

Parallel Hyperedge Replacement Grammars

Graham Campbell^{a,1}

^a*School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, United Kingdom*

Abstract

In 2018, it was shown that all finitely generated virtually Abelian groups have multiple context-free word problems, and it is still an open problem as to where to precisely place the word problems of hyperbolic groups in the formal language hierarchy. Motivated by this, we introduce a new language class, the parallel hyperedge replacement string languages, containing all multiple context-free and ETOL languages. We show that parallel hyperedge replacement grammars can be “synchronised”, which allows us to establish many useful formal language closure results relating to both the hypergraph and string languages generated by various families of parallel hyperedge replacement grammars, laying the foundations for future work in this area.

Keywords: Hyperedge Replacement, Lindenmayer Systems, EOL Languages, ETOL Languages, Multiple Context-Free Languages, Word Problems

1. Introduction

Lindenmayer systems (L systems) originated in the late 1960s to model the development of multicellular organisms [1, 2, 3], and since then, the string generational power of various families of L system has been extensively studied [4, 5]. The family of extended table zero-interaction Lindenmayer (ETOL) string languages is particularly important due to its closure properties and equivalence to many other families of L systems. On the other side of the picture, hyperedge replacement appeared in the early 1970s [6, 7] as a generalisation of context-free string rewriting for hypergraphs, and it was shown that the string generational power of hyperedge replacement grammars is the same as that of multiple context-free grammars [8, 9, 10].

Looking purely at string language generation, one can view ETOL grammars as a generalisation of context-free grammars with parallel rewriting, and similarly, one can view the class of string generating hyperedge replacement grammars as a generalisation of context-free grammars where intermediate states

Email address: g.j.campbell12@newcastle.ac.uk (Graham Campbell)

¹Supported by a Doctoral Training Grant No. (2281162) from the Engineering and Physical Sciences Research Council (EPSRC) in the UK.

can be hypergraphs, rather than just strings. In this paper, we combine both of these ideas, to explore the string generational power of grammars that can exploit the additional power of both hypergraphs and parallel rewriting.

A major motivation of this work is the study of word problems of finitely generated groups. In general, the word problem is the question that asks if two strings (words) represent the same element in some structure. In the case of groups, this is the equivalent to asking if a given string represents the identity element, since if u, v are strings, then they are equal in a group if and only if uv^{-1} represents the identity in the group. Thus, given a presentation $\langle X \mid R \rangle$ for a group G , the word problem is equivalent to the membership problem for the string language $WP_X(G) = \{w \in (X \cup X^{-1})^* \mid w =_G 1_G\}$. Recall that the Cayley graph of a group with respect to a fixed presentation has the group elements as vertices and an edge between two vertices $g, h \in G$ whenever there is a generator $s \in X$ such that $gs = h$. Viewing things geometrically, the word problem of a group can be identified with the set of loops based at the identity in the Cayley graph. A sketch of a finite portion of the Cayley graphs of \mathbb{Z}^2 and F_2 with their usual generating sets is provided in Figure 1.

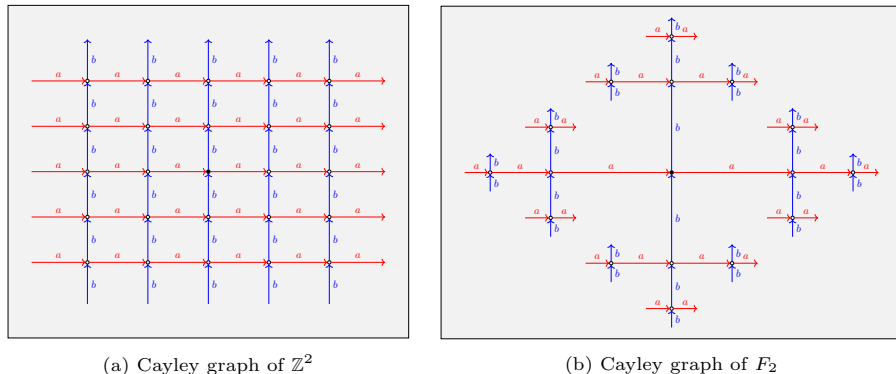


Figure 1: Example Cayley graphs

A natural question to ask is how hard the word problem is, in general, and for specific families of groups. Unsurprisingly, both the universal word problem and the word problem are undecidable in general, even for finite presentations [11]. It is well known that a presentation defines a finite group if and only if it admits a regular word problem [12], and defines a finitely generated virtually free group if and only if it admits a deterministic context-free word problem if and only if it admits a context-free word problem [13]. The multiple context-free (MCF) languages sit strictly in between the context-free and context-sensitive languages [8]. In 2015, a major breakthrough of Salvati was published, showing that the word problem of \mathbb{Z}^2 is an MCF language [14], and in 2018, Ho extended this result to all finitely generated virtually Abelian groups [15]. This is interesting since the MCF languages are exactly the string languages generated by hyperedge replacement grammars [9, 10]. It remains an open problem as to

which other families of groups admit MCF word problems, however, we do at least know that the fundamental group of a hyperbolic three-manifold does not admit an MCF word problem [16].

There are of course, lots of other well-behaved language classes sitting in between the context-free and context-sensitive classes, such as the indexed languages [17] or the subclass of ETOL languages [4]. It is not known if there are any groups with indexed word problems, other than the virtually free groups, but it is known that a particular subclass of the indexed languages, not contained in ETOL, only contains word problems of virtually free groups [18]. We also do not know if any hyperbolic groups have ETOL word problems [19] (other than the virtually free groups), such as the fundamental group of the double torus. It is conjectured that every ETOL group language is admitted by a virtually free group [19]. Figure 2 shows the (group) language hierarchy, where necessarily strict inclusion uses a solid line, and \mathcal{GP} denotes the class of all group languages (the class of word problems of all finitely generated groups). Descriptions of the other language classes can be found in Section 2.

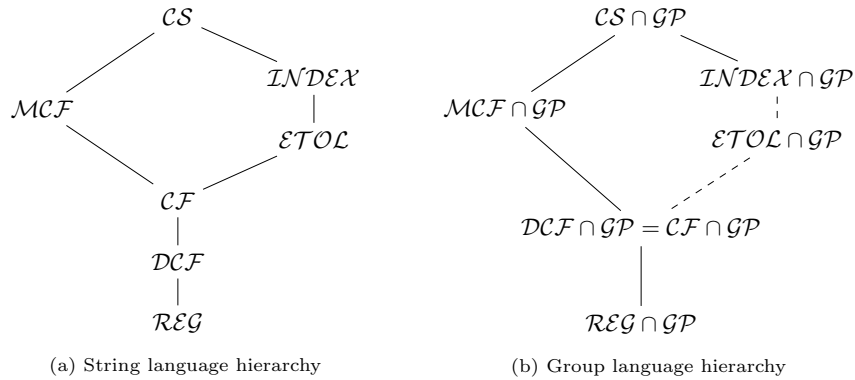


Figure 2: Previously known formal language hierarchies

In this paper, we define and study a new string language class, combining ideas from both ETOL and hyperedge replacement grammars. We call our new class the parallel hyperedge replacement string (PHRS) languages, and show that the class strictly contains both the classes of MCF and ETOL languages, that it is a hyper-algebraically closed super abstract family of languages, and that PHRS group languages are closed under free product. While parallel hyperedge replacement has been considered before, most notably by Habel and Kreowski (separately) [20, 21, 22], the work is not extensive and does not consider repetition-freeness, rational control, number of tables, or string generational power. Figure 3 summarises how the PHRS and repetition-free PHRS languages fit into the string language hierarchy and also how we conjecture the hierarchy collapses when we restrict to group languages. At the end of the paper, we provide much more detailed diagrams, showing both hypergraph and string language hierarchies. Our long term goal is to place the word problem for

as many hyperbolic groups as possible in the PHRS class. Knowledge of (geometric) group theory and word problems is not required to read and understand this paper - it is purely motivational!

This paper is based on, and replaces, the author's TERMGRAPH 2020 workshop proceedings paper [23], with significant revisions, improvements, and new results. Most notably, in the workshop paper, the number of tables was not considered at all. Consideration of the number of tables was a major undertaking, leading to the strengthening of many of the existing results, many of which requiring entirely new proofs, also leading to the creation of Subsection 3.5 looking at synchronisation of grammars and substitution closure. Moreover, we were able to additionally show that the PHRS languages are hyper-algebraically closed. The second most notable addition is showing that the PHRS membership problem is decidable (Subsection 3.6) which provides us with an algorithm for computing membership of PHRS languages (Subsection 4.4). Finally, without the space limitations of a workshop paper, we were able to provide a more self-contained preliminaries section (Section 2), some additional examples in the main part of the paper, and a more detailed handling of WPHRS languages.

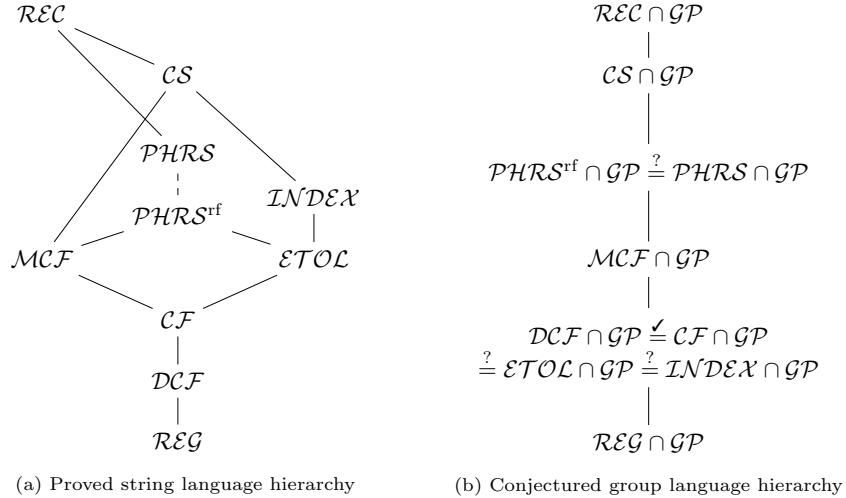


Figure 3: New formal language hierarchies

2. Preliminaries

Before we get started, we must give some preliminary notations and known results. By \mathbb{N} we mean the natural numbers including zero, by \underline{n} we mean $\{1, \dots, n\}$, and \oplus denotes relational override. Given a set S , denote by $\mathcal{P}(S)$ the set of all subsets of S and $\mathcal{P}_0(S) = \mathcal{P}(S) \setminus \{\emptyset\}$. A (finite) sequence on a set S is a function $\sigma : \underline{n} \rightarrow S$, we view strings as sequences on an alphabet. We denote by S^* all sequences on a set S , and denote subsequence (scattered subsequence) by \sqsubseteq (\sqsubseteq_{sc}).

2.1. String Languages

In this paper, all alphabets will be finite. We can equivalently view A^* as the free monoid on A . A string language $L \subseteq A^*$ is simply a set of strings over some alphabet A . It is elementary that monoid homomorphisms $\varphi : A^* \rightarrow B^*$ are totally determined by their values on A . We call φ non-erasing if $\epsilon \notin \varphi(A)$, a coding (letter-to-letter homomorphism) if $\varphi(A) \subseteq B$, and a weak coding if $\varphi(A) \subseteq B \cup \{\epsilon\}$.

By a family of string languages, we mean an isomorphism-closed (closed under renaming of symbols), non-trivial (contains languages other than $\emptyset, \{\epsilon\}$) class of string languages. Given a family of string languages \mathcal{F} and alphabets A, B , an \mathcal{F} -substitution of strings on A is a function $h : A \rightarrow \mathcal{P}_0(B^*)$ such that for each $a \in A$, $h(a) \in \mathcal{F}$. When \mathcal{F} is the family of all finite string languages, we call h a finite substitution of strings. If h is such that for each $a \in A$, $a \in h(a)$, then we call h nested, and if $\epsilon \notin h(A)$, then we call h non-erasing. Given a string $w = x_1 \cdots x_n$ over A , write $h(w)$ for the language $\{w' \in B^* \mid \exists w_1 \in h(x_1), \dots, \exists w_n \in h(x_n), w' = w_1 \cdots w_n\}$, and given a string language L over A , write $h(L)$ for the language $\bigcup_{w \in L} h(w)$.

A family of string languages \mathcal{F} is called a cone if it is closed under homomorphisms, inverse homomorphisms, and intersection with regular languages. That is, if $L \subseteq A^*$ is an \mathcal{F} language, $L' \subseteq B^*$ is an \mathcal{F} language, $K \subseteq A^*$ is a regular language, and $\varphi : A^* \rightarrow B^*$ is a homomorphism, then $\varphi(L) \in \mathcal{F}$, $\varphi^{-1}(L') \in \mathcal{F}$, and $L \cap K \in \mathcal{F}$. The rational operations are union ($L_1 \cup L_2$), concatenation ($L_1 L_2 = \{w_1 w_2 \mid w_1 \in L_1, w_2 \in L_2\}$), and Kleene plus ($L^+ = \bigcup_{i \geq 1} \{w^i \mid w \in L\}$). \mathcal{F} is called a full abstract family of languages (full AFL) if it is both a cone and is closed under the rational operations. That is, if $L_1, L_2 \subseteq A^*$ are \mathcal{F} languages, then $L_1 \cup L_2 \in \mathcal{F}$, $L_1 L_2 \in \mathcal{F}$, and $L_1^+ \in \mathcal{F}$. Call \mathcal{F} substitution closed if for any \mathcal{F} language $L \subseteq A^*$ and \mathcal{F} -substitution h of strings on A , $h(L) \in \mathcal{F}$. Call \mathcal{F} (nested) iterated substitution closed if for any \mathcal{F} language $L \subseteq A^*$ and a (nested) \mathcal{F} -substitution h of strings on A , $\bigcup_{n \in \mathbb{N}} h^n(L) \in \mathcal{F}$, where $h^n(L) = h(\cdots h(L) \cdots)$. A full AFL is called a super AFL if it is additionally substitution and nested iterated substitution closed.

We denote by \mathcal{REG} , \mathcal{DCF} , \mathcal{CF} , \mathcal{CS} , \mathcal{REC} , \mathcal{RE} the families of regular, deterministic context-free (DCF), context-free (CF), context sensitive, recursive, and recursively enumerable string languages, respectively. One can consult any undergraduate textbook on formal languages and automata such as [24] for an introduction to these language families. These language families enjoy the following hierarchy: $\mathcal{REG} \subsetneq \mathcal{DCF} \subsetneq \mathcal{CF} \subsetneq \mathcal{CS} \subsetneq \mathcal{REC} \subsetneq \mathcal{RE}$.

2.2. MCF Languages

The multiple context-free languages were defined by Seki, Matsumura, Fujii, and Kasami in the early 90s [8] as a conservative extension of the context-free language, showing that the family was a substitution closed full AFL, and closure under nested iterated substitution is also obvious by trivial modification of Král's proof for context-free languages [25] (Theorem 2.1). Various alternative descriptions of the MCF languages have now been established (Theorem 2.2).

The actual definition of MCF grammars will not matter to us, since we will only be interested in the hyperedge replacement characterisation. We denote by \mathcal{MCF}_k and \mathcal{MCF} the order at most k multiple context-free languages and the multiple context-free languages, respectively. These families enjoy the following hierarchy: $\mathcal{CF} = \mathcal{MCF}_1 \subsetneq \cdots \subsetneq \mathcal{MCF}_k \subsetneq \mathcal{MCF}_{k+1} \subsetneq \cdots \subsetneq \mathcal{MCF} \subsetneq \mathcal{CS}$.

Theorem 2.1 (MCF Closure Properties [25, 8]). For $k \geq 1$, \mathcal{MCF}_k and \mathcal{MCF} are super AFLs.

Theorem 2.2 (MCF Equivalence). The following families of languages are equivalent, for any $k \geq 1$:

1. $\mathcal{HRS}_{2k} = \mathcal{HRS}_{2k+1} = \mathcal{HRS}_{2k}^{\text{rf}} = \mathcal{HRS}_{2k+1}^{\text{rf}}$: the string languages of (repetition-free) hyperedge replacement grammars of order $2k$ or $2k + 1$ (see Subsection 2.4);
2. $\text{OUT}(\mathcal{DTWT}_k)$: the output languages of deterministic tree walking transducers of crossing number at most k (see [26]);
3. \mathcal{LCFR}_k : the string languages of linear context-free rewriting systems of rank at most k (see [27]);
4. \mathcal{MCF}_k : the languages of k -multiple context-free grammars (see [8]);
5. \mathcal{RTSA}_k : the languages of k -restricted tree stack automata (see [28]).

Proof. $\mathcal{HRS}_k = \mathcal{HRS}_k^{\text{rf}}$ for all $k \geq 2$ is due to Theorem 2.12 and $\mathcal{HRS}_{2k+1}^{\text{rf}} \subseteq \text{OUT}(\mathcal{DTWT}_k) \subseteq \mathcal{HRS}_{2k}$ for all $k \geq 1$ is due to Engelfriet and Heyker (1991) [9], which gives us the all the equalities in (1) and (1) = (2). (2) = (3) is due to Weir (1992) [10], (3) = (4) is due to Seki, Matsumura, Fujii, and Kasami (1991) [8], and (4) = (5) is due to Denkinger (2016) [28]. \square

Call a set $S \subseteq \mathbb{N}^d$ linear if it is of the form $\{p + a_1 p_1 + \cdots + a_n p_n \mid a_1, \dots, a_n \in \mathbb{N}\}$ for some fixed $p, p_1, \dots, p_n \in \mathbb{N}^d$. Call S semilinear if it is a finite union of linear sets. The Parikh vector of a string w over an ordered alphabet $A = (a_1, \dots, a_d)$ is $\psi_A(w) = (|w|_{a_1}, \dots, |w|_{a_d})$ where $|w|_{a_i}$ counts the number of occurrences of a_i in w . A string language L is called linear (semilinear) if the set of Parikh vectors $\psi_A(L)$ is linear (semilinear). Two string languages $L_1, L_2 \subseteq A^*$ are called letter-equivalent if $\psi_A(L_1) = \psi_A(L_2)$. In 1961, Parikh circulated the following famous results, re-published in 1966, showing that all context-free languages are semilinear, and that being semilinear amounts to being letter-equivalent to a regular language:

Theorem 2.3 (Parikh's Theorem I [29]). If $L \in \mathcal{CF}$, then L is semilinear.

Theorem 2.4 (Parikh's Theorem II [29]). A string language L is semilinear if and only if it is letter-equivalent to a regular language.

Corollary 2.5. If $L \subseteq \{a\}^*$ is semilinear, then $L \in \mathcal{REG}$.

