Infiniteness and Boundedness in 0L, DT0L, and T0L Systems

Tim Smith

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

Abstract. We investigate the boundary between finiteness and infiniteness in three types of L systems: 0L, DT0L, and T0L. We establish necessary and sufficient conditions for 0L, DT0L, and T0L systems to be infinite, and characterize the boundedness of finite classes of such systems. First, we give a pumping lemma for these systems, proving that the language of a system is infinite iff the system is pumpable. Next, we show that the number of steps needed to derive any string in any finite 0L or DT0L system is bounded by a function depending only on the size of the alphabet, and not on the production rules or start string. This alphabet boundedness does not hold for finite T0L systems in general. Finally, we show that every infinite 0L system has an infinite D0L subsystem.

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 [6, 2]. 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". In [7], Vitányi gives a necessary and sufficient condition under which a D0L system is finite, and gives an upper bound on the size of a finite D0L language in terms of the size of the alphabet.

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". Generalizing in both directions at once yields the class of T0L systems. Figure 1 depicts the inclusions among these classes. We extend Vitányi's work to these systems.

First, we provide a necessary and sufficient condition under which a T0L system is infinite, in the form of a pumping lemma. In getting this result, we adapt a proof technique used in [5] to obtain a pumping lemma for ET0L systems. It follows from our pumping lemma that every infinite T0L language has an infinite D0L subset.

Next, we look for upper bounds on finite 0L, DT0L, and T0L systems in terms of alphabet size. In contrast to D0L systems, there is no upper bound on the size of a finite 0L or DT0L language in terms of the size of the alphabet alone. However, we show that there is such a bound on the number of steps needed to derive a string in any finite 0L or DT0L system. For finite T0L systems in general, a counterexample shows that no such alphabet-only bound holds. Figure 2 summarizes these results.

Finally, we consider the notion of a D0L subsystem of a 0L system, formed by choosing a single production for each symbol from the finite substitution. We show that every infinite 0L system has an infinite D0L subsystem; this constitutes a necessary and sufficient condition for a 0L system to be infinite. We also consider the notion of a D0L subsystem of a DT0L system, formed by choosing a single table from the set of tables. A simple counterexample shows that not every infinite DT0L system has an infinite D0L subsystem.

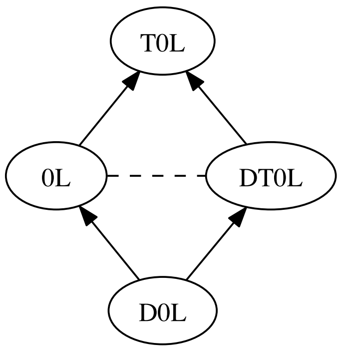

Fig. 1. Inclusion diagram. Arrows indicate proper inclusion of the lower class by the upper class; the dashed line indicates incomparability.

	alphabet size-bounded?	alphabet step-bounded?
D0L	yes	yes
0L	no	yes
DT0L	no	yes
T0L	no	no

Fig. 2. Alphabet boundedness of finite D0L, 0L, DT0L, and T0L systems.

Related Work Finiteness of all the L systems considered in this paper is decidable from Theorem 4.1 of [2]. That the size of the alphabet bounds the number of steps needed to derive λ in a 0L system was known from Lemma 1.3 of [6]; for finite 0L systems, our Theorem 15 generalizes this result to include non-empty strings.

Nishida [3] investigated "quasi-deterministic" 0L systems, those for which there is an integer C such that the cardinality of the set of strings generated in exactly n steps is less than C for every n. Nishida and Salomaa [4] investigated "slender" 0L languages, those for which there is a constant k such that the language has at most k strings of any given length.

Corollary 5, which states our pumping lemma for DT0L systems, can also be proved via a connection with non-negative integer matrices. Each table in a DT0L system can be associated with a "growth matrix" indicating for each production, how many times each symbol appears on the righthand side of that production. Jungers et al. [1] consider the "joint spectral radius" ρ of a finite set of such matrices, distinguishing four cases. In cases (1) and (2) ($\rho = 0$ or $\rho = 1$ with bounded products), the associated DT0L system is finite, whereas in cases (3) and (4) ($\rho > 1$ or $\rho = 1$ with unbounded products), by their Corollary 1 and Proposition 2, assuming every symbol is reachable, the system is pumpable.

