Extensions of Language Families and Canonical Forms for Full AFL-structures*

Peter R.J. Asveld
Department of Applied Mathematics, Twente University of Technology
P.O. Box 217, 7500 AE Enschede, the Netherlands
e-mail: p.r.j.asveld@misc.utwente.nl P.R.J.Asveld@xs4all.nl

Abstract — We consider the following ways of extending a family of languages K to an "enriched " family $X(K)$: (i) hyper-algebraic extension ($X = H$) based on iterated parallel substitution, (ii) algebraic extension ($X = A$) obtained by nested iterated substitution, (iii) rational extension ($X = R$) achieved by not self-embedding nested iterated substitution, and (iv) a few subrational extensions ($X = M, C, S, P, F$) based on several kinds of substitution. We introduce full X-AFL's, i.e., nontrivial families closed under finite substitution, intersection with regular sets and under X, which turn out to be equivalent to well-known full AFL-structures such as full hyper-AFL ($X = H$), full super-AFL ($X = A$), full substitution-closed AFL ($X = R$), full semi-AFL ($X = S$), etc. Then we establish Canonical Forms for the smallest full X-AFL $\hat{\mathcal{X}}(K)$ containing K, i.e., we decompose the operator $\hat{\mathcal{X}}$ into simpler operators. Using Canonical Forms for full X-AFL's we obtain expressions for the smallest full X-AFL containing the result of substituting a family of languages into another family.

1 Introduction

In studying closure properties of language families major progress was made since 1969 when Ginsburg, Greibach & Hopcroft [10] introduced the concept of Abstract Family of Languages (AFL) denoting any family closed under the regular operations (union, concatenation and Kleene +), homomorphism, inverse homomorphism and intersection with regular sets. Using this concept we consider language families as algebras which enables us to deal with closure properties in a more abstract fashion.

As usual in algebra, structures with more restrictive and, on the other hand, more general properties have been considered. Ginsburg & Spanier [11] investigated substitution-closed AFL's generalizing a well-known property of the regular languages, and Greibach

*This research has been supported by the Netherlands Organization for the Advancement of Pure Research (ZWO).

[14, 15] studied AFL's closed under nested iterated substitution (super-AFL's) featuring a property of the context-free languages. In the theory of parallel rewriting (in particular in the case of ETOL languages; cf. [21]) AFL's closed under iterated parallel substitution (hyper-AFL's) were intensively investigated [24, 23, 4, 2]. On the other hand many theorems originally proved for AFL's also hold for weaker structures like semi-AFL (inspired by the linear context-free languages) and trio. For the main results on trio, (semi-)AFL and substitution-closed AFL we refer to a survey by Ginsburg [9].

Van Leeuwen [24] and Salomaa [23] originally introduced (full) hyper-AFL's and a related rewriting system called K-iteration grammar, where K refers to an arbitrary family of languages. A K-iteration grammar is essentially an ETOL system [21] in which each table (i.e., finite substitution) has been replaced by an arbitrary K-substitution. The family of languages generated by K-iteration grammar is called the hyper-algebraic extension of the family K and will be denoted by $H(K)$. According to [24, 23, 2] a nontrivial family K is a full hyper-AFL if and only if (i) K is closed under finite substitution, (ii) K is closed under intersection with regular sets, and (iii) K is hyper-algebraically closed, i.e., $H(K) = K$. In [2] it was shown that the smallest full hyper-AFL $\hat{\mathcal{H}}(K)$ containing the family K equals the hyper-algebraic extension of the smallest prequasoid $\Pi(K)$ (i.e., the smallest nontrivial family satisfying (i) and (ii)) containing K, i.e., $\hat{\mathcal{H}}(K) = H\Pi(K)$. This "Canonical Form Theorem" means that the operator $\hat{\mathcal{H}}$ can be decomposed into a single product (composition) of the simpler operators H and Π.

In this paper we investigate similar ways of connecting other well-known AFL-structures with rewriting systems (like restricted kinds of K-iteration grammars such as, e.g., context-free K-grammars introduced in [25]), in such a way that

(A) the corresponding extension X (i.e., the family of languages generated by that particular kind of restricted iteration grammars) —which may be considered as an operator on families of languages— characterizes the original AFL-structure by means of the conditions: (1) containment of a nontrivial language, (2) closure under finite substitution, (3) closure under intersection with regular sets, and (4) closure under the operator X. A family satisfying (1)–(4) will be called a full X-AFL. Thus K is a full X-AFL when $\Pi(K) = K$ and $X(K) = K$.

(B) a Canonical Form Theorem for $\hat{\mathcal{X}}(K)$ (i.e., the smallest full X-AFL containing K) holds or, equivalently, it is possible to factorize the operator $\hat{\mathcal{X}}$ into a single product of X and Π: $\hat{\mathcal{X}}(K) = X\Pi(K)$.

The particular types of AFL-structures that will be considered in this paper are full trio, full semi-AFL, full pseudo-AFL (i.e., full semi-AFL closed under concatenation), full Kleene-AFL (i.e., full trio closed under Kleene $\star$), full AFL [9], full substitution-closed AFL [11], full super-AFL [14, 15], full hyper(1)-AFL [24, 4, 8] and full hyper-AFL [24, 23, 4, 2].

From (A) it should be clear that $\hat{\mathcal{X}}(K) = \bigcup\{\Omega(K) \mid \Omega \in \{\Pi, X\}^\star\}$, or in a more algebraic notation, that $\hat{\mathcal{X}}(K) = \{\Pi, X\}^\star(K)$. Thus the Canonical Forms as mentioned in (B) imply that instead of the infinite set of operators $\{\Pi, X\}^\star$ we only need to consider the application of one single operator $X\Pi$. Therefore Canonical Forms can be successfully used in proving certain families to be full X-AFLs and in establishing properties of $\hat{\mathcal{X}}(K)$.

All extensions in this paper are obtained either "directly" by natural generalizations of basic families (like ETOL, EOL, context-free and regular languages) or "indirectly" by fixing parameters in extensions already defined in a "direct" way. Investigating extensions of language families instead of basic families enables us to uncover structural features and arguments for the rewriting mechanisms involved, rather than strictly combinatorial proofs usually given in dealing with basic families. This rather algebraic approach to language theory does not only provide a unifying framework for investigating closure properties (to which the present paper is mainly restricted) but has also succesfully been applied in other areas of formal language theory; cf. e.g., [25, 26, 27]. This approach also emphasizes again the principal role played by the concept of prequasoid (or, equivalently, the notion of nondeterministic generalized sequential machine with accepting states; cf. [18], Chapter 9) in dealing with families closed under several kinds of (iterated) substitution [2, 8, 12, 15, 19, 23, 24, 26].

This paper is organized into five sections of which this introductory section is the first. Section 2 contains preliminaries from formal language theory, definitions of the basic rewriting systems, the corresponding extensions and full X-AFL's. We conclude Section 2 with proving basic properties of these extensions and with a few examples. In Section 3 we establish Canonical Forms for full hyper-AFL, full hyper(1)-AFL, full algebraic AFL (i.e., full super-AFL) and full rational AFL (i.e., full substitution-closed AFL). Canonical Forms for the other cases (i.e., ful trio, full semi-AFL, full pseudo-AFL, full Kleene-AFL, full AFL) are proved in Section 4. The results of Section 3 and 4 are applied in Section 5 to the family obtained by substituting languages from a family into languages from another family. Section 5 also contains a few generalizations of results of Section 4 and a example of the application of Caninical Forms and related theorems in proving certain families to be particular kinds of full AFL-structures.

2 Definitions and Basic Properties

For terminology and basic results in formal language theory we refer to [22] or to [18]. Moreover, the reader is assumed to be familiar with standard concepts from the theory of parallel rewriting [17] and from AFL-theory [9].

Let Σ_ω be an arbitrary but fixed countably infinite set of symbols. A *family* (of languages) is a collection of languages over finite subsets, called *alphabets*, of Σ_ω.

Let λ denote the empty word. A language L is nontrivial if L is nonempty and $L \neq \{\lambda\}$. A family is called *nontrivial* when it contains at least one nontrivial language.

Let ONE be the family consisting of all singletons, i.e., ONE $= \{\{w\} \mid w \in \Sigma_\omega^\star\}$, and let SYMBOL be the subfamily containing all singletons of length 1, i.e., SYMBOL $= \{\{\alpha\} \mid \alpha \in \Sigma_\omega\}$. Moreover, we define the families ALPHA and STAR by ALPHA $= \{\{\alpha_1, \ldots, \alpha_n\} \mid n \geq 1,\ \alpha_1, \ldots, \alpha_n \in \Sigma_\omega\}$ and STAR $= \{\{\alpha^\star\} \mid \alpha \in \Sigma_\omega\}$, respectively. We denote the families of finite, regular (rational), context-free (algebraic), OL-, EOL-, TOL-, ETOL- and indexed [1] languages by FIN, REG, CF, OL, EOL, TOL, ETOL and INDEX, respectively.

Let K be a family. A *K-substitution* τ is a function on an alphabet V such that for each α in V, $\tau(\alpha)$ is a language in K. The function τ is extended in the usual way to words by $\tau(\lambda) = \{\lambda\}$, $\tau(\alpha_1 \cdots \alpha_n) = \tau(\alpha_1) \cdots \tau(\alpha_n)$ with $\alpha_1, \ldots, \alpha_n \in V$ $(n \geq 1)$, and to languages by $\tau(L) = \bigcup\{\tau(w) \mid w \in L\}$ for each $L \subseteq V^\star$. A K-substitution $\tau : V \to K$ with $\tau(\alpha) \subseteq V^\star$ for each α in V is called a *K-substitution over V*.

A ONE-substitution is usually called a homomorphism, and in case K equals FIN or REG, τ is called a finite or regular substitution, respectively.