In 1991, Seki et al. showed that all MCF languages are semilinear:

Theorem 2.6 (MCF Semilinear [8]). If $L \in \mathcal{MCF}$, then L is semilinear.

2.3. ETOL Languages

Lindenmayer systems (L systems) were introduced in 1968 by Aristid Lindenmayer [1, 2, 3]. We direct the reader to [4] for a comprehensive introduction to the topic. We are interested in the family of string languages called the ETOL languages. These languages are described by a specific type of L system.

Given a set H of substitutions of strings on Σ and a language $L \subseteq \Sigma^*$, we define $\text{ITER}_H(L) = \bigcup_{n \in \mathbb{N}} \{h_n(\dots h_1(L)\dots) \mid h_1, \dots, h_n \in H\}$. A table over Σ is a left-total finite binary relation $T \subseteq \Sigma \times \Sigma^*$, and can be associated to a finite substitution h_T where $h_T(a) = \{b \in \Sigma^* \mid (a, b) \in T\}$ for each $a \in \Sigma$. An ETOL grammar is a tuple $\mathcal{G} = (\Sigma, A, S, \mathcal{T})$ where Σ is an alphabet, $A \subseteq \Sigma$ is the terminal alphabet, $S \in \Sigma$ is the start symbol and \mathcal{T} is a finite set of tables over Σ . The language generated by \mathcal{G} is $L(\mathcal{G}) = \text{ITER}_{\{h_T \mid T \in \mathcal{T}\}}(\{S\}) \cap A^*$. A language $L \subseteq A^*$ is called an ETOL language if there exists an ETOL grammar \mathcal{G} such that $L(\mathcal{G}) = L$, and is called EOL if the grammar contains only one table. It will be convenient to think of table entries as rules and substitutions as parallel rewriting, and we will take this view in Section 3. We denote the EOL, ETOL, and indexed language families by \mathcal{EOL} , \mathcal{ETOL} , and \mathcal{INDEX} , respectively, and we have the following strict inclusions: $\mathcal{CF} \subsetneq \mathcal{EOL} \subsetneq \mathcal{ETOL} \subsetneq \mathcal{INDEX} \subsetneq \mathcal{CS}$.

Call an ETOL grammar propagating if each table is contained in $\Sigma \times \Sigma^+$ (rather than just $\Sigma \times \Sigma^*$). That is, rules have non-empty right-hand sides. ETOL grammars can be assumed to be propagating:

Theorem 2.7 (Propagating ETOL Generative Power [4]). Given an ETOL grammar \mathcal{G} , one can effectively construct a propagating ETOL grammar \mathcal{G}' with the same number of tables such that $L(\mathcal{G}) \setminus \{\epsilon\} = L(\mathcal{G}')$.

An extremely useful notion in L systems is that of a synchronised system. Call an ETOL grammar $(\Sigma, A, S, \mathcal{T})$ synchronised if any string $w, \in A^*$ and any table $T \in \mathcal{T}$, $h_T(w) \notin A^*$. ETOL grammars can always be synchronised:

Theorem 2.8 (Synchronisation of ETOL Grammars [4]). Given a (propagating) ETOL grammar \mathcal{G} , one can effectively construct a synchronised (propagating) ETOL grammar \mathcal{G}' with the same number of tables such that $L(\mathcal{G}) = L(\mathcal{G}')$.

A family of string languages \mathcal{F} is called hyper-algebraically closed if for any \mathcal{F} language $L \subseteq A^*$ and finite set H of \mathcal{F} -substitutions of strings on A , $\text{ITER}_H(L) \in \mathcal{F}$. Clearly hyper-algebraic closure implies closure under (nested) iterated substitution, viewed as a special case with $|H| = 1$.

Theorem 2.9 (ETOL Closure Properties [30]). \mathcal{ETOL} is a hyper-algebraically closed super AFL. Moreover, it is the smallest hyper-algebraically closed super AFL. \mathcal{EOL} is closed under finite substitution, intersection with regular languages, and rational operations.

For any $n \geq 1$, the order n Dyck language D_n is generated by the context-free grammar with start symbol S , terminal set $\{a_1, a'_1, \dots, a_n, a'_n\}$, and rules $S \rightarrow \epsilon, S \rightarrow SS, S \rightarrow a_1Sa'_1, \dots, S \rightarrow a_nSa'_n$. In 1977, Ehrenfeucht and Rozenberg showed that for all $n \geq 8$, D_n is not an EDTOL language [31]. With not too much work, this gives us the following result:

Theorem 2.10 (MCF and ETOL Incomparable).

1. $K = \{wh(w) \mid w \in D_8\}$ is a 2-MCF language which is not ETOL, where $\bar{\Sigma}$ is a disjoint copy of Σ and $h : \Sigma^* \rightarrow \bar{\Sigma}^*$ is defined by sending each $a \in \Sigma$ to its copy $\bar{a} \in \bar{\Sigma}$.
2. $L = \{a^{2^n} \mid n \geq 0\}$ is an EOL language but not MCF. Moreover, L is not semilinear.

Proof. The first part follows from Theorem 8 of [32], which holds for any language in place of D_8 which is context-free but not EDTOL.

For the second part, it is easy to see that $\mathcal{G} = (\{a\}, \{a\}, a, \{(a, aa)\})$ is an ETOL grammar using one table with $L(\mathcal{G}) = L$. Corollary 2.5 tells us L is semilinear if and only if it is regular, and clearly that language is not regular, so it cannot be semilinear. Theorem 2.6 tells us all MCF languages are semilinear, so it must be the case that L is not MCF. \square

2.4. Hyperedge Replacement

This subsection is mostly based on [20, 33]. By a signature we mean a pair $\mathcal{C} = (\Sigma, \text{type})$ where Σ is some alphabet, called the label set, and $\text{type} : \Sigma \rightarrow \mathbb{N}$ is a typing function which assigns to each label an arity called its type. We usually will assume some arbitrary but fixed signature $\mathcal{C} = (\Sigma, \text{type})$.

A hypergraph is a tuple $H = (V_H, E_H, \text{att}_H, \text{lab}_H, \text{ext}_H)$ where V_H is a finite set of nodes, E_H is a finite set of hyperedges, $\text{att}_H : E_H \rightarrow V_H^*$ is the attachment function, $\text{lab}_H : E_H \rightarrow \Sigma$ is the labelling function, and $\text{ext}_H \in V_H^*$ are the external nodes, such that labelling is compatible with typing ($\text{type} \circ \text{lab}_H = |\cdot| \circ \text{att}_H$). In an abuse of notation, we write $\text{type}(H) = |\text{ext}_H|$ for the type of H , and define $\text{type}_H : E_H \rightarrow \mathbb{N}$ by $\text{type}_H = \text{type} \circ \text{lab}_H$ for the type of a hyperedge. For any hyperedge $e \in E_H$, whenever $m = \text{type}_H(e)$ we call e a type m hyperedge, and call e proper whenever $\text{att}_H(e)$ is injective (contains no repeated nodes). Call H proper if every $e \in E_H$ is proper, repetition-free if ext_H is injective, and well-formed if it is both proper and repetition-free. The class of all hypergraphs (repetition-free hypergraphs) over \mathcal{C} is denoted $\mathcal{H}_{\mathcal{C}}$ ($\mathcal{H}_{\mathcal{C}}^{\text{rf}}$). A hypergraph morphism $g : G \rightarrow H$ between $G, H \in \mathcal{H}_{\mathcal{C}}$ is a pair of functions $(g_V : V_G \rightarrow V_H, g_E : E_G \rightarrow E_H)$ such that $\text{att}_H \circ g_E = g_V^* \circ \text{att}_G$, $\text{lab}_H \circ g_E = \text{lab}_G$, and $g_V \circ \text{ext}_G \sqsubseteq_{\text{sc}} \text{ext}_H$. g is called external node reflecting if $g_V \circ \text{ext}_G = \text{ext}_H$, is called injective (surjective, bijective) whenever both g_V and g_E are injective (surjective, bijective), and is called hyperedge-injective if g_E is injective. We say two hypergraphs $G, H \in \mathcal{H}_{\mathcal{C}}$ are isomorphic ($G \cong H$) there is a bijective external node reflecting hypergraph morphism $G \rightarrow H$.

Given a string $w \in \Sigma^*$ of length n , its string graph is $w^\bullet = (\{v_0, \dots, v_n\}, \{e_1, \dots, e_n\}, \text{att}, \text{lab}, v_0 v_n)$ where $\text{att}(e_i) = v_{i-1} v_i$ and $\text{lab}(e_i) = w(i)$ for all $i \in \underline{n}$ (Figure 4(a)). If $H \cong w^\bullet$ for some $w \in \Sigma^*$, we call H a string graph representing w . We also use the superscript bullet to denote the handle of a label. If $X \in \Sigma$ is of type n , then the handle of X is the hypergraph $X^\bullet = (\{v_1, \dots, v_n\}, \{e\}, \text{att}, \text{lab}, v_1 \cdots v_n)$ where $\text{att}(e) = v_1 \cdots v_n$ and $\text{lab}(e) = X$ (Figure 4(b)). These two definitions coincide for a type 2 label, considered either as a string of length 1 or as a label, so there can be no confusion.

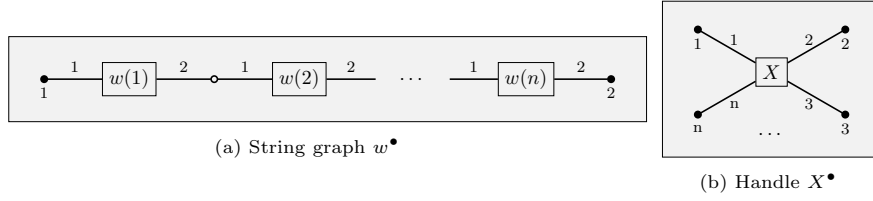


Figure 4: Example hypergraphs

Let $H \in \mathcal{H}_C$ be a hypergraph and $B \subseteq E_H$ be a selection of hyperedges. Then $\sigma : B \rightarrow \mathcal{H}_C$ is called a replacement function if $\text{type} \circ \sigma = \text{type}_H|_B$. The replacement of B in H using σ is denoted by $H[\sigma]$, and is the hypergraph obtained from H by removing B from E_H , disjointly adding the nodes and hyperedges of $\sigma(e)$, for each $e \in B$, and identifying the i -th external node of $\sigma(e)$ with the i -th attachment node of e , for each $e \in B$ and $i \in \text{type}_H(e)$. The external nodes of $H[\sigma]$ remain exactly those of H and all hyperedges keep their original attachments and labels. $H[\sigma]$ exists exactly when $\sigma : B \rightarrow \mathcal{H}_C$ is a replacement function, and is unique up to isomorphism. If $B = \{e_1, \dots, e_n\}$ and $R_i = \sigma(e_i)$ for all $i \in \underline{n}$, then we write $H[e_1/R_1, \dots, e_n/R_n]$ in place of $H[\sigma]$. Figure 5 shows an example replacement.

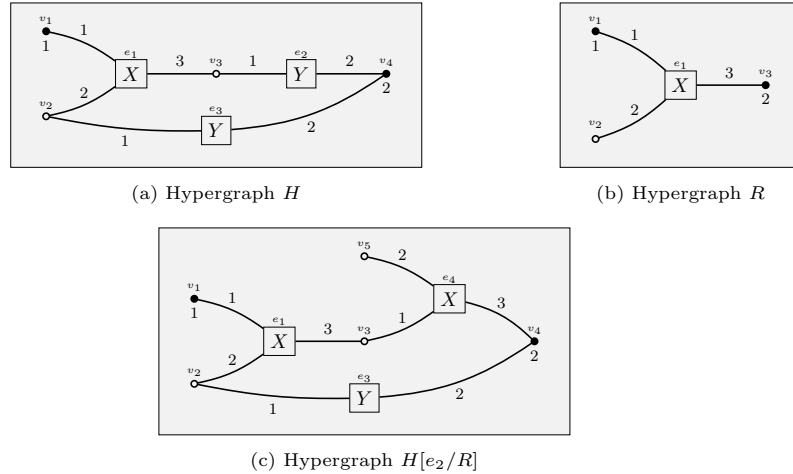


Figure 5: Example hyperedge replacement

Let $N \subseteq \Sigma$ be a set of non-terminals. A type n rule over N is a pair (L, R) with $L \in N$, $R \in \mathcal{H}_C$, and $\text{type}(L) = \text{type}(R) = n$. Call a rule (L, R) repetition-free (proper, well-formed) if R is repetition-free (proper, well-formed). Given a hypergraph $H \in \mathcal{H}_C$ and a set of rules \mathcal{R} , if $e \in E_H$ and $(\text{lab}_H(e), R) \in \mathcal{R}$, then we say that H directly derives $H' \cong H[e/R]$ and write $H \Rightarrow_{\mathcal{R}} H'$. For a given hyperedge e and choice of rule, H' is unique up to isomorphism. Clearly $\Rightarrow_{\mathcal{R}}$ is a binary relation on \mathcal{H}_C . We say $H \in \mathcal{H}_C$ derives H' if there is a sequence

$H \Rightarrow_{\mathcal{R}} H_1 \Rightarrow_{\mathcal{R}} \dots \Rightarrow_{\mathcal{R}} H_k = H'$ for some $k \geq 1$ or $H \cong H'$. We write $H \Rightarrow_{\mathcal{R}}^k H'$ or $H \Rightarrow_{\mathcal{R}}^* H'$. Clearly, (direct) derivations cannot delete nodes, and (direct) derivations made using repetition-free rules cannot merge nodes. We have the following result for repetition-free rules:

Theorem 2.11 (HR Context-Freeness [20]). Let \mathcal{R} be a set of repetition-free rules over N , $H \in \mathcal{H}_{\mathcal{C}}$, $X \in N$, and $k \in \mathbb{N}$. Then there is a derivation $X^{\bullet} \Rightarrow^{k+1} H$ if and only if there is a rule $(X, R) \in \mathcal{R}$ and a mapping $\sigma : \text{lab}_R^{-1}(N) \rightarrow \mathcal{H}_{\mathcal{C}}$ such that $H = R[\sigma]$, $\forall e \in \text{lab}_R^{-1}(N), \text{lab}_R(e)^{\bullet} \Rightarrow^{k(e)} \sigma(e)$, and $\sum_{e \in \text{lab}_R^{-1}(N)} k(e) = k$.

A hyperedge replacement grammar of order k (k -HR grammar) is a tuple $\mathcal{G} = (\mathcal{C}, N, S, \mathcal{R})$ where $\mathcal{C} = (\Sigma, \text{type})$ is a signature, $N \subseteq \Sigma$ is the set of non-terminal labels, $S \in N$ is the start symbol, and \mathcal{R} is a finite set of rules over N , with $\max(\{\text{type}(r) \mid r \in \mathcal{R}\}) \leq k$. We call $\Sigma \setminus N$ the terminal labels and call \mathcal{G} repetition-free (proper, well-formed) if all its rules are repetition-free (proper, well-formed). The language generated by \mathcal{G} is $L(\mathcal{G}) = \{H \in \mathcal{H}_{\mathcal{C}} \mid S^{\bullet} \Rightarrow_{\mathcal{R}}^* H \text{ with } \text{lab}_H^{-1}(N) = \emptyset\} \subseteq \mathcal{H}_{\mathcal{C}}$. $L \subseteq \mathcal{H}_{\mathcal{C}}$ is called a (repetition-free) hyperedge replacement language of order k ((repetition-free) k -HR language) if there is a (repetition-free) k -HR grammar such that $L(\mathcal{G}) = L$. The class of (repetition-free) HR languages is the union of all (repetition-free) k -HR languages for $k \geq 0$. Denote these \mathcal{HR}_k and \mathcal{HR} ($\mathcal{HR}_k^{\text{rf}}$ and \mathcal{HR}^{rf}). All such languages are closed under hypergraph isomorphism and are homogeneous (all hypergraphs have the same type). We define the type of a homogeneous language to be the type of its members.

Theorem 2.12 (Repetition-Free HR Generational Power [9]). Given an HR grammar \mathcal{G} over \mathcal{C} , one can effectively construct a repetition-free HR grammar \mathcal{G}' with $L(\mathcal{G}') = L(\mathcal{G}) \cap \mathcal{H}_{\mathcal{C}}^{\text{rf}}$.

Theorem 2.13 (HR Linear-Growth [20]). Given an infinite HR language L , there exists an infinite sequence of hypergraphs in L , say H_0, H_1, H_2, \dots and constants $c, d \in \mathbb{N}$ with $c + d \geq 1$, such that for all $i \in \mathbb{N}$, $|V_{H_{i+1}}| = |V_{H_i}| + c$ and $|E_{H_{i+1}}| = |E_{H_i}| + d$.

Given a class of hypergraph languages \mathcal{F} , signatures $\mathcal{C} = (\Sigma, \text{type})$ and \mathcal{D} , and $A \subseteq \Sigma$, an \mathcal{F} -substitution of hypergraphs on A is a function $s : A \rightarrow \mathcal{P}_0(\mathcal{H}_{\mathcal{D}})$ such that for each label $X \in A$, $s(X) \in \mathcal{F}$ and $s(X)$ is of type $\text{type}(X)$. When \mathcal{F} is the class of all finite languages of hypergraphs, we call s a finite substitution of hypergraphs. When s is such that for each label $X \in A$, $A^{\bullet} \in s(A)$, then we call s nested. Given a hypergraph H over A with $E_H = \{e_1, \dots, e_n\}$, write $s(H)$ for the language $\{H' \in \mathcal{H}_{\mathcal{D}} \mid \exists R_1 \in s(\text{lab}_H(e_1)), \dots, \exists R_n \in s(\text{lab}_H(e_n)), H' \cong H[e_1/R_1, \dots, e_n/R_n]\}$, and given a hypergraph language L over A , write $s(L)$ for the language $\bigcup_{H \in L} s(H)$. Call \mathcal{F} substitution closed if for any \mathcal{F} language over A and \mathcal{F} -substitution s of hypergraphs on A , $s(L) \in \mathcal{F}$. Call \mathcal{F} (nested) iterated substitution closed if for any \mathcal{F} language over A and (nested) \mathcal{F} -substitution s of hypergraphs on A , $\bigcup_{n \in \mathbb{N}} s^n(L) \in \mathcal{F}$, where $s^n(L) = s(\dots s(L) \dots)$.

Theorem 2.14 (HR Languages Substitution Closed [20]). For $k \geq 0$, the k -HR languages are substitution closed and nested iterated substitution closed.

