On Infinite Words Determined by L Systems

Tim Smith

College of Computer and Information Science, Northeastern University
Boston, MA 02115, USA
smithtim@ccs.neu.edu

Abstract. A deterministic L system generates an infinite word α if each word in its derivation sequence is a prefix of the next, yielding α as a limit. We generalize this notion to arbitrary L systems via the concept of prefix languages. A prefix language is a language L such that for all $x, y \in L$, x is a prefix of y or y is a prefix of x. Every infinite prefix language determines an infinite word. Where C is a class of L systems (e.g. 0L, DT0L), we denote by $\omega(C)$ the class of infinite words determined by the prefix languages in C. This allows us to speak of infinite 0L words, infinite DT0L words, etc. We categorize the infinite words determined by a variety of L systems, showing that the whole hierarchy collapses to just three distinct classes of infinite words: ω(PD0L), ω(D0L), and ω(CD0L).

1 Introduction

L systems are parallel rewriting systems which were originally introduced to model growth in simple multicellular organisms. With applications in biological modelling, fractal generation, and artificial life, L systems have given rise to a rich body of research [11, 9]. L systems can be restricted and generalized in various ways, yielding a hierarchy of language classes.

The simplest L systems are D0L systems (deterministic Lindenmayer systems with 0 symbols of context), in which a morphism is successively applied to a start string or "axiom". The resulting sequence of words comprises the language of the system. If the morphism is prolongable on the start string, then each word in the derivation sequence will be a prefix of the next, yielding an infinite word as a limit. An infinite word obtained in this way is called an infinite D0L word.

Two well-studied generalizations of D0L systems are 0L systems, which introduce nondeterminism by changing the morphism to a finite substitution, and DT0L systems, in which the morphism is replaced by a set of morphisms or "tables". In each case, there is no longer just one possible derivation sequence; rather, there are many possible derivations, depending on which letter substitutions or tables are chosen at each step. This raises the question of under what conditions such a system can be said to determine an infinite word.

We answer this question with the concept of a prefix language. A prefix language is a language L such that for all $x, y \in L$, x is a prefix of y or y is a prefix of x. Every infinite prefix language determines an infinite word. Where C is a class of L systems (e.g. 0L, DT0L), we denote by $\omega(C)$ the class of infinite words determined by the prefix languages in C. This allows us to speak of infinite 0L words, infinite DT0L words, etc.

With this notion in place, we categorize the infinite words determined by a variety of L systems. We consider four production features (D,P,F,T) and five filtering features (E,C,N,W,H). Each production feature may be present or absent, and at most one filtering feature may be present, giving a total of $2^4 \cdot 6 = 96$ classes of L systems. We show that this whole hierarchy collapses to just three classes of infinite words: ω(PD0L), ω(D0L), and ω(CD0L). Our results appear in Figure 1.

The inclusions among these three classes are proper, giving $\omega(\text{PD0L}) \subset \omega(\text{D0L}) \subset \omega(\text{CD0L})$. The class ω(CD0L) contains exactly the morphic words, while ω(D0L) properly contains the pure morphic words.

