.

Claim *Let $\varphi \in \mathrm{FO{+}MOD}$, let $V, X, Y \subseteq \mathbb{V}_1$, let $\delta \colon V \to [-N;\, |v|-N]$, and suppose $\mathrm{FV}(\varphi) = \{x_1, \ldots, x_\ell\} \subseteq V$. There is a formula $\langle\varphi\rangle_{\delta,X,Y}$ with $\mathrm{FV}(\langle\varphi\rangle_{\delta,X,Y}) \subseteq V \setminus (X \cup Y)$ such that $\langle\varphi\rangle_{\delta,X,Y} \leqslant_{\mathcal{F}} \varphi$ for all pad-stable fragments $\mathcal{F}$ and*

$$u, p_1, \ldots, p_\ell \models \langle\varphi\rangle_{\delta,X,Y}(x_1, \ldots, x_\ell) \;\Leftrightarrow\; \sigma_N(\mathrm{pad}_v(u)), p'_1, \ldots, p'_\ell \models \varphi(x_1, \ldots, x_\ell)$$

for all $u \in \Lambda^*$ and all $p_i \in \{1, \ldots, |u|\}$ and all $p'_i \in \{1, \ldots, |\mathrm{pad}_v(u)|\}$, provided that:

1. *$p'_i = (|v|+1)p_i + \delta(x_i)$ for all i with $x_i \notin X \cup Y$,*
2. *$p'_i = \delta(x_i)$ for all i with $x_i \in X$, and*
3. *$p'_i = |\mathrm{pad}_v(u)| + 1 + \delta(x_i)$ for all i with $x_i \in Y$.*

Note that even though the variables in $X \cup Y$ do not actually occur freely in $\langle\varphi\rangle_{\delta,X,Y}$, we consider them nonetheless as free dummy variables. Let us start with an intuition. Every position of the word u has a corresponding position in $\mathrm{pad}_v(u)$. Positions in $\mathrm{pad}_v(u)$ can thus be encoded by a position in u with an offset in the interval $[-N;\, |v|-N]$. This offset is specified by δ, and the encoding property is the first item of the assumptions on the positions. The positions near the border of $\mathrm{pad}_v(u)$ have to be treated differently; these are handled by the sets X and Y, respectively. Such positions in particular need no corresponding position on u and thus variables in $X \cup Y$ are not free in $\langle\varphi\rangle_{\delta,X,Y}$. Now, dragging a sliding windows of diameter $2N+1$ over the word $\mathrm{pad}_v(u)$ does not provide any new information, as in particular any two neighboring positions of u are separated by at least $2N$ positions whose contents are completely determined by v. This means that with the label of the current position, we already know the value of the sliding window at this position.

The lemma follows from this claim by letting $\mathrm{pad}_v^{-1}(\sigma_N^{-1}(\varphi))$ be the sentence $\langle\varphi\rangle_{\emptyset,\emptyset,\emptyset}$, so it truly remains to show the claim. We prove the claim by induction on the structure of the formula. Observe that we may assume X and Y to be disjoint because otherwise the claim is vacuously true.

Let $\langle\varphi\rangle_{\delta,X,Y} := \varphi$ if $\varphi \in \{\top, \bot\}$.

For $B \subseteq \Lambda$ let $B' = \{a \in \Lambda \mid [w] \in B,\ w = (vav)[|va| + \delta(x) - N;\, |va| + \delta(x) + N]\}$. Remember that $(vav)[i;j]$ is the factor of vav induced by all positions in the interval

$[i;j]$. Let $P_d = (\square^N v)[d; d+2N]$ and $Q_d = (v\square^N)[|v\square| + d - N; |v\square| + d + N]$. With this let

$$\langle \lambda(x) \in B\rangle_{\delta,X,Y} := \begin{cases} \lambda(x) \in B' & \text{if } x \notin X \cup Y, \\ \top & \text{if } x \in X \text{ and } [P_{\delta(x)}] \in B, \\ \top & \text{if } x \in Y \text{ and } [Q_{\delta(x)}] \in B, \\ \bot & \text{otherwise.} \end{cases}$$

For the equality predicate and the order predicate $\lesssim \in \{\leqslant, <\}$ the construction in the case $x, y \in V \setminus (X \cup Y)$ is as follows:

$$\langle x = y\rangle_{\delta,X,Y} := \begin{cases} x = y & \text{if } \delta(x) = \delta(y), \\ \bot & \text{otherwise,} \end{cases}$$

$$\langle x \lesssim y\rangle_{\delta,X,Y} := \begin{cases} x \leqslant y & \text{if } \delta(x) \lesssim \delta(y), \\ x < y & \text{otherwise.} \end{cases}$$

For the other case $\{x, y\} \cap (X \cup Y) \neq \emptyset$ let

$$\langle x = y\rangle_{\delta,X,Y} := \begin{cases} \top & \text{if } \delta(x) = \delta(y) \text{ and either } \{x,y\} \subseteq X \text{ or } \{x,y\} \subseteq Y, \\ \bot & \text{otherwise,} \end{cases}$$

$$\langle x \lesssim y\rangle_{\delta,X,Y} := \begin{cases} \top & \text{if } x \in X \text{ and } y \notin X \text{ or if } x \notin Y \text{ and } y \in Y, \\ \top & \text{else if } \delta(x) \lesssim \delta(y) \text{ and either } \{x,y\} \subseteq X \text{ or } \{x,y\} \subseteq Y, \\ \bot & \text{otherwise.} \end{cases}$$

For the successor predicates and its entourage let $\langle \text{empty}\rangle_{\delta,X,Y} := \bot$ and

$$\langle \min(x)\rangle_{\delta,X,Y} := \begin{cases} \top & \text{if } x \in X \text{ and } \delta(x) = 1, \\ \bot & \text{otherwise,} \end{cases}$$

$$\langle \max(x)\rangle_{\delta,X,Y} := \begin{cases} \top & \text{if } x \in Y \text{ and } \delta(x) = -1, \\ \bot & \text{otherwise,} \end{cases}$$

$$\langle \mathrm{suc}(x,y)\rangle_{\delta,X,Y} := \begin{cases} \mathrm{suc}(x,y) & \text{if } x, y \in V \setminus (X \cup Y) \text{ and } \delta(x) - \delta(y) = |v|, \\ \min(y) & \text{if } x \in X \text{ and } \delta(x) - \delta(y) = |v|, \\ \max(x) & \text{if } y \in Y \text{ and } \delta(x) - \delta(y) = |v|, \\ x = y & \text{if } \delta(x) + 1 = \delta(y) \text{ and } x, y \in V \setminus (X \cup Y), \\ \top & \text{if } \delta(x) + 1 = \delta(y) \text{ and if } x, y \in X \text{ or } x, y \in Y, \\ \bot & \text{otherwise.} \end{cases}$$

We come to the modular atomic formula $x \equiv r \pmod{q}$. Suppose v is such that q divides $|v| + 1$. In this case we let

$$\langle x \equiv r \pmod{q}\rangle_{\delta,X,Y} := \begin{cases} \top & \text{if } \delta(x) \equiv r \pmod{q}, \\ \bot & \text{otherwise.} \end{cases}$$

In the case q does not divide $|v| + 1$ suppose first $x \notin X \cup Y$. Let $t = \gcd(|v| + 1, q)$ be the greatest common divisor of $|v| + 1$ and q, let $s = (|v| + 1)/t$, and let $q' = q/t$. Let s^{-1} be a multiplicative inverse of s (modulo q).

If t does not divide $r - \delta(x)$, then we let $\langle x \equiv r \pmod q \rangle_{\delta,X,Y} := \bot$. Otherwise let $r_i = s^{-1}(r - \delta(x) - qi)/t$ and set

$$\langle x \equiv r \pmod q \rangle_{\delta,X,Y} \ := \ \Big(\bigvee_{0 \leqslant i < t} x \equiv r_i \pmod q\Big)$$

The case $x \in X$ is easy enough:

$$\langle x \equiv r \pmod q \rangle_{\delta,X,Y} \ := \ \begin{cases} \top & \text{if } \delta(x) \equiv r \pmod q, \\ \bot & \text{otherwise.} \end{cases}$$

Now for the case $x \in Y$. Let $t = \gcd(|v|, q)$. If t does not divide $r - \delta(x) - 1$, then let $\langle x \equiv r \pmod q \rangle_{\delta,X,Y} := \bot$. Otherwise, if t divides $r - \delta(x) - 1$, set

$$\langle x \equiv r \pmod q \rangle_{\delta,X,Y} \ := \ \Big(\bigvee_{0 \leqslant i < t} \mathrm{len} \equiv r'_i \pmod q\Big),$$

where $r'_i = (|v|/t)^{-1}(r - |v| - \delta(x) - 1 + iq)/t$ with $(|v|/t)^{-1}$ being the multiplicative inverse of $|v|/t$ (modulo q).

For Boolean connectives let

$$\begin{aligned} \langle \varphi \vee \psi \rangle_{\delta,X,Y} \ &:= \ \langle \varphi \rangle_{\delta,X,Y} \vee \langle \psi \rangle_{\delta,X,Y}, \\ \langle \varphi \wedge \psi \rangle_{\delta,X,Y} \ &:= \ \langle \varphi \rangle_{\delta,X,Y} \wedge \langle \psi \rangle_{\delta,X,Y}, \\ \langle \neg\varphi \rangle_{\delta,X,Y} \ &:= \ \neg\langle \varphi \rangle_{\delta,X,Y}. \end{aligned}$$

For existential quantification let

$$\langle \exists x\ \varphi \rangle_{\delta,X,Y} \ := \ \exists x \Big(\bigvee_{-N \leqslant d \leqslant |v| - N} \langle \varphi \rangle_{\delta[x/d],X',Y'}\Big) \vee \bigvee_{1 \leqslant d \leqslant N} \langle \varphi \rangle_{\delta[x/d],X'',Y'} \vee \langle \varphi \rangle_{\delta[x/-d],X',Y''}$$

where $Z' = Z \setminus \{x\}$ and $Z'' = Z \cup \{x\}$ for $Z \in \{X, Y\}$, and $\delta[x/d]$ maps x to d and all other variables to the value under δ. This realizes the case distinction as described in the following. The first big disjunction is for the positions neither on the prefix nor on the suffix of length N. The first term of the second big disjunction is for positions on the prefix of length N, whereas the remaining term handles positions on the suffix of length N. Universal quantification is dual.

Modular quantification is more involved. Let $[N] = \{1, \ldots, N\}$. With this let

$$\langle \exists^{r \bmod q} x\ \varphi \rangle_{\delta,X,Y} \ := \ \bigvee_{J,K \in 2^{[N]}} \varphi_1 \wedge \varphi_2.$$

We give the formal definition of the φ_i in a short while. These formulae depend on J and K and their meaning is as follows. The set J specifies which offsets of the prefix make φ true and K specifies the satisfying offsets of the suffix. The formula φ_1 ensures this meaning of J and K:

$$\begin{aligned} \varphi_1 \ := \ &\bigwedge_{d \in J} \langle \varphi \rangle_{\delta[x/d],X'',Y'} \wedge \bigwedge_{d \in [N] \setminus J} \neg\langle \varphi \rangle_{\delta[x/d],X'',Y'} \wedge \\ &\bigwedge_{d \in K} \langle \varphi \rangle_{\delta[x/-d],X',Y''} \wedge \bigwedge_{d \in [N] \setminus K} \neg\langle \varphi \rangle_{\delta[x/-d],X',Y''} \end{aligned}$$

where Z' and Z'' for $Z \in \{X, Y\}$ are as in the case of existential quantification.

The formula φ_2 has to count the remaining positions in the middle of the word. It has to ensure that (modulo q) there are $r - |J| - |K|$ such positions that make φ true.

This is done by guessing for every offset d how many such positions there are. We have to verify this guess, of course. To formulate this let

$$S = \left\{ s \in \{0, \ldots, q-1\}^{\{-N,\ldots,|v|-N\}} \;\middle|\; \left(|J| + |K| + \textstyle\sum_d s(d)\right) \equiv r \pmod{q} \right\}$$

be the universe of guesses. With this let

$$\varphi_2 \;:=\; \bigvee_{s \in S} \; \bigwedge_{-N \leqslant d \leqslant |v|-N} \; \exists^{s(d) \bmod q} x \; \langle \varphi \rangle_{\delta[x/d],X',Y'}.$$

This concludes the proof of the claim and therefore also of the lemma. □

Let us briefly explain the significance of this lemma. The lemma shows that, not only can we simulate the evaluation of a formula on the padded word, but the padding allows to incorporate the additional information the look-around of the sliding window provides. Applying the lemma to a sliding window formula for φ (which typically eliminates the successor predicate), we thus get a formula that simulates the interpretation of φ on the padded word. On a non-padded word we would still be in trouble as labels with look-around need information about the vicinity. Summarizing, using the sliding window allows to get rid of the successor predicate, while the padding ensures that we do not need to introduce it elsewhere.

Proof of Theorem 8.23. Let c be a neutral letter of $L(\varphi)$. Let $\varphi_N \in \mathcal{F}[=,<,\leqslant,\mathrm{mod}]$ be a sliding window formula of radius N for φ. Choose $v \in \{c\}^+$ such that $|v| \geqslant 2N$ and $x \equiv r \pmod{q}$ being a subformula of φ_N implies that q divides $|v|+1$. Let ψ be the sentence $\mathrm{pad}_v^{-1}(\sigma_N^{-1}(\varphi_N))$ from Lemma 8.24 for the formula φ_N. Note that $\mathcal{F}[=,<,\leqslant,\mathrm{mod}]$ is a pad-stable fragment containing φ_N. The properties of ψ therefore yield $\psi \in \mathcal{F}[=,<,\leqslant]$. For $u \in \Lambda^*$ the following equivalences hold: $u \models \varphi$ if and only if $\mathrm{pad}_v(u) \models \varphi$ if and only if $\sigma_N(\mathrm{pad}_v(u)) \models \varphi_N$ if and only if $u \models \psi$. The first equivalence holds because $v = c^{|v|}$ is neutral for $L(\varphi)$; the second equivalence holds because φ_N is by assumption a sliding window formula for φ; the last equivalence finally is the main property of ψ. □

Algebraic consequences. We now turn to the consequences for the interaction between logic and algebra. For the algebraic terminology refer to Section 2.3. Suppose we are given a monoid M and a fragment $\mathcal{F}$ of FO+MOD that is able to define the languages recognized by M. We are interested what happens if the neutral element of a monoid has some non-empty word as a preimage. Applying the preceding lemma, we show that $\mathcal{F}[=,<,\leqslant]$ suffices to define the recognized languages. Crucial for this is that the fragment must admit sliding window formulae to eliminate the successor predicate.

Lemma 8.25
Let $\mathcal{F}$ be a pad-stable fragment of FO+MOD. *Let $h\colon A^* \to M$ be a homomorphism to a finite monoid with $h(A^+) = M$. Let $P \subseteq M$ and $\varphi \in \mathcal{F}$ with $L(\varphi) \cap A^+ = h^{-1}(P) \cap A^+$. If there is a sliding window formula in $\mathcal{F}[=,<,\leqslant,\mathrm{mod}]$ for φ, then $h^{-1}(P)$ is definable in $\mathcal{F}[=,<,\leqslant,\mathrm{mod}]$.*

Proof. The proof is very similar to that of Theorem 8.23, yet it is slightly more general as not necessarily a single letter is neutral for the language, but only some non-empty word. Because of this, we cannot guarantee the premise of the addendum of Lemma 8.24, which is the reason why the modular predicate cannot be eliminated.

Let $\varphi_N \in \mathcal{F}[=,<,\leqslant,\mathrm{mod}]$ be a sliding window formula of radius N for φ. Let $v \in A^+$ be such that $h(v) = 1$. Every power of v also maps to the neutral element; consequently we may assume without restriction that $|v| \geqslant 2N$. Let $\psi \in \mathcal{F}[=,<,\leqslant,\mathrm{mod}]$ be the sentence $\mathrm{pad}_v^{-1}(\sigma_N^{-1}(\varphi_N))$ from Lemma 8.24 for the formula φ_N. It remains to show that ψ defines $h^{-1}(P)$. For $u \in A^*$ we have $h(u) = h(\mathrm{pad}_v(u)) \in P$ if and only if $\mathrm{pad}_v(u) \models \varphi$ if and only if $\sigma_N(\mathrm{pad}_v(u)) \models \varphi_N$ if and only if $u \models \psi$. The first equivalence holds by the assumption on φ; note that $\mathrm{pad}_v(u)$ is non-empty. This shows $L(\psi) = h^{-1}(P)$. ☐

Suppose we know that all languages recognized by a class of semigroups $\mathbf{V}$ are defined in some fragment $\mathcal{F} \subseteq$ FO+MOD. Can we say something about the subfamily of languages recognized by semigroups that happen to be *monoids* in $\mathbf{V}$? The upcoming lemma answers this questions and shows that such languages are often definable in $\mathcal{F}[=,<,\leqslant,\mathrm{mod}]$ without the successor predicate. Again, the existence of successor-free sliding window formulae is crucial. For a class of semigroups $\mathbf{V}$ let $\mathbf{V_M}$ be the class of monoids in $\mathbf{V}$.

Proposition 8.26
Let $\mathbf{V}$ be a class of semigroups and let $\mathcal{F}$ be a pad-stable fragment of FO+MOD *such that every language recognized in $\mathbf{V}$ is $\mathcal{F}$-definable over non-empty words. If every sentence in $\mathcal{F}$ has a sliding window formula in $\mathcal{F}[=,<,\leqslant,\mathrm{mod}]$, then every language recognized in $\mathbf{V_M}$ is definable in $\mathcal{F}[=,<,\leqslant,\mathrm{mod}]$.*

Proof. Let $L \subseteq A^*$ be recognized by $h\colon A^* \to M$ with $M \in \mathbf{V_M}$. Let $c \notin A$ be a new letter, let $B = A \cup \{c\}$, and let $\hbar\colon B^* \to M$ be the monoid homomorphism given by $c \mapsto 1$ and $a \mapsto h(a)$ for $a \neq c$. Let $P = h(L)$. By assumption there exists $\varphi \in \mathcal{F}$ such that $\mathcal{L}_B(\varphi) \cap B^+ = \hbar^{-1}(P) \cap B^+$. By Lemma 8.25 there exists $\psi \in \mathcal{F}[=,<,\leqslant,\mathrm{mod}]$ such that $\hbar^{-1}(P) = \mathcal{L}_B(\psi)$. Therefore, $L = h^{-1}(P) = \mathcal{L}_A(\psi)$. ☐

9. The Expressive Power of the Minimum and Maximum Predicates

The purpose of this chapter is to analyze what the predicates min and max contribute to the expressive power of a logic fragment. It is intuitive that the minimum and maximum predicates are related to the ability of specifying prefixes and suffixes of models. This chapter substantiates this intuition on a very abstract level. Specifically, it shows that the minimum predicate is superfluous in mod-stable fragments that can specify prefixes and that can ensure a minimum distance to the left border at quantification time. Due to left-right symmetry, we concentrate on the minimum predicate. At the end of the chapter, the dual result for the maximum predicate is made explicit for future reference.

The minimum predicate. The goal of this section is to eliminate the minimum predicate. As an intermediate step, we introduce a new variable x_{min} that is always interpreted by the minimal position, and substitute each atomic formula $\min(x)$ by the equality predicate $x = x_{min}$. Clearly, this substitution yields a formula that is equivalent — provided x_{min} identifies the minimal position. Phrased in this setting, eliminating the minimum predicate thus means to get rid of the variable x_{min}.

It will turn out that on fragment level, we are unable to completely avoid x_{min}, but we can give a normal form that reduces the usage of x_{min} to very specific contexts. More precisely, x_{min} is only necessary to ensure a certain distance between the left border and a newly quantified variable. This is similar to the preceding chapter, where minimum distances between free variables were to be ensured. In fact, we reuse the framework introduced there, in particular generalized quantifiers and the sliding window radius $N(\varphi)$ from Definition 8.13. Remember that an extended formula is allowed to use generalized quantifiers of the form $\exists x \in \mathcal{C}\colon \varphi$, of the form $\forall x \in \mathcal{C}\colon \varphi$, and of the form $\exists^{r \bmod q} x \in \mathcal{C}\colon \varphi$. We also use the distance-stable extension $\widehat{\mathcal{F}}$, defined in Definition 8.5, which, for a mod-stable fragment $\mathcal{F}$, amounts to the ability to replace each ordinary quantifier by its extended version with an arbitrary constraint. We want to restrict these constraints to the form $\{i \mid |x_{min} - i| > d\}$, and we additionally want to bound the parameter d. This is formalized by the following normal form.

Definition 9.1
An extended formula is said to be k-bounded min-normalized *if it does not use the minimum predicate, if the variable x_{min} occurs only in quantifier constraints, and if all quantifier constraints are of the form $\{i \mid |x_{min} - i| > d\}$ with $d \leqslant k$.*

In such contexts, many fragments then allow to avoid the variable x_{min} altogether by guaranteeing the quantifier constraint using other means such as additional quantifier depth. This is true for example for all levels of the two-variable first-order logic that contain either the order predicate or the successor predicate, as shall be illustrated by Corollary 9.6.

We can now formulate the main theorem, which gives a normal form for the usage of the variable x_{min}, which is semantically correct under the assumption that enough information about the prefix of the structure at hand is available.

Theorem 9.2
For every sentence $\varphi \in$ FO+MOD and every $p \in \Lambda^+$ with $|p| > 2N(\varphi)$ there exists a $|p|$-bounded min-normalized extended formula ψ with $\mathrm{FV}(\psi) \subseteq \{x_{min}\}$ such that $\psi \leqslant_{\hat{\mathcal{F}}} \varphi$ for all mod-stable fragments $\mathcal{F}$ and $u \models \varphi$ if and only if $u, 1 \models \psi(x_{min})$ for all $u \in p\Lambda^+$.

The parameter $N(\varphi)$ is the sliding window radius introduced in Definition 8.1. The proof of this theorem employs a rather involved construction, so let us start with a brief intuition. All variables near the left border of the structure are handled by syntactic bookkeeping methods, which makes in particular the minimum predicate superfluous. For a concise terminology call such variables *eliminated* and call the others *free*.

At each quantification, the border of positions up to which variables are to be eliminated is increased. But by how many positions do we have to increase this border? Most numerical predicates need no increase at all, only the successor predicate poses a challenge. Specifically, to assess whether a free variable is the successor of an eliminated variable, we need to guarantee beforehand that there be at least one position in between. This is not sufficient because a subsequent quantification may place a variable precisely in the position in between and the same problem arises. However, a similar argument as for the sliding window construction in the preceding chapter leads to the successor distance $\mathrm{sd}_\varphi(x, y)$: Consider the quantification $\exists x\ \varphi$, for example. If y is eliminated and at position j, then x must not be placed on any position up to position $j + \mathrm{sd}_\varphi(x, y)$. For those positions, x must be eliminated. For all greater positions the variable x may be free. To quantify only over those positions for which x may be free, we use quantifier constraints.

We have to store which variables are eliminated and for eliminated variables also their position. This is formally implemented by a partial mapping $\delta\colon \mathbb{V}_1 \to_p \mathbb{Z}$. The idea is that x is free if $\delta(x)$ is undefined and that it is eliminated and notionally on position $\delta(x) + 1$ if $\delta(x)$ defined. In other words, $\delta(x)$ specifies the offset to the minimal position. To evaluate the label predicate on an eliminated variable x, we need to know the prefix of length at least $1 + \delta(x)$. Therefore, the construction to follow is relative to such a prefix $p \in \Lambda^*$.

For technical reasons that will become apparent later, a first step is to normalize formulae so that they use a new free variable x_{min} instead of the minimum predicate and so that quantifiers are only over formulae that have the quantified variable as a free variable. Specifically, we call a formula φ *preprocessed* if φ does not use the predicate min, the variable x_{min} only occurs in equality predicates, and whenever $\mathrm{Q}x\ \psi$ is a subformula of φ for a quantifier $\mathrm{Q} \in \{\exists,\ \forall,\ \exists^{r \bmod q}\}$, then either $x \in \mathrm{FV}(\psi)$ or $\psi \in \{\top,\ \bot\}$.

In the later application, it will be no restriction to demand that formulae be preprocessed by the following preprocessing: First replace each subformula $\min(x)$ by $x = x_{min}$. Then for every formula ψ with $x \notin \mathrm{FV}(\psi)$ replace each subformula $\exists x\ \psi$ by $\psi \wedge (\exists x\ \top)$, each subformula $\forall x\ \psi$ by $\psi \vee (\forall x\ \bot)$, and each subformula $\exists^{r \bmod q} x\ \psi$ by $(\psi \wedge \exists^{r \bmod q} x\colon \top) \vee (\neg\psi \wedge \exists^{r \bmod q} x\colon \bot)$. If φ' is the formula obtained from φ using this procedure, then φ' is preprocessed, $\varphi' \leqslant_{\mathcal{F}} \varphi$ for every mod-stable fragment $\mathcal{F}$, and $u \models \varphi$ if and only if $u, 1 \models \varphi'(x_{min})$ for all u.

The construction. A mapping $\delta\colon \mathbb{V}_1 \rightarrow_p \mathbb{Z}$ is a x_{min}-*offset function* if $\delta(x_{min})$ is defined with $\delta(x_{min}) = 0$ and if $\delta(x) \geqslant 0$ whenever $\delta(x)$ is defined.

We are now going to construct a min-normalized extended formulae $\langle\varphi\rangle_{p,\delta}$ for x_{min}-offset functions $\delta\colon \mathbb{V}_1 \rightarrow_p \mathbb{Z}$, prefixes $p \in \Lambda^*$, and preprocessed formulae $\varphi \in \mathrm{FO{+}M\small{OD}}$. The properties of this formula will formally be given later. To have an intuition in mind, informally speaking $\langle\varphi\rangle_{p,\delta}$ respects mod-stable fragments and satisfies the equivalence $u \models \varphi$ if and only if $u, 1 \models \langle\varphi\rangle_{p,\delta}(x_{min})$ for all structures u such that p is a prefix of u, eliminated variables are consistent with δ, and certain preconditions on the free variables are met.

Let $\langle\varphi\rangle_{p,\delta} := \varphi$ for $\varphi \in \{\top, \perp, \text{empty}\}$. For the label predicate $\lambda(x) \in B$ with $x \in \mathbb{V}_1$ and $B \subseteq \Lambda$ let

$$\langle\lambda(x) \in B\rangle_{p,\delta} \;:=\; \begin{cases} \lambda(x) \in B & \text{if } \delta(x) \text{ is undefined,}\\ \big[p[\delta(x)+1] \in B\big] & \text{if } \delta(x) \text{ is defined and } \delta(x) < |p|,\end{cases}$$

where $[A]$ for an *a priori* ascertainable assertion A is the logic variant of Iverson's bracket; it is $\top$ if A is true and $\perp$ if A is false.

This leaves the formula $\langle\lambda(x) \in B\rangle_{p,\delta}$ undefined whenever $\delta(x)$ is defined but $\delta(x)+1$ is not a position of p. We henceforth tacitly assume $\langle\varphi\rangle_{p,\delta}$ to be undefined if at any stage in its inductive construction an undefined formula is used.

For the equality predicate $x = y$, the successor predicate $\mathrm{suc}(x,y)$, and the order predicates $x \lesssim y$ with $\lesssim\, \in \{<, \leqslant\}$ and $x, y \in \mathbb{V}_1$ let

$$\langle x = y\rangle_{p,\delta} \;:=\; \begin{cases} x = y & \text{if } \delta(x) \text{ and } \delta(y) \text{ are undefined,}\\ \top & \text{if } \delta(x) \text{ and } \delta(y) \text{ are defined and } \delta(x) = \delta(y),\\ \perp & \text{otherwise,}\end{cases}$$

$$\langle \mathrm{suc}(x,y)\rangle_{p,\delta} \;:=\; \begin{cases} \mathrm{suc}(x,y) & \text{if } \delta(x) \text{ and } \delta(y) \text{ are undefined,}\\ \top & \text{if } \delta(x),\, \delta(y) \text{ are defined and } \delta(x)+1 = \delta(y),\\ \perp & \text{otherwise,}\end{cases}$$

$$\langle x \lesssim y\rangle_{p,\delta} \;:=\; \begin{cases} x \lesssim y & \text{if both } \delta(x) \text{ and } \delta(y) \text{ are undefined,}\\ \top & \text{if } \delta(x) \text{ is defined and } \delta(y) \text{ is undefined,}\\ \top & \text{if } \delta(x),\, \delta(y) \text{ are defined and } \delta(x) \lesssim \delta(y),\\ \perp & \text{otherwise.}\end{cases}$$

For the maximum predicate $\max(x)$ with $x \in \mathbb{V}_1$ let

$$\langle \max(x)\rangle_{p,\delta} \;:=\; \begin{cases} \max(x) & \text{if } \delta(x) \text{ is undefined,}\\ \perp & \text{otherwise.}\end{cases}$$

For the modular predicate $x \equiv r \pmod q$ with $x \in \mathbb{V}_1$ and $r, q \in \mathbb{Z}$ let

$$\langle x \equiv r \pmod q\rangle_{p,\delta} \;:=\; \begin{cases} x \equiv r \pmod q & \text{if } \delta(x) \text{ is undefined,}\\ \big[\delta(x)+1 \equiv r \pmod q\big] & \text{if } \delta(x) \text{ is defined.}\end{cases}$$

Remember that we do not have to consider the minimum predicate, because by construction it was replaced using the new free variable x_{min}. Note that this construction ensures that if $\langle\varphi\rangle_{p,\delta}$ is defined for an atomic formula φ, then it is either $\top$ or $\perp$

whenever δ is defined for any free variable of φ. In particular, $\langle\varphi\rangle_{p,\delta}$ uses neither the variable x_{min} nor the predicate min. For the Boolean connectives we inductively set

$$\begin{aligned}
\langle\neg\varphi\rangle_{p,\delta} &:= \neg\langle\varphi\rangle_{p,\delta},\\
\langle\varphi\vee\psi\rangle_{p,\delta} &:= \langle\varphi\rangle_{p,\delta}\vee\langle\psi\rangle_{p,\delta},\\
\langle\varphi\wedge\psi\rangle_{p,\delta} &:= \langle\varphi\rangle_{p,\delta}\wedge\langle\psi\rangle_{p,\delta}.
\end{aligned}$$

We come to the quantifiers and consider a formula of the form $\mathrm{Q}x\ \varphi$ for some quantifier $\mathrm{Q}\in\{\exists,\ \forall,\ \exists^{r \bmod q}\}$. We distinguish the special cases when $\mathrm{FV}(\varphi)\subseteq\{x\}$ or when there is no $y\in\mathrm{FV}(\varphi)\setminus\{x\}$ such that $\delta(y)$ is defined. In both cases we let $\langle\mathrm{Q}x\ \varphi\rangle_{p,\delta}:=\mathrm{Q}x\ \langle\varphi\rangle_{p,\delta\setminus\{x\}}$, where $\delta\setminus\{x\}$ is the offset function that is undefined on x and maps $y\neq x$ to $\delta(y)$. Assume $\mathrm{FV}(\varphi)\setminus\{x\}\neq\emptyset$ in the remainder of the construction.

Consider now the existential quantification $\exists x\ \varphi$ and the universal quantification $\forall x\ \varphi$ for some $x\in\mathbb{V}_1$ and some formula φ. Depending on x, φ, and δ define

$$\begin{aligned}
d &= \max\bigl(\{\delta(y)+\mathrm{sd}_\varphi(x,y)\mid y\in\mathrm{FV}(\varphi)\setminus\{x\}\text{ and }\delta(y)\text{ is defined}\}\bigr),\\
\mathcal{C} &= \{i\mid |x_{min}-i|>d\}.
\end{aligned}\tag{9.1}$$

Note that d is defined because assuming otherwise, we would have been in one of the above special cases. With this let

$$\begin{aligned}
\langle\exists x\ \varphi\rangle_{p,\delta} &:= \bigl(\exists x\in\mathcal{C}\colon \langle\varphi\rangle_{p,\delta\setminus\{x\}}\bigr)\vee\bigvee_{\ell\in\{0,\dots,d\}}\langle\varphi\rangle_{p,\delta[x/\ell]},\\
\langle\forall x\ \varphi\rangle_{p,\delta} &:= \bigl(\forall x\in\mathcal{C}\colon \langle\varphi\rangle_{p,\delta\setminus\{x\}}\bigr)\wedge\bigwedge_{\ell\in\{0,\dots,d\}}\langle\varphi\rangle_{p,\delta[x/\ell]},
\end{aligned}$$

where $\delta[x/\ell]$ is the function $x\mapsto\ell$ and $y\mapsto\delta(y)$ for all $y\neq x$.

Consider a modular quantification $\exists^{r \bmod q}x\ \varphi$ for some $x\in\mathbb{V}_1$, some $r,q\in\mathbb{Z}$, and some formula φ. The idea is to guess the set of positions Z on the prefix which make φ true, determining in particular how many positions on the remaining word have to make φ true. This is formalized by

$$\langle\exists^{r \bmod q}x\ \varphi\rangle_{p,\delta} := \bigvee_{Z\subseteq\{0,\dots,d\}}\Bigl(\bigwedge_{\ell\in\{0,\dots,d\}}\bigl[(\ell\in Z)\leftrightarrow\varphi(\ell)\bigr]\Bigr)\wedge\bigl(\exists^{s \bmod q}x\in\mathcal{C}\colon\langle\varphi\rangle_{p,\delta\setminus\{x\}}\bigr),$$

where $s=r-|Z|$ and

$$[(\ell\in Z)\leftrightarrow\varphi(\ell)] \;:=\; \begin{cases}\langle\varphi\rangle_{p,\delta[x/\ell]} & \text{if } \ell\in Z \text{ and}\\ \neg\langle\varphi\rangle_{p,\delta[x/\ell]} & \text{if } \ell\notin Z.\end{cases}$$

This concludes the definition of the construction. Let us emphasize again that the formula $\langle\varphi\rangle_{p,\delta}$ is not defined if p is too short. However, note that the only reason for $\langle\varphi\rangle_{p,\delta}$ not being defined lies ultimately in an atomic label predicate.

The next lemma gives sufficient conditions on the length of the prefix for the construction to be defined. For an offset function $\delta\colon\mathbb{V}_1\to_p\mathbb{Z}$ let $\overline{\delta}\colon\mathbb{V}_1\to\mathbb{Z}$ be given by $x\mapsto\delta(x)$ if $\delta(x)$ is defined and $x\mapsto 0$ otherwise.

Lemma 9.3

Let $\varphi\in\mathrm{FO{+}M{\scriptsize OD}}$ be preprocessed, let $\delta\colon\mathbb{V}_1\to_p\mathbb{Z}$ be an x_{min}-offset function, and let $p\in\Lambda^*$. If $|p|>N_{\overline{\delta}}(\varphi)$, then $\langle\varphi\rangle_{p,\delta}$ is defined and $|p|$-bounded min-normalized.

Proof. The proof is by induction on the structure of the formula.

Consider an atomic label predicate of the form $\lambda(x) \in B$ for some $x \in \mathbb{V}_1$ and some $B \subseteq \Lambda$. There is nothing to show if $\delta(x)$ is undefined. If $\delta(x)$ is defined, then $N_{\overline{\delta}}(\lambda(x) \in B) = \overline{\delta}(x) < |p|$ and $\langle \lambda(x) \in B \rangle_{p,\delta}$ is defined. For all other atomic formulae, the construction is always defined. Note that $\langle \varphi \rangle_{p,\delta}$ is vacuously $|p|$-bounded min-normalized for atomic formulae φ as by assumption φ cannot be the minimum predicate by the preprocessing. For the Boolean connectives the claim follows straightforwardly by induction.

It remains to consider a formula of the form $\mathrm{Q}x\ \varphi$ for some $\mathrm{Q} \in \{\exists, \forall, \exists^{r \bmod q}\}$. If $x \notin \mathrm{FV}(\varphi)$, then $\varphi \in \{\top, \bot\}$ by assumption on the formula, and the claim becomes trivial.

Suppose now $x \in \mathrm{FV}(\varphi)$. Let d be the bound defined by (9.1) and let $y \in \mathrm{FV}(\varphi) \setminus \{x\}$ be such that $\delta(y)$ is defined and $d = \delta(y) + \mathrm{sd}_\varphi(x, y)$. The formula $\langle \mathrm{Q}x\ \varphi \rangle_{p,\delta}$ is defined if $\langle \varphi \rangle_{p,\delta \setminus \{x\}}$ and $\langle \varphi \rangle_{p,\delta[x/\ell]}$ are defined for all $\ell \in \{0, \ldots, d\}$.

We have $N_{\overline{\delta \setminus \{x\}}}(\varphi) = N_{\overline{\delta}[x/0]}(\varphi) \leqslant N_{\overline{\delta}}(\mathrm{Q}x\ \varphi) < |p|$. The equality holds because $\overline{\delta \setminus \{x\}} = \overline{\delta}[x/0]$, the first inequality follows by definition of $N_{\overline{\delta}}(\mathrm{Q}x\ \varphi)$, and the second inequality is by assumption. By induction, $\langle \varphi \rangle_{p,\delta \setminus \{x\}}$ is defined and $|p|$-bounded min-normalized. This argument also suffices to show that $\langle \varphi \rangle_{p,\delta \setminus \{x\}}$ is defined and $|p|$-bounded min-normalized in the special case that there is no $y \in \mathrm{FV}(\varphi) \setminus \{x\}$ such that $\delta(y)$ is defined.

We have $N_{\overline{\delta[x/\ell]}}(\varphi) \leqslant N_{\overline{\delta}[x/d]}(\varphi) \leqslant N_{\overline{\delta}}(\mathrm{Q}x\ \varphi) < |p|$. The first inequality holds because $\overline{\delta[x/\ell]} = \overline{\delta}[x/\ell] \leqslant \overline{\delta}[x/d]$, the second inequality holds because $N_{\overline{\delta}}(\mathrm{Q}x\ \varphi) \geqslant N_{\overline{\delta}[x/d]}(\varphi)$ by definition of $N_{\overline{\delta}}(\mathrm{Q}x\ \varphi)$, and the last inequality is by assumption. Induction yields that $\langle \varphi \rangle_{p,\delta[x/\ell]}$ is defined and $|p|$-bounded min-normalized.

This shows that also in the non-special cases, $\langle \mathrm{Q}x\ \varphi \rangle_{p,\delta}$ is defined, and it remains to show $d \leqslant |p|$ in order to prove it $|p|$-bounded min-normalized. Note that all quantifier constraints introduced by proper subformulae are already dealt with. Now, $d \leqslant N_{\overline{\delta}[x/d]}(\varphi) \leqslant N_{\overline{\delta}}(\mathrm{Q}x\ \varphi) < |p|$, where the first inequality holds since $x \in \mathrm{FV}(\varphi)$, the second inequality is as above, and the last is by assumption. □

Correctness. The next definition gives admissible structures for which the above construction is correct. To formulate admissibility, let $\pi\colon (\Lambda \times 2^V)^* \to \Lambda^*$ be the homomorphic projection to the first component of the extended alphabet.

Definition 9.4 (Admissibility $\mathcal{W}_{\varphi,p,\delta}$)
Let $\varphi \in \mathrm{FO{+}M{\small OD}}$, let $\delta\colon \mathbb{V}_1 \to_p \mathbb{Z}$ be an x_{min}-offset function, and let $p \in \Lambda^*$ with $\langle \varphi \rangle_{p,\delta}$ being defined. A structure $u \in \mathcal{U}_V$ with $\mathrm{FV}(\varphi) \subseteq V$ is admissible *for $\langle \varphi \rangle_{p,\delta}$ if*

- *$\pi(u) \in p\Lambda^+$,*
- *$x_{min}(u) = 1$ and $x(u) = x_{min}(u) + \delta(x)$ for all $x \in \mathrm{FV}(\varphi)$ with $\delta(x)$ defined, and*
- *$y(u) - x(u) > \mathrm{sd}_\varphi(x, y)$ for all $x, y \in \mathrm{FV}(\varphi)$ with $\delta(x)$ defined and $\delta(y)$ undefined.*

Let $\mathcal{W}_{\varphi,p,\delta}$ be the set of admissible structures for $\langle \varphi \rangle_{p,\delta}$.

For admissible structures we can guarantee correctness of the construction. This is the next lemma.

Lemma 9.5
For all preprocessed $\varphi \in \mathrm{FO{+}M{\small OD}}$, all x_{min}-offset functions $\delta\colon \mathbb{V}_1 \to_p \mathbb{Z}$, and all $p \in \Lambda^*$ with $\langle\varphi\rangle_{p,\delta}$ being defined we have $u \models \varphi$ if and only if $u \models \langle\varphi\rangle_{p,\delta}$ for all $u \in \mathcal{W}_{\varphi,p,\delta}$.

Proof. The proof is by induction on the structure of the formula. For atomic formulae this follows directly from admissibility. Admissibility ensures that all information about variables for which δ is defined is available. Admissibility also provides sufficient information to assess the relation between variables if δ is defined on one variable but not on the other. As an example consider the successor predicate $\langle \mathrm{suc}(x,y)\rangle_{p,\delta}$. If $\delta(x)$ is defined but $\delta(y)$ is undefined, then admissibility yields $y(u) - x(u) > \mathrm{sd}_{\mathrm{suc}(x,y)}(x,y) = 1$ and in particular $u \not\models \mathrm{suc}(x,y)$. For Boolean combinations the claim follows straightforwardly by induction.

It remains to consider the quantifiers. Consider a quantification of the form $\mathrm{Q}x\ \varphi$ with $\mathrm{Q} \in \{\exists,\ \forall,\ \exists^{r \bmod q}\}$. In the first special case $x \notin \mathrm{FV}(\varphi)$ the preprocessing of the formula guarantees $\varphi \in \{\top, \bot\}$ and the claim is trivially true. The second special case is when there is no $y \in \mathrm{FV}(\varphi) \setminus \{x\}$ with $\delta(y)$ being defined. In this case vacuously $u[x/\ell] \in \mathcal{W}_{\varphi,p,\delta\setminus\{x\}}$ for every position ℓ of u. In particular, $u \models \mathrm{Q}x\ \varphi$ if and only if $\mathrm{Q}x\colon \langle\varphi\rangle_{p,\delta\setminus\{x\}}$.

We need two auxiliary claims to handle the remaining cases. Let $\mathcal{C}$ be the quantifier constraint defined by (9.1). We shall show $\tilde{u} = u[x/\ell] \in \mathcal{W}_{\varphi,p,\delta\setminus\{x\}}$ for all positions ℓ of u that meet the constraint $\mathcal{C}$. For this it suffices to show $y(\tilde{u}) - z(\tilde{u}) > \mathrm{sd}_\varphi(z,y)$ for all $z, y \in \mathrm{FV}(\varphi)$ with $(\delta \setminus \{x\})(z)$ defined and $(\delta \setminus \{x\})(y)$ undefined. Supposing $x \notin \{y,z\}$ yields $y(\tilde{u}) - z(\tilde{u}) = y(u) - z(u) > \mathrm{sd}_{\exists x \varphi}(z,y) \geqslant \mathrm{sd}_\varphi(z,y)$.

Suppose now $x \in \{y,z\}$. We cannot have $x = z$ because assuming so $(\delta \setminus \{x\})(z)$ would be undefined, in contradiction to the choice of z. So let $x = y$. By assumption the position ℓ meets the constraint $\mathcal{C}$, which implies $\ell > x_{min}(u) + \delta(z) + \mathrm{sd}_\varphi(x,z)$ as $\delta(z)$ is defined. Now, $y(\tilde{u}) = x(\tilde{u}) = \ell$ and $x_{min}(u) + \delta(z) = z(\tilde{u})$ which yields $y(\tilde{u}) - z(\tilde{u}) > \mathrm{sd}_\varphi(y,z)$.

The second auxiliary claim is $\tilde{u} = u[x/1+\ell] \in \mathcal{W}_{\varphi,p,\delta[x/\ell]}$ for all $\ell \leqslant d$, where d is defined by (9.1). For this it suffices to show $y(\tilde{u}) - z(\tilde{u}) > \mathrm{sd}_\varphi(z,y)$ for all $z, y \in \mathrm{FV}(\varphi)$ with $\delta[x/\ell](z)$ defined and $\delta[x/\ell](y)$ undefined.

Supposing $x \notin \{y,z\}$ yields $y(\tilde{u}) - z(\tilde{u}) = y(u) - z(u) > \mathrm{sd}_{\exists x \varphi}(z,y) \geqslant \mathrm{sd}_\varphi(z,y)$. It remains to consider $x \in \{y,z\}$. We cannot have $x = y$ because assuming so would imply that $\delta[x/\ell](y)$ is defined, contradicting the choice of y. Let therefore $x = z$. Suppose $d = \delta(t) + \mathrm{sd}_\varphi(x,t)$ for some $t \in \mathrm{FV}(\varphi) \setminus \{x\}$ with $\delta(t)$ defined. By admissibility of $u \in \mathcal{W}_{\exists x \varphi,p,\delta}$ we have $y(\tilde{u}) - t(\tilde{u}) > \mathrm{sd}_{\exists x \varphi}(y,t)$.

Using $t(\tilde{u}) = x_{min}(\tilde{u}) + \delta(t)$ and $\mathrm{sd}_{\exists x \varphi}(y,t) = \mathrm{sd}_\varphi(x,t) + \mathrm{sd}_\varphi(x,y)$, this implies $y(\tilde{u}) - x_{min}(\tilde{u}) > \delta(t) + \mathrm{sd}_\varphi(x,t) + \mathrm{sd}_\varphi(x,y) = d + \mathrm{sd}_\varphi(x,y)$. This in turn implies $y(\tilde{u}) - (x_{min}(\tilde{u}) + \ell) > \mathrm{sd}_\varphi(x,y)$. Therefore the desired inequality $y(\tilde{u}) - z(\tilde{u}) > \mathrm{sd}_\varphi(z,y)$ follows with $x_{min}(\tilde{u}) + \ell = x(\tilde{u}) = z(\tilde{u})$.

The first claim yields $u[x/1+\ell] \models \varphi$ if and only if $u[x/1+\ell] \models \langle\varphi\rangle_{p,\delta\setminus\{x\}}$ for every $\ell \in \{d+1, \dots, |u|-1\}$. This equivalence holds by induction and the first auxiliary claim above; note that a position i of u meets $\mathcal{C}$ if and only if $i > d+1$.

The second claim yields $u[x/1+\ell] \models \varphi$ if and only if $u[x/1+\ell] \models \langle\varphi\rangle_{p,\delta[x/\ell]}$ if and only if $u \models \langle\varphi\rangle_{p,\delta[x/\ell]}$ for $\ell \in \{0, \dots, d\}$. The first equivalence holds by induction and the second auxiliary claim above; the second equivalence holds since $x \notin \mathrm{FV}(\langle\varphi\rangle_{p,\delta[x/\ell]})$.

Consider now the existential quantification $\exists x\ \varphi$. The observations yield $u \models \exists x\ \varphi$ if and only if there exists $\ell \in \{0, \ldots, |u|-1\}$ such that $u[x/1+\ell] \models \varphi$ if and only if either there exists $\ell \in \{0, \ldots, d\}$ such that $u \models \langle\varphi\rangle_{p,\delta[x/\ell]}$, or $u \models \exists x \in \mathcal{C}\colon \langle\varphi\rangle_{p,\delta\setminus\{x\}}$. This shows the correctness of the construction for existential quantifiers. Universal quantifiers are dual.

It remains to consider the modular counting quantifier $\exists^{r \bmod q} x\ \varphi$. The construction makes the set Z of satisfying positions in the range $\{1, \ldots, 1+d\}$ explicit: It postulates $Z \subseteq \{0, \ldots, d\}$ such that for all positions $\ell \in \{0, \ldots, d\}$ we have $u[x/1+\ell] \models \varphi$ if and only if $\ell \in Z$. That Z indeed is this set is captured by the subformula $\bigwedge_{\ell\in\{0,\ldots,d\}}\big[(\ell \in Z) \leftrightarrow \varphi(\ell)\big]$. This follows using the above observations.

Knowing that there are $|Z|$ small satisfying positions, we can infer that in order to make $\exists^{r \bmod q} x\ \varphi$ true, there must be $s = r - |Z|$ positions (modulo q) that satisfy φ and meet the constraint $\mathcal{C}$. This is expressed by the subformula $\exists^{s \bmod q} x \in \mathcal{C}\colon \langle\varphi\rangle_{p,\delta\setminus\{x\}}$, as can be seen, again using the above observations. $\square$

Proof of Theorem 9.2. Let φ' be the formula obtained from φ using the preprocessing procedure described at the construction. For φ' we have $\mathrm{FV}(\varphi') \subseteq \{x_{min}\}$, the formula φ' is preprocessed, $\varphi' \leqslant_{\mathcal{F}} \varphi$ for every mod-stable fragment $\mathcal{F}$, and $u, 1 \models \varphi'(x_{min}) \Leftrightarrow u \models \varphi$ for all $u \in \Lambda^*$.

We claim that we can set ψ to be the formula $\langle\varphi'\rangle_{p,\{x_{min}\mapsto 0\}}$, where $\{x_{min} \mapsto 0\}$ is the x_{min}-offset function that maps x_{min} to 0 and is undefined for all other variables.

First, we show that ψ is defined and $|p|$-bounded min-normalized. For this it suffices to show $N_{\{x_{min}\mapsto 0\}}(\varphi') \leqslant 2N(\varphi)$ by Lemma 9.3. Without restriction we may assume that φ is of the form $\mathrm{Q}x\ \psi$ with $x \in \mathrm{FV}(\psi)$. We have $N_{\{x_{min}\mapsto 0\}}(\varphi') \leqslant \mathrm{sd}_\psi(x, x_{min}) + N(\varphi)$. The claim follows because $\mathrm{sd}_\psi(x, x_{min}) \leqslant N_{\{x\mapsto 0\}}(\psi) = N(\varphi)$. To show the inequality of this statement, we more generally claim that $N_\delta(\psi) \geqslant |\delta(x)| + \mathrm{sd}_\psi(x, x_{min})$ for all preprocessed formulae ψ, all $\delta\colon \mathrm{FV}(\psi) \to \mathbb{Z}$, and all $x \in \mathrm{FV}(\psi)$.

The proof of this claim is by induction on the structure of ψ. If ψ is atomic, then $\mathrm{sd}_\psi(x, x_{min}) = 0$ because x_{min} cannot occur in a successor predicate in a preprocessed formula. Since $x \in \mathrm{FV}(\psi)$, we have $N_\delta(\psi) \geqslant |\delta(x)| = |\delta(x)| + \mathrm{sd}_\psi(x, x_{min})$ by definition. Boolean combinations are canonical.

Consider a quantification $\mathrm{Q}z\ \psi$ for some $z \notin \{x, x_{min}\}$, some $\mathrm{Q} \in \{\exists,\ \forall,\ \exists^{r \bmod q}\}$, and some $\psi \in \mathrm{FO{+}M{\small OD}}$. Note that we do not need to consider $z \in \{x, x_{min}\}$ since $x \in \mathrm{FV}(\mathrm{Q}z\ \psi)$ and x_{min} is never quantified in a preprocessed formula. Remember that $\mathrm{sd}_{\mathrm{Q}z\,\psi}(x, x_{min}) = \max\{\mathrm{sd}_\psi(x, x_{min}),\ \mathrm{sd}_\psi(x, z) + \mathrm{sd}_\psi(z, x_{min})\}$.