The partial function $\text{STR} : \mathcal{H}_C \rightarrow \Sigma^*$ sends string graphs to the strings they represent, and is undefined elsewhere. A language $L \subseteq \mathcal{H}_C$ is said to be a string graph language if it only contains string graphs. Given an HR grammar \mathcal{G} that generates a string graph language, we write $\text{STR}(L(\mathcal{G}))$ for the actual string language it generates. A string language $L \subseteq A^*$ is called a (repetition-free) hyperedge replacement string language of order k ((repetition-free) k -HRS language) if there is a (repetition-free) k -HR grammar \mathcal{G} such that \mathcal{G} generates a string graph language and $\text{STR}(L(\mathcal{G})) = L \setminus \{\epsilon\}$. The class of (repetition-free) HRS languages is the union of all (repetition-free) k -HRS languages for $k \geq 2$. Denote these \mathcal{HRS}_k and \mathcal{HRS} ($\mathcal{HRS}_k^{\text{rf}}$ and $\mathcal{HRS}^{\text{rf}}$).

3. PHR Languages

In this section, we will formally define parallel hyperedge replacement and parallel hyperedge replacement grammars, showing that they generalise hyperedge replacement. Our definitions are equivalent to those defined by Habel in Chapter VIII.3 of [20]. After that, we investigate the effects of restricting the maximum number of tables, show that rational control of traces does not change generative power when the number of tables is not restricted, establish some substitution closure results, and finally, show that the emptiness problem for parallel hyperedge replacement grammars is decidable.

3.1. Definitions and Foundations

The most fundamental notion to us is that of a parallel direct derivation, where every hyperedge is necessarily replaced. In order to ensure progress can always be made, we are only interested in sets of rules that are tables:

Definition 3.1 (Table). Given $\mathcal{C} = (\Sigma, \text{type})$, a table T over Σ is a finite set of rules over Σ such that for each $L \in \Sigma$ there is at least one $R \in \mathcal{H}_C$ with $(L, R) \in T$. Call T repetition-free (proper, well-formed) if all its rules are repetition-free (proper, well-formed).

Definition 3.2 (Parallel Direct Derivation). Given $\mathcal{C} = (\Sigma, \text{type})$, $H \in \mathcal{H}_C$ with $E_H = \{e_1, \dots, e_n\}$, and T a table over Σ , if for each $e_i \in E_H$, there is a $R_i \in \mathcal{H}_C$ such that $(\text{lab}_H(e_i), R_i) \in T$, then we say that H parallelly directly derives $H' \cong H[e_1/R_1, \dots, e_n/R_n]$, and write $H \Rightarrow_T H'$.

Direct derivations enjoy a restriction property, and it is easy to see that parallel direct derivations also enjoy this same property. Given hypergraphs G and H , write $G \subseteq H$ if there is an injective hypergraph morphism $G \rightarrow H$ and $G \sqsubseteq H$ if there is a hyperedge-injective hypergraph morphism $G \rightarrow H$.

Lemma 3.3 (PHR Derivation Restriction). Given $\mathcal{C} = (\Sigma, \text{type})$, a table T over Σ , and hypergraphs $H, H' \in \mathcal{H}_C$ such that $H \Rightarrow_T H'$ and $G \subseteq H$, then there is a hypergraph $G' \in \mathcal{H}_C$ such that $G \Rightarrow_T G'$ and $G' \subseteq H'$. Moreover, if T is repetition-free, we can replace both occurrences of \sqsubseteq with \subseteq .

An equivalent way to view a parallel direct derivation is the application of a finite substitution of hypergraphs. To any table T over Σ , we can associate a finite substitution of hypergraphs s_T where $s_T(L) = \{R \mid (L, R) \in T\}$ for each $L \in \Sigma$. Next, we define parallel derivations to be a sequence of parallel direct derivations. Equivalently, we could view these as the application of a composition of substitutions, like in Section 2.3.

Definition 3.4 (Parallel Derivation). Given $\mathcal{C} = (\Sigma, \text{type})$, $H, H' \in \mathcal{H}_{\mathcal{C}}$, and a finite set of tables $\mathcal{T} = \{T_i \mid i \in I\}$ over Σ indexed by I , we say H parallelly derives H' if there is a sequence $H \Rightarrow_{T_{i_1}} H_1 \Rightarrow_{T_{i_2}} \cdots \Rightarrow_{T_{i_k}} H_k = H'$ or $H \cong H'$. We write $H \Rightarrow_{\mathcal{T}}^{i_1 i_2 \cdots i_k} H'$, $H \Rightarrow_{\mathcal{T}}^k H'$, or $H \Rightarrow_{\mathcal{T}}^* H'$. Call $i_1 i_2 \cdots i_k \in I^*$ the trace of the derivation, defined to be ϵ when $H \cong H'$.

Rather than replacing only non-terminals, as is usual in hyperedge replacement grammars, we allow all hyperedges to be replaced, and have a special set of terminal symbols to allow us to say when it is that a hypergraph is terminally labelled, just like ETOL grammars. This decision is more than just a matter of taste. Disallowing rewriting of terminal symbols can change generational power of systems with parallel rewriting. For example, the EOL language $\{a^{n^2} \mid n \geq 1\}$ would not be EOL if we disallowed rewriting of terminal symbols, however, it would at least be ETOL without needing to re-write terminals to something other than themselves. Combining this observation with Theorem 4.4 tells us that $\mathcal{PHR}_{2,1}$ would be different if we were to disallow replacement of terminals.

Definition 3.5 (PHR Grammar). A parallel hyperedge replacement grammar of order k using at most l tables is a tuple $\mathcal{G} = (\mathcal{C}, A, S, \mathcal{T})$ where $\mathcal{C} = (\Sigma, \text{type})$ is a signature, $A \subseteq \Sigma$ is the set of terminal labels, $S \in \Sigma$ is the start symbol, and $\mathcal{T} = \{T_i \mid i \in I\}$ is a non-empty, finite set of tables over Σ indexed by I with $\max(\{\text{type}(r) \mid r \in \bigcup_{T_i \in \mathcal{T}} T_i\}) \leq k$ and $|\mathcal{T}| \leq l$. Call $\Sigma \setminus A$ the non-terminal labels and call \mathcal{G} repetition-free (proper, well-formed) if all its tables are repetition-free (proper, well-formed). Also call \mathcal{G} a (k, l) -PHR grammar, k -PHR grammar, or PHR grammar. The language generated by \mathcal{G} is $L(\mathcal{G}) = \{H \in \mathcal{H}_{\mathcal{C}} \mid S^{\bullet} \Rightarrow_{\mathcal{T}}^* H \text{ with } \text{lab}_H^{-1}(A) = E_H\} \subseteq \mathcal{H}_{\mathcal{C}}$.

Definition 3.6 (PHR Language). $L \subseteq \mathcal{H}_{\mathcal{C}}$ is called a (repetition-free) parallel hyperedge replacement language of order k using at most l tables if there is a (repetition-free) (k, l) -PHR grammar \mathcal{G} such that $L(\mathcal{G}) = L$. The class of (repetition-free) k -PHR languages is the union of all (repetition-free) (k, l) -PHR languages for $l \geq 1$, and the class of (repetition-free) PHR languages is the union of all (repetition-free) k -PHR languages for $k \geq 0$. Denote these $\mathcal{PHR}_{k,l}$, \mathcal{PHR}_k , and \mathcal{PHR} ($\mathcal{PHR}_{k,l}^{\text{rf}}$, $\mathcal{PHR}_k^{\text{rf}}$, and $\mathcal{PHR}^{\text{rf}}$).

Just like languages generated by hyperedge replacement, parallel hyperedge replacement languages are closed under hypergraph isomorphism and are homogeneous in the sense that all hypergraphs in a language have the same type.

Our next theorem confirms that PHR languages strictly contain the HR languages, as expected. We note that our statement is actually stronger than

Theorem VIII.3.3 of Habel's book [20], which sketches a proof for only $k \geq 2$, and did not explicitly consider the number of tables required.

Theorem 3.7 (PHR Generalises HR). For all $k \geq 0$, $\mathcal{HR}_k \subsetneq \mathcal{PHR}_{k,1}$ and $\mathcal{HR}_k^{\text{rf}} \subsetneq \mathcal{PHR}_{k,1}^{\text{rf}}$.

Proof. Suppose $\mathcal{C} = (\Sigma, \text{type})$ and $\mathcal{G} = (\mathcal{C}, N, S, \mathcal{R})$ is a (repetition-free) k -HR grammar with $\mathcal{R} = \{r_1, \dots, r_n\}$. Then we construct a (repetition-free) $(k, 1)$ -PHR grammar $\mathcal{G}' = (\mathcal{C}, A, S, \mathcal{T})$ with $A = \Sigma \setminus N$ and $\mathcal{T} = \{T_1\}$ where $T_1 = \mathcal{R} \cup \{(X, X^\bullet) \mid X \in \Sigma\}$. Clearly every parallel direct derivation with start hypergraph G can be decomposed into at most $|E_G|$ direct derivations where if an edge is replaced by itself, we omit it, and if a genuine replacement from \mathcal{R} occurs, we use that. So, by induction on derivation length, we see that every parallel derivation in \mathcal{G}' can actually be written as a derivation in \mathcal{G} . Similarly, every direct derivation in \mathcal{G} can be lifted to a parallel derivation in \mathcal{G}' by replacing all but one edge by itself. So, by induction on derivation length, we see that every derivation in \mathcal{G} can be written as a parallel derivation in \mathcal{G}' . Thus, together with the fact that the terminal symbols and start symbol coincide, we have that $L(\mathcal{G}) = L(\mathcal{G}')$.

To see strictness, we are inspired by the fact that the string language $\{a^{2^n} \mid n \in \mathbb{N}\}$ is E0L but not MCF (Theorem 2.10). We will show there is a repetition-free $(0, 1)$ -PHR language that is not k -HR for any $k \geq 0$. Let $\mathcal{G} = (\mathcal{C}, A, S, \mathcal{T})$ be the repetition-free $(0, 1)$ -PHR grammar with $\mathcal{C} = (\{\square\}, \{(\square, 0)\})$, $A = \{\square\}$, $S = \square$, and $\mathcal{T} = \{T_1\}$ where $T_1 = \{(\square, \square^\bullet \sqcup \square^\bullet)\}$ (\sqcup denotes disjoint union of hypergraphs). Clearly $L(\mathcal{G})$ is the language of hypergraphs over \mathcal{C} with 2^n hyperedges. By Theorem 2.13, $L(\mathcal{G})$ is not a HR language, as required. \square

Corollary 3.8. PHR languages need not have only linear growth, in the sense of Theorem 2.13.

Finishing this subsection, we remark that it is easy to see that proper PHR languages can be generated by proper PHR grammars:

Lemma 3.9 (Constructing Proper PHR Grammars). For any $k \geq 0, l \geq 1$, given a (repetition-free) (k, l) -PHR grammar \mathcal{G} generating a proper language, one can effectively construct a (repetition-free) proper (k, l) -PHR grammar \mathcal{G}' such that $L(\mathcal{G}) = L(\mathcal{G}')$.

Lemma 3.9 will be useful to us when we look at the string generational power of PHR grammars in Section 4, since string graphs are proper hypergraphs.

3.2. Example PHR Languages

The proof of Theorem 3.7 tells us all HR languages are PHR languages, and also provides us with an example repetition-free $(0, 1)$ -PHR language which is not HR. An example derivation in this grammar is shown in Figure 6. At each stage, the type 0 hyperedges are simultaneously replaced by two such hyperedges.

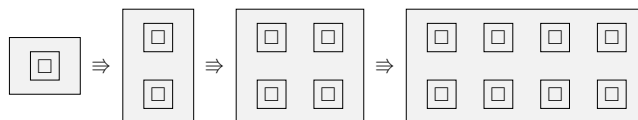


Figure 6: Example parallel derivation

For our next example, recall that one can view rooted unlabelled trees as directed graphs, with the root an external node. Such a tree is called a full binary tree if each node has exactly zero or two children, and if a node has no children, then all other nodes on the same level must also have no children. We denote the language of all such hypergraphs FBT, using a type 2 symbol \square to label all the edges. It is easy to check that the repetition-free $(2, 1)$ -PHR grammar $\mathcal{G} = ((\Sigma, \text{type}), \{\square\}, S, \{T_1\})$ where $\Sigma = \{S, X, Y, F, \square\}$, $\text{type}(S) = 1$, $\text{type}(X) = \text{type}(Y) = \text{type}(F) = \text{type}(\square) = 2$, $T_1 = \{r_0, \dots, r_7\}$, and the eight rules are defined in Figure 7, generates FBT.

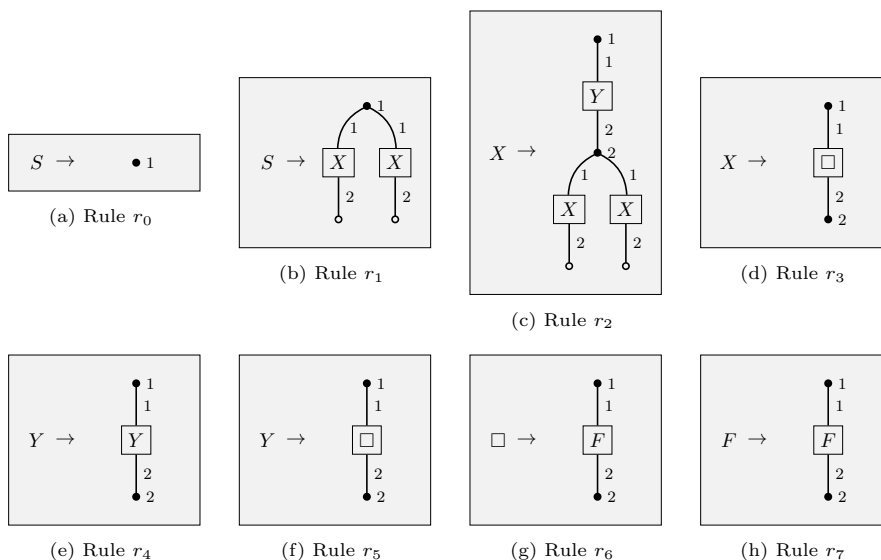


Figure 7: Full binary tree rules

The rule r_6 is responsible for ensuring that whenever terminals are derived, if any non-terminals remained, then the hypergraph can never reach a terminally labelled state in the future, effectively controlling when we stop adding levels to our tree. This trick with the terminals and the F label will re-appear in Subsection 3.5, when we discuss synchronisation. Figures 8 and 9 show positive and negative derivations, by which we mean, derivations that lead to a terminally labelled hypergraph, and derivations that are not a prefix of a derivation that leads to a terminally labelled hypergraph, respectively.

Kreowski showed that his notion of PHR grammars could generate the lan-

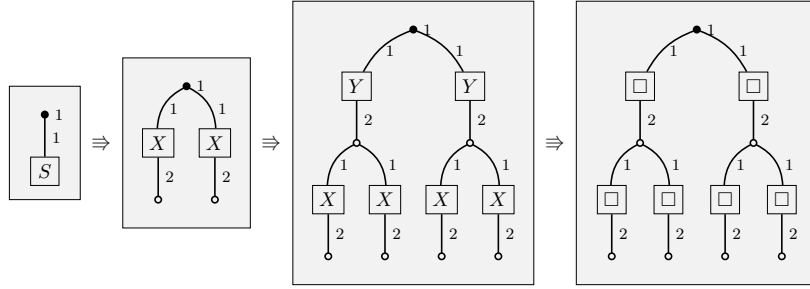


Figure 8: Example positive derivation

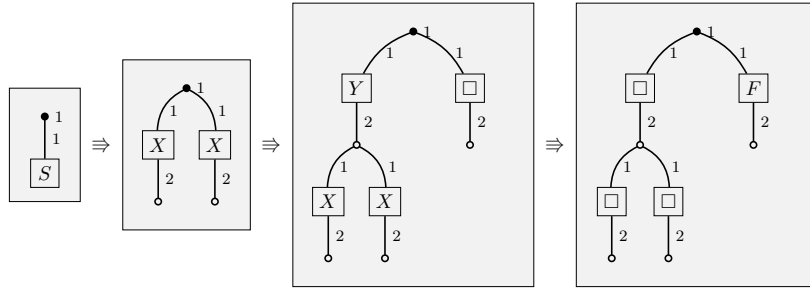


Figure 9: Example negative derivation

guage of Sierpinski triangles, using two tables [21]. We can improve upon this result, showing that only one table is needed, constructing a repetition-free $(2, 1)$ -PHR grammar \mathcal{G} generating the language of all directed \square -labelled Sierpinski triangles. Let $\mathcal{G} = ((\Sigma, \text{type}), \{\square\}, S, \{T_1\})$ where $\Sigma = \{S, X, F, \square\}$, $\text{type}(S) = \text{type}(X) = 3$, $\text{type}(F) = \text{type}(\square) = 2$, $T_1 = \{r_0, \dots, r_5\}$, and the fix rules are defined in Figure 10. Once again, our trick to avoiding a second table is to have the terminal rewrite to F , so that derivations leading to a terminally labelled output must write all the terminals at the same time. Rewriting of terminals was not possible in Kreowski's definition of PHR grammars, and we believe this is precisely what prevented him from using only one table.

3.3. The Power of Tables

In this subsection, we show that for each $k \geq 0$, we in general, have only two possibly distinct classes: the $(k, 1)$ -PHR languages and the k -PHR languages. That is, the availability of three or more tables does not increase generative power beyond what was possible with only two tables. This is exactly the same situation one finds with EOL and ETOL languages.

Theorem 3.10 (PHR Table Power). For all $k \geq 0$, $\text{PHR}_{k,2} = \text{PHR}_k$ and $\text{PHR}_{k,2}^{\text{rf}} = \text{PHR}_k^{\text{rf}}$. Moreover, given a (repetition-free) (proper) (k, l) -PHR grammar \mathcal{G} for any $k \geq 0, l \geq 2$, one can effectively construct a (repetition-free) (proper) $(k, 2)$ -PHR grammar \mathcal{G}' such that $L(\mathcal{G}) = L(\mathcal{G}')$.