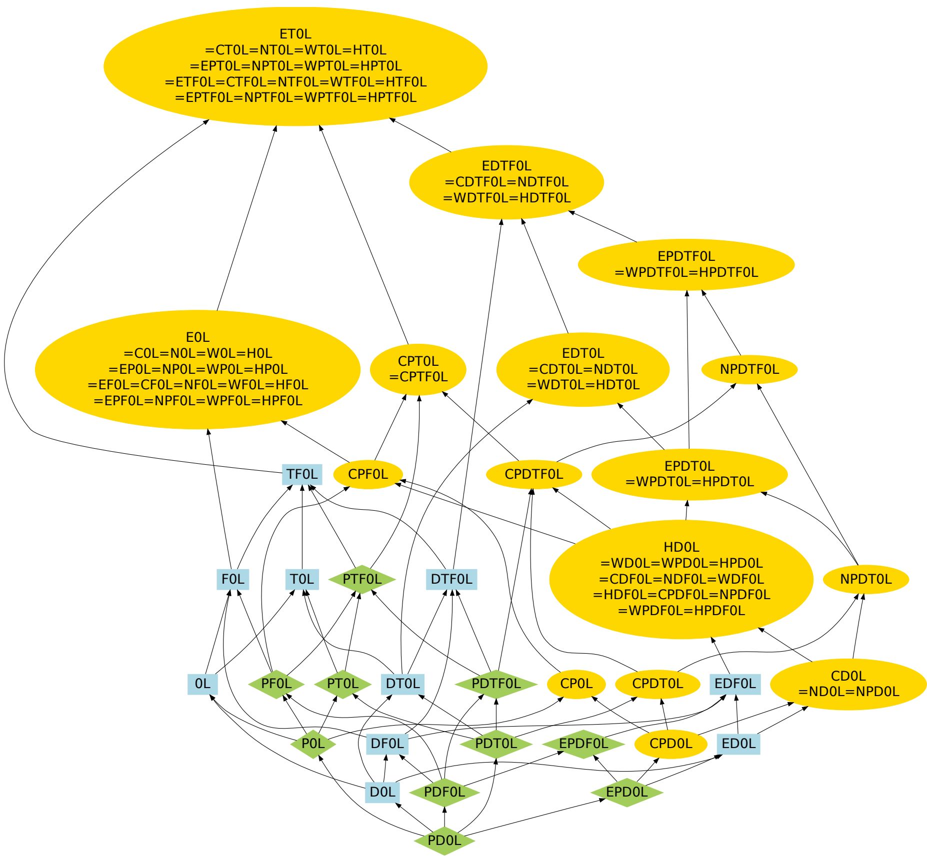

Fig. 1. Inclusion diagram showing classes of L systems colored by the infinite words they determine. Green classes (diamonds) determine exactly ω(PD0L), blue classes (rectangles) determine exactly ω(D0L), and yellow classes (ellipses) determine exactly ω(CD0L). Inclusions and equalities are from [9].

Proof techniques We obtain our categorization results by showing that all infinite languages in certain classes of L systems have infinite subsets in certain smaller classes of L systems. This limits the infinite words of the larger class to the infinite words of the smaller class. That every infinite T0L language has an infinite D0L subset was shown in [12] using a pumping lemma for T0L languages. With this result, we show that every infinite ET0L language has an infinite CD0L subset, and we make further use of the pumping lemma to show that every infinite PT0L language has an infinite PD0L subset. A separate argument shows that every infinite ED0L (EPD0L) language has an infinite D0L (PD0L) subset.

Related work Prefix languages were investigated by Book [3], who formulated a "prefix property" intended to allow languages to "approximate" infinite sequences, and showed that for certain classes of languages, if a language in the class has the prefix property, then it is regular. Languages whose complement is a prefix language, called "coprefix languages", have also been studied [2].

The iterative processes underlying L systems have been investigated in connection with infinite words. Pansiot [10] considered various classes of infinite words obtained by iterated mappings. Culik & Karhumäki [5] considered iterative devices generating infinite words. Culik & Salomaa [6] investigated infinite words associated with D0L and DT0L systems; their notion of "strong uniform convergence" is equivalent to our notion of a language "determining" an infinite word.

Our results on infinite subsets can be restated in the framework of set immunity [13]. For a language class C, a language L is C-immune iff L is infinite and no infinite subset of L is in C. For example, our result that every infinite ET0L language has an infinite CD0L subset could be stated: no ET0L language is CD0L-immune. In addition to categorizing the infinite words determined by L systems, our results characterize the immunity relationships among these systems.

Outline of paper The paper is organized as follows. Section 2 gives preliminary definitions and propositions. Section 3 gives results on infinite subsets of certain classes of L systems. Section 4 categorizes the infinite words determined by the hierarchy of L systems. Section 5 separates and characterizes the classes ω(PD0L), ω(D0L), and ω(CD0L). Section 6 gives our conclusions.

2 Preliminaries

An **alphabet** A is a finite set of symbols. A **string** (or finite word) is an element of A^*. We denote the empty string by λ. A **language** is a subset of A^*. An **infinite word** (or stream) is an element of A^ω. A (symbolic) **sequence** S is an element of $A^* \cup A^\omega$. A **prefix** of S is a string x such that $S = xS'$ for some sequence S'. A **subword** (or factor) of S is a string x such that $S = wxS'$ for some string w and sequence S'. For a nonempty string x, x^ω denotes the infinite word $xxx\cdots$. Such a word is called **purely periodic**. An infinite word of the form xy^ω, where x and y are strings and $y \neq \lambda$, is called **ultimately periodic**.

A **morphism** on an alphabet A is a map h from A^* to A^* such that for all $x, y \in A^*$, $h(xy) = h(x)h(y)$. Notice that $h(\lambda) = \lambda$. The morphism h is **nonerasing** if for all $a \in A$, $h(a) \neq \lambda$. The morphism h is a **coding** if for all $a \in A$, $|h(a)| = 1$. The morphism h is a **weak coding** if for all $a \in A$, $|h(a)| \leq 1$. The morphism h is an **identity** if for all $a \in A$, $h(a) = a$. For a language L, we define $h(L) = \{h(x) \mid x \in L\}$. A string $x \in A^*$ is **mortal** (for h) if there is an $m \geq 0$ such that $h^m(x) = \lambda$. The morphism h is **prolongable** on a symbol a if $h(a) = ax$ for some $x \in A^*$, and x is not mortal. In this case $h^\omega(a)$ denotes the infinite word $a\ x\ h(x)\ h^2(x)\ \cdots$. An infinite word α is **pure morphic** if there is a morphism h and symbol a such that h is prolongable on a and $\alpha = h^\omega(a)$. An infinite word α is **morphic** if there is a morphism h, coding e, and symbol a such that h is prolongable on a and $\alpha = e(h^\omega(a))$. Every purely periodic word is pure morphic, and every ultimately periodic word is morphic. For results on morphic words, see [1].