With K_1/K_2 we denote the family obtained by all K_1-substitutions applied to K_2-languages, i.e., $K_1/K_2 = \{\tau(L) \mid \tau$ is a K_1-substitution and $L \in K_2\}$. This binary operation on families is neither commutative, nor associative, i.e., neither $K_1/K_2 = K_2/K_1$, nor $K_1/(K_2/K_3) = (K_1/K_2)/K_3$ holds in general [11]. But the inclusion $K_1/(K_2/K_3) \subseteq (K_1/K_2)/K_3$ does hold [11], whereas equality holds whenever K_2 is closed under isomorphism (i.e., renaming of symbols) [11]. In the sequel the operator / will be often applied to families closed under isomorphism and consequently parentheses in expressions will be omitted which results for instance in $K_1/K_2/K_3$.

A family K_2 is closed under K_1-substitution if $K_1/K_2 \subseteq K_2$. The closure of K_2 under K_1-substitution is the smallest family containing K_2 and closed under K_1-substitution. The family K_1 is closed under substitution into the family K_2 if $K_1/K_2 \subseteq K_1$. The closure of K_1 under substitution into K_2 is the smallest family containing K_1 and closed under substitution into K_2. A family is substitution closed when $K/K \subseteq K$. The substitution closure of K is the smallest substitution-closed family K_∞ containing K.

Let $[K_1/]^0 K_2 = K_2$ and by induction $[K_1/]^{n+1} K_2 = K_1/(\bigcup_{i=0}^{n} [K_1/]^i K_2)$ $(n \geq 0)$ and let $[K_1/]^\star K_2 = \bigcup_{n=0}^{\infty} [K_1/]^n K_2$. Similarly let $K_1[/K_2]^0 = K_1$ and by induction $K_1[/K_2]^{n+1} = (\bigcup_{i=0}^{n} K_1[/K_2]^i)/K_2$ $(n \geq 0)$ and let $K_1[/K_2]^\star = \bigcup_{n=0}^{\infty} K_1[/K_2]^n$.

Lemma 2.1. [11]

(1) $[K_1/]^\star K_2$ *is the closure of* K_2 *under* K_1*-substitution.*

(2) $K_1[/K_2]^\star$ *is the closure of* K_1 *under substitution into* K_2*.*

(3) $K_\infty = K[/K]^\star$*.*

(4) *If* $K \subseteq K/K$*, then* $K_\infty = [K/]^\star K$*.* □

We define the following operators on families of languages: $\Theta(K) = \mathrm{ONE}/K$, $\Phi(K) = \mathrm{FIN}/K$, and $\Delta(K) = \{L \cap R \mid L \in K,\ R \in \mathrm{REG}\}$.

The notion of prequasoid was defined in [2] as a slightly weaker variant of the quasoid introduced by Van Leeuwen [24].

Definition. A family K is a *prequasoid* if

(1) K is nontrivial,

(2) K is closed under finite substitution, i.e., $\Phi(K) \subseteq K$,

(3) K is closed under intersection with regular languages, i.e., $\Delta(K) \subseteq K$.

A *quasoid* is a prequasoid containing at least one infinite language. □

Each (pre)quasoid contains all regular (finite, respectively) languages. Consequently REG (FIN respectively) is the smallest (pre)quasoid, and FIN is the only prequasoid which is not a quasoid [2]. For the smallest prequasoid $\Pi(K)$ containing K we have

Lemma 2.2. *If K is a nontrivial family, then* $\Pi(K) = \Theta\Delta\Phi(K)$.

Proof. It is well known that a nontrivial family is a prequasoid if and only if it is closed under mappings induced by a-NGSM's (nondeterministic generalized sequential machines with accepting states; cf. [18] for definitions and details). Moreover each a-NGSM mapping T can be written as $T(L) = h(f(L) \cap R)$, where h is a homomorphism, R is a regular set and f is a finite substitution [18]. Since a-NGSM mappings are closed under composition [7], we have $\Pi(K) = \bigcup_{n=0}^{\infty} (\Theta\Delta\Phi)^n(K) = \Theta\Delta\Phi(K)$. □

The relation between substitution and the operator Δ is given by the following result established by Ginsburg & Spanier [11].

Lemma 2.3. [11]

(1) $\Delta(K_1/K_2) \subseteq \Delta(K_1)/\Delta\Phi(K_2)$,

(2) $\Delta(K_1)/\Delta(K_2) \subseteq \Theta\Delta((\mathrm{FIN} \cup K_1)/\Phi(K_2))$. □

The family K is closed under *iterated* (parallel) *substitution* if for all finite sets U of K-substitutions and for all L in K, $U^\star(L) = \bigcup\{\tau_1 \cdots \tau_n(L) \mid n \geq 0,\ \tau_1, \ldots, \tau_n \in U\}$ is in K. The family K is closed under *single iterated* (parallel) *substitution*, when $U^\star(L)$ is in K for each L in K and for each singleton $U = \{\tau\}$, i.e., when $\bigcup\{\tau^n(L) \mid n \geq 0\} \in K$. A K-substitution τ over V is called *nested*, if for each α in V, $\alpha \in \tau(\alpha)$. If for each L in a family K and each nested K-substitution τ, $\bigcup\{\tau^n(L) \mid n \geq 0\}$ is in K, then K is said to be closed under *nested iterated substitution*.

A *full trio* is a nontrivial family closed under homomorphism, inverse homomorphism and intersection with regular sets. A *full semi-AFL* is a full trio closed under union. We call a full semi-AFL closed under concatenation a *full pseudo-AFL*. Thus a *full AFL* is a full pseudo-AFL closed under Kleene $\star$. We call a full trio closed under Kleene $\star$ a *full Kleene-AFL*. In [9] it has been shown that each full Kleene-AFL closed under either union or concatenation is a full AFL. Similarly it is straightforward to show that each full trio closed under concatenation is a full pseudo-AFL. A *full substitution-closed AFL* [*full super-AFL, full hyper(1)-AFL, full hyper-AFL*] is a full AFL closed under substitution [nested iterated substitution, single iterated substitution, iterated substitution, respectively]. Let $\hat{\mathcal{M}}$, $\hat{\mathcal{C}}$, $\hat{\mathcal{S}}$, $\hat{\mathcal{P}}$, $\hat{\mathcal{F}}$, $\hat{\mathcal{R}}$, $\hat{\mathcal{A}}$, $\hat{\mathcal{H}}_1$ and $\hat{\mathcal{H}}$ denote the least full trio, full Kleene-AFL, full semi-AFL, full pseudo-AFL, full AFL, full substitution-closed AFL, full super-AFL, full hyper(1)-AFL and full hyper-AFL containing K, respectively. In Sections 3 and 4 we will show that all these AFL-structures are different.

Ginsburg & Spanier [11] proved the following

Lemma 2.4. *If K is a quasoid, then K_∞ is a full substitution-closed AFL.* □

We now come to the main definitions. They paraphrase and slightly generalize the original definitions introduced by Van Leeuwen [24, 25] and Salomaa [23] (cf. [2]), in order to fit standard AFL-structures in the framework of family extensions; cf. Section 3 and 4.

Definition. Let K_1, K_2 and K_3 be families. A language L is called *hyper-algebraic* over (K_1, K_2, K_3) when there exist (i) an alphabet V, (ii) a terminal alphabet Σ, (iii) an initial language $L_0 \subseteq V^\star$ in K_1, (iv) an initial K_2-substitution τ_0 over V, and (v) a finite set U of K_3-substitutions over V such that $L = U^\star(\tau_0(L)) \cap \Sigma^\star$. If each τ in U is nested, then L is called *algebraic* over (K_1, K_2, K_3), whereas such an algebraic L is called *rational* over (K_1, K_2, K_3) when U is not *self-embedding*, i.e., if for all u in $U^\star$ and for all α in V the implication

$$w_1 \alpha w_2 \in u(\alpha) \Rightarrow (w_1 = \lambda \text{ or } \mathrm{w}_2 = \lambda)$$

holds, where $w_1, w_2 \in V^\star$. The *hyper-algebraic* [*algebraic, rational*, respectively] *extension* $\underline{H}(K_1, K_2, K_3)$ [$\underline{A}(K_1, K_2, K_3)$, $\underline{R}(K_1, K_2, K_3)$] of (K_1, K_2, K_3) is the family of all languages hyper-algebraic [algebraic, rational] over (K_1, K_2, K_3). □

We are mainly interested in several less general extensions which are easily obtained as particular cases.

Definition. Let K be a family. The *hyper-algebraic, algebraic* and *rational extension* of K are respectively

$$H(K) = \underline{H}(\mathrm{SYMBOL}, \mathrm{SYMBOL}, K),$$
$$A(K) = \underline{A}(\mathrm{SYMBOL}, \mathrm{SYMBOL}, K), \text{ and}$$
$$R(K) = \underline{R}(\mathrm{SYMBOL}, \mathrm{SYMBOL}, K).$$
□

Let $K_1 = K_2 = \mathrm{SYMBOL}$ and $K_3 = K$. Let V, Σ, U, τ_0 and L_0 be as in the previous definition. The construct $G = (V, \Sigma, U, S)$, where $\{S\} = \tau_0(L_0)$ is called a *K-iteration grammar*, if U is a finite set of K-substitutions over V. We call G a *context-free K-grammar* [*regular K-grammar*] if each τ in U is a nested K-substitution over V [and U is not self-embedding]. The grammar G generates a language $L(G) = U^\star(S) \cap \Sigma^\star$ which is a member of $H(K)$ [$A(K)$, $R(K)$, respectively].

The m-restricted extensions $H_m(K)$, $A_m(K)$ and $R_m(K)$ are the subfamilies of $H(K)$, $A(K)$ and $R(K)$ respectively, generated by grammars containing at most m K-substitutions ($m \geq 1$).