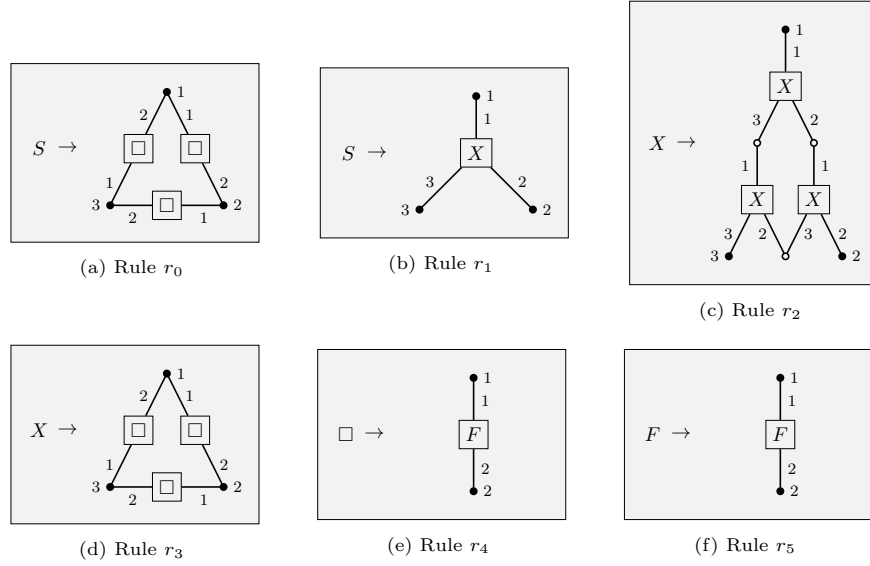


Figure 10: Sierpinski triangle rules

Proof. $\mathcal{PHR}_{k,2} \subseteq \mathcal{PHR}_{k,l}$ and $\mathcal{PHR}_{k,2}^{\text{rf}} \subseteq \mathcal{PHR}_{k,l}^{\text{rf}}$ by definition. To see the reverse inclusions, we follow the proof of Theorem V.1.3 of [4] with modifications lifting the proof from ET0L grammars to PHR grammars. We must show that given some (k, l) -PHR grammar $\mathcal{G} = ((\Sigma, \text{type}), A, S, \{T_1, \dots, T_l\})$ with $l \geq 2$, we can construct a $(k, 2)$ -PHR grammar $\mathcal{G}' = ((\Sigma', \text{type}'), A, S, \{T'_1, T'_2\})$, with $L(\mathcal{G}) = L(\mathcal{G}')$.

We setup the signature as follows. Let $\Sigma' = \{[A, i] \mid A \in \Sigma, i \in \underline{l}\} \cup \Sigma$, $\text{type}'([A, i]) = \text{type}(A)$ for all $A \in \Sigma, i \in \underline{l}$, and $\text{type}'(A) = \text{type}(A)$ for all $A \in \Sigma$. The first table, T'_1 , contains exactly the rules $(A, [A, 1]^\bullet)$ for all $A \in \Sigma$, $([A, i], [A, i+1]^\bullet)$ for all $A \in \Sigma, i \in \underline{l-1}$, and $([A, l], [A, 1]^\bullet)$ for all $A \in \Sigma$. The second table, T'_2 , contains exactly the rules (A, A^\bullet) for all $A \in \Sigma$ and $([A, i], R)$ for all $i \in \underline{l}, (A, R) \in T_i$.

This construction means that the table T'_1 rewrites a hypergraph $H \in \mathcal{H}_C$ into the hypergraph $H_{(1)}$ resulting from H by replacing every label A by $[A, 1]$, then $H_{(1)}$ is rewritten to $H_{(2)}$ by replacing every $[A, 1]$ by $[A, 2]$, etc, with $[A, l]$ looping back round to $[A, 1]$. If at any moment, table T'_2 can be applied, it rewrites $H_{(i)}$ to H' if and only if $H \xRightarrow{T_i} H'$ where H is $H_{(i)}$ with each $[A, i]$ replaced by A . One should also note that the typing function type' is set up to be compatible with all these replacements, so all our rules are well-typed.

By repeating the above process, we can see all derivations in \mathcal{G} are simulated by \mathcal{G}' , so $L(\mathcal{G}) \subseteq L(\mathcal{G}')$, since the terminal labels and start states are the same. The reverse inclusion is obvious, so we have $L(\mathcal{G}) = L(\mathcal{G}')$, as required.

Finally, it is easy to see that all of our transformations preserved if the grammar was repetition-free or proper. \square

3.4. Rational Control of Traces

It is often convenient to restrict the sequences of allowed traces when defining a language using a PHR grammar, leading to better readability of grammars and possibly shorter proofs. A popular choice in L systems is so-called rational control, and was considered for ET0L in 1975 by Nielsen [34] and later by Asveld [35]. We will make precise a notion of rational control for PHR grammars, and show that generational power actually remains the same whenever we have at least two tables, because we can always encode the rational control.

Definition 3.11 (Controlled Parallel Derivation). Given $\mathcal{C} = (\Sigma, \text{type})$, $H, H' \in \mathcal{H}_{\mathcal{C}}$, a finite set of tables $\mathcal{T} = \{T_i \mid i \in I\}$ over Σ indexed by I , and \mathcal{M} an FSA over I , we say H (\mathcal{M} -)parallelly derives H' if H parallelly derives H' with trace $i_1 i_2 \cdots i_k \in L(\mathcal{M})$. We write $H \Rightarrow_{\mathcal{T}}^{i_1 i_2 \cdots i_k} H'$, $H \Rightarrow_{\mathcal{T}}^k H'$, or $H \Rightarrow_{\mathcal{T}}^{\mathcal{M}} H'$.

Definition 3.12 (PHR Grammar with Control). A (repetition-free) parallel hyperedge replacement grammar with control of order k using at most l tables is a tuple $\mathcal{G} = (\mathcal{C}, A, S, \mathcal{T}, \mathcal{M})$ where $(\mathcal{C}, A, S, \mathcal{T})$ is a (repetition-free) (k, l) -PHR grammar (called the underlying grammar) with \mathcal{T} indexed by I , and \mathcal{M} is an FSA over I (called the rational control). The language generated by \mathcal{G} is $L(\mathcal{G}) = \{H \in \mathcal{H}_{\mathcal{C}} \mid S^{\bullet} \Rightarrow_{\mathcal{T}}^{\mathcal{M}} H \text{ with } \text{lab}_H^{-1}(A) = E_H\} \subseteq \mathcal{H}_{\mathcal{C}}$.

Theorem 3.13 (PHR Grammar Control Removal). Given a (repetition-free) k -PHR grammar with control \mathcal{G} , one can effectively construct a (repetition-free) k -PHR grammar \mathcal{G}' such that $L(\mathcal{G}) = L(\mathcal{G}')$.

Proof. Let $\mathcal{G} = ((\Sigma, \text{type}), A, S, \{T_1, \dots, T_l\}, \mathcal{M})$. Without loss of generality, we can assume that $\mathcal{M} = (Q, \underline{n}, \delta, i, F)$ is deterministic and full, and $Q \cap \Sigma = \emptyset$.

We construct the (repetition-free) k -PHR grammar $\mathcal{G}' = ((\Sigma', \text{type}'), A, S', \mathcal{T}')$. First, make a disjoint (from Σ_1) copy of A , \bar{A} , and to each $X \in \bar{A}$, associate a unique $\bar{X} \in A$. Moreover, given a hypergraph H , denote by \bar{H} the same hypergraph but with its labelling function composed with the function that sends $X \in A$ to \bar{X} and leaves everything else in place. Next, choose some additional fresh symbols: S', F_0, \dots, F_k . Using these, we define $\Sigma' = \{S'\} \cup Q \cup \Sigma \cup \bar{A} \cup \{F_0, \dots, F_l\}$ where $\text{type}(S') = \text{type}(S)$, $\text{type}(q) = 0$ for all $q \in Q$, $\text{type}'(X) = \text{type}(X)$ for all $X \in \Sigma$, $\text{type}'(\bar{X}) = \text{type}(X)$ for all $X \in A$, and $\text{type}'(F_j) = j$ for all $j \in \underline{k}$.

Finally, let $\mathcal{T}' = \{T'_0, T'_1, \dots, T'_l\}$ where $T'_0 = \{(S', \bar{S}^{\bullet} \sqcup i^{\bullet})\} \cup \{(q, \emptyset) \mid q \in F\} \cup \{(\bar{X}, X^{\bullet}) \mid X \in A\} \cup \{(X, F_{\text{type}'(X)}) \mid X \in (Q \setminus F) \cup \Sigma \cup \{F_0, \dots, F_k\}\}$ and for each $j \in \underline{l}$, $T'_j = \{(\bar{L}, \bar{R}) \mid (L, R) \in T_j \wedge L \in A\} \cup \{(L, \bar{R}) \mid (L, R) \in T_j \wedge L \notin A\} \cup \{(q, \delta(q, j)^{\bullet}) \mid q \in Q\} \cup \{(X, X^{\bullet}) \mid X \in \{S'\} \cup A \cup \{F_0, \dots, F_k\}\}$.

We can see that the purpose of table 0 is to start and stop the derivation process, with the rest of the tables simulating the original system, while also simulating the automaton. So, if $S^{\bullet} \Rightarrow_{\mathcal{T}}^w H$ is a derivation in \mathcal{G} with H terminally labelled and $w \in L(\mathcal{M})$, then there is a corresponding derivation $S'^{\bullet} \Rightarrow_{\mathcal{T}'}^{0w0} H$ in \mathcal{G}' . That is $L(\mathcal{G}) \subseteq L(\mathcal{G}')$. To see the reverse inclusion, we analyse all derivations of the form $S'^{\bullet} \Rightarrow_{\mathcal{T}'}^w H$. If w does not contain 0 at least twice, then H is necessarily not terminally labelled. So, useful derivations must have trace $x0y0z$ where

$x, y \in \{1, \dots, l\}^*$, $z \in \{0, \dots, l\}^*$. Clearly if $S'^\bullet \Rightarrow_{\mathcal{T}}^x H'$, then $H' \cong S'^\bullet$, so we can assume $x = \epsilon$. Similarly, if $S'^\bullet \Rightarrow_{\mathcal{T}}^{0y^0} H'$ and $H' \Rightarrow_{\mathcal{T}}^z H$, then either $H' \cong H$ if H was terminally labelled, and so we could assume $z = \epsilon$, or H' is labelled by at least non-terminal which has no terminally labelled successor hypergraph, and so it doesn't matter what H is. Finally, we analyse y . If $y \in L(\mathcal{M})$, then we proceed as in the analysis of the other direction of inclusion. If $y \notin L(\mathcal{M})$, then the final step sends the type zero symbol tracking the machine state to F_0 which forces all successors of the hypergraph to be not terminally labelled. \square

Thus, the hypergraph languages that can be generated by (repetition-free) k -PHR grammars are exactly those that can be generated by (repetition-free) k -PHR grammars with control, since certainly no control can be simulated. It is unclear if this also holds for $(k, 1)$ -PHR grammars, but we conjecture that it does not.

3.5. Synchronisation and Substitution Closure

Recall from Subsection 2.3 that an extremely useful notion in L systems is that of a synchronised system. We will now define this notion of PHR grammars, and show that all PHR grammars can be synchronised. This will allow us to show some substitution closure results for various relevant classes of hypergraph languages and hypergraph substitutions.

Definition 3.14 (Synchronised PHR Grammar). Call a PHR grammar $(\mathcal{C}, A, S, \mathcal{T})$ synchronised if for all hypergraphs H, H' over A , there cannot exist a direct derivation $H \Rightarrow_{\mathcal{T}} H'$, for any table $T \in \mathcal{T}$. Moreover, call a symbol $X \in \Sigma$, Y -final if every rule in any table in \mathcal{T} of the form (X, R) is such that $R \cong Y^\bullet$. We call an X -final symbol X , final.

The following result shows that grammars can be synchronised (just like ET0L grammars), and that without loss of generality, we can assume the start symbol essentially does not appear in the RHS of any rules (just like context-free grammars).

Lemma 3.15 (Synchronisation of PHR Grammars). Given a (repetition-free) (proper) (k, l) -PHR grammar $\mathcal{G} = ((\Sigma, \text{type}), A, S, \{T_1, \dots, T_l\})$, one can effectively construct a synchronised (repetition-free) (proper) (k, l) -PHR grammar $\mathcal{G}' = ((\Sigma', \text{type}'), A, S', \mathcal{T}')$ with $\Sigma \subseteq \Sigma'$ and fresh symbols $S', F_0, \dots, F_k \subseteq \Sigma' \setminus \Sigma$ such that:

1. $\text{type}'|_{\Sigma} = \text{type}$ and $L(\mathcal{G}) = L(\mathcal{G}')$;
2. S' is such that $\forall T \in \mathcal{T}', \forall (L, R) \in T, S' \notin \text{lab}_R(E_R)$;
3. $\forall i \in \underline{k}, \text{type}(F_i) = i$ and each F_i is F_i -final;
4. each $X \in A$ is $F_{\text{type}(X)}$ -final.

Proof. We will construct \mathcal{G}' in multiple phases, each preserving the properties established by the phases prior. Property (1) is already satisfied by \mathcal{G} . It is easy to see that properties (3) and (4) imply that the grammar is synchronised,

since any terminally labelled hypergraph is immediately sent to one labelled by only the F_i s, and cannot escape.

We first construct the grammar $\mathcal{G}_0 = ((\Sigma_0, \text{type}_0), A, S', \mathcal{T}_0)$. We have $\Sigma_0 = \Sigma \cup \{S'\}$ where S' is some fresh label, $\text{type}_0(X) = \text{type}(X)$ for all $X \in \Sigma$, and $\text{type}_0(S') = \text{type}(S)$. Finally, set $\mathcal{T}_0 = \{T_{01}, \dots, T_{0l}\}$ where each $T_{0i} = T_i \cup \{(S', S^\bullet)\}$. Clearly \mathcal{G}_0 now additionally satisfies property (2).

Next, we construct the grammar $\mathcal{G}_1 = ((\Sigma_1, \text{type}_1), A, S', \mathcal{T}_1)$. We have $\Sigma_1 = \Sigma_0 \cup \{F_0, \dots, F_l\}$ where the F_i are fresh labels, $\text{type}_1(X) = \text{type}_0(X)$ for all $X \in \Sigma_0$, and $\text{type}_1(F_i) = i$ for all $0 \leq i \leq l$. Finally, set $\mathcal{T}_1 = \{T_{11}, \dots, T_{1l}\}$ where each $T_{1i} = T_{0i} \cup \{(F_j, F_j^\bullet) \mid 0 \leq j \leq l\}$. Clearly \mathcal{G}_1 now additionally satisfies property (3).

Next, we construct the grammar $\mathcal{G}' = ((\Sigma', \text{type}'), A, S', \mathcal{T}')$. First, make a disjoint (from Σ_1) copy of A , \bar{A} and to each $X \in \bar{A}$, associate a unique $\bar{X} \in \bar{A}$. Moreover, given a hypergraph H , denote by \bar{H} the same hypergraph but with its labelling function composed with the function that sends $X \in A$ to \bar{X} and leaves everything else in place. Now, set $\Sigma' = \Sigma_1 \cup \bar{A}$, $\text{type}'(X) = \text{type}_1(X)$ for all $X \in \Sigma_1$, and $\text{type}'(\bar{X}) = \text{type}_1(X)$ for all $X \in A$. Finally, set $\mathcal{T}' = \{T'_1, \dots, T'_l\}$ where each $T'_i = \{(X, F_{\text{type}'(X)}^\bullet), (\bar{X}, X^\bullet) \mid X \in A\} \cup \{(\bar{L}, \bar{R}) \mid (L, R) \in T_{1i} \wedge L \in A\} \cup \{(L, \bar{R}) \mid (L, R) \in T_{1i} \wedge L \in \Sigma_1 \setminus A\}$. Clearly \mathcal{G}' now additionally satisfies property (4). \square

Habel observed that every PHR language can be embedded in an HR language [20]. Using synchronisation, we can quickly establish this result:

Corollary 3.16 (PHR Embedding in HR). Every (repetition-free) (proper) k -PHR language can be embedded in a (repetition-free) (proper) k -HR language, for any $k \geq 0$. Moreover, given a (repetition-free) (proper) k -PHR grammar \mathcal{G} , we can effectively construct a (repetition-free) (proper) k -HR grammar \mathcal{G}' such that $L(\mathcal{G}) \subseteq L(\mathcal{G}')$.

Proof. Lemma 3.15 tells us how to construct a k -PHR grammar $\mathcal{G}_0 = ((\Sigma, \text{type}), A, S, \{T_1, \dots, T_l\})$ satisfying the conditions of the lemma with $L(\mathcal{G}_0) = L(\mathcal{G})$. We now construct the k -HR grammar $\mathcal{G}' = ((\Sigma, \text{type}), N, S, \mathcal{R})$ where $N = \Sigma \setminus A$ and $\mathcal{R} = \{(L, R) \mid (L, R) \in \bigcup \mathcal{T} \wedge L \notin A\}$. Clearly \mathcal{G}' is well-defined and $L(\mathcal{G}) \subseteq L(\mathcal{G}')$ via decomposition of parallel derivations. \square

We can also use synchronisation to show that PHR languages are substitution closed, a property also enjoyed by HR languages (Theorem 2.14). Specifically, we show that a k -HR language under a $(k, 1)$ -PHR substitution is a $(k, 1)$ -PHR language, a $(k, 1)$ -PHR language under a finite substitution is a $(k, 1)$ -PHR language, and that a k -PHR language under a k -PHR substitution is a k -PHR language. Our proof is constructive, starting with grammars, and constructing a grammar generating the image under the substitution. Finally, we also note that being proper, a graph language, or a string graph language is preserved under substitutions that send terminals to languages also satisfying these properties. We will later use our result in Theorem 4.7.

Theorem 3.17 (PHR Languages Substitution Closed). Given $\mathcal{C} = (\Sigma, \text{type})$ and $A \subseteq \Sigma$:

1. for any $k \geq 0$, if L is a k -HR language over A and s is a $\mathcal{PHR}_{k,1}$ -substitution of hypergraphs on A , then $s(L)$ is a $(k, 1)$ -PHR language;
2. for any $k \geq 0$, if L is a $(k, 1)$ -PHR language over A and s is a finite substitution of hypergraphs on A , then $s(L)$ is a $(k, 1)$ -PHR language;
3. for any $k \geq 0$, if L is a k -PHR language over A , and s is a \mathcal{PHR}_k -substitution of hypergraphs on A , then $s(L)$ is a k -PHR language.

Moreover, if \mathcal{P} is the property of being repetition-free, proper, a graph language, or a string graph language, if L is \mathcal{P} and $s(X)$ is \mathcal{P} for each $X \in A$, then $s(L)$ is \mathcal{P} .

Proof. Let $A = \{X_1, \dots, X_m\}$ and $B = \bigcup_{H \in s(A)} \text{lab}_H(E_H)$ throughout this proof, and without loss of generality, we can assume $A \cap B = \emptyset$, by simply renaming items in A and adjusting s accordingly.