A **finite substitution** on A is a map σ from A^* to 2^{A^*} such that (1) for all $x \in A^*$, $\sigma(x)$ is finite and nonempty, and (2) for all $x, y \in A^*$, $\sigma(xy) = \{x'y' \mid x'$ is in $\sigma(x)$ and y' is in $\sigma(y)\}$. Notice that $\sigma(\lambda) = \{\lambda\}$. σ is **nonerasing** if for all $a \in A$, $\sigma(a) \not\ni \lambda$. For a language L, we define $\sigma(L) = \{x' \mid x'$ is in $\sigma(x)$ for some $x \in L\}$.

2.1 Prefix languages

A **prefix language** is a language L such that for all $x, y \in L$, x is a prefix of y or y is a prefix of x. A language L **determines** an infinite word α iff L is infinite and every $x \in L$ is a prefix of α. For example, the infinite prefix language $\{\lambda$, `ab`, `abab`, `ababab`, $\ldots\}$ determines the infinite word $(\texttt{ab})^\omega$. The following propositions are basic consequences of the definitions.

Proposition 1. *A language determines at most one infinite word.*

Proposition 2. *A language L determines an infinite word iff L is an infinite prefix language.*

Notice that while a language determines at most one infinite word, an infinite word may be determined by more than one language. In particular, we will make use of the following fact.

Proposition 3. *If a language L determines an infinite word α and L' is an infinite subset of L, then L' determines α.*

For a language class C, let $\omega(C) = \{\alpha \mid \alpha$ is an infinite word and some $L \in C$ determines $\alpha\}$.

2.2 L systems

Many classes of L systems appear in the literature. Following [9], we consider four production features (D,P,F,T) and five filtering features (E,C,N,W,H). Each production feature may be present or absent, and at most one filtering feature may be present, for a total of $2^4 \cdot 6 = 96$ classes of L systems.

Feature	Meaning	Example
none		A **0L system** is a tuple $G = (A, \sigma, w)$ where A is an alphabet, σ is a finite substitution on A, and w is in A^*. The language of G is $L(G) = \{s \in \sigma^i(w) \mid i \geq 0\}$.
D	Deterministic	A **D0L system** is a tuple $G = (A, h, w)$ where A is an alphabet, h is a morphism on A, and w is in A^*. The language of G is $L(G) = \{h^i(w) \mid i \geq 0\}$.
P	Propagating	A **PD0L** system is a D0L system (A, h, w) such that h is nonerasing.
F	Finite axiom set	A **DF0L system** is a tuple $G = (A, h, F)$ where A is an alphabet, h is a morphism on A, and F is a finite set of strings in A^*. The language of G is $L(G) = \{h^i(f) \mid f \in F \text{ and } i \geq 0\}$.
T	Tables	A **DT0L system** is a tuple $G = (A, H, w)$ where A is an alphabet, H is a finite nonempty set of morphisms on A (called "tables"), and w is in A^*. The language of G is $L(G) = \{s \mid h_i \cdots h_1(w) = s$ for some $h_1, \ldots, h_i \in H$.
E	Extended	An **ED0L system** is a tuple $G = (A, h, w, B)$ where A and B are alphabets and $B \subseteq A$, h is a morphism on A, and w is in A^*. The language of G is $L(G) = \{s \in B^* \mid h^i(w) = s \text{ for some } i \geq 0\}$.
H	Homomorphism	An **HD0L system** is a tuple $G = (A, h, w, g)$ such that $G' = (A, h, w)$ is a D0L system and g is a morphism on A. The language of G is $L(G) = \{g(s) \mid s$ is in $L(G')\}$.
C	Coding	A **CD0L** system is an HD0L system (A, h, w, g) such that g is a coding.
N	Nonerasing	An **ND0L** system is an HD0L system (A, h, w, g) such that g is nonerasing.
W	Weak coding	A **WD0L** system is an HD0L system (A, h, w, g) such that g is a weak coding.

These features combine to form complex L systems. For example, an **EPD0L system** is an ED0L system (A, h, w, B) such that h is nonerasing. A **T0L system** is a tuple $G = (A, T, w)$ where A is an alphabet, T is a finite nonempty set of finite substitutions on A (called "tables"), and w is in A^*. The language of G is $L(G) = \{s \mid \sigma_i \cdots \sigma_1(w) \ni s \text{ for some } \sigma_1, \ldots, \sigma_i \in T\}$. If for all $\sigma \in T$, σ is nonerasing, then G is a **PT0L system**. See [11] and [9] for more definitions.