Outline of Paper The paper is organized as follows. Section 2 gives preliminary definitions. Section 3 presents our pumping lemma for T0L systems. Section 4 examines alphabet boundedness for finite 0L, DT0L, and T0L systems. Section 5 studies D0L subsystems of 0L and DT0L systems. Section 6 gives our conclusions.

2 Definitions

An **alphabet** A is a finite set of symbols. A **string** is an element of A^*. λ denotes the empty string. A **language** is a subset of A^*. A **morphism** on 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 $h(\lambda) = \lambda$. h is **nonerasing** if for every $c \in A$, $h(c) \neq \lambda$. 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 $\sigma(\lambda) = \{\lambda\}$. For a language L, we define $\sigma(L) = \{x' \mid x'$ is in $\sigma(x)$ for some $x \in L\}$.

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^*. For $x, y \in A^*$ and $i \geq 0$, we write $x \xrightarrow{i} y$ iff $h^i(x) = y$.

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^*. For $x, y \in A^*$ and $i \geq 0$, we write $x \xrightarrow{i} y$ iff $\sigma^i(x) \ni y$.

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^*. For $x, y \in A^*$ and $i \geq 0$, we write $x \xrightarrow{i} y$ iff $h_i \cdots h_1(x) = y$ for some $h_1, \ldots, h_i \in H$.

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^*. For $x, y \in A^*$ and $i \geq 0$, we write $x \xrightarrow{i} y$ iff $\sigma_i \cdots \sigma_1(x) \ni y$ for some $\sigma_1, \ldots, \sigma_i \in T$.

For any of the above systems G, w is called the "axiom" or "start string". The language of G is $L(G) = \{s \mid w \xrightarrow{i} s$ for some $i \geq 0\}$. Call G finite iff $L(G)$ is finite. Intuitively, a derivation in G means a sequence of steps, starting with w unless otherwise specified, each consisting of a string together with the precise table and/or productions used to derive it from the previous step. For formal definitions, see [6]. **D0L**, **0L**, **DT0L**, and **T0L** are the classes of D0L, 0L, DT0L, and T0L languages, respectively. Clearly D0L $\subseteq$ 0L $\subseteq$ T0L and D0L $\subseteq$ DT0L $\subseteq$ T0L. In fact, 0L and DT0L are incomparable, making all of these inclusions proper [6].

A D0L (0L, DT0L, T0L) system G with axiom w is **step-bounded by** n iff for every $s \in L(G)$, there is an $m \leq n$ such that $w \xrightarrow{m} s$. Let C be any class of D0L (0L, DT0L, T0L) systems. C is **alphabet size-bounded** iff for every alphabet A, there is an $n \geq 0$ such that for every $G \in C$ for which the alphabet of G is A, $|L(G)| \leq n$. C is **alphabet step-bounded** iff for every alphabet A, there is an $n \geq 0$ such that for every $G \in C$ for which the alphabet of G is A, G is step-bounded by n. Clearly if C is alphabet-size bounded, then C is alphabet step-bounded, since the same n will suffice.

3 Pumping Lemma for T0L Systems

A T0L system $G = (A, T, w)$ is **pumpable** iff there are $x, y \in A$ such that (1) some $s_0 \in L(G)$ contains x, and (2) for some composition t of tables from T, $t(x)$ includes a string s_1 containing distinct occurrences of x and y and $t(y)$ includes a string s_2 containing y.

Lemma 1. *Suppose the T0L system $G = (A, T, w)$ is pumpable. Then $L(G)$ is infinite.*

Proof. Since s_0 is in $L(G)$ and t is a composition of tables from T, $t^i(s_0) \subseteq L(G)$ for every $i \geq 0$. A simple induction shows that for all $i \geq 0$, $t^i(s_0)$ includes a string containing x and at least i copies of y. Hence $L(G)$ is infinite. □

Lemma 2. *Suppose the T0L system $G = (A, T, w)$ is infinite. Then G is pumpable.*

Proof. We assume a familiarity with [5], particularly the notions of an ET0L system, derivation tree, marked node, and branch node. G can be treated as an ET0L system in which the alphabet and terminal alphabet are identical. Following the proof of Theorem 15 in [5], for any node in a derivation tree, consider the "marked set", or set of marked symbols which appear on the same level of the tree. As shown in that proof, since G is infinite, there is an $x \in A$ such that some derivation tree in G of a string in which every position is marked has a path with two branch nodes labelled by x, with one an ancestor of the other, with the same marked set. Call the strings in which the ancestor and descendant nodes appear w_1 and w_2, respectively. Let t be the composition of tables which was applied to w_1 to derive w_2.