To see (1), there must exist $(k, 1)$ -PHR grammars $\mathcal{G}_i = ((\Sigma_i, \text{type}_i), B, S_i, \{T_i\})$ such that $L(\mathcal{G}_i) = s(X_i)$ for all $i \in \underline{m}$, each \mathcal{G}_i satisfies the conditions Lemma 3.15, the final symbols F_0, \dots, F_k are shared between the grammars, and the remaining non-terminals are disjoint (for all $i \neq j \in \underline{m}$, $(\Sigma_i \setminus (\{F_0, \dots, F_k\} \cup A_i)) \cap \Sigma_j \setminus (\{F_0, \dots, F_k\} \cup A_j) = \emptyset$). There must also exist a k -HR grammar $\mathcal{G} = (\mathcal{C}, N, S, \mathcal{R})$ such that $L(\mathcal{G}) = L$, $\Sigma \setminus N = A$, and $N \cap (\bigcup_{i \in \underline{k}} \Sigma_i) = \emptyset$.

We now use the proof of Theorem 3.7 to construct a $(k, 1)$ -PHR grammar $\mathcal{G}_0 = (\mathcal{C}, A, S, \{T\})$ such that $L(\mathcal{G}_0) = L(\mathcal{G}) = L$. Recall this construction set $T = \mathcal{R} \cup \{(X, X^\bullet) \mid X \in \Sigma\}$. We now construct a new $(k, 1)$ -PHR grammar $\mathcal{G}' = ((\Sigma', \text{type}'), B, S, \{T'\})$. Set $\Sigma' = \Sigma \cup \Sigma_1 \cup \dots \cup \Sigma_m$ and the type function accordingly. Set $T' = T \cup \{(X_i, S_i) \mid 1 \leq i \leq m\} \cup \{(F_i, F_i^\bullet) \mid 0 \leq i \leq k\} \cup \{(X, F_{\text{type}'(X)}^\bullet) \mid X \in B\} \cup \bigcup_{i \in \underline{m}} \{(L, R) \in T_i \mid L \in \Sigma \setminus (B \cup \{F_0, \dots, F_k\})\}$ where the T_1, \dots, T_m are the single tables from the $\mathcal{G}_1, \dots, \mathcal{G}_m$, respectively. It is now easy to see that derivations in our new grammar simulate derivations in \mathcal{G} , and then arbitrarily re-write terminals (in A) to the start symbol of the grammar generating the language that terminal is to be substituted for. Note that because the rule (X, X^\bullet) is in the table for all $X \in A$, each derivation step need not immediately re-write a terminal to the start symbol of the associated grammar. The effect of this is that when these grammars are then simulated, they need not run for the same number of steps as each other. It is now clear that $L(\mathcal{G}') = s(L(\mathcal{G})) = s(L)$, as required.

To see (2), let $\mathcal{G} = (\mathcal{C}, A, S, \{T\})$ be a $(k, 1)$ -PHR grammar generating L . Without loss of generality, we can assume it satisfies the conditions of Lemma 3.15 with final symbols F_0, \dots, F_k .

We now construct a new $(k, 1)$ -PHR grammar $\mathcal{G}' = ((\Sigma', \text{type}'), B, S, \{T'\})$ where $\Sigma' = \Sigma \cup B$ and the typing function is defined in the obvious way. Next, we define the finite substitution s' on Σ by $s'(X) = s(X)$ for all $X \in A$ and $s'(X) = \{X^\bullet\}$ for all $X \in \Sigma \setminus A$. Finally, let $T' = (\bigcup \{(L, H) \mid H \in s'(R)\} \mid (L, R) \in T \wedge L \in \Sigma \setminus A\}) \cup \{(X, F_{\text{type}'(X)}) \mid X \in B\}$. It is now clear that this new grammar simulates the original, only with all terminals in RHS replaced

with all combinations of their replacements under s , and the synchronisation properties are preserved.

To see (3), there must exist k -PHR grammars $\mathcal{G}_i = ((\Sigma_i, \text{type}_i), B, S_i, \mathcal{T}_i)$ such that $L(\mathcal{G}_i) = s(X_i)$ for all $i \in \underline{m}$, each \mathcal{G}_i satisfies the conditions Lemma 3.15, the final symbols F_0, \dots, F_k are shared between the grammars, and the remaining non-terminals are disjoint (for all $i \neq j \in \underline{m}$, $(\Sigma_i \setminus (\{F_0, \dots, F_k\} \cup B)) \cap \Sigma_j \setminus (\{F_0, \dots, F_k\} \cup A_j) = \emptyset$). There must also exist a k -PHR grammar $\mathcal{G} = (\mathcal{C}, A, S, \mathcal{T})$ such that $L(\mathcal{G}) = L$, and $\Sigma \cap (\bigcup_{i \in \underline{k}} \Sigma_i) = \emptyset$. For each $i \in \underline{m}$, let $\bar{\Sigma}_i$ be a copy of Σ_i consisting of fresh symbols, and identify each $X \in \Sigma_i$ with its copy $\bar{X} \in \bar{\Sigma}_i$. Moreover, given a hypergraph H , by \bar{H} we mean H but with its labelled function composed with the function which takes any $X \in \Sigma_i$ to \bar{X} and leaves everything else fixed.

We now construct the k -PHR grammar $\mathcal{G}' = (\mathcal{C}', B, \mathcal{T}', S)$ such that $L(\mathcal{G}') = S(L)$. Let $\mathcal{C}' = (\Sigma', \text{type}')$ be the union of all the above signatures also including the disjoint copies, where copies are of the same type as their original symbols. Finally, let $\mathcal{R} = \{(X, X^\bullet) \mid X \in \Sigma'\}$, $\mathcal{F} = \{(X, F_{\text{type}'(X)}^\bullet) \mid X \in \Sigma'\}$, and $\mathcal{T}' = \bigcup_{0 \leq i \leq m} \mathcal{T}'_i$, where $\mathcal{T}'_0 = \{\mathcal{R} \oplus T \mid T \in \mathcal{T}\} \cup \{\mathcal{F} \oplus \{(X_i, S_i^\bullet) \mid i \in \underline{m}\}\}$, and for each $i \in \underline{m}$ let $\mathcal{T}'_i = \{\mathcal{R} \oplus (\{(\bar{L}, \bar{R}) \mid (L, R) \in T\} \cup \{(\bar{S}_i, \bar{S}_i^\bullet)\}) \mid T \in \mathcal{T}_i\} \cup \{\mathcal{R} \oplus (\{(\bar{X}, X^\bullet) \mid X \in B\} \cup \{(\bar{X}, F_{\text{type}'(X)}^\bullet) \mid X \in \Sigma_i \setminus (B \cup \{S_i\})\})\}$.

One can see that derivations that make progress start with S^\bullet then apply tables from the first part of \mathcal{T}'_0 , simulating \mathcal{G} . At some point, the final table of \mathcal{T}'_0 may be applied, which immediately rewrites all the terminals to encoded start symbols for their respective grammars for substitution and sends all non-terminals to failure non-terminals. If a hypergraph contains any failure non-terminals at this point, non terminally labelled hypergraph can be derived in future. Derivations can now simulate the \mathcal{G}_i totally independently, with choice of delaying start, giving total freedom over the simulated derivation sequences for each instance of the encoded start symbol. Finally, the encoded systems can end their simulation at any point, just like before, by sending their encoded terminals to real terminals in B and their encoded non-terminals to failure non-terminals. It is now clear that $L(\mathcal{G}') = s(L)$, as required.

Finally, preservation of the property of being repetition-free, proper, a graph language, or a string graph language is easy to see. \square

In light of the above result, we can view the language FBT from Subsection 3.2 as the image of the repetition-free (2, 1)-PHR grammar $((\{S, X, Y\}, \langle S \mapsto 1, X \mapsto 2, Y \mapsto 2 \rangle), \{X, Y\}, S, \{\{r_0, r_1, r_2, r_3\}\})$ under the repetition-free finite substitution of hypergraphs $h(X) = h(Y) = \square^\bullet$, where the r_i are defined in Figure 11. Moreover, we can generate the language of full binary trees over an arbitrary label set $\{X_1, \dots, X_m\}$ using the substitution $h(X) = h(Y) = \{X_1^\bullet, \dots, X_m^\bullet\}$. A similar trick is also possible with the grammar that generated the language of all directed Sierpinski triangles.

Finally, we can generalise Theorem 2.9 for PHR languages:

Theorem 3.18 (PHR Languages Hyper-Algebraically Closed). Given $\mathcal{C} = (\Sigma, \text{type})$, $A \subseteq \Sigma$, and $k \geq 0$, if L is a k -PHR language over A , and \mathcal{S} is a finite set

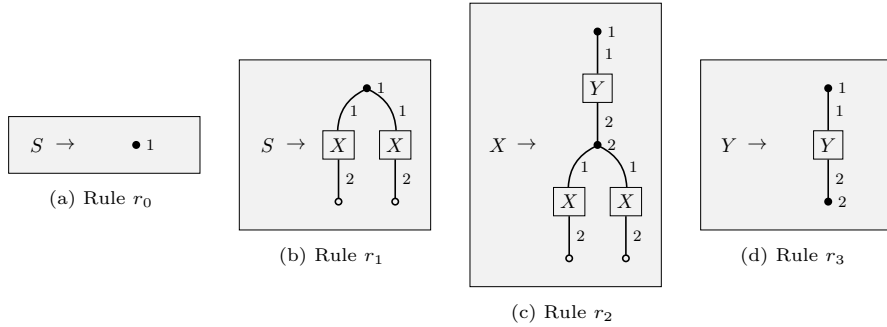


Figure 11: Full binary tree rules revisited

of \mathcal{PHR}_k -substitutions of hypergraphs on A , then $L' = \bigcup_{n \in \mathbb{N}} \{s_n(\cdots s_1(L) \cdots) \mid s_1, \dots, s_n \in \mathcal{S}\}$ is a k -PHR language. Moreover, if \mathcal{P} is the property of being repetition-free, proper, a graph language, or a string graph language, if L is \mathcal{P} and $s(X)$ is \mathcal{P} for each $s \in \mathcal{S}$ and $X \in A$, then L' is \mathcal{P} .

Proof. The proof is very similar to that of part (3) of Theorem 3.17. The only difference is the addition of another table which can be used to send all terminals to restart the process and all non-terminals to a failure symbol. \square

Corollary 3.19 (PHR Languages Iterated Substitution Closed). For all $k \geq 0$, \mathcal{PHR}_k ($\mathcal{PHR}_k^{\text{rf}}$) is iterated substitution closed.

3.6. The Emptiness Problem

An important and useful decision problem is the emptiness problem. That is, given a PHR grammar, decide if it generates a non-empty language or not. In this subsection, we show this problem is decidable, in general. This result also has utility when eliminating unreachable symbols from PHR grammars, which we tackle at the end of this subsection, but also for providing us with an explicit algorithm for solving the membership problem for PHRS languages (Subsection 4.4).

Before we present and prove Theorem 3.25, we first must define label sets, and derivations on label sets. Soundness and completeness of simulation of PHR derivations by label set derivations, together with the fact that analysis of label sets is easy, enables us to show how to decide emptiness for PHR grammars.

Definition 3.20 (Label Set). Given some signature $\mathcal{C} = (\Sigma, \text{type})$ let $\mathcal{L}_{\mathcal{C}} = \mathcal{P}(\Sigma)$ and labels $: \mathcal{H}_{\mathcal{C}} \rightarrow \mathcal{L}_{\mathcal{C}}$ be the function defined by $H \mapsto \text{lab}_H(E_H)$. Call $X \in \mathcal{L}_{\mathcal{C}}$ a label set and define the hypergraph:

$$H(X) = \bigsqcup_{A \in X} \text{untype}(A^\bullet)$$

where $\text{untype}(H)$ is H with ext_H replaced by \emptyset .

Definition 3.21 (Direct Derivation). If $X, X' \in \mathcal{L}_{\mathcal{C}}$ and T is a table over \mathcal{C} , then write $X \Rightarrow_T X'$ whenever $\exists H' \in \mathcal{H}_{\mathcal{C}}, H(X) \Rightarrow_T H' \wedge \text{labels}(H') = X'$.

The following three lemmata are easy to see:

Lemma 3.22 (Restriction of Label Set Direct Derivations). Given \mathcal{C} and a table T over \mathcal{C} , then $\forall X, Y, Y' \in \mathcal{L}_{\mathcal{C}}$:

$$(Y \Rightarrow_T Y' \wedge X \subseteq Y) \Rightarrow (\exists X' \in \mathcal{L}_{\mathcal{C}}, X \Rightarrow_T X' \wedge X' \subseteq Y').$$

Lemma 3.23 (Completeness of Label Set Direct Derivations). Given \mathcal{C} , and a table T over \mathcal{C} , then $\forall H, H' \in \mathcal{H}_{\mathcal{C}}$:

$$(H \Rightarrow_T H') \Rightarrow (\exists X' \in \mathcal{L}_{\mathcal{C}}, \text{labels}(H) \Rightarrow_T X' \wedge X' \subseteq \text{labels}(H')).$$

Lemma 3.24 (Soundness of Label Set Direct Derivations). Given \mathcal{C} and a table T over \mathcal{C} , then $\forall X, X' \in \mathcal{L}_{\mathcal{C}}, \forall H \in \mathcal{H}_{\mathcal{C}}$:

$$(X \Rightarrow_T X' \wedge \text{labels}(H) = X) \Rightarrow (\exists H' \in \mathcal{H}_{\mathcal{C}}, H \Rightarrow_T H' \wedge X' = \text{labels}(H')).$$

Theorem 3.25 (Decidable PHR Emptiness Problem). The following problem is decidable, with an explicit algorithm:

Instance: A PHR grammar $\mathcal{G} = (\mathcal{C}, A, S, \mathcal{T})$.

Question: Is $L(\mathcal{G}) = \emptyset$?

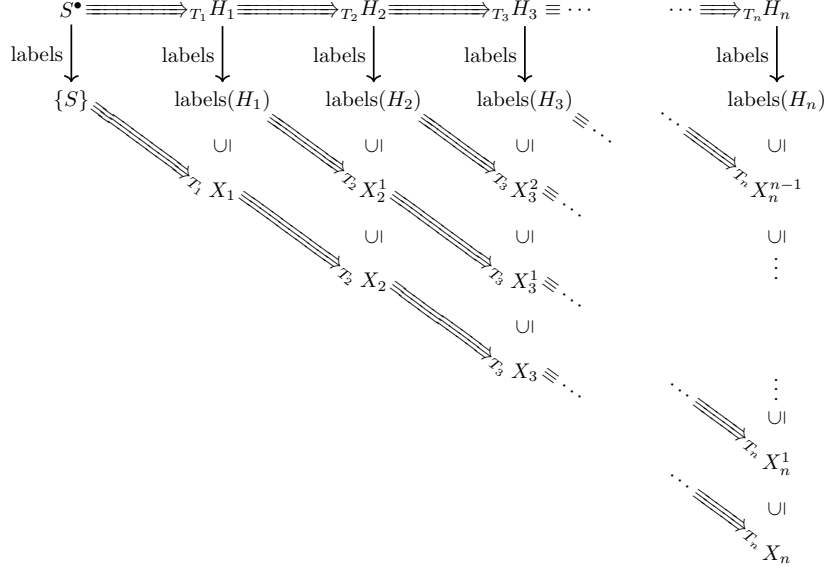
Proof. Write $X \Rightarrow_{\mathcal{T}}^+ X'$ if there exists a sequence of direct derivations $X \Rightarrow_{T_1} X_1 \Rightarrow_{T_2} \cdots \Rightarrow_{T_n} X'$ with $n \geq 1, T_i \in \mathcal{T}, X, X', X_i \in \mathcal{L}_{\mathcal{C}}$, and define $\text{succ}_{\mathcal{T}}(X) = \{X' \in \mathcal{L}_{\mathcal{C}} \mid X \Rightarrow_{\mathcal{T}}^+ X'\}$, computed by iteratively computing label set direct successors (breadth first search). Eventually this process must terminate due to the fact that $\mathcal{L}_{\mathcal{C}}$ is finite. We claim that $\text{succ}_{\mathcal{T}}(\{S\})$ contains a label set containing only terminals if and only if $L(\mathcal{G}) \neq \emptyset$.

Suppose first that $L(\mathcal{G})$ is non-empty. Then there must be a hypergraph derivation sequence $S^{\bullet} \Rightarrow_{T_1} H_1 \Rightarrow_{T_2} \cdots \Rightarrow_{T_n} H_n$ with $n \geq 1$ and $\text{lab}_H(E_H) \subseteq A$, which, by Lemmata 3.22 and 3.23, induces the corresponding label set derivation sequence $\{S\} \Rightarrow_{T_1} X_1 \Rightarrow_{T_2} \cdots \Rightarrow_{T_n} X_n$ such that $X_n \subseteq \text{labels}(H_n) \subseteq A$ and $X_n \in \text{succ}_{\mathcal{T}}(X)$, as per Figure 12(a).

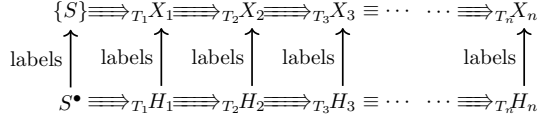
Now suppose that there is a label set containing only terminals in $\text{succ}_{\mathcal{T}}(S^{\bullet})$. Then there must be a label set derivation sequence $\{S\} \Rightarrow_{T_1} X_1 \Rightarrow_{T_2} \cdots \Rightarrow_{T_n} X_n$ with $n \geq 1$ and $X_n \subseteq A$. By Lemma 3.24, there must be a corresponding derivation sequence $S^{\bullet} \Rightarrow_{T_1} H_1 \Rightarrow_{T_2} \cdots \Rightarrow_{T_n} H_n$, as per Figure 12(b). Since $\text{labels}(H_n) = X_n \subseteq A$, we know $H_n \in L(\mathcal{G})$ so $L(\mathcal{G})$ is non-empty. \square

Finally, the above view of derivations informs us of how to eliminate unreachable labels from a PHR grammar. Elimination of unreachable symbols will be needed later in the proof of Theorem 4.4.

Definition 3.26 (Unreachable Symbol). Given a PHR grammar over (Σ, type) , $X \in \Sigma$ is called unreachable if there is no derivation starting at S^{\bullet} containing a hypergraph with a hyperedge labelled by X .



(a) First part



(b) Second part

Figure 12: Diagrams for proof of Theorem 3.25

Lemma 3.27 (Elimination of Unreachable Labels). For $k \geq 2, l \geq 1$, given a (repetition-free) (proper) (k, l) -PHR grammar \mathcal{G} , one can effectively construct a (repetition-free) (proper) (k, l) -PHR grammar \mathcal{G}' with no unreachable symbols and $L(\mathcal{G}) = L(\mathcal{G}')$.