Another type of less general extensions are the following, which we call the *subrational extensions* of K:

$$M(K) = \underline{R}(\mathrm{SYMBOL}, K, \mathrm{FIN}),$$
$$C(K) = \underline{R}(\mathrm{STAR}, K, \mathrm{FIN}),$$
$$S(K) = \underline{R}(\mathrm{ALPHA}, K, \mathrm{FIN}),$$
$$P(K) = \underline{R}(\mathrm{FIN}, K, \mathrm{FIN}),$$
$$F(K) = \underline{R}(\mathrm{REG}, K, \mathrm{FIN}).$$

We call a family K *σ-simple* when (i) K contains a SYMBOL-language, and (ii) K is closed under isomorphism ("remaining of symbols"). A σ-simple family closed under union with SYMBOL-languages is called *α-simple*. Clearly, SYMBOL [ALPHA, respectively] is the smallest σ-simple [α-simple] family and it is contained in each σ-simple [α-simple] family. Obviously, each prequasoid is σ-simple. Consider a nontrivial language $L \subseteq \Sigma^\star$ from a prequasoid K. Let f be a finite substitution and h a homomorphism defined by $f(a) = \{a, b, e\}$ for each a in Σ, $e \notin \Sigma$ and $h(e) = \lambda$, $h(\alpha) = \alpha$ for $\alpha \in \Sigma \cup \{b\}$. Then $L \cup \{b\} = h(f(L) \cap (\Sigma^\star \cup be^\star))$ is in K. Hence each prequasoid is α-simple.

Our first result deals with the relation between extensions and the corresponding m-restricted variants.

Theorem 2.5.

(1) *If K is σ-simple, then $H_2(K) = H_m(K) = H(K)$ for each $m \geq 2$.*

(2) *If K is α-simple, then $A_1(K) = A_m(K) = A(K)$ and $R_1(K) = R_m(K) = R(K)$ for each $m \geq 1$.*

Proof. (1) was already estiblished in [2]. Thus it remains to show that for each $m \geq 1$, $A(K) \subseteq A_m(K) \subseteq A_1(K)$ and similarly for R, since the converse inclusions are obvious.

Let $G = (V, \Sigma, U, S)$ be a context-free K-grammar with $U = \{\tau_1, \ldots, \tau_m\}$ for some $m \geq 2$. Define for each k with $1 \leq k \leq m$ an isomorphism $\varphi_k(\alpha) = \alpha_k$ (α in V; all α_k's are new symbols) and extend these isomorphisms in the usual way to words and languages. Define a new alphabet $V_0 = V \cup \{\varphi_k(\alpha) \mid \alpha \in V,\ 1 \leq k \leq m\}$. Consider the context-free K-grammar $G_0 = (V_0, \Sigma, \tau, S)$ where the nested K-substitution τ over V_0 is defined by

$$\begin{array}{lll} \tau(\alpha) = \{\alpha, \alpha_1\} & \alpha \in V, & \\ \tau(\alpha_k) = \{\alpha_k\} \cup C_k \cup \tau_k(\alpha) & \text{with } C_k = \varnothing & \text{iff } k = m, \\ & \text{and } C_k = \{\alpha_{k+1}\} & \text{iff } \leq k \leq m{-}1. \end{array}$$

The basic idea of the simulation of G by G_0 is the following: each occurrence of each symbol β in V_0 may be object to the following replacements (in an "asynchronous way"): (i) changing into α_{k+1} (if $\beta = \alpha_k$, $1 \leq k \leq m{-}1$) or into α_1 (if $\beta = \alpha$), i.e., from from α we can reach α_j for each j ($1 \leq j \leq m$), (ii) substituting $\tau_j(\alpha)$ into that particular instance of $\beta = \alpha_j$, i.e., we simulate the application of τ_j on that occurrence of α_j while the subscript j is removed.

By this construction we obtain $L(G_0) = L(G)$ and hence $A(K) \subseteq A_m(K) \subseteq A_1(K)$ for each $m \geq 1$. Since the construction preserves the not self-embedding property of the grammar, a similar conclusion holds in the rational case. □

Note that 2.5(2) is obvious, whenever K is closed under union. In the hyper-algebraic case 2.5(1) is the best possible result, i.e., in general one cannot reduce the number of substitutions in a K-iteration grammar to one [21, 8].

By relating subrational extensions to (finitely repeated) substitutions we obtain a considerable simplification.

Theorem 2.6. $\underline{R}(K_1, K, \mathrm{FIN}) = \mathrm{REG}/(K/K_1)$*, and consequently*

$M(K) = \mathrm{REG}/K,$

$C(K) = \mathrm{REG}/(K/\mathrm{STAR}),$

$S(K) = \mathrm{REG}/(K/\mathrm{ALPHA}),$

$P(K) = \text{REG}/(K/\text{FIN})$, *and*
$F(K) = \text{REG}/(K/\text{REG})$,

whereas the parentheses may be omitted whenever K is closed under isomorphism.

Proof. Suppose $L \in \text{REG}/(K/K_1)$ and let $L = \rho\tau_0(L_0)$ where $L \subseteq \Delta^\star$ is in K_1, $\tau_0 : \Delta \to K$ is a K-substitution with $\tau(\alpha) \subseteq B^\star$ for each α in Δ, and $\rho : B \to \text{REG}$ is a regular substitution with $\rho(\beta) \subseteq \Sigma_\beta^\star$ ($\beta \in B$), and $\Sigma = \bigcup\{\Sigma_\beta \mid \beta \in B\}$. For each $\beta \in B$, let $(V_\beta, \Sigma_\beta, P_\beta, \beta)$ be an ordinary regular grammar generating $\rho(\beta)$ such that all nonterminal alphabets $V_\beta - \Sigma_\beta$ are mutually disjoint. Let $U = \{\tau\}$ where τ is a substitution over $\bigcup\{V_\beta \mid \beta \in B\} \cup \Delta$ defined by

$$\tau(\gamma) = \{\gamma\} \cup \{w \mid (\gamma, w) \in P_\beta\} \in \text{FIN} \quad \text{iff } \gamma \in V_\beta - \Sigma_\beta \ (\beta \in B),$$
$$\tau(\gamma) = \{\gamma\} \quad \text{otherwise.}$$

Clearly, U is not self-embedding and therefore we have $L = \rho\tau_0(L_0) = U^\star(\tau_0(L_0)) \cap \Sigma^\star \in \underline{R}(K_1, K, \text{FIN})$.

Conversely, let $L \in \underline{R}(K_1, K, \text{FIN})$, i.e., there exist alphabets V and Σ, a language L_0 in K_1, a K-substitution τ_0 over V, and a finite set U of finite substitutions over V, such that $L = U^\star(\tau_0(L_0)) \cap \Sigma^\star$. Since FIN satisfies the conditions required in 2.5(2) we can (by a similar argument as in the proof of 2.5(2)) restrict our attention to the case $U = \{\tau\}$. Consider for each α in V the grammar $G_\alpha = (V, \Sigma, \tau, \alpha)$. Since each G_α is not self-embedding and τ is a finite substitution, $L(G) = (\bigcup\{\tau^n(\alpha) \mid n \geq 0\}) \cap \Sigma^\star$ is regular (cf., e.g., [22]). Hence there exists a regular substitution $\rho : V \to \text{REG}$ with $\rho(\alpha) = L(G_\alpha)$ ($\alpha \in V$) and consequently $L = \rho\tau_0(L_0) \in \text{REG}/(K/K_1)$. □

With each extension introduced above we associate a full AFL-structure as follows.

Definition. Let X be a symbol from $\{H, H_1, A, R, F, P, S, C, M\}$. A family K is a *full X-AFL* if

(i) K is a prequasoid, i.e., $\Pi(K) = K$, and
(ii) K is equal to its X-extension, i.e., $X(K) = K$. □

In the subsequent two sections we relate full X-AFL's to well-known AFL-structures.

Examples. (1) $H(\text{ONE}) = \text{EDTOL}$; $H(\text{FIN}) = H(\text{REG}) = \text{ETOL}$; $H(\text{INDEX}) = \text{INDEX}$, hence ETOL and INDEX are full H-AFL's [24, 23, 4].
(2) $H_1(\text{ONE}) = \text{EDOL}$; $H_1(\text{FIN}) = \text{EOL}$ [24].
(3) $A(\text{FIN}) = A(\text{REG}) = A(\text{CF}) = \text{CF}$ [25, 26, 14, 15], hence CF is a full A-AFL.
(4) $R(\text{FIN}) = R(\text{REG}) = \text{REG}$ (cf. 2.6 and 3.3 below), hence REG is a full R-AFL.
(5) For each X from $\{M, C, S, P, F\}$, we have $X(\text{FIN}) = X(\text{REG}) = \text{REG}$ (cf. 2.6) and consequently REG is a full X-AFL. □

3 Canonical Forms for Full H-, H_1-, A- and R-AFL's

In [2] it was already shown that a family is a full H-AFL if and only if it is a full hyper-AFL (cf. Section 1, [24, 23]). By a similar argument we obtain that a family is a full H_1-AFL if and only if it is a full hyper(1)-AFL. It is straightforward to show that the notions of full algebraic AFL (i.e., full A-AFL) and full super-AFL are equivalent; cf. [14]. However, establishing the equivalence between the concepts of full rational AFL (i.e., full R-AFL) and of full substitution-closed AFL is more complicated. The proof is based on two lemmas which are of some interest on their own.

Lemma 3.1. *If K is α-simple, then $K_\infty \subseteq R(K)$.*

Proof. Since SYMBOL $\subseteq K$ we have $K \subseteq K/K$ and hence by 2.1(4), $K_\infty = [K/]^\star K$. Moreover $K \subseteq K/K$ also implies $[K/]^{n+1}K = K/(\bigcup_{i=0}^{n}[K/]^i K) = K/[K/]^n K$.

In order to establish the inclusion $K_\infty \subseteq R(K)$ we show that for all natural numbers p, $[K/]^p K \subseteq R(K)$. The proof is by induction on p.

Initial step ($p = 0$): $[K/]^0 K = K$. Let $L_0 \subseteq \Sigma^\star$ be in K and consider the context-free K-grammar $G = (V, \Sigma, \{\tau\}, S)$ where $V = \Sigma \cup \{S\}$, $S \notin \Sigma$, $\tau(S) = \{S\} \cup L_0$ and $\tau(a) = \{a\}$ for each a in Σ. Clearly, G is not self-embedding and $L(G) = L_0$ is in $R(K)$.