Since the ancestor node labelled by x in w_1 is a branch node, its descendant string in w_2 contains, in addition to the descendant node labelled by x, a marked node labelled by some $e \in A$. Now, since w_1 and w_2 have the same marked set, every $c \in A$ which labels a marked node in w_2 also labels a marked node in w_1. By definition, every marked node in w_1 has a marked descendant in w_2. A simple induction then shows that for every $i \geq 0$, there is a $c \in A$ such that w_2 contains a marked node labelled by c, and some $s \in t^i(e)$ contains c. Hence for every $i \geq 0$, $t^i(e)$ contains a non-empty string. So there are $j \geq 0$, $k \geq 1$ and $y \in A$ such that $t^j(e)$ includes a string containing y and $t^k(y)$ includes a string containing y. Then since $t(x)$ includes a string containing distinct occurrences of x and e, $t^{j+1}(x)$ includes a string containing distinct occurrences of x and y. Then $t^{k(j+1)}(x)$ includes a string containing distinct occurrences of x and y and $t^{k(j+1)}(y)$ includes a string containing y. So G is pumpable. □

Theorem 3. *A T0L system is infinite iff it is pumpable.*

Proof. Immediate from Lemmas 1 and 2. □

Corollary 4. *A 0L system $G = (A, \sigma, w)$ is infinite iff there are $x, y \in A$ such that (1) some $s \in L(G)$ contains x, and (2) for some $i \geq 0$, $\sigma^i(x)$ includes a string containing distinct occurrences of x and y and $\sigma^i(y)$ includes a string containing y.*

Corollary 5. *A DT0L system $G = (A, H, w)$ is infinite iff there are $x, y \in A$ such that (1) some $s \in L(G)$ contains x, and (2) for some composition h of morphisms from H, $h(x)$ contains distinct occurrences of x and y and $h(y)$ contains y.*

Corollary 6. *Every infinite T0L language has an infinite D0L subset.*

Proof. Take any infinite T0L language L with T0L system $G = (A, T, w)$. By Theorem 3, G is pumpable. Let h be a morphism on A such that $h(x) = s_1$, $h(y) = s_2$ unless $x = y$, and for every other $c \in A$, $h(c) = s$ for some $s \in t(c)$. Then the language of the D0L system (A, h, s_0) is an infinite subset of L. □

4 Alphabet Boundedness

In this section we examine the alphabet size-boundedness and step-boundedness of 0L, DT0L, and T0L systems. For D0L, Corollary 4 of [7] implies the following.

Theorem 7 (Vitányi). *The class of finite D0L systems is alphabet size-bounded and alphabet step-bounded.*

4.1 0L

We first give a simple counterexample to show that the class of finite 0L systems is not alphabet size-bounded.

Theorem 8. *The class of finite 0L systems is not alphabet size-bounded.*

Proof. Let $A = \{\mathtt{a}, \mathtt{b}\}$ and take any $n \geq 0$. Let $w = \mathtt{a}$. Let σ be a finite substitution on A such that $\sigma(\mathtt{a}) = \{\mathtt{b}, \mathtt{bb}, \mathtt{bbb}, \ldots, \mathtt{b}^n\}$ and $\sigma(\mathtt{b}) = \{\mathtt{b}\}$. Let $G = (A, \sigma, w)$. Then $L(G) = \{\mathtt{a}, \mathtt{b}, \mathtt{bb}, \mathtt{bbb}, \ldots, \mathtt{b}^n\}$. So $L(G)$ is finite, but $|L(G)| > n$. So the class of finite 0L systems is not alphabet size-bounded. □

Next we will show that the class of finite 0L systems is alphabet step-bounded. We begin with some definitions. Take any 0L system (A, σ, w) and any $c \in A$. For any $s \in A^*$, c is **reachable from** s iff for some $i \geq 0$, $\sigma^i(s)$ includes a string which contains c. c is **reachable** iff c is reachable from w. Let $L(s)$ be the language of the 0L system (A, σ, s). c is **mortal** (c is in M) iff $\sigma^i(c) = \{\lambda\}$ for some $i \geq 0$. c is **vital** (c is in V) iff c is not in M. c is **recursive** (c is in R) iff c is reachable from some $s \in \sigma(c)$. c is **monorecursive** (c is in MR) iff for every

s such that for some $i \geq 0$, $\sigma^i(c)$ includes s and s contains c, s is in M^*cM^*. For each $i \geq 0$, let $reach_c(i) = \{s \in \sigma^i(c) \mid c$ is reachable from $s\}$.