Proof. Similar to the proof of Theorem 3.25, if a symbol appears in a hypergraph derivation sequence starting at S^\bullet , then it appears in a label set derivation sequence starting at $\{S\}$, and vice versa. So determining if a symbol is unreachable is just a finite process.

Removal of an unreachable symbol from a grammar is an easy process. Suppose that we have identified $X \in \Sigma$ as an unreachable symbol in the PHR grammar $\mathcal{G} = ((\Sigma, \text{type}), A, S, \mathcal{T})$. Then the grammar $\mathcal{G}' = ((\Sigma \setminus \{X\}, \text{type}|_{\Sigma \setminus \{X\}}), A \setminus \{X\}, S, \{(L, R) \mid (L, R) \in T \wedge L \neq X \wedge X \notin \text{lab}_R(E_R)\} \mid T \in \mathcal{T})$ no longer contains the symbol X and is such that $L(\mathcal{G}) = L(\mathcal{G}')$. We can repeat this process to remove all unreachable symbols. It is easy to see that after we have completed this process, the resultant grammar will have no unreachable symbols. \square

4. PHRS Languages

We now turn our attention to string languages. We believe the family of parallel hyperedge replacement string languages is a genuinely new family, containing all MCF and ET0L languages. It is not simply equal to the (parallel) MCF languages because these are known to be incomparable with ET0L [32]. Recall that the hyperedge replacement string languages are exactly the MCF languages. We will confirm that parallel hyperedge replacement string (PHRS) languages contain all MCF languages and also all of the ET0L languages. We then go on to show various closure properties including that the family of PHRS languages is a hyper-algebraically closed super AFL with decidable membership.

4.1. Definitions and Foundations

Definition 4.1 (PHR String Language). A string language $L \subseteq A^*$ is called a (repetition-free) parallel hyperedge replacement string language of order k with at most l tables if there is a (repetition-free) (k, l) -PHR grammar \mathcal{G} such that \mathcal{G} generates a string graph language and $\text{STR}(\text{L}(\mathcal{G})) = L \setminus \{\epsilon\}$. The family of (repetition-free) k -PHRS languages is the union of all (repetition-free) (k, l) -PHRS languages for $l \geq 1$, and the family of (repetition-free) PHRS languages is the union of all (repetition-free) k -PHRS languages for $k \geq 2$. Denote these $\mathcal{PHRS}_{k,l}$, \mathcal{PHRS}_k , and \mathcal{PHRS} ($\mathcal{PHRS}_{k,l}^{\text{rf}}$, $\mathcal{PHRS}_k^{\text{rf}}$, and $\mathcal{PHRS}^{\text{rf}}$).

Notice how we exclude the cases where $k < 2, l < 1$. This is to ensure the classes form a family of string languages, as per the definition in Subsection 2.1, since $\mathcal{PHRS}_0^{\text{rf}} = \mathcal{PHRS}_0 = \mathcal{PHRS}_1^{\text{rf}} = \mathcal{PHRS}_1 = \{\emptyset, \{\epsilon\}\}$. Notice also that $\mathcal{PHRS}_{k,l}^{\text{rf}} \subseteq \mathcal{PHRS}_{k,l}$ for $k \geq 2, l \geq 1$, and Theorem 3.10 tells us:

Theorem 4.2 (PHRS Table Power). For all $k \geq 2$, $\mathcal{PHRS}_{k,2}^{\text{rf}} = \mathcal{PHRS}_k^{\text{rf}}$ and $\mathcal{PHRS}_{k,2} = \mathcal{PHRS}_k$.

We now show that the PHRS languages generalise both the ET0L and MCF languages. In particular, the $(2, 1)$ -PHRS languages are exactly the E0L languages, and the 2-PHR languages, ET0L.

Lemma 4.3. Given a (k, l) -PHR grammar \mathcal{G} generating a string graph language, one can effectively construct a proper (k, l) -PHR grammar \mathcal{G}' such that there are no unreachable symbols, all terminals are type 2, all non-terminals are type at least 2, and $\text{L}(\mathcal{G}) = \text{L}(\mathcal{G}')$.

Proof. Since all string graph languages are proper, Lemmata 3.9 and 3.27 tell us how to construct a proper (k, l) -PHR grammar \mathcal{G}' such that there are no unreachable symbols and $\text{L}(\mathcal{G}) = \text{L}(\mathcal{G}')$. All terminals can be assumed to be type 2 since the grammar generates a string graph language. Finally, all type 0 non-terminals can be converted to type 1, since the start symbol must be of type 2, so there will always be a node they can attached to. Next all type 1 non-terminals, including those we just converted from type 0, can be converted to type 2 symbols in the obvious way, and all rules that use them updated so that the extra node is merged upon deletion of the symbol. \square

Theorem 4.4 (PHRS Generalises ETOL). For all $k \geq 2$, $\mathcal{EOL} \subseteq \mathcal{PHRS}_{k,1}^{\text{rf}}$ and $\mathcal{ETOL} \subseteq \mathcal{PHRS}_k^{\text{rf}}$. When $k \geq 4$, $\mathcal{EOL} \subsetneq \mathcal{PHRS}_{k,1}^{\text{rf}}$ and $\mathcal{ETOL} \subsetneq \mathcal{PHRS}_k^{\text{rf}}$. Moreover, $\mathcal{EOL} = \mathcal{PHRS}_{2,1}^{\text{rf}} = \mathcal{PHRS}_{2,1}$ and $\mathcal{ETOL} = \mathcal{PHRS}_2^{\text{rf}} = \mathcal{PHRS}_2$.

Proof. First we show $\mathcal{EOL} = \mathcal{PHRS}_{2,1}^{\text{rf}}$ and $\mathcal{ETOL} \subseteq \mathcal{PHRS}_k^{\text{rf}}$ for all $k \geq 2$. Suppose L is an E(T)OL language, then by Theorem 2.7 there exists a propagating ETOL grammar $\mathcal{G} = (\Sigma, A, S, \{T_i \mid i \in I\})$ such that $L \setminus \{\epsilon\} = L(\mathcal{G})$. It follows that every rule can be encoded as a hyperedge replacement rule over $\mathcal{C}' = (\Sigma, \Sigma \times \{2\})$ giving us a repetition-free 2-PHR grammar $\mathcal{G}' = (\mathcal{C}', A, S, \{(L, R^\bullet) \mid (L, R) \in T_i \mid i \in I\})$ with $L(\mathcal{G}) = \text{STR}(L(\mathcal{G}'))$.

Next, we show that $\mathcal{PHRS}_{2,1} \subseteq \mathcal{EOL}$ and $\mathcal{PHRS}_2 \subseteq \mathcal{ETOL}$. Suppose L is a $(2, l)$ -PHRS language, then there is a $(2, l)$ -PHR grammar $\mathcal{G} = (\mathcal{C}, A, S, \{T_i \mid i \in I\})$ generating a string graph language such that $L \setminus \{\epsilon\} = \text{STR}(L(\mathcal{G}))$. Lemma 4.3 allows us to assume a lot about the form of RHSs of rules. It is easy to see that all RHSs must actually be string graphs, or could be transformed to string graphs, since any non-conformant pieces can just be inlined into the string graph because it will ultimately be deleted and the nodes merged in any terminally labelled derived hypergraph. Thus, the system can be converted into an ETOL grammar using at most l tables.

Thus, we have $\mathcal{PHRS}_{2,1} \subseteq \mathcal{EOL} \subseteq \mathcal{PHRS}_{2,1}^{\text{rf}}$ and $\mathcal{PHRS}_{2,1}^{\text{rf}} \subseteq \mathcal{PHRS}_{2,1}$, so all these inclusions must be equalities. Similarly, we have $\mathcal{PHRS}_{2,l} \subseteq \mathcal{ETOL} \subseteq \mathcal{PHRS}_{2,l}^{\text{rf}}$ and $\mathcal{PHRS}_{2,l}^{\text{rf}} \subseteq \mathcal{PHRS}_{2,l}$, so all these inclusions must be equalities.

Finally, strictness follows from Theorems 2.2, 2.10, and 3.7. That is, we can construct a repetition-free $(4, 1)$ -PHR grammar \mathcal{G}' generating a string graph language with $\text{STR}(L(\mathcal{G}')) = K \setminus \{\epsilon\}$ (from Theorem 2.10) which is not EOL. \square

Corollary 4.5. There are repetition-free $(2, 1)$ -PHRS languages that are not semilinear.

Proof. Theorem 2.10 tells us $L = \{a^{2^n} \mid n \in \mathbb{N}\}$ is an EOL language which is not semilinear. Theorem 4.4 tells us L is a repetition-free $(2, 1)$ -PHRS language. \square

Theorem 4.6 (PHRS Generalises MCF). For all $k \geq 2$, $\mathcal{HRS}_k^{\text{rf}} \subsetneq \mathcal{PHRS}_k^{\text{rf}}$.

Proof. This follows from Theorem 2.2, and Theorem 3.7 and its proof. We get strictness from Theorem 2.10 together with Theorem 4.4. \square

4.2. Formal Language Closure Properties

In this subsection, we show that the family of PHRS languages is a hyper-algebraically closed super AFL, and that the family of repetition-free PHRS languages is closed under non-erasing (iterated) substitution, with closure under rational operations and non-erasing homomorphisms following from this as a corollary.

Theorem 4.7 (PHRS Closed Under Substitutions). Let $L \subseteq A^*$. Then:

1. for any $k \geq 2$, if L is a k -HRS language and h is a $\mathcal{PHRS}_{k,1}$ -substitution (non-erasing $\mathcal{PHRS}_{k,1}^{\text{rf}}$ -substitution) on A , then $h(L)$ is a $(k, 1)$ -PHRS language (rep.-free $(k, 1)$ -PHRS language);

2. for any $k \geq 2$, if L is a $(k, 1)$ -PHRS language (rep.-free $(k, 1)$ -PHRS language) and h is a finite substitution (non-erasing finite substitution) on A , then $h(L)$ is a $(k, 1)$ -PHRS language (rep.-free $(k, 1)$ -PHRS language);
3. for any $k \geq 2$, if L is a k -PHRS language (rep.-free k -PHRS language) and h is a \mathcal{PHRS}_k -substitution (non-erasing $\mathcal{PHRS}_k^{\text{rf}}$ -substitution) on A , then $h(L)$ is a k -PHRS language (rep.-free k -PHRS language).

Proof. First, recall that all k -HRS languages can be generated by repetition-free k -HRS grammars by Theorem 2.2. Now, for all $k \geq 2, l \geq 1$, all (repetition-free) $\mathcal{PHRS}_{k,l}$ -substitutions h of strings on A can be converted to (repetition-free) $\mathcal{PHR}_{k,l}$ -substitutions s of hypergraphs on A viewing A as part of the signature (A, type) where $\text{type}(X) = 2$ for all $X \in A$, defining $s(X) = \{w^\bullet \mid w \in h(X)\}$ for all $X \in A$. The rest then follows from Theorem 3.17, translating back from string graphs to strings. \square

Corollary 4.8 (PHRS Closed Under Homomorphisms). Let $L \subseteq A^*$ be a (k, l) -PHRS language (repetition-free (k, l) -PHRS language) for any $k \geq 2, l \geq 1$ and $\varphi : A^* \rightarrow B^*$ be a homomorphism (non-erasing homomorphism). Then $\varphi(L)$ is a (k, l) -PHRS language (repetition-free (k, l) -PHRS language).

Proof. If $l = 1$, then we can use Theorem 4.7(2) since a (non-erasing) homomorphism can be trivially viewed as a (non-erasing) finite substitution. If $l \geq 2$, then we can similarly use Theorem 4.7(3). \square

Theorem 4.9 (PHRS Hyper-Algebraically Closed). Let $L \subseteq A^*$. Then, for any $k \geq 2$, if L is a k -PHRS language (repetition-free k -PHRS language) and H is a finite set of \mathcal{PHRS}_k -substitutions (non-erasing $\mathcal{PHRS}_k^{\text{rf}}$ -substitutions) on A , then $\text{ITER}_H(L)$ is a k -PHRS language (repetition-free k -PHRS language).

Proof. This follows directly from Theorem 3.18, realising substitutions as in the proof of Theorem 4.7. \square

Next, we show closure under rational operations, which can be seen via the following general result:

Lemma 4.10. Let \mathcal{F} be a family of string languages such that if $L \subseteq A^*$ is regular and h is a non-erasing \mathcal{F} -substitution on A , then $h(L) \in \mathcal{F}$. Then for \mathcal{F} languages $L_1, L_2 \subseteq A^*$:

1. $L_1 \cup L_2$ is an \mathcal{F} language; (closure under union)
2. $L_1 L_2$ is an \mathcal{F} language; (closure under concatenation)
3. L_1^+ is an \mathcal{F} language. (closure under Kleene plus)

Proof. To see (1), notice that $L_1 \cup L_2$ is simply $h(K)$ where $K = \underline{\text{if}} \epsilon \in L_1 \cup L_2 \underline{\text{then}} \{X, Y, \epsilon\} \underline{\text{else}} \{X, Y\}$, and h is a non-erasing \mathcal{F} -substitution defined by $h(X) = L_1 \setminus \{\epsilon\}$ and $h(Y) = L_2 \setminus \{\epsilon\}$. Thus we have $L_1 \cup L_2 = h(K)$, and since, in either case, K is a regular language, $h(K) \in \mathcal{F}$ due to our assumptions about \mathcal{F} , as required.

To see (2), we again construct a regular language K . If $\epsilon \in X$ and $\epsilon \in Y$, then set $K = \{XY, X, Y, \epsilon\}$. If only $\epsilon \in X$, set $K = \{XY, Y\}$. If only $\epsilon \in Y$, set $K = \{XY, X\}$. Otherwise, set $K = \{XY\}$. Finally, h is exactly as before, giving us $L_1L_2 = h(K) \in \mathcal{F}$ as required.

Finally, to see (3), set $K = \underline{\text{if}} \epsilon \in L_1 \underline{\text{then}} \{X^n \mid n \geq 0\} \underline{\text{else}} \{X^n \mid n \geq 1\}$, and set $h(X) = L_1$, giving us $L_1^+ = h(K) \in \mathcal{F}$ as required. \square

Theorem 4.11 (PHRS Closed Under Rational Operations). Let $L_1, L_2 \subseteq A_1^*$ be (repetition-free) (k, l) -PHRS languages for any $k \geq 2, l \geq 1$. Then:

1. $L_1 \cup L_2$ is a (r.-free) (k, l) -PHRS language; (closure under union)
2. L_1L_2 is a (r.-free) (k, l) -PHRS language; (closure under concatenation)
3. L_1^+ is a (r.-free) (k, l) -PHRS language. (closure under Kleene plus)

Proof. If $l = 1$, then Theorem 4.7(1) tells us that $\mathcal{PHRS}_{k,1}$ satisfies the condition of Lemma 4.10 since $\mathcal{REG} \subseteq \mathcal{HRS}_k$. If $l \geq 2$ then Theorem 4.7(3) tells us that \mathcal{PHRS}_k satisfies the condition of Lemma 4.10 since $\mathcal{REG} \subseteq \mathcal{PHRS}_k$. \square

Corollary 4.12 (PHRS More Than ET0L and MCF). We have the following strict inclusions: $\mathcal{MCF}_2 \cup \mathcal{EOL} \subsetneq \mathcal{PHRS}_{4,1}^{\text{rf}}$ and $\mathcal{MCF}_2 \cup \mathcal{ETOL} \subsetneq \mathcal{PHRS}_4^{\text{rf}}$.

Proof. Inclusion follows from Theorems 4.4 and 4.6, but is not the interesting part of this Theorem. The interesting part is that there are repetition-free $(2, 1)$ -PHRS (repetition-free 2-PHRS) languages that are not just E0L (ET0L) or 2-MCF. Take $K \subseteq A^*$ and $L \subseteq B^*$ to be as in Theorem 2.10, over disjoint alphabets. Then $L' = K \cup L$ is repetition-free $(2, 1)$ -PHRS by Theorem 4.11, but is neither E(T)0L nor MCF since if it were, then $L' \cap A^*$ and $L' \cap B^*$ would be both 2-MCF and E0L since these language families are closed under rational intersection (Theorems 2.1 and 2.9), but Theorem 2.10 tells us that $L' \cap A^*$ is not ET0L and $L' \cap B^*$ is not MCF. \square

We now show closure under rational intersection, inspired by the proof of Theorem V.1.7(iv) of [4]:

Theorem 4.13 (PHRS Closed Under Rational Intersection). Let $L \subseteq A^*$ be a (repetition-free) (k, l) -PHRS language and $K \subseteq B^*$ be a regular language, for any $k \geq 2, l \geq 1$. Then $L \cap K$ is a (repetition-free) (k, l) -PHRS language.

Proof. By Theorem 4.2, it will suffice to show that when L is a (repetition-free) $(k, 1)$ -PHRS language, $L \cap K$ is a (repetition-free) $(k, 1)$ -PHRS language, and when L is a (repetition-free) $(k, 2)$ -PHRS language, $L \cap K$ is a (repetition-free) k -PHRS language. Our proof will set up some commonalities for both cases, and then we will analyse the cases $l = 1$ and $l = 2$ separately.

There must exist a (repetition-free) (k, l) -PHR grammar $\mathcal{G} = ((\Sigma, \text{type}), A, S, \{T_1, \dots, T_l\}, S)$ such that $L(\mathcal{G})$ is a string graph language and $\text{STR}(L(\mathcal{G})) = L \setminus \{\epsilon\}$. We assume without loss of generality that $(\Sigma \setminus A) \cap B = \emptyset$ and that \mathcal{G} is of the form specified by Lemma 3.15. There must also be a deterministic full FSA $\mathcal{M} = (Q, B, \delta, p, F)$ such that $L(\mathcal{M}) = K$.

We introduce the intermediate notion of a hypergraph with node labels because we want to label the nodes by states of \mathcal{M} . A node labelled hypergraph over (\mathcal{C}, Q) is a pair (H, l) where H is a hypergraph over $\mathcal{C} = (\Sigma, \text{type})$ and l is a function $V_H \rightarrow Q$. Notice that any such node labelled hypergraph can be encoded as a hypergraph over $(\Delta, \text{type}'|_{\Delta})$: for each $e \in E_H$, the new hyperedge labelling function is defined by sending e to $(\text{lab}_H(e), q_1, \dots, q_i)$ where $i = \text{type}_H(e)$ and $q_j = l(\text{att}_H(e)(j))$ for $1 \leq j \leq i$. Call this injective encoding function enc . Next, given a type t hypergraph H over \mathcal{C} and a sequence $\sigma : \underline{t} \rightarrow Q$, define $\text{CHOICES}_Q(H, \sigma) = \{\text{enc}((H, l)) \mid l : V_H \rightarrow Q, l \circ \text{ext}_H = \sigma\}$. Given a rule (L, R) over \mathcal{C} , define $\text{AUGMENT}_Q(L, R) = \{((L, \sigma(1), \dots, \sigma(t)), H) \mid t = \text{type}(L), \sigma : \underline{t} \rightarrow Q, H \in \text{CHOICES}_Q(R, \sigma)\}$.