First, suppose $\mathrm{sd}_{\mathrm{Q}z\,\psi}(x, x_{min}) = \mathrm{sd}_\psi(x, x_{min})$. In this case we have

$$N_\delta(\mathrm{Q}z\ \psi) \geqslant N_{\delta[z/0]}(\psi) \geqslant |\delta(x)| + \mathrm{sd}_\psi(x, x_{min}) = |\delta(x)| + \mathrm{sd}_{\mathrm{Q}z\,\psi}(x, x_{min}).$$

The first inequality is by definition of $N_\delta(\psi)$. The second is by induction (note that $x \in \mathrm{FV}(\mathrm{Q}z\ \psi) \subseteq \mathrm{FV}(\psi)$ and in particular $x \neq z$).

Supposing otherwise, $\mathrm{sd}_{\mathrm{Q}z\,\psi}(x, x_{min}) = \mathrm{sd}_\psi(x, z) + \mathrm{sd}_\psi(z, x_{min})$ yields

$$N_\delta(\mathrm{Q}z\ \psi) \geqslant N_{\delta[z/|\delta(x)|+\mathrm{sd}_\psi(x,z)]}(\psi) \geqslant |\delta(x)| + \mathrm{sd}_\psi(x, z) + \mathrm{sd}_\psi(z, x_{min}).$$

The first inequality is by definition of $N_\delta(\mathrm{Q}z\ \psi)$. The second is by induction. Note that we may assume $z \in \mathrm{FV}(\psi)$ because ψ is preprocessed. Thus in this case also $N_\delta(\mathrm{Q}z\ \psi) \geqslant |\delta(x)| + \mathrm{sd}_{\mathrm{Q}z\,\psi}(x, x_{min})$ as desired. This concludes the proof of the claim.

It is straightforward to verify $\psi \leqslant_{\widehat{\mathcal{F}}} \varphi' \leqslant_{\mathcal{F}} \varphi$ for all mod-stable fragments. That $\mathrm{FV}(\psi) \subseteq \{x_{min}\}$ follows from $\mathrm{FV}(\langle\xi\rangle_{p,\delta}) \subseteq (\mathrm{FV}(\xi) \setminus \{x \in \mathbb{V}_1 \mid \delta(x) \text{ defined}\}) \cup \{x_{min}\}$ for all $\xi \in \mathrm{FO{+}M{\scriptstyle OD}}$ and all x_{min}-offset functions δ. The latter statement can be verified by a straightforward induction. Finally by Lemma 9.5, $u \models \varphi$ if and only if $u, 1 \models \varphi'(x_{min})$ if and only if $u, 1 \models \psi(x_{min})$ for all $u \in p\Lambda^+$. To see that Lemma 9.5 truly applies, note that $u[x_{min}/1] \in \mathcal{W}_{\varphi',p,\{x_{min}\mapsto 0\}}$. ☐

The following is a consequence of Theorem 9.2 for the alternation hierarchy within the two-variable first order logic FO^2. It will be used later in Section 12.3 as an intermediate step of a decidability result.

Corollary 9.6
Let $\mathcal{N}$ be a signature containing $\{<, \mathrm{suc}\}$, and let $m, n \geqslant 2$ be integers. For each sentence $\varphi \in \Sigma^2_{m,n}[\mathcal{N}, \min]$ and each finite alphabet $A \subseteq \Lambda$ there exists a sentence $\psi \in \Sigma^2_{m,2n}[\mathcal{N}]$ such that $\mathcal{L}_A(\varphi) = \mathcal{L}_A(\psi)$.

Proof. Distinguishing between words of length at most $2n-1$ and longer words, we can write the language $\mathcal{L}_A(\varphi)$ as a finite disjunction

$$\mathcal{L}_A(\varphi) = L' \cup \bigcup_{p \in A^{2n-1}} L_p,$$

where $L' = \{u \in \mathcal{L}_A(\varphi) \mid |u| \leqslant 2n-1\}$ and $L_p = pA^+ \cap \mathcal{L}_A(\varphi)$. For each word $u \in L'$, the singleton language $\{u\}$ is definable in $\Sigma^2_{2,2n}[<, \mathrm{suc}]$. Since L' is finite, closure under conjunction yields that L' is definable in $\Sigma^2_{2,2n}[<, \mathrm{suc}]$.

It remains to show that the languages L_p are $\Sigma^2_{m,2n}[\mathcal{N}]$-definable. Because pA^+ is $\Sigma^2_{2,2n}[<, \mathrm{suc}]$-definable for $|p| = 2n-1$, it suffices to give a sentence $\psi_p \in \Sigma^2_{m,2n}[\mathcal{N}]$ such that $\mathcal{L}_A(\psi_p) \cap pA^+ = \mathcal{L}_A(\varphi) \cap pA^+$.

We have $N(\varphi) \leqslant n-1$ by Proposition 8.15. Therefore $|p| = 2n-1 > 2N(\varphi)$ and Theorem 9.2 applies. Let ψ'_p be the $|p|$-bounded min-normalized formula for φ postulated by Theorem 9.2. If we now replaced all generalized quantifiers in ψ'_p by their ordinary counterparts, we would end up with a formula in $\Sigma^2_{m,n}[\mathcal{N}]$ — which would of course not be equivalent to ψ'_p in general.

We have to be more careful in the replacement to ensure equivalence. For this we introduce auxiliary formulae $x \geqslant x_{min} + d$ with free variable x for $d \in \mathbb{N}$, which ensures that there are at least d positions strictly to the left of x. Note that in this formula, x_{min} is actually just a formal symbol to convey the meaning and not a variable. The formula is given by the recursive definition

- $(x \geqslant x_{min} + 0) \ := \ \top$,
- $(x \geqslant x_{min} + 2d + 1) \ := \ \exists y \, \big(y < x \wedge (y \geqslant x_{min} + 2d)\big)$,
- $(x \geqslant x_{min} + 2d + 2) \ := \ \exists y \, \big(y < x \wedge \neg\,\mathrm{suc}(y, x) \wedge (y \geqslant x_{min} + 2d)\big)$.

The formulae $y \geqslant x_{min} + \ell$ with free variable y are obtained as usual by switching the roles of x and y. It is clear by definition that $x \geqslant x_{min} + d$ is in $\Sigma^2_{1,\lceil d/2 \rceil}[<, \mathrm{suc}]$.

Let ψ_p be the sentence obtained from ψ'_p as the fixed point of the replacement rules

$$\mu\big(\exists x \in \{i \mid |x_{min} - i| > d\} \colon \xi\big) \quad \mapsto \quad \mu\big(\exists x \; (x \geqslant x_{min} + d + 1) \wedge \xi\big)$$

for all extended contexts μ, all $x \in \mathbb{V}_1$, all extended formulae ξ, and all $d \in \mathbb{N}$.

By definition of Σ^2_m, the formula φ and ψ'_p only use existential quantifiers. Moreover, all occurrences of the variable x_{min} in ψ'_p are in some quantifier constraint. Hence ψ_p is a sentence without occurrence of the variable x_{min}. Because ψ'_p is $(2n-1)$-bounded, the replacement procedure increases the quantifier depth by at most n and $\psi_p \in \Sigma^2_{m,2n}[\mathcal{N}]$. By the semantics of the formulae $x \geqslant x_{min} + d$, we have $u \models \psi_p$ if and only if $u, 1 \models \psi'_p(x_{min})$ if and only if $u \models \varphi$ for all $u \in pA^+$. $\square$

An immediate consequence of this is the following result for the full levels of the alternation hierarchy.

Corollary 9.7
Let $\mathcal{N}$ be a signature containing $\{<, \mathrm{suc}\}$, and let $m, n \geqslant 2$ be integers. For each sentence $\varphi \in \mathrm{FO}^2_{m,n}[\mathcal{N}, \min]$ and each finite alphabet $A \subseteq \Lambda$ there exists a sentence ψ in $\mathrm{FO}^2_{m,2n}[\mathcal{N}]$ such that $\mathcal{L}_A(\varphi) = \mathcal{L}_A(\psi)$. $\square$

The maximum predicate. For completeness, we record the left-right dual of the Minimum Elimination Theorem 9.2. Let an extended formula be *k-bounded max-normalized* if it does not use the maximum predicate, the variable x_{max} occurs only in quantifier constraints, and all quantifier constraints are of the form $\{i \mid |x_{max} - i| > d\}$ with $d \leqslant k$.

Theorem 9.8
For every sentence $\varphi \in \mathrm{FO{+}M{\small OD}}$ and every $q \in \Lambda^+$ with $|q| > 2N(\varphi)$ there exists a $|q|$-bounded max-normalized extended formula ψ with $\mathrm{FV}(\psi) \subseteq \{x_{max}\}$ such that $\psi \leqslant_{\widehat{\mathcal{F}}} \varphi$ for every mod-stable fragment $\mathcal{F}$ and such that $u \models \varphi$ if and only if $u, |u| \models \psi(x_{max})$ for all $u \in \Lambda^+ q$. $\square$

10. Requantification-Free Normal Form within Fragments

The goal of this chapter is to formally prove the following fact about quantifiers: It is not reasonable to requantify the same variable by two successive quantifiers. More precisely, we show that there is an equivalent normal form where the inner quantification is moved to the same level as the outer quantification. Moreover, this normal form is not more complicated than the original formula with respect to prefragments. We give formal justification on an abstract level for this intuitively plausible fact. The normal form will be used later in Chapter 11.

We shall use the notion of subformulae a lot in this chapter, and we denote by $\psi \leqslant \varphi$ that ψ is a subformula of φ, and by $\psi < \varphi$ that ψ is a *proper subformula* of φ; *i.e.*, a subformula of φ that is not φ itself. Recall that a formula ψ is a subformula of φ if there exists a context μ such that $\varphi = \mu(\psi)$. Here, a context μ is a formula with a unique occurrence of a 0-ary placeholder $\circ$, and $\mu(\psi)$ substitutes ψ for $\circ$.

For a uniform presentation, we denote by $\mathfrak{z} \in \mathbb{V}_1 \cup \mathbb{V}_2$ a variable that might *a priori* be first-order or second-order. A formula φ is said to be *root quantifying* over $\mathfrak{z}$ if it is of the form $\exists \mathfrak{z}\ \psi$, $\forall \mathfrak{z}\ \psi$, or $\exists^{r \bmod q} \mathfrak{z}\ \psi$. The following definition formalizes what we mean by a successive requantification of the same variable.

Definition 10.1 (Requantification)
A formula φ is requantifying *if there exists a subformula $\psi < \varphi$ and a variable $\mathfrak{z} \in \mathbb{V}_1 \cup \mathbb{V}_2$ such that ψ and φ are root quantifying over $\mathfrak{z}$ and there exists no root quantifying formula χ with $\psi < \chi < \varphi$.*

In terms of the parse-tree of φ this means that the root is a quantifier-node over $\mathfrak{z}$ and there exists a path without any intermediate quantifier-nodes to another quantifier-node over the same variable $\mathfrak{z}$. A formula φ is *requantification-free* if there does not exist a subformula ψ of φ that is requantifying.

The following stability notion formalizes that quantifiers may be deleted if the formula they quantify does not depend on the quantified variable.

Definition 10.2 (Quantifier-stability)
A prefragment $\mathcal{F}$ is quantifier-stable *if for all φ, all $x \in \mathbb{V}_1$, all $X \in \mathbb{V}_2$, and all $q, r \in \mathbb{Z}$:*

1. $\varphi \leqslant_{\mathcal{F}} (\exists x\ \varphi)$ *and* $\varphi \leqslant_{\mathcal{F}} (\forall x\ \varphi)$ *whenever* $x \notin \mathrm{FV}(\varphi)$*.*
2. $\varphi \leqslant_{\mathcal{F}} (\exists X\ \varphi)$ *and* $\varphi \leqslant_{\mathcal{F}} (\forall X\ \varphi)$ *whenever* $X \notin \mathrm{FV}(\varphi)$*.*
3. $\varphi \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$ *and* $\neg\varphi \leqslant_{\mathcal{F}} (\exists^{r \bmod q} x\ \varphi)$ *whenever* $x \notin \mathrm{FV}(\varphi)$*.*

One might wonder about the negation introduced by axiom (3). We shall see at the end of this chapter that this cannot be avoided; *cf.* Example 10.4.

The following is the main proposition of this chapter. It gives a requantification-free normal form within quantifier-stable prefragments. Actually it shows a stronger assertion as there is even one normal form that works for all quantifier-stable prefragments. Two formulae φ and ψ are said to be *equivalent* if $\mathrm{FV}(\varphi) = \mathrm{FV}(\psi) = V$ and $[\varphi]_V = [\psi]_V$.

Proposition 10.3 (Requantification-free normal form)
For every formula φ there exists an equivalent requantification-free formula $\hat{\varphi}$ such that $\hat{\varphi} \leqslant_{\mathcal{F}} \varphi$ for every quantifier-stable prefragment $\mathcal{F}$.

Proof. We proceed by induction on the structure of the formula. If φ is already requantification-free, then there is nothing to show. Note that in particular atomic formulae are requantification-free.

By induction it suffices to consider the case when φ is itself requantifying. By definition there exists a subformula $\psi < \varphi$ and a variable $\mathfrak{z}$ such that ψ and φ are root quantifying over $\mathfrak{z}$ and there exists no root quantifying formula χ with $\psi < \chi < \varphi$. By induction we may assume ψ to be requantification-free. We can decompose φ as $\mathrm{Q}\mathfrak{z}\colon \mu(\psi)$ for some $\mathrm{Q} \in \{\exists, \forall, \exists^{r \bmod q}\}$ and some context μ.

We define a derived context μ', which is obtained from μ by pruning all branches of the parse tree not containing the placeholder $\circ$. More formally, let $\circ' = \circ$, let $(\neg\nu)' := \neg(\nu')$, and let

$$(\nu_1 \vee \nu_2)' \;:=\; (\nu_1 \wedge \nu_2)' \;:=\; \begin{cases} \nu_1' & \text{if } \circ \text{ occurs in } \nu_1, \\ \nu_2' & \text{if } \circ \text{ occurs in } \nu_2. \end{cases}$$

What remains in μ' are the negations on the path from the root to $\circ$ in the parse tree of μ. Let $\|\mu'\|$ be the number of negations occurring in μ'. By construction this number coincides with the negation depth over $\circ$ in μ. Let

$$[\mu'(\top)] \;:=\; \begin{cases} \top & \text{if } \|\mu'\| \text{ is even}, \\ \bot & \text{if } \|\mu'\| \text{ is odd}. \end{cases}$$

This notation will be convenient to formalize the following uniformly. Interpreting $\mu'(\top)$ as $\neg\cdots\neg\top$ with $\|\mu'\|$ negations, this is the logic counterpart of the usual Iverson bracket $[A]$ which is $\top$ if the assertion A is true and $\bot$ else. Accordingly $[\mu'(\bot)]$ is $\bot$ if $\|\mu'\|$ is even and $\top$ otherwise.

First consider the case of modular counting quantifiers $\varphi = \exists^{r \bmod q}\mathfrak{z}\colon \mu(\psi)$. Let

$$\hat{\varphi} \;:=\; \big(\mu'(\psi) \wedge \mathrm{Q}\mathfrak{z}\colon \mu([\mu'(\top)])\big) \vee \big(\neg\mu'(\psi) \wedge \mathrm{Q}\mathfrak{z}\colon \mu([\mu'(\bot)])\big).$$

This formula realizes a straightforward case distinction: The left branch of the disjunction is for the case of ψ being true, the right branch for ψ being false. To see semantic correctness, note that the truth value of ψ is the same in the context $\circ$ as in the context $\mathrm{Q}\mathfrak{z}\colon \mu(\circ)$. In the parse tree of μ there are no quantifiers on the path from the root to $\circ$. Therefore, the interpretation of all free variables of ψ is the same in both contexts. Knowing the truth value of $\mu'(\psi)$, we can infer that of ψ on syntax level. All we need is the parity of the negation depth over $\circ$ in μ', which is precisely what $[\mu'(\top)]$ and $[\mu'(\bot)]$ do.

For first-order and second-order quantifiers $\varphi = \exists\mathfrak{z}\ \mu(\psi)$ or $\varphi = \forall\mathfrak{z}\ \mu(\psi)$ let

$$\hat{\varphi} \;:=\; \big(\mu'(\psi) \wedge \mathrm{Q}\mathfrak{z}\colon \mu([\mu'(\top)])\big) \vee \big(\mathrm{Q}\mathfrak{z}\colon \mu([\mu'(\bot)])\big).$$

Let us establish semantic correctness of this construction. Let $V = \mathrm{FV}(\varphi)$ and consider a structure $u \in \mathcal{U}_V$. Note that $\mathrm{FV}(\hat{\varphi}) \subseteq V$. First consider the case $u \not\models \mu'(\psi)$. In this case $\hat{\varphi}$ is equivalent to its right branch of the disjunction, that is to say, $u \models \hat{\varphi}$ if and only if $u \models \mathrm{Q}\mathfrak{z}\colon \mu([\mu'(\bot)])$. But $[\mu'(\bot)]$ is just the truth value of ψ in the context

$Q\mathfrak{z}\colon \mu(\circ)$. Formally, $[\mu'(\bot)] = \top$ if and only if $u[\mathfrak{z}/\mathfrak{i}] \models \psi$ for all interpretations $\mathfrak{i}$ of $\mathfrak{z}$.[1] Of course, syntactically inserting the truth value of the subformula does not alter the truth value of the original formula.

Suppose now $u \models \mu'(\psi)$. The left branch of the disjunction of $\hat{\varphi}$ is true if and only if φ is true. Formally, $u \models Q\mathfrak{z}\colon \mu([\mu'(\top)])$ if and only if $u \models \varphi$. This follows by the same reasoning as in the previous case, because in this case $[\mu'(\top)]$ is precisely the truth value of ψ in the context $Q\mathfrak{z}\colon \mu(\circ)$. In particular, if $u \models \varphi$, then $u \models \hat{\varphi}$. It remains to show the converse implication.

Suppose $u \models \hat{\varphi}$. If the left branch of the disjunction in $\hat{\varphi}$ is true, then we already saw that $u \models \varphi$. Assume now the right branch to be true; *i.e.*, assume $u \models Q\mathfrak{z}\colon \mu([\mu'(\bot)])$. We shall show that if $\mathfrak{i}$ is an interpretation of $\mathfrak{z}$ such that $u[\mathfrak{z}/\mathfrak{i}] \models \mu([\mu'(\bot)])$, then $u[\mathfrak{z}/\mathfrak{i}] \models \mu(\psi)$. This then yields the claim. Note that $[\mu'(\bot)]$ is by construction the negation of the truth value $u[\mathfrak{z}/\mathfrak{i}] \models \psi$.

Suppose $u[\mathfrak{z}/\mathfrak{i}] \models \mu([\mu'(\bot)])$ and assume on the contrary $u[\mathfrak{z}/\mathfrak{i}] \not\models \mu(\psi)$. The pruning process to obtain μ' can also be thought of as replacing the branches not containing $\circ$ by the neutral truth value $\bot$ for disjunction and $\top$ for conjunction. For the difference in the truth value of $[\mu'(\bot)]$ and ψ to come through in the context μ, all branches not containing $\circ$ must evaluate to the neutral truth value, or, in other words, μ behaves just like μ'. To illustrate this point consider the context $\nu_1 \vee \nu_2$, for example. Suppose the right branch ν_2 contains $\circ$, and that the left branch ν_1 evaluates to $\top$; *i.e.*, $u[\mathfrak{z}/\mathfrak{i}] \models \nu_1$. In this case $\nu_1 \vee \nu_2$ also evaluates to true, completely independent of the truth value of ν_2. The only cause of a different outcome is thus obliterated.

Formally this means $u[\mathfrak{z}/\mathfrak{i}] \models \mu(\psi)$ if and only if $u[\mathfrak{z}/\mathfrak{i}] \models \mu'(\psi)$. Truth of $\mu'(\psi)$ is independent of the interpretation of $\mathfrak{z}$ because $\mathfrak{z} \notin \mathrm{FV}(\mu'(\psi))$. Altogether, this shows

$$u[\mathfrak{z}/\mathfrak{i}] \models \mu(\psi) \text{ if and only if } u \models \mu'(\psi),$$

a contradiction as the current case is $u[\mathfrak{z}/\mathfrak{i}] \not\models \mu(\psi)$ and $u \models \mu'(\psi)$.

Note that $\mu'(\psi) \leqslant_{\mathcal{F}} \mu(\psi)$ and $\mathfrak{z} \notin \mathrm{FV}(\mu'(\psi))$. Using these facts, a straightforward verification shows $\hat{\varphi} \leqslant_{\mathcal{F}} \varphi$ for all quantifier-stable prefragments $\mathcal{F}$. $\square$

Note that in the preceding proof the same construction as for ordinary quantifiers does not work for modular counting quantifiers. The whole argument for ordinary quantifiers is based on the following monotony property: Let $\mathfrak{I}_{u,\mathfrak{z}}(\zeta)$ consist of all interpretations $\mathfrak{i}$ of $\mathfrak{z}$ such that $u[\mathfrak{z}/\mathfrak{i}] \models \zeta$. The first-order and second-order quantifiers $Q \in \{\exists, \forall\}$ are *monotonic* in the sense that $u \models Q\mathfrak{z}\colon \zeta$ implies $u \models Q\mathfrak{z}\colon \chi$ for all formulae ζ and χ such that $\mathfrak{I}_{u,\mathfrak{z}}(\zeta) \subseteq \mathfrak{I}_{u,\mathfrak{z}}(\chi)$.

This monotony property fails to hold for modular counting quantifiers as the quantifier may well switch from true to false, even though more interpretations make the quantified formula true. Consider, for example, the quantifier $\exists^{0 \bmod 2}x$ and let u be a structure of odd length. We have $u \models \exists^{0 \bmod 2}x\ \bot$ and $u \not\models \exists^{0 \bmod 2}x\ \top$ even though clearly $\mathfrak{I}_{u,x}(\bot) \subseteq \mathfrak{I}_{u,x}(\top)$.

Because modular counting quantifiers are not monotonic in this sense, negations are included in axiom (3) of quantifier-stability. This is truly necessary and not merely for want of a better construction as the following example shows.

[1] The type of the interpretation $\mathfrak{i}$ depends on the type of $\mathfrak{z}$, of course. For a first-order variable $\mathfrak{z} \in \mathbb{V}_1$ it is a position in $\{1, \ldots, |u|\}$. For a second-order variable $\mathfrak{z} \in \mathbb{V}_2$ it is a subset of $\{1, \ldots, |u|\}$.

Example 10.4
Let $\mathcal{F}$ be the prefragment generated by the formula $\varphi := \exists^{0 \bmod 2} x\ (\exists^{0 \bmod 2} x \colon \lambda(x) = a)$. Note that $\mathcal{F}$ is not quantifier-stable. The formula φ is requantifying and holds if the length is even or if there is an odd number of a-positions. Every requantification-free formula in $\mathcal{F}$ is equivalent to $\top$, $\bot$, or one of the following formulae:

- $\exists^{0 \bmod 2} x \colon \top$,
- $\exists^{0 \bmod 2} x \colon \lambda(x) = a$,
- $(\exists^{0 \bmod 2} x \colon \top) \wedge (\exists^{0 \bmod 2} x \colon \lambda(x) = a)$,
- $(\exists^{0 \bmod 2} x \colon \top) \vee (\exists^{0 \bmod 2} x \colon \lambda(x) = a)$.

Now, any of these formulae is inequivalent to φ: This is obvious for $\top$; for all other formulae observe that, in contrast to φ, neither of them is satisfied by the single letter word a. This in particular shows that there is no equivalent requantification-free formula $\hat{\varphi}$ with $\hat{\varphi} \leqslant_{\mathcal{F}} \varphi$. $\diamond$

Part II.

Quantifier Alternation in Two-Variable First-Order Logic

This part of the thesis studies the *quantifier alternation hierarchy* or just *alternation hierarchy* of two-variable first-order logic. This hierarchy is depicted in Figure 10.1. It consists of the *full levels* FO^2_m and the intermediate *half levels* Σ^2_m. Recall that for these levels at most $m-1$ alternations between existential and universal quantifiers are allowed on any path in the parse tree of the formula. For the full levels the type of the first quantifier does not matter, whereas the half levels restrict formulae to start with existential quantifiers. It is actually more convenient for formal reasoning to disregard universal quantifiers and count the nesting depth of negations. This faithfully corresponds to the alternation hierarchy as a negation is semantically an alternation between the two quantifier types. Refer to Section 4.2 for a formal definition of the quantifier alternation hierarchy.

Chapter 11 establishes a combinatorial characterization of the alternation hierarchy in terms of so-called rankers. Rankers are a means to navigate upon a word using simple instructions as atomic building blocks.

Chapter 12 gives an affirmative answer to the question whether definability in a certain level of the alternation hierarchy is decidable. Even though the ranker characterization of the alternation hierarchy does not show decidability, it is a useful tool in the proofs.

In contrast to two-variable logic, the quantifier alternation hierarchy of full first-order logic FO corresponds to the so-called dot-depth hierarchy of star-free languages, for which very little is known about decidability: Only the levels Σ_1, FO_1, and Σ_2 are known to be decidable [Pin95; PW97; Sim75; Kna83; Arf91; GS08]; refer also to Section 12.2 for proofs of decidability of Σ_1 and FO_1. To date, decidability of all other levels remain open.[2]

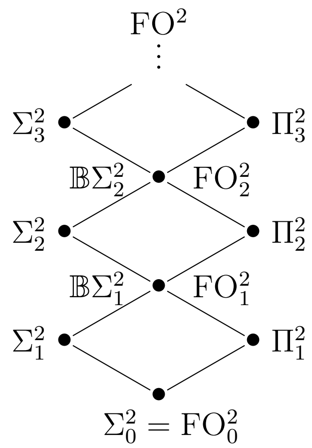

Figure 10.1.: The alternation hierarchy within two-variable first-order logic FO^2. The Boolean closure of Σ^2_m is denoted by $\mathbb{B}\Sigma^2_m$ and Π^2_m consists of negations of Σ^2_m-formulae.

[2]A paper that was published after this thesis was submitted shows that FO_2 and Σ_3 are also decidable [PZ14].

11. Rankers for the Quantifier Alternation Hierarchy

The present chapter characterizes the expressive power of the m^{th} half level Σ^2_m and the m^{th} full level FO^2_m of the two-variable alternation hierarchy in terms of rankers.

Rankers are sequences of simple instructions like “go to the next a-position” or “go to the previous a-position”. If a ranker is defined on a word, it identifies a unique position. The main result of this chapter is that the expressive power of the half levels as well as that of the full levels is completely determined by the ability to express certain order comparisons between certain pairs of rankers. More concretely, the m^{th} level $\mathrm{FO}^2_m[<]$ can require from a word that two rankers r and s have the same relative order (*i.e.*, smaller, equal, greater), provided that r and s alternate at most $m-1$ times the direction and end in different directions. The level $\Sigma^2_m[<]$ is more asymmetric in that it can only ensure that r is left of s whenever r ends right moving and s ends left moving; in general, however, it cannot enforce that r is right of s for such rankers. Including a successor predicate, $\mathrm{FO}^2_m[<, \mathrm{suc}]$ and $\Sigma^2_m[<, \mathrm{suc}]$ admit similar ranker characterizations when rankers are allowed a look-around.

The first explicit occurrence of rankers is in [STV02] under the name *turtle program*, where they are used as an intermediate step towards an automaton model for $\mathrm{FO}^2[<]$. Over finite words two variable first-order logic $\mathrm{FO}^2[<]$ is expressively complete for the family of languages $\Delta_2[<]$ consisting of all languages that are $\Sigma_2[<]$-definable and whose complement is also $\Sigma_2[<]$-definable. This is not true over infinite words as there are languages definable in $\mathrm{FO}^2[<]$ over infinite words that are not in $\Delta_2[<]$; *cf.* [DK11]. Rankers are also the main tool to extend Schwentick, Thérien, and Vollmer’s result to infinite words, providing a similar Büchi-automaton model for $\Delta_2[<]$ over infinite words [KL11c], building on an earlier characterization of $\Delta_2[<]$ over infinite words in terms of rankers [DKL10].

The term *ranker* itself was coined by Weis and Immerman [WI09], who strengthened the correspondence between rankers and FO^2 to a characterization of the full levels $\mathrm{FO}^2_{m,n}[<]$ and $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$, with an additional bound n on the quantifier depth. This chapter refines this correspondence even further to include the half levels $\Sigma^2_{m,n}[<]$ and $\Sigma^2_{m,n}[<, \mathrm{suc}]$, whence the characterizations of the full levels immediately follow. Moreover, the result of Weis and Immerman is erroneous in the case with successor predicate, as their “successor rankers” are actually too weak to capture quantifier depth. This is rectified and a proper look-around ranker description of $\Sigma^2_{m,n}[<, \mathrm{suc}]$ and $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$ is given.

The organization of this chapter is as follows. Rankers are introduced in Section 11.1. There are two kinds of rankers, with and without look-around. Rankers without look-around are used for the characterization of the $\mathrm{FO}^2[<]$-alternation hierarchy given in Section 11.3. Look-around rankers yield the description of the $\mathrm{FO}^2[<, \mathrm{suc}]$-alternation hierarchy of Section 11.4. A convenient intermediate step in our proofs, which is also interesting in its own right, is *temporal logic*, introduced in Section 11.2.

11.1. Rankers and Look-Around Rankers

Let henceforth A be an arbitrary but fixed finite alphabet over which word models are built. A *ranker* is a non-empty word over the alphabet $\{\mathrm{X}_a,\, \mathrm{Y}_a \mid a \in A\}$. The modality X_a is short for *neXt-a* and means "go to the next a-position"; the modality Y_a is short for *Yesterday-a* and means "go to the previous a-position". A ranker is a sequence of such instructions, which is evaluated from left to right; it is undefined if at some stage an instruction cannot be executed. The first modality, which is not provided with a designated starting position, goes to the first (if it is X_a) or last (if it is Y_a) a-position. Intuitively this means that rankers starting with an X-modality begin their evaluation in front of the word, whereas rankers starting with a Y-modality begin behind the word.

More precisely, for a word $u \in A^*$ and a position $i \in \{1, \ldots, |u|\}$ let

$$\begin{aligned} \mathrm{X}_a(u,i) &= \min\{j \in \{i+1, \ldots, |u|\} \mid u[j] = a\}, \\ \mathrm{Y}_a(u,i) &= \max\{j \in \{1, \ldots, i-1\} \mid u[j] = a\}, \end{aligned}$$

with minimum and maximum of the empty set being undefined. The symbol Z often denotes one of X or Y. For a composite ranker $r = \mathrm{Z}_a s$ with $\mathrm{Z} \in \{\mathrm{X}, \mathrm{Y}\}$ let inductively

$$r(u,i) = s(u, \mathrm{Z}_a(u,i))$$

be the position reached by the ranker s from the starting position $\mathrm{Z}_a(u,i)$; *i.e.*, rankers are evaluated from left to right. If either of $\mathrm{Z}_a(u,i)$ or s is undefined, then $r(u,i)$ is also undefined. It is sometimes convenient to also consider the *empty ranker* ε, for which $\varepsilon(u,i) = i$.

This describes the behavior of a ranker on a word with a designated starting position. The *evaluation* of a ranker r on a word u is $r(u) = s(u, \mathrm{Z}_a(u))$ if $r = \mathrm{Z}_a s$, where

$$\begin{aligned} \mathrm{X}_a(u) &= \min\{j \in \{1, \ldots, |u|\} \mid u[j] = a\}, \\ \mathrm{Y}_a(u) &= \max\{j \in \{1, \ldots, |u|\} \mid u[j] = a\}. \end{aligned}$$

This means that we can think of the first modality as starting in front of the word (if it is an X-modality) or behind the word (if it is a Y-modality).

If a ranker r is defined on u, then $r(u)$ identifies a unique position. For example, for the ranker $r = \mathrm{X}_a \mathrm{Y}_b \mathrm{Y}_c$ we have $r(cbca) = 1$, but r is undefined on the word $acbca$ since there is no b before the first a. For an arbitrary set of rankers R and a word u let $R(u)$ consist of all rankers in R which are defined on the given word u:

$$R(u) = \{r \in R \mid r(u) \text{ is defined}\}.$$

The *length* $|r|$ of a ranker r is its number of modalities; *i.e.*, its length as a word over the alphabet $\{\mathrm{X}_a,\, \mathrm{Y}_a \mid a \in A\}$.

Whether a ranker is defined on a word or not, gives the first link to formal languages: Languages of the form $L(r) = \{u \in A^* \mid r(u) \text{ is defined}\}$ are probably the most natural means to describe languages by rankers. For example, a language is $\mathrm{FO}^2[<]$-definable if and only if it is a Boolean combination of languages of the form $L(r)$; *cf.* [STV02] where rankers appear under the name *turtle program*. Coming back to the above example, $L(\mathrm{X}_a \mathrm{Y}_b \mathrm{Y}_c) = \{b,c\}^* c \{b\}^* b \{c\}^* a \{a,b,c\}^*$ over the alphabet $\{a,b,c\}$. Such languages are too coarse to capture the alternation hierarchy within $\mathrm{FO}^2[<]$, however. Section 11.3 gives a proper ranker description of all its levels.

Look-around rankers. With the successor predicate available, formulae may peek at the surrounding positions. This fact is reflected by equipping rankers with a look-around, formalized by a triple $w = (p, a, q)$ with $p, q \in A^*$ and $a \in A$. This leads to modalities X_w and Y_w that scan for the next/previous a-position with p directly to the left and q directly to the right. Another distinction from rankers without look-around is that we introduce modalities $\mathrm{X\!\!X}_w$ and $\mathrm{Y\!\!Y}_w$, which skip the neighboring position in the scanning process. This is necessary because $\mathrm{FO}^2[<, \mathrm{suc}]$ can express, for example, $x \geqslant y + 2$ on atomic level, namely by the formula $(y < x) \wedge \neg\,\mathrm{suc}(y, x)$. The following formalizes these extensions.

A *look-around ranker* is a non-empty word over the (infinite) alphabet

$$\{\mathrm{X}_w,\ \mathrm{Y}_w,\ \mathrm{X\!\!X}_w,\ \mathrm{Y\!\!Y}_w \mid w \in A^* \times A \times A^*\}.$$

The symbols of the form X_w and $\mathrm{X\!\!X}_w$ are called X-*modalities*, whereas Y_w and $\mathrm{Y\!\!Y}_w$ are Y-*modalities*. As before, we often write Z_w to stand for either an X-modality or a Y-modality. When no misunderstandings may arise, we also drop the attribute "look-around" and just write "ranker" for short.

We need some more notation to define how look-around rankers are evaluated. A *context* is a triple $(p, a, q) \in A^* \times A \times A^*$. An n-*context* is a context with $|p|, |q| \leqslant n$. Contexts are typically denoted by w in the following. A position i of a word u is a w-*position* for a context $w = (p, a, q)$ whenever $u[i - |p|; i + |q|] = paq$. Remember that $u[i; j]$ denotes the factor of u induced by all positions of u in the closed interval $[i; j]$. In other words, i is a w-position if it is an a-position, the prefix up to position $i - 1$ has p as suffix, and the suffix starting with position $i + 1$ has q as prefix.

For a word $u \in A^*$, a position i, and a context w define

$$\begin{aligned}
\mathrm{X}_w(u, i) &= \min\{j \in \{i+1, \ldots, |u|\} \mid j \text{ is a } w\text{-position of } u\},\\
\mathrm{X\!\!X}_w(u, i) &= \min\{j \in \{i+2, \ldots, |u|\} \mid j \text{ is a } w\text{-position of } u\},\\
\mathrm{Y}_w(u, i) &= \max\{j \in \{1, \ldots, i-1\} \mid j \text{ is a } w\text{-position of } u\},\\
\mathrm{Y\!\!Y}_w(u, i) &= \max\{j \in \{1, \ldots, i-2\} \mid j \text{ is a } w\text{-position of } u\}.
\end{aligned}$$

Remember that the minimum and the maximum of the empty set is undefined. In particular, $\mathrm{X}_w(u, i)$ and $\mathrm{Y}_w(u, i)$ are undefined if there is no w-position greater (respectively, smaller) than i. A way to memorize $\mathrm{X\!\!X}_w$ is as the contraction of $\mathrm{X}\,\mathrm{X}_w$, where X is a modality for the "neXt" position. Similarly, $\mathrm{Y\!\!Y}_w$ is a contraction of $\mathrm{Y}\,\mathrm{Y}_w$, where Y stands for the "Yesterday" position.

As usual, the *evaluation* of a composite look-around ranker is from left to right: If $r = \mathrm{Z}_w s$ with $\mathrm{Z} \in \{\mathrm{X}, \mathrm{X\!\!X}, \mathrm{Y}, \mathrm{Y\!\!Y}\}$, then $r(u, i) = s(u, \mathrm{Z}_w(u, i))$ and $r(u) = s(u, \mathrm{Z}_w(u))$, where

$$\begin{aligned}
\mathrm{X}_w(u) = \mathrm{X\!\!X}_w(u) &= \min\{j \in \{1, \ldots, |u|\} \mid j \text{ is a } w\text{-position of } u\},\\
\mathrm{Y}_w(u) = \mathrm{Y\!\!Y}_w(u) &= \max\{j \in \{1, \ldots, |u|\} \mid j \text{ is a } w\text{-position of } u\}.
\end{aligned}$$

Similarly to rankers without look-around, we can think of X-modalities as starting in front of the word and of Y-modalities as starting behind the word; and a look-around ranker r is also either undefined on a word u or $r(u)$ identifies a unique position of u.

We extend the definitions of rankers and let $\widetilde{R}(u) = \{r \in \widetilde{R} \mid r(u) \text{ is defined}\}$ for an arbitrary set of look-around rankers $\widetilde{R}$ and a word u, and we let the *length* $|r|$ be the number of modalities of r.

Rankers without look-around can canonically be interpreted as look-around rankers by identifying a letter a with the 0-context $(\varepsilon, a, \varepsilon)$. The modalities Z_a and $\mathrm{Z}_{(\varepsilon,a,\varepsilon)}$ for $\mathrm{Z} \in \{\mathrm{X}, \mathrm{Y}\}$ have precisely the same semantics.

A note on notation: For rankers with look-around we often use curly denominations throughout this chapter. For example, $\widetilde{R}$ generally denotes a set of look-around rankers, and $\approx$ as well as $\preccurlyeq$ generally refer to relations whose definition is based on look-around rankers. A way to memorize this is to think of the tilde as a rotated and scaled "S" for *successor*. Our look-around rankers are called *successor rankers* by Weis and Immerman, with the difference that they did not include the stricter modalities XX_w and YY_w, *cf.* [WI09]. As we shall see, these modalities are truly essential to the ranker description of quantifier alternation within $\mathrm{FO}^2[<, \mathrm{suc}]$, and not merely spurious relicts to simplify the proof.

11.2. Temporal Logic

In contrast to predicate logic considered so far, temporal logic does not have variables and quantifiers to speak about positions. Instead there is one distinguished current position on the word, and *modalities* can be used to modify this position. The name "temporal logic" derives from the intuition of thinking of the word as a series of events (*i.e.*, letters) over discrete time (*i.e.*, positions). In this intuition, the current position is the present, all positions to the right comprise the future, and those to the left represent the past. This motivates names for modalities like F for *Future* and P for *Past*, or X for *neXt* and Y for *Yesterday*.

The syntax of *temporal logic* formulae is given by the following inductive rules (where $a \in A$, and φ and ψ are already temporal formulae):

$$\begin{aligned}&\top \mid \bot \mid a \mid \min \mid \max \mid \neg\varphi \mid \varphi \vee \psi \mid \varphi \wedge \psi \mid\\ &\mathrm{X}\varphi \mid \mathrm{Y}\varphi \mid \mathrm{F}\varphi \mid \mathrm{P}\varphi \mid \mathrm{XF}\varphi \mid \mathrm{YP}\varphi \mid \mathrm{XXF}\varphi \mid \mathrm{YYP}\varphi\end{aligned}$$

Let TL be the set of all temporal logic formulae. The atomic formulae min and max are included for technical reasons; they will be needed mainly in the next chapter.

To define truth $u, i \models \varphi$ of a TL-formula φ over a word u with a current position i, we identify φ with a FO^2-formula $\varphi(x)$ in one free variable as follows: Atomic formulae $\top$ and $\bot$ as well as Boolean combinations are canonical. Let $a(x)$ be $\lambda(x) = a$, and let $(\min)(x) := \min(x)$ and $(\max)(x) := \max(x)$. For the modalities X, Y, F, and P let

$$\begin{aligned}(\mathrm{X}\varphi)(x) &:= \exists y\colon \mathrm{suc}(x,y) \wedge \varphi(y),\\ (\mathrm{Y}\varphi)(x) &:= \exists y\colon \mathrm{suc}(y,x) \wedge \varphi(y),\\ (\mathrm{F}\varphi)(x) &:= \exists y\colon x \leqslant y \wedge \varphi(y),\\ (\mathrm{P}\varphi)(x) &:= \exists y\colon y \leqslant x \wedge \varphi(y),\end{aligned}$$

where $\varphi(y)$ is obtained from $\varphi(x)$ as usual by interchanging x and y. For example, $u, i \models \mathrm{X}\varphi$ if and only if $i+1$ is a position of u (*i.e.*, $i + 1 \leqslant |u|$) and $u, i+1 \models \varphi$; or $u, i \models \mathrm{F}\varphi$ if and only if there exists a position $j \geqslant i$ of u such that $u, j \models \varphi$.

The remaining modalities are composites of the above modalities that are given by the equivalences $\mathrm{XF}\varphi \equiv \mathrm{X}(\mathrm{F}\varphi)$, $\mathrm{XXF}\varphi \equiv \mathrm{X}(\mathrm{X}(\mathrm{F}\varphi))$, $\mathrm{YP}\varphi \equiv \mathrm{Y}(\mathrm{P}\varphi)$, and $\mathrm{YYP}\varphi \equiv \mathrm{Y}(\mathrm{Y}(\mathrm{P}\varphi))$, which are valid on words with a current position. So XF, for example, is really just a contraction of X and F. These composite modalities are nonetheless included in the syntax because they count as only one modality. This will be important in order to capture the quantifier depth of first-order formulae by temporal logic fragments.

In order to define languages by temporal formulae, we need to define truth $u \models \varphi$ over words without a current position. The intuition is that we start at an unlabeled position far outside u. This leads to the following definition: We let $u \models \varphi$, and for each of the formulae $\varphi \in \{a, \bot, \mathrm{X}\psi, \mathrm{Y}\psi, \min, \max\}$, where $a \in A$ and $\psi \in \mathrm{TL}$, we let $u \not\models \varphi$. Boolean combinations are as usual, and for $\varphi \in \{\mathrm{F}\psi, \mathrm{XF}\psi, \mathrm{XXF}\psi, \mathrm{P}\psi, \mathrm{YP}\psi, \mathrm{YYP}\psi\}$ let $u \models \varphi$ if and only if $u \models \exists x\ \psi(x)$, where $\psi(x)$ is the two-variable formula from above. As usual, for $\varphi \in \mathrm{TL}$ let the *language defined* by φ be $L(\varphi) = \{u \in A^* \mid u \models \varphi\}$.

Note that without a current position, the modalities XF, XXF, YP, and YYP are *not* the contractions of X, F, Y, and P from above; these would not be of much use, as $\mathrm{X}(\mathrm{F}\psi)$ is always false, for example. Also, the modalities XF, XXF, YP, and YYP are equivalent on the outermost level. We may in particular assume that, when interpreting over words, formulae begin with a future modality, and never start with an X-modality or a Y-modality. Readers familiar with temporal logic might wonder about these semantics for X and Y as outermost modalities, as they are commonly defined to move to the first and last position, respectively. Although not common, this will be crucial for our later characterizations.

In order to characterize the quantifier alternation hierarchy within FO^2, the following fragments of TL-formulae will be especially important in this thesis:

- TL_m^+ consists of all formulae with negation nesting depth of at most m.
- $\mathrm{TL}_{m,n}^+$, the fragment of TL_m^+ with modality nesting depth at most n.
- TL_m and $\mathrm{TL}_{m,n}$, the Boolean closures of TL_m^+ and $\mathrm{TL}_{m,n}^+$, respectively.

More specifically, let $\mathrm{TL}_{m,n}^+$ be defined as follows: If $m = 0$ or $n = 0$, then let $\mathrm{TL}_{m,n}^+$ contain all Boolean combinations of atomic formulae a, $\top$, $\bot$, min, and max. For $m, n \geqslant 1$ inductively let $\mathrm{TL}_{m,n}^+$ be the smallest set of TL-formulae such that

- $\{\varphi, \neg\varphi\} \subseteq \mathrm{TL}_{m,n}^+$ for all $\varphi \in \mathrm{TL}_{m-1,n}^+$,
- $\{\varphi \vee \psi, \varphi \wedge \psi\} \subseteq \mathrm{TL}_{m,n}^+$ for all $\varphi, \psi \in \mathrm{TL}_{m,n}^+$, and
- $\{\varphi, \mathrm{X}\varphi, \mathrm{Y}\varphi, \mathrm{F}\varphi, \mathrm{P}\varphi, \mathrm{XF}\varphi, \mathrm{YP}\varphi, \mathrm{XXF}\varphi, \mathrm{YYP}\varphi\} \subseteq \mathrm{TL}_{m,n}^+$ for all $\varphi \in \mathrm{TL}_{m,n-1}^+$.

With this let $\mathrm{TL}_m^+ = \bigcup_{n \geqslant 0} \mathrm{TL}_{m,n}^+$. Note that negations over atomic formulae can be avoided by using De Morgan's laws to move negations towards atoms and then eliminating negations by replacing $\neg a$ by the disjunction $\bigvee_{b \in A \setminus \{a\}} b$.

For $\mathcal{F} \subseteq \mathrm{TL}$ and $\mathcal{P} \subseteq \{\mathrm{X}, \mathrm{Y}, \mathrm{F}, \mathrm{P}, \mathrm{XF}, \mathrm{YP}, \mathrm{XXF}, \mathrm{YYP}, \min, \max\}$ let $\mathcal{F}[\mathcal{P}]$ be the set of formulae in $\mathcal{F}$ that only use modalities in $\mathcal{P}$.

Let us conclude this section on temporal logic with some concrete temporal logic fragments and, partly anticipating some of the results to come, their connection to two-variable first-order logic:

- $\mathrm{TL}[\mathrm{XF}, \mathrm{YP}]$ is known to be expressively complete for $\mathrm{FO}^2[<]$; *cf.* [EVW02].
- $\mathrm{TL}[\mathrm{X}, \mathrm{XF}, \mathrm{Y}, \mathrm{YP}]$ is known to be expressively complete for $\mathrm{FO}^2[<, \mathrm{suc}]$; *cf.* [EVW02].
- $\mathrm{TL}_{m,n}^+[\mathrm{XF}, \mathrm{YP}]$ and $\mathrm{TL}_{m,n}[\mathrm{XF}, \mathrm{YP}]$ are fragments that are expressively complete for $\Sigma_{m,n}^2[<]$ and $\mathrm{FO}_{m,n}^2[<]$, respectively; *cf.* Theorem 11.3 and Corollary 11.6 below.
- $\mathrm{TL}_{m,n}^+[\mathrm{X}, \mathrm{XXF}, \mathrm{Y}, \mathrm{YYP}]$ and $\mathrm{TL}_{m,n}[\mathrm{X}, \mathrm{XXF}, \mathrm{Y}, \mathrm{YYP}]$ are expressively complete for $\Sigma_{m,n}^2[<, \mathrm{suc}]$ and $\mathrm{FO}_{m,n}^2[<, \mathrm{suc}]$, respectively; *cf.* Theorem 11.16 and Corollary 11.19.

From the very definition of the semantics of TL-formulae it is clear that, in all these instances, a language definable in the TL-fragment is definable in the respective FO^2-fragment. So it suffices to construct a TL-formula for a FO^2-formula that is equivalent and obeys the depth restrictions.

The construction employed in the paper of Etessami, Vardi, and Wilke does not respect negations and, therefore, also not quantifier alternations. Moreover, it does not preserve quantifier depth for $\mathrm{FO}^2[<, \mathrm{suc}]$, basically because modalities XF and YP are used instead of XXF and YYP. The latter can of course be expressed in terms of the former by using X and Y; this is the reason why [EVW02, Theorem 1] has a depth blowup by a factor of two.

11.3. Rankers for Quantifier Alternation in $\mathrm{FO}^2[<]$

We turn to two-variable first-order logic $\mathrm{FO}^2[<]$ with linear order as the only numerical predicate and characterize the expressive power of the half level $\Sigma^2_{m,n}[<]$ as well as of the full level $\mathrm{FO}^2_{m,n}[<]$ of its alternation hierarchy.

The following definition introduces the sets of rankers $R_{m,n}$, $R^{\mathrm{X}}_{m,n}$, and $R^{\mathrm{Y}}_{m,n}$, which are used frequently throughout this section.

Definition 11.1
Let $m, n \geqslant 0$. If $m = 0$ or if $n = 0$, then let $R_{m,n} = R^{\mathrm{X}}_{m,n} = R^{\mathrm{Y}}_{m,n} = \emptyset$. If $m, n \geqslant 1$, let $R^{\mathrm{X}}_{m,n}$ and $R^{\mathrm{Y}}_{m,n}$ be the smallest sets of rankers such that

$$r \in \{\varepsilon\} \cup R^{\mathrm{X}}_{m,n-1} \cup R^{\mathrm{Y}}_{m-1,n-1} \text{ implies } r\mathrm{X}_a \in R^{\mathrm{X}}_{m,n},$$
$$r \in \{\varepsilon\} \cup R^{\mathrm{Y}}_{m,n-1} \cup R^{\mathrm{X}}_{m-1,n-1} \text{ implies } r\mathrm{Y}_a \in R^{\mathrm{Y}}_{m,n}$$

for all $a \in A$. Let $R_{m,n} = R^{\mathrm{X}}_{m,n} \cup R^{\mathrm{Y}}_{m,n}$.

This means that $R_{m,n}$ contains those rankers with at most n modalities (*i.e.*, with length n) and with at most $m-1$ alternations between different modalities (*i.e.*, with at most m alternating blocks of modalities). We also say m is the *alternation parameter* and n is the *depth parameter* of a ranker in $R_{m,n}$. The sets $R^{\mathrm{X}}_{m,n}$ and $R^{\mathrm{Y}}_{m,n}$ are those subsets containing rankers ending on an X-modality or a Y-modality, respectively.

The following preorder on words is the central combinatorial description of the expressive power of $\Sigma^2_{m,n}[<]$ over words.

Definition 11.2
Let $u, v \in A^*$ and $m, n \geqslant 0$. Let $u \leqslant^R_{m,n} v$ if either $m = 0$ or $n = 0$, or if $v \leqslant^R_{m-1,n} u$ and all of the following hold:

1. *$R_{m,n}(v) \subseteq R_{m,n}(u)$,*
2. *$r(v) < s(v)$ implies $r(u) < s(u)$ and $r(v) \leqslant s(v)$ implies $r(u) \leqslant s(u)$ for all $r \in R^{\mathrm{X}}_{m,n}(v)$ and all $s \in R^{\mathrm{X}}_{m-1,n-1}(v)$,*
3. *$r(v) > s(v)$ implies $r(u) > s(u)$ and $r(v) \geqslant s(v)$ implies $r(u) \geqslant s(u)$ for all $r \in R^{\mathrm{Y}}_{m,n}(v)$ and all $s \in R^{\mathrm{Y}}_{m-1,n-1}(v)$,*
4. *$r(v) < s(v)$ implies $r(u) < s(u)$ and $r(v) \leqslant s(v)$ implies $r(u) \leqslant s(u)$ for all $r \in R^{\mathrm{X}}_{m,n}(v)$ and all $s \in R^{\mathrm{Y}}_{m,n}(v)$ with $|r| + |s| < 2n$.*

The following is the main result of this section, characterizing $\Sigma^2_{m,n}[<]$ in terms of the above preorder and in terms of an expressively complete fragment of temporal logic.