We now build up a series of lemmas toward our result that the class of finite 0L systems is alphabet step-bounded. Lemmas 9 and 10 are given without proof but are not difficult to verify.

Lemma 9. *Suppose c is in $R - MR$. Then $L(c)$ is infinite.*

Lemma 10. *For all $c \in M$, $\sigma^{|M|}(c) = \{\lambda\}$.*

Lemma 11. *Suppose c is in MR and s is in $reach_c(i)$ for some i. Then s is in M^*dM^* for some $d \in MR$.*

Proof. Since c is reachable from s, s must contain at least one symbol in V. s cannot contain more than one symbol in V, otherwise c would not be in MR. So s contains exactly one symbol d in V. Since c is reachable from d and d is reachable from c, d is in R. Now suppose d is in $R - MR$. Then $L(d)$ contains a string s' which includes d and a symbol in V. Since c is reachable from d, $L(s')$ contains a string which includes c and a symbol in V. Then $L(d)$ contains such a string. But then $L(c)$ contains such a string, a contradiction, since c is in MR. So d is in MR. Then s is in M^*dM^*. □

Lemma 12. *Suppose c is in MR. Then there is a k such that $1 \leq k \leq |MR|$ and some string in $\sigma^k(c)$ contains c.*

Proof. Since c is in MR, there is a $k \geq 1$ such that some string in $\sigma^k(c)$ contains c. Take the smallest such k. Then there is an s in M^*cM^* and derivation D of s from c in k steps. Suppose $k > |MR|$. Then the c in s has $> |MR|$ ancestors in D. Take any such ancestor d. c is reachable from d, so by Lemma 11, since d is not in M, d is in MR. So every ancestor of the c in s is in MR. But then one such ancestor must repeat, and the derivation could have been shortened to yield a k' such that $1 \leq k' < k$. Therefore $k \leq |MR|$. □

Lemma 13. *Suppose c is in A and $L(c)$ is finite. Then there is a k such that $1 \leq k \leq |A|$ and $reach_c(|A|^2) = reach_c(|A|^2 + k)$.*

Proof. Suppose c is not in MR. Then by Lemma 9, c is not in R. So c is not reachable from any s in $\sigma(c)$. Then for every $i \geq 1$, $reach_c(i) = \{\}$. So say c is in MR. From Lemma 12, there is a k such that $1 \leq k \leq |MR|$ and some $s \in \sigma^k(c)$ contains c. Then s is in $reach_c(k)$. Let $Set(i) = \{a \in MR \mid$ some string in $reach_c(i)$ includes $a\}$. Take any $i \geq 0$ and $a \in Set(ki)$. Some $s' \in reach_c(ki)$ includes a. Then s' is in $\sigma^{ki}(c)$. Since s contains c, s' is a substring of some s'' in $\sigma^{ki}(s)$. s'' is in $\sigma^{k(i+1)}(c)$. Since c is reachable from s', c is reachable from s''. So s'' is in $reach_c(k(i+1))$. Hence a is in $Set(k(i+1)$. So for all $i \geq 0$, $Set(ki)$ is a subset of $Set(k(i+1))$. Hence there is an $i < |MR|$ such that $Set(ki) = Set(k(i+1))$. Let $m = ki + |M|$ and $n = k(i+1) + |M|$. We will show that $reach_c(m) = reach_c(n)$.

Take any $s \in reach_c(m)$. There is some $s' \in \sigma^{ki}(c)$ such that s is in $\sigma^{|M|}(s')$. Then s' is in $reach_c(ki)$, so by Lemma 11, s' is in M^*dM^* for some $d \in MR$.

Then by Lemma 10, $\sigma^{|M|}(d)$ includes s. Now d is in $Set(ki)$, hence in $Set(k(i+1))$. Then by Lemma 11, $reach_c(k(i+1))$ contains an $s'' \in M^*dM^*$. So s is in $\sigma^{|M|}(s'')$, hence in $\sigma^{k(i+1)+|M|}(c)$, hence in $reach_c(n)$.