We call an L system G infinite iff $L(G)$ is infinite. When speaking of language classes, we denote the class of D0L languages simply by D0L, and similarly with other classes. An **L system feature set** is a subset of {D,P,F,T} ∪ {E,C,N,W,H} containing at most one of {E,C,N,W,H}. Let $\mathscr{L}(S)$ be the language class of

L systems with feature set S. For example, $\mathscr{L}(\{\text{C,D,T}\}) = \text{CDT0L}$. From the definitions of the features, we have the following inclusions.

Proposition 4 (Structural inclusions).
Let S be an L system feature set. Then:

- $\mathscr{L}(S \cup \{D\}) \subseteq \mathscr{L}(S)$,
- $\mathscr{L}(S \cup \{P\}) \subseteq \mathscr{L}(S)$,
- $\mathscr{L}(S) \subseteq \mathscr{L}(S \cup \{F\})$, *and*
- $\mathscr{L}(S) \subseteq \mathscr{L}(S \cup \{T\})$.

Let S be an L system feature set containing none of $\{E,C,N,W,H\}$. Then:

- $\mathscr{L}(S) \subseteq \mathscr{L}(S \cup \{E\})$,
- $\mathscr{L}(S) \subseteq \mathscr{L}(S \cup \{C\})$,
- $\mathscr{L}(S \cup \{C\}) \subseteq \mathscr{L}(S \cup \{N\}) \subseteq \mathscr{L}(S \cup \{H\})$, *and*
- $\mathscr{L}(S \cup \{C\}) \subseteq \mathscr{L}(S \cup \{W\}) \subseteq \mathscr{L}(S \cup \{H\})$.

Beyond these structural inclusions, many relationships are known among the language classes; see [9]. In comparing L system classes, [9] considers two languages to be equal if they differ by the empty word only; otherwise, propagating classes would be automatically different from nonpropagating ones. See Figure 1 for a depiction of the known inclusions and equalities.

3 Infinite Subsets of L Systems

In this section we show that all infinite languages in certain classes of L systems have infinite subsets in certain smaller classes of L systems. This limits the infinite words of the larger class to the infinite words of the smaller class. We make use of a pumping lemma for T0L systems from [12]. A T0L system $G = (A, T, w)$ is **pumpable** iff there are $a, b \in A$ such that (1) some $s_0 \in L(G)$ contains a, and (2) for some composition t of tables from T, $t(a)$ includes a string s_1 containing distinct occurrences of a and b and $t(b)$ includes a string s_2 containing b. The next two theorems appear in [12].

Theorem 5 (Smith). *A T0L system is infinite iff it is pumpable.*

Theorem 6 (Smith). *Every infinite T0L language has an infinite D0L subset.*

Theorem 7. *Every infinite PT0L language has an infinite PD0L subset.*

Proof. Take any infinite PT0L language L with PT0L system $G = (A, T, w)$. By Theorem 5, G is pumpable for some $a, b \in A$, $s_0, s_1, s_2 \in A^*$, and composition t of tables from T. Let h be a morphism on A such that $h(a) = s_1$, $h(b) = s_2$ unless $a = b$, and for every other $c \in A$, $h(c) = s$ for some $s \in t(c)$. Since t is a composition of tables from T, t is nonerasing, hence h is nonerasing. Further, for all $i \geq 0$, $h^i(s_0)$ is in $t^i(s_0)$, so $h^i(s_0)$ is in L. A simple induction shows that for all $i \geq 0$, $h^i(s_0)$ contains a and at least i copies of b. Hence the language of the PD0L system (A, h, s_0) is an infinite subset of L. □

Theorem 8. *Let $G = (A, h, w, B)$ be an infinite ED0L system. Then there are $a \geq 0, b \geq 1$ such that the language of the D0L system $(A, h^b, h^a(w))$ is an infinite subset of $L(G)$.*

Proof. Let alph(s) be the set of symbols which appear in the string s. Since $L(G)$ is infinite, there is an $m \geq 0$ such that the sequence w, $h(w)$, $h^2(w)$, $\ldots$, $h^m(w)$ contains more than $2^{|B|}$ strings in $L(G)$. For every $s \in L(G)$, alph$(s) \subseteq B$. Hence there is a $C \subseteq B$ and i, j such that $0 \leq i < j \leq m$ and alph$(h^i(w)) =$ alph$(h^j(w)) = C$. Then for any string s such that alph$(s) = C$, alph$(h^{j-i}(s)) = C$. Let $a = i$ and $b = j - i$. Then for every $n \geq 0$, alph$(h^{a+bn}(w)) = C$. Hence for every $n \geq 0$, $h^{a+bn}(w)$ is in $L(G)$. So take the D0L system $G' = (A, h^b, h^a(w))$. We have $L(G') \subseteq L(G)$. Suppose some string s occurs twice in the derivation sequence of G'. Then s occurs twice in the derivation sequence of G, making $L(G)$ finite, a contradiction. So $L(G')$ is infinite. Therefore $L(G')$ is an infinite subset of $L(G)$. □