Theorem 11.3
Let $L \subseteq A^*$, and let $m, n \geqslant 1$. The following are equivalent:

1. *L is definable in $\Sigma^2_{m,n}[<]$.*
2. *L is definable in $\mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$.*
3. *L is an $\leqslant^R_{m,n}$-order ideal.*

Remember that L is a $\leqslant^R_{m,n}$-order ideal if $u \leqslant^R_{m,n} v$ and $v \in L$ always implies $u \in L$. A proof of the theorem will be given shortly. Let us first ponder some implications.

We rephrase Theorem 11.3. For this we need to introduce two more preorders on A^*. Let $u \leqslant^{\mathrm{FO}^2}_{m,n} v$ for $u, v \in A^*$ if $v \in L(\varphi)$ implies $u \in L(\varphi)$ for all sentences $\varphi \in \Sigma^2_{m,n}[<]$. Let $u \leqslant^{\mathrm{TL}}_{m,n} v$ for $u, v \in A^*$ if $v \in L(\varphi)$ implies $u \in L(\varphi)$ for all $\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$.

Corollary 11.4
Let $m, n \geqslant 1$, and let $u, v \in A^*$. The following are equivalent:

1. *$u \leqslant^{\mathrm{FO}^2}_{m,n} v$.*
2. *$u \leqslant^{\mathrm{TL}}_{m,n} v$.*
3. *$u \leqslant^R_{m,n} v$.*

Proof. (1) $\Leftrightarrow$ (2): This follows directly from Theorem 11.3.

(1) $\Rightarrow$ (3): For a word v consider $L = \{u \in A^* \mid u \leqslant^R_{m,n} v\}$. This is an $\leqslant^R_{m,n}$-order ideal by definition, and as such definable in $\Sigma^2_{m,n}[<]$ by Theorem 11.3. Since $v \in L$, the assumption yields $u \in L$; in other words $u \leqslant^R_{m,n} v$.

(3) $\Rightarrow$ (1): Let ψ_v be the conjunction of all $\varphi \in \Sigma^2_{m,n}[<]$ with $v \in L(\varphi)$. We have $\psi_v \in \Sigma^2_{m,n}[<]$. Theorem 11.3 shows that $L(\psi_v)$ is an $\leqslant^R_{m,n}$-order ideal. Since v is in $L(\psi_v)$, so is u. Therefore, $u \in L(\varphi)$ for all $\varphi \in \Sigma^2_{m,n}[<]$ with $v \in L(\varphi)$. $\square$

Theorem 11.3 and Corollary 11.4 are original contributions of this thesis and were not previously known in the literature.

Let us turn to the full levels of the alternation hierarchy. For integers i and j the *order type* $\mathrm{ord}(i, j) \in \{<, 0, >\}$ specifies the relative ordering of i and j; i.e.,

$$\mathrm{ord}(i,j) = \begin{cases} < & \text{if } i < j, \\ 0 & \text{if } i = j, \\ > & \text{if } i > j. \end{cases}$$

The following equivalence relation captures the combinatorial structure of $\mathrm{FO}^2_{m,n}[<]$.

Definition 11.5
Let $u, v \in A^*$, and let $m, n \geqslant 0$. Let $u \equiv^R_{m,n} v$ if either $m = 0$ or $n = 0$ or all of the following hold:

1. *$R_{m,n}(u) = R_{m,n}(v)$,*
2. *$\mathrm{ord}(r(u), s(u)) = \mathrm{ord}(r(v), s(v))$ for all $(r, s) \in R^{\mathrm{X}}_{m,n}(u) \times R^{\mathrm{X}}_{m-1,n-1}(u)$,*
3. *$\mathrm{ord}(r(u), s(u)) = \mathrm{ord}(r(v), s(v))$ for all $(r, s) \in R^{\mathrm{Y}}_{m,n}(u) \times R^{\mathrm{Y}}_{m-1,n-1}(u)$,*
4. *$\mathrm{ord}(r(u), s(u)) = \mathrm{ord}(r(v), s(v))$ for all $(r, s) \in R^{\mathrm{X}}_{m,n}(u) \times R^{\mathrm{Y}}_{m,n}(u)$ with $|r| + |s| < 2n$.*

An elementary verification shows that $u \equiv^R_{m,n} v$ if and only if $u \leqslant^R_{m,n} v$ and $v \leqslant^R_{m,n} u$. The following corollaries to Theorem 11.3 describe the structure of the full level $\mathrm{FO}^2_{m,n}[<]$. They are direct corollaries to Theorem 11.3 and do not require further proof.

Corollary 11.6
Let $L \subseteq A^*$, and let $m, n \geqslant 1$. The following are equivalent:

1. *L is definable in $\mathrm{FO}^2_{m,n}[<]$.*
2. *L is definable in $\mathrm{TL}_{m,n}[\mathrm{XF}, \mathrm{YP}]$.*
3. *L is a union of $\equiv^R_{m,n}$-classes.* □

Corollary 11.7
Let $m, n \geqslant 1$, and let $u, v \in A^*$. The following are equivalent:

1. *$v \in L(\varphi)$ if and only if $u \in L(\varphi)$ for all $\varphi \in \mathrm{FO}^2_{m,n}[<]$.*
2. *$v \in L(\varphi)$ if and only if $u \in L(\varphi)$ for all $\varphi \in \mathrm{TL}_{m,n}[\mathrm{XF}, \mathrm{YP}]$.*
3. *$u \equiv^R_{m,n} v$.* □

The equivalence (1) $\Leftrightarrow$ (3) in Corollary 11.7 is the ranker characterization of $\mathrm{FO}^2_{m,n}[<]$ due to Weis and Immerman [WI09, Theorem 4.5]. The equivalence of $\mathrm{FO}^2_{m,n}[<]$ and $\mathrm{TL}_{m,n}[\mathrm{XF}, \mathrm{YP}]$ is new. A coarser result in this direction is that $\mathrm{FO}^2[<]$ and $\mathrm{TL}[\mathrm{XF}, \mathrm{YP}]$ have the same expressiveness, shown by Etessami, Vardi, and Wilke [EVW02, Theorem 2]. Their construction did not preserve negations, however.

The remainder of this section is dedicated to the proof of Theorem 11.3, starting with the implication from first-order logic to temporal logic.

Lemma 11.8
Let $m, n \geqslant 0$. Every language definable in $\Sigma^2_{m,n}[<]$ is definable in $\mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$.

Proof. The proof idea is that, of the two variables of the FO^2-formula, the interpretation of one variable is given by the current position of the TL-formula, while all necessary information about the other variable is stored syntactically. Consistency of this information is ensured at quantification time by a suitable decomposition of the quantifier with respect to the order between the two variables. For this to work we need the requantification-free normal form of Chapter 10.

The construction is by induction on the structure of the formula. For the inductive step, we have to handle free variables. Let $\varphi(x, y) \in \Sigma^2_{m,n}[<]$ be in requantification-free normal form such that all outermost quantifiers bind the variable y. We show that for $\tau \in \{<, 0, >\}$ and $a \in A$ there exists $\langle\varphi\rangle_{\tau,a}(x)$ in $\mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$ such that $u, p, q \models \varphi(x, y)$ if and only if $u, p \models \langle\varphi\rangle_{\tau,a}(x)$ for all $u \in A^*$ and all positions p and q of u with $u[q] = a$ and $\mathrm{ord}(p, q) = \tau$.

This means that the current position of the TL-formula corresponds to x, and the parameters a and τ provide the label of y and the order of y relative to x. Given the consistency of these parameters, the formula $\langle\varphi\rangle_{\tau,a}(x)$ is equivalent to φ. The construction is by induction on the structure of the formula. In the inductive procedure we shall also use the dual version with interchanged roles of x and y; *i.e.*, the free variable corresponds to y and the parameters determine all necessary information about x.

Let $\langle \psi \rangle_{\tau,a}(x) := \psi$ for $\psi \in \{\top, \bot\}$, let $\langle \lambda(x) = b \rangle_{\tau,a}(x) := b$, and for the other atomic formulae let

$$\langle \lambda(y) = b \rangle_{\tau,a}(x) := \begin{cases} \top & \text{if } a = b, \\ \bot & \text{otherwise,} \end{cases}$$

$$\langle x < y \rangle_{\tau,a}(x) := \begin{cases} \top & \text{if } \tau \text{ is } <, \\ \bot & \text{otherwise,} \end{cases}$$

$$\langle y < x \rangle_{\tau,a}(x) := \begin{cases} \top & \text{if } \tau \text{ is } >, \\ \bot & \text{otherwise.} \end{cases}$$

For negation set $\langle \neg\varphi \rangle_{\tau,a}(x) := \neg\langle \varphi \rangle_{\tau,a}(x)$. Conjunction and disjunction are given by $\langle \varphi \wedge \psi \rangle_{\tau,a}(x) := \langle \varphi \rangle_{\tau,a}(x) \wedge \langle \psi \rangle_{\tau,a}(x)$ and $\langle \varphi \vee \psi \rangle_{\tau,a}(x) := \langle \varphi \rangle_{\tau,a}(x) \vee \langle \psi \rangle_{\tau,a}(x)$. For existential quantification let

$$\langle \exists y\, \varphi \rangle_{\tau,a}(x) := \bigvee_{b \in A} b \wedge \big(\mathrm{YP} \langle \varphi \rangle_{<,b}(y) \vee \langle \varphi \rangle_{0,b}(y) \vee \mathrm{XF} \langle \varphi \rangle_{>,b}(y) \big).$$

Here, the formulae $\langle \varphi \rangle_{\tau',b}(y)$ are the dual formulae obtained by induction with the roles of x and y interchanged. Note that we do not have to handle quantification over x by requantification-freeness.

We can now proof the lemma. Let $\varphi \in \Sigma^2_{m,n}[<]$ be a sentence. By Proposition 10.3, we assume without restriction that φ is in requantification-free normal form. Renaming variables in subformulae we may further assume all outermost quantifiers to bind the same variable. This allows to use the above construction, and we inductively define an equivalent $\mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$-formula φ'. Boolean connectives are straightforward: $(\varphi \wedge \psi)' := \varphi' \wedge \psi'$, $(\varphi \vee \psi)' := \varphi' \vee \psi'$, and $(\neg\varphi)' := \neg\varphi'$. For quantification we let $(\exists y \colon \varphi)' := \mathrm{XF}(\langle \varphi \rangle_{\tau,a}(x))$, where $\tau \in \{<, 0, >\}$ and $a \in A$ are arbitrary. This shows that for every sentence $\varphi \in \Sigma^2_{m,n}[<]$ there exists $\varphi' \in \mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$ with $L(\varphi) = L(\varphi')$. $\square$

We come to the implication from temporal logic to the ranker condition in Theorem 11.3. To account for the free variable of the temporal formula, we extend the preorder $\leqslant^R_{m,n}$ over words to words with a designated position. To avoid tedious case distinctions, we also include a pseudo-position $-\infty$, which models the situation at the beginning where we imagine to start the evaluation of the temporal formula (far) in front of the model. The usual arithmetic involving $\pm\infty$ applies; *i.e.*, $i \pm \infty = \pm\infty$ and $-\infty < i < \infty$ for all integers i.

Let $u, v \in A^*$, let $m, n \geqslant 0$ be integers, let x be a position of u or $-\infty$, and let x' be a position of v or $-\infty$. Put $(u, x') \leqslant^R_{m,n} (v, x)$ if $u \leqslant^R_{m,n} v$, and $v[x] = u[x']$, and if further either $m = 0$, or $n = 0$, or all of the following hold for $d \in \{0, 1\}$:

1. $s(v) \leqslant x - d$ implies $s(u) \leqslant x' - d$ for every $s \in R^{\mathrm{X}}_{m,n}(v) \cup R^{\mathrm{Y}}_{m-1,n}(v)$,
2. $s(v) \geqslant x + d$ implies $s(u) \geqslant x' + d$ for every $s \in R^{\mathrm{Y}}_{m,n}(v) \cup R^{\mathrm{X}}_{m-1,n}(v)$.

The intuition of $(u, x') \leqslant^R_{m,n} (v, x)$ is that the position x' in u is at least as hard to distinguish for $\mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$-formulae than the position x in v. Formally, this means that $v, x \models \varphi$ implies $u, x' \models \varphi$ for all $\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$. The next lemma is central to proving this intuition.

Lemma 11.9
Let $u, v \in A^*$, let $m, n \geqslant 1$, let x be a position of v or $x = -\infty$, and let x' be a position of u or $x' = -\infty$. If $(u, x') \leqslant^R_{m,n} (v, x)$, then for every position y of v there exists a position y' of u with $\mathrm{ord}(x, y) = \mathrm{ord}(x', y')$ and $(u, y') \leqslant^R_{m,n-1} (v, y)$.

Proof. For $y = x$ take $y' = x'$. By left-right symmetry it suffices to consider the case $y > x$. Let $a = v[y]$, and let

$$R_{\mathit{right}} = \{r \mid r\mathrm{Y}_a \in R^{\mathrm{Y}}_{m,n}(v) \text{ and } r\mathrm{Y}_a(v) \geqslant y\},$$
$$R_{\mathit{left}} = \{r \mid r\mathrm{X}_a \in R^{\mathrm{X}}_{m-1,n-1}(v) \text{ and } r\mathrm{X}_a(v) \geqslant y\}.$$

Note that $R_{\mathit{right}}\,\mathrm{Y}_a \cup R_{\mathit{left}}\,\mathrm{X}_a \subseteq R_{m,n}(v) \subseteq R_{m,n}(u)$. Let $r \in R_{\mathit{right}}\,\mathrm{Y}_a \cup R_{\mathit{left}}\,\mathrm{X}_a$ be such that $r(u)$ is minimal. We claim that we can choose $y' = r(u)$. By $r(v) \geqslant y > x$ and condition (2) in the definition of $(u, x') \leqslant^R_{m,n} (v, x)$ we conclude $y' = r(u) > x'$.

Next, we show $(u, y') \leqslant^R_{m,n-1} (v, y)$. The alphabetic condition is clear.

Condition (1): Let $s \in R^{\mathrm{X}}_{m,n-1}(v) \cup R^{\mathrm{Y}}_{m-1,n-1}(v)$ with $s(v) < y$. In particular, we have $s(v) < r(v)$. The assumption implies $u \leqslant^R_{m,n} v$ which yields $s(u) < r(u) = y'$. (If $s \in R^{\mathrm{X}}_{m,n-1}$ and $r \in R^{\mathrm{Y}}_{m,n}$, then condition (4) of Definition 11.2 is applied. If $s \in R^{\mathrm{X}}_{m,n-1}$ and $r \in R^{\mathrm{X}}_{m-1,n-1}$, then condition (2) is applied. If $s \in R^{\mathrm{Y}}_{m-1,n-1}$ and $r \in R^{\mathrm{Y}}_{m,n}$, then condition (3) is applied. If $s \in R^{\mathrm{Y}}_{m-1,n-1}$ and $r \in R^{\mathrm{X}}_{m-1,n-1}$, then condition (4) for $v \leqslant^R_{m-1,n} u$ is applied.) The case $s(v) \leqslant y$ is similar.

Condition (2): Let $s \in R^{\mathrm{Y}}_{m,n-1}(v) \cup R^{\mathrm{X}}_{m-1,n-1}(v)$. Suppose $s(v) > y$. Since $s \in R_{\mathit{right}}$, we see that $s(u) > s\mathrm{Y}_a(u) \geqslant r(u) = y'$ by choice of r. Suppose $s(v) \geqslant y$. By the previous case we may assume $s(v) = y$; in particular, s ends with one of the modalities X_a or Y_a. First suppose $s = s'\mathrm{X}_a$. Then $s' \in R_{\mathit{left}}$ and thus $s(u) = s'\mathrm{X}_a(u) \geqslant r(u) = y'$ by choice of r. Finally suppose $s = s'\mathrm{Y}_a$. Then $s' \in R_{\mathit{right}}$ and thus $s(u) = s'\mathrm{Y}_a(u) \geqslant r(u) = y'$ by choice of r. □

The next lemma gives a way to handle negations. It shows that the roles of u and v can be interchanged by investing one negation.

Lemma 11.10
Let $u, v \in A^*$, and let $m, n \geqslant 1$. If $(u, x') \leqslant^R_{m,n} (v, x)$, then $(v, x) \leqslant^R_{m-1,n} (u, x')$.

Proof. We may assume $m \geqslant 2$ and $n \geqslant 1$, since the claim is trivial otherwise. The alphabetic condition is clear and $v \leqslant^R_{m-1,n} u$ is a requirement of $u \leqslant^R_{m,n} v$.

To show condition (1), let $s \in R^{\mathrm{X}}_{m-1,n}(u) \cup R^{\mathrm{Y}}_{m-2,n}(u)$, and first suppose $s(u) < x'$. We have to show $s(v) < x$. Suppose $s(v) \geqslant x$ for the sake of contradiction. By condition (2) of $(u, x') \leqslant^R_{m,n} (v, x)$ we obtain $s(u) \geqslant x'$, contradicting $s(u) < x'$. For $s(u) \leqslant x'$ a similar reasoning applies and condition (2) is left-right symmetric. □

The following lemma shows that every $\mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$-definable language is a $\leqslant^R_{m,n}$-order ideal. It combines the two preceding Lemmas 11.9 and 11.10 by a straightforward induction.

Lemma 11.11
Let $u, v \in A^*$, and let $m, n \geqslant 0$. If $u \leqslant^R_{m,n} v$, then $u \leqslant^{\mathrm{TL}}_{m,n} v$.

Proof. We show that if $(u, x') \leqslant^R_{m,n} (v, x)$, then $v, x \models \varphi$ implies $u, x' \models \varphi$ for all formulae $\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$. The proof of this claim is by structural induction. For atomic formulae this follows by the alphabetic condition in the definition of $\leqslant^R_{m,n}$. In particular we may assume $m, n \geqslant 1$.

For Boolean connectives the claim follows by induction; in the case of negation this relies on Lemma 11.10 and $\varphi \in \mathrm{TL}^+_{m-1,n}[\mathrm{XF}, \mathrm{YP}]$ whenever $\neg\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$.

Suppose $v, x \models \mathrm{XF}\varphi$. Then there exists a position $y > x$ of v such that $v, y \models \varphi$. By Lemma 11.9 there exists a position $y' > x'$ of u such that $(u, y') \leqslant^R_{m,n-1} (v, y)$. Because we have $\varphi \in \mathrm{TL}^+_{m,n-1}[\mathrm{XF}, \mathrm{YP}]$, induction yields $u, y' \models \varphi$ and finally $u, x' \models \mathrm{XF}\varphi$. The remaining case $\mathrm{YP}\varphi$ is symmetric to $\mathrm{XF}\varphi$. This concludes the proof of the claim.

Suppose now $v \models \varphi$ for some $\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XF}, \mathrm{YP}]$. Replacing outermost YP-modalities by XF yields a formula which defines the same language. We may therefore assume without loss of generality that $v \models \varphi$ if and only if $v, -\infty \models \varphi$. If $u \leqslant^R_{m,n} v$, then $(u, -\infty) \leqslant^R_{m,n} (v, -\infty)$ and the claim thus yields $u, -\infty \models \varphi$. This shows $u \leqslant^{\mathrm{TL}}_{m,n} v$. $\square$

We now come to the last implication, closing the loop back from the ranker conditions to first-order logic. As an intermediate step for this, we give two-variable formulae for order comparisons with ranker positions. These formulae show that a direction alternation in a ranker can be handled using a negation in first-order logic.

Lemma 11.12
For each $r \in R^{\mathrm{X}}_{m,n} \cup R^{\mathrm{Y}}_{m-1,n}$ there exist formulae $\langle x \geqslant r \rangle$ and $\langle x > r \rangle$ in $\Sigma^2_{m,n}[<]$ with free variable x such that for all words u with $r(u)$ defined and all positions i of u we have

$$u, i \models \langle x \geqslant r \rangle \quad \textit{if and only if} \quad i \geqslant r(u),$$
$$u, i \models \langle x > r \rangle \quad \textit{if and only if} \quad i > r(u).$$

Symmetrically, for each $r \in R^{\mathrm{Y}}_{m,n} \cup R^{\mathrm{X}}_{m-1,n}$ there exist formulae $\langle x \leqslant r \rangle$ and $\langle x < r \rangle$ in $\Sigma^2_{m,n}[<]$ with free variable x such that for all words u with $r(u)$ defined and for all positions i of u we have

$$u, i \models \langle x \leqslant r \rangle \quad \textit{if and only if} \quad i \leqslant r(u),$$
$$u, i \models \langle x < r \rangle \quad \textit{if and only if} \quad i < r(u).$$

Proof. The proof is by induction on m and n. If $m = 0$ or if $n = 0$ the claim is vacuously true, so let $m, n \geqslant 1$. It suffices to give $\langle x \geqslant r \rangle$ and $\langle x > r \rangle$ for rankers $r \in R^{\mathrm{X}}_{m,n} \cup R^{\mathrm{Y}}_{m-1,n}$. The formulae $\langle x \leqslant r \rangle$ and $\langle x < r \rangle$ for rankers in $R^{\mathrm{Y}}_{m,n} \cup R^{\mathrm{X}}_{m-1,n}$ are left-right symmetric.

For $r \in R^{\mathrm{Y}}_{m-1,n}$ we set $\langle x > r \rangle := \neg\langle x \leqslant r \rangle$ and $\langle x \geqslant r \rangle := \neg\langle x < r \rangle$, where $\langle x \leqslant r \rangle$ and $\langle x < r \rangle$ are obtained by induction on m. It remains to consider rankers $R^{\mathrm{X}}_{m,n}$. Let

$$\langle x > r\mathrm{X}_a \rangle \;:=\; \exists y < x \, \big(\lambda(y) = a \wedge \langle y > r \rangle\big),$$
$$\langle x \geqslant r\mathrm{X}_a \rangle \;:=\; \exists y \leqslant x \, \big(\lambda(y) = a \wedge \langle y > r \rangle\big).$$

Here, r is a possibly empty ranker. The formula $\langle y > r \rangle$ is to be read as $\top$ if r is empty and obtained by induction on n otherwise. As usual, the formula $\langle y > r \rangle$ with free variable y is obtained by interchanging x and y. Note that $y \leqslant x$ can be rewritten as $\neg(x < y)$. $\square$

The formulae in Lemma 11.12 yield suitable two-variable sentences to express definedness of rankers as well as order comparisons between rankers.

Lemma 11.13
Let $u, v \in A^*$, and let $m, n \geqslant 0$ be integers. If $u \leqslant^{\mathrm{FO2}}_{m,n} v$, then $u \leqslant^{R}_{m,n} v$.

Proof. The proof is by induction on m with the trivial base case $m = 0$. Suppose $u \leqslant^{\mathrm{FO2}}_{m,n} v$. In particular $v \leqslant^{\mathrm{FO2}}_{m-1,n} u$ and hence $v \leqslant^{R}_{m-1,n} u$ by induction. We show conditions (1) to (4) of Definition 11.2 one after another. The proof makes extensive use of Lemma 11.12.

Condition (1): Suppose $r \in R_{m,n}$ is defined on v but not on u. Let $r = r'\mathrm{Z}_a r''$ with $\mathrm{Z} \in \{\mathrm{X}, \mathrm{Y}\}$ for some $a \in A$ be such that $r'\mathrm{Z}_a$ is the shortest prefix of r which is undefined on u. Note that $\mathrm{alph}(v) \subseteq \mathrm{alph}(u)$ and thus r' cannot be empty. Consider the formula

$$\langle r'\mathrm{Z}_a\rangle \ := \ \exists x\colon \big(\lambda(x) = a\big) \wedge \begin{cases} \langle x > r'\rangle & \text{if } \mathrm{Z} = \mathrm{X}, \\ \langle x < r'\rangle & \text{if } \mathrm{Z} = \mathrm{Y}. \end{cases}$$

In both cases $\langle r'\mathrm{Z}_a\rangle \in \Sigma^2_{m,n}[<]$, and $\langle r'\mathrm{Z}_a\rangle$ is true on a word if and only if $r'\mathrm{Z}_a$ is defined. In particular, $r'\mathrm{Z}_a$ is defined on u, which contradicts the definition of r'. Therefore, $r(u)$ is defined. This shows $R_{m,n}(v) \subseteq R_{m,n}(u)$.

Condition (2): Consider rankers $r \in R^{\mathrm{X}}_{m,n}(v)$ and $s \in R^{\mathrm{X}}_{m-1,n-1}(v)$. Let $r = r'\mathrm{X}_a$ for $a \in A$ and let $\lesssim \in \{<, \leqslant\}$. The formula

$$\langle r \lesssim s\rangle \ := \ \exists x\colon \big(\lambda(x) = a\big) \wedge \langle x > r'\rangle \wedge \langle x \lesssim s\rangle$$

is in $\Sigma^2_{m,n}[<]$. Here we set $\langle x > r'\rangle = \top$ whenever r' is empty. Moreover $\langle r \lesssim s\rangle$ is true on a word if and only if $r \lesssim s$ on that word. Hence $r(v) < s(v)$ implies $r(u) < s(u)$, and $r(v) \leqslant s(v)$ implies $r(u) \leqslant s(u)$.

Condition (3): This is symmetric to condition (2).

Condition (4): Consider rankers $r \in R^{\mathrm{X}}_{m,n}(v)$ and $s \in R^{\mathrm{Y}}_{m,n}(v)$ with $|r| + |s| < 2n$. Let $\lesssim \in \{<, \leqslant\}$ and let $\gtrsim$ be its inverse relation. If $r = r'\mathrm{X}_a$ and $s \in R_{m,n-1}(v)$, then let

$$\langle r \lesssim s\rangle \ := \ \exists x\colon \big(\lambda(x) = a\big) \wedge \langle x > r'\rangle \wedge \langle x \lesssim s\rangle$$

and if $r \in R_{m,n-1}(v)$ and $s = s'\mathrm{Y}_a$, then let

$$\langle r \lesssim s\rangle \ := \ \exists x\colon \big(\lambda(x) = a\big) \wedge \langle x \gtrsim r\rangle \wedge \langle x < s'\rangle.$$

In both cases $\langle r \lesssim s\rangle \in \Sigma^2_{m,n}[<]$ and $\langle r \lesssim s\rangle$ is true on a word if and only if $r \lesssim s$ on that word. Therefore $r(v) < s(v)$ implies $r(u) < s(u)$, and $r(v) \leqslant s(v)$ implies $r(u) \leqslant s(u)$. $\square$

Proof of Theorem 11.3. (1) $\Rightarrow$ (2): This is Lemma 11.8.

(2) $\Rightarrow$ (3): Let L be definable in $\mathrm{TL}^{+}_{m,n}[\mathrm{XF}, \mathrm{YP}]$. Suppose $u \leqslant^{R}_{m,n} v$ and $v \in L$. Lemma 11.11 yields $u \leqslant^{\mathrm{TL}}_{m,n} v$ and consequently $u \in L$. Thus L is a $\leqslant^{R}_{m,n}$-order ideal.

(3) $\Rightarrow$ (1): Let L be a $\leqslant^{R}_{m,n}$-order ideal. Suppose $u \leqslant^{\mathrm{FO2}}_{m,n} v$ and $v \in L$. By Lemma 11.13 we see $u \leqslant^{R}_{m,n} v$ and thus $u \in L$. This shows that L is a $\leqslant^{\mathrm{FO2}}_{m,n}$-order ideal. For a word v let L_v be the intersection of all $\Sigma^2_{m,n}[<]$-definable languages containing v. This intersection is finite because, up to equivalence, there are only finitely many formulae in $\Sigma^2_{m,n}[<]$. This shows that the language L_v is $\Sigma^2_{m,n}[<]$-definable; moreover, there are only finitely many languages of the form L_v. If $v \in L$, then $v \in L_v \subseteq L$. Therefore, $L = \bigcup_{v \in L} L_v$ is $\Sigma^2_{m,n}[<]$-definable. $\square$

11.4. Look-Around Rankers for Quantifier Alternation in FO^2[$<$, suc]

This section gives combinatorial structure theorems for $\Sigma^2_{m,n}[<, \mathrm{suc}]$ and $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$. Compared to the preceding section, the successor predicate is allowed for first-order formulae, leading to rankers with a look-around as a combinatorial means. Once the correct extensions of the concepts to the case with successor are defined, the ideas behind the proofs are largely similar to the case without successor. Still, coming up with these correct extensions is non-trivial, and the proofs are even more technical and differ largely in detail.

We start by introducing sets of look-around rankers that are used extensively throughout this section.

Definition 11.14
Let $\widetilde{R}^{\mathrm{X}}_{m,n} = \widetilde{R}^{\mathrm{Y}}_{m,n} = \widetilde{R}_{m,n} = \emptyset$ if either $m = 0$ or $n = 0$. For integers $m, n \geqslant 1$ let $\widetilde{R}_{m,n} = \widetilde{R}^{\mathrm{X}}_{m,n} \cup \widetilde{R}^{\mathrm{Y}}_{m,n}$, where $\widetilde{R}^{\mathrm{X}}_{m,n}$ and $\widetilde{R}^{\mathrm{Y}}_{m,n}$ are the smallest sets of look-around rankers such that

$$r \in \{\varepsilon\} \cup \widetilde{R}^{\mathrm{X}}_{m,n-1} \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n-1} \text{ implies } \{r\mathrm{X}_w,\ r\mathbb{X}_w\} \subseteq \widetilde{R}^{\mathrm{X}}_{m,n},$$
$$r \in \{\varepsilon\} \cup \widetilde{R}^{\mathrm{Y}}_{m,n-1} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1} \text{ implies } \{r\mathrm{Y}_w,\ r\mathbb{Y}_w\} \subseteq \widetilde{R}^{\mathrm{Y}}_{m,n}$$

for all $w \in A^* \times A \times A^*$ with $w = (p, a, q)$ and $|p|, |q| \leqslant n - 1$.

A ranker is said to have an *alternation parameter* of m and a *depth parameter* of n if it is in $\widetilde{R}_{m,n}$. A look-around ranker r of length ℓ is thus in $\widetilde{R}_{m,n}$ only if $\ell \leqslant n$ and it can be written as $r = \mathrm{Z}^{(\ell)}_{(p_\ell,a_\ell,q_\ell)} \cdots \mathrm{Z}^{(1)}_{(p_1,a_1,q_1)}$ for some $|p_i|, |q_i| \leqslant n - i$, some $a_i \in A$, and some $\mathrm{Z}^{(i)} \in \{\mathrm{X}, \mathbb{X}, \mathrm{Y}, \mathbb{Y}\}$. The look-around of a modality may thus *increase*, the *deeper* inside a ranker it is. Rankers in $\widetilde{R}_{m,n}$ have at most n modalities with at most $m - 1$ alternations between X-modalities and Y-modalities.

A simple observation is that the alternation parameter as well as the depth parameter sum when concatenating rankers; i.e., $rs \in \widetilde{R}_{m+m',n+n'}$ if $r \in \widetilde{R}_{m,n}$ and $s \in \widetilde{R}_{m',n'}$. Let us consider what happens if we split a look-around ranker into a prefix and a suffix. The depth parameter of the prefix decreases with the length of the suffix which is cut off. The suffix, on the other hand, does not gain anything for the depth parameter in general. Specifically, consider a composite ranker rs with depth parameter n. The ranker r has depth parameter $|r|$. In general, all we can conclude for s is that it also has depth parameter n (and not $n - 1$ or less). In fact, consider an arbitrary ranker for s which ends on an X-modality $\mathrm{X}_{(p,a,q)}$ such that $|p| = |q| = n - 1$; even though it has length 1, its smallest depth parameter is n.

The next definition is the proper extension of $\leqslant^R_{m,n}$ from the preceding section to look-around rankers. Apart from adding look-around as well as modalities $\mathbb{X}_w$ and $\mathbb{Y}_w$, there are also additional order comparisons between pairs of rankers available.

Definition 11.15
Let $u, v \in A^*$ and $m, n \geqslant 0$. Let $u \preccurlyeq^R_{m,n} v$ if either $m = 0$ or $n = 0$, or if $v \preccurlyeq^R_{m-1,n} u$ and all of the following hold for $d \in \{-1, 0, 1, 2\}$:

1. $\widetilde{R}_{m,n}(v) \subseteq \widetilde{R}_{m,n}(u)$,
2. $r(v) \leqslant s(v) - d$ *implies* $r(u) \leqslant s(u) - d$
   for all $r \in \widetilde{R}^{\mathrm{X}}_{m,n}(v)$ *and* $s \in \widetilde{R}^{\mathrm{Y}}_{m,n-1}(v) \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1}(v)$,
3. $r(v) \geqslant s(v) + d$ *implies* $r(u) \geqslant s(u) + d$
   for all $r \in \widetilde{R}^{\mathrm{Y}}_{m,n}(v)$ *and* $s \in \widetilde{R}^{\mathrm{X}}_{m,n-1}(v) \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}(v)$.

Very roughly speaking, the relation $u \preccurlyeq^R_{m,n} v$ means that v may be replaced by u without $\widetilde{R}_{m,n}$-rankers noticing. The converse need not be true: In general, the word u cannot be replaced by v without $\widetilde{R}_{m,n}$-rankers noticing in general.

If $u \preccurlyeq^R_{m,n} v$, then every $\widetilde{R}_{m,n}$-ranker that is defined on v is also defined on u. Moreover, conditions (2) and (3) allow to convert certain order comparisons between rankers from v to u. Let us clarify what we mean by way of an example. Consider condition (2), *e.g.*, and choose arbitrary rankers $r \in \widetilde{R}^{\mathrm{X}}_{m,n}$ and $s \in \widetilde{R}^{\mathrm{Y}}_{m,n-1} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1}$ that are defined on v. If $r(v) \leqslant s(v)$, then using $d = 0$ this condition shows that also $r(u) \leqslant s(u)$. The parameter d controls the "strictness" of the order:

- we have $\ll$ for $d = 2$,
- we have $<$ for $d = 1$,
- we have $\leqslant$ for $d = 0$, and
- we have a loose variant of $\leqslant$ for $d = -1$ (at most the next position).

What is remarkable about the rankers in that definition is that both r and s may have full alternation parameter — provided they end with different directions. Note, however, that not both have full depth parameter and that for r ending on an X-modality, being less than s converts from v to u, whereas for r ending on a Y-modality, being greater than s converts.

This relation allows the following characterization of $\Sigma^2_{m,n}[<, \mathrm{suc}]$. It is an original contribution of this thesis.

Theorem 11.16
Let $L \subseteq A^*$, and let $m, n \geqslant 1$. The following are equivalent:

1. *L is definable in* $\Sigma^2_{m,n}[<, \mathrm{suc}]$.
2. *L is definable in* $\mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$.
3. *L is an* $\preccurlyeq^R_{m,n}$*-order ideal.*

Recall that L is a $\preccurlyeq^R_{m,n}$-order ideal if $u \preccurlyeq^R_{m,n} v$ and $v \in L$ implies $u \in L$. Note that by the semantics of temporal logic formulae, it does not make sense to have X or Y as outermost modalities: Formulae of the form $u \models \mathrm{X}\varphi$ or $u \models \mathrm{Y}\varphi$ are always false. More common in the literature is that $\mathrm{X}\varphi$ holds on u if φ holds on the first position of u (and similarly $\mathrm{Y}\varphi$ goes to the last position). In other words, one imagines being just outside the word and an outer X-modality or Y-modality move to the first or last position. Although this semantics also yields Theorem 11.16 for $m, n \geqslant 2$, it does *not* for $m = 1$ (basically because $\Sigma^2_1[<, \mathrm{suc}]$ is unable to specify prefixes or suffixes).

For $u, v \in A^*$ let $u \preccurlyeq^{\mathrm{FO}^2}_{m,n} v$ if $v \in L(\varphi)$ implies $u \in L(\varphi)$ for all $\varphi \in \Sigma^2_{m,n}[<, \mathrm{suc}]$; and let $u \preccurlyeq^{\mathrm{TL}}_{m,n} v$ if $v \in L(\varphi)$ implies $u \in L(\varphi)$ for all $\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$. As in the case without successor, we can rephrase Theorem 11.16.

Corollary 11.17
Let $m, n \geqslant 1$, and let $u, v \in A^*$. The following are equivalent:
1. $u \preccurlyeq^{\mathrm{FO}^2}_{m,n} v$.
2. $u \preccurlyeq^{\mathrm{TL}}_{m,n} v$.
3. $u \preccurlyeq^{R}_{m,n} v$.

Proof. This proof is completely analogous to that of Corollary 11.4, using Theorem 11.16 instead of Theorem 11.3. □

We postpone the proof of the theorem and first consider the consequences for the full levels $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$. With successor the neighborhood can be specified, leading to the following refinement of the order type: For integers i and j the *successor order type* $\mathrm{ord}_S(i, j) \in \{\ll, -1, 0, +1, \gg\}$ is defined by

$$\mathrm{ord}_S(i,j) = \begin{cases} \ll & \text{if } i < j-1, \\ -1 & \text{if } i = j-1, \\ 0 & \text{if } i = j, \\ +1 & \text{if } i = j+1, \\ \gg & \text{if } i > j+1. \end{cases}$$

The following definition captures the expressive power of the full level $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$ as an equivalence relation on words that is defined by means of rankers.

Definition 11.18
Let $u, v \in A^*$ and $m, n \geqslant 0$. Let $u \approx^{R}_{m,n} v$ if either $m = 0$, or $n = 0$, or all of the following hold:
1. $\widetilde{R}_{m,n}(u) = \widetilde{R}_{m,n}(v)$,
2. $\mathrm{ord}_S(r(u), s(u)) = \mathrm{ord}_S(r(v), s(v))$
   for all $r \in \widetilde{R}^{\mathrm{X}}_{m,n}(u)$ *and all* $s \in \widetilde{R}^{\mathrm{Y}}_{m,n-1}(u) \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1}(u)$,
3. $\mathrm{ord}_S(r(u), s(u)) = \mathrm{ord}_S(r(v), s(v))$
   for all $r \in \widetilde{R}^{\mathrm{Y}}_{m,n}(u)$ *and all* $s \in \widetilde{R}^{\mathrm{X}}_{m,n-1}(u) \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}(u)$.

It can be verified that $u \approx^{R}_{m,n} v$ if and only if $u \preccurlyeq^{R}_{m,n} v$ and $v \preccurlyeq^{R}_{m,n} u$. We get the following corollaries to Theorem 11.16 for the full levels.

Corollary 11.19
Let $L \subseteq A^*$ and $m, n \geqslant 1$. The following are equivalent:
1. *L is definable in* $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$.
2. *L is definable in* $\mathrm{TL}_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$.
3. *L is a union of $\approx^{R}_{m,n}$-classes.* □

As for the half levels this can be rephrased in terms of equivalence relations on words that are given by agreement on formulae in the fragment.

Corollary 11.20
Let $m, n \geqslant 1$, and let $u, v \in A^*$. The following are equivalent:
1. $v \in L(\varphi)$ *if and only if* $u \in L(\varphi)$ *for all* $\varphi \in \mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$.
2. $v \in L(\varphi)$ *if and only if* $u \in L(\varphi)$ *for all* $\varphi \in \mathrm{TL}_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$.
3. $u \approx^R_{m,n} v$. □

The equivalence of $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$ and $\mathrm{TL}_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$ is an original contribution. As in the case without successor, Etessami, Vardi, and Wilke showed the coarse result that $\mathrm{FO}^2[<, \mathrm{suc}]$ and $\mathrm{TL}[\mathrm{F}, \mathrm{P}, \mathrm{X}, \mathrm{Y}]$ are equivalent [EVW02, Theorem 1]. In a joint paper with Kufleitner it was shown that FO^2-formulae can be translated to $\mathrm{TL}[\mathrm{F}, \mathrm{P}, \mathrm{X}, \mathrm{Y}]$-formulae respecting negations [KL13]. This translation does not preserve the quantifier depth, thus only showing the equivalence of $\mathrm{FO}^2_m[<, \mathrm{suc}]$ and $\mathrm{TL}_m[\mathrm{F}, \mathrm{P}, \mathrm{X}, \mathrm{Y}]$.

The equivalence (1) ⇔ (3) in Corollary 11.20 rectifies a similar but erroneous characterization of Weis and Immerman [WI09, Theorem 5.5]. On first glance it looks just like a slight reformulation thereof, but there is a huge difference in the definition of look-around rankers. Unlike Weis and Immerman, we allow additional modalities XX_w and YY_w.

The following example shows that those modalities are truly necessary. More precisely, it disproves the "successor ranker" characterizations of Weis and Immerman for $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$, showing that these conditions are to weak. Indeed, they fail to specify the length of models by a factor of roughly 2.

Example 11.21
Consider the language $L \subseteq A^*$ consisting of all words of length at least $4n - 3$, where $n \geqslant 2$ is arbitrary but fixed. We shall see shortly that L is definable in $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$ for all $m \geqslant 1$. On the other hand, the words $a^{4n-4} \notin L$ and $a^{4n-3} \in L$ are distinguished by neither Theorem 5.4 (i) nor Theorem 5.5 (i) in [WI09], no matter what $m \geqslant 1$. More precisely, to distinguish these words with those conditions, rankers with a length up to $2n - 1$ have to be considered. Consequently, they fail to characterize the quantifier depth by a factor of about 2.

To see definability, let us first consider the following formulae $\varphi_n(x)$ in $\Sigma^2_{1,n}[<, \mathrm{suc}]$. Let $\varphi_0(x) := \top$ and

$$\varphi_n(x) \ := \ \exists y \left(y \ll x \wedge \varphi_{n-1}(y) \right)$$

for $n \geqslant 1$, where $\varphi_{n-1}(y)$ is obtained inductively by interchanging variables x and y. Here $y \ll x$ is short for $y < x \wedge \neg \, \mathrm{suc}(y, x)$. The formula $\varphi_n(x)$ ensures that there are at least $2n$ positions strictly smaller than x. If $\varphi'_n(x)$ is the left-right dual formula ensuring $2n$ positions strictly greater than x, then

$$\exists x \left(\varphi_{n-1}(x) \wedge \varphi'_{n-1}(x) \right)$$

is in $\Sigma^2_{1,n}[<, \mathrm{suc}]$ and defines L.

Let us finally show that our characterization is able to distinguish $u = a^{4n-4}$ from $v = a^{4n-3}$. Letting

$$r = (\mathrm{XX}_a)^n \in \widetilde{R}^{\mathrm{X}}_{1,n} \quad \text{and} \quad s = (\mathrm{YY}_a)^{n-1} \in \widetilde{R}^{\mathrm{Y}}_{1,n-1}$$

we have $r(v) = r(u) = 2n - 1$ and $|v| - s(v) = |u| - s(u) = 2n - 4$. In particular $r(v) \leqslant s(v) - 2$ and $r(u) \nleqslant s(u) - 2$. Condition (2) of Definition 11.15 is thus violated and $u \not\approx^R_{1,n} v$. ◇

The remainder of this section is dedicated to the proof of Theorem 11.16. We start with the implication (1) $\Rightarrow$ (2), that is, every formula in $\Sigma^2_{m,n}[<,\mathrm{suc}]$ has an equivalent formula in $\mathrm{TL}^+_{m,n}[\mathrm{XXF},\mathrm{YYP},\mathrm{X},\mathrm{Y}]$. The proof idea much resembles that without successor predicate, but more order types have to be made explicit.

Lemma 11.22
Let $m, n \geqslant 0$. Every $\Sigma^2_{m,n}[<,\mathrm{suc}]$-definable language is $\mathrm{TL}^+_{m,n}[\mathrm{XXF},\mathrm{YYP},\mathrm{X},\mathrm{Y}]$-definable.

Proof. The proof is an adaption of that for Lemma 11.22 to include the successor predicate. The idea is that of the two variables of an FO^2-formula, one variable is given by the current position of the TL-formula. The necessary information about the other variable is queried at quantification time by splitting the quantifier with respect to the successor order type to the other variable. This information is stored syntactically for later reference. The key to the splitting of the quantifiers is the requantification-free normal form discussed in Chapter 10.

We proceed by induction on the structure of the formula. In the course of the proof we have to cope with free variables. Let $\varphi(x,y) \in \Sigma^2_{m,n}[<,\mathrm{suc}]$ be in requantification-free normal form with all outermost quantifiers binding y. For $\tau \in \{\ll, -1, 0, +1, \gg\}$ and $a \in A$ we construct a $\mathrm{TL}^+_{m,n}[\mathrm{XXF},\mathrm{YYP},\mathrm{X},\mathrm{Y}]$-formula $\langle\varphi\rangle_{\tau,a}(x)$ such that $u,p,q \models \varphi(x,y)$ if and only if $u,p \models \langle\varphi\rangle_{\tau,a}(x)$ for all $u \in A^*$ and all positions p and q of u with $u[q] = a$ and $\mathrm{ord}_S(p,q) = \tau$.

Let $\langle\psi\rangle_{\tau,a}(x) := \psi$ for $\psi \in \{\top,\bot\}$ and $\langle\lambda(x) = b\rangle_{\tau,a}(x) := b$. The remaining atomic formulae are as follows:

$$\langle\lambda(y) = b\rangle_{\tau,a}(x) \;:=\; \begin{cases} \top & \text{if } a = b,\\ \bot & \text{otherwise,}\end{cases}$$

$$\langle x < y\rangle_{\tau,a}(x) \;:=\; \begin{cases} \top & \text{if } \tau \in \{\ll, -1\},\\ \bot & \text{otherwise,}\end{cases}$$

$$\langle y < x\rangle_{\tau,a}(x) \;:=\; \begin{cases} \top & \text{if } \tau \in \{\gg, +1\},\\ \bot & \text{otherwise,}\end{cases}$$

$$\langle \mathrm{suc}(x,y)\rangle_{\tau,a}(x) \;:=\; \begin{cases} \top & \text{if } \tau = -1,\\ \bot & \text{otherwise,}\end{cases}$$

$$\langle \mathrm{suc}(y,x)\rangle_{\tau,a}(x) \;:=\; \begin{cases} \top & \text{if } \tau = +1,\\ \bot & \text{otherwise.}\end{cases}$$

For Boolean combinations let

$$\begin{aligned}\langle\varphi \wedge \psi\rangle_{\tau,a}(x) &:= \langle\varphi\rangle_{\tau,a}(x) \wedge \langle\psi\rangle_{\tau,a}(x),\\ \langle\varphi \vee \psi\rangle_{\tau,a}(x) &:= \langle\varphi\rangle_{\tau,a}(x) \vee \langle\psi\rangle_{\tau,a}(x),\\ \langle\neg\varphi\rangle_{\tau,a}(x) &:= \neg\langle\varphi\rangle_{\tau,a}(x).\end{aligned}$$

For quantification let $\langle\exists y\, \varphi\rangle_{\tau,a}(x)$ be the disjunction $\bigvee_{c\in A} c \wedge \psi_c$, where

$$\psi_c \;:=\; \mathrm{YYP}\langle\varphi\rangle_{\ll,c}(y) \vee \mathrm{Y}\langle\varphi\rangle_{-1,c}(y) \vee \langle\varphi\rangle_{0,c}(y) \vee \mathrm{X}\langle\varphi\rangle_{+1,c}(y) \vee \mathrm{XXF}\langle\varphi\rangle_{\gg,c}(y).$$

The formulae $\langle\varphi\rangle_{\tau',c}(y)$ in this definition are the dual formulae obtained by induction with the roles of x and y interchanged; *i.e.*, y is given by the current position and the parameters τ' and b determine all relevant information about x. Note that quantification over x does not occur by assumption.

This allows to prove the lemma. Let $\varphi \in \Sigma^2_{m,n}[<, \mathrm{suc}]$ be a sentence. Proposition 10.3 shows that φ can be assumed without restriction to be in requantification-free normal form with all outermost quantifiers binding y. We inductively construct an equivalent $\mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$-formula φ'. Let $(\varphi \wedge \psi)' := \varphi' \wedge \psi'$, let $(\varphi \vee \psi)' := \varphi' \vee \psi'$, and let $(\neg\varphi)' := \neg\varphi'$. Quantification is given by $(\exists y \colon \varphi)' := \mathrm{XXF}(\langle\varphi\rangle_{\tau,a}(x))$, where $\tau \in \{\ll, -1, 0, +1, \gg\}$ and $a \in A$ can be chosen arbitrarily. This shows that for each sentence $\varphi \in \Sigma^2_{m,n}[<, \mathrm{suc}]$ there is $\varphi' \in \mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$ with $L(\varphi) = L(\varphi')$. $\square$

We can prove the following result which takes minimum and maximum predicates into account along the same lines. Although we are presently only interested in the signature $[<, \mathrm{suc}]$, we shall use it in the next chapter.

Lemma 11.23
Let $m, n \geqslant 0$. Every language that is definable in $\Sigma^2_{m,n}[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$ is definable in $\mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}, \mathrm{min}, \mathrm{max}]$.

Proof. The proof is very similar to that of Lemma 11.22. We only sketch the differences, starting with an intuition. The idea was that one free variable is represented by the current position of the TL-formula and all relevant information about the other variable is stored syntactically. The notion of *relevant information* has to be extended for the signature including min and max: We additionally have to remember whether the dismissed variable was minimal or maximal. This is done by two Boolean flags b_{min} and b_{max} as additional parameters.

Let us render this formally. For requantification-free $\varphi(x, y) \in \Sigma^2_{m,n}[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$ with all outermost quantifiers binding y, an order type $\tau \in \{\ll, -1, 0, +1, \gg\}$, a letter $a \in A$, and Boolean flags $b_{\mathrm{min}}, b_{\mathrm{max}} \in \{\bot, \top\}$ we give a formula $\langle\varphi\rangle_{\tau,a,b_{\mathrm{min}},b_{\mathrm{max}}}(x)$ in $\mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}, \mathrm{min}, \mathrm{max}]$. The semantics of this formula is as follows: Suppose $u \in A^*$, and suppose that p and q are positions of u such that

- $u[q] = a$,
- $\mathrm{ord}_S(p, q) = \tau$,
- $q = 1 \Leftrightarrow b_{\mathrm{min}} = \top$, and $q = |u| \Leftrightarrow b_{\mathrm{max}} = \top$.

Under these premises $u, p, q \models \varphi(x, y)$ if and only if $u, p \models \langle\varphi\rangle_{\tau,a,b_{\mathrm{min}},b_{\mathrm{max}}}(x)$.