Now take any $s \in reach_c(n)$. There is some $s' \in \sigma^{k(i+1)}(c)$ such that s is in $\sigma^{|M|}(s')$. Then s' is in $reach_c(k(i+1))$, so by Lemma 11, s' is in M^*dM^* for some $d \in MR$. Then by Lemma 10, $\sigma^{|M|}(d)$ includes s. Now d is in $Set(k(i+1))$, hence in $Set(ki)$. Then by Lemma 11, $reach_c(ki)$ contains an $s'' \in M^*dM^*$. So s is in $\sigma^{|M|}(s'')$, hence in $\sigma^{ki+|M|}(c)$, hence in $reach_c(m)$.

Therefore $reach_c(m) = reach_c(n)$. Then for all $i \geq m$, $reach_c(i) = reach_c(i+k)$. Then since $m \leq |MR| \cdot (|MR|-1) + |M| \leq |A|^2$, $reach_c(|A|^2) = reach_c(|A|^2 + k)$. □

Theorem 14. *For every alphabet A, there are $f \geq 1, g \geq 0$ such that for every finite 0L system (A, σ, w), $\sigma^g(w) = \sigma^{g+f}(w)$.*

Proof. Let $f(0) = 1$ and for every $x \geq 1$, $f(x) = x!f(x-1)$. Let $g(0) = 0$ and for every $x \geq 1$, $g(x) = x^2 + g(x-1) + f(x)$. Take any finite 0L system $G = (A, \sigma, w)$. We will show by induction on $|A|$ that $\sigma^{g(|A|)}(w) = \sigma^{g(|A|)+f(|A|)}(w)$.

Take the base case of $|A| = 0$. Then $w = \lambda$. Then for all $i \geq 0$, $\sigma^i(w) = \{\lambda\}$. So $\sigma^{g(0)}(w) = \sigma^{g(0)+f(0)}(w)$.

So say $|A| \geq 1$ and $w \neq \lambda$. Suppose for induction that for every finite 0L system (A', σ', w') such that $|A'| < |A|$, $\sigma'^{g(|A'|)}(w') = \sigma'^{g(|A'|)+f(|A'|)}(w')$. Take any c in w. By Lemma 13, there is a k' such that $1 \leq k' \leq |A|$ and $reach_c(|A|^2) = reach_c(|A|^2 + k')$. Let $k = |A!|$ and $t = |A|^2$. Then since k is divisible by k', $reach_c(t) = reach_c(t+k)$. Let $x = f(|A|-1)$ and $y = g(|A|-1)$. We will show that $\sigma^{t+y+kx}(c) = \sigma^{t+y+2kx}(c)$.

Take any $s \in \sigma^{t+y+kx}(c)$. Then there is an $r \in \sigma^t(c)$ such that s is in $\sigma^{y+kx}(r)$. Suppose c is reachable from r. Then r is in $reach_c(t)$. Since $reach_c(t) = reach_c(t+kx)$, r is in $\sigma^{t+kx}(c)$. Then s is in $\sigma^{t+y+2kx}(c)$. So say c is not reachable from r. Then by the induction hypothesis, $\sigma^y(r) = \sigma^{y+x}(r)$. Hence $\sigma^{y+kx}(r) = \sigma^{y+2kx}(r)$. Then s is in $\sigma^{y+2kx}(r)$. So s is in $\sigma^{t+y+2kx}(c)$.

Now take any $s \in \sigma^{t+y+2kx}(c)$. Then there is an $r \in \sigma^{t+kx}(c)$ such that s is in $\sigma^{y+kx}(r)$. Suppose c is reachable from r. Then r is in $reach_c(t+kx)$. Since $reach_c(t+kx) = reach_c(t)$, r is in $\sigma^t(c)$. Then s is in $\sigma^{t+y+kx}(c)$. So say c is not reachable from r. Then by the induction hypothesis, $\sigma^y(r) = \sigma^{y+x}(r)$. Hence $\sigma^y(r) = \sigma^{y+kx}(r)$. Then s is in $\sigma^y(r)$. So s is in $\sigma^{t+y+kx}(c)$.