Induction step: Suppose $[K/]^p K \subseteq R(K)$, then we have to show that $[K/]^{p+1}K \subseteq R(K)$. By the induction hypothesis we obtain $[K/]^{p+1}K = K/[K/]^p K \subseteq K/R(K)$. Thus it only remains to prove that $K/R(K) \subseteq R(K)$.

Let $L' \subseteq \Sigma^\star$ be a language in $R(K)$ generated by the regular K-grammar $G = (V, \Sigma, \{\tau\}, S)$, and let g be a K-substitution on Σ with $\bigcup\{g(\alpha) \mid \alpha \in \Sigma\} \subseteq \Sigma_0^\star$ for some alphabet Σ_0. Without loss of generality we may assume that $V \cap \Sigma_0 = \varnothing$. Consider the context-free K-grammar $G_0 = (V_0, \Sigma_0, U_0, S)$ where $V_0 = V \cup \Sigma_0$, $U_0 = \{\tau_0, \tau_1\}$ with

$$\begin{aligned} \tau_0(\alpha) &= \{\alpha\}, && \text{for } \alpha \text{ in } V_0 - \Sigma,\\ \tau_0(\alpha) &= \{\alpha\} \cup g(\alpha) \quad, && \text{for } \alpha \text{ in } \Sigma,\\ \tau_1(\alpha) &= \tau(\alpha), && \text{for } \alpha \text{ in } V,\\ \tau_1(\alpha) &= \{\alpha\}, && \text{for } \alpha \text{ in } \Sigma_0. \end{aligned}$$

Then clearly, G_0 is a not self-embedding context-free K-grammar and $L(G_0) = g(L(G)) = g(L') \in R(K)$.

This completes the induction and establishes the inclusion $K_\infty \subseteq R(K)$. □

Lemma 3.2. *If K is a prequasoid, then $R(K) \subseteq K_\infty \cup \mathrm{REG}$.*

Proof. If K is a prequasoid but not a quasoid, then $K = \mathrm{FIN}$ and hence $R(\mathrm{FIN}) = \mathrm{REG} \subseteq K_\infty \cup \mathrm{REG}$.

Let K be a quasoid, then $\mathrm{REG} \subseteq K \subseteq K_\infty$. By 2.5(2) we only have to prove that $R_1(K) \subseteq K_\infty$. We first show that $R_1(K_\infty) \subseteq K_\infty$. The proof of this inclusion is a generalization of the argument that each non-self-embedding context-free grammar generates a regular language; cf. [22].

Consider a 1-restricted regular K_∞-grammar $G = (V, \Sigma, \{\tau\}, S)$. Without loss of generality we may assume that $\tau(a) = \{a\}$ for each a in Σ. (Otherwise, we introduce for each a in Σ a new nonterminal symbol A_a and we define an isomorphism $\varphi(a) = A_a$ ($a \in \Sigma$), $\varphi(\alpha) = \alpha$ ($\alpha \in V - \Sigma$). We replace τ by τ_0 with $\tau_0(\alpha) = \varphi\tau(\alpha)$ iff $\alpha \in V - \Sigma$, $\tau_0(a) = \{a\}$ iff $a \in \Sigma$, $\tau_o(A_a) = \{A_a, a\} \cup \varphi\tau(a)$ for each a in Σ.)

Moreover we assume that for each α in V there is a sequence u in $\tau^\star$ such that there is a word in $u(S)$ that contains an occurrence of α. Otherwise we can remove α from V and intersect all languages involved in G with $(V - \{\alpha\})^\star$ (due to the fact that K_∞ is a full AFL by 2.4), without affecting $L(G)$. We consider the following cases:

Case 1: For each A in $V - \Sigma$, there is a sequence u in $\tau^\star$ such that $u(A)$ contains a word in which S occurs.

If $w \in \tau(A)$ is an arbitrary word containing a nonterminal symbol, then it is of one of the four forms: (i) $w = \varphi B \psi$, (ii) $w = \varphi B$, (iii) $w = B\psi$, or (iv) $w = B$, where φ and ψ are nonempty words over V, and $B \in V - \Sigma$. If w satisfies (i) we must have by the assumption of Case 1: there exist sequences u_1 and u_2 in $\tau^\star$ such that

$$u_1 u_2 \tau(A) \supseteq u_1 u_2(\varphi B \psi) \supseteq u_1(\varphi\varphi_1 S \psi_1 \psi) \supseteq \{\varphi\varphi_1\varphi_2 A \psi_2 \psi_1 \psi\}$$

for some (possibly empty) words φ_1, φ_2, ψ_1 and ψ_2 over V. But then, since φ and ψ are nonempty, G is a self-embedding context-free K_∞-grammar which contradicts the assumption that G is regular. We obtain the same contradiction if $\tau(A)$ contains words of both forms (ii) and (iii). Thus if $\tau(A)$ contains a word of the form (ii), then all words of the form (ii) contained in $\tau(A)$ the word φ is in $\Sigma^\star$. (Otherwise $\tau(A)$ would also contain a word of the form (i) or (iii).) Hence G is a "right-linear" context-free K_∞-grammar, i.e., for each $A \in V - \Sigma$, we have $\tau(A) \subseteq \Sigma^\star(V - \Sigma) \cup \Sigma^\star$. By a similar argument we conclude that if $\tau(A)$ contains a word of the form (iii), then G is "left linear", i.e., for each A in $V - \Sigma$, $\tau(A) \subseteq (V - \Sigma)\Sigma^\star \cup \Sigma^\star$.

So if G is right linear, then we have $\tau(A) = \bigcup\{L_{AX}\{X\} \mid X \in V - \Sigma\} \cup L_A$ for each A in $V - \Sigma$, where L_A and L_{AX} are languages over Σ. (The left-linear case is similar.) Note that L_{AX} is empty whenever X does not occur in $\tau(A)$. Since K_∞ is a full AFL (by 2.4), we have that L_A, and each (nonempty) L_{AX} are in K_∞ and, consequently, $R_1(K_\infty) \subseteq K_\infty/\mathrm{REG} = K_\infty$; cf. 2.4 and [9, 11].

Case 2: There exists a nonterminal symbol A such that for no words φ and ψ in $V^\star$ and for no sequence u in $\{\tau\}^\star$, $\varphi S \psi \in u(A)$.

The proof of $L(G)$ being in K_∞ proceeds by induction on the number of nonterminal symbols m. For $m = 1$, there is nothing to prove because $\lambda(S) = \{S\}$.

Assume that the assertion holds for $m = n$. Let the number of nonterminal symbols in $V - \Sigma$ be $n + 1$. Consider the context-free K_∞-grammar $G_1 = (V - \{S\}, \Sigma, \tau_1, A)$ with

$$\tau_1(\alpha) = \tau(\alpha) \cap (V - \{S\})^\star \qquad \alpha \neq S$$

and the context-free K_∞-grammar $G_2 = (V, \Sigma \cup \{A\}, \tau_2, S)$ with

$$\tau_2(A) = \{A\},$$
$$\tau_2(\alpha) = \tau(\alpha) \qquad \alpha \neq A.$$

Then both G_1 and G_2 are not self-embedding grammars having n nonterminal symbols. Now both of the languages $L(G_1)$ and $L(G_2)$ are in K_∞; either by the induction hypothesis or by Case 1. But $L(G)$ is the result of substituting $L(G_1)$ for A in $L(G_2)$. The fact that K_∞ is substitution closed implies that $L(G)$ is in K_∞. This completes the induction and establishes the inclusion $R_1(K_\infty) \subseteq K_\infty$.

Finally, together with $K \subseteq K_\infty$, this yields $R_1(K) \subseteq R_1(K_\infty) \subseteq K_\infty$. □

Theorem 3.3.

(1) *If K is a prequasoid, then* $R(K) = K_\infty \cup \mathrm{REG}$.

(2) *K is a full R-AFL if and only if K is a full substitution-closed AFL.*

Proof. (1) Lemmas 3.1 and 3.2.

(2) Let K be a full R-AFL. Then $\mathrm{FIN} \subseteq K$ and consequently $\mathrm{REG} \subseteq R(K) = K$. Now 3.3(1) implies $R(K) = K = K_\infty$. Hence by 2.4 K is a full substitution-closed AFL.

Conversely, let K be a full substitution-closed AFL. Obviously, K is a prequasoid and by (1): $K = K_\infty = R(K)$, i.e., K is a full R-AFL. □

Let $H_1^0(K) = K$, $H_1^{n+1}(K) = H_1 H_1^n(K)$ for $n \geq 0$ and $H_1^\star(K) = \bigcup_{n=0}^\infty H_1^n(K)$. In the remaining part of this section X will denote a symbol from $\{H, H_1^\star, A, R\}$. Let $\hat{\mathcal{X}}(K)$ be the least corresponding full X-AFL containing K (i.e., $\hat{\mathcal{X}}$ equals $\hat{\mathcal{H}}$, $\hat{\mathcal{H}}_1$, $\hat{\mathcal{A}}$ and $\hat{\mathcal{R}}$, respectively).

We now prove the main result of this section, i.e., we will decompose the operator $\hat{\mathcal{X}}$ into simpler operators like X, Π, Δ and Φ. Some other factorizations were already known. E.g., Greibach [14] established a result that may be interpreted as $\hat{\mathcal{A}}(K) = \Delta A \hat{\mathcal{F}}(K)$. Similarly, from 2.4 due to Ginsburg & Spanier [11] we can infer that $\hat{\mathcal{R}}(K)$ equals the substitution closure of $\Pi(K)$ provided K contains an infinite language.

Theorem 3.4. (Canonical Form Theorem for Full H-, $H_1^\star$-, A- and R-AFL's)

(1) $\hat{\mathcal{X}}(K) = X\Pi(K)$.

(2) *If K is nontrivial, then* $\hat{\mathcal{X}}(K) = X\Delta\Phi(K)$.