The construction is mostly along the same lines as that for Lemma 11.22. But of course we have to handle minimum and maximum predicates (over both variables), and we also have to adapt quantification to store the necessary information. Let

$$\begin{aligned}
\langle\mathrm{min}(x)\rangle_{\tau,a,b_{\mathrm{min}},b_{\mathrm{max}}}(x) &:= \mathrm{min}, \\
\langle\mathrm{max}(x)\rangle_{\tau,a,b_{\mathrm{min}},b_{\mathrm{max}}}(x) &:= \mathrm{max}, \\
\langle\mathrm{min}(y)\rangle_{\tau,a,b_{\mathrm{min}},b_{\mathrm{max}}}(x) &:= b_{\mathrm{min}}, \\
\langle\mathrm{max}(y)\rangle_{\tau,a,b_{\mathrm{min}},b_{\mathrm{max}}}(x) &:= b_{\mathrm{max}}.
\end{aligned}$$

For the quantification $\langle\exists y\ \varphi\rangle_{\tau,a,b_{\mathrm{min}},b_{\mathrm{max}}}(x)$, we adapt the construction to ensure the premises for semantic correctness. First let

$$\begin{aligned}
\psi_{c,b'_{\mathrm{min}},b'_{\mathrm{max}}} := \ & \mathrm{YYP}\langle\varphi\rangle_{\ll,c,b'_{\mathrm{min}},b'_{\mathrm{max}}}(y) \vee \mathrm{Y}\langle\varphi\rangle_{-1,c,b'_{\mathrm{min}},b'_{\mathrm{max}}}(y) \vee \langle\varphi\rangle_{0,c,b'_{\mathrm{min}},b'_{\mathrm{max}}}(y) \vee \\
& \mathrm{XXF}\langle\varphi\rangle_{\gg,c,b'_{\mathrm{min}},b'_{\mathrm{max}}}(y) \vee \mathrm{X}\langle\varphi\rangle_{+1,c,b'_{\mathrm{min}},b'_{\mathrm{max}}}(y)
\end{aligned}$$

and with that set

$$\langle \exists y\ \varphi \rangle_{\tau,a,b_{\min},b_{\max}}(x) \ := \bigvee_{\substack{c\in A\\ b'_{\min},b'_{\max}\in\{\bot,\top\}}} c \wedge [b'_{\min} \leftrightarrow \min] \wedge [b'_{\max} \leftrightarrow \max] \wedge \psi_{c,b'_{\min},b'_{\max}},$$

where for $P \in \{\min, \max\}$ the symbol $[b'_P \leftrightarrow P]$ is short for $(P \wedge b'_P) \vee (\neg P \wedge \neg b'_P)$. Note that the negations involved are on atomic level and thus do not increase the alternation parameter. All other constructs are similar to those for Lemma 11.22.

The subsequent construction on the sentence level of $\Sigma^2_{m,n}[<, \mathrm{suc}, \min, \max]$ is a straightforward adaption of the proof without minimum and maximum. $\square$

We next come to the implication (2) ⇒ (3) in Theorem 11.16, *i.e.*, every language definable in $\mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$ is an $\preccurlyeq^R_{m,n}$-order ideal.

We start with an auxiliary lemma, providing a means to displace a look-around ranker by a fixed amount of positions without losing too much resources. For integers i, j, and k and a word v let $v[j;i;k]$ be the context $(v[j;i-1], v[i], v[i+1;k])$ be obtained by slicing the intervals $[j;i-1]$, $[i;i]$, and $[i+1;k]$ out off the word v. Note that $v[j;i;k]$ is a $\max(i-j, k-j)$-context.

Lemma 11.24
Let $m, n \geqslant 0$, $k \in \mathbb{Z}$, $v \in A^*$, and $r \in \widetilde{R}_{m,n-|k|}(v)$. If $r(v)+k$ is a position of v, then there exists $\langle r+k\rangle \in \widetilde{R}_{m,n}$ such that for all $u \in A^*$ with $u \preccurlyeq^R_{m,n} v$ the following hold:

1. *$r \in \widetilde{R}\mathrm{Z}_{w'}$ implies $\langle r+k\rangle \in \widetilde{R}\mathrm{Z}_w$, where $w = v[i-n+1; i; i+n-1]$ and $i = r(v)+k$,*
2. *$\langle r+k\rangle(v) = r(v)+k$,*
3. *$\langle r+k\rangle(u) \geqslant r(u)+k$ if $r \in \widetilde{R}^{\mathrm{X}}_{m,n-|k|} \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n-|k|}$,*
4. *$\langle r+k\rangle(u) \leqslant r(u)+k$ if $r \in \widetilde{R}^{\mathrm{Y}}_{m,n-|k|} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-|k|}$.*

Proof. By assumption w' is an $(n-|k|-1)$-context. Intuitively, when shifting the $(n-1)$-context w by k positions to the left, it encompasses the context w'. Moreover, we make w exhaust its look-around of $n-1$ positions by filling it up with the content of v around $r(v)+k$.

It suffices to give the construction only for $k \leqslant 0$ due to symmetry. If $r = \mathrm{Z}_{w'}$, then let $\langle r+k\rangle = \mathrm{Z}_w$. Suppose $r = r'\mathrm{Z}_{w'}$ for some ranker r' and some $\mathrm{Z} \in \{\mathrm{X}, \mathrm{XX}, \mathrm{Y}, \mathrm{YY}\}$. If $r'(v)+k$ is not a position of v, then $r'(v) \leqslant k$. In particular $\mathrm{Z} \in \{\mathrm{X}, \mathrm{XX}\}$, because $r'\mathrm{Z}_{w'}$ is defined. In this case we set $\langle r+k\rangle = \mathrm{Z}_w$. Otherwise, if $r'(v)+k$ is a position of v, then by induction there exists $\langle r'+k\rangle$ and we let $\langle r+k\rangle = \langle r'+k\rangle\mathrm{Z}_w$.

Properties (1) and (2) are clear by construction. Note that as w shifts w' by k positions, we also have to displace the starting position of the modality Z_w by k positions.

Property (3): For $r \in \widetilde{R}^{\mathrm{Y}}_{m-1,n-|k|}$ follows from condition (3) in Definition 11.15 for $u \preccurlyeq^R_{m,n} v$. For $r \in \widetilde{R}^{\mathrm{X}}_{m,n-|k|}$ note that increasing the size of the look-around can only lead to longer jumps of the modalities.

Property (4) is symmetric (and for $r \in \widetilde{R}^{\mathrm{X}}_{m-1,n-|k|}$ relies on condition (2) in Definition 11.15 for $u \preccurlyeq^R_{m,n} v$). $\square$

Note that $\langle r+k\rangle$ depends on the word v as well as on the depth parameter n, both of which are understood implicitly from context in the later applications. For rankers $r \in \widetilde{R}_{m-1,n-|k|}$ not using the full alternation parameter this means $\langle r+k\rangle(u) = r(u)+k$.

For the last alternation $r \in \widetilde{R}_{m,n-|k|}$ on the other hand, the ranker $\langle r+k \rangle$ does not necessarily identify the position $r(u)+k$. However, the error is one-sided: Depending on whether the last modality is an X-modality or a Y-modality, $\langle r+k \rangle(u)$ is either at least or at most $r(u)+k$.

To cope with the free variable of a temporal formula, we extend the relation $\preccurlyeq^R_{m,n}$ over words to pairs of a word with a designated current position. For a uniform presentation we also include a pseudo-position $-\infty$ for the situation at the beginning of the evaluation of temporal formulae (where we imagine to be at an unlabeled position far outside of the word).

Let $u, v \in A^*$, let $m, n \geqslant 0$, let x be a position of u or $x = -\infty$, and let x' be a position of v or $x' = -\infty$. Define the relation $(u, x') \preccurlyeq^R_{m,n} (v, x)$ if $u \preccurlyeq^R_{m,n} v$ and $u[x'-n; x'+n] = v[x-n; x+n]$, and if further either $m = 0$, or $n = 0$, or all of the following hold for $d \in \{-1, 0, 1, 2\}$:

1. $s(v) \leqslant x - d \Rightarrow s(u) \leqslant x' - d$ if $s \in \widetilde{R}^{\mathrm{X}}_{m,n}(v) \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n}(v)$,
2. $s(v) \geqslant x + d \Rightarrow s(u) \geqslant x' + d$ if $s \in \widetilde{R}^{\mathrm{Y}}_{m,n}(v) \cup \widetilde{R}^{\mathrm{X}}_{m-1,n}(v)$.

Extending our notation, we let $u[x'-n; x'+n]$ and $v[x-n; x+n]$ be empty if $x' = -\infty$ and $x = -\infty$, respectively. Note that thus $(u, -\infty) \preccurlyeq^R_{m,n} (v, -\infty)$ if $u \preccurlyeq^R_{m,n} v$.

The intuition of $(u, x') \preccurlyeq^R_{m,n} (v, x)$ is that, for $\mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$, the position x' on u is harder to distinguish than the position x on v. Formally this means that for $(u, x') \preccurlyeq^R_{m,n} (v, x)$ we have that $v, x \models \varphi$ implies $u, x' \models \varphi$ for all temporal formulae $\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$. The following is the key to prove this intuition: Every move of the current position on v can be countered on u with the same order type, only losing one for the depth parameter.

Lemma 11.25
Let $u, v \in A^*$, let $m, n \geqslant 1$, let x be a position of v or $x = -\infty$, and let x' be a position of u or $x' = -\infty$. If $(u, x') \preccurlyeq^R_{m,n} (v, x)$, then for every position y of v there exists a position y' of u such that $\mathrm{ord}_S(x, y) = \mathrm{ord}_S(x', y')$ and $(u, y') \preccurlyeq^R_{m,n-1} (v, y)$.

Proof. We distinguish the different order types for $\mathrm{ord}_S(x, y)$. For $y = x$ let $y' = x'$, and, by left-right symmetry, it suffices to consider the cases $y = x+1$ and $y \gg x$.

Suppose $y = x + 1$. We shall show that $y' = x' + 1$ satisfies $(u, y') \preccurlyeq^R_{m,n-1} (v, y)$. Note that $x' + 1$ actually is a position of u. The alphabetic condition is immediate and it remains to show conditions (1) and (2). The idea is to use the ranker $\langle s - 1 \rangle$ from Lemma 11.24 with depth n which identifies the position $s(v) - 1$ and to apply the assumption on this ranker.

Condition (1): Let $s \in \widetilde{R}^{\mathrm{X}}_{m,n-1}(v) \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}(v)$, and let $d \in \{-1, 0, 1, 2\}$. Suppose $s(v) \leqslant x + 1 - d$. Specifically, $s(v) - 1 = \langle s - 1 \rangle(v) \leqslant x - d$ and due to $\langle s - 1 \rangle \in \widetilde{R}^{\mathrm{X}}_{m,n} \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n}$, the assumption yields $s(u) - 1 \leqslant \langle s - 1 \rangle(u) \leqslant x' - d$. This shows $s(u) \leqslant x' + 1 - d$.

Condition (2): Let $s \in \widetilde{R}^{\mathrm{Y}}_{m,n-1}(v) \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1}(v)$, and let $d \in \{-1, 0, 1, 2\}$. Suppose $s(v) \geqslant x+1+d$. Then $s(v)-1 = \langle s - 1 \rangle(v) \geqslant x+d$ implies $s(u)-1 \geqslant \langle s - 1 \rangle(u) \geqslant x'+d$. This shows $s(u) \geqslant x' + 1 + d$.

Suppose $y \gg x$. Be warned, gentle reader; we have to do a lot of preparatory work before we are able to attack what we actually want to show. For a concise presentation of the following we start by extending look-around rankers. For $w = (p, a, q)$ and $d \in \{-1, 0, 1, 2\}$ we introduce modalities X^d_w and Y^{-d}_w, whose semantics are given by

$$\mathrm{X}^d_w(u, i) = \min \{i + d \leqslant j \leqslant |u| \mid u[j - |p|; j + |q|] = paq\},$$
$$\mathrm{Y}^{-d}_w(u, i) = \max \{1 \leqslant j \leqslant i - d \mid u[j - |p|; j + |q|] = paq\}$$

for a word $u \in A^*$ and a position i of u. Without designated starting position let $\mathrm{X}^d_w(u) = \mathrm{X}_w(u)$ and $\mathrm{Y}^{-d}_w(u) = \mathrm{Y}_w(u)$. The evaluation of composite rankers with these modalities is as usual from left to right. Intuitively, the modalities offset the current position by d and start matching w to the right (for X-modalities) or to the left (for Y-modalities). Note that we may identify $\mathrm{X}_w = \mathrm{X}^1_w$, $\mathbb{X}_w = \mathrm{X}^2_w$, $\mathrm{Y}_w = \mathrm{Y}^{-1}_w$, and $\mathbb{Y}_w = \mathrm{Y}^{-2}_w$. We use these symbols interchangeably.

With the $(n-1)$-context $w = v[y - n + 1; y; y + n - 1]$ let

$$\widetilde{R}^d_{\mathit{left}} = \left\{\ell' \mathrm{X}^d_w \;\middle|\; \ell' \in \{\varepsilon\} \cup \widetilde{R}^{\mathrm{X}}_{m,n-1} \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n-1} \text{ and } \ell' \mathrm{X}^d_w(v) \leqslant y\right\},$$
$$\widetilde{R}^d_{\mathit{right}} = \left\{r' \mathrm{Y}^d_w \;\middle|\; r' \in \{\varepsilon\} \cup \widetilde{R}^{\mathrm{Y}}_{m,n-1} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1} \text{ and } r' \mathrm{Y}^d_w(v) \geqslant y\right\},$$

and let $\widetilde{R}_{\mathit{left}} = \bigcup_{d=-1}^{2} \widetilde{R}^d_{\mathit{left}}$ and $\widetilde{R}_{\mathit{right}} = \bigcup_{d=-2}^{1} \widetilde{R}^d_{\mathit{right}}$.

Let $\ell \in \widetilde{R}_{\mathit{left}}$ such that $\ell(u)$ is maximal, and if there is more than one possible choice, we give preference to rankers ending on X^d_w with maximal d. Symmetrically, let $r \in \widetilde{R}_{\mathit{right}}$ such that $r(u)$ is minimal, and if there is more than one possibility, we give preference to rankers ending on Y^d_w with minimal d. We are going to show that we can choose $y' = r(u)$. For this we shall prove $\ell(u) \leqslant r(u)$. Note that $\ell(v) \leqslant r(v)$.

We start with a preliminary observation which helps to avoid the "new" modalities X^0_w and X^{-1}_w as well as Y^0_w and Y^1_w thus enabling the usage of $(u, x') \preccurlyeq^R_{m,n} (v, x)$: We claim that if $\ell = \ell' \mathrm{X}^d_w$ for some $d \in \{-1, 0\}$ and some ℓ', then $\ell' \in \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}$ and $\ell'(v) \geqslant y$. Note that ℓ' cannot be empty by choice of ℓ. For the sake of contradiction first assume $\ell'(v) < y$. Then $\ell' \mathrm{X}_w \in \widetilde{R}_{\mathit{left}}$ and $\ell' \mathrm{X}_w(u) \geqslant \ell' \mathrm{X}^0_w(u) \geqslant \ell' \mathrm{X}^{-1}_w(u)$. By choice of ℓ this contradicts $\ell = \ell' \mathrm{X}^d_w$ with $d \in \{-1, 0\}$. Suppose now $\ell'(v) \geqslant y$ and $\ell' \in \widetilde{R}^{\mathrm{X}}_{m,n-1}$. Note that $\ell'(v) \geqslant y$ actually means $\ell'(v) = y$ or $\ell'(v) = y + 1$. Let $\ell' = \ell'' \mathrm{Z}_{w'}$ for some $\mathrm{Z} \in \{\mathrm{X}, \mathbb{X}\}$ and some possibly empty ℓ''. If $\ell'(v) = y$, then $\ell'' \mathrm{Z}_w(v) = y$, *i.e.*, $\ell'' \mathrm{Z}_w \in \widetilde{R}_{\mathit{left}}$. Moreover, $\ell'' \mathrm{Z}_w(u) \geqslant \ell'' \mathrm{Z}_{w'} \mathrm{X}^d_w(u)$ because w' appears as a factor of w (meaning that we have $w' = (w'_1, a, w'_2)$ and $w = (w_1 w'_1, a, w'_2 w_2)$ for some $w'_i, w_i \in A^*$). Suppose $\ell'(v) = y + 1$, *i.e.*, $\ell = \ell' \mathrm{X}^{-1}_w$. Using the ranker from Lemma 11.24 we have $\langle \ell' - 1 \rangle(v) = y$ and $\langle \ell' - 1 \rangle \in \widetilde{R}^{\mathrm{X}}_{m,n}$. In particular $\langle \ell' - 1 \rangle \in \widetilde{R}_{\mathit{left}}$ and $\langle \ell' - 1 \rangle(u) \geqslant \ell'(u) - 1$. Since the last modality of $\langle \ell' - 1 \rangle$ jumps to a w-position, we see $\langle \ell' - 1 \rangle(u) \geqslant \ell' \mathrm{X}^{-1}_w(u) = \ell(u)$. By choice of ℓ this contradicts $\ell = \ell' \mathrm{X}^{-1}_w$.

We thus have $\ell = \ell' \mathrm{X}^d_w$ with $d \in \{-1, 0\}$ only if $\ell' \in \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}$; moreover, if $d = 0$, then $\ell'(v) = y$, and if $d = -1$, then $\ell'(v) = y + 1$. Symmetrically, $r = r' \mathrm{Y}^d_w$ with $d \in \{0, 1\}$ only if $r' \in \widetilde{R}^{\mathrm{X}}_{m-1,n-1}$; and if $d = 0$, then $r'(v) = y$, and if $d = 1$, then $r'(v) = y - 1$.

Next, we show that $\ell(u) \leqslant r(u)$. We distinguish between the last modalities of ℓ and r. Suppose $\ell \in \widetilde{R}^{\mathrm{X}}_{m,n}$ (*i.e.*, $\ell = \ell' \mathrm{X}^1_w$ or $\ell = \ell' \mathrm{X}^2_w$). Let $r = r' \mathrm{Y}^d_w$ with $d \in \{-2, \ldots, 1\}$. If r' is empty, then clearly $\ell(u) \leqslant r' \mathrm{Y}^d_w(u) = \mathrm{Y}_w(u)$. Otherwise $\ell(v) \leqslant r'(v) + d$. Condition (2) of $u \preccurlyeq^R_{m,n} v$ yields $\ell(u) \leqslant r'(u) + d$. Therefore, $\ell(u) \leqslant r' \mathrm{Y}^d_w(u)$ which shows $\ell(u) \leqslant r(u)$. For $r \in \widetilde{R}^{\mathrm{Y}}_{m,n}$ this follows left-right symmetrically.

It remains to consider $\ell = \ell' \mathrm{X}^{d'}_w$ and $r = r' \mathrm{Y}^d_w$ for $d' \in \{-1, 0\}$ and $d \in \{0, 1\}$. First suppose $\ell = \ell' \mathrm{X}^0_w$ and let $r = r' \mathrm{Y}^d_w$ with $d \in \{0, 1\}$. By the above observation we have

$\ell'(v) = y = r'(v) + d$ with $\ell' \in \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}$ and $r' \in \widetilde{R}^{\mathrm{X}}_{m-1,n-1}$. Let us first see that $\ell'(u)$ is a w-position of u. Let $\ell' = \ell''\mathrm{Z}_{w'}$ with possibly empty ℓ'' and $\mathrm{Z} \in \{\mathrm{Y}, \mathrm{YY}\}$. Since w' is a factor of w, we trivially have $\ell''\mathrm{Z}_w(u) \leqslant \ell'(u)$. On the other hand $\ell''\mathrm{Z}_w(v) \geqslant \ell'(v)$ and condition (3) of $u \preccurlyeq^R_{m,n} v$ yields $\ell''\mathrm{Z}_w(u) \geqslant \ell'(u)$. This shows that $\ell'(u) = \ell''\mathrm{Z}_w(u)$ is a w-position of u. Now, assuming $\ell'(u) \geqslant r'(u) + d + 1$ yields $\ell'(v) \geqslant r'(v) + d + 1$ using condition (2) of $v \preccurlyeq^R_{m-1,n} u$. This is a contradiction. Thus $\ell'(u) \leqslant r'(u) + d$ which implies $\ell(u) = \ell'\mathrm{X}^0_w(u) \leqslant r'\mathrm{Y}^d_w(u) = r(u)$. The argument for $\ell = \ell'\mathrm{X}^{-1}_w$ and $r = r'\mathrm{Y}^0_w$ is left-right symmetric to the case of $\ell = \ell'\mathrm{X}^0_w$ and $r = r'\mathrm{Y}^1_w$.

The last remaining case is $\ell = \ell'\mathrm{X}^{-1}_w$ and $r = r'\mathrm{Y}^1_w$. By the above observation we have $\ell'(v) - 1 = y = r'(v) + 1$ for some $\ell' \in \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}$ and some $r' \in \widetilde{R}^{\mathrm{X}}_{m-1,n-1}$. Hence $\ell'\mathrm{Y}_w(v) = y$. Note that $\ell'\mathrm{Y}_w(v) \geqslant \ell'(v) - 1$ by condition (3) of $u \preccurlyeq^R_{m,n} v$ and thus $\ell'\mathrm{Y}_w(u) = \ell'(u) - 1$. For the sake of contradiction assume $\ell'(u) - 1 \geqslant r'(u) + 2$, that is, $\ell'\mathrm{Y}_w(u) \geqslant r'(u) + 2$. But then $\ell'\mathrm{Y}_w(v) \geqslant r'(v) + 2$ by condition (3) of $v \preccurlyeq^R_{m-1,n} u$ because $\ell'\mathrm{Y}_w \in \widetilde{R}^{\mathrm{Y}}_{m-1,n}$. Since $\ell'\mathrm{Y}_w(v) = y$ this contradicts $y = r'(v) + 1$. Thus $\ell'(u) - 1 \leqslant \ell'\mathrm{Y}_w(u) \leqslant r'(u) + 1$ which implies $\ell'\mathrm{X}^{-1}_w(u) \leqslant r'\mathrm{Y}^1_w(u)$. This concludes the proof that $\ell(u) \leqslant r(u)$.

As was already mentioned, we want to show that we can choose $y' = r(u)$. We have to show $y' \gg x'$ and $(u, y') \preccurlyeq^R_{m,n-1} (v, y)$. Suppose $r \in \widetilde{R}^{\mathrm{Y}}_{m,n}$, that is, $r = r'\mathrm{Y}^d_w$ with $d \in \{-2, -1\}$. Since $r(v) \geqslant y \gg x$, condition (2) of $(u, x') \preccurlyeq^R_{m,n} (u, x)$ yields $y' = r(u) \gg x'$. Suppose $r = r'\mathrm{Y}^0_w$. Then $r' \in \widetilde{R}^{\mathrm{X}}_{m-1,n-1}$ and $r'(v) = y$. A similar argument as above shows that $r'(u)$ is a w-position of u. This yields $y' = r(u) = r'(u)$. Since $y = r'(v) \gg x$, condition (2) of $(u, x') \preccurlyeq^R_{m,n} (u, x)$ yields $y' = r'(u) \gg x'$. Suppose $r = r'\mathrm{Y}^1_w$. Then $r' \in \widetilde{R}^{\mathrm{X}}_{m-1,n-1}$ and $r'(v) = y - 1$. Hence $\langle r' + 1\rangle(v) = y \gg x$, where $\langle r' + 1\rangle$ is the ranker from Lemma 11.24. As $\langle r' + 1\rangle \in \widetilde{R}^{\mathrm{X}}_{m-1,n}$, condition (2) of $(u, x') \preccurlyeq^R_{m,n} (u, x)$ yields $\langle r' + 1\rangle(u) \gg x'$. Moreover, $\langle r' + 1\rangle(u) = r'(u) + 1$ is a w-position of u. We therefore have $y' = r'\mathrm{Y}^1_w(u) = r'(u) + 1 \gg x'$.

It remains to show $(u, y') \preccurlyeq^R_{m,n-1} (v, y)$, the alphabetic condition of which is clear by choice of w.

Condition (1): Let $d \in \{-1, 0, 1, 2\}$ and $s \in \widetilde{R}^{\mathrm{X}}_{m,n-1} \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}$ such that $s(v) \leqslant y - d$. With $s(v) \leqslant y - d$ we see $s\mathrm{X}^d_w(v) \leqslant y$. The choice of ℓ yields $s\mathrm{X}^d_w(u) \leqslant \ell(u) \leqslant r(u) = y'$. This implies $s(u) \leqslant y' - d$.

Condition (2): Let $d \in \{-1, 0, 1, 2\}$ and $s \in \widetilde{R}^{\mathrm{Y}}_{m,n-1} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1}$ such that $s(v) \geqslant y + d$. Since $s(v) \geqslant y + d$ we see that $s\mathrm{Y}^d_w(v) \geqslant y$. The choice of r yields $s\mathrm{Y}^d_w(u) \geqslant r(u) = y'$. This implies $s(u) \geqslant y' + d$. □

The following lemma shows that the roles of u and v can be interchanged by investing one negation.

Lemma 11.26
Let $u, v \in A^*$, and let $m, n \geqslant 1$. If $(u, x') \preccurlyeq^R_{m,n} (v, x)$, then $(v, x) \preccurlyeq^R_{m-1,n} (u, x')$.

Proof. Suppose $m \geqslant 2$ and $n \geqslant 1$ as otherwise the claim is trivial. The alphabetic condition is clear and $v \preccurlyeq^R_{m-1,n} u$ follows directly from $u \preccurlyeq^R_{m,n} v$.

Condition (1): Let $s \in \widetilde{R}^{\mathrm{X}}_{m-1,n}(u) \cup \widetilde{R}^{\mathrm{Y}}_{m-2,n}(u)$ and $d \in \{-1, 0, 1, 2\}$ with $s(u) \leqslant x' - d$. For the sake of contradiction assume $s(v) > x - d$. Then $s(v) \geqslant x + (1 - d)$. Now, since $s \in \widetilde{R}(v)$ and $1 - d \in \{-1, 0, 1, 2\}$, condition (2) for $(u, x') \preccurlyeq^R_{m,n} (v, x)$ implies $s(u) \geqslant x' + (1 - d)$, contradicting $s(u) \leqslant x' - d$. This shows $s(v) \leqslant x - d$.

Condition (2) is left-right symmetrical. □

Next we combine Lemma 11.25 and 11.26 by a straightforward induction showing that every language definable by a sentence in $\mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$ is a $\preccurlyeq^R_{m,n}$-order ideal.

Lemma 11.27
Let $u, v \in A^*$, and let $m, n \geqslant 0$. If $u \preccurlyeq^R_{m,n} v$, then $u \preccurlyeq^{\mathrm{TL}}_{m,n} v$.

Proof. We claim that if $(u, x') \preccurlyeq^R_{m,n} (v, x)$, then $v, x \models \varphi$ implies $u, x' \models \varphi$ for all formulae $\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$. The proof of this claim is by structural induction. For atomic formulae this follows from the alphabetic condition of $u \preccurlyeq^R_{m,n} v$. In particular we may assume $m, n \geqslant 1$. For Boolean connectives the claim follows by induction. (For negation this relies on Lemma 11.26.)

Suppose $v, x \models \mathrm{X}\varphi$, *i.e.*, $x + 1$ is a position of v and $v, x + 1 \models \varphi$. Lemma 11.25 yields that $x' + 1$ is a position of u such that $(u, x' + 1) \preccurlyeq^R_{m,n-1} (v, x + 1)$. By induction $u, x' + 1 \models \varphi$ and thus $u, x' \models \mathrm{X}\varphi$. Suppose $v, x \models \mathrm{XXF}\varphi$. Then there exists a position y of v such that $y \gg x$ and $v, y \models \varphi$. By Lemma 11.25 there exists a position y' of u such that $y' \gg x'$ and $(u, y') \preccurlyeq^R_{m,n-1} (v, y)$. Because $\varphi \in \mathrm{TL}^+_{m,n-1}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$, induction yields $u, y' \models \varphi$ and thus $u, x' \models \mathrm{XXF}\varphi$. The remaining cases $\mathrm{YYP}\varphi$ and $\mathrm{Y}\varphi$ are symmetric to $\mathrm{XXF}\varphi$ and $\mathrm{X}\varphi$, respectively. This concludes the proof of the claim.

Suppose now $v \models \varphi$ with $\varphi \in \mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$. We may assume that neither X nor Y appear as outermost modalities. Replacing outermost YYP-modalities by XXF yields a formula defining the same language. We may therefore assume without loss of generality that $v \models \varphi$ if and only if $v, -\infty \models \varphi$. Now $(u, -\infty) \preccurlyeq^R_{m,n} (v, -\infty)$ whenever $u \preccurlyeq^R_{m,n} v$. The claim thus yields $u, -\infty \models \varphi$. This shows $u \preccurlyeq^{\mathrm{TL}}_{m,n} v$. □

This will suffice to prove the implication (2) ⇒ (3) of Theorem 11.16 from temporal logic to the ranker characterization.

We now close the circle by showing that every $\preccurlyeq^R_{m,n}$-order ideal is definable in $\Sigma^2_{m,n}[<, \mathrm{suc}]$. The next lemma is an intermediate step towards this. It shows that some order comparisons between a variable and certain rankers are expressible in $\Sigma^2_{m,n}$.

Lemma 11.28
If $r \in \widetilde{R}^{\mathrm{X}}_{m,n} \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n}$ and $d \in \{-1, 0, 1, 2\}$, then there exists a formula $\langle x \geqslant r + d \rangle$ in $\Sigma^2_{m,n}[<, \mathrm{suc}]$ with free variable x such that for all words u with $r(u)$ being defined and for all positions i of u we have

$$u, i \models \langle x \geqslant r + d \rangle \quad \text{if and only if} \quad i \geqslant r(u) + d.$$

Symmetrically, for $r \in \widetilde{R}^{\mathrm{Y}}_{m,n} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n}$ and $d \in \{-1, 0, 1, 2\}$ there exists $\langle x \leqslant r - d \rangle$ in $\Sigma^2_{m,n}[<, \mathrm{suc}]$ with free variable x such that for all words u with $r(u)$ being defined and for all positions i of u we have

$$u, i \models \langle x \leqslant r - d \rangle \quad \text{if and only if} \quad i \leqslant r(u) - d.$$

Proof. We perform an induction on m and n. If $m = 0$ or if $n = 0$, there is nothing to show, so let $m, n \geqslant 1$. It suffices to give $\langle x \geqslant r + d \rangle$ for rankers $r \in \widetilde{R}^{\mathrm{X}}_{m,n} \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n}$. The formulae $\langle x \leqslant r - d \rangle$ for rankers in $\widetilde{R}^{\mathrm{Y}}_{m,n} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n}$ are left-right symmetric.

If $r \in \widetilde{R}^{\mathrm{Y}}_{m-1,n}$, then we set $\langle x \geqslant r + d \rangle := \neg \langle x \leqslant r + (d - 1) \rangle$, where $\langle x \leqslant r + (d - 1) \rangle$ is obtained by induction on m. It remains to consider $r \in \widetilde{R}^{\mathrm{X}}_{m,n}$. Suppose either $r = s\mathrm{XX}_w$

and $d' = 2$, or $r = s\mathrm{X}_w$ and $d' = 1$ for some $w = (p, a, q)$ with words p, q such that $|p|, |q| \leqslant n - 1$ and some $s \in \{\varepsilon\} \cup \widetilde{R}^{\mathrm{X}}_{m,n-1} \cup \widetilde{R}^{\mathrm{Y}}_{m-1,n-1}$. Let

$$\begin{aligned}
\langle x \geqslant r+2 \rangle &:= \exists y \ll x \left(\lambda(y) = w \wedge \langle y \geqslant s + d' \rangle\right),\\
\langle x \geqslant r+1 \rangle &:= \exists y < x \left(\lambda(y) = w \wedge \langle y \geqslant s + d' \rangle\right),\\
\langle x \geqslant r \rangle &:= \exists y \leqslant x \left(\lambda(y) = w \wedge \langle y \geqslant s + d' \rangle\right),\\
\langle x \geqslant r-1 \rangle &:= \exists y \leqslant x+1 \left(\lambda(y) = w \wedge \langle y \geqslant s + d' \rangle\right).
\end{aligned}$$

Here we use the shortcuts $(y \ll x) := (y < x \wedge \neg\,\mathrm{suc}(y, x))$, $(y \leqslant x) := \neg(x < y)$ and $(y \leqslant x + 1) := (y \leqslant x \vee \mathrm{suc}(x, y))$. The extended label $\lambda(x) = w$ is defined by (8.1). The formula $\langle y \geqslant s + d' \rangle$ is obtained by induction on n with the convention that for $s = \varepsilon$ it is to be read as $\top$. As usual, the formula $\langle y \geqslant s + d' \rangle$ with free variable y is obtained by interchanging x and y. $\square$

For a transparent presentation, the notation of the previous lemma is adapted in a canonical way, and we write, *e.g.*, $\langle x \geqslant r \rangle$, $\langle x > r \rangle$, or $\langle x \gg r \rangle$, specifying the parameter d implicitly. We also write $\langle x \geqslant r - d \rangle$ instead of $\langle x \geqslant r + (-d) \rangle$. Similar remarks apply to $\langle x \leqslant r - d \rangle$.

Lemma 11.29
Let $u, v \in A^*$, and let $m, n \geqslant 0$. If $u \preccurlyeq^{\mathrm{FO2}}_{m,n} v$, then $u \preccurlyeq^{R}_{m,n} v$.

Proof. We perform an induction on m. The base case for $m = 0$ is trivial. Suppose $m \geqslant 1$ and $u \preccurlyeq^{\mathrm{FO2}}_{m,n} v$.

In particular $v \preccurlyeq^{\mathrm{FO2}}_{m-1,n} u$ and hence $v \preccurlyeq^{R}_{m-1,n} u$ by induction. We now show the remaining conditions (1) to (3) of Definition 11.15.

Condition (1): Suppose $r \in \widetilde{R}_{m,n}$ is defined on v but not on u. Let $r = r'\mathrm{Z}_w r''$ with $\mathrm{Z} \in \{\mathrm{X}, \mathrm{Y}, \mathrm{XX}, \mathrm{YY}\}$ and $w = (p, a, q)$ for some $p, q \in A^*$ and some $a \in A$ such that $r'\mathrm{Z}_w$ is the shortest prefix of r which is not defined on u. Note that r' cannot be empty because by assumption every factor of v length at most $2n - 1$ is also a factor of u. Let

$$\langle r'\mathrm{Z}_w \rangle := \exists x\colon \left(\lambda(x) = w\right) \wedge \begin{cases} \langle x > r' \rangle & \text{if } \mathrm{Z} = \mathrm{X},\\ \langle x \gg r' \rangle & \text{if } \mathrm{Z} = \mathrm{XX},\\ \langle x < r' \rangle & \text{if } \mathrm{Z} = \mathrm{Y},\\ \langle x \ll r' \rangle & \text{if } \mathrm{Z} = \mathrm{YY} \end{cases}$$

using the formulae from Lemma 11.28. The formula $\langle r'\mathrm{Z}_w \rangle$ is true on a word if and only if $r'\mathrm{Z}_w$ is defined. In particular $\langle r'\mathrm{Z}_w \rangle$ is true on v, and because $\langle r'\mathrm{Z}_w \rangle$ is in $\Sigma^2_{m,n}[<, \mathrm{suc}]$, the assumption yields that $r'\mathrm{Z}_w$ is defined on u. This contradicts the choice of r' and shows that r is defined on u. Thus $\widetilde{R}_{m,n}(v) \subseteq \widetilde{R}_{m,n}(u)$.

Condition (2): Let $d \in \{-1, 0, 1, 2\}$, $r \in \widetilde{R}^{\mathrm{X}}_{m,n}(v)$, and $s \in \widetilde{R}^{\mathrm{X}}_{m-1,n-1}(v) \cup \widetilde{R}^{\mathrm{Y}}_{m,n-1}(v)$. Suppose $r = r'\mathrm{X}_w$ or $r = r'\mathrm{XX}_w$ for some $w = (p, a, q)$; in the former case let $d' = 1$ and in the latter case let $d' = 2$. Define the formula

$$\langle r \leqslant s - d \rangle := \exists x\colon \left(\lambda(x) = w\right) \wedge \langle x \geqslant r' + d' \rangle \wedge \langle x \leqslant s - d \rangle,$$

where we use the formulae from Lemma 11.28; we let $\langle x \geqslant r' + d' \rangle := \top$ whenever r' is empty. The extended label $\lambda(x) = w$ is defined by (8.1). With this we have $\langle r \leqslant s - d \rangle \in \Sigma^2_{m,n}[<, \mathrm{suc}]$. By construction $\langle r \leqslant s - d \rangle$ is true on v if and only if

$r(v) \leqslant s(v) - d$. Using the assumption, we see that then $\langle r \leqslant s - d \rangle$ is true on u and therefore $r(u) \leqslant s(u) - d$.

Condition (3) is left-right symmetric to condition (2). ☐

Proof of Theorem 11.16. (1) ⇒ (2): Lemma 11.22.

(2) ⇒ (3): Let L be defined by some formula in $\mathrm{TL}^+_{m,n}[\mathrm{XXF}, \mathrm{YYP}, \mathrm{X}, \mathrm{Y}]$ and suppose $u \preccurlyeq^R_{m,n} v$ and $v \in L$. Lemma 11.27 yields $u \preccurlyeq^{\mathrm{TL}}_{m,n} v$ and consequently $u \in L$. This shows that L is an $\preccurlyeq^R_{m,n}$-order ideal.

(3) ⇒ (1): Let L be an $\preccurlyeq^R_{m,n}$-order ideal. By Lemma 11.29, L is a $\preccurlyeq^{\mathrm{FO2}}_{m,n}$-order ideal. For a word v let L_v be the intersection of all $\Sigma^2_{m,n}[<, \mathrm{suc}]$-definable languages containing v. This intersection is finite because, up to equivalence, there are only finitely many formulae in $\Sigma^2_{m,n}[<, \mathrm{suc}]$. Hence, the language L_v is $\Sigma^2_{m,n}[<, \mathrm{suc}]$-definable. If $v \in L$, then we have $v \in L_v \subseteq L$. Therefore $L = \bigcup_{v \in L} L_v$ is $\Sigma^2_{m,n}[<, \mathrm{suc}]$-definable because there are only finitely many languages of the form L_v. ☐

12. Decidability of the Quantifier Alternation Hierarchy

This chapter studies quantifier alternation within FO^2 from a computational point of view. A fragment $\mathcal{F}$ is called *decidable* if its definability problem is decidable; *i.e.*, if it is decidable whether a given regular language is definable in $\mathcal{F}$. The central question of this chapter is whether the levels of the alternation hierarchy are decidable in this sense. The means to tackle this question is an algebraic approach to regular languages. We assume the reader familiar with basic concepts of the algebraic language theory introduced in Section 2.3.

Depending on the predicates available, the quantifier alternation hierarchy has several variants. This thesis considers

- quantifier alternation within $\mathrm{FO}^2[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$,
- quantifier alternation within $\mathrm{FO}^2[<, \mathrm{suc}, \mathrm{min}]$,
- quantifier alternation within $\mathrm{FO}^2[<, \mathrm{suc}, \mathrm{max}]$,
- quantifier alternation within $\mathrm{FO}^2[<, \mathrm{suc}]$,
- quantifier alternation within $\mathrm{FO}^2[<]$,

and for each of these alternation hierarchies effective conditions characterizing the half levels Σ^2_m and the full levels FO^2_m are given. In particular, definability is decidable in any given level for any relational signature from above. Note that the fragments $\Sigma^2_{m,n}$ and $\mathrm{FO}^2_{m,n}$ with bounded quantifier depth are not interesting with respect to this decidability question: Every such fragment is trivially decidable because, up to equivalence, there are only finitely many formulae with bounded quantifier depth. Also notice that the ranker characterizations of the previous chapter do not *per se* yield decidability of the alternation hierarchies. Still, they are a convenient means in proving the effective characterizations of this chapter.

The question whether a language is definable in full FO^2 is known to be decidable for all signatures: Thérien and Wilke [TW98] showed that a language is definable in $\mathrm{FO}^2[<]$ if and only if its syntactic monoid is in $\mathbf{DA}$, and by a result of Almeida [Alm96], a language is definable in $\mathrm{FO}^2[<, \mathrm{suc}]$ if and only if its syntactic semigroup is in $\mathbf{LDA}$. Disregarding alternation, minimum and maximum predicates can be eliminated using additional quantifiers. Therefore, $\mathrm{FO}^2[<, \mathrm{suc}]$ and $\mathrm{FO}^2[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$ have the same expressive power. These results show that it is decidable whether a language is in the alternation hierarchy at all. The results in this chapter refine this as follows: On input of a language, it is possible to compute the minimal level required to define this language. The characterizations to accomplish this are of a similar kind as those for full FO^2, imposing an effective criterion on the syntactic semigroup or the syntactic monoid.

The main theorems of this chapter shall now be given, with references to the propositions in the sections to come that obtain them.

For $m \in \{1/2, 1\} + \mathbb{N}$ define sequences of omega-terms U_m and V_m

$$U_{1/2} = e^\omega z e^\omega, \quad U_1 = (e^\omega s f^\omega x_1)^\omega e^\omega s f^\omega (y_1 e^\omega t f^\omega)^\omega, \quad U_m = (U_{m-1} x_m)^\omega U_{m-1} (y_m U_{m-1})^\omega,$$
$$V_{1/2} = e^\omega, \qquad V_1 = (e^\omega s f^\omega x_1)^\omega e^\omega t f^\omega (y_1 e^\omega t f^\omega)^\omega, \quad V_m = (U_{m-1} x_m)^\omega V_{m-1} (y_m U_{m-1})^\omega,$$

over variables $e, f, s, t, x_i, y_i \in \Xi$. The omega-terms U_m and V_m intuitively differ only in the very center. With increasing m, the difference between U_m and V_m becomes more and more screened by idempotents of increasing alphabet.

The omega-terms with fractional index yield an identity for the half levels Σ^2_m, whereas the omega-terms with integer index yield an identity for the full levels FO^2_m:

Theorem 12.1
For $L \subseteq A^+$ and integer $m \geqslant 1$ the following are equivalent:

1. *L is $\Sigma^2_m[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$-definable.*
2. *S_L is in **LDA** and satisfies $U_{m-1/2} \leqslant V_{m-1/2}$.*

Theorem 12.2
For $L \subseteq A^+$ and integer $m \geqslant 1$ the following are equivalent:

1. *L is $\mathrm{FO}^2_m[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$-definable.*
2. *S_L satisfies $U_m = V_m$.*

The identity in Theorem 12.1 for $m = 1$ follows from a result due to Pin and Weil [PW97]. This thesis presents a new proof for this result in Proposition 12.20. The identities for the higher half levels are an original contribution of this thesis. They will be established by Proposition 12.51. Remember that the syntactic semigroup S_L is naturally ordered by the syntactic preorder.

By a result of Almeida, **LDA** is the algebraic limit of the union of all half levels [Alm96]. We cannot omit the requirement that S_L be in **LDA** for the half levels, because the syntactic semigroup of the language $A^* \setminus A^* a c^* a A^*$ over $A = \{a, b, c\}$, for example, is not in **LDA** but satisfies $U_{m-1/2} \leqslant V_{m-1/2}$ for all $m \geqslant 2$.

The identity for the first full level is due to Knast's theorem on dot-depth one languages [Kna83]. A new and self-contained proof for Knast's theorem is given in Proposition 12.30. The identities for the higher full levels are a contribution of this thesis, established by Proposition 12.50. They were published in a slightly more complicated form in [KL13].

For the proof it is most natural to start with the alternation hierarchy over the relational signature $\{<, \mathrm{suc}, \mathrm{min}, \mathrm{max}\}$ for two reasons: First, the characterizations over the signature $\{<\}$ are obtained by a reduction to those over the signature $\{<, \mathrm{suc}, \mathrm{min}, \mathrm{max}\}$ using the neutral letter approach developed in Section 8.2. Second, Corollary 5.12 shows that all levels over the signature $\{<, \mathrm{suc}, \mathrm{min}, \mathrm{max}\}$ define positive $+$-varieties of languages. This syntactic argument does not apply to the other signatures containing the successor predicate, and the levels $1/2$ and 1 actually do not form positive $+$-varieties when the predicates min or max are not available.

Using the methods developed in Chapter 9 it turns out, however, that the absence of the predicates min and max does not restrict the expressive power of the levels $3/2$ and higher because those can specify prefixes and suffixes. This leads to the following.

Proposition 12.3
Let $m \geqslant 2$ be an integer, and let $L \subseteq A^+$.

1. *L is $\Sigma^2_m[<, \mathrm{suc}]$-definable if and only if L is $\Sigma^2_m[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$-definable.*
2. *L is $\mathrm{FO}^2_m[<, \mathrm{suc}]$-definable if and only if L is $\mathrm{FO}^2_m[<, \mathrm{suc}, \mathrm{min}, \mathrm{max}]$-definable.*

This follows from Proposition 12.50 and Proposition 12.51. In particular, the alternation hierarchies over the signature $\{<, \mathrm{suc}, \min\}$ and over the signature $\{<, \mathrm{suc}, \max\}$ coincide with that over $\{<, \mathrm{suc}, \min, \max\}$ for level $3/2$ and above.

For the levels $1/2$ and 1, the unavailability of the minimum and maximum predicates prevents the specification of prefixes and suffixes. This leads to ideals for the first half level, and to Green's relations for the first full level.

Theorem 12.4
For $L \subseteq A^+$ the following are equivalent:

1. *L is definable in $\Sigma_1^2[<, \mathrm{suc}, \min]$.*
2. *S_L satisfies $U_{1/2} \leqslant V_{1/2}$ and $h_L(L)$ is a right ideal.*

Theorem 12.5
For $L \subseteq A^+$ the following are equivalent:

1. *L is definable in $\mathrm{FO}_1^2[<, \mathrm{suc}, \min]$.*
2. *S_L satisfies $U_1 = V_1$ and $h_L(L)$ is a union of $\mathcal{R}$-classes.*

Remember that $h_L \colon A^+ \to S_L$ is the syntactic homomorphism. These characterizations are established in Proposition 12.27 and Proposition 12.37, respectively. For completeness, the variants with only max instead of min are stated explicitly. They follow immediately by left-right symmetry.

Theorem 12.6
For $L \subseteq A^+$ the following are equivalent:

1. *L is definable in $\Sigma_1^2[<, \mathrm{suc}, \max]$.*
2. *S_L satisfies $U_{1/2} \leqslant V_{1/2}$ and $h_L(L)$ is a left ideal.*

Theorem 12.7
For $L \subseteq A^+$ the following are equivalent:

1. *L is definable in $\mathrm{FO}_1^2[<, \mathrm{suc}, \max]$.*
2. *S_L satisfies $U_1 = V_1$ and $h_L(L)$ is a union of $\mathcal{L}$-classes.*

In view of these results, it is hardly surprising that having neither the minimum predicate nor the maximum predicate available leads to two-sided ideals for Σ_1^2 and $\mathcal{J}$-classes for FO_1^2.

Theorem 12.8
For $L \subseteq A^+$ the following are equivalent:

1. *L is definable in $\Sigma_1^2[<, \mathrm{suc}]$.*
2. *S_L satisfies $U_{1/2} \leqslant V_{1/2}$ and $h_L(L)$ is an ideal.*

Theorem 12.9
For $L \subseteq A^+$ the following are equivalent:

1. *L is definable in $\mathrm{FO}_1^2[<, \mathrm{suc}]$.*
2. *S_L satisfies $U_1 = V_1$ and $h_L(L)$ is a union of $\mathcal{J}$-classes.*

These results are obtained in Proposition 12.29 and Proposition 12.40, respectively. Theorems 12.4 to 12.9 are contributions of this thesis, published in [KL11a; KL12a].

When recognizing languages with semigroups, the natural setting is non-empty words: The inverse image of a semigroup homomorphism $h \colon A^+ \to S$ never contains the empty word. Consequently, the natural framework is to evaluate formulae over non-empty words.[1] The next remark clarifies that the above results also hold when *definable over A^+* is used instead of the implicit *definable over A^**.

Remark 12.10
All of the above fragments of Σ_m^2 and FO_m^2 can express non-emptiness of a model, independently of the signature. As the formula $\exists x \top$ shows, it is indeed even expressible in $\Sigma_{1,1}^2$ using no predicate at all. In particular, a language $L \subseteq A^+$ is definable over A^* in any of the above fragments $\mathcal{F}$ if and only if it is definable in $\mathcal{F}$ over A^+. $\diamond$

[1] Remember that formulae are evaluated over A^* which includes the empty word, *cf.* Section 3.1. Therefore, *a priori* the language $L(\varphi)$ defined by φ may contain the empty word.

The only signature not yet considered is $\{<\}$, which leads to monoids instead of semigroups as recognizers. For a monoid the neutral element can be substituted for omega-terms e^ω and f^ω with $e, f \in \Xi$. The omega-terms U_m and V_m thus simplify to

$$U'_{1/2} = z, \quad U'_1 = (sx_1)^\omega s(y_1t)^\omega, \quad U'_m = (U'_{m-1}x_m)^\omega U'_{m-1}(y_m U'_{m-1})^\omega,$$
$$V'_{1/2} = 1, \quad V'_1 = (sx_1)^\omega t(y_1t)^\omega, \quad V'_m = (U'_{m-1}x_m)^\omega V'_{m-1}(y_m U'_{m-1})^\omega,$$

where $m \in \{3/2, 2\} + \mathbb{N}$. These omega-terms yield the following theorems for $\mathrm{FO}^2[<]$.

Theorem 12.11
For $L \subseteq A^*$ and integer $m \geqslant 1$ the following are equivalent:
1. *L is definable in $\Sigma^2_m[<]$.*
2. *M_L is in* **DA** *and satisfies $U'_{m-1/2} \leqslant V'_{m-1/2}$.*

Theorem 12.12
For $L \subseteq A^*$ and integer $m \geqslant 1$ the following are equivalent:
1. *L is definable in $\mathrm{FO}^2_m[<]$.*
2. *M_L satisfies $U'_m = V'_m$.*

The identity for $\Sigma^2_1[<]$ is due to Pin [Pin95], and the identity for $\mathrm{FO}^2_1[<]$ is due to Simon's theorem on piecewise testable languages [Sim75]. Both results are therefore a corollary to these theorems. For the full levels $\mathrm{FO}^2_m[<]$ there are already different identities known [KS12b], whereas the identities for the higher half levels $\Sigma^2_m[<]$ were not previously known in the literature. They will be presented in [FKL14].

The syntactic semigroup and the syntactic monoid of a regular language are finite and computable if the language is given, for instance, as an automaton. Moreover, the recognizing set $h_L(L)$ can also be computed and all algebraic characterizations given in the above theorems can thus be verified effectively for a given regular language. This yields the following decidability corollary.

Corollary 12.13
For $m \geqslant 1$ the definability problem is decidable for all of the following fragments:
- $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$,
- $\Sigma^2_m[<, \mathrm{suc}, \min]$,
- $\Sigma^2_m[<, \mathrm{suc}, \max]$,
- $\Sigma^2_m[<, \mathrm{suc}]$,
- $\Sigma^2_m[<]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \min]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}, \max]$,
- $\mathrm{FO}^2_m[<, \mathrm{suc}]$,
- $\mathrm{FO}^2_m[<]$. □

The alternation hierarchies are in fact decidable in a stronger sense: Not only are their levels decidable for a fixed level, but one can even compute the minimal level in which a given regular language is definable. Note that on input of a number m, the omega-terms U_m and V_m can be computed easily.