Corollary 9. *Every infinite ED0L language has an infinite D0L subset.*

Corollary 10. *Every infinite EPD0L language has an infinite PD0L subset.*

Proof. Take any infinite EPD0L system $G = (A, h, w, B)$. By Theorem 8, there are $a \geq 0, b \geq 1$ such that the language of the D0L system $G' = (A, h^b, h^a(w))$ is an infinite subset of $L(G)$. Since h is nonerasing, h^b is nonerasing. Hence $L(G')$ is an infinite PD0L subset of $L(G)$. □

Theorem 11. *Every infinite ET0L language has an infinite CD0L subset.*

Proof. Take any infinite ET0L language L. By Theorem 2.7 of [9], ET0L = CT0L. Hence there is a coding e and T0L language L' such that $L = e(L')$. Since L is infinite, L' is infinite. Then by Theorem 6, L' has an infinite D0L subset L''. Since L'' is infinite and e is a coding, $e(L'')$ is infinite. Since $L'' \subseteq L'$, $e(L'') \subseteq e(L')$. Therefore $e(L'')$ is an infinite CD0L subset of L. □

Theorem 12. *Let S be an L system feature set not containing F. Then every infinite $\mathscr{L}(S \cup \{F\})$ language has an infinite $\mathscr{L}(S)$ subset.*

Proof. Take any infinite L system G with feature set $S \cup \{\mathrm{F}\}$. Since G has a finite axiom set, $L(G)$ is a finite union of $\mathscr{L}(S)$ languages. Then since $L(G)$ is infinite, one of these $\mathscr{L}(S)$ languages is infinite. Therefore $L(G)$ has an infinite $\mathscr{L}(S)$ subset. □

Theorem 13. *Let C and D be language classes such that every infinite language in C has an infinite subset in D. Then $\omega(C) \subseteq \omega(D)$.*

Proof. Take any $\alpha \in \omega(C)$. Some $L \in C$ determines α. Then L is infinite, so L has an infinite subset L' in D. Then L' determines α. So α is in $\omega(D)$. Hence $\omega(C) \subseteq \omega(D)$. □

4 Categorizations

In this section we categorize the infinite words determined by each class of L systems. We partition the 96 classes into three sets, called Set_1, Set_2, and Set_3, and show that for every $C_1 \in Set_1$, $C_2 \in Set_2$, and $C_3 \in Set_3$, $\omega(C_1) = \omega(\text{PD0L})$, $\omega(C_2) = \omega(\text{D0L})$, and $\omega(C_3) = \omega(\text{CD0L})$.

4.1 PD0L classes

Let $Set_1 = \{$PD0L, PDF0L, P0L, PF0L, PDT0L, PDTF0L, PT0L, PTF0L, EPD0L, EPDF0L$\}$.

Theorem 14. *For every $C \in Set_1$, every infinite C language has an infinite PD0L subset.*

Proof. Take any $C \in Set_1$. By structural inclusion, $C \subseteq$ PTF0L or $C \subseteq$ EPDF0L. By Theorem 12, every infinite PTF0L language has an infinite PT0L subset. By Theorem 7, every infinite PT0L language has an infinite PD0L subset. Hence every infinite PTF0L language has an infinite PD0L subset. By Theorem 12, every infinite EPDF0L language has an infinite EPD0L subset. By Corollary 10, every infinite EPD0L language has an infinite PD0L subset. Hence every infinite EPDF0L language has an infinite PD0L subset. Hence every infinite C language has an infinite PD0L subset. □

Theorem 15. *For every $C \in Set_1$, $\omega(C) = \omega(\text{PD0L})$.*

Proof. Take any $C \in Set_1$. By structural inclusion, PD0L $\subseteq$ C. Hence $\omega(\text{PD0L}) \subseteq \omega(C)$. By Theorem 14, every infinite C language has an infinite PD0L subset. Then by Theorem 13, $\omega(C) \subseteq \omega(\text{PD0L})$. Therefore $\omega(C) = \omega(\text{PD0L})$. □

4.2 D0L classes

Let $Set_2 = \{$D0L, DF0L, 0L, F0L, DT0L, DTF0L, T0L, TF0L, ED0L, EDF0L$\}$.