Proof. (1) We distinguish the following cases:

$\underline{X = H}$: In [2] we proved that $H(K)$ is a full H-AFL, provided K is a prequasoid. Hence $\hat{\mathcal{H}}(K) \subseteq H\Pi(K)$. Conversely, we have $K \subseteq \hat{\mathcal{H}}(K)$ which implies $H\Pi(K) \subseteq H\Pi\hat{\mathcal{H}}(K) = \hat{\mathcal{H}}(K)$.

$\underline{X = H_1^\star}$: Similar to the previous case it suffices to show that $H_1^\star(K)$ is a full H_1-AFL, whenever K is a prequasoid. So let K be a prequasoid. Obviously, $H_1^\star(K)$ is H_1-closed and thus it remains to prove that $H_1^\star$ is a prequasoid. This will be done by induction. Since $\mathrm{FIN} \subseteq K$, we have $H_1(\mathrm{EOL}) \subseteq H_1^2(K)$. Applying a result from [23] yields that $H_1^2(K)$ is a full AFL. Suppose $H_1^n(K)$, $n \geq 2$ is a full AFL. Then $H_1^{n+1}(K) = H_1 H_1^n(K)$ is also a full AFL [23], which completes the induction.

$\underline{X = A}$: Let K be a prequasoid. We will show that $A(K)$ is a full A-AFL. Since the proof that $A(K)$ is a prequasoid is rather standard (cf. [26, 2]) and the inclusion $A(K) \subseteq AA(K)$ is trivial, it remains to establish that $AA(K) \subseteq A(K)$.

Let $G = (V, \Sigma, \tau, S)$ be a context-free $A(K)$-grammar, i.e., τ is a nested $A(K)$-substitution. For each α in V, let $G_\alpha = (V_\alpha, V, \tau_\alpha, S_\alpha)$ be a context-free K-grammar (i.e., each τ_α is a nested K-substitution) such that $L(G_\alpha) = \tau(\alpha)$. Obviously, we may assume that all nonterminal alphabets $V_\alpha - V$ are mutually disjoint. Thus we have to show that $L(G) \in A(K)$.

We modify each G_α in such a way that $\tau_\alpha(\beta) = \{\beta\}$ for each β in V (cf. the proof of

3.2). Consider the context-free K-grammar $G_0 = (V_0, \Sigma, U_0, S)$ where $V_0 = \bigcup\{V_\alpha \mid \alpha \in V\}$ and $U_0 = \{\sigma_\alpha \mid \alpha \in V\}$ with for each α in V:

$$\begin{aligned} \sigma_\alpha(\beta) &= \tau_\alpha(\beta) && \beta \in V_\alpha - V, \\ \sigma_\alpha(\beta) &= \{\beta, S_\beta\} && \beta \in V, \\ \sigma_\alpha(\beta) &= \{\beta\} && \beta \in V_0 - V_\alpha. \end{aligned}$$

Clearly, $L(G_0) = L(G)$ and hence $L(G)$ is in $A(K)$.

$\underline{X = R}$: The fact that $R(K)$ is a full R-AFL whenever K is a prequasoid directly follows from 3.3 and 2.4.

(2) By (1) and 2.2, we have $\hat{\mathcal{H}}(K) = H\Theta\Delta\Phi(K)$. Therefore it suffices to show that $H\Theta\Delta\Phi(K) \subseteq H\Delta\Phi(K)$.

Let $L \subseteq \Omega^\star$ be a nontrivial language in the family K. Define a finite substitution f by $f(\alpha) = \{\lambda, \omega\}$ for each α in Ω, where ω is an arbitrary but fixed word. Then clearly $\{\omega\} = f(L) \cap \{\omega\}$ is in $\Delta\Phi(K)$, i.e., ONE $\subseteq \Delta\Phi(K)$, which enables us to replace homomorphisms (according to Θ) by $\Delta\Phi(K)$-substitutions as follows. Let $G = (V, \Sigma, U, S)$ be a $\Theta\Delta\Phi(K)$-iteration grammar and consider the $\Delta\Phi(K)$-iteration grammar $G_0 = (V_0, \Sigma, U_0, S)$ where $V_0 = \bigcup\{\varphi_{\alpha\tau}(V) \mid \alpha \in V;\ \tau \in U\}$ with each $\varphi_{\alpha\tau}$ is an isomorphism such that all alphabets in this union are mtually disjoint. For each τ in U, we define a substitution τ' in U_0 by $\tau'(\alpha) = L_{\alpha\tau}$ if and only if $\alpha \in V$, $L_{\alpha\tau} \subseteq (\varphi_{\alpha\tau}(V))^\star$ is in $\Delta\Phi(K)$ (Note that $\Delta\Phi(K)$ is closed under isomorphism.) and $h_{\alpha\tau}(L_{\alpha\tau}) = \tau(\alpha)$ where $h_{\alpha\tau}$ is the relative homomorphism according to Θ:

$$\begin{aligned} \tau'(\alpha) &= h_{\alpha\tau}(\alpha) && \text{if } \alpha \in \varphi_{\alpha\tau}(V), \\ \tau'(\beta) &= \{\beta\} && \text{if } \beta \in V_0 - V - \varphi_{\alpha\tau}(V). \end{aligned}$$

By this construction we have $L(G_0) = L(G)$ and consequently $\hat{\mathcal{H}}(K) = H\Delta\Phi(K)$ for each nontrivial K.

Since this construction preserves the number of substitutions and the not self-embedding property, whereas nesting can also be incorporated, the same conclusion holds in the other cases. □

As each prequasoid contains FIN, and FIN is the smallest prequasoid we have

Corollary 3.5. *Let X be a symbol from $\{H, H_1^\star, A, R\}$. Then*

(1) *each full X-AFL contains the family X(FIN);*

(2) X(FIN) *is the smallest full X-AFL.* □

This corollary implies that $H(\text{FIN}) = \text{ETOL}$ is the smallest full H-AFL [4], $A(\text{FIN}) = \text{CF}$ is the smallest full A-AFL [14], and $R(\text{FIN}) = \text{REG}$ is the smallest full R-AFL [11].

An improvement in the H_1-case to $\hat{\mathcal{H}}_1(K) = H_1\Pi(K)$ instead of $H_1^\star\Pi(K)$ is impossible because $H_1(\text{FIN}) = \text{EOL}$ is not even a full AFL [16]. Starting with quasoids rather than prequasoids indeed yields full AFL's [23] but no full H_1-AFL's or even full R-AFL's [4]. Recently Engelfriet [8] showed that the iteration of H_1 (applied to $\Pi(K)$) cannot be reduced to a finite power since $H_1^\star(\text{FIN})$ (i.e., the smallest full H_1-AFL) gives rise to an infinite hierarchy of full AFL's:

$H_1^2(\text{FIN}) \subset H_1^3(\text{FIN}) \subset \cdots \subset H_1^n(\text{FIN}) \subset \cdots$.

For X equal to H, H_1, A and R we denote the class of all full X- AFL's over Σ_ω by $\underline{X\text{-AFL}}$.

Corollary 3.6. $\underline{R\text{-AFL}} \supset \underline{A\text{-AFL}} \supset \underline{H_1\text{-AFL}} \supset \underline{H\text{-AFL}}$.

Proof. From the definitions in 2 it is clear that

$\underline{R\text{-AFL}} \supseteq \underline{A\text{-AFL}} \supseteq \underline{H_1\text{-AFL}} \supseteq \underline{H\text{-AFL}}$.

Since the smallest full X-AFL's for $X \in \{R, A, H_1, H\}$ are mutually different (i.e., REG $\subset$ CF $\subset H_1^\star(\text{FIN}) \subset$ ETOL (cf. 3.5 and [21, 8]) the inclusions are also proper. □

4 Canonical Forms for Subrationally Closed Full AFL's

Let X be a symbol from $\{M, C, S, P, F\}$. First we relate full X-AFL's to well-known AFL-structures.

In this section we only give detailed proofs for a few typical cases. The other arguments are obtained by straightforward modifications and are left as simple exercises.

Theorem 4.1.

(1) *K is a full M-AFL if, and only if, K is a full trio.*

(2) *K is a full C-AFL if, and only if, K is a full Kleene-AFL.*

(3) *K is a full S-AFL if, and only if, K is a full semi-AFL.*

(4) *K is a full P-AFL if, and only if, K is a full pseudo-AFL.*

(5) *K is a full F-AFL if, and only if, K is a full AFL.*

Proof. (5) Let K be a full F-AFL, i.e., K is a prequasoid and $K = \underline{R}(\text{REG}, K, \text{FIN})$. By 2.6 we have $K = \text{REG}/K/\text{REG}$. This implies (i) K is closed under Δ, (ii) REG $\subseteq K$, because SYMBOL $\subseteq K$, (iii) REG$/K \subseteq K$ and $K/$REG $\subseteq K$, since SYMBOL $\subseteq$ REG, which means that K is a full AFL [11].

Conversely, let K be a full AFL. Clearly, K is a prequasoid. Moreover, $K/\text{REG} \subseteq K$ and $\text{REG}/K \subseteq K$ hold [11]. Applying 2.6 yields $\underline{R}(\text{REG}, K, \text{FIN}) = \text{REG}/K/\text{REG} \subseteq K$, whereas the opposite inclusion is obvious. □

Let $\hat{\mathcal{X}}(K)$ denote the least full X-AFL containing K, i.e., if X equals M, C, S, P or F, then $\hat{\mathcal{X}}$ is euqal to $\hat{\mathcal{M}}$, $\hat{\mathcal{C}}$, $\hat{\mathcal{S}}$, $\hat{\mathcal{P}}$ or $\hat{\mathcal{F}}$, respectively. With respect to Canonical Forms the case $X = C$ differs from the other cases and will be treated separalely; cf. 4.4.

Theorem 4.2. (First Canonical Form Theorem for Full M-, S-, P- and F-AFL's.) *Let X be a symbol from $\{M, S, P, F\}$ and let K be a family of languages. Then*

(1) $\hat{\mathcal{X}}(K) = X\Pi(K)$.

(2) *If K is nontrivial, then* $\hat{\mathcal{X}}(K) = X\Delta\Phi(K)$.