Decidability of the full level hierarchy over the signature $\{<\}$ was first shown by Kufleitner and Weil using completely different means [KW12]: They showed that $\mathrm{FO}^2_m[<]$ corresponds to a decidable level of the so-called Trotter-Weil hierarchy.[2] Independently, Krebs and Straubing showed decidability of $\mathrm{FO}^2_m[<]$ by a characterization in terms of omega-term identities [KS12b]. As mentioned above, decidability of Σ^2_1 and FO^2_1 over the relational signatures $\{<, \mathrm{suc}, \min, \max\}$ and $\{<\}$ is well-known [PW97; Kna83; Pin95; Sim75].

[2] See also [Kuf13] for a treatise on the Trotter-Weil hierarchy and its connections to quantifier alternation in two-variable logic.

The organization of the remainder of this chapter is as follows: First necessary algebraic foundations for the latter proofs are given in Section 12.1. Section 12.2 considers the first half level and the first full level of the alternation hierarchies with a successor predicate. For those, the unavailability of the minimum and maximum predicates matters. Section 12.3 establishes the results for the higher levels of the alternation hierarchies with a successor predicate. Section 12.4 finally derives the results concerning the alternation hierarchy over the signature $\{<\}$.

12.1. Algebraic Foundations

Refer to Section 2.3 for basic algebraic terms. The next lemma is an important means to obtain stabilizing idempotents within short factors of a semigroup. An element e is a *stabilizer* from the right of an element x if $x = xe$.

Lemma 12.14 (Stabilizer Lemma)
Let S be a non-empty finite semigroup, and let $x_1, \ldots, x_{|S|} \in S$. There exist an index $i \in \{1, \ldots, |S|\}$ and an idempotent $e \in S$ such that $x_1 \cdots x_i = x_1 \cdots x_i e$.

Proof. Choose some $x_{|S|+1} \in S$ arbitrarily. There exist $i < j \leqslant |S| + 1$ such that $x_1 \cdots x_i = x_1 \cdots x_j$ by the pigeonhole principle. In particular $i \leqslant |S|$. Defining the idempotent $e = (x_{i+1} \cdots x_j)^\omega$ yields $x_1 \cdots x_i = x_1 \cdots x_i e$. □

There is also a left-right symmetric version, which is not stated explicitly, to obtain idempotents stabilizing from the left.

The proofs to come heavily use combinatorial semigroup theory; *i.e.*, a semigroup is investigated by considering preimage words of its elements. One of the most fundamental tools in our approach are the following factorizations of these preimages concerning Green's relations.

Definition 12.15 ($\mathcal{R}$-factorization)
*Let $h\colon A^+ \to S$ be a semigroup homomorphism. The $\mathcal{R}$-*factorization *of a word $u \in A^+$ with respect to h is the unique factorization $u = a_1 u_1 \cdots a_k u_k$ with $k \geqslant 1$, $a_i \in A$, and $u_i \in A^*$ such that for all $i \in \{1, \ldots, k\}$ the following hold:*

1. $h(a_1 u_1 \cdots u_{i-1}) >_{\mathcal{R}} h(a_1 u_1 \cdots u_{i-1} a_i)$, and

2. $h(a_1 u_1 \cdots a_i) \mathrel{\mathcal{R}} h(a_1 u_1 \cdots a_i u_i)$.

*Let $D_{\mathcal{R}}(u)$ be the set of $\mathcal{R}$-*descents*. More precisely, let $D_{\mathcal{R}}(u) = \{p_1, \ldots, p_k\}$, where $p_i = |a_1 u_1 \cdots u_{i-1} a_i|$ is the position of the i^{th} $\mathcal{R}$-descent.*

As usual, the expression $a_1 u_1 \cdots u_{i-1}$ in the preceding definition is to be read as the empty word if $i = 1$. Note that the number of $\mathcal{R}$-descents $|D_{\mathcal{R}}(u)|$ is bounded by the cardinality of S.

There is an intuitive way to think of the $\mathcal{R}$-factorization in terms of the right Cayley graph of S: For a fixed homomorphism h, every word u induces a unique path $x_1 \to x_2 \to \cdots \to x_{|u|}$, where $x_i = h(u[1;i])$ is the image of the prefix of u up to position i and each edge $x_{i-1} \to x_i$ in this path is labeled by $h(u[i])$. Traversing this path, some transitions leave the current strongly connected component of the right Cayley graph. The set $D_{\mathcal{R}}(u)$ specifies precisely those prefixes that enter a new strongly connected component.

The *$\mathcal{L}$-factorization* of $v \in A^+$ is the left-right symmetric factorization $v = v_1 b_1 \cdots v_\ell b_\ell$ with $\ell \geqslant 1$, $b_i \in A$, and $v_i \in A^*$ such that

$$h(b_{i-1} v_i b_i \cdots v_\ell b_\ell) <_{\mathcal{L}} h(v_i b_i \cdots v_\ell b_\ell) \mathrel{\mathcal{L}} h(b_i \cdots v_\ell b_\ell) \quad \text{for all } i.$$

For this factorization let $D_{\mathcal{R}}(u) = \{q_1, \ldots, q_\ell\}$, with $q_i = |b_i v_{i+1} \cdots v_\ell b_\ell|$. A similar intuition connecting the $\mathcal{L}$-factorization and the left Cayley graph holds.

We already mentioned that **LDA** is the algebraic counterpart of $\mathrm{FO}^2[<, \mathrm{suc}]$. This variety of semigroups plays a central role for these hierarchies as it is the upper bound for the alternation hierarchies involving the successor predicate. The following is an algebraic formulation of the most important property of **LDA** for our purposes. Together with its combinatorial reinterpretations below, it allows to control the descents in the above $\mathcal{R}$- and $\mathcal{L}$-factorizations.

Lemma 12.16
Let $S \in \mathbf{LDA}$, let $x, y, z \in S$, and let $e \in S$ be idempotent.

1. *If $xe \mathrel{\mathcal{R}} ye$ in S, then $xe \mathrel{\mathcal{R}} xez$ if and only if $ye \mathrel{\mathcal{R}} yez$.*
2. *If $ex \mathrel{\mathcal{L}} ey$ in S, then $ex \mathrel{\mathcal{L}} zex$ if and only if $ey \mathrel{\mathcal{L}} zey$.*

Proof. Since **LDA** is left-right symmetric, it suffices to show (1). Suppose $xe \mathrel{\mathcal{R}} xez$. Since $ye \mathrel{\mathcal{R}} xe \mathrel{\mathcal{R}} xez$ there exist s, t such that $xe = yes$ and $ye = xezt$. We get $ye = ye(esezte)$. Pumping the factor in the parentheses and using $S \in \mathbf{LDA}$ yields $ye = ye(esezte)^\omega = ye(esezte)^\omega ezte(esezte)^\omega \in yezS$. ☐

The next lemma rephrases this combinatorially. It states that descending in the strongly connected component in the right Cayley graph (or the left Cayley graph) only depends on short factors of the path label.

Lemma 12.17 (LDA Descending Lemma)
Let $h\colon A^+ \to S$ be a homomorphism with $S \in \mathbf{LDA}$, let $u, v \in A^+$ with $|v| \geqslant |S|$, and let $s, t \in A^*$ with $\mathrm{alph}_{|S|+1}(vs) = \mathrm{alph}_{|S|+1}(vt)$.

1. *If $h(u) \mathrel{\mathcal{R}} h(uv)$, then $h(u) \mathrel{\mathcal{R}} h(uvs)$ if and only if $h(u) \mathrel{\mathcal{R}} h(uvt)$.*
2. *If $h(u) \mathrel{\mathcal{L}} h(vu)$, then $h(u) \mathrel{\mathcal{L}} h(svu)$ if and only if $h(u) \mathrel{\mathcal{L}} h(tvu)$.*

Proof. Since **LDA** is left-right symmetric, it suffices to show (1). Assume $h(u) \mathrel{\mathcal{R}} h(uv)$ and $h(u) \mathrel{\mathcal{R}} h(uvs)$. We have to show $h(u) \mathrel{\mathcal{R}} h(uvt)$. This is trivial if t is the empty word. Otherwise factorize $vt = pwz$ so that $|w| < |wz| = |S| + 1$ with $h(w) = h(w)e$ for some idempotent e of S. This is possible by Lemma 12.14. Since vs and vt have the same factors of length $|S| + 1$, we find a factorization $vs = s_1 wzs_2$. Let $x = us_1 w$ and $y = upw$. By induction, $h(u) \mathrel{\mathcal{R}} h(y)$ and thus $h(x)e = h(x) \mathrel{\mathcal{R}} h(y) = h(y)e$. Moreover, $h(x)e \mathrel{\mathcal{R}} h(x)eh(z)$. By Lemma 12.16 we see $h(y)e \mathrel{\mathcal{R}} h(y)eh(z)$. This implies the claim. ☐

Consider $h\colon A^+ \to S$ with $S \in \mathbf{LDA}$. In particular, if $u, v \in A^+$ with $|v| \geqslant |S|$, and $a \in A$, then $h(u) \mathrel{\mathcal{R}} h(uv) >_{\mathcal{R}} h(uva)$ implies $\mathrm{alph}_{|S|+1}(v) \neq \mathrm{alph}_{|S|+1}(va)$. This follows immediately from Lemma 12.17 by choosing s to be the empty word and $t = a$. Intuitively, this means that a descent in the $\mathcal{R}$-class when reading a word from left to right can only happen if there is a new short factor (*i.e.*, of length at most $|S| + 1$) since the last $\mathcal{R}$-descent.

The claim of Theorem 12.2 is that the identity $U_m = V_m$ is the algebraic counterpart of the level $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$. This necessarily implies that semigroups satisfying $U_m = V_m$ have to be in **LDA**. This is shown in the next lemma.

Lemma 12.18
Let $m \geqslant 1$ and S be a finite semigroup. If S satisfies $U_m = V_m$, then $S \in \mathbf{LDA}$.

Proof. Let S be a finite semigroup and let $\omega \geqslant 1$ be an integer such that x^ω is idempotent for all $x \in S$. Let $x, y \in S$ and let $e \in S$ be idempotent. Setting $e_1 = f_1 = s = e$, $x_1 = xey$, $y_1 = x$, $t = y$ we get $U_1 = (exeye)^\omega$ and $V_1 = (exeye)^\omega eye(exeye)^\omega$. Setting all other variables occurring in U_m or in V_m to be e, we see $U_m = (exeye)^\omega$ and $V_m = (exeye)^\omega eye(exeye)^\omega$. Thus, if S satisfies $U_m = V_m$ and $e \in E(S)$, then eSe satisfies the identity $(xy)^\omega = (xy)^\omega y(xy)^\omega$, i.e., $S \in \mathbf{LDA}$. $\square$

The next lemma yields that the identity $U_{1/2} \leqslant V_{1/2}$ describes a variety below **LDA**. Note that this is not true for the identities of the half levels above $1/2$.

Lemma 12.19
Let S be a finite ordered semigroup. If S satisfies $e^\omega z e^\omega \leqslant e^\omega$, then S satisfies $U_1 = V_1$.

Proof. We have $f \geqslant fy(exfy)^{\omega-1}esf$ for all $s, x, y \in S$ and all idempotents $e, f \in S$. Hence $(exfy)^\omega exf(tesf)^\omega \geqslant (exfy)^\omega ex(fy(exfy)^{\omega-1}esf)(tesf)^\omega = (exfy)^\omega esf(tesf)^\omega$. By symmetry $(exfy)^\omega exf(tesf)^\omega \leqslant (exfy)^\omega esf(tesf)^\omega$. $\square$

12.2. The Low Levels of the Alternation Hierarchy with Successor Predicate

The purpose of this section is to give algebraic characterizations for the lowest half level Σ^2_1 and the lowest full level FO^2_1 of the alternation hierarchy within two-variable first-order logic. These results will serve in Section 12.3 as a base case for an induction over the level parameter. For the lowest level, the availability of only two variables does not restrict the expressive power. In this section we thus use Σ_1 instead of Σ^2_1 and FO_1 instead of FO^2_1.

The following inclusion diagram depicts the connection of the expressive power of Σ_1 over several relational signatures.

$$\begin{array}{ccccccc}
\Sigma_1[<] & \subsetneq & & \subsetneq & \Sigma_1[<, \mathrm{suc}, \min] & \subsetneq & \\
& & \Sigma_1[<, \mathrm{suc}] & & & & \Sigma_1[<, \mathrm{suc}, \min, \max] \\
\Sigma_1[\mathrm{suc}] & \subsetneq & & \subsetneq & \Sigma_1[<, \mathrm{suc}, \max] & \subsetneq &
\end{array}$$

For the first full level FO_1 we get the following inclusion diagram for the expressive power over several relational signatures.

$$\begin{array}{ccccccc}
\mathrm{FO}_1[<] & \subsetneq & & \subsetneq & \mathrm{FO}_1[<, \mathrm{suc}, \min] & \subsetneq & \\
& & \mathrm{FO}_1[<, \mathrm{suc}] & & & & \\
\mathrm{FO}_1[\mathrm{suc}] & \subsetneq & & \subsetneq & \mathrm{FO}_1[<, \mathrm{suc}, \max] & \subsetneq & \mathrm{FO}_1[<, \mathrm{suc}, \min, \max] \\
& \subsetneq & & \mathrm{FO}_1[\mathrm{suc}, \min, \max] = \mathrm{FO}[\mathrm{suc}] & & \subsetneq &
\end{array}$$

Let us shortly summarize what is already known about the various fragments with respect to decidability. For $\Sigma_1[<]$, $\Sigma_1[\mathrm{suc}]$, and $\mathrm{FO}_1[\mathrm{suc}]$ decidability is due to Pin [Pin95;

Pin05]. Pin and Weil showed that $\Sigma_1[<, \mathrm{suc}, \min, \max]$ is decidable via the identity $e^\omega z e^\omega \leqslant e^\omega$ [PW97]. The fragments $\mathrm{FO}_1[<]$ and $\mathrm{FO}_1[<, \mathrm{suc}, \min, \max]$ are decidable by results due to Simon and Knast, respectively [Sim75; Kna83]. Thomas [Tho82] showed that $\mathrm{FO}_1[\mathrm{suc}, \min, \max]$ already has the full expressive power of $\mathrm{FO}[\mathrm{suc}]$. Both fragments define precisely so-called locally threshold testable languages, which are known to be decidable; *cf. e.g.* [Str94, Theorem VI.3.1].

For the fragments over the full signature $\{<, \mathrm{suc}, \min, \max\}$ we contribute new proofs for Pin and Weil's result on Σ_1 in Section 12.2.1 and for Knast's result on FO_1 in Section 12.2.3. As depicted above, the levels Σ_1 and FO_1 drop in expressive power when minimum or maximum predicates are unavailable. Consequently, it makes sense to consider the signatures $\{<, \mathrm{suc}, \min\}$, $\{<, \mathrm{suc}, \max\}$, and $\{<, \mathrm{suc}\}$. We give effective algebraic characterizations for both Σ_1 and FO_1 over any of these signatures in Sections 12.2.2 and 12.2.4. Decidability of these fragments was not previously known. The contents of this section have been published in [KL11a; KL12a].

12.2.1. Existential First-Order Logic and Dot-Depth 1/2

A *monomial* of degree n is a language of A^+ of the form $w_1 A^* w_2 \cdots A^* w_\ell$, where $w_i \in A^*$ with $|w_1 \cdots w_\ell| = n$. A language has *dot-depth 1/2* if it is a positive Boolean combination of monomials. The main result of this section is as follows.

Proposition 12.20 (Pin and Weil [PW97], Thomas [Tho82])
Let $L \subseteq A^+$. The following are equivalent:

1. *L is definable in $\Sigma_1^2[<, \mathrm{suc}, \min, \max]$.*
2. *L is definable in $\Sigma_1[<, \mathrm{suc}, \min, \max]$.*
3. *L is a finite union of monomials $w_1 A^* w_2 \cdots A^* w_\ell$.*
4. *S_L satisfies $e^\omega z e^\omega \leqslant e^\omega$.*

The remainder of this section is devoted to the proof of this proposition. We start by the implication from logic to the description in terms of monomials.

Lemma 12.21
If $L \subseteq A^+$ is definable by a $\Sigma_1[<, \mathrm{suc}, \min, \max]$-sentence with a total of n quantifiers, then L is a finite union of languages $w_1 A^+ w_2 \cdots A^+ w_\ell$ with $|w_1 \cdots w_\ell| \leqslant n$. In particular, L is a finite union of monomials of the form $w_1 A^* w_2 \cdots A^* w_\ell$ of degree at most $2n + 1$.

Proof. Let $L = L(\varphi)$ for $\varphi \in \Sigma_1[<, \mathrm{suc}, \min, \max]$. Renaming variables so that no two different quantifiers bind the same variable, we can rewrite φ by the usual procedure to get an equivalent formula in prenex normal form. So we may assume that

$$\varphi \;:=\; \exists x_1 \cdots \exists x_n \; \psi(x_1, \ldots, x_n)$$

with a quantifier-free formula ψ. Suppose $u \models \varphi$, *i.e.*, there exist positions $j_1, \ldots, j_n$ of u such that $u, j_1, \ldots, j_n \models \psi(x_1, \ldots, x_n)$. We say that a position j of u is *marked* if $j = j_i$ for some i. Assume that the first and the last position of u are marked. Let $u = w_1 u_1 w_2 \cdots u_{\ell-1} w_\ell$ for $u_i \in A^+$ such that the factors w_i consist of the marked positions. Now, for $P_u = w_1 A^+ w_2 \cdots A^+ w_\ell$ we have $|w_1 \cdots w_\ell| \leqslant n$ and $u \in P_u$. Moreover, $P_u \subseteq L(\varphi)$ since the satisfying assignment of u can be adapted to all $v \in P_u$. Suppose now that the first position is marked but the last position is not marked, and let $u = w_1 u_1 \cdots w_\ell u_\ell$ for $u_i \in A^+$ such that the factors w_i consist of the marked positions. For the language $P_u = w_1 A^+ \cdots w_\ell A^+$ we have $|w_1 \cdots w_\ell| \leqslant n$ and $u \in P_u \subseteq L(\varphi)$. In case the last position is marked but the first position is not, we take the language $P_u = A^+ w_1 \cdots A^+ w_\ell$ and if both positions are not marked we take $P_u = A^+ w_1 \cdots A^+ w_\ell A^+$. It follows $L(\varphi) = \bigcup_{u \models \varphi} P_u$ and this union is finite since there are only finitely many languages of the form $w_1 A^+ w_2 \cdots A^+ w_\ell$ with $|w_1 \cdots w_\ell| \leqslant n$.

Making all possibilities for the first letter between the markers explicit, we see that $w_1 A^+ \cdots w_{\ell-1} A^+ w_\ell$ is a union of monomials of the form $w_1 a_1 A^* \cdots w_{\ell-1} a_{\ell-1} A^* w_\ell$ for $a_1, \ldots, a_{\ell-1} \in A$. At the worst $w_1 = w_\ell = 1$ and $w_i \in A$ for $1 < i < \ell$. $\square$

The following gives the reverse implication back from monomials to logic. On our way, we also handle the other signatures.

Lemma 12.22
Let $P = w_1 A^* w_2 \cdots A^* w_\ell$, where $w_i \in A^+$, and let $n = |w_1 \cdots w_\ell|$.

1. *P is definable in $\Sigma^2_{1,n}[<, \mathrm{suc}, \min, \max]$.*
2. *PA^* is definable in $\Sigma^2_{1,n}[<, \mathrm{suc}, \min]$.*
3. *A^*PA^* is definable in $\Sigma^2_{1,n}[<, \mathrm{suc}]$.*

Proof. We introduce the following formula, which extends the label predicate. For a formula $\psi(x)$ the formula $\lambda(x) = (\vec{w}_1, \ldots, \vec{w}_\ell, \psi)$ with free variable x requires that, starting from the position x, there be occurrences of factors $w_1, \ldots, w_\ell$ in this order such that on the last position j of the occurrence of w_ℓ the formula ψ holds. More precisely, for every $u \in A^+$ and every position i of u we have $u, i \models (\lambda(x) = (\vec{w}_1, \ldots, \vec{w}_\ell, \psi))$ if and only if there exist $v_i \in A^+$, $v_\ell \in A^*$, and $1 \leqslant i < \ell$ such that $u[i; |u|] = w_1 v_1 \cdots w_\ell v_\ell$ and $u, j \models \psi(x)$ for $j = i + |w_1 v_1 \cdots w_\ell| - 1$. Formally, for $w = aw'$ with $w' \in A^*$ we define $\lambda(x) = (\vec{w}, \psi)$ to be

$$\lambda(x) = a \wedge \begin{cases} \exists y \, \big(\mathrm{suc}(x, y) \wedge \lambda(y) = (\vec{w}', \psi)\big) & \text{if } |w'| > 0, \\ \psi(x) & \text{else.} \end{cases}$$

Extend this to sequences by $(\lambda(x) = (\vec{w}_1, \ldots, \vec{w}_\ell, \psi)) := (\lambda(x) = (\vec{w}_1, \psi'))$, where

$$\psi'(x) \; := \; \exists y > x \, \big(\neg \, \mathrm{suc}(x, y) \wedge \lambda(y) = (\vec{w}_2, \ldots, \vec{w}_\ell, \psi(y))\big).$$

As usual, $\psi(y)$ is obtained by interchanging the variables x and y in $\psi(x)$. With this formula let

$$\begin{aligned} \varphi_1 \; &:= \; \exists x \; \min(x) \wedge \lambda(x) = (\vec{w}_1, \ldots, \vec{w}_\ell, \max(x)), \\ \varphi_2 \; &:= \; \exists x \; \min(x) \wedge \lambda(x) = (\vec{w}_1, \ldots, \vec{w}_\ell, \top), \\ \varphi_3 \; &:= \; \exists x \; \lambda(x) = (\vec{w}_1, \ldots, \vec{w}_\ell, \top). \end{aligned}$$

Clearly, $\varphi_1 \in \Sigma^2_{1,n}[<, \mathrm{suc}, \min, \max]$ defines P, the sentence $\varphi_2 \in \Sigma^2_{1,n}[<, \mathrm{suc}, \min]$ defines PA^*, and the sentence $\varphi_3 \in \Sigma^2_{1,n}[<, \mathrm{suc}]$ defines A^*PA^*. $\square$

The next lemma shows that the syntactic semigroup of monomials satisfy the identity, establishing the direction from monomials to algebra.

Lemma 12.23
If $P \subseteq A^+$ is a monomial of the form $w_1 A^* w_2 \cdots A^* w_\ell$, then S_P satisfies $e^\omega z e^\omega \leqslant e^\omega$.

Proof. Consider a monomial $P = w_1 A^* w_2 \cdots A^* w_\ell$, and let $m \geqslant \max\{|w_1|, \ldots, |w_\ell|\}$. Let $x, y \in A^+$ and $u, v \in A^*$ be such that $ux^m v \in P$. Let i be maximal such that $ux^m \in w_1 A^* \cdots w_i A^* = Q_i$, and let j be minimal such that $x^m v \in A^* w_j \cdots A^* w_\ell = R_j$. By the choice of m we have $j \leqslant i+1$. Therefore, $ux^m y x^m v \in Q_i R_j \subseteq P$. ☐

The following lemma is the main combinatorial ingredient of the proof. It constructs a factorization that splits a given word at short factors which are stabilized in pairs by the same idempotent.

Lemma 12.24
Let $h\colon A^+ \to S$ be a homomorphism to a finite semigroup. For every $u \in A^+$ there exists a factorization $u = x_1 u_1 y_1 \cdots x_\ell u_\ell y_\ell s$ such that the following hold:

1. *$0 \leqslant \ell \leqslant |E(S)|$ and $|x_1 y_1 \cdots x_\ell y_\ell s| < (2|E(S)|+1)|S|$.*
2. *$u_i, s \in A^*$ and $x_i, y_i \in A^+$ and $|y_i| \leqslant |S|$ for all $i \in \{1, \ldots, \ell\}$.*
3. *For all $i \in \{1, \ldots, \ell\}$ there is $e_i \in E(S)$ with $h(x_i) = h(x_i)e_i$ and $h(y_i) = h(y_i)e_i$.*

Proof. Let $E(v)$ for $v \in A^*$ consist of all $e \in E(S)$ such that there exists a factor $x \in A^+$ of v with $|x| \leqslant |S|$ and $h(xe) = h(x)$. We replace condition (1) by the statement

$$0 \leqslant \ell \leqslant |E(u)| \text{ and } |x_1 y_1 \cdots x_\ell y_\ell s| < (2|E(u)|+1)|S|$$

and prove this stronger result by induction on $|E(u)|$.

Suppose $|E(u)| = 0$. By Lemma 12.14 we have $|u| < |S|$. Hence, we can choose $\ell = 0$ and $s = u$. If $|E(u)| \geqslant 1$, then Lemma 12.14 yields a non-empty prefix x of u with $|x| \leqslant |S|$ such that $h(x)e = h(x)$ for some idempotent $e \in E(S)$. Let $u = xu'$. We have to distinguish two cases.

The first case is $e \notin E(u')$. By induction, we can write $u' = x_1 u_1 y_1 \cdots x_\ell u_\ell y_\ell s$ with $\ell \leqslant |E(u')| < |E(u)|$ and $|x_1 y_1 \cdots x_\ell y_\ell s| < 2|S| \cdot |E(u')| + |S|$ satisfying conditions (2) and (3). If $\ell \geqslant 1$, then $u = (xx_1) u_1 y_1 \cdots x_\ell u_\ell y_\ell s$ is a desired factorization of u. If $u' = s$, then the factorization is $u = xs$ with $\ell = 0$.

The second case is $e \in E(u')$. Let $u' = u_0 y_0 u''$ such that $y_0 \in A^+$, $|y_0| \leqslant |S|$, $h(y_0)e = h(y_0)$ and $e \notin E(u'')$; i.e., we let y_0 be the last short factor of u' which is stabilized by e. By induction, there exists a factorization $u'' = x_1 u_1 y_1 \cdots x_\ell u_\ell y_\ell s$. Now, $u = x_0 u_0 y_0 \cdots x_\ell u_\ell y_\ell s$ with $x_0 = x$ is a factorization of u of the desired form. ☐

The next lemma is auxiliary. Phrased in more algebraic terms it shows that semigroups satisfying $U_1 = V_1$ (and by Lemma 12.19 also those satisfying $e^\omega z e^\omega \leqslant e^\omega$) are what is called locally $\mathcal{R}$-trivial.

Lemma 12.25
Let S satisfy $U_1 = V_1$, let $e \in E(S)$, and let $u, v \in S$. If $ue = u \mathrel{\mathcal{R}} v = ve$, then $u = v$.

Proof. Let $x, y \in S$ be such that $v = ux$ and $u = vy$. The identity $U_1 = V_1$ yields $v = u(exey)^\omega exe(eeee)^\omega = u(exey)^\omega eee(eeee)^\omega = u$. ☐

The following is the heart of the proof. It shows that it is possible to pass from a semigroup satisfying the identity $e^\omega z e^\omega \leqslant e^\omega$ to monomials.

Proposition 12.26
If $L \subseteq A^+$ is recognized by an ordered semigroup S that satisfies $e^\omega z e^\omega \leqslant e^\omega$, then L is a finite union of languages $w_1 A^* w_2 \cdots A^* w_n$ with $|w_1 \cdots w_n| < 2|S|^3 + |S|^2$ and $n \leqslant |S|^2$.

Proof. Let $h\colon A^+ \to S$ be a homomorphism recognizing L. The order ideal of S generated by a subset $P \subseteq S$ is $\downarrow P = \{x \in S \mid x \leqslant y \text{ for some } y \in P\}$. We define the *depth* of a word $u \in A^+$ to be $d(u) = |\{s \in S \mid h(u) \leqslant_{\mathcal{R}} s\}|$. For every $u \in A^+$ we are going to construct a monomial $P_u = w_1 A^* w_2 \cdots A^* w_n$ with $n \leqslant d(u)|S|$ and $|w_1 \cdots w_n| < 2d(u)|S|^2 + d(u)|S|$ such that $u \in P_u \subseteq h^{-1}\big(\downarrow h(u)\big)$. With this we have $L = \bigcup_{u \in L} P_u$, which is a finite union as there are only finitely many such monomials.

Write $u = vw$ with $v \in A^*$ and $w \in aA^*$ such that $h(va) \mathrel{\mathcal{R}} h(u)$ and either $v = 1$ or $h(v) >_{\mathcal{R}} h(va)$. By Lemma 12.24 we find a factorization $w = x_1 w_1 y_1 \cdots x_\ell w_\ell y_\ell s$ such that $|x_1 y_1 \cdots x_\ell y_\ell s| < 2|S|^2 + |S|$ and for all $i \in \{1, \ldots, \ell\}$ there exists an idempotent e_i with $h(x_i)e_i = h(x_i)$ and $h(y_i)e_i = h(y_i)$. Using Lemma 12.19 and Lemma 12.25, we get $h(u) = h(vw) = h(vx_1 \cdots x_\ell)$.

If $v = 1$, then we set $P_u = x_1 A^* y_1 \cdots x_\ell A^* y_\ell$. The degree of P_u is less than $2|S|^2 + |S|$. By construction, we have $u = w \in P_u$. Consider $w' \in P_u$ with $w' = x_1 w_1' y_1 \cdots x_\ell w_\ell' y_\ell s$. We have $h(x_i) = h(x_i)e_i \geqslant h(x_i)e_i h(w_i' y_i)e_i = h(x_i w_i' y_i)$ since $eze \leqslant e$ for all $z \in S$ and all $e \in E(S)$. Therefore, $h(u) = h(x_1 \cdots x_\ell s) \geqslant h(w')$. This shows $P_u \subseteq h^{-1}\big(\downarrow h(u)\big)$.

Let now $v \neq 1$. Then $d(v) < d(u)$ and thus, by induction, there exists a monomial P_v with $v \in P_v \subseteq h^{-1}(\downarrow h(v))$ of degree less than $2d(u)|S|^2 + d(u)|S| - 2|S|^2 - |S|$. We set $P_u = P_v\, x_1 A^* y_1 \cdots x_\ell A^* y_\ell$. The degree of P_u is less than $2d(u)|S|^2 + d(u)|S|$. Note that $u \in P_u$. Suppose $v'w' \in P_u$ with $v' \in P_v$ and $w' = x_1 w_1' y_1 \cdots x_\ell w_\ell' y_\ell s$. Then we have $h(v') \leqslant h(v)$, and we see $h(x_i) \geqslant h(x_i w_i' y_i)$ as before. Therefore, $h(x_1 \cdots x_\ell s) \geqslant h(w')$ and $h(u) = h(vx_1 \cdots x_\ell s) \geqslant h(v'w')$. □

The main feature of this proposition is that the degree $|w_1 \cdots w_n|$ is polynomially bounded, whereas the original proof of Pin and Weil yields an exponential bound. It is also on a more elementary level as it does not use so-called factorization forests.

Proof of Proposition 12.20. The implication (1) ⇒ (2) is trivial. (2) ⇒ (3): This is Lemma 12.21, and (3) ⇒ (1) follows from property (1) of Lemma 12.22 and the fact that $\Sigma_1[<, \mathrm{suc}, \min, \max]$ is closed under union. The implication (3) ⇒ (4) follows from Lemma 12.23. Finally, (4) ⇒ (3) follows immediately from Proposition 12.26. □

12.2.2. Existential First-Order Logic without min and max

We turn to existential first-order logic without minimum and maximum predicates. The unavailability of the maximum predicate prohibits the specification of suffixes, leading to right ideals for the recognizing set.

Proposition 12.27
Let $L \subseteq A^+$. The following are equivalent:

1. *L is definable in $\Sigma_1^2[<, \mathrm{suc}, \min]$.*
2. *L is definable in $\Sigma_1[<, \mathrm{suc}, \min]$.*
3. *L is a finite union of monomials $w_1 A^* \cdots w_\ell A^*$.*
4. *S_L satisfies $e^\omega z e^\omega \leqslant e^\omega$ and $h_L(L)$ is a right ideal of S_L.*

Proof. (1) ⇒ (2): This is trivial.

(2) ⇒ (3): Let $L = L(\varphi)$ for $\varphi \in \Sigma_1[<, \mathrm{suc}, \min]$. The language L is a finite union of monomials $w_1 A^* w_2 \cdots A^* w_\ell$ by Proposition 12.20. If $u \in L(\varphi)$, then the

same assignment of the variables that makes φ true on u also satisfies φ on uv for every $v \in A^*$. Therefore, $LA^* \subseteq L$. Since concatenation distributes over union, *i.e.*, $(P \cup Q)A^* = PA^* \cup QA^*$, it follows that L is a finite union of monomials $w_1 A^* \cdots w_\ell A^*$.

(3) $\Rightarrow$ (1): This is statement (2) in Lemma 12.22.

(3) $\Rightarrow$ (4): The syntactic semigroup S_L satisfies $e^\omega z e^\omega \leqslant e^\omega$ by Proposition 12.20. The language L is a right ideal of A^+. Therefore, $h_L(L)$ is a right ideal of S_L, because the image of a right ideal under a surjective homomorphism is again a right ideal.

(4) $\Rightarrow$ (3): The language L is a union of monomials of the form $P = w_1 A^* w_2 \cdots A^* w_\ell$ by Proposition 12.20. Suppose $P \subseteq L$ for a monomial of this form. Right ideals are closed under inverse homomorphisms, and we have $PA^* \subseteq L$. Therefore, L is a union of monomials of the form $w_1 A^* w_2 \cdots A^* w_\ell A^*$. □

The following is the variant without minimum predicate. It follows immediately by left-right symmetry.

Corollary 12.28
Let $L \subseteq A^+$. The following are equivalent:

1. *L is definable in $\Sigma_1^2[<, \mathrm{suc}, \max]$.*
2. *L is definable in $\Sigma_1[<, \mathrm{suc}, \max]$.*
3. *L is a finite union of monomials $A^* w_1 \cdots A^* w_\ell$.*
4. *S_L satisfies $e^\omega z e^\omega \leqslant e^\omega$ and $h_L(L)$ is a left ideal of S_L.* □

With neither minimum nor maximum predicates, the recognizing set must be a two-sided ideal. The proof can be adapted straightforwardly and is thus omitted.

Proposition 12.29
Let $L \subseteq A^+$. The following are equivalent:

1. *L is definable in $\Sigma_1^2[<, \mathrm{suc}]$.*
2. *L is definable in $\Sigma_1[<, \mathrm{suc}]$.*
3. *L is a finite union of monomials $A^* w_1 \cdots A^* w_\ell A^*$.*
4. *S_L satisfies $e^\omega z e^\omega \leqslant e^\omega$ and $h_L(L)$ is an ideal of S_L.* □

12.2.3. Alternation-Free First-Order Logic and Dot-Depth One

A language has *dot-depth one* if it is a Boolean combination of monomials. This section provides a new proof of the following proposition on dot-depth one languages.

Proposition 12.30 (Knast [Kna83], Thomas [Tho82])
Let $L \subseteq A^+$. The following are equivalent:

1. *L is definable in $\mathrm{FO}_1^2[<, \mathrm{suc}, \min, \max]$.*
2. *L is definable in $\mathrm{FO}_1[<, \mathrm{suc}, \min, \max]$.*
3. *L is a Boolean combination of monomials $w_1 A^* w_2 \cdots A^* w_\ell$.*
4. *S_L satisfies $U_1 = V_1$.*

The equivalence of (3) and (4) is due to Knast [Kna83], the equivalence of (2) and (3) is due to Thomas [Tho82]. The original proof of Knast's result as well as the simpler proof given latter by Thérien [Thé88] both use extensions of semigroups called *finite categories*, *cf.* [Til87]. Our proof, in contrast, is rather combinatorial and, as for dot-depth $1/2$, its main benefit are more explicit bounds on the size of the monomials. A similar technique allowed to obtain a characterization for languages of dot-depth one over infinite words [KL11b]. The remainder of this section gives the proof of the proposition.

Suppose u and v are words that have both $x, y \in A^+$ as factors. Consider the concrete occurrence of these factors given by the decomposition $u = u_0 x u_1 = u'_0 y u'_1$ and $v = v_0 x v_1 = v'_0 y v'_1$ for some $u_i, u'_i, v_i, v'_i \in A^*$. Let $\Delta_u = |u| - |u_0 u'_1|$ and $\Delta_v = |v| - |v_0 v'_1|$. We say that *the relative order of x and y is the same in u and v* if one of the following conditions applies:

- $\Delta_u > |xy|$ and $\Delta_v > |xy|$, *i.e.*, in each of the words u and v, all positions of x are to the left of all positions of y,
- $\Delta_u < 0$ and $\Delta_v < 0$, *i.e.*, in each of the words u and v, all positions of x are to the right of all positions of y,
- $\Delta_u = \Delta_v$, *i.e.*, if none of the previous conditions applies, then the occurrences of x and y have the same overlap in both words u and v.

The following lemma about relative orders is the main combinatorial ingredient for the present proof. It generalizes an idea of Klíma [Klí11] to factors of words. The determinacy mechanism is similar to unambiguous interval logic with look-around [LPS10].

Lemma 12.31
Let k, ℓ be positive integers, let $x_i, y_i, u_i, u'_i, v_i, v'_i \in A^+$ and $u_k, v_k, u'_1, v'_1 \in A^*$, and let

$$u = x_1 u_1 \cdots x_k u_k = u'_1 y_1 \cdots u'_\ell y_\ell$$
$$v = x_1 v_1 \cdots x_k v_k = v'_1 y_1 \cdots v'_\ell y_\ell$$

such that $x_1 u_1 \cdots x_k$ (respectively, $x_1 v_1 \cdots x_k$) is the shortest prefix of u (respectively, v) in $x_1 A^+ x_2 \cdots A^+ x_k$, and $y_1 \cdots u'_\ell y_\ell$ (respectively, $y_1 \cdots v'_\ell y_\ell$) is the shortest suffix of u (respectively, v) in $y_1 A^+ y_2 \cdots A^+ y_\ell$. If u and v are contained in the same languages $w_1 A^+ w_2 \cdots A^+ w_n$ with $n \leqslant k + \ell$ and $|w_1 \cdots w_n| \leqslant |x_1 \cdots x_k\, y_1 \cdots y_\ell|$, then the relative orders of x_k and y_1 are the same in u and v.

Proof. Let $\Delta_u = |u| - |x_1 \cdots u_{k-1}| - |u'_2 \cdots y_\ell|$ and $\Delta_v = |v| - |x_1 \cdots v_{k-1}| - |v'_2 \cdots y_\ell|$. First suppose that $\Delta_u > |x_k y_1|$. In this case $x_1 u_1 \cdots x_k$ is a proper prefix of u'_1. Therefore $u \in x_1 A^+ \cdots x_k A^+ y_1 \cdots A^+ y_\ell$. This implies $v \in x_1 A^+ \cdots x_k A^+ y_1 \cdots A^+ y_\ell$ and $\Delta_v > |x_k y_1|$. By symmetry we conclude that $\Delta_u > |x_k y_1|$ if and only if $\Delta_v > |x_k y_1|$.

Let now $0 \leqslant \Delta_u \leqslant |x_k y_1|$. We can assume $\Delta_v \leqslant |x_k y_1|$. Now, u is contained in $P = x_1 A^+ \cdots x_i A^+ z A^+ y_j \cdots A^+ v_\ell$, where z is the factor of u comprising all $x_{i+1}, \ldots, x_k$ that are overlapping with (or adjacent to) y_1 and all $y_1, \ldots, y_{j-1}$ which are overlapping with (or adjacent to) x_k. Since $v \in P$, we conclude that x_k is not further to the right of y_1 in the word u as in the word v, *i.e.*, we have $\Delta_u \leqslant \Delta_v$. By symmetry, this shows that $0 \leqslant \Delta_u \leqslant |x_k y_1|$ if and only if $0 \leqslant \Delta_v \leqslant |x_k y_1|$. Moreover, if $0 \leqslant \Delta_u \leqslant |x_k y_1|$, then $\Delta_u = \Delta_v$.

We see that $\Delta_u < 0$ if and only if $\Delta_v < 0$ by the above two cases. This shows that x_k and y_1 have the same relative order in u and v. $\square$

The next lemma is the crucial property of semigroups satisfying $U_1 = V_1$. It shows that, in certain contexts, an element x may be replaced by s without changing the value in the semigroup.

Lemma 12.32
Let S be a semigroup satisfying $U_1 = V_1$, let $u, v, x, s \in S$, and let $e, f \in S$ be idempotent. If $u \mathrel{\mathcal{R}} uexf$ and $esfv \mathrel{\mathcal{L}} v$, then $uexfv = uesfv$.

Proof. Since $u \mathrel{\mathcal{R}} uexf$ and $v \mathrel{\mathcal{L}} esfv$ there exist $y, t \in S$ such that $u = uexfy$ and $v = tesfv$. In particular, $u = u(exfy)^\omega$ and $v = (tesf)^\omega v$. Therefore, we can conclude $uexfv = u(exfy)^\omega exf(tesf)^\omega v = u(exfy)^\omega esf(tesf)^\omega v = uesfv$, where the second equality uses the identity $U_1 = V_1$. $\square$

The following is the key to the proof. It states that the preimage of any element of a semigroup satisfying $U_1 = V_1$ is a union of monomials of small degree.

Proposition 12.33
Let $L \subseteq A^+$ be recognized by a homomorphism $h\colon A^+ \to S$ to a finite semigroup S satisfying $U_1 = V_1$, and let $u, v \in A^+$. If u and v are contained in the same languages $w_1 A^+ w_2 \cdots A^+ w_n$ with $n \leqslant 2|S|$ and $|w_1 \cdots w_n| \leqslant 4|S|^2 - 2|S|$, then $h(u) = h(v)$.

Proof. The outline is as follows. First, we take the $\mathcal{R}$-factorization of u and the $\mathcal{L}$-factorization of v, and incorporate the factors surrounding the descents. Then, we transfer the factorization of u to v and *vice versa* such that the respective orders of the factors in u and v are the same. Finally, we use the resulting factorization to transform v into u by a sequence of h-invariant substitutions.

Consider the $\mathcal{R}$-factorization $u = a_1 u_1 \cdots a_k u_k$ with $a_i \in A$ such that

$$h(a_1 u_1 \cdots a_i) \mathrel{\mathcal{R}} h(a_1 u_1 \cdots a_i u_i) >_{\mathcal{R}} h(a_1 u_1 \cdots a_i u_i a_{i+1}) \text{ for all } i.$$

We have $k \leqslant |S|$. Let j_i be the position of a_i in the above factorization. We color red all positions of u in all the intervals $[\, j_i - |S|;\, j_i + |S| - 1\,]$. In particular, the a_i-positions j_i are red. Moreover in general, there is a neighborhood of size $2|S|$ around each a_i which contains only red positions. (In the worst case, a_1 is the sole exception.) Hence, there are at most $2|S|^2 - |S|$ red positions in u. Let R_i be the i^{th} consecutive factor consisting of red positions. Then $u = R_1 u'_1 \cdots R_{k'} u'_{k'}$ for some $u'_i \in A^+$, $i < k'$, and $u'_{k'} \in A^*$. Note that $k' \leqslant k$ because some intervals could overlap. We have $S \in \mathbf{LDA}$ by Lemma 12.18, and, with Lemma 12.17, we see that for each i the word $R_1 u'_1 \cdots u'_{i-1} R_i$ is the shortest prefix of u contained in $R_1 A^+ \cdots A^+ R_i$.

Symmetrically, let $v = v_1 b_1 \cdots v_\ell b_\ell$ with $b_i \in A$ be the $\mathcal{L}$-factorization such that

$$h(b_{i-1} v_i b_i \cdots v_\ell b_\ell) <_{\mathcal{L}} h(v_i b_i \cdots v_\ell b_\ell) \mathrel{\mathcal{L}} h(b_i \cdots v_\ell b_\ell) \text{ for all } i.$$

Let j'_i be the position of b_i in the above factorization. We color blue all positions of v in all the intervals $[\, j'_i - |S| + 1;\, j'_i + |S|\,]$. As before, there are at most $2|S|^2 - |S|$ blue positions. Let B_i be the i^{th} consecutive factor of blue positions. Then $v = v'_1 B_1 \cdots v'_{\ell'} B_{\ell'}$ for $\ell' \leqslant |S|$ and some $v'_i \in A^+$, $i > 1$, and $v'_1 \in A^*$. As before, $B_i v'_{i+1} \cdots v'_{\ell'} B_{\ell'}$ is the shortest suffix of v contained in $B_i A^+ \cdots A^+ B_{\ell'}$.

Next, we transfer the red positions of u to v, and we transfer the blue positions of v to u. By assumption $v \in R_1 A^+ \cdots R_{k'} A^+$ and therefore, there exists a factorization $v = R_1 v''_1 \cdots R_{k'} v''_{k'}$ such that $R_1 v''_1 \cdots v''_{i-1} R_i$ is the shortest prefix of v contained in

$R_1A^+ \cdots A^+R_i$. We color the positions of the R_i's in v red. Similarly, there exists a factorization $u = u''_1 B_1 \cdots u''_{\ell'} B_{\ell'}$ such that $B_i u''_{i+1} \cdots u''_{\ell'} B_{\ell'}$ is the shortest suffix of u contained in $B_i A^+ \cdots A^+ B_{\ell'}$. We color the positions of the B_i's in u blue. Now, colored positions in u and v are either red or blue or both. By Lemma 12.31, the colored positions in u have the same order as the colored positions in v. Thus if w_i is the i^{th} consecutive factor of colored (*i.e.*, red or blue) positions, then

$$\begin{aligned} u &= w_1 x_1 \cdots w_{n-1} x_{n-1} w_n, \\ v &= w_1 s_1 \cdots w_{n-1} s_{n-1} w_n. \end{aligned}$$

The next step is the construction of idempotent stabilizers near the beginning and near the end of each w_i. We do this from the inside to the outside by considering the first and the last $|S|$ letters in every word w_i: By Lemma 12.14 and its left-right dual, there exist idempotents $e_1, \ldots, e_{n-1} \in E(S)$ and $f_2, \ldots, f_n \in E(S)$ such that each w_i admits a factorization $w_i = p_i r'_i r''_i q_i$ with $|p_i r'_i| = |S|$, $|r'_i| \geqslant 1$ and $|q_i| \leqslant |S| - 1$ satisfying

$$\begin{aligned} h(r'_i) &= f_i\, h(r'_i) \quad \text{for all } 1 < i \leqslant n, \\ h(r''_i) &= h(r''_i)\, e_i \quad \text{for all } 1 \leqslant i < n. \end{aligned}$$

In particular, we can assume $p_1 = q_n = \varepsilon$. Let $x'_i = q_i x_i p_{i+1}$ and $s'_i = q_i s_i p_{i+1}$ for $1 \leqslant i < n$, and let $r_i = r'_i r''_i$ for $1 \leqslant i \leqslant n$. Then

$$\begin{aligned} u &= r_1 x'_1 r_2 \cdots x'_{n-1} r_n, \\ v &= r_1 s'_1 r_2 \cdots s'_{n-1} r_n. \end{aligned}$$

By construction, every position of the $\mathcal{R}$-factorization of u lies within some r''_i. We thus have $h(r_1 x'_1 \cdots r_i) \mathrel{\mathcal{R}} h(r_1 x'_1 \cdots r_i x'_i r'_{i+1}) = h(r_1 x'_1 \cdots r_i) \cdot e_i h(x'_i) f_{i+1} \cdot h(r'_{i+1})$ for all $1 \leqslant i < n$. Therefore, for all $1 \leqslant i < n$ we get

$$h(r_1 x'_1 \cdots r_i) \mathrel{\mathcal{R}} h(r_1 x'_1 \cdots r_i) \cdot e_i h(x'_i) f_{i+1}.$$

A symmetric argument shows

$$h(r_{i+1} \cdots s'_n r_n) \mathrel{\mathcal{L}} e_i h(s'_i) f_{i+1} \cdot h(r_{i+1} \cdots s'_n r_n).$$

By an $(n-1)$-fold application of Lemma 12.32, we obtain

$$\begin{aligned} h(v) &= h(r_1 s'_1 r_2 s'_2 r_3 \cdots s'_{n-1} r_n) \\ &= h(r_1 x'_1 r_2 s'_2 r_3 \cdots s'_{n-1} r_n) \\ &= h(r_1 x'_1 r_2 x'_2 r_3 \cdots s'_{n-1} r_n) \\ &\;\;\vdots \\ &= h(r_1 x'_1 r_2 x'_2 r_3 \cdots x'_{n-1} r_n) = h(u). \end{aligned}$$

Note that the substitution rules $s'_i \to x'_i$ are h-invariant in their respective contexts only when applied from left to right. $\square$

When using monomials of the form $w_1 A^* w_2 \cdots A^* w_n$ instead of $w_1 A^+ w_2 \cdots A^+ w_n$, the degree increases slightly.

Corollary 12.34
Let $L \subseteq A^+$ be recognized by a finite semigroup S satisfying $U_1 = V_1$, and let $u, v \in A^+$. If u and v are contained in the same monomials $w_1 A^* w_2 \cdots A^* w_n$ with $n \leqslant 2|S|$ and degree $|w_1 \cdots w_n| < 4|S|^2$, then $h(u) = h(v)$.

Proof. Every monomial $w_1 A^+ \cdots w_{n-1} A^+ w_n$ is a finite union of monomials of the form $w_1 a_1 A^* \cdots w_{n-1} a_{n-1} A^* w_n$, where $a_i \in A$. Proposition 12.33 thus yields the claim. $\square$

Proof of Proposition 12.30. (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3): This follows from Proposition 12.20.

(3) $\Rightarrow$ (4): The syntactic semigroup of every monomial $w_1 A^* w_2 \cdots A^* w_\ell$ satisfies $e^\omega z e^\omega \leqslant e^\omega$ by Lemma 12.23, and by Lemma 12.19 it satisfies $U_1 = V_1$. Hence, L is recognizable by a direct product of semigroups satisfying $U_1 = V_1$. Therefore, S_L satisfies $U_1 = V_1$ because S_L is a divisor of this direct product.

(4) $\Rightarrow$ (3): Let L be recognized by $h\colon A^+ \to S$ with S satisfying $U_1 = V_1$. Let $u \approx v$ if u and v are contained in the same monomials of the form $w_1 A^* w_2 \cdots A^* w_n$ with $|w_1 \cdots w_n| < 4|S|^2$. We have $L = h^{-1}(P)$ for $P = h(L)$. Corollary 12.34 shows that every set $h^{-1}(p)$ is a union of $\approx$-classes. Moreover, $\approx$ has finite index since there are only finitely many monomials of bounded degree. Every $\approx$-class is a finite Boolean combination of the required form by specifying which monomials hold and which do not hold. $\square$

12.2.4. Alternation-Free First-Order Logic without min and max

We first omit the maximum predicate. As for existential first-order logic this prohibits the specification of suffixes. Because complementation is available, this leads to $\mathcal{R}$-classes instead of right ideals for the recognizing set.

An auxiliary fact that we need later is that the residual of a monomial is itself a union of monomials of the same degree.

Lemma 12.35
Let $P = w_1 A^* w_2 \cdots A^* w_n$, and let $uq \in P$. There exists a monomial $Q = v_1 A^* v_2 \cdots A^* v_\ell$ with $|v_1 \cdots v_\ell| \leqslant |w_1 \cdots w_n|$ and $\ell \leqslant n$ such that $u \in Q \subseteq Pq^{-1}$.