We now handle the case $l = 1$. We will construct a (repetition-free) $(k, 1)$ -PHR grammar $\mathcal{G}' = ((\Sigma', \text{type}'), A \cap B, \{T'\}, S)$ such that $L(\mathcal{G}')$ is a string graph language and $\text{STR}(L(\mathcal{G}')) = (L \cap K) \setminus \{\epsilon\}$, thus proving that $L \cap K$ is a (repetition-free) $(k, 1)$ -PHRS language. Let $\Delta = (\bigcup_{0 \leq i \leq k} \{\text{type}^{-1}(\{i\}) \times Q^i\}) \cup \{F_0, \dots, F_k\}$, $\Sigma' = \Delta \cup \{S\} \cup (A \cap B) \cup \{F_0, \dots, F_k\}$, $\text{type}'((X, q_1, \dots, q_i)) = \text{type}(X)$ for all $(X, q_1, \dots, q_i) \in \Delta$, $\text{type}'(S) = 2$, $\text{type}'(X) = 2$ for all $X \in A \cap B$, and $\text{type}'(F_i) = i$ for all $0 \leq i \leq k$. Define $T' = \mathcal{F} \oplus (T'_1 \cup T'_2 \cup T'_3)$ where:

1. $T'_1 = \{((X, q_1, q_2), Y^\bullet) \mid X \in A \cap B, \delta(q_1, X) = q_2\}$;
2. $T'_2 = \bigcup_{(L, R) \in T} \text{AUGMENT}_Q(L, R)$;
3. $T'_3 = \{(S, (S, p, q)^\bullet) \mid q \in F\}$;
4. $\mathcal{F} = \{(X, F_{\text{type}'(X)}^\bullet) \mid X \in \Sigma'\}$.

It is not hard to check this grammar will produce a terminal string graph $(x_1 x_2 \cdots x_m)^\bullet$ if and only if the previous hypergraph in the derivation was a string graph of the form $((x_1, q_1, q_2)(x_2, q_2, q_3) \cdots (x_m, q_m, q_{m+1}))^\bullet$ and $\delta(q_i, x_i) = q_{i+1}$ for $1 \leq i \leq m$, $q_1 = p$, and $q_{m+1} \in F$, that is, traced out an accepting path in the FSA \mathcal{M} , having simulated \mathcal{G} .

Finally, we handle the case $l = 2$. We will construct a (repetition-free) k -PHR grammar $\mathcal{G}' = ((\Sigma', \text{type}'), A \cap B, \{T'_0, T'_1, T'_2\}, S)$ such that $L(\mathcal{G}')$ is a string graph language and $\text{STR}(L(\mathcal{G}')) = (L \cap K) \setminus \{\epsilon\}$, thus proving that $L \cap K$ is a (repetition-free) k -PHRS language. Let Δ and (Σ, type') be as before. Define (for $i = 1, 2$):

1. $T'_0 = \mathcal{F} \oplus \{((X, q_1, q_2), Y^\bullet) \mid X \in A \cap B, \delta(q_1, X) = q_2\}$;
2. $T'_i = \mathcal{R} \oplus (\bigcup_{(L, R) \in T_i} \text{AUGMENT}_Q(L, R) \cup \{(S, (S, p, q)^\bullet) \mid q \in F\})$;

where $\mathcal{F} = \{(X, F_{\text{type}'(X)}^\bullet) \mid X \in \Sigma'\}$ and $\mathcal{R} = \{(X, X^\bullet) \mid X \in \Sigma'\}$. Clearly progress in derivations is made when the tables T'_1 and T'_2 are used, and application of the table T'_0 will produce a terminal string graph $(x_1 x_2 \cdots x_m)^\bullet$ if and only if the previous hypergraph was a string graph of the form $((x_1, q_1, q_2)(x_2, q_2, q_3) \cdots (x_m, q_m, q_{m+1}))^\bullet$ and $\delta(q_i, x_i) = q_{i+1}$ for $1 \leq i \leq m$, $q_1 = p$, and $q_{m+1} \in F$, as before. \square

Finally, we show closure under inverse homomorphisms, via the following general result:

Lemma 4.14. Let \mathcal{F} be a family of string languages which is closed under rational substitution and rational intersection. Let $L \subseteq A^*$ be an \mathcal{F} language and $\varphi : B^* \rightarrow A^*$ a homomorphism. Then $\varphi^{-1}(L)$ is an \mathcal{F} language too.

Proof. Let \bar{B} be a copy of B such that $(A \cup B) \cap \bar{B} = \emptyset$, and let $\bar{\cdot} : B \rightarrow \bar{B}$ identify each $b \in B$ with its copy $\bar{b} \in \bar{B}$. For each $a \in A$, define the regular language $L_a = \{w_1 a w_2 \mid w_1, w_2 \in \bar{B}^*\} \subseteq (A \cup \bar{B})^*$ and the rational substitution h on A by $a \mapsto L_a$. Also define $K = \bigcup_{n \in \mathbb{N}} \{\varphi(x_1) \bar{x}_1 \varphi(x_2) \bar{x}_2 \cdots \varphi(x_n) \bar{x}_n \mid x_1, x_2, \dots, x_n \in B\}$ and the homomorphism $\psi : (A \cup \bar{B})^* \rightarrow B^*$ by $\psi(a) = \epsilon$ for each $a \in A$ and $\psi(\bar{b}) = b$ for each $b \in B$.

Notice $h(L) \cap K = \bigcup_{n \in \mathbb{N}} \{\varphi(x_1) \bar{x}_1 \varphi(x_2) \bar{x}_2 \cdots \varphi(x_n) \bar{x}_n \mid x_1, \dots, x_n \in B \text{ and } \varphi(x_1) \varphi(x_2) \cdots \varphi(x_n) \in L\}$, so we have $\varphi^{-1}(L) = \psi(h(L) \cap K)$. Now, $h(L)$ is an \mathcal{F} language since \mathcal{F} is closed under rational substitution, $h(L) \cap K$ is an \mathcal{F} language since \mathcal{F} is closed under rational intersection, and $\psi(h(L) \cap K)$ is an \mathcal{F} language since \mathcal{F} is closed under homomorphisms (a special case of rational substitution). Thus, $\varphi^{-1}(L)$ is an \mathcal{F} language, as required. \square

Theorem 4.15 (PHRS Closed Under Inverse Homomorphisms). For all $k \geq 2$, \mathcal{PHRS}_k is closed under inverse homomorphisms.

Proof. The result follows from Theorems 4.7 and 4.13 and Lemma 4.14. \square

4.3. Group Word Problem Closure Properties

Since the family of k -PHRS languages is a full AFL for any $k \geq 2$, it satisfies the following important properties:

Theorem 4.16 (WP Independent Of Presentation [36]). Let \mathcal{F} be a family of string languages which is closed under inverse homomorphisms, and let $\langle X \mid R \rangle$ be a presentation of a group G such that $\text{WP}_X(G)$ is an \mathcal{F} language. Then all presentations $\langle X' \mid R' \rangle$ of G are such that $\text{WP}_{X'}(G)$ is an \mathcal{F} language.

Theorem 4.17 (WP Subgroup and Supergroup Closure [16]). Let \mathcal{F} be a full AFL and G be a group with word problem in \mathcal{F} . Then every finitely generated subgroup and every finite index supergroup of G has word problem in \mathcal{F} .

In 2019, Kropholler and Spriano showed that a graph of groups with vertex groups with MCF word problem and edge groups finite, yields a group with an MCF word problem [37]. A special case of this construction is a free product of groups. We now show that a free product of groups with (repetition-free) PHRS word problems is a group with a (repetition-free) PHRS word problem. Our strategy is entirely different to Kropholler and Spriano's approach, which relied on Denkinger's automata characterisation of MCF languages (Theorem 2.2).

The following easy lemma, where presentations of groups are written as monoid presentations, gives us a recursive description of the word problem of free products, enabling us to prove Theorem 4.19.

Lemma 4.18. Let G_1, G_2 be finitely generated groups over disjoint alphabets $A_1 = \{a_1, \dots, a_n\}$, $A_2 = \{b_1, \dots, b_m\}$, respectively. If $X = A_1 \cup A_2$ and $L_i = \text{WP}_{A_i}(G_i)$ for $i = 1, 2$, then $\text{WP}_X(G_1 * G_2)$ is the smallest set L such that $\epsilon \in L$ and $\forall i \in \{1, 2\}, \forall w \in L_i, \forall u, v \in X^*, uv \in L \Rightarrow u w v \in L$.

Theorem 4.19 (WP Free Product Closure). Let \mathcal{F} be a family of string languages containing all finite languages, closed under union and concatenation, and closed under nested iterated substitution. Then if G_1, G_2 are groups with presentations admitting a \mathcal{F} word problem, $G_1 * G_2$ has a presentation admitting a \mathcal{F} word problem.

Proof. Let A_1, A_2, X, L_1, L_2, L be as in Lemma 4.18, then it is immediate that iterated application of the nested non-erasing \mathcal{F} -substitution h of strings on $A_1 \cup A_2$, defined by $h(a_i) = \{a_i\} \cup a_i L_1 \cup L_1 a_i$ and $h(b_j) = \{b_j\} \cup b_j L_2 \cup L_2 b_j$ for all $i \in \underline{n}, j \in \underline{m}$, to L , gives us exactly $\text{WP}_X(G_1 * G_2)$. The result then follows from the assumed closure properties. \square

4.4. The Membership Problem

We show that the universal membership problem for PHRS grammars is decidable. That is, there is an algorithm that, when given a PHR grammar \mathcal{G} generating a string graph language and a string w , can decide if $w \in \text{STR}(\text{L}(\mathcal{G}))$.

Theorem 4.20 (Decidable PHRS Membership). The following problem is decidable, with an explicit algorithm:

- Instance:** A PHR grammar $\mathcal{G} = (\mathcal{C}, A, S, \mathcal{T})$ generating a string graph language and a string $w \in A^*$.
Question: Is $w \in \text{STR}(\text{L}(\mathcal{G}))$?

Proof. Theorem 4.13 tells us how to turn the PHR grammar \mathcal{G} and the regular language $\{w\}$ into a PHR grammar \mathcal{G}' generating a string graph language such that $\text{STR}(\text{L}(\mathcal{G}')) = \text{STR}(\text{L}(\mathcal{G})) \cap \{w\}$. Then:

$$\begin{aligned} w \notin \text{STR}(\text{L}(\mathcal{G})) & \\ \Leftrightarrow \text{STR}(\text{L}(\mathcal{G})) \cap \{w\} = \emptyset & \\ \Leftrightarrow \text{STR}(\text{L}(\mathcal{G}')) = \emptyset & \\ \Leftrightarrow \text{L}(\mathcal{G}') = \emptyset. & \end{aligned}$$

Finally, Theorem 3.25 tells us how to decide emptiness of $\text{L}(\mathcal{G}')$. \square

Corollary 4.21 (PHRS is Recursive). $\text{PHRS} \subseteq \text{REC}$.

Proof. Inclusion follows from decidable membership (Theorem 4.20). Strictness from decidable emptiness (Theorem 3.25). \square

4.5. Weak-Coded PHRS Languages

Theorem 2.12 tells us that the string generational power of HR grammars is not restricted by requiring rules to be repetition-free. It is not clear if a similar result holds in the parallel replacement setting. If it turns out that there is no

such result, there is still a middle-ground where one can obtain all of the closure properties we have shown in this section, but without allowing merging of nodes by derivations:

Definition 4.22 (WPHR String Language). Call the below equivalent families the repetition-free weak-coded (k, l) -PHR string $((k, l)$ -WPHRS) languages:

1. The family of string languages generated by repetition-free (k, l) -PHR grammars under the image of some weak coding.
2. The family of string languages generated by repetition-free (k, l) -PHR grammars with a special type 2 label **empty**, interpreted as the empty string by STR.

For $k \geq 2$ and $l \geq 1$, we define $WPHRS_{k,l}^{\text{rf}}$, $WPHRS_k^{\text{rf}}$, and $WPHRS^{\text{rf}}$ in the obvious way.

Theorem 4.23 (WPHRS Hierarchy Position). For all $k \geq 2$, we have:

1. $WPHRS_{k,2}^{\text{rf}} = WPHRS_k^{\text{rf}}$;
2. $\mathcal{EOL} \subseteq \mathcal{PHRS}_{k,1}^{\text{rf}} \subseteq WPHRS_{k,1}^{\text{rf}} \subseteq \mathcal{PHRS}_{k,1}$
3. $\mathcal{ETOL} \subseteq \mathcal{PHRS}_k^{\text{rf}} \subseteq WPHRS_k^{\text{rf}} \subseteq \mathcal{PHRS}_k$.

Moreover:

1. $\mathcal{EOL} = \mathcal{PHRS}_{2,1}^{\text{rf}} = WPHRS_{2,1}^{\text{rf}} = \mathcal{PHRS}_{2,1}$;
2. $\mathcal{ETOL} = \mathcal{PHRS}_2^{\text{rf}} = WPHRS_2^{\text{rf}} = \mathcal{PHRS}_2$.

Proof. The first part follows from Theorem 3.10. For the second part, $\mathcal{EOL} \subseteq \mathcal{PHRS}_{k,1}^{\text{rf}}$ follows from Theorem 4.4, $\mathcal{PHRS}_{k,1}^{\text{rf}} \subseteq WPHRS_{k,1}^{\text{rf}}$ follows by definition using the identity coding, and $WPHRS_{k,1}^{\text{rf}} \subseteq \mathcal{PHRS}_{k,1}$ follows using the fact that $\mathcal{PHRS}_{k,1}$ is closed under weak codings (Theorem 4.8). The third part follows in a similar way to the second. The final two statements follow from Theorem 4.4 and the earlier inclusions we established in this proof. \square

Using the results and proofs from Subsection 4.2, it is not too difficult to see that $WPHRS_k^{\text{rf}}$ is a hyper-algebraically closed full AFL, for all $k \geq 2$. We finish this subsection by giving the fine details of the closure properties.

Theorem 4.24 (WPHRS Closed Under Substitutions). Let $L \subseteq A^*$. Then:

1. for any $k \geq 2$, if L is a k -HRS language and h is a $WPHRS_{k,1}^{\text{rf}}$ -substitution on A , then $h(L)$ is a repetition-free $(k, 1)$ -WPHRS language;
2. for any $k \geq 2$, if L is a repetition-free $(k, 1)$ -WPHRS language and h is a finite substitution on A , then $h(L)$ is a repetition-free $(k, 1)$ -WPHRS language;
3. for any $k \geq 2$, if L is a repetition-free k -PHRS language and h is a $WPHRS_k^{\text{rf}}$ -substitution on A , then $h(L)$ is a repetition-free k -WPHRS language.

Corollary 4.25 (WPHRS Closed Under Homomorphisms). Let $L \subseteq A^*$ be a repetition-free (k, l) -WPHRS language for any $k \geq 2, l \geq 1$ and $\varphi : A^* \rightarrow B^*$ be a homomorphism. Then $\varphi(L)$ is a repetition-free (k, l) -WPHRS language.

Theorem 4.26 (WPHRS Hyper-Algebraically Closed). Let $L \subseteq A^*$. Then, for any $k \geq 2$, if L is a repetition-free k -WPHRS language and H is a finite set of $\text{WPHRS}_k^{\text{rf}}$ -substitutions on A , then $\text{ITER}_H(L)$ is a repetition-free k -WPHRS language.

Theorem 4.27 (WPHRS Closed Under Rational Operations). For all $k \geq 2, l \geq 1$, $\text{WPHRS}_k^{\text{rf}}$ is closed under rational operations.

Theorem 4.28 (WPHRS Closed Under Rational Intersection). For all $k \geq 2, l \geq 1$, $\text{WPHRS}_k^{\text{rf}}$ is closed under rational intersection.

Theorem 4.29 (WPHRS Closed Under Inverse Homomorphisms). For all $k \geq 2, l \geq 1$, $\text{WPHRS}_k^{\text{rf}}$ is closed under inverse homomorphisms.

5. Conclusion and Future Work

We have shown some foundational properties of parallel hyperedge replacement grammars, with a focus on string generational power, showing that the family of PHRS languages is a hyper-algebraically closed super AFL, containing all MCF and ETOL languages. In Subsection 4.5 we discuss the possible gap between the string generative power of repetition-free and not necessarily repetition-free PHR grammars. We conjecture that there is no gap:

Conjecture 5.1 (PHR String Generational Power). For all $k \geq 2, l \geq 1$, we have the following: $\text{PHRS}_{k,l}^{\text{rf}} = \text{WPHRS}_{k,l}^{\text{rf}} = \text{PHRS}_{k,l}$.

It remains future work to show that the family of PHRS languages is a strict subclass of the context-sensitive languages. We conjecture this to be true, and we also conjecture that only even increments in order increase string generative power. Figure 13 summarises the formal language closure properties we know, and Figures 14, 15, 16 and 17 summarise the key language hierarchies, where $\text{PHR}_{\geq l} = \bigcup_{k \geq 2} \text{PHR}_{k,l}$ and $\text{PHRS}_{\geq l} = \bigcup_{k \geq 2} \text{PHRS}_{k,l}$ for any $l \geq 1$.

Conjecture 5.2 (CS Generalises PHRS). $\text{PHRS} \subseteq \text{CS}$.

Conjecture 5.3 (PHRS Grouping). For all $k \geq 1$, $\text{PHRS}_{2k} = \text{PHRS}_{2k+1}$.

Because PHRS is closed under inverse homomorphisms, we know that the property of having a PHRS word problem is independent of the presentation. We have additionally shown that PHRS groups are closed under free product. We also conjecture the following, which has a wide-reaching corollary:

Conjecture 5.4 (PHRS WP Double Torus). The fundamental group of the double torus admits a PHRS word problem which is neither an MCF nor ETOL language.

Corollary 5.5. If Conjecture 5.4 is true, then the word problem of any surface group is PHRS.