Proof. (1) $\underline{X = S}$: Similar as in the proof of 3.4 it suffices to show that $S(K)$ is a full S-AFL provided K is a prequasoid. So let K be a prequasoid, then by 2.6 $S(K) = \text{REG}/K/\text{ALPHA}$ and hence

(i) $\Phi S(K) = \text{FIN}/\text{REG}/K/\text{ALPHA} = \text{REG}/K/\text{ALPHA} = S(K)$;

(ii) $\Delta S(K) = \Delta(\text{REG}/K/\text{ALPHA}) \subseteq \Delta(\text{REG})/\Delta\Phi(K/\text{ALPHA}) = \text{REG}/\Delta(K/\text{ALPHA}) = \text{REG}/\{(\bigcup_i L_i) \cap R \mid L_i \in K; R \in \text{REG}\} = \text{REG}/\{\bigcup_i (L_i \cap R) \mid L_i \in K; R \in \text{REG}\} = \text{REG}/K/\text{ALPHA} = S(K)$ (The inclusion is obtained by 2.3(1); i goes through any finite index set.);

(iii) $SS(K) = \text{REG}/\text{REG}/K/\text{ALPHA}/\text{ALPHA} = \text{REG}/K/\text{ALPHA} = S(K)$.

(i), (ii) and (iii) imply that $S(K)$ is a full S-AFL.

<u>$X = F$</u>: Apart from taking REG instead of ALPHA, the cases (i) and (iii) are the same as above. So it remains to show that $\Delta F(K) \subseteq F(K)$. Applying 2.3(1) twice yields: $\Delta F(K) = \Delta(\text{REG}/K/\text{REG}) \subseteq \Delta(\text{REG})/\Delta\Phi(K/\text{REG}) = \text{REG}/\Delta(K/\text{REG}) \subseteq \text{REG}/\Delta(K)/\Delta\Phi(\text{REG}) = \text{REG}/K/\text{REG} = F(K)$.

(2) Let K_X be SYMBOL, ALPHA, FIN and REG for X equal to M, S, P and F, respectively. By (1), 2.2 and the fact that $\Delta\Phi(K)$ is closed under isomorphism we obtain $\hat{\mathcal{X}}(K) = X\Pi(K) = X\Theta\Delta\Phi(K) = \text{REG}/\text{ONE}/\Delta\Phi(K)/K_X = \text{REG}/\Delta\Phi(K)/K_X = X\Delta\Phi(K)$. □

From 2.6 and 4.2(1) we may infer other Canonical Forms like $\hat{\mathcal{F}}(K) = \hat{\mathcal{M}}(K)/\text{REG}$ or $\hat{\mathcal{S}}(K) = \hat{\mathcal{M}}(K)/\text{ALPHA}$ which were originally established in [10] and [19] respectively; cf. [12, 20].

Theorem 4.3. (Second Canonical Form Theorem for Full M-, S-, P- and F-AFL's.) *If K is σ-simple, then $\hat{\mathcal{X}}(K) = \Pi X(K) = \Theta\Delta X(K)$ for each X from $\{M, S, P, F\}$.*

Proof. The inclusion $K \subseteq \Pi(K)$ and 4.2 imply $X(K) \subseteq X\Pi(K) = \hat{\mathcal{X}}(K)$ and hence $\Theta\Delta X(K) \subseteq \Pi X(K) \subseteq \Pi\hat{\mathcal{X}}(K) = \hat{\mathcal{X}}(K)$. Thus it remains to show that $\hat{\mathcal{X}}(K) \subseteq \Theta\Delta X(K)$. Distinguish the following cases.

<u>$X = M$</u>: According to 4.2, 2.6 and 2.3(2) we have

$$\hat{\mathcal{M}}(K) = \text{REG}/\Delta\Phi(K) \subseteq \Theta\Delta((\text{REG} \cup \text{FIN})/\Phi\Phi(K)) = \Theta\Delta M(K).$$

<u>$X = S$</u>: By 4.2 and the previous case we have $\hat{\mathcal{S}}(K) = \{\bigcup_i L_i \mid L_i \in \hat{\mathcal{M}}(K)\}; 1 \le i \le n\} = \{\bigcup_i L_i \mid L_i \in \Theta\Delta M(K); 1 \le i \le n\}$, i.e., $L_i = h_i(L_i' \cap R_i)$ with $L_i' \in M(K)$, $R_i \in \text{REG}$ and h_i is an arbitrary homomorphism. We may assume that $L_i \subseteq \Sigma_i^\star$ where all the Σ_i are mutually disjoint alphabets. Let $R = \bigcup_i (R_i \cap \Sigma_i^\star)$, $1 \le i \le n$ and let h be the homomorphism defined by $h(\alpha) = h_i(\alpha)$ for each α in $\bigcup_i \Sigma_i$. It is straightforward to show that $\bigcup_i h_i(L_i' \cap R_i) = h((\bigcup_i L_i) \cap R)$, $1 \le i \le n$, which implies that $\hat{\mathcal{S}}(K) \subseteq \Theta\Delta\{\bigcup_i L_i' \mid L_i' \in M(K); 1 \le i \le n\} = \Theta\Delta S(K)$.

<u>$X = P$</u> (and similar for $X = F$): From 4.2 and 2.6 we obtain $\hat{\mathcal{P}}(K) = \text{REG}/\Delta\Phi(K)/\text{FIN}$. Applying 2.3(2) twice yields: $\hat{\mathcal{P}}(K) \subseteq \text{REG}/\Theta\Delta((\Phi(K) \cup \text{FIN})/\Phi(\text{FIN})) = \text{REG}/\Delta(\Phi(K)/\text{FIN}) \subseteq \Theta\Delta((\text{REG} \cup \text{FIN})/\Phi(\Phi(K)/\text{FIN})) = \Theta\Delta(\text{REG}/K/\text{FIN}) = \Theta\Delta P(K)$.

□

Other Canonical Forms like $\hat{\mathcal{F}}(K) = \hat{\mathcal{M}}(K/\text{REG})$ and $\hat{\mathcal{S}}(K) = \hat{\mathcal{M}}(K/\text{ALPHA})$ (cf. [19, 20]) for σ-simple K, directly follow from 4.3.

In case $X = C$ it is possible to establish a Canonical Form like $\hat{\mathcal{C}}(K) = \Pi C(K) = \Theta\Delta C(K)$ (cf. 4.3) provided K is σ-simple and $L\alpha$ is in K for each L in K and each symbol α not occurring in any word of L. In order to avoid further restrictions on K we will however follow another approach.

A substitution $\tau : V \to K$ is called *marked* if $\tau(\alpha) \subseteq \Delta^\star\alpha$ for each α in V, and the alphabets Δ and V are disjoint. Let $K_1\underline{/}K_2$ denote the family obtained by applying marked K_1-substitutions on languages from K_2, i.e., $K_1\underline{/}K_2 = \{\tau(L) \mid L \in K_2; \tau \text{ is a marked } K_1\text{-substitution}\}$; cf. [19]. We now redefine the C-extension by $C(K) = \mathrm{REG}/(K\underline{/}\mathrm{STAR})$; cf. 2.6. Notice that 4.1(2) remains valid. In the sequel we only consider this redefined extension, for we which we prove

Theorem 4.4. (Canonical Form Theorem for Full C-AFL's.)
$\hat{\mathcal{C}} = \Pi C(K) = \Theta\Delta C(K)$.

Proof. Since $\hat{\mathcal{C}}(K)$ is closed under Π and C, we have $\Pi C(K) \subseteq \hat{\mathcal{C}}(K)$. In order to establish the converse inclusion it suffices to show that $\Pi C(K)$ is a full C-AFL containig K. Obviously, $\Pi C(K)$ is a prequasoid and it is easy to prove that $K \subseteq \Pi C(K)$. Thus it remains to show that $\Pi C(K)$ is closed under C. The proof is based on 2.2, 2.3(2) and the following inclusions (i) $\Delta(K)\underline{/}\mathrm{STAR} \subseteq \Theta\Delta(K\underline{/}\mathrm{STAR})$, (ii) $(K\underline{/}\mathrm{STAR})\underline{/}\mathrm{STAR} \subseteq \Delta\Phi(K\underline{/}\mathrm{STAR})$.

Let $L \subseteq \Sigma^\star$ be in K, $R \in \mathrm{REG}$ and let a, b and c be symbols not in Σ. Define a homomorphism h and a finite substitution f by $h(b) = a$, $h(\alpha) = \alpha$ for $\alpha \in \Sigma$, and $f(c) = \{a, ab\}$, $f(\alpha) = \{\alpha\}$ for $\alpha \in \Sigma$. It is straightforward to prove that $((L \cap R)a)^\star = h((Lb)^\star \cap (Rb)^\star)$ and $((La)^\star b)^\star = f((Lc)^\star) \cap (\Sigma \cup \{a, b\})^\star b$, which establish (i) and (ii), respectively.

Applying respectively 2.2, (i), 2.3(2), (ii) and 2.3(2) yields

$$\begin{aligned} C\Pi C(K) &= \mathrm{REG}/\Delta\Phi C(K)\underline{/}\mathrm{STAR} \subseteq \mathrm{REG}/\Theta\Delta(\Phi C(K)\underline{/}\mathrm{STAR} \subseteq \\ &\Theta\Delta(\mathrm{REG}/\Phi C(K)\underline{/}\mathrm{STAR}) = \Pi(\mathrm{REG}/(K\underline{/}\mathrm{STAR})\underline{/}\mathrm{STAR}) \subseteq \\ &\Pi(\mathrm{REG}/\Delta\Phi(K\underline{/}\mathrm{STAR})) \subseteq \Pi(\mathrm{REG}/\Phi(K\underline{/}\mathrm{STAR})) = \Pi C(K), \end{aligned}$$

i.e., $\Pi C(K)$ is C-closed and hence $\hat{\mathcal{C}}(K) = \Pi C(K)$. Finally, we have

$$\Pi C(K) = \Theta\Delta(\mathrm{FIN}/\mathrm{REG}/(K\underline{/}\mathrm{STAR})) = \Theta\Delta C(K).$$ □

From 4.2–4.4 we obtain

Corollary 4.5. *Let X be a symbol from $\{M, S, P, C, F\}$. Then*
(1) *each full X-AFL contains the family $X(\mathrm{FIN})$;*
(2) $X(\mathrm{FIN}) = \mathrm{REG}$ *is the smallest full X-AFL.* □

By $\underline{X\text{-AFL}}$ we will again denote the class of all full X-AFL's over Σ_ω (for $X \in \{M, C, S, P, F\}$).