Theorem 16. *For every $C \in Set_2$, every infinite C language has an infinite D0L subset.*

Proof. Take any $C \in Set_2$. By structural inclusion, $C \subseteq$ TF0L or $C \subseteq$ EDF0L. By Theorem 12, every infinite TF0L language has an infinite T0L subset. By Theorem 6, every infinite T0L language has an infinite D0L subset. Hence every infinite TF0L language has an infinite D0L subset. By Theorem 12, every infinite EDF0L language has an infinite ED0L subset. By Corollary 9, every infinite ED0L language has an infinite D0L subset. Hence every infinite EDF0L language has an infinite D0L subset. Hence every infinite C language has an infinite D0L subset. □

Theorem 17. *For every $C \in Set_2$, $\omega(C) = \omega(\text{D0L})$.*

Proof. Take any $C \in Set_2$. By structural inclusion, D0L $\subseteq$ C. Hence $\omega(\text{D0L}) \subseteq \omega(C)$. By Theorem 16, every infinite C language has an infinite D0L subset. Then by Theorem 13, $\omega(C) \subseteq \omega(\text{D0L})$. Therefore $\omega(C) = \omega(\text{D0L})$. □

4.3 CD0L classes

Let Set_3 = {CD0L, ND0L, WD0L, HD0L, CPD0L, NPD0L, WPD0L, HPD0L, CDF0L, NDF0L, WDF0L, HDF0L, CPDF0L, NPDF0L, WPDF0L, HPDF0L, E0L, C0L, N0L, W0L, H0L, EP0L, CP0L, NP0L, WP0L, HP0L, EF0L, CF0L, NF0L, WF0L, HF0L, EPF0L, CPF0L, NPF0L, WPF0L, HPF0L, EDT0L, CDT0L, NDT0L, WDT0L, HDT0L, EPDT0L, CPDT0L, NPDT0L, WPDT0L, HPDT0L, EDTF0L, CDTF0L, NDTF0L, WDTF0L, HDTF0L, EPDTF0L, CPDTF0L, NPDTF0L, WPDTF0L, HPDTF0L, ET0L, CT0L, NT0L, WT0L, HT0L, EPT0L, CPT0L, NPT0L, WPT0L, HPT0L, ETF0L, CTF0L, NTF0L, WTF0L, HTF0L, EPTF0L, CPTF0L, NPTF0L, WPTF0L, HPTF0L}.

Theorem 18. *For every $C \in Set_3$, every infinite C language has an infinite CD0L subset.*

Proof. Take any $C \in Set_3$. By structural inclusion, $C \subseteq$ ETF0L or $C \subseteq$ HTF0L. By Theorem 2.7 of [9], ETF0L = HTF0L = ET0L. So $C \subseteq$ ET0L. By Theorem 11, every infinite ET0L language has an infinite CD0L subset. Hence every infinite C language has an infinite CD0L subset. □

Theorem 19. *For every $C \in Set_3$, $\omega(C) = \omega$(CD0L).*

Proof. Take any $C \in Set_3$. By Theorem 18, every infinite C language has an infinite CD0L subset. Then by Theorem 13, $\omega(C) \subseteq \omega$(CD0L).

Next, by structural inclusion, CPD0L $\subseteq C$ or EP0L $\subseteq C$ or EPDT0L $\subseteq C$. By Theorem 2.4 of [9], EP0L = C0L, so CPD0L $\subseteq$ EP0L. By Theorem 2.6 of [9], CPDT0L $\subseteq$ EPDT0L, so CPD0L $\subseteq$ EPDT0L. Hence CPD0L $\subseteq C$. Now by Theorem 2.3 of [9], CPDF0L = CDF0L. Hence CD0L $\subseteq$ CPDF0L. Hence ω(CD0L) $\subseteq \omega$(CPDF0L). By Theorem 12, every infinite CPDF0L language has an infinite CPD0L subset. Then by Theorem 13, ω(CPDF0L) $\subseteq \omega$(CPD0L). Hence ω(CD0L) $\subseteq \omega$(CPD0L) $\subseteq \omega(C)$.

Therefore $\omega(C) = \omega$(CD0L). □

5 ω(PD0L), ω(D0L), and ω(CD0L)

In this section, we separate the three classes of infinite words obtained in the previous section, giving ω(PD0L) $\subset \omega$(D0L) $\subset \omega$(CD0L). We observe that ω(D0L) properly contains the pure morphic words and we show that ω(CD0L) contains exactly the morphic words.

5.1 Separating the classes

From Theorem 2.3 of [10], the infinite words generated by iterating nonerasing morphisms are a proper subset of the pure morphic words, which in turn are a proper subset of the morphic words. Our classes ω(PD0L), ω(D0L), and ω(CD0L) are defined more generally using prefix languages, but similar arguments serve to separate them.