Proof. Let $uq = w_1 s_1 w_2 \cdots s_{n-1} w_n$. First consider the case $u = w_1 s_1 \cdots w_{i-1} s_{i-1} v$, $w_i = vv'$ and $q = v' s_i w_{i+1} \cdots s_{n-1} w_n$ for some i. Setting $Q = w_1 A^* \cdots w_{i-1} A^* v$ yields the claim. In the other case we have $u = w_1 s_1 \cdots s_{i-1} w_i t$, $s_i = tt'$ and $q = t' w_{i+1} \cdots s_{n-1} w_n$ for some i. In this case, we set $Q = w_1 A^* w_2 \cdots w_i A^*$. $\square$

Remember that we saw in the preceding section that with the maximum predicate available, the preimage of an element in semigroup satisfying $U_1 = V_1$ is a union of monomials of small degree. When considering the preimage of a whole $\mathcal{R}$-class instead of a single element, the next lemma shows that the monomials can be chosen so as to not specify the suffix.

Lemma 12.36
Let $h\colon A^+ \to S$ with S satisfying $U_1 = V_1$, and let $u, v \in A^+$. If u and v are contained in the same monomials $w_1 A^* \cdots w_n A^*$ with $|w_1 \cdots w_n| < 8|S|^2$, then $h(u) \mathrel{\mathcal{R}} h(v)$.

Proof. We write $u \approx_m v$ if u and v are contained in the same monomials $w_1 A^* w_2 \cdots A^* w_n$ of degree $|w_1 \cdots w_n| \leqslant m$. Analogously, we write $u \sim_m v$ if u and v are contained in the same monomials $w_1 A^* \cdots w_n A^*$ of degree $|w_1 \cdots w_n| \leqslant m$. If $u \approx_m v$ for $m = 4|S|^2 - 1$, then by Corollary 12.34 we have $h(u) = h(v)$.

Let $u \sim_{2m} v$. We want to show $h(u) \mathrel{\mathcal{R}} h(v)$. We can assume $|u|, |v| \geqslant 2m$ because otherwise $u = v$. Let $u = u'q$ with $|q| = m$. Consider the factorization $v = v'qx$ such that qx is the shortest suffix of v admitting q as a factor, *i.e.*, v is factorized at the last occurrence of q. This factorization exists since both u and v belong to A^*qA^*. We claim that $u \approx_m v'q$ and therefore, $h(v) \leqslant_{\mathcal{R}} h(v'q) = h(u)$. Symmetry then yields $h(u) \mathrel{\mathcal{R}} h(v)$.

We now prove $u \approx_m v'q$. Let $P = w_1A^*w_2 \cdots A^*w_n$ with $|w_1 \cdots w_n| \leqslant m$. First suppose that $v'q \in P$. Then $v \in PA^*$ and $u \in PA^*$. Since w_n is a suffix of q, we conclude $u \in P$.

For the converse suppose $u \in P$. Lemma 12.35 yields a monomial $Q = v_1A^*v_2 \cdots A^*v_\ell$ with $|v_1 \cdots v_\ell| \leqslant |w_1 \cdots w_n|$ and $u' \in Q \subseteq Pq^{-1}$. Since $u'q \in QqA^*$ and the degree of the monomial QqA^* is at most $2m$, we obtain $v \in QqA^*$. By choice of x we have $v'q \in QqA^* \subseteq PA^*$. Since w_n is a suffix of q, we conclude $v'q \in w_1A^*w_2 \cdots A^*w_n$. □

This yields the following effective characterization of $\mathrm{FO}_1[<, \mathrm{suc}, \mathrm{min}]$.

Proposition 12.37
Let $L \subseteq A^+$. The following are equivalent:

1. *L is definable in* $\mathrm{FO}_1^2[<, \mathrm{suc}, \mathrm{min}]$.
2. *L is definable in* $\mathrm{FO}_1[<, \mathrm{suc}, \mathrm{min}]$.
3. *L is a Boolean combination of monomials $w_1A^* \cdots w_\ell A^*$.*
4. *S_L satisfies $U_1 = V_1$ and $h_L(L)$ is a union of $\mathcal{R}$-classes.*

Proof. The equivalences (1) ⇔ (2) ⇔ (3) follow from Proposition 12.27.

(3) ⇒ (4): We have that S_L satisfies $U_1 = V_1$ by Proposition 12.30. The set $h_L(w_1A^* \cdots w_\ell A^*)$ is a right ideal. Hence, $h_L(L)$ is a Boolean combination of right ideals. The claim follows since every Boolean combination of right ideals is a union of $\mathcal{R}$-classes.

(4) ⇒ (3): By Lemma 12.36, there exists $m \in \mathbb{N}$ such that $h_L(u) \mathrel{\mathcal{R}} h_L(v)$ if u and v are contained in the same languages of the form $w_1A^* \cdots w_\ell A^*$ with $|w_1 \cdots w_\ell| \leqslant m$. Therefore, for each $\mathcal{R}$-class R of S_L, the language $h_L^{-1}(R)$ is a Boolean combination of languages $w_1A^* \cdots w_\ell A^*$ with $|w_1 \cdots w_\ell| \leqslant m$. The claim follows, since L is a union of languages of the form $h_L^{-1}(R)$. □

The following is concerned with the left-right dual version for FO_1 without minimum predicate. It follows immediately from the previous proposition by left-right symmetry.

Corollary 12.38
Let $L \subseteq A^+$. The following are equivalent:

1. *L is definable in* $\mathrm{FO}_1^2[<, \mathrm{suc}, \mathrm{max}]$.
2. *L is definable in* $\mathrm{FO}_1[<, \mathrm{suc}, \mathrm{max}]$.
3. *L is a Boolean combination of monomials $A^*w_1 \cdots A^*w_\ell$.*
4. *S_L satisfies $U_1 = V_1$ and $h_L(L)$ is a union of $\mathcal{L}$-classes.* □

In view of the preceding results it is hardly surprising that disregarding minimum as well as maximum predicates leads to $\mathcal{J}$-classes. To prove this, we have to describe preimages of $\mathcal{J}$-classes in terms of monomials that do specify neither prefix nor suffix.

Lemma 12.39
Let $h\colon A^+ \to S$ with S satisfying $U_1 = V_1$, and let $u, v \in A^+$. If u and v are contained in the same monomials $A^* w_1 A^* \cdots w_n A^*$ with $|w_1 \cdots w_n| < 12|S|^2$, then $h(u) \mathrel{\mathcal{J}} h(v)$.

Proof. The proof is along the same lines as Lemma 12.36. The main difference is that we need to consider the factorization $u = pu'q$ with $|p| = |q| = m$ as well as the factorization $v = spv'qx$ such that sp is the shortest prefix of v admitting p as a factor and qx is the shortest suffix of v admitting q as a factor, *i.e.*, v is factorized at the first occurrence of p and the last occurrence of q. □

With this it is straightforward to adapt the proof of Proposition 12.37 with (4) ⇒ (3) relying on Lemma 12.39 to show the following characterization of $\mathrm{FO}_1[<, \mathrm{suc}]$.

Proposition 12.40
Let $L \subseteq A^+$. The following are equivalent:

1. *L is definable in* $\mathrm{FO}^2_1[<, \mathrm{suc}]$.
2. *L is definable in* $\mathrm{FO}_1[<, \mathrm{suc}]$.
3. *L is a Boolean combination of monomials $A^* w_1 \cdots A^* w_\ell A^*$.*
4. *S_L satisfies $U_1 = V_1$ and $h_L(L)$ is a union of $\mathcal{J}$-classes.* □

Beauquier and Pin already used the requirement that the recognizing set be a union of $\mathcal{J}$-classes in their effective characterization of strongly locally testable languages [BP91]. A general framework to investigate regular ideal languages was developed in [JKL12].

Green's relation $\mathcal{J}$ is the finest equivalence relation containing $\mathcal{R}$ and $\mathcal{L}$; *cf.* [Pin86]. This yields the following rather surprising result on logic level: A language is definable in $\mathrm{FO}_1[<, \mathrm{suc}]$ if and only if it is definable in both $\mathrm{FO}_1[<, \mathrm{suc}, \min]$ and $\mathrm{FO}_1[<, \mathrm{suc}, \max]$.

12.3. The Higher Levels of the Alternation Hierarchy with Successor Predicate

The preceding section handled the lowest levels over the various signatures. This section now turns to the higher levels Σ^2_m and FO^2_m with $m \geqslant 2$. The goal is to show that:

- $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$ corresponds to the identity $U_m = V_m$.
- $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$ corresponds to the identity $U_{m-1/2} \leqslant V_{m-1/2}$.

It is easier to handle half and full levels simultaneously, because many intermediate technical constructions for one fragment can be re-used for the other. Moreover, the general proof idea is largely similar.

The proofs for this will be by induction on the level parameter m, going down one level in each inductive step. To get the inductive scheme working, it is a convenient intermediate step to prove $\mathrm{FO}^2_m[<, \mathrm{suc}]$ and $\Sigma^2_m[<, \mathrm{suc}]$ expressively complete for $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$ and $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$, respectively. This also shows straight away that the unavailability of minimum and maximum predicates does not restrict the fragments semantically — in contrast to FO^2_1 and Σ^2_1 as seen in the preceding sections.

Recall that rankers are a simple means specify positions on words. The next lemmas show that formulae can be restricted to factors whose boundaries are given by rankers. Moreover, the overhead required by the relativization involved is sufficiently small .

To formulate this concisely let us introduce some terminology. Let $u \in A^*$, and let r and s be rankers that are defined on u. Let $u[1;r)$ be the word $u[1;r(u)-1]$, and let $u(r;s)$ be the word $u[r(u)+1;s(u)-1]$.

The following lemma gives a relativization $\Sigma^2_{m,n}[<,\mathrm{suc}]$ to the factor $u[1;r)$.

Lemma 12.41
Let $m,n,k \geqslant 0$, let $r \in \widetilde{R}^{\mathrm{Y}}_{1,k}$, and let $\varphi \in \Sigma^2_{m,n}[<,\mathrm{suc}]$. There exists a formula $\langle\varphi\rangle_{<r} \in \Sigma^2_{m,n+k}[<,\mathrm{suc}]$ such that for every $u \in A^*$ with $r(u)$ defined we have

$$u \models \langle\varphi\rangle_{<r} \text{ if and only if } u[1;r) \models \varphi.$$

Proof. The construction of $\langle\varphi\rangle_{<r}$ is by induction on the structure of φ. For atomic φ set $\langle\varphi\rangle_{<r} := \varphi$. For disjunction we set $\langle\psi_1 \vee \psi_2\rangle_{<r} := \langle\psi_1\rangle_{<r} \vee \langle\psi_2\rangle_{<r}$, for conjunction let $\langle\psi_1 \wedge \psi_2\rangle_{<r} := \langle\psi_1\rangle_{<r} \wedge \langle\psi_2\rangle_{<r}$, and for negation let $\langle\neg\psi\rangle_{<r} := \neg\langle\psi\rangle_{<r}$. For existential quantification let $\langle\exists x\, \psi\rangle_{<r}$ be $\exists x\, \big(\langle x < r\rangle \wedge \langle\psi\rangle_{<r}\big)$, where $\langle x < r\rangle$ in $\Sigma^2_{1,k}[<,\mathrm{suc}]$ is the formula from Lemma 11.28. □

Note that the alternation parameter of the formula does not change because r has no direction alternation. Of course, there is also a left-right dual of this lemma, which restricts the interpretation of a formula to the suffix to the right of a ranker in $\widetilde{R}^{\mathrm{X}}_{1,k}$.

We next restrict the interpretation of the formula to the factor $u(r;s)$ between two unidirectional rankers r and s.

Lemma 12.42
Let $m,n,k \geqslant 0$, let $r,s \in \widetilde{R}_{1,k}$, and let $\varphi \in \Sigma^2_{m,n}[<,\mathrm{suc}]$. There exists a formula $\langle\varphi\rangle_{(r;s)} \in \Sigma^2_{m+1,n+k}[<,\mathrm{suc}]$ such that for every $u \in A^*$ with $r(u) < s(u)$ we have

$$u \models \langle\varphi\rangle_{(r;s)} \text{ if and only if } u(r;s) \models \varphi.$$

Proof. We construct $\langle\varphi\rangle_{(r;s)}$ by structural induction. Atomic formulae and Boolean combinations are straightforward; they are handled as in the proof of Lemma 12.41. For existential quantifiers let $\langle\exists x\, \psi\rangle_{(r;s)} := \exists x\, \big(\langle x > r\rangle \wedge \langle x < s\rangle \wedge \langle\psi\rangle_{(r;s)}\big)$, where $\langle x > r\rangle$ and $\langle x < s\rangle$ are the formulae in $\Sigma^2_{2,k}[<,\mathrm{suc}]$ from Lemma 11.28. □

Notice that this relativization loses one alternation. The following two lemmas rephrase the preceding two lemmas in a more global terminology, relativizing the relation $\preccurlyeq^{\mathrm{FO2}}_{m,n}$ to certain parts of the models.

Lemma 12.43
Let $m \geqslant 2$, let $n,k \geqslant 0$, and let $r \in \widetilde{R}^{\mathrm{Y}}_{1,k}(v)$. If $u \preccurlyeq^{\mathrm{FO2}}_{m,n+k} v$, then $u[1;r) \preccurlyeq^{\mathrm{FO2}}_{m,n} v[1;r)$.

Proof. Suppose $v[1;r) \in L(\varphi)$ for $\varphi \in \Sigma^2_{m,n}[<,\mathrm{suc}]$. Using Lemma 12.42, we have $v \in L(\langle\varphi\rangle_{<r})$. Since $\langle\varphi\rangle_{<r}$ is in $\Sigma^2_{m,n+k}[<,\mathrm{suc}]$, the assumption yields $u \in L(\langle\varphi\rangle_{<r})$, which in turn implies $u[1;r) \models \varphi$. □

Again, there is also a left-right dual of this lemma, relativizing to the part right of a ranker in $\widetilde{R}^{\mathrm{X}}_{1,k}$. We use it only implicitly to argue with symmetry and do not state it explicitly.

Lemma 12.44
Let $m \geqslant 2$, let $n,k \geqslant 0$, and let $r,s \in \widetilde{R}_{1,k}(v)$ such that $r(v) < s(v)$ and $r(u) < s(u)$. If $u \preccurlyeq^{\mathrm{FO2}}_{m,n+k} v$, then $u(r;s) \preccurlyeq^{\mathrm{FO2}}_{m-1,n} v(r;s)$.

Proof. Suppose $v(r;s) \models \varphi$ for $\varphi \in \Sigma^2_{m-1,n}[<, \mathrm{suc}]$. We have $v \models \langle\varphi\rangle_{(r;s)}$ for the formula from Lemma 12.42. Since $\langle\varphi\rangle_{(r;s)}$ is in $\Sigma^2_{m,n+k}[<, \mathrm{suc}]$, the assumption yields $u \models \langle\varphi\rangle_{(r;s)}$. This in turn implies $u(r;s) \models \varphi$. $\square$

The following is the main combinatorial ingredient for the proof of the characterizations for both FO^2_m and Σ^2_m. It combines the $\mathcal{L}$-factorization of one word with a variant of the $\mathcal{R}$-factorization of another and uses **LDA** to show that all descents are easily reached by look-around rankers.

Lemma 12.45
*Let $h\colon A^+ \to S$ be a surjective homomorphism with $S \in$ **LDA**, and let $u \preccurlyeq^{\mathrm{FO}^2}_{2,|S|^2+1} v$. There exists $\ell \leqslant 2|S|$ and factorizations $u = a_1 s_1 a_2 \cdots s_{\ell-1} a_\ell$ and $v = a_1 t_1 a_2 \cdots t_{\ell-1} a_\ell$ for some $a_i \in A$ and some $s_i, t_i \in A^*$ such that for each $i \in \{1, \ldots, \ell\}$ there exists a ranker $r_i \in \widetilde{R}_{1,|S|^2}$ with $r_i(u) = |a_1 s_1 a_2 \cdots s_{i-1} a_i|$ and $r_i(v) = |a_1 t_1 a_2 \cdots t_{i-1} a_i|$, and for each $i \in \{1, \ldots, \ell - 1\}$ the following properties hold:*

1. $h(a_1 t_1 a_2 \cdots t_{i-1} a_i) \mathrel{\mathcal{R}} h(a_1 t_1 a_2 \cdots t_{i-1} a_i s_i)$,
2. $h(a_{i+1} s_{i+1} \cdots s_{\ell-1} a_\ell) \mathrel{\mathcal{L}} h(s_i a_{i+1} \cdots s_{\ell-1} a_\ell)$.

Note that in property (1) it is not a misprint that we $\mathcal{R}$-invariantly append s_i and not t_i. This will be essential in the later application for the half levels. The existence of the rankers r_i means that the marker a_i is easily reached (*i.e.*, by means of r_i) on both u and v.

Proof of Lemma 12.45. The construction of the factorization is in three stages. The first stage takes care of $\mathcal{L}$-descents, ultimately leading to property (2). The second stage is similar (yet not just left-right symmetric) and takes care of $\mathcal{R}$-descents to get property (1). In the third and last step the factorizations of the first two stages are shuffled into one another.

Before we start with the construction, note that the assumption $u \preccurlyeq^{\mathrm{FO}^2}_{2,|S|^2+1} v$ implies $u \preccurlyeq^{R}_{2,|S|^2+1} v$ by Corollary 11.17. In particular $u \approx^{R}_{1,|S|^2+1} v$.

We come to the first stage. To get an induction working, we shall construct the following more general factorization. We show by induction on $\ell' = |D_{\mathcal{L}}(uq)| - |D_{\mathcal{L}}(q)|$ that for $q \in A^+$ and $u, v \in A^*$ with $u \preccurlyeq^{\mathrm{FO}^2}_{2,\ell'(|S|+1)} v$ we can factorize

$$u = s'_1 b_1 \cdots s'_{\ell'} b_{\ell'} s'_{\ell'+1} \quad \text{and} \quad v = t'_1 b_1 \cdots t'_{\ell'} b_{\ell'} t'_{\ell'+1}$$

with $b_i \in A$ and $s'_i, t'_i \in A^*$ such that

$$h(b_i s'_{i+1} \cdots b_{\ell'} s'_{\ell'+1} q) \mathrel{\mathcal{L}} h(s'_i b_i s'_{i+1} \cdots b_{\ell'} s'_{\ell'+1} q)$$

for each $i \in \{1, \ldots, \ell' + 1\}$. Moreover, for each $i \in \{1, \ldots, \ell'\}$ there exists $\bar{r}_i \in \widetilde{R}^{\mathrm{Y}}_{1,\ell'(|S|+1)}$ with $\bar{r}_i(u) = |s'_1 b_1 \cdots s'_i b_i|$ and $\bar{r}_i(v) = |t'_1 b_1 \cdots t'_i b_i|$.

The induction base $\ell' = 0$ is vacuously true, so suppose $\ell' \geqslant 1$. By the assumption $u \preccurlyeq^{\mathrm{FO}^2}_{2,|S|+1} v$, the words u and v have the same prefixes and suffixes of length at most $|S| + 1$. Let $u = u'bs$ with $b \in A$ and $|bs|$ minimal such that $h(q) >_{\mathcal{L}} h(bsq)$. We distinguish between $|s| < |S|$ and $|s| \geqslant |S|$ to construct a certain factorization $v = v'bt$ as follows.

Suppose first $|s| < |S|$. The suffix of u of length $|s| + 1$ is also a suffix of v and thus $v = v'bt$ for some word v' and $t = s$. In particular there exists a ranker $r \in \widetilde{R}^{\mathrm{Y}}_{1,|S|+1}$ such that $r(u) = |u'b|$ and $r(v) = |v'b|$.

In the other case suppose $|s| \geqslant |S|$. By the **LDA** Descending Lemma 12.17 the prefix of length $|S|+1$ of bs is not a factor of s and consequently $r(u) = |u'b|$, where $r = \mathrm{Y}_w$ with w being the $|S|$-context of the position $|u'b|$; that is, $w = u[|u'b|-|S|; |u'b|; |u'b|+|S|]$. In particular $r \in \widetilde{R}^{\mathrm{Y}}_{1,|S|+1}$. The ranker r is also defined on v. Let $v = v'bt$ such that $r(v) = |v'b|$.

In both cases this yields a factorization $v = v'bt$ and a ranker $r \in \widetilde{R}^{\mathrm{Y}}_{1,|S|+1}$ such that $r(u) = |u'b|$ and $r(v) = |v'b|$. Lemma 12.43 shows $u' \preccurlyeq^{\mathrm{FO}^2}_{2,(\ell'-1)(|S|+1)} v'$. Induction (with $q' = bsq$) yields $u' = s'_1 b_1 \cdots s'_{\ell'-1} b_{\ell'-1} s'_{\ell'}$ and $v' = t'_1 b_1 \cdots t'_{\ell'-1} b_{\ell'-1} t'_{\ell'}$ with $b_i \in A$ and $s'_i, t'_i \in A^*$ such that for all i we have:

– $r'_i(u') = |s'_1 b_1 \cdots s'_i b_i|$ and $r'_i(v') = |t'_1 b_1 \cdots t'_i b_i|$ for some $r'_i \in \widetilde{R}^{\mathrm{Y}}_{1,(\ell'-1)(|S|+1)}$, and
– $h(b_i s'_{i+1} \cdots b_{\ell'-1} s'_{\ell'} q') \mathrel{\mathcal{L}} h(s'_i b_i s'_{i+1} \cdots b_{\ell'-1} s'_{\ell'} q')$.

This extends to the desired factorization by setting $b_{\ell'} = b$ as well as $s'_{\ell'+1} = s$ and $t'_{\ell'+1} = t$. The rankers $\bar{r}_i$ are given by $rr'_i \in \widetilde{R}^{\mathrm{Y}}_{1,\ell'(|S|+1)}$.

We now come to the second stage, and give a similar factorization on v with respect to $\mathcal{R}$. This is mostly left-right symmetric to the first-stage and we confine ourselves to pointing out the major differences to the first stage.

We prove by induction on $\bar{\ell} = |D_{\mathcal{R}}(pv)| - |D_{\mathcal{R}}(p)|$ that for $p \in A^+$ and $u, v \in A^*$ with $u \preccurlyeq^{\mathrm{FO}^2}_{2,\bar{\ell}(|S|+1)} v$ there exists an integer $\ell'' \leqslant \bar{\ell}$ and factorizations

$$u = s''_1 c_1 \cdots s''_{\ell''} c_{\ell''} s''_{\ell''+1} \quad \text{and} \quad v = t''_1 c_1 \cdots t''_{\ell''} c_{\ell''} t''_{\ell''+1}$$

with $c_i \in A$ and $s''_i, t''_i \in A^*$ such that $h(pt''_1 c_1 \cdots t''_{i-1} c_{i-1}) \mathrel{\mathcal{R}} h(pt''_1 c_1 \cdots t''_{i-1} c_{i-1} s''_i)$ for each $i \in \{1, \ldots, \ell''+1\}$. Moreover, for each $i \in \{1, \ldots, \ell''\}$ there exists a ranker $\bar{r}'_i \in \widetilde{R}^{\mathrm{X}}_{1,\ell''(|S|+1)}$ with $\bar{r}'_i(u) = |s''_1 c_1 \cdots s''_i c_i|$ and $\bar{r}'_i(v) = |t''_1 c_1 \cdots t''_i c_i|$.

The claim is vacuously true for $\bar{\ell} = 0$. Suppose $\bar{\ell} \geqslant 1$, that is, $h(p) >_{\mathcal{R}} h(pv)$. The **LDA** Descending Lemma 12.17 implies that then also $h(p) >_{\mathcal{R}} h(pu)$. Note that we have $\mathrm{alph}_{|S|+1}(u) = \mathrm{alph}_{|S|+1}(v)$.

Consider $u = scu''$, where $c \in A$ and $|sc|$ is minimal such that $h(p) >_{\mathcal{R}} h(psc)$. We continue left-right symmetrically as in stage one to obtain a factorization $v = tcv''$ and a ranker $r \in \widetilde{R}^{\mathrm{X}}_{1,|S|+1}$ such that $r(u) = |sc|$ and $r(v) = |tc|$. The main difference to stage one is to show that we can use induction; *i.e.*, we have to show $h(p) >_{\mathcal{R}} h(ptc)$. This is trivial if $|s| < |S|$, because then $s = t$ by construction. If $|s| \geqslant |S|$, then the construction of r ensures that it ends with X_w or $\mathbb{X}_w$, where the left factor of the $|S|$-context w exhausts the full possible length $|S|$. Using this observation, Lemma 12.17 yields $h(p) >_{\mathcal{R}} h(ptc)$.

Setting $p'' = ptc$, this yields $|D_{\mathcal{R}}(p''v)| - |D_{\mathcal{R}}(p'')| < |D_{\mathcal{R}}(pv)| - |D_{\mathcal{R}}(p)|$. Therefore, induction is indeed applicable, and we obtain $\ell'' \leqslant |D_{\mathcal{R}}(pv)| - |D_{\mathcal{R}}(p)|$, the factorizations $u'' = s''_2 c_2 \cdots s''_{\ell''} c_{\ell''} s''_{\ell''+1}$ and $v'' = t''_2 c_2 \cdots t''_{\ell''} c_{\ell''} t''_{\ell''+1}$ as well as rankers the $\bar{r}''_i \in \widetilde{R}^{\mathrm{X}}_{1,(\ell''-1)(|S|+1)}$. The desired factorization is obtained by setting $s''_1 = s$ and $t''_1 = t$ as well as $c_1 = c$. The rankers $\bar{r}'_i$ are given by $r\bar{r}''_i \in \widetilde{R}^{\mathrm{X}}_{1,\ell''(|S|+1)}$.

Stage three combines the factorizations of the two preceding stages. Let $b, c \in A$ be such that $u, v \in A^*b \cap cA^*$. Let

$$\begin{aligned} u &= s'_1 b_1 \cdots s'_{\ell'} b_{\ell'} s'_{\ell'+1} b = cs''_1 c_1 \cdots s''_{\ell''} c_{\ell''} s''_{\ell''+1}, \\ v &= t'_1 b_1 \cdots t'_{\ell'} b_{\ell'} t'_{\ell'+1} b = ct''_1 c_1 \cdots t''_{\ell''} c_{\ell''} t''_{\ell''+1}, \end{aligned}$$

and rankers $\bar{r}_i \in \widetilde{R}^{\mathrm{Y}}_{1,|S|^2-1}$ and $\bar{r}'_i \in \widetilde{R}^{\mathrm{X}}_{1,|S|^2-1}$ be given by stage one (with $q = b$) and stage two (with $p = c$), respectively. The parameters ℓ' and ℓ'' are bounded by $|S| - 1$. Note that $ub^{-1} \preccurlyeq^{\mathrm{FO2}}_{2,|S|^2} vb^{-1}$ by Lemma 12.43; by symmetry $c^{-1}u \preccurlyeq^{\mathrm{FO2}}_{2,|S|^2} c^{-1}v$. Stage one can thus be applied to ub^{-1} and vb^{-1}, and stage two can be applied to $c^{-1}u$ and $c^{-1}v$, because $|S|^2 \geqslant (|S| - 1)(|S| + 1)$.

The next step is to shuffle the two different factorizations of the same word together. This yields the same factorization because the markers have the same relative order on u and v; *i.e.*, we have $\mathrm{ord}(\mathrm{Y}_b\bar{r}_i(u), \mathrm{X}_c\bar{r}'_j(u)) = \mathrm{ord}(\mathrm{Y}_b\bar{r}_i(v), \mathrm{X}_c\bar{r}'_j(v))$ for all $i \in \{1, \ldots, \ell'\}$ and $j \in \{1, \ldots, \ell''\}$. To see this, notice that the rankers $\mathrm{Y}_b\bar{r}_i$ and $\mathrm{X}_c\bar{r}'_j$ are all in $\widetilde{R}_{1,|S|^2}$. Using $u \approx^R_{1,|S|^2+1} v$ the claim follows. This allows to write

$$u = a_1 s_1 a_2 \cdots s_{\ell-1} a_\ell \quad \text{and} \quad v = a_1 t_1 a_2 \cdots t_{\ell-1} a_\ell$$

for some $\ell \leqslant 2|S|$ with $a_i \in A$ and $s_i, t_i \in A^*$ such that for each i there exists a ranker $r_i \in \widetilde{R}_{1,|S|^2}$ with $r_i(u) = |a_1 s_1 a_2 \cdots s_{i-1} a_i|$ and $r_i(v) = |a_1 t_1 a_2 \cdots t_{i-1} a_i|$, and such that properties (1) and (2) hold. The markers a_i in this factorization cover precisely the positions of the b_i and c_i; the properties are inherited, as we only refine the original factorizations by splitting factors s'_i and t'_i, respectively, factors s''_i and t''_i. □

Building on the factorizations of the previous lemma, we also see that $u \preccurlyeq^{\mathrm{FO2}}_{m,n+k} v$ transfers to $s_i \preccurlyeq^{\mathrm{FO2}}_{m-1,n} t_i$ on the factors. This uses that the markers are reached by unidirectional rankers to employ the relativization techniques that we have seen before. Moreover, the factors have a common prefix and a common suffix. This will be important to recover minimum and maximum predicates for the induction.

Lemma 12.46
Let $m \geqslant 2$, let $n \geqslant 0$, and let $u = a_1 s_1 a_2 \cdots s_{\ell-1} a_\ell$ and $v = a_1 t_1 a_2 \cdots t_{\ell-1} a_\ell$, where $\ell \geqslant 1$, $a_i \in A$, and $s_i, t_i \in A^*$. Let $k \geqslant 1$, and let $r_i \in \widetilde{R}_{1,k}$ for $i \in \{1, \ldots, \ell\}$ such that $r_i(u) = |a_1 s_1 a_2 \cdots s_{i-1} a_i|$ and $r_i(v) = |a_1 t_1 a_2 \cdots t_{i-1} a_i|$. If $u \preccurlyeq^{\mathrm{FO2}}_{m,n+k} v$, then for all $i \in \{1, \ldots, \ell - 1\}$ the following hold:

1. *s_i and t_i have the same prefixes and suffixes of length at most $n - 1$, and*
2. *$s_i \preccurlyeq^{\mathrm{FO2}}_{m-1,n} t_i$.*

Proof. Let u and v be factorized according to the premise of the lemma and suppose $u \preccurlyeq^{\mathrm{FO2}}_{m,n+k} v$. Property (2) immediately follows from Lemma 12.44.

It remains to show (1). For $n = 0$ this is vacuously true, so suppose $n \geqslant 1$ in the following. Fix an index i and consider the ranker r_i. Let $j = r_i(v)$ be the position of the marker a_i in v, and let $w = v[j - n + 1; j; j + n - 1]$ be the $(n - 1)$-context of j on v. Let further $\langle x \geqslant r_i \rangle$ and $\langle x \leqslant r_i \rangle$ be the $\Sigma^2_{2,k}[<, \mathrm{suc}]$-formulae from Lemma 11.28. With these define a formula $\psi_i \in \Sigma^2_{2,n+k}[<, \mathrm{suc}]$ by

$$\psi_i \;:=\; \exists x \, \big(\langle x \geqslant r_i \rangle \wedge \langle x \leqslant r_i \rangle \wedge \lambda(x) = w\big),$$

which requires that the $(n - 1)$-context of the position reached by r_i be w. Remember that $\lambda(x) = w$ is defined by (8.1).

By construction $v \in L(\psi_i)$ and thus, using the assumption, $u \in L(\psi_i)$. Informally speaking, this shows that for each i, the $(n - 1)$-context of the marker a_i is the same on u and v. Supposing both s_i and t_i have length at least n, this already suffices for s_i and t_i to have the same prefixes and suffixes of length at most $n-1$. In general, however, this might not suffice as *a priori* one of the factors might be short, whereas the other might not be.

We shall now show that this cannot happen: If s_i is short (shorter than n letters), then t_i is short, too, and has the same length — and *vice versa*. Using property (2) we see that $|t_i| \geqslant n$ implies $|s_i| \geqslant n$. Suppose now $|t_i| = n' \leqslant n-1$. We shall show $|s_i| = |t_i|$. Property (2) shows that $|s_i| \geqslant |t_i|$ in this case, and it suffices to show $|s_i| \leqslant |t_i|$. In the following, we shall use $u \preccurlyeq^{R}_{2,n+k} v$, which is equivalent to $u \preccurlyeq^{\mathrm{FO2}}_{2,n+k} v$ by Corollary 11.20. Let $r' = \langle r_i + n' \rangle$ be the ranker in $\widetilde{R}_{1,n'+k}$ from Lemma 11.24. For this ranker we have $r'(v) = r_i(v) + n'$ and $r'(u) = r_i(u) + n'$. Therefore $r'(v) \geqslant r_{i+1}(v) - 1$ yields $r'(u) \geqslant r_{i+1}(u) - 1$. This shows $|s_i| \leqslant n'$ as desired.

Summarizing, $|s_i| \leqslant n-1$ implies $|t_i| \leqslant n-1$ which in turn implies $|s_i| = |t_i|$. Together with the above we conclude $s_i = t_i$ if either of s_i or t_i is shorter than n. $\square$

Property (1) in particular yields $|s_i| < n$ if and only if $|t_i| < n$ and in this case $s_i = t_i$. The assumption on the factorizations is symmetric, therefore interchanging u and v to get $v \preccurlyeq^{\mathrm{FO2}}_{m,n+k} u$ in the premise, we can infer (1) and $t_i \preccurlyeq^{\mathrm{FO2}}_{m-1,n} s_i$. We shall exploit this fact for the full levels in the next proposition.

Let $\mathcal{S}_m$ be the signature $\{<, \mathrm{suc}, \min, \max\}$ if $m = 1$ and $\{<, \mathrm{suc}\}$ otherwise. Let $u \cong_{m,n} v$ if $u \models \varphi \Leftrightarrow v \models \varphi$ for all $\mathrm{FO}^2_{m,n}[\mathcal{S}_m]$. Let $u \approx^{\mathrm{FO2}}_{m,n} v$ if $u \preccurlyeq^{\mathrm{FO2}}_{m,n} v$ and $v \preccurlyeq^{\mathrm{FO2}}_{m,n} u$. In particular, the relations $\cong_{m,n}$ and $\approx^{\mathrm{FO2}}_{m,n}$ coincide whenever $m \geqslant 2$.

The following proposition shows that the preimage of any element of a semigroup satisfying $U_m = V_m$ is a union of $u \cong_{m,n} v$ for some sufficiently large n. For $m = 1$ we have to include minimum and maximum predicates for this to be true. See Section 12.2.4 as to why we cannot omit those predicates.

Proposition 12.47
Let $m \geqslant 1$ be an integer, let $h\colon A^+ \to S$ be a surjective homomorphism onto a finite semigroup S that satisfies $U_m = V_m$. There exists a positive integer n such that $u \cong_{m,n} v$ implies $h(u) = h(v)$ for all $u, v \in A^+$.

Proof. We perform an induction on m. Proposition 12.30 yields the base case $m = 1$. To see this, consider the languages $L_v = h^{-1}(h(v))$ for $v \in A^+$. The homomorphism h clearly recognizes those languages. By Proposition 12.30 there exists an integer n such that for all $v \in A^+$ the language L_v is definable in $\mathrm{FO}^2_{1,n}[<, \mathrm{suc}, \min, \max]$. Consider $u, v \in A^+$. By definition $v \in L_v$ and $u \in L_u$. Now, $u \cong_{1,n} v$ implies $u \in L_v$; i.e., $h(u) = h(v)$.

So let $m \geqslant 2$ in the following, and take some integer $\omega \geqslant |S|$ such that x^ω is idempotent for all $x \in S$. Define a string rewriting relation $\to$ on A^+ by $t \to s$ if $h(s) = h(t)$, or if $t = pv_{m-1}q$ and $s = pu_{m-1}q$ for some $p, q \in A^*$ and some $v_i, u_i \in A^+$, which for $i \geqslant 2$ comply with the recursive scheme:

$$v_1 = (e^\omega z f^\omega x_1 e^\omega)^\omega \bar{z} (f^\omega y_1 e^\omega \bar{z} f^\omega)^\omega, \quad v_i = (u_{i-1} x_i)^\omega v_{i-1} (y_i u_{i-1})^\omega,$$
$$u_1 = (e^\omega z f^\omega x_1 e^\omega)^\omega z (f^\omega y_1 e^\omega \bar{z} f^\omega)^\omega, \quad u_i = (u_{i-1} x_i)^\omega u_{i-1} (y_i u_{i-1})^\omega$$

for some $x_i, y_i, e, f, z, \bar{z} \in A^+$. Intuitively this means that the words v_i and u_i are obtained from the omega-terms V_i and U_i, respectively, by substituting suitable non-empty words for the variables and replacing the formal symbol ω by the number of the same name. We also write $s \leftarrow t$ instead of $t \to s$. Let $\leftrightarrow$ be the symmetric closure of $\to$, and let $\overset{*}{\leftrightarrow}$ be the transitive closure of $\leftrightarrow$. In other words, we have $t \overset{*}{\leftrightarrow} s$ if there exists a chain $t = w_1 \to w_2 \leftarrow w_3 \to \cdots \leftarrow w_\ell = s$ for some $\ell \geqslant 1$ and $w_i \in A^+$. Note that $t \overset{*}{\leftrightarrow} s$ implies $\hat{p} t \hat{q} \overset{*}{\leftrightarrow} \hat{p} s \hat{q}$ for all $\hat{p}, \hat{q} \in A^*$ and thus $\overset{*}{\leftrightarrow}$ is a congruence on A^+.

Observe that if $t \leftrightarrow s$, then s and t have the same prefix and the same suffix of length $|S|$, and the factors of length $|S|+1$ are the same in s and t. Therefore, the **LDA** Descending Lemma 12.17 shows $h(u) \mathrel{\mathcal{R}} h(us)$ if and only if $h(u) \mathrel{\mathcal{R}} h(ut)$ as well as $h(v) \mathrel{\mathcal{L}} h(sv)$ if and only if $h(v) \mathrel{\mathcal{L}} h(tv)$. Note that S is indeed in **LDA** by Lemma 12.18.

We now claim that, within certain contexts, we can lift the rewriting steps of $t \overset{*}{\leftrightarrow} s$ to S without changing the value in S.

Claim *Let* $u, v, s, t \in A^+$ *with* $t \overset{*}{\leftrightarrow} s$*. If* $h(u) \mathrel{\mathcal{R}} h(us)$ *and* $h(v) \mathrel{\mathcal{L}} h(sv)$*, then* $h(usv) = h(utv)$*.*

The claim is trivial if $h(t) = h(s)$; otherwise the proof proceeds by induction on the length of a shortest $\leftrightarrow$-chain from t to s. Assume $t \overset{*}{\leftrightarrow} t' \leftrightarrow s$. This means that we either have $t' = pv_{m-1}q$ and $s = pu_{m-1}q$, or $s = pv_{m-1}q$ and $t' = pu_{m-1}q$.

In either case, using the observation right before the claim, we get $h(u) \mathrel{\mathcal{R}} h(upu_{m-1}q)$ and $h(v) \mathrel{\mathcal{L}} h(pu_{m-1}qv)$. Therefore, there exists $x \in A^*$ such that $h(u) = h(upu_{m-1}x)$, and there exists $y \in A^*$ such that $h(v) = h(ypu_{m-1}qv)$. Pumping these identities yields $h(u) = h\big(u(pu_{m-1}qx)^\omega\big)$ and $h(v) = h\big((ypu_{m-1}q)^\omega v\big)$.

To avoid issues with empty x and y, we insert e^ω and f^ω suitably without changing the h-image. This exploits the fact that u_{m-1} has e^ω as a prefix and f^ω as a suffix. This yields $h(u) = h\big(u(pu_{m-1}f^\omega qx)^\omega\big)$ and $h(v) = h\big((ype^\omega u_{m-1}q)^\omega v\big)$. Letting $x_m = f^\omega qxp$ and $y_m = qype^\omega$, the identity $U_m = V_m$ in S yields

$$\begin{aligned} h(usv) &= h\big(up(u_{m-1}x_m)^\omega u_{m-1}(y_m u_{m-1})^\omega qv\big) \\ &= h\big(up(u_{m-1}x_m)^\omega v_{m-1}(y_m u_{m-1})^\omega qv\big) = h(ut'v). \end{aligned}$$

To see this, observe $(pu_{m-1}f^\omega qx)^\omega p = p(u_{m-1}f^\omega qxp)^\omega = p(u_{m-1}x_m)^\omega$ and, similarly, $q(ype^\omega u_{m-1}q)^\omega = (y_m u_{m-1})^\omega q$. By induction $h(ut'v) = h(utv)$ and thus $h(usv) = h(utv)$. Note that induction indeed applies as $h(u) \mathrel{\mathcal{R}} h(ut')$ and $h(v) \mathrel{\mathcal{L}} h(t'v)$. This completes the proof of the claim.

The quotient set $S' = (A^+/\overset{*}{\leftrightarrow})$ is naturally equipped with a semigroup structure because $\overset{*}{\leftrightarrow}$ is a congruence on A^+. The semigroup S' is a quotient of S and, in particular, S' is a finite semigroup in **LDA** and x^ω is idempotent for all $x \in S'$. Let $h' \colon A^+ \to S'$ be the canonical homomorphism, mapping a word to its equivalence class modulo $\overset{*}{\leftrightarrow}$. By construction S' satisfies the identity $U_{m-1} = V_{m-1}$.

To be able to use induction, however, we still have to have minimum and maximum predicates available for the very first level. Induction yields an integer $\bar{n}$ such that $s \cong_{m-1,\bar{n}} t$ implies $h'(s) = h'(t)$ for all $s, t \in A^+$. Let $n = 4\bar{n} + |S|^2 + 2$ if $m = 2$ and let $n = \bar{n} + |S|^2$ otherwise. We shall show that $u \cong_{m,n} v$ implies $h(u) = h(v)$ for all $u, v \in A^+$. (Note that the fragment behind $\cong_{m,n}$ has neither minimum nor maximum predicates.)

Let $u, v \in A^+$ and $u \cong_{m,n} v$ (i.e., $u \approx^{\mathrm{FO2}}_{m,n} v$ as $m \geqslant 2$). Let $u = a_1 s_1 a_2 \cdots s_{\ell-1} a_\ell$ and $v = a_1 t_1 a_2 \cdots t_{\ell-1} a_\ell$ be the factorizations from Lemma 12.45. Plugging this factorization into Lemma 12.46, we get the following for all i:

- $h(a_1 t_1 a_2 \cdots t_{i-1} a_i) \mathrel{\mathcal{R}} h(a_1 t_1 a_2 \cdots t_{i-1} a_i s_i)$,
- $h(a_{i+1} s_{i+1} \cdots s_{\ell-1} a_\ell) \mathrel{\mathcal{L}} h(s_i a_{i+1} \cdots s_{\ell-1} a_\ell)$,
- s_i and t_i have the same prefixes and suffixes of length at most $4\bar{n}+1$, and
- $s_i \approx^{\mathrm{FO2}}_{m-1,n-|S|^2} t_i$.

We claim that $s_i \cong_{m-1,\bar{n}} t_i$. For $m \geqslant 3$ this is trivial as $\bar{n} = n - |S|^2$ and thus $\approx^{\text{FO2}}_{m-1,n-|S|^2}$ coincides with $\cong_{m-1,\bar{n}}$. So let $m = 2$ in the following. We distinguish between short factors and long factors. By the third property we know that if either of s_i or t_i has length at most $4\bar{n}$, then we even have $s_i = t_i$. So we may assume that s_i and t_i both have length at least $4\bar{n}+1$. Consider a sentence φ in $\mathrm{FO}^2_{1,\bar{n}}[<, \mathrm{suc}, \min, \max]$. Using Corollary 9.7 and its left-right symmetric version, there exists a sentence ψ in $\mathrm{FO}^2_{1,4\bar{n}}[<, \mathrm{suc}]$ such that $w \in L(\varphi) \Leftrightarrow w \in L(\psi)$ if w has length at least $4\bar{n}+1$ as well as the same prefix and suffix of length $4\bar{n}$ as s_i (or equivalently t_i). Thus $t_i \in L(\varphi)$ implies $t_i \in L(\psi)$, which yields $s_i \in L(\psi)$ by $s_i \approx^{\text{FO2}}_{1,4\bar{n}} t_i$. Therefore, $s_i \in L(\varphi)$. Note that $n - |S|^2 > 4\bar{n}$.

This shows that we have $s_i \cong_{m-1,\bar{n}} t_i$ independently of m, which implies $t_i \overset{*}{\leftrightarrow} s_i$ by choice of $\bar{n}$ and definition of h'. Applying the above claim repeatedly to substitute s_i with t_i for increasing $i \in \{0, \ldots, \ell\}$ yields the following chain of identities:

$$\begin{aligned}
h(u) &= h(a_1 s_1 a_2 s_2 \cdots a_{\ell-1} s_{\ell-1} a_\ell) \\
&= h(a_1 t_1 a_2 s_2 \cdots a_{\ell-1} s_{\ell-1} a_\ell) \\
&= h(a_1 t_1 a_2 t_2 \cdots a_{\ell-1} s_{\ell-1} a_\ell) \\
&\;\;\vdots \\
&= h(a_1 t_1 a_2 t_2 \cdots a_{\ell-1} t_{\ell-1} a_\ell) = h(v).
\end{aligned}$$

This concludes the proof. $\square$

We obtain a similar proposition for the half level $\Sigma^2_m[<, \mathrm{suc}]$ very much along the same lines of the preceding proof. To employ an induction on the alternation level m we again use string rewriting to obtain an ordered semigroup one level lower. This time the rewriting relation is not symmetric which ultimately leads to an order of its equivalence classes.

Let $u \lesssim_{m,n} v$ if $v \models \varphi$ implies $u \models \varphi$ for all $\Sigma^2_{m,n}[\mathcal{S}_m]$, where $\mathcal{S}_m$ is defined as above.

Proposition 12.48
*Let $m \geqslant 1$ be an integer, let $h\colon A^+ \to S$ be a surjective homomorphism onto an ordered semigroup $S \in$ **LDA** that satisfies $U_{m-1/2} \leqslant V_{m-1/2}$. There exists a positive integer n such that $u \lesssim_{m,n} v$ implies $h(u) \leqslant h(v)$ for all $u, v \in A^+$.*

Proof. We perform an induction on m. For the base case $m = 1$ consider the order ideals I generated by $h(v)$ for $v \in A^+$ and let $L = h^{-1}(I)$. The languages L are all recognized by h. Proposition 12.20 yields an integer n such that all such languages are definable in $\Sigma^2_{1,n}[<, \mathrm{suc}, \min, \max]$. Consider $u \in A^+$ with $u \lesssim_{1,n} v$. By definition $v \in L$ and thus $u \in L$. This shows $h(u) \leqslant h(v)$.

Let $m \geqslant 2$ in the following, and fix some integer $\omega \geqslant |S|$ such that x^ω is idempotent for all $x \in S$. To avoid fractional indices, let $\bar{m} = m - 1/2$. We introduce a string rewriting relation $\to$ on A^+ by letting $t \to s$ if $h(s) = h(t)$, or if $t = p\, v_{\bar{m}-1}\, q$ and $s = p\, u_{\bar{m}-1}\, q$ for some $p, q \in A^*$ and some $v_i, u_i \in A^+$ obeying the following recursion for $i - 3/2 \in \mathbb{N}$:

$$\begin{aligned}
v_{1/2} &= e^\omega, & v_i &= (u_{i-1} x_i)^\omega v_{i-1} (y_i u_{i-1})^\omega, \\
u_{1/2} &= e^\omega z e^\omega, & u_i &= (u_{i-1} x_i)^\omega u_{i-1} (y_i u_{i-1})^\omega
\end{aligned}$$

for some $x_i, y_i, e, z \in A^+$. In other words, v_i and u_i arise from V_i and U_i, respectively, by substituting the formal power ω with the number of that name, and by substituting non-empty words for all the variables.

Let $\xrightarrow{*}$ be the transitive closure of $\to$, *i.e.*, let $t \xrightarrow{*} s$ if there exists a chain $t = w_1 \to w_2 \to \cdots \to w_\ell = s$ for some $\ell \geqslant 1$ and $w_i \in A^+$. Note that $t \xrightarrow{*} s$ implies $\hat{p}t\hat{q} \xrightarrow{*} \hat{p}s\hat{q}$ for all $\hat{p}, \hat{q} \in A^*$ and thus $\xrightarrow{*}$ is a compatible preorder of A^+.

We claim that, within certain contexts, we can lift the rewriting steps of $t \xrightarrow{*} s$ to S in an order respecting way.

Claim Let $u, v, s, t \in A^+$ *with* $t \xrightarrow{*} s$. *If* $h(u) \mathrel{\mathcal{R}} h(us)$ *and* $h(v) \mathrel{\mathcal{L}} h(sv)$, *then* $h(usv) \leqslant h(utv)$.

The proof of the claim is by induction on the length of a shortest $\to$-chain from t to s. The claim is trivial if $h(t) = h(s)$. Suppose $t \xrightarrow{*} t' \to s$ and $t' = p\, v_{\bar{m}-1}\, q$ and $s = p\, u_{\bar{m}-1}\, q$. Since $h(u) \mathrel{\mathcal{R}} h(us)$ there exists $x \in A^*$ such that $h(u) = h(usx)$; and since $h(v) \mathrel{\mathcal{L}} h(sv)$ there exists $y \in A^*$ such that $h(v) = h(ysv)$. Now $h(u) = h\big(u(pu_{\bar{m}-1}qx)^\omega\big)$ and $h(v) = h\big((ypu_{\bar{m}-1}q)^\omega v\big)$. To avoid issues with empty x and y, we insert e^ω without changing the h-image. This exploits the fact that $u_{\bar{m}-1}$ has e^ω both as a prefix and a suffix. This yields $h(u) = h\big(u(pu_{\bar{m}-1}e^\omega qx)^\omega\big)$ and $h(v) = h\big((ype^\omega u_{\bar{m}-1}q)^\omega v\big)$.

Letting $x_{\bar{m}} = e^\omega qxp$ and $y_{\bar{m}} = qype^\omega$, the identity $U_{\bar{m}} \leqslant V_{\bar{m}}$ in S yields

$$\begin{aligned} h(usv) &= h\big(up(u_{\bar{m}-1}x_{\bar{m}})^\omega u_{\bar{m}-1}(y_{\bar{m}}u_{\bar{m}-1})^\omega qv\big) \\ &\leqslant h\big(up(u_{\bar{m}-1}x_{\bar{m}})^\omega v_{\bar{m}-1}(y_{\bar{m}}u_{\bar{m}-1})^\omega qv\big) = h(ut'v). \end{aligned}$$

To see this, observe $(pu_{\bar{m}-1}e^\omega qx)^\omega p = p(u_{\bar{m}-1}e^\omega qxp)^\omega = p(u_{\bar{m}-1}x_{\bar{m}})^\omega$ and, similarly, $q(ype^\omega u_{\bar{m}-1}q)^\omega = (y_{\bar{m}}u_{\bar{m}-1})^\omega q$. By induction $h(ut'v) \leqslant h(utv)$ and consequently $h(usv) \leqslant h(utv)$. Note that by construction every factor of length at most $|S| + 1$ of t' is also a factor of s, and moreover, t' and s have the same prefix and suffix of length $|S|$. Lemma 12.17 shows that $h(u) \mathrel{\mathcal{R}} h(us)$ implies $h(u) \mathrel{\mathcal{R}} h(ut')$, and $h(v) \mathrel{\mathcal{L}} h(sv)$ implies $h(v) \mathrel{\mathcal{L}} h(t'v)$. This completes the proof of the claim.