Theorem 4.6.
(1) $\underline{M\text{-AFL}} \supset \underline{S\text{-AFL}} \supset \underline{P\text{-AFL}} \supset \underline{F\text{-AFL}} \supset \underline{R\text{-AFL}}$;
(2) $\underline{M\text{-AFL}} \supset \underline{C\text{-AFL}} \supset \underline{F\text{-AFL}}$;
(3) $\underline{C\text{-AFL}}$ *is incomparable with* $\underline{S\text{-AFL}}$ *and* $\underline{P\text{-AFL}}$.

Proof. We will establish the existence of a full C-AFL K_0 which is not closed under union. Together with well-known results (cf. [9, 10]) this implies 4.6(1)–(3).

Let K_1 and K_2 be incomparable full AFL's. (The existence of K_1 and K_2 is guaranteed by [13].) Let $L_1 \subseteq \Sigma_1^\star$ be in $K_1 - K_2$ and let $L_2 \subseteq \Sigma_2^\star$ be in $K_2 - K_1$ such that $\Sigma_1 \cap \Sigma_2 = \varnothing$. For $i = 1,\ 2$, define homomorphisms h_i on $\Sigma_1 \cup \Sigma_2$ by $h_i(\alpha) = \alpha$ if $\alpha \in \Sigma_i$ and $h_i(\alpha) = \lambda$ if $\alpha \notin \Sigma_i$.

We define $K_0 = K_1 \cup K_2$, which is a full C-AFL [3]. Suppose K_0 is closed under union. Then according to [9, 10], K_0 is also closed under concatenation and, consequently, $L_1L_2 \in K_1$ or $L_1L_2 \in K_2$. But then we have $h_2(L_1L_2) = L_2 \in K_1$ or $h_1(L_1L_2) = L_1 \in K_2$, respectively, contradicting the choice of L_1 and L_2. □

5 Substituting Families into Families

In this section we apply Canonical Forms of Sections 3 and 4 to the family obtained by substituting K_1-languages into languages from K_2 (5.1–5.3). Then we discuss a generalization of 4.2–4.4 and 5.2 (5.4–5.5) and finally, we consider an example of the application of Canonical Forms and related results in proving a certain family to be a full X-AFL (5.7).

We first consider substituting a prequasoid into a prequasoid.

Lemma 5.1. *If K_1 is σ-simple and if K_2 is nontrivial, then*

$$\Pi(K_1/K_2) \subseteq \Pi(K_1)/\Delta\Phi(K_2) = \Pi(K_1)/\Pi(K_2) = \Pi(K_1/\mathrm{FIN}/K_2),$$

whereas equality holds whenever, either K_1 is closed under substitution into FIN *(i.e., under union and concatenation), or K_2 is closed under finite substitution.*

Proof. An application of 2.2 and 2,3(1) yields $\Pi(K_1/K_2) = \Theta\Delta(\Phi(K_1)/K_2) \subseteq \mathrm{ONE}/\Delta\Phi(K_1)/\Delta\Phi(K_2) = \Pi(K_1)/\Delta\Phi(K_2) \subseteq \Pi(K_1)/\Pi(K_2)$. By 2.2 and 2.3(2) we have

$$\mathrm{ONE}/\Delta\Phi(K_1)/\mathrm{ONE}/\Delta\Phi(K_2) \subseteq \Theta\Delta(\Phi(K_1)/\Phi(\mathrm{ONE}))/\Delta\Phi(K_2) =$$
$$\Theta\Delta(\Phi(K_1)/\mathrm{FIN})/\Delta\Phi(K_2) \subseteq \Theta\Delta(\Phi(K_1))/\mathrm{FIN}/\Phi(K_2) = \Pi(K_1/\mathrm{FIN}/K_2).$$

Using the former inclusion we just established with FIN/K_2 instead of K_2, we obtain $\Pi(K_1/\mathrm{FIN}/K_2) \subseteq \Pi(K_1)/\Delta\Phi(K_2)$. □

From 5.1 it directly follows that substitutiong a prequasoid into a prequasoid yields a prequasoid. A similar well-known conclusion for full trio, full semi-AFL, full AFL [9, 11] and also for full Kleene-AFL and full pseudo-AFL can be inferred immediately from

Theorem 5.2. *Let K_1 and K_2 be σ-simple families. If K_2 is also closed under finite substitution, then*

(1) $\hat{\mathcal{C}}(K_1/K_2) = \hat{\mathcal{M}}(K_1)/\Delta(K_2\underline{/}\mathrm{STAR})$,

(2) $\hat{\mathcal{M}}(K_1/K_2) = \hat{\mathcal{M}}(K_1)/\Delta(K_2)$,

(3) $\hat{\mathcal{S}}(K_1/K_2) = \hat{\mathcal{M}}(K_1)/\Delta(K_2)/\mathrm{ALPHA}$,

(4) $\hat{\mathcal{P}}(K_1/K_2) = \hat{\mathcal{M}}(K_1)/\Delta(K_2)/\mathrm{FIN}$,

(5) $\hat{\mathcal{F}}(K_1/K_2) = \hat{\mathcal{M}}(K_1)/\Delta(K_2)/\mathrm{REG}$.

Proof. (1) By 4.4 and 5.1 we have $\hat{\mathcal{C}}(K_1/K_2) = \Pi(\mathrm{REG}/K_1/K_2\underline{/}\mathrm{STAR}) = \Pi(\mathrm{REG}/K_1)/\Delta\Phi(K_2\underline{/}\mathrm{STAR})$, and hence by 4.3, $\hat{\mathcal{C}}(K_1/K_2) = \hat{\mathcal{M}}(K_1)/\Delta(K_2\underline{/}\mathrm{STAR})$.

(2) According to 4.2 and 5.1 we obtain $\hat{\mathcal{M}}(K_1/K_2) = \mathrm{REG}/\Pi(K_1/K_2) = \mathrm{REG}/\Pi(K_1)/\Delta\Phi(K_2) = \hat{\mathcal{M}}(K_1)/\Delta(K_2)$.

Together with 2.6 and 4.3 this implies (3)–(5). □

In the remaining cases we even obtain

Theorem 5.3. *Let X be a symbol from $\{R, A, H_1^\star, H\}$. Then*
(1) $\hat{\mathcal{X}}(K_1 \cup K_2) = X(\Pi(K_1) \cup \Pi(K_2)) = X(\Pi(K_1)/\Pi(K_2)) = X(\Pi(K_2)/\Pi(K_1))$.
(2) $\hat{\mathcal{X}}(K_1/K_2) = \hat{\mathcal{X}}(K_2/K_1) = \hat{\mathcal{X}}(K_1 \cup K_2)$ *provided both K_1 and K_2 are σ-simple.*

Proof. (1) For reasons of symmetry it suffices to show the former two equalities. By 3.4(1) and [3] we have $\hat{\mathcal{X}}(K_1 \cup K_2) = X\Pi(K_1 \cup K_2) = X(\Pi(K_1) \cup \Pi(K_2))$. From 3.6 and 3.3(2) we obtain that each full X-AFL is a substitution-closed prequasoid (for $X \in \{R, A, H_1^\star, H\}$), which implies $\Pi(K_1)/\Pi(K_2) \subseteq \hat{\mathcal{X}}(K_1 \cup K_2)$. Since each prequasoid includes SYMBOL, we also have $K_1 \cup K_2 \subseteq \Pi(K_1)/\Pi(K_2) \subseteq \hat{\mathcal{X}}(K_1 \cup K_2)$. Applying 3.4(1) and 5.1 (with $\Pi(K_i)$ instead of K_i, $i = 1,\ 2$) yields $\hat{\mathcal{X}}(K_1 \cup K_2) = X\Pi(\Pi(K_1)/\Pi(K_2)) \subseteq X(\Pi(K_1)/\Pi(K_2)) \subseteq \hat{\mathcal{X}}(K_1 \cup K_2)$.
(2) If both K_1 and K_2 are σ-simple, then $K_1 \cup K_2 \subseteq K_1/K_2 \subseteq \hat{\mathcal{X}}(K_1 \cup K_2)$. Hence $\hat{\mathcal{X}}(K_1 \cup K_2) = \hat{\mathcal{X}}(K_1/K_2)$. □

We may also consider 4.2–4.4 and 5.2 as particular instances of more general results (cf. 5.4 and 5.5 below). We call a family K a *full* $[K', K'']$*-structure* if K is a prequasoid which is closed under K'-substitution and under substitution into K'', i.e., $K'/K/K'' \subseteq K$; cf. [19]. If $K'' = \mathrm{STAR}$, we demand, as in Section 4, that K is closed under marked substitution into STAR, i.e., $K'/K\underline{/}\mathrm{STAR}$. Let $[K', K''](K)$ denote the smallest full $[K', K'']$-structure containing K.

Theorem 5.4. *Let K' be a prequasoid and let K'' be either a prequasoid or equal to* SYMBOL *or* ALPHA. *Then*
(1) $[K', \mathrm{STAR}](K) = \Pi(K'_\infty/(K\underline{/}\mathrm{STAR})) = \Theta\Delta(K'_\infty/(K\underline{/}\mathrm{STAR}))$;
(2) $[K', K''](K) = K'_\infty/\Pi(K)/K''_\infty = K'_\infty/\Delta\Phi(K)/K''_\infty = \Pi(K'_\infty/K/K''_\infty) = \Theta\Delta(K'_\infty/K/K''_\infty)$, provided K is σ-simple. □

The proof of 5.4 consists of a straightforward modification of the arguments used in establishing 4.2–4.4 and it will therefore be omitted. Note that SYMBOL and ALPHA are substitution closed, i.e., $K''_\infty \subseteq K''$.