Theorem 20. *ω(PD0L) $\subset$ ω(D0L).*

Proof. By structural inclusion, ω(PD0L) $\subseteq$ ω(D0L). To separate the two classes, we use an infinite word from [4]. Let $A = \{0, 1, 2\}$. Let f be a morphism on A such that $f(0) = 01222$, $f(1) = 10222$, and $f(2) = \lambda$. Let $\alpha = f^\omega(0) = 0122210222102220 1222\ldots$ Then α is a pure morphic word, hence α is in ω(D0L). In [4] it is shown that there is no nonerasing morphism g on A such that $g^\omega(0) = \alpha$. We generalize this result to show that α is not in ω(PD0L). First, we show that if g is a nonerasing morphism on A and $g(\alpha) = \alpha$, then g is an identity morphism. We adapt the proof of Example 3 in [4].

Let τ be the Thue-Morse word $\tau = 01101001\ldots = u^\omega(0)$, where u is a morphism on $\{0, 1\}$ such that $u(0) = 01$ and $u(1) = 10$. Let d be a morphism on A such that $d(0) = 0$, $d(1) = 1$, and $d(2) = \lambda$. As observed by [4], $d(\alpha) = \tau$. Notice that the only subwords of α in $\{0, 1\}^*$ are in $\{\lambda, 0, 1, 01, 10\}$ and the only subwords of α in $\{2\}^*$ are in $\{\lambda, 2, 22, 222\}$. Notice also that α does not contain the subword 212.

Suppose g is a nonerasing morphism on A and $g(\alpha) = \alpha$. Suppose $g(2)$ is not in 2^*. Let $s = d(g(2))$. Then s is not empty. Since 222 is a subword of α and $g(\alpha) = \alpha$, $g(222)$ is a subword of α. Then since $d(\alpha) = \tau$, τ contains $d(g(222)) = sss$, a contradiction, since τ is known to be cubefree. So $g(2)$ is in 2^*. Then since α contains $g(222)$, and 2222 is not a subword of α, and g is nonerasing, $g(2) = 2$.

Suppose $g(0) \neq 0$. Then since α starts with 0, $g(0) = 01x$ for some $x \in A^*$. Since 1222 is a subword of α, $g(1222) = g(1)\ 222$ is a subword of α. Then since 2222 is not a subword of α, $g(1)$ cannot end with 2. So $g(1) = ya$ for some $y \in A^*$ and $a \in \{0, 1\}$. Now since 10 is a subword of α, so is $g(10) = ya01x$. But α contains no subword of the form $a01$, a contradiction. So $g(0) = 0$.

Suppose $g(1) \neq 1$. Then since α begins with 012, $g(1) = 12z$ for some $z \in A^*$. Since 2221 is a subword of α, $g(2221) = 22212z$ is a subword of α, a contradiction, since α does not contain the subword 212. So $g(1) = 1$. Then g is an identity morphism.

So suppose α is in ω(PD0L). Then there is a PD0L system $G = (A, h, w)$ such that $L(G)$ determines α. Since h is nonerasing, $h(\alpha)$ is an infinite word. Suppose $h(\alpha) \neq \alpha$. Then there is a prefix p of α such that $h(p)$ is not a prefix of α. Since $L(G)$ determines α, p is a prefix of some s in $L(G)$. Then $h(p)$ is a prefix of $h(s)$. But then since $h(s)$ is in $L(G)$, $h(p)$ is a prefix of α, a contradiction. So $h(\alpha) = \alpha$. Then from above, h is an identity morphism. But then $h(w) = w$, so $L(G)$ is finite, a contradiction. Therefore α is not in ω(PD0L). Hence ω(PD0L) $\subset$ ω(D0L). □

Theorem 21. *ω(D0L) $\subset$ ω(CD0L).*

Proof. By structural inclusion, ω(D0L) $\subseteq$ ω(CD0L). Let $\alpha = \mathtt{abba}^\omega$. Since α is ultimately periodic, α is morphic, hence α is in ω(CD0L). Suppose α is in ω(D0L). Then there is a D0L system $G = (A, h, w)$ such that $L(G)$ determines α. Clearly $h(\mathtt{a})$ cannot include $\mathtt{b}$, and if $h(\mathtt{a}) = \lambda$, $L(G)$ is finite, a contradiction. So since $h(\mathtt{a})$ must be a prefix of α, $h(\mathtt{a}) = \mathtt{a}$. Then $\mathtt{a}\ h(\mathtt{b})\ h(\mathtt{b})$ is a prefix of α, hence $h(\mathtt{b}) = \lambda$ or $h(\mathtt{b}) = \mathtt{b}$. But then $L(G)$ is finite, a contradiction. So α is not in ω(D0L). Hence ω(D0L) $\subset$ ω(CD0L). □

5.2 Characterizing the words in each class

That ω(D0L) includes every pure morphic word is immediate from the definitions. In [8], the infinite word $\mathtt{aab}^\omega$ is given as an example of an infinite D0L word which is not pure morphic. Hence ω(D0L) properly contains the pure morphic words. Next, we show that ω(CD0L) contains exactly the morphic words. The **adherence** of a language L, denoted Adherence(L), is the set $\{\alpha \mid \alpha$ is an infinite word and for every prefix p of α, there is an $s \in L$ such that p is a prefix of $s\}$.