Let $t \sim s$ if $t \xrightarrow{*} s$ and $s \xrightarrow{*} t$. This is a congruence on A^+, and $S' = (A^+/{\sim})$ becomes a semigroup. Let $h' \colon A^+ \to S'$ be the canonical homomorphism, mapping a word to its equivalence class modulo $\sim$. The preorder $\xrightarrow{*}$ on A^+ induces a partial order on S' (which is also denoted by $\leqslant$ for conciseness) by letting $h'(u) \leqslant h'(v)$ if $v \xrightarrow{*} u$. Moreover, S' is a (non-ordered) quotient of S and, in particular, S' is finite and in **LDA**, and x^ω is idempotent for all $x \in S'$.

By construction S' satisfies the identity $U_{\bar{m}-1} \leqslant V_{\bar{m}-1}$. Still, to use induction we have to invest some more work for the very first level to have the minimum and maximum predicates available. By induction there exists an integer $\bar{n}$ such that $s \lesssim_{m-1,\bar{n}} t$ implies $h'(s) \leqslant h'(t)$ for all $s, t \in A^+$. Let $n = 4\bar{n} + |S|^2 + 2$ if $m = 2$, and let $n = \bar{n} + |S|^2$ otherwise. We shall show that $u \lesssim_{m,n} v$ implies $h(u) \leqslant h(v)$ for all $u, v \in A^+$.

Let $u, v \in A^+$ with $u \lesssim_{m,n} v$ (in other words $u \preccurlyeq^{\mathrm{FO2}}_{m,n} v$ as $m \geqslant 2$). Consider the factorizations $u = a_1s_1a_2 \cdots s_{\ell-1}a_\ell$ and $v = a_1t_1a_2 \cdots t_{\ell-1}a_\ell$ from Lemma 12.45. Lemma 12.46 shows the following for all i:

- $h(a_1t_1a_2 \cdots t_{i-1}a_i) \mathrel{\mathcal{R}} h(a_1t_1a_2 \cdots t_{i-1}a_is_i)$,
- $h(a_{i+1}s_{i+1} \cdots s_{\ell-1}a_\ell) \mathrel{\mathcal{L}} h(s_ia_{i+1} \cdots s_{\ell-1}a_\ell)$,
- s_i and t_i have the same prefixes and suffixes of length at most $4\bar{n} + 1$, and
- $s_i \preccurlyeq^{\mathrm{FO2}}_{m-1,n-|S|^2} t_i$.

We claim that $s_i \lesssim_{m-1,\bar{n}} t_i$. If $m \geqslant 3$, this is trivial since $\lesssim_{m-1,\bar{n}}$ and $\preccurlyeq^{\mathrm{FO2}}_{m-1,n-|S|^2}$ describe the same relation; remember that $\bar{n} = n - |S|^2$ by definition. So let $m = 2$ in the following. We distinguish between short factors and long factors. By the third property we know that if either of s_i or t_i has length at most $4\bar{n}$, then we even have $s_i = t_i$. So we may assume that s_i and t_i both have length at least $4\bar{n} + 1$. Let $\varphi \in \Sigma^2_{1,\bar{n}}[<, \mathrm{suc}, \min, \max]$. By Corollary 9.6 and its left-right symmetric version, there exists $\psi \in \Sigma^2_{1,4\bar{n}}[<, \mathrm{suc}]$ such that $w \in L(\varphi) \Leftrightarrow w \in L(\psi)$ if w has length at least $4\bar{n} + 1$ as well as the same prefix and suffix of length $4\bar{n}$ as s_i (or equivalently t_i). Thus $t_i \in L(\varphi)$ implies $t_i \in L(\psi)$. Therefore, $s_i \in L(\psi)$ by $s_i \preccurlyeq^{\mathrm{FO2}}_{1,4\bar{n}} t_i$, which finally yields $s_i \in L(\varphi)$. Note that $n - |S|^2 > 4\bar{n}$.

This shows that, independently of m, we have $s_i \lesssim_{m-1,\bar{n}} t_i$, which implies $t_i \xrightarrow{*} s_i$. Repeatedly applying the above claim to substitute t_i for s_i for increasing $i \in \{0, \ldots, \ell\}$ yields the following chain of inequalities:

$$\begin{aligned}
h(u) &= h(a_1 s_1 a_2 s_2 \cdots a_{\ell-1} s_{\ell-1} a_\ell) \\
&\leqslant h(a_1 t_1 a_2 s_2 \cdots a_{\ell-1} s_{\ell-1} a_\ell) \\
&\leqslant h(a_1 t_1 a_2 t_2 \cdots a_{\ell-1} s_{\ell-1} a_\ell) \\
&\;\;\vdots \\
&\leqslant h(a_1 t_1 a_2 t_2 \cdots a_{\ell-1} t_{\ell-1} a_\ell) = h(v).
\end{aligned}$$

This concludes the proof. □

These results will suffice to establish the direction from semigroups to logic. What is missing for the proof is the reverse implication showing that all languages defined by formulae satisfy the identity. For this we show the following lifting lemma for look-around rankers. It shows that, under certain conditions, we can transfer $u \preccurlyeq^R_{m,n} v$ to $px^{2n}uy^{2n}q \preccurlyeq^R_{m+1,n} px^{2n}vy^{2n}q$ one alternation level higher. Applying this lifting lemma to a suitable instantiation of the identity yields the claim.

The proof is very technical and involved. However, there is a similar lifting lemma (Lemma 12.53) for ordinary rankers without look-around below. This lemma is much less technical, and it may be a good idea to jump ahead. Bearing in mind that, broadly speaking, factors are for look-around rankers what letters are for ordinary rankers, the reader is afterwards probably able to come up with an adaptation of the proof to get the following.

Lemma 12.49 (Ranker Lifting)
Let $m \geqslant 0$, and let $n \geqslant 1$ be integers. Let $u, v, x, y \in A^*$ be of length at least $2n$. Suppose $\mathrm{alph}_{2n-1}(u) \cup \mathrm{alph}_{2n-1}(v) \subseteq \mathrm{alph}_{2n-1}(x) \cap \mathrm{alph}_{2n-1}(y)$, suppose that x, u, and v have a common prefix of length $2n$, and suppose that y, u, and v have a common suffix of length $2n$. If $u \preccurlyeq^R_{m,n} v$, then $px^{2n}uy^{2n}q \preccurlyeq^R_{m+1,n} px^{2n}vy^{2n}q$ for all $p, q \in A^*$.

Proof. Let $\tilde{u} = px^{2n}uy^{2n}q$ and $\tilde{v} = px^{2n}vy^{2n}q$. In informal explanations, the positions corresponding to the middle factors u and v are called the *center* of $\tilde{u}$ and $\tilde{v}$, respectively. Let us start by outlining the proof in rough terms. Basically the idea is to reduce $\tilde{u} \preccurlyeq^R_{m+1,n} \tilde{v}$ to $u \preccurlyeq^R_{m,n} v$. For this we need some auxiliary observations. All rankers involved in $\tilde{u} \preccurlyeq^R_{m+1,n} \tilde{v}$ are in $\widetilde{R}_{m+1,n}$. For such rankers, split off the first direction (*i.e.*, the longest prefix that consists solely of X-modalities or of Y-modalities). Since every factor of length $2n - 1$ in u or v also occurs in both x and y, such unidirectional rankers cannot lead far into the center of $\tilde{u}$ and $\tilde{v}$, respectively. Consider some ranker r arising from some ranker in $\widetilde{R}_{m+1,n}$ as the suffix after the first direction; its alternation parameter is thus at most m. We want to use this ranker on u and v. However, it starts from the position the first direction led to, and of course, r may use positions outside the center during its evaluation on $\tilde{u}$ and $\tilde{v}$. So we cannot simply take this ranker and evaluate it on u and on v.

Suppose for the moment that r has alternation parameter strictly smaller than m. In this case the ranker behaves the same on $\tilde{u}$ and $\tilde{v}$ in the sense that, roughly speaking, it leaves the center on one word if it does so on the other, in which case it leads to the same position in the prefix ps^{2n} or the suffix $y^{2n}q$. So if r leaves the center on $\tilde{u}$ and $\tilde{v}$ and re-enters later on, then we can take a shortcut (*i.e.*, a suffix of r) to simulate r on u and v.

Suppose now that r has full alternation parameter m. If r leaves the center on $\tilde{u}$ and $\tilde{v}$ before its very last alternation, then with the above observation it can be shortened to get a corresponding ranker on u and v. We may thus assume that up until its very last alternation, r stays inside the center. On the other hand, if it leaves the center on $\tilde{u}$ and $\tilde{v}$ with no alternation left, then it cannot do so on $\tilde{u}$ before it does on $\tilde{v}$. (Assuming the contrary, we could construct a ranker that disproves $u \preccurlyeq^R_{m,n} v$.) Roughly speaking, this means that if r ends on an X-modality, then r on $\tilde{u}$ is less than or equal to r on $\tilde{v}$; and symmetrically, if r ends on a Y-modality, then r on $\tilde{u}$ is greater than or equal to r on $\tilde{v}$. These observations then allow to reduce $\tilde{u} \preccurlyeq^R_{m+1,n} \tilde{v}$ to $u \preccurlyeq^R_{m,n} v$. There are still some technicalities to overcome, though. For instance the look-around of a ranker may offset the actual position of a ranker a bit, so the first direction may actually lead into the center (yet not very far). In what follows, we formalize the proof.

Let $\mathrm{pos}(\tilde{u})$ and $\mathrm{pos}(\tilde{v})$ be the sets of positions of $\tilde{u}$ and $\tilde{v}$, respectively. We define environments (or *balls*) B_k of positions of $\tilde{u}$ and $\tilde{v}$ of decreasing size as k increases, and we introduce their complement C_k. Specifically, let

$$B_k(\tilde{u}) = \{i \in \mathrm{pos}(\tilde{u}) \mid |px^{2n}| + k < i \leqslant |px^{2n}u| - k\},$$
$$B_k(\tilde{v}) = \{i \in \mathrm{pos}(\tilde{u}) \mid |px^{2n}| + k < i \leqslant |px^{2n}v| - k\},$$

where $0 \leqslant k \leqslant n$, and let $C_k(\tilde{u}) = \mathrm{pos}(\tilde{u}) \setminus B_k(\tilde{u})$ and $C_k(\tilde{v}) = \mathrm{pos}(\tilde{v}) \setminus B_k(\tilde{v})$. There is a unique order-preserving bijection π between $C_n(\tilde{u})$ and $C_n(\tilde{v})$, and by abuse of notation, we henceforth identify $C_n(\tilde{u})$ and $C_n(\tilde{v})$ via π. Using this convention allows to drop the argument for C_k by setting $C_k = C_k(\tilde{u}) = C_k(\tilde{v})$. Whenever ambiguities might arise, we shall make clear whether by $i \in C_n$ we mean a position of $\tilde{u}$ or a position of $\tilde{v}$.

Claim 1 Let $k \leqslant n$. If $r \in \widetilde{R}_{1,k}$, then $r(\tilde{u}) = r(\tilde{v}) \in C_{k-1}$.

It suffices to consider $r \in \widetilde{R}^{\mathrm{X}}_{1,k}$ by symmetry. We actually prove a slightly stronger statement using induction on the length of the ranker. To formulate this, we have to introduce notation. Let $C = C_0$, *i.e.*, C comprises all positions outside the center, and subdivide C according to the number of factors x^2 which lie between a position and the center. Specifically, let $D_\ell = \{i \in C \cup \{-\infty\} \mid i \leqslant |px^{2\ell}|\}$. Let r be a ranker in $\widetilde{R}^{\mathrm{X}}_{1,k}$ with $|r| + \ell \leqslant k$ and $i \in D_\ell$. We claim $r(\tilde{u}, i) = r(\tilde{v}, i) \in C_{k-1}$. This implies the claim as $-\infty \in D_0$.

Let $r = \mathbb{XX}_w r'$ or $r = \mathrm{X}_w r'$ for some $(k-1)$-context w. In the following we consider only the case $r = \mathbb{XX}_w r'$; the case $r = \mathrm{X}_w r'$ is similar and actually slightly easier. The assumption implies $\mathrm{alph}_{2n-1}(xu) \cup \mathrm{alph}_{2n-1}(xv) \subseteq \mathrm{alph}_{2n-1}(xx)$, and we thus either have $\mathrm{X\!\!X}_w(\tilde{u}, i) = \mathrm{X\!\!X}_w(\tilde{v}, i) \in D_{\ell+1}$ or $\mathrm{X\!\!X}_w(\tilde{u}, i) + k - 1 = \mathrm{X\!\!X}_w(\tilde{v}, i) + k - 1 \in C$, that is, $\mathrm{X\!\!X}_w$ either simultaneously performs a small jump on both words, leading to a position in $D_{\ell+1}$, or it simultaneously makes a huge leap over the center. In the latter case $\mathrm{X\!\!X}_w$ necessarily involves a position of the suffix $y^{2n}q$ and thus the resulting position has to be near the right part of the environment C.

Note that the modality $\mathrm{X\!\!X}_w$ skips position $i + 1$ which by chance might be the only position with context w within the next $|x|$ positions. In such a case $\mathrm{X\!\!X}_w$ would indeed skip a whole factor x, but, in any case, it gets stuck at the latest in the second repetition of the factor x thereafter.[3] A starting modality X_w, in contrast, always gets stuck in the first factor. In this sense, this case is easier, and in particular the same reasoning is also valid for $r = \mathrm{X}_w r'$.

If $\mathrm{X\!\!X}_w(\tilde{u}, i) = \mathrm{X\!\!X}_w(\tilde{v}, i) \in D_{\ell+1}$, then induction yields the claim. Otherwise, if $j = \mathrm{X\!\!X}_w(\tilde{u}, i) = \mathrm{X\!\!X}_w(\tilde{v}, i)$ with $j+k-1 \in C$, then also $r'(\tilde{u}, j)+k-1 = r'(\tilde{v}, j)+k-1 \in C$ as r' consists solely of X-modalities. In particular, $r(\tilde{u}) = r(\tilde{v}) \in C_{k-1}$. This concludes the proof of the claim.

This claim already implies $\tilde{u} \preccurlyeq^R_{1,n} \tilde{v}$ and thus the statement of the lemma for $m = 0$. We therefore suppose $m \geqslant 1$ henceforth. We know what happens with the first direction of a ranker. For the remaining m alternations, the ranker may enter and leave the center of the words $\tilde{u}$ and $\tilde{v}$. To control this is the concern of statements (1) and (2) of the following claim. Statement (3) gives a means to transfer the evaluation of certain rankers from $\tilde{u}$ and $\tilde{v}$ to u and v.

Claim 2 Let $1 \leqslant \ell \leqslant n$, let $r \in \widetilde{R}_{m,\ell}$, let $k = \ell - |r|$, and let $i \in C_{k-1}$.

1. If $r(\tilde{v}, i)$ is defined, then $r(\tilde{u}, i)$ is defined and the following hold:

– $r(\tilde{v}, i) \in B_{\ell-1}(\tilde{v})$ implies $r(\tilde{u}, i) \in B_{\ell-1}(\tilde{u})$, and

– $r(\tilde{v}, i) \in C_{\ell-1}$ implies $r(\tilde{u}, i) \leqslant r(\tilde{v}, i)$ if $r \in \widetilde{R}^{\mathrm{X}}$, and $r(\tilde{u}, i) \geqslant r(\tilde{v}, i)$ if $r \in \widetilde{R}^{\mathrm{Y}}$, where $r(\tilde{v}, i)$ is interpreted as a position of $\tilde{u}$.

2. If $r \in \widetilde{R}_{m-1,\ell}$, then $r(\tilde{v}, i)$ is defined if and only if $r(\tilde{u}, i)$ is defined and:

– $r(\tilde{v}, i) \in B_{\ell-1}(\tilde{v})$ if and only if $r(\tilde{u}, i) \in B_{\ell-1}(\tilde{u})$, and

– $r(\tilde{v}, i) \in C_{\ell-1}$ implies $r(\tilde{u}, i) = r(\tilde{v}, i)$.

3. If $r \in \widetilde{R}^{\mathrm{Z}}_{m,\ell}$ with $\mathrm{Z} \in \{\mathrm{X}, \mathrm{Y}\}$ and $r(\tilde{v}, i) \in B_{\ell-1}(\tilde{v})$, then there exists a ranker $\hat{r} \in \widetilde{R}^{\mathrm{Z}}_{m,\ell}(v)$ such that $\hat{r}(u)$ is defined and the following hold:

$$r(\tilde{u}, i) = \hat{r}(u) + |px^{2n}|,$$
$$r(\tilde{v}, i) = \hat{r}(v) + |px^{2n}|.$$

The proof is by induction on m. Statement (1) inductively yields (2). (Note that the latter is trivial for $m = 1$ and that $u \preccurlyeq^R_{m-1,n} v$ as well as $v \preccurlyeq^R_{m-1,n} u$.) We shall show (2) $\Rightarrow$ (3) and (2) & (3) $\Rightarrow$ (1).

(2) $\Rightarrow$ (3): Suppose there exists a non-empty prefix s of r such that $s(\tilde{v}, i) \in C_{k+|s|-1}$. There also exists such a prefix with $s \in \widetilde{R}_{m-1,\ell}$, which yields $s(\tilde{v}, i) = s(\tilde{v}, i)$ by (2). In

[3]Maybe an example clarifies more than a thousand words: Consider the word $u = (baa)^3$ and the position 3 of the second a. For visual presentation of this example, we represent the position by an underscore. With $w = (a, b, a)$ the modality $\mathrm{X\!\!X}_w$ skips over a complete factor baa, and we get $\mathrm{X\!\!X}_w(ba\underline{a}baabaa) = baabaa\underline{b}aa$, whereas $\mathrm{X}_w(ba\underline{a}baabaa) = baa\underline{b}aabaa$.

this case the ranker $\hat{r}$ is obtained by induction on the length of the remaining suffix of the ranker r.

We may thus assume that $s(\tilde{v}, i) \in B_{k+|s|-1}(\tilde{v})$ for all non-empty prefixes s of r. Intuitively this means that the evaluation of $r(\tilde{v}, i)$ never involves a position outside the environment $B_0(\tilde{v})$ — not even as a look-around of a modality. In other words the evaluation of $r(\tilde{v}, i)$ is oblivious of the prefix px^{2n} and the suffix $y^{2n}q$ of $\tilde{v}$, and the ranker can be transferred to v.

We still have to handle the first modality, which we have to start from the right position. First suppose $r = \mathrm{X}_w s$ for some w and some s. Let $j = i - |px^{2n}|$, and let $a_1 \cdots a_k$ be the prefix of length k of v (or equivalently of u). By assumption, the position i must lie to the left of the center, *i.e.*, $j \leqslant k - 1$. If $j \leqslant 0$, then set $\hat{r} = r$; otherwise set $\hat{r} = \mathrm{X}_{a_1} \cdots \mathrm{X}_{a_j} r$. Note that $\hat{r} \in \widetilde{R}_{m,\ell}$ as $j < k$ and that $\mathrm{X}_{a_1} \cdots \mathrm{X}_{a_j}(u) = \mathrm{X}_{a_1} \cdots \mathrm{X}_{a_j}(v) = j$. By construction $\hat{r}(v) = r(\tilde{v}, i) - |px^{2n}|$.

Suppose there exists a prefix $s\mathrm{Z}_w$ of r with $\mathrm{Z} \in \{\mathrm{X}, \mathrm{XX}, \mathrm{Y}, \mathrm{YY}\}$ such that the last modality Z_w involves a position in C_0 (potentially for the look-around). Consider the shortest such prefix. Disregarding the last modality, $s(\tilde{u}, i)$ is oblivious of the positions in C by minimality of $|s\mathrm{Z}_w|$. In other words $\hat{s}(u) = s(\tilde{u}, i) - |px^{2n}|$. Hence $\hat{s}\mathrm{Z}_w(u)$ must be undefined by choice of $s\mathrm{Z}_w$. This contradicts $u \preccurlyeq^R_{m,n} v$, because the ranker $\hat{s}\mathrm{Z}_w \in \widetilde{R}_{m,n}$ is defined on v, but not on u. Therefore, there is no such prefix of r, the evaluation of $r(\tilde{u}, i)$ is oblivious of the positions in C, and $\hat{r}(u) = r(\tilde{u}, i) - |px^{2n}|$.

A similar argument applies if $r = \mathrm{XX}_w s$, which we thus state only briefly. Consider $j = 1 + i - |px^{2n}|$. The offset $+1$ compensates for the fact that XX_w skips position $i + 1$. If $j \leqslant 0$, then set $\hat{r} = \mathrm{X}_w s$; else set $\hat{r} = \mathrm{X}_{a_1} \cdots \mathrm{X}_{a_j} \mathrm{X}_w s$. Note that $j \leqslant k$ and thus $\hat{r} \in \widetilde{R}_{m,\ell}$. The constructions for the cases $r = \mathrm{Y}_w s$ and $r = \mathrm{YY}_w s$ are left-right symmetric.

(2) & (3) $\Rightarrow$ (1): Due to symmetry it suffices to consider the case $r \in \widetilde{R}^{\mathrm{X}}_{m,\ell}$ only. Let $r(\tilde{v}, i)$ be defined. First suppose $r(\tilde{v}, i) \in B_{\ell-1}(\tilde{v})$. We have to show $r(\tilde{u}, i) \in B_{\ell-1}(\tilde{u})$. Let $\hat{r}$ be the ranker from (3) for r, so that $r(\tilde{u}, i) = \hat{r}(u) + |px^{2n}|$. By $u \preccurlyeq^R_{m,n} v$ it is easy to see that, for $k \leqslant n$, if $\hat{r}(v)$ has distance at least $k - 1$ to the last position of v, then $\hat{r}(u)$ has also distance at least $k - 1$ to the last position of u. More formally, we have that

$$\hat{r}(v) + k - 1 \leqslant |v| \text{ implies } \hat{r}(u) + k - 1 \leqslant |u| \text{ for } k \in \{1, \ldots, n\}. \tag{12.1}$$

This yields $r(\tilde{u}, i) \in B_{\ell-1}(\tilde{u})$.

Suppose now $r(\tilde{v}, i) \in C_{\ell-1}$. We have to show $r(\tilde{u}, i) \leqslant r(\tilde{v}, i)$.

Suppose further that there exists a prefix s of r such that $s(\tilde{v}, i) \in B_{k+|s|-1}(\tilde{v})$. Let s be a longest such prefix, and suppose $\mathrm{Z}_w \in \{\mathrm{X}_w, \mathrm{XX}_w, \mathrm{Y}_w, \mathrm{YY}_w\}$ is such that $s\mathrm{Z}_w$ is a prefix of r. If $s\mathrm{Z}_w \in \widetilde{R}_{m-1,\ell}$, then $s\mathrm{Z}_w(\tilde{u}, i) = s\mathrm{Z}_w(\tilde{v}, i)$ by (2), and induction yields $r(\tilde{u}, i) \leqslant r(\tilde{v}, i)$. Hence $s\mathrm{Z}_w$ uses the full alternation parameter m; in particular $\mathrm{Z}_w \in \{\mathrm{X}_w, \mathrm{XX}_w\}$. Let $\hat{s}$ be the ranker from (3) for s. We have that $\hat{s}(v)$ as well as $\hat{s}(u)$ are defined, and $\hat{s}\mathrm{Z}_w(v)$ is undefined by choice of $s\mathrm{Z}_w$.

We distinguish whether $\hat{s}\mathrm{Z}_w(u)$ is defined or not. The first case is $\hat{s}\mathrm{Z}_w(u)$ being undefined. This means that $s\mathrm{Z}_w(\tilde{v}, i)$ and $s\mathrm{Z}_w(\tilde{u}, i)$ both leave the center simultaneously with the last modality. Since the words $\tilde{u}$ and $\tilde{v}$ are the same outside the center, we have $s\mathrm{Z}_w(\tilde{u}, i) = s\mathrm{Z}_w(\tilde{v}, i)$. The second case is $\hat{s}\mathrm{Z}_w(u)$ being defined. This means $s\mathrm{Z}_w(\tilde{u}, i)$ stays in the center, while $s\mathrm{Z}_w(\tilde{v}, i)$ leaves it with an X-modality. Therefore, interpreting

$s\mathrm{Z}_w(\tilde{v}, i)$ as a position of $\tilde{u}$, we get $s\mathrm{Z}_w(\tilde{u}, i) \leqslant s\mathrm{Z}_w(\tilde{v}, i)$. Consider the suffix r' of r such that $r = s\mathrm{Z}_w r'$. The suffix r' is unidirectional and $r'(\tilde{u}, s\mathrm{Z}_w(\tilde{v}, i)) = r'(\tilde{v}, s\mathrm{Z}_w(\tilde{v}, i)) = r(\tilde{v}, i)$. In both cases we thus have $r(\tilde{u}, i) \leqslant r(\tilde{v}, i)$.

By these considerations, we may assume $s(\tilde{v}, i) \in C_{k+|s|-1}$ for all prefixes s of r. If for every prefix s of r we further have $s(\tilde{u}, i) \in C_{k+|s|-1}$, then $s(\tilde{u}, i) = s(\tilde{v}, i)$. In particular $r(\tilde{u}, i) = r(\tilde{v}, i)$. It remains to consider $r = sr'$ such that $s(\tilde{u}, i) \in B_{k+|s|-1}(\tilde{u})$ for some s. We know $s \notin \widetilde{R}_{m-1,k+|s|}$ by (2). This means $s \in \widetilde{R}^{\mathrm{X}}_{m,k+|s|}$ and $r' \in \widetilde{R}^{\mathrm{X}}_{1,\ell}$. We thus have $s(\tilde{u}, i) \leqslant s(\tilde{v}, i)$, where $s(\tilde{v}, i)$ is interpreted as a position of $\tilde{u}$. As the suffix r' of r is unidirectional, $r(\tilde{u}, i) = r'(\tilde{u}, s(\tilde{u}, i)) \leqslant r'(\tilde{u}, s(\tilde{v}, i)) = r(\tilde{v}, i)$.

That concludes the proof of the claim.

After these preparatory claims, we finally come to prove the lemma. We show $\tilde{u} \preccurlyeq^R_{m+1,n} \tilde{v}$ by proving conditions (1) to (3) of Definition 11.15 one after another. For this consider rankers $r \in \widetilde{R}_{m+1,n}(\tilde{v})$ in the following, from which we split off the first direction; *i.e.*, let $r = r'r''$ such that $r' \in \widetilde{R}_{1,n}$ and $r'' \in \{\varepsilon\} \cup \widetilde{R}_{m,n}$. By Claim 1 we have $r'(\tilde{u}) = r'(\tilde{v})$ and setting $i = r'(\tilde{v})$ we have $i \in C_{|r'|-1}$.

Definition 11.15 (1): We have to show that r is defined on $\tilde{u}$. There is nothing to show if $r'' = \varepsilon$. Suppose otherwise $r'' \in \widetilde{R}_{m,n}$ has length at most $n - |r'|$. By assumption $r''(\tilde{v}, i)$ is defined and hence so is $r(\tilde{u}) = r''(\tilde{u}, i)$ by statement (1) in Claim 2.

Definition 11.15 (2): Assume that r ends on an X-modality, and consider a ranker $s \in \widetilde{R}^{\mathrm{X}}_{m,n-1}(\tilde{v}) \cup \widetilde{R}^{\mathrm{Y}}_{m+1,n-1}(\tilde{v})$ and $d \in \{-1, \ldots, 2\}$. Suppose $r(\tilde{v}) \leqslant s(\tilde{v}) - d$. We have to show $r(\tilde{u}) \leqslant s(\tilde{u}) - d$. For s we also split off the first direction and let $s = s's''$ be such that $s' \in \widetilde{R}_{1,n}$ and $s'' \in \{\varepsilon\} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1} \cup \widetilde{R}^{\mathrm{Y}}_{m,n-1}$. By Claim 1, setting $j = s'(\tilde{u}) = s'(\tilde{v})$ is well-defined and $j \in C_{|s'|-1}$. We shall show $r''(\tilde{u}, i) \leqslant s''(\tilde{u}, j) - d$, which yields $r(\tilde{u}) \leqslant s(\tilde{u}) - d$ as desired. We distinguish cases depending on whether or not $r''(\tilde{v}, i) \in B_{n-1}(\tilde{v})$ and whether or not $s''(\tilde{v}, j) \in B_{n-2}(\tilde{v})$

First suppose $r''(\tilde{v}, i) \in B_{n-1}(\tilde{v})$ and $s''(\tilde{v}, j) \in B_{n-2}(\tilde{v})$. Let $\hat{r}$ and $\hat{s}$ be the rankers corresponding to r'' and s'', respectively, by statement (3) in Claim 2. For $w \in \{u, v\}$ we have $r''(\tilde{w}, i) = \hat{r}(w) + |px^{2n}|$ and $s''(\tilde{w}, j) = \hat{s}(w) + |px^{2n}|$. Since $r(\tilde{v}) \leqslant s(\tilde{v}) - d$, we have $\hat{r}(v) \leqslant \hat{s}(v) - d$. This in turn yields $\hat{r}(u) \leqslant \hat{s}(u) - d$ by the assumption $u \preccurlyeq^R_{m,n} v$; note that $\hat{r} \in \widetilde{R}^{\mathrm{X}}_{m,n}$ and $\hat{s} \in \widetilde{R}^{\mathrm{Y}}_{m,n-1} \cup \widetilde{R}^{\mathrm{X}}_{m-1,n-1}$. This shows $r''(\tilde{u}, i) \leqslant s''(\tilde{u}, j) - d$.

Next suppose that not both $r''(\tilde{v}, i) \in B_{n-1}(\tilde{v})$ and $s''(\tilde{v}, j) \in B_{n-2}(\tilde{v})$. By Claim 2 we see that

- if $r''(\tilde{v}, i) \in C_{n-1}$, then $r''(\tilde{u}, i) \leqslant r''(\tilde{v}, i)$ and
- if $s''(\tilde{v}, j) \in C_{n-2}$, then $s''(\tilde{v}, j) \leqslant s''(\tilde{u}, j)$.

Supposing $r''(\tilde{v}, i) \in C_{n-1}$ as well as $s''(\tilde{v}, j) \in C_{n-2}$, this immediately yields that $r''(\tilde{u}, i) \leqslant r''(\tilde{v}, i) \leqslant s''(\tilde{v}, j) - d \leqslant s''(\tilde{u}, j) - d$. Suppose now $r''(\tilde{v}, i) \in B_{n-1}(\tilde{v})$ and $s''(\tilde{v}, j) \in C_{n-2}$. Using equation (12.1) and the ranker $\hat{r}$ corresponding to r'' given by statement (3) in Claim 2, it is straightforward to verify $r''(\tilde{u}, i) \leqslant s''(\tilde{u}, j) - d$. The last remaining case ($r''(\tilde{v}, i) \in C_{n-1}$ and $s''(\tilde{v}, j) \in B_{n-2}(\tilde{v})$) follows by a similar argument.

Definition 11.15 (3) is left-right symmetric (2). That concludes the proof of the lemma. □

We are now able to state and prove the main results. For the full levels FO^2_m this is the following.

Proposition 12.50
Let $L \subseteq A^+$, and let $m \geqslant 2$ be an integer. The following are equivalent:

1. *L is definable in* $\mathrm{FO}^2_m[<, \mathrm{suc}]$.
2. *L is definable in* $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$.
3. *The ordered syntactic semigroup S_L satisfies $U_m = V_m$.*

Proof. We show (1) ⇔ (2) and (1) ⇔ (3).

(1) ⇒ (2): This is trivial.

(2) ⇒ (1): This follows from Corollary 9.7. Note that $\mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$ for $m, n \geqslant 2$ can specify prefixes and suffixes of length up to $2n - 1$.

(1) ⇒ (3): For $i \geqslant 1$ define words u_i and v_i, parameterized by n, by the recursion

$$u_1 = (e^{2n} s f^{2n} x_1)^{2n} e^{2n} s f^{2n} (y_1 e^{2n} t f^{2n})^{2n}, \quad u_i = (u_{i-1} x_i)^{2n} u_{i-1} (y_i u_{i-1})^{2n},$$
$$v_1 = (e^{2n} s f^{2n} x_1)^{2n} e^{2n} t f^{2n} (y_1 e^{2n} t f^{2n})^{2n}, \quad v_i = (u_{i-1} x_i)^{2n} v_{i-1} (y_i u_{i-1})^{2n},$$

where $e, f, s, t, x_i, y_i \in A^+$. This definition is guided by the omega-terms U_i and V_i in the sense that u_i and v_i are indeed instantiations of the omega-terms U_i and V_i with words substituted for the variables, and with the concrete power $2n$ substituted for the formal symbol ω.

We claim that there exists n such that $pu_iq \approx^R_{i,n} pv_iq$. The proof is by induction on i. The case $i = 1$ follows from Proposition 12.30. Suppose $i \geqslant 2$. The recursion ensures

$$\mathrm{alph}_{2n-1}(v_{i-1}) = \mathrm{alph}_{2n-1}(u_{i-1}) \subseteq \mathrm{alph}_{2n-1}(u_{i-1}x_i) \cap \mathrm{alph}_{2n-1}(y_i u_{i-1})$$

because the power $2n$ is sufficiently large. Also, the words u_{i-1} and v_{i-1} have a common prefix of length $2n$ as well as a common suffix of length $2n$. Induction yields $u_{i-1} \approx^R_{i-1,n} v_{i-1}$, and Lemma 12.49 shows $pu_iq \preccurlyeq^R_{i,n} pv_iq$ as well as $pv_iq \preccurlyeq^R_{i,n} pu_iq$ for all $p, q \in A^*$. This proves the claim.

In particular we see $pu_mq \approx^R_{m,n} pv_mq$ for some n. Choose such an n that in addition is large enough so that L is defined by some sentence $\varphi \in \mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$. By Corollary 11.20, $pu_mq \in L(\varphi)$ if and only if $pv_mq \in L(\varphi)$. This shows that S_L satisfies $U_m = V_m$.

(3) ⇒ (1): Let S satisfy $U_m = V_m$, and let $h \colon A^+ \to S$ be a surjective homomorphism recognizing L. Proposition 12.47 yields an integer n such that every preimage $h^{-1}(P)$ is a union of $\approx^{\mathrm{FO}^2}_{m,n}$-classes. Hence the language L is a union of $\approx^{\mathrm{FO}^2}_{m,n}$-classes. It is thus definable as the union over all $L(\varphi)$ with $L(\varphi) \subseteq L$ and $\varphi \in \mathrm{FO}^2_{m,n}[<, \mathrm{suc}]$. This union yields a finite disjunction, because there are only finitely many inequivalent formulae of bounded quantifier depth. □

For the half levels Σ^2_m we get the following.

Proposition 12.51
Let $L \subseteq A^+$, and let $m \geqslant 2$ be an integer. The following are equivalent:

1. *L is definable in* $\Sigma^2_m[<, \mathrm{suc}]$.
2. *L is definable in* $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$.
3. *The ordered syntactic semigroup S_L is in* **LDA** *and satisfies $U_{m-1/2} \leqslant V_{m-1/2}$.*

Proof. We shall show the equivalences (1) ⇔ (2) and (1) ⇔ (3).

(1) ⇒ (2): This is trivial.

(2) $\Rightarrow$ (1): This follows from Corollary 9.6. Note that $\Sigma^2_{m,n}[<, \mathrm{suc}]$ is sufficiently expressive to specify prefixes and suffixes of length up to $2n-1$.

(1) $\Rightarrow$ (3): Every $\Sigma^2_m[<, \mathrm{suc}]$-definable language is $\mathrm{FO}^2_m[<, \mathrm{suc}]$-definable. By Proposition 12.50 and Lemma 12.18 the syntactic semigroup is in **LDA**. It remains to show the identity $U_{m-1/2} \leqslant V_{m-1/2}$. This is done similarly to the previous proof.

For $i - 1/2 \in \mathbb{N}$ define words u_i and v_i depending on n by

$$\begin{aligned} U_{1/2} &= e^{2n} z e^{2n}, & u_i &= (u_{i-1}x_i)^{2n} u_{i-1} (y_i u_{i-1})^{2n}, \\ V_{1/2} &= e^{2n}, & v_i &= (u_{i-1}x_i)^{2n} v_{i-1} (y_i u_{i-1})^{2n}, \end{aligned}$$

where $e, z, x_i, y_i \in A^+$.

We show that there exists n such that $pu_iq \preccurlyeq^R_{i+1/2,n} pv_iq$. The case $i = 1/2$ follows from Proposition 12.20. Suppose $i \geqslant 3/2$. We have

$$\mathrm{alph}_{2n-1}(v_{i-1}) \subseteq \mathrm{alph}_{2n-1}(u_{i-1}) \subseteq \mathrm{alph}_{2n-1}(u_{i-1}x_i) \cap \mathrm{alph}_{2n-1}(y_i u_{i-1}).$$

Moreover, u_{i-1} and v_{i-1} have a common prefix of length $2n$ as well as a common suffix of length $2n$. Induction yields $pu_{i-1}q \preccurlyeq^R_{i-1/2,n} pv_{i-1}q$, and Lemma 12.49 shows $pu_iq \preccurlyeq^R_{i+1/2,n} pv_iq$ for all $p, q \in A^*$.

Therefore, $pu_{m-1/2}q \preccurlyeq^R_{m,n} pv_{m-1/2}q$ for some n. Choose such an n such that, moreover, L is defined by some $\varphi \in \Sigma^2_{m,n}[<, \mathrm{suc}]$. Corollary 11.17 yields $pu_{m-1/2}q \in L(\varphi)$ if $pv_{m-1/2}q \in L(\varphi)$. This shows that S_L satisfies $U_{m-1/2} \leqslant V_{m-1/2}$.

(3) $\Rightarrow$ (1): Let $h\colon A^+ \to S$ be a surjective homomorphism recognizing L for any ordered semigroup S in **LDA** which satisfies $U_{m-1/2} \leqslant V_{m-1/2}$; *i.e.*, there exists a $\leqslant$-order ideal I such that $L = h^{-1}(I)$. By Proposition 12.48 there exists an integer n such that every preimage of a $\leqslant$-order ideal is a $\preccurlyeq^{\mathrm{FO}^2}_{m,n}$-order ideal. In particular, L is a $\preccurlyeq^{\mathrm{FO}^2}_{m,n}$-order ideal and thus definable as the union of all $L(\varphi)$ over $\varphi \in \Sigma^2_{m,n}[<, \mathrm{suc}]$ such that $L(\varphi) \subseteq L$. This union is finite because there are only finitely many languages definable in $\Sigma^2_{m,n}[<, \mathrm{suc}]$. □

12.4. Quantifier Alternation without Successor Predicate

This section proves Theorem 12.11 (namely that $\Sigma^2_m[<]$ corresponds to monoids in **DA** that satisfy $U'_{m-1/2} \leqslant V'_{m-1/2}$) as well as Theorem 12.12 (namely that $\mathrm{FO}^2_m[<]$ corresponds to monoids satisfying $U'_m = V'_m$).

The step from algebra to logic (typically considered the hard part in such correspondences) is done by a reduction to Theorem 12.1 and Theorem 12.2, using the neutral letter approach given in Section 8.2. For this approach to work, the existence of sliding window formulae in $\Sigma^2_m[<]$ for formulae in $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$ has to be established. This is most conveniently done by a detour via temporal logic.

The other direction from logic to algebra is very much along the same lines as in the case with successor, albeit much less technical. For this, we use the ranker description of the alternation hierarchy from Section 11.3.

Let us start by showing that there are sliding window formulae in $\Sigma^2_m[<]$ for formulae in $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$.

Lemma 12.52
For every $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$-sentence φ there exists an integer $N \geqslant 1$ and a sliding window formula $\varphi_N \in \Sigma^2_m[<]$ for φ of radius N.

Proof. Lemma 11.23 shows that there exists a TL-formula φ' that is equivalent to φ such that $\varphi' \in \mathrm{TL}_m^+[\mathrm{XF}, \mathrm{YP}, \mathrm{X}, \mathrm{Y}, \min, \max]$. Note that $\mathrm{XXF}\varphi$ is equivalent to $\mathrm{X}(\mathrm{XF}\varphi)$, and $\mathrm{YYP}\varphi$ is equivalent to $\mathrm{Y}(\mathrm{YP}\varphi)$.

The next step is to pull the modalities X and Y inwards towards atomic formula. For X this is done using repeatedly the distributive properties

$$\begin{aligned} \mathrm{X}(\varphi \vee \psi) &\equiv \mathrm{X}\varphi \vee \mathrm{X}\psi, \\ \mathrm{X}(\varphi \wedge \psi) &\equiv \mathrm{X}\varphi \wedge \mathrm{X}\psi, \end{aligned}$$

the equivalences

$$\begin{aligned} \mathrm{X}(\neg\varphi) &\equiv \mathrm{X}\top \wedge \neg\mathrm{X}\varphi, \\ \mathrm{X}(\mathrm{XF}(\varphi) &\equiv \mathrm{XF}(\mathrm{X}\varphi), \end{aligned}$$

and the annihilation rules

$$\begin{aligned} \mathrm{XY}\varphi &\equiv \mathrm{X}\top \wedge \varphi, \\ \mathrm{X}(\mathrm{YP}\varphi) &\equiv \mathrm{X}\top \wedge (\varphi \vee \mathrm{YP}\varphi). \end{aligned}$$

To move Y inwards, the dual variants with future and past modalities interchanged are used. This yields an equivalent formula φ'' with the property that if $\mathrm{X}\varphi$ is a subformula, then φ is of the form $\mathrm{X}^k\psi$ with $\psi \in \{\top, \bot, a, \min, \max\}$ for some $a \in A$. Symmetrically, if $\mathrm{Y}\varphi$ is a subformula, then φ is of the form $\mathrm{Y}^k\psi$ for such a ψ. In other words, X and Y may only occur as a cluster right before atomic formulae and only in pure form. Intuitively, this allows to displace atomic formulae by a certain amount, and all necessary information is provided by the look-around of the sliding window.

We can now give the desired sliding window formula. Let N be such that $\mathrm{X}^k\psi$ or $\mathrm{Y}^k\psi$ being a subformula of φ'' implies $N \geqslant k+1$. A formula $\varphi \in \mathrm{TL}$ can be viewed as a FO^2-formula $\langle\varphi\rangle(x)$ in one free variable. Specifically, we have

$$\begin{aligned} \langle \mathrm{XF}\psi\rangle(x) &:= \exists y\, \big(y > x \wedge \langle\psi\rangle(y)\big), \\ \langle \mathrm{YP}\psi\rangle(x) &:= \exists y\, \big(y < x \wedge \langle\psi\rangle(y)\big) \end{aligned}$$

and as usual $\langle\varphi \vee \psi\rangle(x) := \langle\varphi\rangle(x) \vee \langle\psi\rangle(x)$ and $\langle\varphi \wedge \psi\rangle(x) := \langle\varphi\rangle(x) \wedge \langle\psi\rangle(x)$ as well as $\langle\neg\varphi\rangle(x) := \neg\langle\varphi\rangle(x)$.

To express X and Y in terms of first-order logic without losing an alternation, we normally need a successor predicate. The normal form ensures that these modalities occur only in the following very specific contexts:

$$\begin{aligned} \langle \mathrm{X}^k \bot\rangle(x) &:= \bot, \\ \langle \mathrm{X}^k \top\rangle(x) &:= (\lambda(x+k) \neq \square), \\ \langle \mathrm{X}^k a\rangle(x) &:= (\lambda(x+k) = a), \\ \langle \mathrm{X}^k \min\rangle(x) &:= (\lambda(x+k) \neq \square) \wedge (\lambda(x+k-1) = \square), \\ \langle \mathrm{X}^k \max\rangle(x) &:= (\lambda(x+k) \neq \square) \wedge (\lambda(x+k+1) = \square). \end{aligned}$$

Here, $\lambda(x+k) = a$ for $k \in \{-N, \ldots, N\}$ and $a \in A \cup \{\square\}$ is an abbreviation for the label $\lambda(x) \in B$ over the sliding window alphabet

$$B = \{[uav] \mid u, v \in (A \cup \{\square\})^*,\ |u| = N+k,\ |v| = N-k\}.$$

Formulae of the form $\mathrm{Y}^k\psi$ are dealt with similarly. (Actually, they may be identified with $\mathrm{X}^{-k}\psi$ in the above construction).

Negations only come from negations, and the only numerical predicate used to define $\langle\varphi\rangle(x)$ is the order predicate. Hence, starting with a formula φ in $\Sigma_m^2[<, \mathrm{suc}, \min, \max]$, this procedure yields a sliding window formula for φ in $\Sigma_m^2[<]$. $\square$

The preceding lemma implies that sliding window formulae exist also for every full level. Indeed, any $\mathrm{FO}_m^2[<, \mathrm{suc}, \min, \max]$-sentence φ can be seen as a Boolean combination of $\Sigma_m^2[<, \mathrm{suc}, \min, \max]$-sentences ψ_i. Increasing the sliding window diameter is no problem. Hence, Lemma 12.52 yields an integer N such that each ψ_i possesses a sliding window formula $\psi_{i,N} \in \Sigma_m^2[<]$ of radius N. The Boolean combination of the $\psi_{i,N}$ according to φ thus yields a sliding window formula $\varphi_N \in \mathrm{FO}_m^2[<]$ for φ.

This suffices to show the implication from algebra to logic. The reverse implication is handled similarly as in the case with successor. We again give a lifting lemma for rankers similar to Lemma 12.49, this time for ordinary rankers. As no look-around is available, the formulation and the proof are much less technical; *e.g.*, we do not need to take factors into account but only single letter alphabets.

Lemma 12.53 (Ranker Lifting)
Let $m \geqslant 0$, $n \geqslant 1$, and $u, v, x, y \in A^*$. Suppose $\mathrm{alph}(u) \cup \mathrm{alph}(v) \subseteq \mathrm{alph}(x) \cap \mathrm{alph}(y)$. *If $u \leqslant_{m,n}^R v$, then $px^nuy^nq \leqslant_{m+1,n}^R px^nvy^nq$ for all $p, q \in A^*$.*

Proof. Let $\tilde{u} = px^nuy^nq$ and $\tilde{v} = px^nvy^nq$. The middle factors u and v are the *center* of $\tilde{u}$ and $\tilde{v}$, respectively. The idea is to reduce $\tilde{u} \leqslant_{m+1,n}^R \tilde{v}$ to $u \leqslant_{m,n}^R v$. Observe that all rankers involved in $\tilde{u} \leqslant_{m+1,n}^R \tilde{v}$ are in $R_{m+1,n}$. We study the behavior of such rankers on $\tilde{u}$ and $\tilde{v}$.

Take a ranker in $R_{m+1,n}$. The first direction of modalities is consumed to get near the center: The number of repetitions of x is large enough so that a ranker consisting solely of X-modalities gets stuck in the prefix px^n or in the suffix y^nq. What is more, such a unidirectional ranker leads to the same position in the prefix or suffix in both $\tilde{u}$ and $\tilde{v}$.

After processing the first block, the ranker has thus m blocks left. Consider such a ranker $r \in R_{m,n}$. This ranker cannot directly be taken for evaluation on u and v. For one thing, r starts outside the center. And for another, r may intermediately jump to positions outside the center; in the extreme it may actually never enter u or v in the center.

The greater part of the effort goes into controlling the behavior of $r \in R_{m,n}$ with respect to the center. For example a ranker $r \in R_{m-1,n}$ is in the center of $\tilde{u}$ anytime during its evaluation if and only if it is in the center of $\tilde{v}$. If it later on re-enters the center, we may thus take a shortcut (namely a suffix of r). For a ranker r with full alternation parameter m this is no longer true. But r leaves the center of $\tilde{u}$ at the earliest when r leaves the center of $\tilde{v}$. By the observations we just saw, we may assume that this happens with the last direction. If r leaves the center with an X-modality, this means that it is not greater on $\tilde{u}$ than on $\tilde{v}$. (If r leaves with a Y-modality, it is not smaller on $\tilde{u}$ than on $\tilde{v}$.) The unidirectional remainder of the ranker preserves this as an invariant. In what follows, we formalize the proof.

Let the sets of positions of $\tilde{u}$ and $\tilde{v}$ be denoted by $\mathrm{pos}(\tilde{u})$ and $\mathrm{pos}(\tilde{v})$, respectively. The *center* B of $\tilde{u}$ and $\tilde{v}$ are the positions corresponding to the occurrence of u and v in the center, *i.e.*,

$$B(\tilde{u}) = \{i \in \mathrm{pos}(\tilde{u}) \mid |px^n| < i \leqslant |px^nu|\},$$
$$B(\tilde{v}) = \{i \in \mathrm{pos}(\tilde{u}) \mid |px^n| < i \leqslant |px^nv|\}.$$

We also let $C(\tilde{u}) = \mathrm{pos}(\tilde{u}) \setminus B(\tilde{u})$ and $C(\tilde{v}) = \mathrm{pos}(\tilde{v}) \setminus B(\tilde{v})$ be the positions outside the center. We henceforth identify the positions $C(\tilde{u})$ and $C(\tilde{v})$ outside the center in the canonical way, and we often drop the argument for C by setting $C_k = C(\tilde{u}) = C(\tilde{v})$. We shall be explicit whether by $i \in C$ we mean a position of $\tilde{u}$ or $\tilde{v}$ whenever ambiguities might arise.

The first block of direction cannot lead into the center.

Claim 3 Let $k \leqslant n$. If $r \in R_{1,k}$, then $r(\tilde{u}) = r(\tilde{v}) \in C$.

There is little to show for this claim, but let us give the main arguments. It suffices to consider $r \in \widetilde{R}^{\mathrm{X}}_{1,k}$ by symmetry. All letters of u and of v also occur in x. There are two cases of what can happen with a modality X_a. If $a \in \mathrm{alph}(x)$, then on both $\tilde{u}$ and $\tilde{v}$ the jump leads at most as far as to the next factor x. The center thus cannot be reached as there are to many occurrences of x in front of the center in both words. If $a \notin \mathrm{alph}(x)$, then $a \notin \mathrm{alph}(u)$ and $a \notin \mathrm{alph}(v)$ by assumption. Hence, the modality leads into the suffix $y^n q$ in both $\tilde{u}$ and $\tilde{v}$. Moreover, as the suffix is the same in $\tilde{u}$ and $\tilde{v}$, the same position is reached by X_a on both words. The ranker cannot change the direction, so it stays in these suffixes. This concludes the proof of the claim.

This already suffices to show $\tilde{u} \leqslant^R_{1,n} \tilde{v}$, and we may assume $m \geqslant 1$ in the following. In particular $\mathrm{alph}(v) \subseteq \mathrm{alph}(u)$.

We already know what happens with the first direction of a ranker. With the remaining m alternations the ranker may enter or leave the center at will. There is an interdependence between leaving $\tilde{u}$ and $\tilde{v}$, which are given by (1) and (2) in the following claim. Statement (3) shows that we can transfer the evaluation of certain rankers from $\tilde{u}$ and $\tilde{v}$ to u and v.

Claim 4 Let $1 \leqslant \ell \leqslant n$, let $r \in R_{m,\ell}$, and let $i \in C$.

1. If $r(\tilde{v}, i)$ is defined, then $r(\tilde{u}, i)$ is defined and the following hold:
– $r(\tilde{v}, i) \in B(\tilde{v})$ implies $r(\tilde{u}, i) \in B(\tilde{u})$, and
– $r(\tilde{v}, i) \in C$ implies $r(\tilde{u}, i) \leqslant r(\tilde{v}, i)$ for r ending on an X-modality, and $r(\tilde{u}, i) \geqslant r(\tilde{v}, i)$ for r ending on a Y-modality. Here, $r(\tilde{v}, i)$ is interpreted as a position of $\tilde{u}$.