From 5.1 and 5.4 we can infer in a way similar to 5.2

Theorem 5.5. *Let K' be a prequasoid and let K'' be either a prequasoid or equal to* SYMBOL *or* ALPHA. *If K_1 and K_2 are σ-simple families and, if K_2 is closed under finite substitution, then*
(1) $[K', \mathrm{STAR}](K_1/K_2) = \Pi(K'_\infty/K_1/K_2\underline{/}\mathrm{STAR}) = \Theta\Delta(K'_\infty/K_1/K_2\underline{/}\mathrm{STAR})$;
(2) $[K', K''](K_1/K_2) = K'_\infty/\Pi(K_1/K_2)/K''_\infty = K'_\infty/\Pi(K_1)/\Delta(K_2)/K''_\infty$. □

Taking K' (and K'') equal to a quasoid instead of a prequasoid implies that K'_∞ (and K''_∞) is a full R-AFL (cf. 2.4) and that 5.4 and 5.5 yield full (pseudo-)AFL. In particular we have $\hat{\mathcal{P}}(K) = [\text{REG}, \text{FIN}](K)$ and $\hat{\mathcal{F}}(K) = [\text{REG}, \text{REG}](K)$, whereas full M-AFL, full S-AFL and full C-AFL correspond to the cases $K' = \text{REG}$, $K'' = \text{SYMBOL}$, ALPHA and STAR, respectively (provided we apply a marked substitution in the last case).

When we take $K' = \text{FIN}$ instead of REG, we can obtain (1) prequasoids ($K'' =$ SYMBOL), (2) prequasoids closed under union ($K'' =$ ALPHA), (3) prequasoids closed under concatenation [and union] ($K'' =$ FIN), (4) quasoids closed under Kleene $\star$ ($K'' =$ STAR), and (5) quasoids closed under Kleene $\star$, union and/or concatenation ($K'' =$ REG). It is straightforward to show that results similar to 4.2–4.6 and 5.2 also hold for these structures.

Canonical Forms and theorems like 5.2–5.5 could serve as a useful tool in showing certain families to be a (particular kind of) full AFL or a full $[K', K'']$-structure. We conclude this paper with an application.

A language L is called *hyper-sentential* [27] over a family K, if there exist (1) an alphabet V, (2) an initial language $L_0 \subseteq V^\star$ in K, (3) a finite set U of K-substitutions over V, such that $L = U^\star(L_0)$. The *hyper-sentential extension* $\$(K)$ of K consists of all languages hyper-sentential over K. Clearly, $L \cap \Sigma^\star$ is in $H(K)$ when L is in $\$(K)$. The hyper-algebraic extension $H(K)$ has been characterized in terms of $\$(K)$ by Van Leeuwen & Wood [27].

Theorem 5.6. *Let K be a family closed under isomorphism and under intersection with $\Sigma^\star$ for each finite alphabet Σ. If* ONE $\subseteq K$*, then* $H(K) = (\text{SYMBOL} \cup \{\lambda\})/\$(K)$. □

In the finite case even a stronger result holds: ETOL = SYMBOL/TOL and EOL = SYMBOL/OL (cf. [5, 6]; note that $\$(\text{FIN}) = \text{TOL}$). A SYMBOL- [or a SYMBOL $\cup \{\lambda\}$-] substitution is sometimes referred to as a [weak] coding.

We show how to obtain full AFL's from $\$(K)$.

Theorem 5.7. *Let K be a full trio. If K_0 is a prequasoid, then $K/\$(K_0)$ is a* $[\text{REG}, \hat{\mathcal{H}}(K_0)]$*-structure and consequently, it is a full AFL.*

Proof.

$[\text{REG}, \hat{\mathcal{H}}(K_0)](K/\$(K_0)) =$	(K is a full trio)
$[\text{REG}, \hat{\mathcal{H}}(K_0)](K/(\text{SYMBOL} \cup \{\lambda\})/\$(K_0) =$	(5.6)
$[\text{REG}, \hat{\mathcal{H}}(K_0)](K/H(K_0)) =$	(5.5(2), 3.4)
$\text{REG}/K/H(K_0)/H(K_0) =$	(K is a full trio, 3.4)
$K/H(K_0) =$	(5.6)
$K/(\text{SYMBOL} \cup \{\lambda\})/\$(K_0) = K/\$(K_0)$	(K is a full trio),

i.e., $K/\$(K_0)$ is a $[\text{REG}, \hat{\mathcal{H}}(K_0)]$-structure. □

If K is contained in $H(K_0)$, then clearly $K/\$(K_0) = H(K_0)$ and consequently it is a full H-AFL. Thus 5.7 is only interesting if K is not included in $H(K_0)$. When $K_0 = \text{FIN}$, 5.7 yields: K/TOL is a full AFL provided K is a full trio, which was originally established by Salomaa [23, 4]. Using EOL = SYMBOL/OL [5] one can prove in a way similar to 5.7 that K/OL is a full AFL whenever K is a full trio [23, 4].

Acknowledgements

I am grateful to Joost Engelfriet and Leo Verbeek for stimulating discussions and for their critical remarks on an earlier version of this paper.

References

1. A.V. Aho, Indexed grammars – An extension of context-free grammars, *J. Assoc. Comp. Mach.* **15** (1968) 647–671.
2. P.R.J. Asveld, Controlled iteration grammars and full hyper-AFL's, *Inform. Contr.* **34** (1977) 248–269.
3. P.R.J. Asveld, Incomparable elements in algebraic lattices with an application to AFL-theory, TW-memorandum no. 202, Twente University of Technology, Enschede, The Netherlands.
4. P.A. Christensen, Hyper-AFL's and ETOL systems, DAIMI PB-35 (1974), University of Aarhus, Denmark; cf. pp. 254–257 in G. Rozenberg & A. Salomaa (eds.), *L Systems* (1974), Springer-Verlag, Berlin.
5. A. Ehrenfeucht & G. Rozenberg, The equality of EOL languages and codings of OL languages, *Inter. J. Comp. Math.* **4** (1974) 95–104.
6. A. Ehrenfeucht & G. Rozenberg, Nonterminals versus homomorphisms in defining languages for some classes of rewriting systems, *Acta Inform.* **3** (1974) 265–283.
7. C.C. Elgot & J.E. Mezei, On relations defined by generalized finite automata, *IBM J. Res. Develop.* **9** (1965) 47–65.
8. J. Engelfriet, Iterating iterated substitution, *Theor. Comp. Sci.* **5** (1977) 85–100.
9. S. Ginsburg, *Algebraic and Automata-Theoretic Properties of Formal Languages* (1975), North-Holland, Amsterdam.
10. S. Ginsburg, S.A. Greibach & J.E. Hopcroft, *Studies in Abstract Families of Languages*, Memoirs of the American Mathematical Society, No. 87 (1969).
11. S. Ginsburg & E.H. Spanier, Substitution in families of languages, *Infrom. Sci.* **2** (1970) 83–100.
12. S. Ginsburg & E.H. Spanier, AFL with the semlinear property, *J. Comp. System Sci.* **5** (1971) 365–396.
13. S. Ginsburg & E.H. Spanier, On incomparable Abstract Families of Languages (AFL), *J. Comp. System Sci.* **9** (1974) 88–108.

14. S.A. Greibach, Full AFL's and nested iterated substitution, *Inform. Contr.* **16** (1970) 7–35.

15. S.A. Greibach, A generalization of Parikh's semilinear theorem, *Discrete Math.* **2** (1972) 347–355.

16. G.T. Herman, Closure properties of some families of languages associated with biological systems, *Inform. Contr.* **24** (1974) 101-121.

17. G.T. Herman & G. Rozenberg, *Developmental Systems and Languages* (1975), North-Holland, Amsterdam.

18. J.E. Hopcroft & J.D. Ullman, *Formal Languages and Their Relation to Automata* (1969), Addison-Wesley, Reading, MA.

19. D.J. Lewis, Closure of families of languages under substitution operators, pp. 100–108 in Second Annual ACM symposium on Theory of Computing (1970).

20. M. Nivat, Opérateurs sur les familles de langages, Rapport de Recherche no. 106 (1975), IRIA, Rocquencourt, France.

21. G. Rozenberg, Extensions of tabled OL-systems and languages, *Intern. J. Comp. Inform. Sci.* **2** (1973) 311–336.

22. A. Salomaa, *Formal Languages* (1973), Academic Press, New York.

23. A. Salomaa, Macros, iterated substitution and Lindenmayer-AFL's, DAIMI PB-18 (1974), University of Aarhus, Denmark; cf. pp. 250-253 in A. Salomaa & G. Rozenberg (eds.), *L systems* (1974), Springer-Verlag, Berlin.

24. J. van Leeuwen, F-iteration languages, Memorandum (1973), University of California, Berkeley, CA.

25. J. van Leeuwen, A generalization of Parikh's theorem in formal language theory, pp. 17-26 in J. Loeckx (ed.), *Automata, Languages and Programming, 2nd Colloquium* (1974), Springer-Verlag, Berlin.

26. J. van Leeuwen, Effective constructions in well-partially-ordered monoids, TR no. 203 (1976), Pennsylvania State University.

27. J. van Leeuwen & D. Wood, A decomposition theorem for hyper-algebraic extensions of language families, *Theor. Comp. Sci.* **1** (1976) 199–214.

Notes

The original typescript of this report consisted of 34 pages; the present LaTeX version reduced this number to 20.

One of the original references has been published (in slightly different form) as

3. P.R.J. Asveld, An algebraic approach to incomparable families of formal languages, in: G. Rozenberg & A. Salomaa (eds.), *Lindenmayer Systems — Impacts on Theoretical Computer Science, Computer Graphics, and Developmental Biology* (1992), Springer-Verlag, Berlin, etc., pp. 455–475.