Proof. By a *surface* here, we mean a closed, connected, orientable, 2-manifold, and by a *surface group*, we mean the fundamental group of a surface. Any surface always has a finite genus. The genus 0 surface (the sphere) gives us the trivial group, and 1 (the torus), \mathbb{Z}^2 (see for example [38]). We know both of these groups are regular, 2-MCF [15], respectively, so certainly PHRS (Theorem 4.6). For higher genres, it follows from the Fundamental Theorem of Covering Spaces (Theorem 1.38 of [39]) that the fundamental group appears as a finitely generated subgroup of the fundamental group of a genus 2 surface such as a double torus. Because \mathcal{PHRS} is a full AFL, if the double torus has fundamental group with PHRS word problem, then all its finitely generated subgroups do too (Theorem 4.17). \square

Highly related to the word problem is the consideration of sets of solutions of more general equations over groups or other structures. It is a recent result that solutions set (of fixed normal forms) of finite systems of equations in hyperbolic groups are EDTOL languages [40]. We are yet to consider deterministic parallel hyperedge replacement, but it may be possible to establish that other classes of groups have solution sets that are deterministic parallel hyperedge replacement string languages.

Recall from Theorem 2.7 that the ETOL languages are the smallest hyper-algebraically closed super AFL. It is possible that the PHRS languages could be the smallest hyper-algebraically closed super AFL containing the MCF languages. Other more general future work would include investigating both the tree and graph generational power of PHR grammars, and investigating more decidability and complexity results for basic problems relating to PHR grammars. We do not know if the finiteness problem is decidable, but we conjecture that this can be decided, in general. If Conjecture 5.1 fails, it will also be worthwhile to fill out some of the question marks in Figure 13.

Conjecture 5.6 (Decidable PHR Finiteness). The following problem is decidable, with an explicit algorithm:

Instance: A PHR grammar $\mathcal{G} = (\mathcal{C}, A, S, \mathcal{T})$.

Question: Does $L(\mathcal{G})$ contain only finitely many non-isomorphic hyper-graphs?

Acknowledgements

I should like to thank Detlef Plump for introducing me to graph transformation and teaching me to write papers for this audience, my supervisors Sarah Rees and Andrew Duncan for their guidance, Annegret Habel and Meng-Che Ho for their helpful email discussions regarding hyperedge replacement and surface groups, respectively, and Murray Elder for introducing me to MCF languages. I am also grateful to the anonymous reviewers for their comments on the earlier TERMGRAPH workshop proceedings version of this paper.

Operation/Family	HRS_k^{sf} $= HRS_k$	$PHRS_{k,1}^{sf}$	$WPHRS_{k,1}^{sf}$	$PHRS_{k,1}$	$PHRS_k^{sf}$	$WPHRS_k^{sf}$	$PHRS_k$
Rational Operations	✓	✓	✓	✓	✓	✓	✓
Rational Intersection	✓	✓	✓	✓	✓	✓	✓
Inverse Homomorphisms	✓	?	?	?	?	✓	✓
Non-Erasing Homomorphisms	✓	✓	✓	✓	✓	✓	✓
Arbitrary Homomorphisms	✓	?	✓	✓	?	✓	✓
Non-Erasing Finite Substitutions	✓	✓	✓	✓	✓	✓	✓
Arbitrary Finite Substitutions	✓	?	✓	✓	?	✓	✓
Non-Erasing Substitutions	✓	?	?	?	✓	✓	✓
Arbitrary Substitutions	✓	?	?	?	?	✓	✓
Iterated Nested Non-Erasing Substitutions	✓	?	?	?	✓	✓	✓
Iterated Nested Arbitrary Substitutions	✓	?	?	?	?	✓	✓
Iterated Non-Erasing Substitutions	✗	?	?	?	✓	✓	✓
Iterated Arbitrary Substitutions	✗	?	?	?	?	✓	✓
Hyper-Algebraic Closure	✗	✗	✗	✗	?	✓	✓

Figure 13: Summary of formal language closure properties ($k \geq 2$)

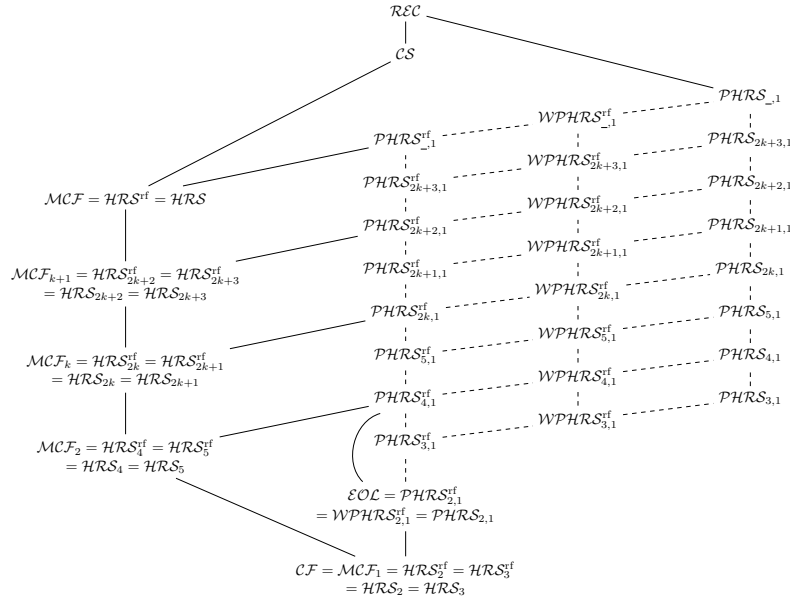


Figure 14: Proved single table string language hierarchy ($k \geq 3$)

References

- [1] A. Lindenmayer, Mathematical models for cellular interactions in development I. filaments with one-sided inputs, Journal of Theoretical Biology 18 (3) (1968) 280–299. doi:10.1016/0022-5193(68)90079-9.
- [2] A. Lindenmayer, Mathematical models for cellular interactions in development II. simple and branching filaments with two-sided inputs, Journal of Theoretical Biology 18 (3) (1968) 300–315. doi:10.1016/0022-5193(68)90080-5.
- [3] A. Lindenmayer, Developmental systems without cellular interactions, their languages and grammars, Journal of Theoretical Biology 30 (3) (1971) 455–484. doi:10.1016/0022-5193(71)90002-6.

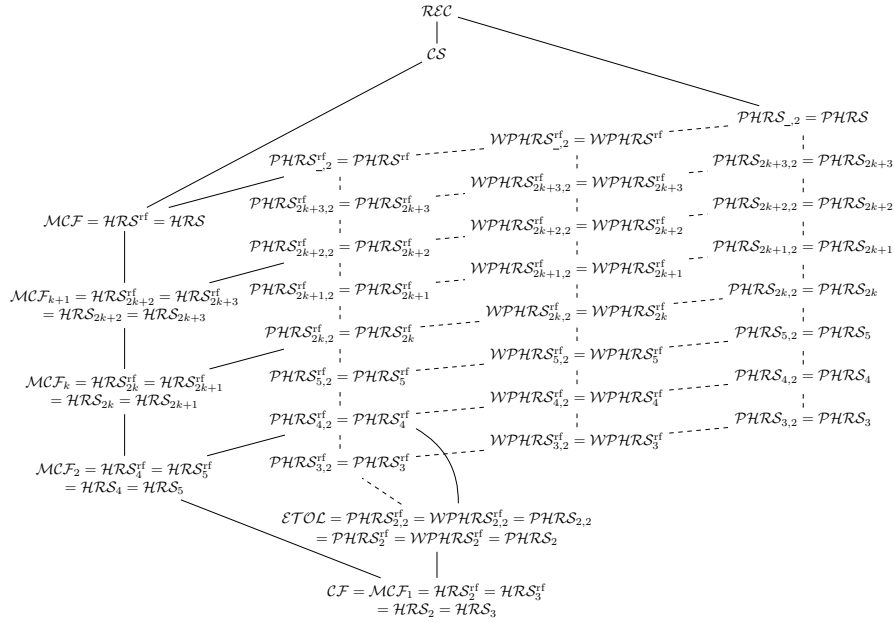


Figure 15: Proved multiple tables string language hierarchy ($k \geq 3$)

- [4] G. Rozenberg, A. Salomaa, The Mathematical Theory of L Systems, Vol. 90 of Pure and Applied Mathematics, Academic Press, 1980.
- [5] L. Kari, G. Rozenberg, A. Salomaa, L Systems, Springer, 1997, pp. 253–328. doi:10.1007/978-3-642-59136-5_5.
- [6] J. Feder, Plex languages, Information Sciences 3 (3) (1971) 225–241. doi:10.1016/S0020-0255(71)80008-7.
- [7] T. Pavlidis, Linear and context-free graph grammars, Journal of the ACM 19 (1) (1972) 11–22. doi:10.1145/321679.321682.
- [8] H. Seki, T. Matsumura, M. Fujii, T. Kasami, On multiple context-free grammars, Theoretical Computer Science 88 (2) (1991) 191–229. doi:10.1016/0304-3975(91)90374-B.
- [9] J. Engelfriet, L. Heyker, The string generating power of context-free hypergraph grammars, Journal of Computer and System Sciences 43 (2) (1991) 328–360. doi:10.1016/0022-0000(91)90018-Z.
- [10] D. Weir, Linear context-free rewriting systems and deterministic tree-walking transducers, in: H. Thompson (Ed.), Proc. 30th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 1992, pp. 136–143. doi:10.3115/981967.981985.

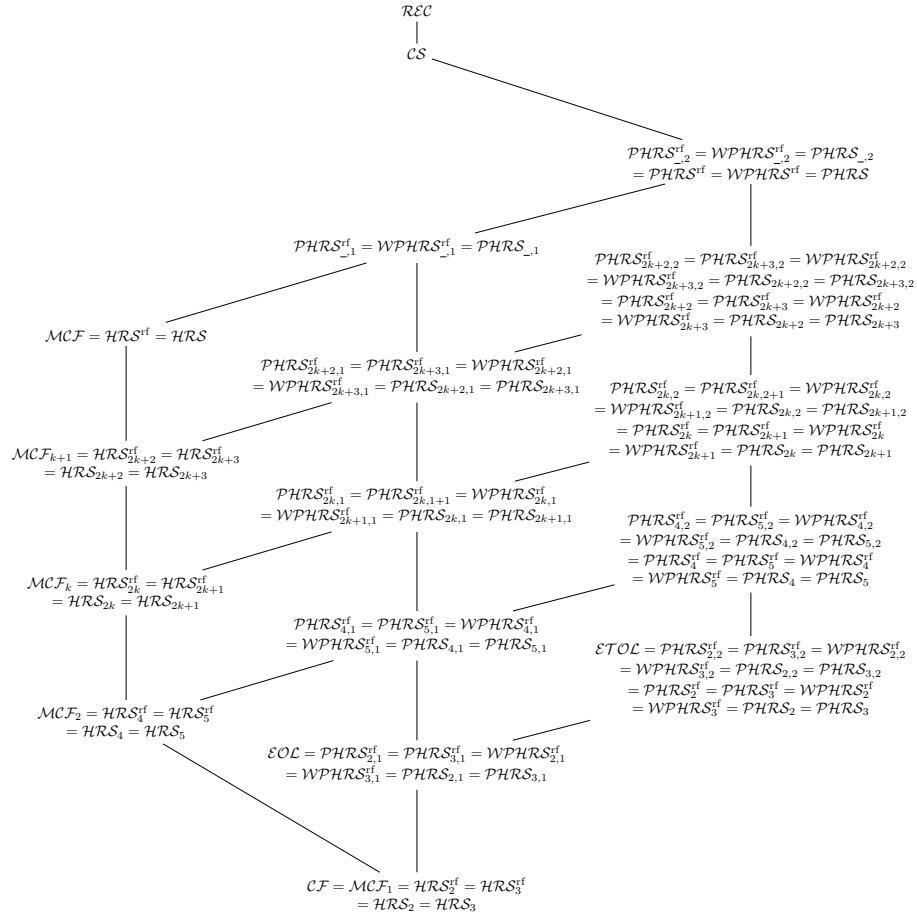


Figure 16: Conjectured string language hierarchy ($k \geq 3$)

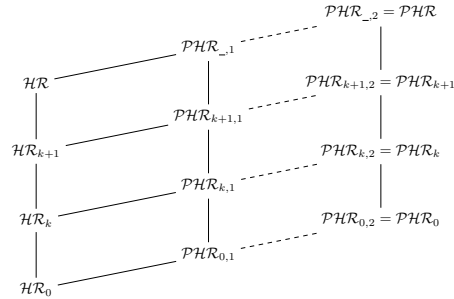


Figure 17: Proved hypergraph language hierarchy ($k \geq 1$)

- [11] P. Novikov, Über die algorithmische unentscheidbarkeit des wortproblems in der gruppentheorie, *Trudy Matematicheskogo Instituta imeni V.A. Steklova* 44 (1955) 1–143.
- [12] A. Anisimov, Group languages, *Cybernetics* 7 (1971) 594–601. doi:10.1007/BF01071030.
- [13] D. Muller, P. Schupp, Groups, the theory of ends, and context-free languages, *Journal of Computer and System Sciences* 26 (3) (1983) 295–310. doi:10.1016/0022-0000(83)90003-X.
- [14] S. Salvati, Mix is a 2-mcfl and the word problem in \mathbb{Z}^2 is captured by the io and the oi hierarchies, *Journal of Computer and System Sciences* 81 (7) (2015) 1252–1277. doi:10.1016/j.jcss.2015.03.004.
- [15] M.-C. Ho, The word problem of \mathbb{Z}^n is a multiple context-free language, *Groups Complexity Cryptology* 10 (1) (2018) 9–15. doi:10.1515/gcc-2018-0003.
- [16] R. Gilman, R. Kropholler, S. Schleimer, Groups whose word problems are not semilinear, *Groups Complexity Cryptology* 10 (2) (2018) 53–62. doi:10.1515/gcc-2018-0010.
- [17] A. Aho, Indexed grammars – an extension of context-free grammars, *Journal of the ACM* 15 (4) (1968) 647–671. doi:10.1145/321479.321488.
- [18] R. Gilman, M. Shapiro, On groups whose word problem is solved by a nested stack automaton (1998). arXiv:math/9812028. URL <https://arxiv.org/abs/math/9812028>
- [19] L. Ciobanu, M. Elder, M. Ferov, Applications of L systems to group theory, *International Journal of Algebra and Computation* 28 (2) (2018) 309–329. doi:10.1142/S0218196718500145.
- [20] A. Habel, *Hyperedge Replacement: Grammars and Languages*, Vol. 643 of *Lecture Notes in Computer Science*, Springer, 1992. doi:10.1007/BFb0013875.
- [21] H.-J. Kreowski, *Parallel Hyperedge Replacement*, Springer, 1992, pp. 271–282. doi:10.1007/978-3-642-58117-5_17.
- [22] H.-J. Kreowski, Five facets of hyperedge replacement beyond context-freeness, in: Z. Ésik (Ed.), *Proc. 9th International Conference on Fundamentals of Computation Theory (FCT 1993)*, Vol. 710 of *Lecture Notes in Computer Science*, Springer, 1993, pp. 69–86. doi:10.1007/3-540-57163-9_5.
- [23] G. Campbell, Parallel hyperedge replacement string languages, in: P. Bahr (Ed.), *Proc. 11th International Workshop on Computing with Terms and*

Graphs (TERMGRAPH 2020), Electronic Proceedings in Theoretical Computer Science, Open Publishing Association, 2021, to appear.
URL <https://cdn.gjcampbell.co.uk/2021/PHRS-Languages.pdf>

- [24] J. Hopcroft, R. Motwani, J. Ullman, Introduction to Automata Theory, Languages, and Computation, 3rd Edition, Addison-Wesley, 2006.
- [25] J. Král, A modification of a substitution theorem and some necessary and sufficient conditions for sets to be context-free, *Mathematical systems theory* 4 (1970) 129–139. doi:10.1007/BF01691097.
- [26] A. Aho, J. Ullman, Translations on a context free grammar, *Information and Control* 19 (5) (1972) 439–475. doi:10.1016/S0019-9958(71)90706-6.
- [27] V. Shanker, D. Weir, A. Joshi, Characterizing structural descriptions produced by various grammatical formalisms, in: C. Sidner (Ed.), Proc. 25th Annual Meeting on Association for Computational Linguistics (ACL '87), Association for Computational Linguistics, 1987, pp. 104–111. doi:10.3115/981175.981190.
- [28] T. Denking, An automata characterisation for multiple context-free languages, in: S. Brlek, C. Reutenauer (Eds.), Proc. International Conference on Developments in Language Theory (DLT 2016), Vol. 9840 of Lecture Notes in Computer Science, Springer, 2016, pp. 138–150. doi:10.1007/978-3-662-53132-7_12.
- [29] R. Parikh, On context-free languages, *Journal of the ACM* 13 (4) (1966) 570–581. doi:10.1145/321356.321364.
- [30] G. Rozenberg, A. Salomaa, L Systems, Vol. 15 of Lecture Notes in Computer Science, Springer, 1974. doi:10.1007/3-540-06867-8.
- [31] A. Ehrenfeucht, G. Rozenberg, On some context-free languages that are not deterministic ET0L languages, *R.A.I.R.O. Informatique théorique* 11 (4) (1977) 273–291. doi:10.1051/ita/1977110402731.
- [32] T. Nishida, S. Seki, Grouped partial ET0L systems and parallel multiple context-free grammars, *Theoretical Computer Science* 246 (1–2) (2000) 131–150. doi:10.1016/S0304-3975(99)00076-6.
- [33] F. Drewes, H.-J. Kreowski, A. Habel, Hyperedge Replacement Graph Grammars, World Scientific, 1997, pp. 95–162. doi:10.1142/9789812384720_0002.
- [34] M. Nielsen, EOL systems with control devices, *Acta Informatica* 4 (1975) 373–386. doi:10.1007/BF00289618.
- [35] P. Asveld, Controlled iteration grammars and full hyper-AFL's, *Information and Control* 34 (3) (1977) 248–269. doi:10.1016/S0019-9958(77)90308-4.

- [36] T. Herbst, R. Thomas, Group presentations, formal languages and characterizations of one-counter groups, *Theoretical Computer Science* 112 (2) (1993) 187–213. doi:10.1016/0304-3975(93)90018-0.
- [37] R. Kropholler, D. Spriano, Closure properties in the class of multiple context-free groups, *Groups Complexity Cryptology* 11 (1) (2019) 1–15. doi:10.1515/gcc-2019-2004.
- [38] W. Massey, *Algebraic Topology: An Introduction*, Vol. 56 of Graduate Texts in Mathematics, Springer, 1977.
- [39] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.
- [40] L. Ciobanu, M. Elder, Solutions sets to systems of equations in hyperbolic groups are EDT0L in PSPACE, in: C. Baier, I. Chatzigiannakis, P. Flocchini, S. Leonardi (Eds.), *Proc. 46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, Vol. 132 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2019, pp. 110:1–110:15. doi:10.4230/LIPIcs.ICALP.2019.110.