2. If $r \in R_{m-1,\ell}$, then $r(\tilde{v}, i)$ is defined if and only if $r(\tilde{u}, i)$ is defined and:
– $r(\tilde{v}, i) \in B(\tilde{v})$ if and only if $r(\tilde{u}, i) \in B(\tilde{u})$, and
– $r(\tilde{v}, i) \in C$ implies $r(\tilde{u}, i) = r(\tilde{v}, i)$.

3. If $\mathrm{Z} \in \{\mathrm{X}, \mathrm{Y}\}$ such that $r \in R^{\mathrm{Z}}_{m,\ell}$ and if $r(\tilde{v}, i) \in B(\tilde{v})$, then there exists a ranker $\hat{r} \in R^{\mathrm{Z}}_{m,\ell}(v)$ such that $\hat{r}(u)$ is defined and the following hold:

$$
\begin{aligned}
r(\tilde{u}, i) &= \hat{r}(u) + |px^n|,\\
r(\tilde{v}, i) &= \hat{r}(v) + |px^n|.
\end{aligned}
$$

The proof is by induction on m. Statement (1) inductively yields (2). (Note that the latter is trivial for $m = 1$ and that $u \leqslant^R_{m-1,n} v$ as well as $v \leqslant^R_{m-1,n} u$.) We are going to show (2) $\Rightarrow$ (3) and (2) & (3) $\Rightarrow$ (1).

(2) $\Rightarrow$ (3): Suppose that there exists a non-empty prefix s of r such that $s(\tilde{v}, i) \in C$. Necessarily, s is in $R_{m-1,\ell}$. This yields $s(\tilde{v}, i) = s(\tilde{v}, i)$ by (2), and $\hat{r}$ is obtained by induction on the length.

We may thus assume that $s(\tilde{v}, i) \in B(\tilde{v})$ for all non-empty prefixes s of r. This means that $r(\tilde{v}, i)$ never leaves the center, and the evaluation of $r(\tilde{v}, i)$ is oblivious of the prefix

px^n and the suffix $y^n q$ of $\tilde{v}$. Therefore, the ranker can be transferred to v in the sense that $r(\tilde{v}, i) = r(v) + |px^n|$. We claim that $r(\tilde{u}, i) = r(u) + |px^n|$ so that we may choose $\hat{r} = r$ to show the claim. Indeed, suppose otherwise and take the minimal prefix s of r such that $s(\tilde{u}, i) \neq s(u) + |px^n|$. Then $s(u)$ is necessarily undefined, contradicting the assumption $u \leqslant^R_{m,n} v$.

(2) & (3) $\Rightarrow$ (1): First suppose $r(\tilde{v}, i) \in B(\tilde{v})$. Let $\hat{r}$ be the ranker from (3) for r, so that $r(\tilde{u}, i) = \hat{r}(u) + |px^n|$. This implies $r(\tilde{u}, i) \in B(\tilde{u})$.

Suppose now $r(\tilde{v}, i) \in C$. Due to symmetry, it suffices to consider that r ends on an X-modality, in which case we have to show $r(\tilde{u}, i) \leqslant r(\tilde{v}, i)$.

Suppose further that there exists a prefix s of r such that $s(\tilde{v}, i) \in B(\tilde{v})$. Let s be a longest such prefix, and suppose $\mathrm{Z}_a \in \{\mathrm{X}_a, \mathrm{Y}_a\}$ is such that $s\mathrm{Z}_a$ is a prefix of r. If $s\mathrm{Z}_a \in \widetilde{R}_{m-1,\ell}$, then $s\mathrm{Z}_a(\tilde{u}, i) = s\mathrm{Z}_a(\tilde{v}, i)$ by (2), and induction yields $r(\tilde{u}, i) \leqslant r(\tilde{v}, i)$. Hence $s\mathrm{Z}_a$ uses the full alternation parameter m; in particular $\mathrm{Z}_a = \mathrm{X}_a$. Let $\hat{s}$ be the ranker from (3) for s. We have that $\hat{s}(v)$ as well as $\hat{s}(u)$ are defined, and $\hat{s}\mathrm{X}_a(v)$ is undefined by choice of $s\mathrm{X}_a$.

We distinguish whether $\hat{s}\mathrm{X}_a(u)$ is defined or not. Let us first consider $\hat{s}\mathrm{X}_a(u)$ being undefined. This means that $s\mathrm{X}_a(\tilde{v}, i)$ and $s\mathrm{X}_a(\tilde{u}, i)$ both leave the center simultaneously with the last modality. Since the words $\tilde{u}$ and $\tilde{v}$ are the same outside the center, we have $s\mathrm{X}_a(\tilde{u}, i) = s\mathrm{X}_a(\tilde{v}, i)$. The second case is $\hat{s}\mathrm{X}_a(u)$ being defined. This means $s\mathrm{X}_a(\tilde{u}, i)$ stays in the center, whereas $s\mathrm{X}_a(\tilde{v}, i)$ leaves it with an X-modality. Therefore, interpreting $s\mathrm{X}_a(\tilde{v}, i)$ as a position of $\tilde{u}$, we get $s\mathrm{X}_a(\tilde{u}, i) \leqslant s\mathrm{X}_a(\tilde{v}, i)$. Consider the suffix r' of r such that $r = s\mathrm{X}_a r'$. The suffix r' is unidirectional by the above observations, and $r'(\tilde{u}, s\mathrm{X}_a(\tilde{v}, i)) = r'(\tilde{v}, s\mathrm{X}_a(\tilde{v}, i)) = r(\tilde{v}, i)$. Therefore $r(\tilde{u}, i) \leqslant r(\tilde{v}, i)$ in both cases.

By these considerations, we may assume that $s(\tilde{v}, i)$ is outside the center for all prefixes s of r. If further $s(\tilde{u}, i)$ is outside the center for every prefix s of r, then $s(\tilde{u}, i) = s(\tilde{v}, i)$. In particular $r(\tilde{u}, i) = r(\tilde{v}, i)$. It remains to consider $r = sr'$ for some s with $s(\tilde{u}, i) \in B(\tilde{u})$. We know $s \notin \widetilde{R}_{m-1,k+|s|}$ by (2). This means $s \in \widetilde{R}^{\mathrm{X}}_{m,k+|s|}$ and $r' \in \widetilde{R}^{\mathrm{X}}_{1,\ell}$. We thus have $s(\tilde{u}, i) \leqslant s(\tilde{v}, i)$, where $s(\tilde{v}, i)$ is interpreted as a position of $\tilde{u}$. As the suffix r' of r is unidirectional, $r(\tilde{u}, i) = r'(\tilde{u}, s(\tilde{u}, i)) \leqslant r'(\tilde{u}, s(\tilde{v}, i)) = r(\tilde{v}, i)$.

That concludes the proof of the claim.

Using these claims we can prove the lemma. We prove $\tilde{u} \leqslant^R_{m+1,n} \tilde{v}$ by showing conditions (1) to (3) of Definition 11.2 one after another.

Let in the following $r \in R_{m+1,n}(\tilde{v})$, and let r' be the first direction block of r. More precisely, let $r = r'r''$ such that $r' \in R_{1,n}$ and $r'' \in \{\varepsilon\} \cup R_{m,n}$. By Claim 3 setting $i = r'(\tilde{u}) = r'(\tilde{v})$ is a well-defined position with $i \in C$.

Definition 11.2 (1): We have to show that r is defined on $\tilde{u}$. There is nothing to show if $r'' = \varepsilon$. Suppose otherwise $r'' \in R_{m,n}$. By assumption $r''(\tilde{v}, i)$ is defined and hence so is $r(\tilde{u}) = r''(\tilde{u}, i)$ by statement (1) in Claim 4.

Definition 11.2 (2): Assume that r ends on an X-modality, and consider a ranker $s \in R^{\mathrm{X}}_{m,n-1}(\tilde{v}) \cup R^{\mathrm{Y}}_{m+1,n-1}(\tilde{v})$ and $d \in \{0, 1\}$. Suppose $r(\tilde{v}) \leqslant s(\tilde{v}) - d$. We have to show $r(\tilde{u}) \leqslant s(\tilde{u}) - d$.

For s we also split off the first direction and let $s = s's''$ be such that $s' \in R_{1,n}$ and $s'' \in \{\varepsilon\} \cup R^{\mathrm{X}}_{m-1,n-1} \cup R^{\mathrm{Y}}_{m,n-1}$. By Claim 1, setting $j = s'(\tilde{u}) = s'(\tilde{v})$ is well-defined and $j \in C$. We shall show $r''(\tilde{u}, i) \leqslant s''(\tilde{u}, j) - d$, which yields $r(\tilde{u}) \leqslant s(\tilde{u}) - d$ as desired. We distinguish cases depending on whether or not $r''(\tilde{v}, i) \in B(\tilde{v})$ and whether or not $s''(\tilde{v}, j) \in B(\tilde{v})$.

First suppose $r''(\tilde{v}, i) \in B(\tilde{v})$ and $s''(\tilde{v}, j) \in B(\tilde{v})$. Let $\hat{r}$ and $\hat{s}$ be the rankers corresponding to r'' and s'', respectively, in statement (3) of Claim 4. For $w \in \{u, v\}$ we have $r''(\tilde{w}, i) = \hat{r}(w) + |px^n|$ and $s''(\tilde{w}, j) = \hat{s}(w) + |px^n|$. Since $r(\tilde{v}) \leqslant s(\tilde{v}) - d$, we have $\hat{r}(v) \leqslant \hat{s}(v) - d$. This in turn yields $\hat{r}(u) \leqslant \hat{s}(u) - d$ by the assumption $u \leqslant^R_{m,n} v$; note that $\hat{r} \in R^{\mathrm{X}}_{m,n}$ and $\hat{s} \in R^{\mathrm{Y}}_{m,n-1} \cup R^{\mathrm{X}}_{m-1,n-1}$. This shows $r''(\tilde{u}, i) \leqslant s''(\tilde{u}, j) - d$.

Next suppose that not both $r''(\tilde{v}, i) \in B(\tilde{v})$ and $s''(\tilde{v}, j) \in B(\tilde{v})$. Claim 4 yields

- if $r''(\tilde{v}, i) \in C$, then $r''(\tilde{u}, i) \leqslant r''(\tilde{v}, i)$ and
- if $s''(\tilde{v}, j) \in C$, then $s''(\tilde{v}, j) \leqslant s''(\tilde{u}, j)$.

If $r''(\tilde{v}, i) \in C$ and $s''(\tilde{v}, j) \in C$, then $r''(\tilde{u}, i) \leqslant r''(\tilde{v}, i) \leqslant s''(\tilde{v}, j) - d \leqslant s''(\tilde{u}, j) - d$. Suppose $r''(\tilde{v}, i) \in B(\tilde{v})$ and $s''(\tilde{v}, j) \in C$. By statement (1) in Claim 4 we see $r''(\tilde{u}, i) \in B(\tilde{u})$. Thus $r''(\tilde{u}, i) < s''(\tilde{v}, j) \leqslant s''(\tilde{u}, j)$. The last remaining case, when $r''(\tilde{v}, i) \in C$ and $s''(\tilde{v}, j) \in B(\tilde{v})$, follows by a similar argument.

Definition 11.2 (3) is left-right symmetric. That concludes the proof of the lemma. □

We can now prove Theorem 12.11 and Theorem 12.12. The implication from the algebraic description to formulae relies on the respective Theorems 12.1 and 12.2 including successor. Note that these theorems are proven by now, so this is not a circular reasoning.

Proof of Theorem 12.11. (1) ⇒ (2): Let $L \subseteq A^*$ be defined by $\varphi \in \Sigma^2_{m,n}[<]$. We have to show that $M_L \in \mathbf{DA}$ and that M_L satisfies $U'_{m-1/2} \leqslant V'_{m-1/2}$. We verify these identities in M_L, starting with $\mathbf{DA}$. Let $p, q, x, y \in A^*$, and let $u = p(xy)^n x(xy)^n q$ and $v = p(xy)^{2n} q$. The words u and v have the same scattered subwords of length $2n$. This implies $u \equiv^R_{1,n} v$. Corollary 11.7 yields $u \in L(\varphi)$ if and only if $v \in L(\varphi)$.

Let us next verify the identity $U'_{m-1/2} \leqslant V'_{m-1/2}$ in M_L. For $i - 1/2 \in \mathbb{N}$ let $z, x_i, y_i \in A^*$, and recursively define

$$u_{1/2} = z, \quad u_i = (u_{i-1}x_i)^{2n} u_{i-1} (y_i u_{i-1})^{2n},$$
$$v_{1/2} = 1, \quad v_i = (u_{i-1}x_i)^{2n} v_{i-1} (y_i u_{i-1})^{2n}.$$

The words u_i and v_i are instantiations of the omega-terms U'_i and V'_i with words substituted for the variables, and with the concrete power $2n$ substituted for the formal symbol ω. We shall show $pu_iq \leqslant^R_{i+1/2,n} pv_iq$ for all $p, q \in A^*$ by induction on i. For the base case observe that every scattered subword of length $2n$ in $pv_{1/2}q$ also appears in $pu_{1/2}q$. This implies $pu_{1/2}q \leqslant^R_{1,n} pv_{1/2}q$. Suppose now $i \geqslant 3/2$. Induction yields $u_{i-1} \leqslant^R_{i-1/2,n} v_{i-1}$, and Lemma 12.53 yields $pu_iq \leqslant^R_{i+1/2,n} pv_iq$ as desired. Note that $\mathrm{alph}(v_{i-1}) \subseteq \mathrm{alph}(u_{i-1}) \subseteq \mathrm{alph}(u_{i-1}x_i) \cap \mathrm{alph}(y_i u_{i-1})$. This in particular shows $pu_{m-1/2}q \leqslant^R_{m,n} pv_{m-1/2}q$. By Corollary 11.4 we see that $pv_{m-1/2}q \in L(\varphi)$ implies $pu_{m-1/2}q \in L(\varphi)$. This shows that the identity $U'_{m-1/2} \leqslant V'_{m-1/2}$ is satisfied by M_L.

(2) ⇒ (1): Let $\mathbf{V}$ be the class of semigroups in $\mathbf{LDA}$ satisfying $U_{m-1/2} \leqslant V_{m-1/2}$. Theorem 12.1 shows that every language recognized by semigroups in $\mathbf{V}$ is definable in $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$. Lemma 12.52 shows that every formula in $\Sigma^2_m[<, \mathrm{suc}, \min, \max]$ has a sliding window formula in $\Sigma^2_m[<]$. Let now $M_L \in \mathbf{DA}$ satisfy $U'_{m-1/2} \leqslant V'_{m-1/2}$. Treating M_L as a semigroup we have $M_L \in \mathbf{V_M}$. Therefore, Proposition 8.26 is applicable and L is $\Sigma^2_m[<]$-definable. □

Proof of Theorem 12.12. This proof is very much along the same lines as the previous one. We give it for completeness' sake.

(1) ⇒ (2): Let $L \subseteq A^*$ be defined by the sentence $\varphi \in \mathrm{FO}^2_{m,n}[<]$. We show that M_L satisfies $U'_m = V'_m$. For $i \geqslant 1$ let $s, t, x_i, y_i \in A^*$, and

$$u_1 = (sx_1)^{2n} s (y_1 t)^{2n}, \quad u_i = (u_{i-1} x_i)^{2n} u_{i-1} (y_i u_{i-1})^{2n},$$
$$v_1 = (sx_1)^{2n} t (y_1 t)^{2n}, \quad v_i = (u_{i-1} x_i)^{2n} v_{i-1} (y_i u_{i-1})^{2n}.$$

We claim $pu_iq \equiv^R_{i,n} pv_iq$ for all $p, q \in A^*$. For the base case observe that pu_1q and pv_1q have the same scattered subword of length $2n$, implying $pu_1q \equiv^R_{1,n} pv_1q$. Suppose now $i \geqslant 2$. By induction we get $u_{i-1} \equiv^R_{i-1,n} v_{i-1}$. Lemma 12.53 yields $pu_iq \leqslant^R_{i,n} pv_iq$ as well as $pv_iq \leqslant^R_{i,n} pu_iq$. Note that $\mathrm{alph}(v_{i-1}) = \mathrm{alph}(u_{i-1}) \subseteq \mathrm{alph}(u_{i-1}x_i) \cap \mathrm{alph}(y_iu_{i-1})$.

In particular $pu_mq \equiv^R_{m,n} pv_mq$. By Corollary 11.7 we conclude $pv_mq \in L(\varphi)$ if and only if $pu_mq \in L(\varphi)$. This shows that M_L satisfies $U'_m = V'_m$.

(2) ⇒ (1): Consider the class of semigroups satisfying $U_m = V_m$. By Theorem 12.2 every language recognized in $\mathbf{V}$ is $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$-definable. Lemma 12.52 implies that every formula in $\mathrm{FO}^2_m[<, \mathrm{suc}, \min, \max]$ has a sliding window formula in $\mathrm{FO}^2_m[<]$. If M_L satisfies $U'_m = V'_m$, then $M_L \in \mathbf{V_M}$, and applying Proposition 8.26 shows that L is $\mathrm{FO}^2_m[<]$-definable. □

Conclusion and Future Work

The first part of this thesis developed an abstract theory of logic fragments, based on natural axiomatic closure properties on the syntax level of formulae. Such syntactic closure properties were then used to prove natural semantic closure properties of the language family defined by logic fragments. The axiomatic theory of logic fragments was further exemplified by investigations of the influence of predicates for successor, minimum, and maximum on the expressive power.

An obvious continuation of this work is to extend the logic framework further to include function symbols in atomic formulae, other quantifiers used in the literature such as Lindström or monoidal quantifiers, and more numerical predicates such as addition $x+y=z$ and multiplication $x \cdot y = z$. The latter extension regarding numerical predicates might be approached even more generally by considering relational symbols and axiomatizing their numerical interpretation. In such an extension, suc-stability, for example, might turn out to be just an instance of a valid numerical interpretation of the relational symbols.

Another direction for future research can be to use the theory of logic fragments presented in this thesis to prove other semantic properties such as closure under shuffle product. Moreover, the approach of Section 8.2 using neutral letters to obtain Crane-Beach-type results is most likely not completely exploited, and the theory of logic fragments might help in proving the Crane Beach conjecture for many more natural logic fragments. Also expanding this thesis in the semantic direction, one could consider more general structures than finite words and interpret formulae over infinite words, traces, trees, data words, or other data structures.

This thesis focused on predicate logic. For problems of algorithmic software verification such as model checking, temporal logic is more common than predicate logic. Extending the axiomatic approach of logic fragments to other types of logic such as temporal logic would be desirable.

The second part of this thesis investigated the quantifier alternation hierarchy within two-variable first-order logic. First, a combinatorial description in terms of rankers for each level of the hierarchies over the signatures $\{<\}$ and $\{<, \mathrm{suc}\}$ was provided. After this, decidability of any level of the hierarchy over several signatures between $\{<\}$ and $\{<, \mathrm{suc}, \mathrm{min}, \mathrm{max}\}$ was proved. The precise computational complexity of deciding membership in a level of the alternation hierarchy remains open. We suspect that the identities established in this thesis can be transformed into forbidden patterns for deterministic automata. This might lead to a nondeterministic logarithmic space-bounded algorithm for deciding membership in a fixed level on input of a deterministic automaton. See [GS08; Gla01] for more information about forbidden patterns.

Only recently, decidability of the full levels of the FO^2-alternation hierarchy over the signature $\{<, \mathrm{mod}\}$ was proved [DP14]. It would be interesting to extend this result to include half levels, more signatures such as $\{<, \mathrm{suc}, \mathrm{mod}\}$, or even modular counting quantifiers.

Acknowledgments

Only through hard work and perseverance can one truly suffer.

— Unknown origin

I owe my deepest gratitude to my supervisors Manfred Kufleitner and Volker Diekert. I am particularly grateful to Manfred Kufleitner for the excellent mentoring and the fruitful collaboration, and also for being incredulous that $\mathrm{FO}^2_{m,n}$ with bounded quantifier depth is a variety, the proof of which ultimately developed into the theory of logic fragments presented in the first part of this thesis. I want to thank Volker Diekert for offering me a position as a research assistant and for his unfathomable patience with my thesis project. Thanks also to Chrisitan Glaßer for agreeing to examine my thesis. Further, I acknowledge the support of the German Research Foundation (DFG), grant DI 435/5-1, under which most of the work presented in this thesis was done.

A thanks is also due to all current and former colleagues of the theoretical computer science department for the pleasant working atmosphere and the many cakes they supplied: Volker Diekert, Ulrich Hertrampf, Benjamin Hoffman, Jonathan Kausch, Steffen Kopecki, Manfred Kufleitner, Jürn-Jochen Laun, Heike Photien, Horst Prote, Tobias Walter, and Armin Weiß.

I am deeply indebted to the many persons who agreed to read through the sometimes enigmatic strings of cryptic symbols of this thesis, namely Heinz Isengard, Andreas Lauser, Erne Lauser, Sonja Mannhardt, Lisa Speidel, and Daniel Zimmermann. In this context, I wish to thank in particular my brother Andreas who has proved a most meticulous, yet patient and always benevolent proofreader. For all remaining misprints I assume full responsibility.

Finally, I am most grateful to my mother Erne Lauser and my father Helmut Lauser for their unconditional love and the beautiful childhood they provided. And last but not at all least a big thank you to Sonja for all the wonderful time together and her unwavering patience during the arduous process of writing this thesis.

Thanks a lot!

Bibliography

[ABR92] Douglas Albert, Robert Baldinger, and John Rhodes. "Undecidability of the Identity Problem for Finite Semigroups." In: *Journal of Symbolic Logic* 57.1 (1992), pp. 179–192.

[Alm96] Jorge Almeida. "A syntactical proof of locality of **DA**." In: *International Journal of Algebra and Computation* 6.2 (1996), pp. 165–177.

[Alm94] Jorge Almeida. *Finite Semigroups and Universal Algebra.* World Scientific, 1994.

[Arf91] Mustapha Arfi. "Opérations polynomiales et hiérarchies de concaténation." In: *Theoretical Computer Science* 91.1 (1991), pp. 71–84.

[BCST92] David Arno Mix Barrington, Kevin Jay Compton, Howard Straubing, and Denis Thérien. "Regular Languages in NC^1." In: *Journal of Computer and System Sciences* 44.3 (1992), pp. 478–499.

[BILST05] David Arno Mix Barrington, Neil Immerman, Clemens Lautemann, Nicole Schweikardt, and Denis Thérien. "First-order expressibility of languages with neutral letters or: The Crane Beach conjecture." In: *Journal of Computer and System Sciences* 70.2 (2005), pp. 101–127.

[BP91] Danièle Beauquier and Jean-Éric Pin. "Languages and Scanners." In: *Theoretical Computer Science* 84.1 (1991), pp. 3–21.

[Bir35] Garrett Birkhoff. "On the structure of abstract algebras." In: *Proceedings of the Cambridge Philosophical Society* 31 (1935), pp. 433–454.

[BK78] Janusz Antoni Brzozowski and Robert Knast. "The Dot-Depth Hierarchy of Star-Free Languages is Infinite." In: *Journal of Computer and System Sciences* 16.1 (1978), pp. 37–55.

[Büc60] Julius Richard Büchi. "Weak Second-Order Arithmetic and Finite Automata." In: *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 6 (1960), pp. 66–92.

[CPS06] Laura Chaubard, Jean-Éric Pin, and Howard Straubing. "First Order Formulas with Modular Predicates." In: *Proceedings of the 21st Annual IEEE Symposium on Logic in Computer Science (LICS 2006).* IEEE Computer Society, 2006, pp. 211–220.

[Chu36] Alonzo Church. "An unsolvable problem of elementary number theory." In: *American Journal of Mathematics* 58 (1936), pp. 345–363.

[Cod70] Edgar Frank Codd. "A Relational Model of Data for Large Shared Data Banks." In: *Communications of the ACM* 13.6 (1970), pp. 377–387.

[Cod90] Edgar Frank Codd. *The Relational Model for Database Management, Version 2.* Addison-Wesley, 1990.

[CB71] Rina Sara Cohen and Janusz Antoni Brzozowski. "Dot-Depth of Star-Free Events." In: *Journal of Computer and System Sciences* 5.1 (1971), pp. 1–16.

[CB12] Carlo Comin and Maria Paola Bianchi. "Algebraic Characterization of the Class of Languages recognized by Measure Only Quantum Automata." In: *Computing Research Repository* arXiv:1206.1702 [cs.FL] (2012), pp. 1–7.

[DKL10] Luc Dartois, Manfred Kufleitner, and Alexander Lauser. "Rankers over Infinite Words (Extended Abstract)." In: *Proceedings of the 14th International Conference on Developments in Language Theory (DLT 2010)*. Ed. by Yuan Gao, Hanlin Lu, Shinnosuke Seki, and Sheng Yu. Vol. 6224. Lecture Notes in Computer Science. Springer, 2010, pp. 148–159.

[DP14] Luc Dartois and Charles Paperman. "Adding modular predicates." In: *Computing Research Repository* arXiv:1401.6576 [cs.LO] (2014), pp. 1–17.

[DP13] Luc Dartois and Charles Paperman. "Two-variable first order logic with modular predicates over words." In: *Proceedings of the 30th International Symposium on Theoretical Aspects of Computer Science (STACS 2013)*. Ed. by Natacha Portier and Thomas Wilke. Vol. 20. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2013, pp. 329–340.

[DG08] Volker Diekert and Paul Gastin. "First-order definable languages." In: *Logic and Automata: History and Perspectives*. Ed. by Jörg Flum, Erich Grädel, and Thomas Wilke. Texts in Logic and Games. Amsterdam University Press, 2008, pp. 261–306.

[DGK08] Volker Diekert, Paul Gastin, and Manfred Kufleitner. "A Survey on Small Fragments of First-Order Logic over Finite Words." In: *International Journal of Foundations of Computer Science* 19.03 (2008), pp. 513–548.

[DK11] Volker Diekert and Manfred Kufleitner. "Fragments of First-Order Logic over Infinite Words." In: *Theory of Computing Systems* 48.3 (2011), pp. 486–516.

[EF95] Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite Model Theory*. Perspectives in Mathematical Logic. Springer, 1995.

[Eil76] Samuel Eilenberg. *Automata, Languages, and Machines*. Vol. B. Academic Press, 1976.

[Elg61] Calvin Creston Elgot. "Decision problems of finite automata design and related arithmetics." In: *Transactions of the American Mathematical Society* 98 (1961), pp. 21–51.

[EVW02] Kousha Etessami, Moshe Ya'akov Vardi, and Thomas Wilke. "First-Order Logic with Two Variables and Unary Temporal Logic." In: *Information and Computation* 179.2 (2002), pp. 279–295.

[Fag74] Ronald Fagin. "Generalized first-order spectra and polynomial-time recognizable sets." In: *Proceedings of the American Mathematical Society*. Vol. 7. Complexity of Computation. American Mathematical Society, 1974, pp. 43–73.

[FKL14] Lukas Fleischer, Manfred Kufleitner, and Alexander Lauser. "Block Products and Negation Nesting in FO^2." In: *Proceedings of the 9th International Computer Science Symposium in Russia (CSR 2014)*. Ed. by Edward Hirsch, Sergei Kuznetsov, Jean-Éric Pin, and Nikolay Vereshchagin. Vol. 8476. Lecture Notes in Computer Science. Springer, 2014, pp. 176–189.

[GN08] Wouter Gelade and Frank Neven. "Succinctness of the Complement and Intersection of Regular Expressions." In: *Proceedings of the 25th International Symposium on Theoretical Aspects of Computer Science (STACS 2008)*. Ed. by Susanne Albers and Pascal Weil. Vol. 1. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2008, pp. 325–336.

[Gla01] Christian Glaßer. "Forbidden-Patterns and Word Extensions for Concatenation Hierarchies." Dissertation. Universität Würzburg, 2001.

[GS08] Christian Glaßer and Heinz Schmitz. "Languages of Dot-Depth 3/2." In: *Theory of Computing Systems* 42.2 (2008), pp. 256–286.

[GH08] Hermann Gruber and Markus Holzer. "Finite automata, digraph connectivity, and regular expression size (Extended abstract)." In: *Proceedings of the 35th International Colloquium on Automata, Languages and Programming (ICALP 2008), Part II*. Ed. by Luca Aceto, Ivan Damgård, LeslieAnn Goldberg, Magnús Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz. Vol. 5126. Lecture Notes in Computer Science. Springer, 2008, pp. 39–50.

[HK14] Martin Huschenbett and Manfred Kufleitner. "Ehrenfeucht-Fraïssé Games on Omega-Terms." In: *Proceedings of the 31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014)*. Ed. by Ernst Wilhelm Mayr and Natacha Portier. Vol. 25. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2014, pp. 374–385.

[Imm99] Neil Immerman. *Descriptive complexity*. Graduate texts in computer science. Springer, 1999, pp. I–XVI, 1–268.

[JKL12] Franz Jahn, Manfred Kufleitner, and Alexander Lauser. "Regular Ideal Languages and Their Boolean Combinations." In: *Proceedings of the 17th International Conference on Implementation and Application of Automata (CIAA 2012)*. Ed. by Nelma Moreira and Rogério Reis. Vol. 7381. Lecture Notes in Computer Science. Springer, 2012, pp. 205–216.

[KKL11] Jakub Kallas, Manfred Kufleitner, and Alexander Lauser. "First-order Fragments with Successor over Infinite Words." In: *Proceedings of the 28th International Symposium on Theoretical Aspects of Computer Science (STACS 2011)*. Ed. by Thomas Schwentick and Christoph Dürr. Vol. 9. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2011, pp. 356–367.

[Kam68] Johan Anthony Willem Kamp. "Tense Logic and the Theory of Linear Order." PhD thesis. University of California, 1968.

[Kle56] Steven Cole Kleene. “Representation of events in nerve nets and finite automata.” In: *Automata Studies*. Ed. by Claude Elwood Shannon and John McCarthy. Annals of Mathematics Studies 34. Princeton University Press, 1956, pp. 3–40.

[Klí11] Ondřej Klíma. “Piecewise testable languages via combinatorics on words.” In: *Discrete Mathematics* 311.20 (2011), pp. 2124–2127.

[KP08] Ondřej Klíma and Libor Polák. “On varieties of literally idempotent languages.” In: *RAIRO. Theoretical Informatics and Applications* 42.3 (2008), pp. 583–598.

[Kna83] Robert Knast. “A Semigroup Characterization of Dot-Depth one Languages.” In: *R.A.I.R.O. — Informatique Théorique et Applications* 17.4 (1983), pp. 321–330.

[KS12a] Andreas Krebs and A. V. Sreejith. “Non-definability of Languages by Generalized First-order Formulas over $(\mathbb{N}, +)$.” In: *Proceedings of the 27th Annual IEEE Symposium on Logic in Computer Science (LICS 2012)*. IEEE Computer Society, 2012, pp. 451–460.

[KS12b] Andreas Krebs and Howard Straubing. “An effective characterization of the alternation hierarchy in two-variable logic.” In: *Proceedings of the 34th Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2012)*. Ed. by Deepak D’Souza, Telikepalli Kavitha, and Jaikumar Radhakrishnan. Vol. 18. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2012, pp. 86–98.

[Kuf13] Manfred Kufleitner. “The Trotter-Weil Hierarchy.” Habilitationsschrift. Universität Stuttgart, 2013.

[KL12a] Manfred Kufleitner and Alexander Lauser. “Around Dot-Depth One.” In: *International Journal of Foundations of Computer Science* 23.6 (2012), pp. 1323–1339.

[KL11a] Manfred Kufleitner and Alexander Lauser. “Around Dot-Depth One (Extended Abstract).” In: *Proceedings of the 13th International Conference on Automata and Formal Languages (AFL 2011)*. Ed. by Pál Dömösi and Szabolcs Iván. College of Nyíregyháza, 2011, pp. 255–269.

[KL11b] Manfred Kufleitner and Alexander Lauser. “Languages of Dot-Depth One over Infinite Words.” In: *Proceedings of the 26th Annual IEEE Symposium on Logic in Computer Science (LICS 2011)*. IEEE Computer Society, 2011, pp. 23–32.

[KL12b] Manfred Kufleitner and Alexander Lauser. “Lattices of Logical Fragments over Words.” In: *Proceedings of the 39th International Colloquium on Automata, Languages and Programming (ICALP 2012), Part II*. Ed. by Artur Czumaj, Kurt Mehlhorn, Andrew Pitts, and Roger Wattenhofer. Vol. 7392. Lecture Notes in Computer Science. Springer, 2012, pp. 275–286.

[KL11c] Manfred Kufleitner and Alexander Lauser. "Partially Ordered Two-way Büchi Automata." In: *International Journal of Foundations of Computer Science* 22.8 (2011), pp. 1861–1876.

[KL13] Manfred Kufleitner and Alexander Lauser. "Quantifier Alternation in Two-Variable First-Order Logic with Successor Is Decidable." In: *Proceedings of the 30th International Symposium on Theoretical Aspects of Computer Science (STACS 2013)*. Ed. by Natacha Portier and Thomas Wilke. Vol. 20. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2013, pp. 305–316.

[KL12c] Manfred Kufleitner and Alexander Lauser. "The join levels of the Trotter-Weil hierarchy are decidable." In: *Proceedings of the 37th International Symposium Mathematical Foundations of Computer Science (MFCS 2012)*. Ed. by Branislav Rovan, Vladimiro Sassone, and Peter Widmayer. Vol. 7464. Lecture Notes in Computer Science. Springer, 2012, pp. 603–614.

[KL12d] Manfred Kufleitner and Alexander Lauser. "The Join of the Varieties of R-trivial and L-trivial Monoids via Combinatorics on Words." In: *Discrete Mathematics & Theoretical Computer Science* 14.1 (2012), pp. 141–146.

[KW13] Manfred Kufleitner and Tobias Walter. "One Quantifier Alternation in First-Order Logic with Modular Predicates." In: *Computing Research Repository* arXiv:1310.5043 [cs.FL] (2013), pp. 1–22.

[KW12] Manfred Kufleitner and Pascal Weil. "The FO^2 alternation hierarchy is decidable." In: *Proceedings of the 26th International Workshop & 21st Annual Conference of the EACSL on Computer Science Logic (CSL 2012)*. Ed. by Patrick Cégielski and Arnaud Durand. Vol. 16. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2012, pp. 426–439.

[Kun03] Michal Kunc. "Equational description of pseudovarieties of homomorphisms." In: *RAIRO. Theoretical Informatics and Applications* 37 (2003), pp. 243–254.

[LTT06] Clemens Lautemann, Pascal Tesson, and Denis Thérien. "An algebraic point of view on the Crane Beach property." In: *Proceedings of the 20th International Workshop & 15th Annual Conference of the EACSL on Computer Science Logic (CSL 2006)*. Ed. by Zoltán Ésik. Vol. 4207. Lecture Notes in Computer Science. Springer, 2006, pp. 426–440.

[LPS10] Kamal Lodaya, Paritosh K. Pandya, and Simoni S. Shah. "Around dot depth two." In: *Proceedings of the 14th International Conference on Developments in Language Theory (DLT 2010)*. Ed. by Yuan Gao, Hanlin Lu, Shinnosuke Seki, and Sheng Yu. Vol. 6224. Lecture Notes in Computer Science. Springer, 2010, pp. 303–315.

[MP71] Robert McNaughton and Seymour Papert. *Counter-Free Automata*. The MIT Press, 1971.

[Pin95] Jean-Éric Pin. "A variety theorem without complementation." In: *Russian Mathematics (Izvestija VUZov. Matematika)*. Vol. 39. 1995, pp. 80–90.

[Pin12] Jean-Éric Pin. "Equational descriptions of languages." In: *International Journal of Foundations of Computer Science* 23.6 (2012), pp. 1227–1240.

[Pin05] Jean-Éric Pin. "Expressive power of existential first-order sentences of Büchi's sequential calculus." In: *Discrete Mathematics* 291.1-3 (2005), pp. 155–174.

[Pin97] Jean-Éric Pin. "Syntactic semigroups." In: *Handbook of Formal Languages*. Vol. 1. Springer, 1997, pp. 679–746.

[Pin86] Jean-Éric Pin. *Varieties of Formal Languages*. North Oxford Academic, 1986, p. 138.

[PS05] Jean-Éric Pin and Howard Straubing. "Some results on $\mathcal{C}$-varieties." In: *RAIRO. Theoretical Informatics and Applications* 39 (2005), pp. 239–262.

[PW97] Jean-Éric Pin and Pascal Weil. "Polynomial Closure and Unambiguous Product." In: *Theory of Computing Systems* 30.4 (1997), pp. 383–422.

[PW02] Jean-Éric Pin and Pascal Weil. "The wreath product principle for ordered semigroups." In: *Communications in Algebra* 30.12 (2002), pp. 5677–5713.

[PZ14] Thomas Place and Marc Zeitoun. "Going Higher in the First-Order Quantifier Alternation Hierarchy on Words." In: *Proceedings of the 41st International Colloquium on Automata, Languages and Programming (ICALP 2014), Part II*. Ed. by Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias. Vol. 8573. Lecture Notes in Computer Science. Springer, 2014, pp. 342–353.

[Rei82] Jan Reiterman. "The Birkhoff theorem for finite algebras." In: *Algebra Universalis* 14 (1982), pp. 1–10.

[Sch65] Marcel Paul Schützenberger. "On finite monoids having only trivial subgroups." In: *Information and Control* 8 (1965), pp. 190–194.

[STV02] Thomas Schwentick, Denis Thérien, and Heribert Vollmer. "Partially-Ordered Two-Way Automata: A New Characterization of DA." In: *Proceedings of the 5th International Conference on Developments in Language Theory (DLT 2001)*. Ed. by Werner Kuich, Grzegorz Rozenberg, and Arto Salomaa. Vol. 2295. Lecture Notes in Computer Science. Springer, 2002, pp. 239–250.

[Sim75] Imre Simon. "Piecewise testable events." In: *Automata Theory and Formal Languages, 2nd GI Conference*. Ed. by Helmut Barkhage. Vol. 33. Lecture Notes in Computer Science. Springer, 1975, pp. 214–222.

[Sto74] Larry Joseph Stockmeyer. "The complexity of decision problems in automata theory and logic." PhD thesis. Massachusetts Institute of Technology, 1974.

[Str81] Howard Straubing. "A Generalization of the Schützenberger Product of Finite Monoids." In: *Theoretical Computer Science* 13 (1981), pp. 137–150.

[Str94] Howard Straubing. *Finite Automata, Formal Logic, and Circuit Complexity*. Birkhäuser, 1994.

[Str85] Howard Straubing. "Finite semigroup varieties of the form $\mathbf{V} * \mathbf{D}$." In: *Journal of Pure and Applied Algebra* 36.1 (1985), pp. 53–94.

[Str01] Howard Straubing. "Languages defined with modular counting quantifiers." In: *Information and Computation* 166.2 (2001), pp. 112–132.

[Str02] Howard Straubing. "On Logical Descriptions of Regular Languages." In: *Proceedings of the 5th Latin American Theoretical Informatics Conference (LATIN 2002)*. Ed. by Sergio Rajsbaum. Vol. 2286. Lecture Notes in Computer Science. Springer, 2002, pp. 528–538.

[STT95] Howard Straubing, Denis Thérien, and Wolfgang Thomas. "Regular languages defined with generalized quantifiers." In: *Information and Computation* 118.2 (1995), pp. 289–301.

[TT02] Pascal Tesson and Denis Thérien. "Diamonds are Forever: The Variety **DA**." In: *Semigroups, Algorithms, Automata and Languages*. Ed. by Gracinda Maria dos Gomes Moreira da Cunha, Pedro Ventura Alves da Silva, and Jean-Éric Pin. World Scientific, 2002, pp. 475–500.

[TT07] Pascal Tesson and Denis Thérien. "Logic Meets Algebra: The Case of Regular Languages." In: *Logical Methods in Computer Science* 3.1 (2007), pp. 1–37.

[Thé88] Denis Thérien. "Categories Et Langages De Dot-Depth Un." In: *R.A.I.R.O. — Informatique Théorique et Applications* 22.4 (1988), pp. 437–445.

[Thé81] Denis Thérien. "Classification of finite monoids: The language approach." In: *Theoretical Computer Science* 14.2 (1981), pp. 195–208.

[TW98] Denis Thérien and Thomas Wilke. "Over Words, Two Variables Are as Powerful as One Quantifier Alternation." In: *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC 1998)*. ACM Press, 1998, pp. 234–240.

[Tho82] Wolfgang Thomas. "Classifying Regular Events in Symbolic Logic." In: *Journal of Computer and System Sciences* 25 (1982), pp. 360–376.

[Til87] Bret Tilson. "Categories as algebras: An essential ingredient in the theory of monoids." In: *Journal of Pure and Applied Algebra* 48 (1987), pp. 83–198.

[Tra61] Boris Avraamovich Trakhtenbrot. "Finite automata and logic of monadic predicates (in Russian)." In: *Doklady Akademii Nauk SSSR* 140 (1961), pp. 326–329.

[Tur36] Alan Mathison Turing. "On computable numbers, with an application to the Entscheidungsproblem." In: *Proceedings of the London Mathematical Society, Second Series* 42 (1936), pp. 230–265.

[Var07] Moshe Ya'akov Vardi. "Automata-Theoretic Model Checking Revisited." In: *Proceedings of the 8th International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI 2007)*. Ed. by Byron Cook and Andreas Podelski. Vol. 4349. Lecture Notes in Computer Science. Springer, 2007, pp. 137–150.

[Wei02] Pascal Weil. "Profinite methods in semigroup theory." In: *International Journal of Algebra and Computation* 12 (2002), pp. 137–178.

[WI09] Philipp Weis and Neil Immerman. "Structure theorem and strict alternation hierarchy for FO^2 on words." In: *Logical Methods in Computer Science* 5.3 (2009), pp. 1–23.

Index

Appendix

A. Syntax and Semantics of Monadic Second-Order Logic over Words

This appendix gives the missing formal definitions of concepts introduced in Section 3.1; in particular, a rigorous definition of the truth value and the formal semantics are given. Remember that Λ is the set of labels and $\mathbb{V} = \mathbb{V}_1 \dot\cup \mathbb{V}_2$ is the set of (first-order and second-order) variables. Let us give a more concise definition of the syntax of formulae. Formulae are given by the following inductive grammar-like rules (where $x, y \in \mathbb{V}_1$, $X \in \mathbb{V}_2$, $B \subseteq \Lambda$, $r, q \in \mathbb{Z}$, and φ, ψ are already formulae):

$$\begin{aligned}
&\top \mid \bot \mid \mathrm{empty} \mid \lambda(x) \in B \mid x \in X \mid \\
&x = y \mid x < y \mid x \leqslant y \mid \mathrm{suc}(x, y) \mid \min(x) \mid \max(x) \mid x \equiv r \pmod{q} \mid \\
&\neg\varphi \mid \varphi \vee \psi \mid \varphi \wedge \psi \mid \exists x\, \varphi \mid \forall x\, \varphi \mid \exists X\, \varphi \mid \forall X\, \varphi \mid \exists^{r \bmod q} x\, \varphi
\end{aligned}$$

We read $\lambda(x) = a$ for $a \in \Lambda$ as an abbreviation for $\lambda(x) \in \{a\}$. Parentheses may be used to disambiguate and to increase readability. Many constructions employed throughout this thesis are by induction on the structure of formulae and are thus closely linked to this definition. An example of such an induction is the definition of the set $\mathrm{FV}(\varphi)$ of *free variables*, which is defined as usual by

$$\begin{aligned}
\mathrm{FV}(\mathrm{empty}) &= \mathrm{FV}(\top) = \mathrm{FV}(\bot) = \emptyset, \\
\mathrm{FV}(x \in X) &= \{x, X\}, \\
\mathrm{FV}(\lambda(x) \in B) &= \mathrm{FV}(\min(x)) = \mathrm{FV}(\max(x)) = \mathrm{FV}(x \equiv r \pmod{q}) = \{x\}, \\
\mathrm{FV}(x = y) &= \mathrm{FV}(x < y) = \mathrm{FV}(x \leqslant y) = \mathrm{FV}(\mathrm{suc}(x, y)) = \{x, y\}
\end{aligned}$$

for atomic formulae, and inductively

$$\begin{aligned}
\mathrm{FV}(\neg\varphi) &= \mathrm{FV}(\varphi), \\
\mathrm{FV}(\varphi \vee \psi) &= \mathrm{FV}(\varphi \wedge \psi) = \mathrm{FV}(\varphi) \cup \mathrm{FV}(\psi), \\
\mathrm{FV}(\exists x\, \varphi) &= \mathrm{FV}(\forall x\, \varphi) = \mathrm{FV}(\exists^{r \bmod q} x\, \varphi) = \mathrm{FV}(\varphi) \setminus \{x\}, \\
\mathrm{FV}(\exists X\, \varphi) &= \mathrm{FV}(\forall X\, \varphi) = \mathrm{FV}(\varphi) \setminus \{X\}.
\end{aligned}$$

We now define the *formal semantics* $[\![\varphi]\!]_V$ of a formula φ with respect to any set of variables V containing all free variables of φ. We do this by set theoretic means. By definition, the semantics will be a subset of words over the infinite alphabet $\Lambda \times 2^V$, where 2^V denotes the power set of V. The intuitive idea of the additional second component is already given in Section 3.2, so we concentrate on the formal definition.

In the following let $u = (a_1, J_1) \cdots (a_n, J_n)$ with $a_i \in \Lambda$ and $J_i \subseteq V$. All first-order variables occurring freely in φ must have a well-defined interpretation. Hence, every such variable x must occur exactly once within a second component, *i.e.*, there must be a unique position i such that $x \in J_i$. In this case we let $x(u) = i$ be the *interpretation* of x and say that $x(u)$ is defined; otherwise $x(u)$ is undefined. The interpretation $X(u)$ of a second-order variable X is the set of positions $i \in \{1, \ldots, n\}$ such that $X \in J_i$.

The *universe of structures* over which formulae are interpreted in this thesis is $\mathcal{U}_V$. It consists of all words with a well-defined interpretation of every first-order variable in V:

$$\mathcal{U}_V = \big\{ u \in \big(\Lambda \times 2^V\big)^* \;\big|\; x(u) \text{ is defined for all first-order variables } x \big\}.$$

Notice that $\varepsilon \in \mathcal{U}_\emptyset \setminus \mathcal{U}_V$ if V contains at least one first-order variable.

For a formula φ and a set of variables V with $\mathrm{FV}(\varphi) \subseteq V$ we next define the formal semantics $[\![\varphi]\!]_V \subseteq \mathcal{U}_V$. We let $[\![\top]\!]_V = \mathcal{U}_V$ and $[\![\bot]\!]_V = \emptyset$, and for the other atomic formulae we put:

$$\begin{aligned}
[\![\mathrm{empty}]\!]_V &= \{u \in \mathcal{U}_V \mid u = \varepsilon\},\\
[\![\lambda(x) \in B]\!]_V &= \{u \in \mathcal{U}_V \mid u[x(u)] \in B \times 2^V\},\\
[\![x \in X]\!]_V &= \{u \in \mathcal{U}_V \mid x(u) \in X(u)\}.\\
[\![x = y]\!]_V &= \{u \in \mathcal{U}_V \mid x(u) = y(u)\},\\
[\![x < y]\!]_V &= \{u \in \mathcal{U}_V \mid x(u) < y(u)\},\\
[\![x \leqslant y]\!]_V &= \{u \in \mathcal{U}_V \mid x(u) \leqslant y(u)\},\\
[\![\mathrm{suc}(x,y)]\!]_V &= \{u \in \mathcal{U}_V \mid x(u) + 1 = y(u)\},\\
[\![\min(x)]\!]_V &= \{u \in \mathcal{U}_V \mid x(u) = 1\},\\
[\![\max(x)]\!]_V &= \{u \in \mathcal{U}_V \mid x(u) = |u|\},\\
[\![x \equiv r \pmod{q}]\!]_V &= \{u \in \mathcal{U}_V \mid x(u) - r \in q\mathbb{Z}\}.
\end{aligned}$$

Boolean combinations are given inductively by:

$$\begin{aligned}
[\![\neg\varphi]\!]_V &= \mathcal{U}_V \setminus [\![\varphi]\!]_V,\\
[\![\varphi \vee \psi]\!]_V &= [\![\varphi]\!]_V \cup [\![\psi]\!]_V,\\
[\![\varphi \wedge \psi]\!]_V &= [\![\varphi]\!]_V \cap [\![\psi]\!]_V.
\end{aligned}$$

For the semantics of the quantifiers we need some more notation. For $I \subseteq \mathbb{Z}$ and a variable $X \in \mathbb{V}$ let $u[X/I] = (a_1, J_1') \cdots (a_n, J_n')$ with

$$J_i' = \begin{cases} J_i \cup \{X\} & \text{if } i \in I,\\ J_i \setminus \{X\} & \text{if } i \notin I, \end{cases}$$

that is, I specifies which positions are to contain X. Extending this notation, for $x \in \mathbb{V}_1$ we denote by $u[x/i]$ the word $u[x/\{i\}]$. (We agree upon the convention that if i is not a position of u, then no position of $u[x/i]$ contains x in its second component; this will come in handy in some technical discussions.) With this we define

$$\begin{aligned}
[\![\exists x\, \varphi]\!]_V &= \big\{u \in \mathcal{U}_V \;\big|\; \text{there exists } i \in \{1,\ldots,|u|\} \text{ with } u[x/i] \in [\![\varphi]\!]_{\{x\}\cup V}\big\},\\
[\![\exists X\, \varphi]\!]_V &= \big\{u \in \mathcal{U}_V \;\big|\; \text{there exists } I \subseteq \{1,\ldots,|u|\} \text{ with } u[X/I] \in [\![\varphi]\!]_{\{X\}\cup V}\big\}.
\end{aligned}$$

For a formula φ and a first-order variable x let

$$I_u(x,\varphi) = \big\{i \in \{1,\ldots,|u|\} \;\big|\; u[x/i] \in [\![\varphi]\!]_{\{x\}\cup V}\big\}$$

be the set of positions i of u such that φ holds if x is interpreted by position i. With this we define the modular quantifier by

$$[\![\exists^{r \bmod q} x\, \varphi]\!]_V = \{u \in \mathcal{U}_V \mid |I_u(x,\varphi)| - r \in q\mathbb{Z}\}.$$

Universal quantifiers are given by the dual semantics $[\![\forall x\, \varphi]\!]_V = [\![\neg\exists x\, \neg\varphi]\!]_V$ and $[\![\forall X\, \varphi]\!]_V = [\![\neg\exists X\, \neg\varphi]\!]_V$; in particular, semantically, we may view universal quantifiers as an abbreviation for negated existential quantifiers of the negated formula.

Note that in case $q = 0$, the modular predicate degenerates to equality and the modular counting quantifier counts the exact number of positions, *i.e.*, for $u \in \mathcal{U}_V$ we have $u \in [\![x \equiv r \pmod{0}]\!]_V$ if and only if $x(u) = r$ and $u \in [\![\exists^{r \bmod 0} x\, \varphi]\!]_V$ if and only if $|I_u(x, \varphi)| = r$. In particular, we may view the classical quantifier $\exists! x\, \varphi$ for "there exists exactly one position" as an abbreviation for $\exists^{1 \bmod 0} x\, \varphi$.