Lemma 22. *Suppose L is in D0L and α is in Adherence(L). Then α is morphic.*

Proof. From [7], either (1) α is ultimately periodic, or (2) $\alpha = w\ x\ h(x)\ h^2(x)\ \cdots$ for some morphism h and strings w, x such that $h(w) = wx$ and x is not mortal. If (1), α is morphic. If (2), α is an infinite D0L word, so by Proposition 10.2.2 of [8], α is morphic. □

Theorem 23. *α is in ω(CD0L) iff α is morphic.*

Proof. That ω(CD0L) includes every morphic word is immediate from the definitions. So take any $\alpha \in \omega$(CD0L). Then there is a CD0L system $G = (A, h, w, e)$ such that $L(G)$ determines α. Then $L(G)$ is infinite. Hence the language L of the D0L system (A, h, w) is infinite. As noted in [7], a language has empty adherence iff the language is finite. Therefore there is an $\alpha' \in$ Adherence(L). By Lemma 22, α' is morphic. Now for any prefix p of α', there is a string s in L with p as a prefix. Then $e(p)$ is a prefix of $e(s)$. Then since $e(s)$ is in $L(G)$, $e(p)$ is a prefix of α. So for every prefix p of α', $e(p)$ is a prefix of α. Since e is a coding, $e(\alpha')$ is infinite. So $e(\alpha') = \alpha$. Then because a coding of a morphic word is still a morphic word, α is morphic. Hence α is in ω(CD0L) iff α is morphic. □

6 Conclusion

In this paper we have categorized the infinite words determined by L systems, showing that a variety of classes of L systems collapse to just three classes of infinite words. To associate L systems with infinite words, we used the concept of prefix languages. This concept can be applied not just to L systems, but to arbitrary language classes, offering many opportunities for further research. That is, where C is any language class, we denote by $\omega(C)$ the class of infinite words determined by the prefix languages in C. Then for a given language class, we can ask what class of infinite words it determines. From the other direction, for a given infinite word, we can ask in what language classes it can be determined. It is hoped that work in this area will help to build up a theory of the complexity of infinite words with respect to what language classes can determine them.

Acknowledgments. I want to thank my advisor, Rajmohan Rajaraman, for supporting this work, encouraging me, and offering many helpful comments and suggestions.

References

1. Allouche, J.P., Shallit, J.: Automatic Sequences: Theory, Applications, Generalizations. Cambridge University Press, New York, NY, USA (2003)
2. Berstel, J.: Properties of infinite words : Recent results. In: Monien, B., Cori, R. (eds.) STACS 89, Lecture Notes in Computer Science, vol. 349, pp. 36–46. Springer Berlin Heidelberg (1989)
3. Book, R.V.: On languages with a certain prefix property. Mathematical Systems Theory 10, 229–237 (1977)
4. Cassaigne, J., Nicolas, F.: Quelques propriétés des mots substitutifs. Bulletin of the Belgian Mathematical Society-Simon Stevin 10(5), 661–677 (2003)
5. Culik, K., Karhumäki, J.: Iterative devices generating infinite words. Int. J. Found. Comput. Sci. 5(1), 69–97 (1994)
6. Culik, K., Salomaa, A.: On infinite words obtained by iterating morphisms. Theoretical Computer Science 19(1), 29–38 (1982)
7. Head, T.: Adherences of D0L languages. Theoretical Computer Science 31(1–2), 139–149 (1984)
8. Honkala, J.: The equality problem for purely substitutive words. In: Combinatorics, automata, and number theory, pp. 505–529. Cambridge: Cambridge University Press (2010)
9. Kari, L., Rozenberg, G., Salomaa, A.: In: Rozenberg, G., Salomaa, A. (eds.) Handbook of Formal Languages, Vol. 1, chap. L systems, pp. 253–328. Springer-Verlag New York, Inc., New York, NY, USA (1997)
10. Pansiot, J.J.: On various classes of infinite words obtained by iterated mappings. In: Nivat, M., Perrin, D. (eds.) Automata on Infinite Words, Lecture Notes in Computer Science, vol. 192, pp. 188–197. Springer Berlin Heidelberg (1985)
11. Rozenberg, G., Salomaa, A.: Mathematical Theory of L Systems. Academic Press, Inc., Orlando, FL, USA (1980)
12. Smith, T.: Infiniteness and boundedness in 0L, DT0L, and T0L systems. In: Dediu, A.H., Martín-Vide, C., Truthe, B. (eds.) Language and Automata Theory and Applications, Lecture Notes in Computer Science, vol. 7810, pp. 535–546. Springer Berlin Heidelberg (2013)
13. Yamakami, T., Suzuki, T.: Resource bounded immunity and simplicity. Theoretical Computer Science 347(12), 90–129 (2005)