So for all c in w, $\sigma^{t+y+kx}(c) = \sigma^{t+y+2kx}(c)$. Hence $\sigma^{t+y+kx}(w) = \sigma^{t+y+2kx}(w)$. Now $t + y + kx = g(|A|)$ and $kx = f(|A|)$, completing the induction. □

Theorem 15. *The class of finite 0L systems is alphabet step-bounded.*

Proof. Take any alphabet A. Take any f, g meeting the conditions of Theorem 14 for A. Let $n = f + g$. Take any finite 0L system $G = (A, \sigma, w)$ and any $s \in L(G)$. Then there is a lowest $i \geq 0$ such that s is in $\sigma^i(w)$. Suppose $i > n$. By Theorem 14, $\sigma^g(w) = \sigma^{g+f}(w)$. Then $\sigma^i(w) = \sigma^{i-f}(w)$. Then s is in $\sigma^{i-f}(w)$, a contradiction. So G is step-bounded by n. Hence the class of finite 0L systems is alphabet step-bounded. □

4.2 DT0L

In this subsection, we first give a simple counterexample to show that the class of finite DT0L systems is not alphabet size-bounded. We then show that this class is alphabet step-bounded, first proving a lemma about a more restricted class of systems.

Theorem 16. *The class of finite DT0L systems is not alphabet size-bounded.*

Proof. Let $A = \{\mathtt{a}, \mathtt{b}\}$ and take any $n \geq 0$. Let $w = \mathtt{a}$. For every $1 \leq i \leq n$, let h_i be a morphism on A such that $h_i(\mathtt{a}) = \mathtt{b}^i$ and $h_i(\mathtt{b}) = \mathtt{b}$. Let $H = \{h_1, \ldots, h_n\}$. Let $G = (A, H, w)$. Then $L(G) = \{\mathtt{a}, \mathtt{b}, \mathtt{bb}, \mathtt{bbb}, \ldots, \mathtt{b}^n\}$. So $L(G)$ is finite, but $|L(G)| > n$. So the class of finite DT0L systems is not alphabet size-bounded. □

Take any DT0L system $G = (A, H, w)$. For any $s \in A^*$, let $L(s)$ be the language of the DT0L system (A, H, s). If for every $h \in H$, h is nonerasing, G is called a propagating DT0L system or **PDT0L system** [2].

Lemma 17. *For every alphabet A, there is an $m \geq 0$ such that for every finite PDT0L system $G = (A, H, w)$, for any $h_1, \ldots, h_n \in H$ such that $n > m$, there are j, k such that $0 \leq j < k \leq n$ and $h_j \cdots h_1(w) = h_k \cdots h_1(w)$.*

Proof. Take any alphabet A. Take any $m > (1 + (|A| + 1)!) \cdot |A|^{|A|}$. Take any finite PDT0L system $G = (A, H, w)$. For any $A' \subseteq A$, call $S = \{s_0, s_1, s_2, \ldots, s_n\}$ relevant to A' if every s_i is in A'^*, $L(s_0)$ is finite, and for every $1 \leq i \leq n$, there is an $h \in H$ such that $h(s_{i-1}) = s_i$. Notice that since G is a PDT0L system, for every i, $|s_i| \leq |s_{i+1}|$. Let $Jumps(A', S) = |\{i \mid |s_i| < |s_{i+1}|\}|$. We will show by induction on $|A|$ that for every S relevant to A, $Jumps(A, S) \leq (|A| + 1)!$. Take any $S = \{s_0, s_1, s_2, \ldots, s_n\}$ relevant to A.

For the base case, suppose $|A| = 0$. Then for every $s_i \in S$, $s_i = \lambda$. So $Jumps(A, S) = 0$.

Now suppose for induction that for every $A' \subsetneq A$, for every S' relevant to A', $Jumps(A', S') \leq (|A'| + 1)!$. If $s_0 = \lambda$, clearly $Jumps(A, S) = 0$. So say $s_0 \neq \lambda$. Take any c in s_0. Let $S' = \{s'_0, s'_1, s'_2, \ldots, s'_n\}$, where $s'_0 = c$ and each s'_i is the descendant string in s_i of the c in s_0. We will show that $Jumps(A, S') \leq 1 + |A|!$. If there is no j such that $|s'_j| < |s'_{j+1}|$, then $Jumps(A, S') = 0$. So say there is such a j. Take the first such j. Then $s'_j = d$ for some $d \in A$, since $|s'_0| = 1$. Suppose there is a $k > j$ such that s'_k contains d. Then there is a composition h of morphisms from H such that $h(d)$ contains d and $|h(d)| > 1$. Then for every $i \geq 0$, $|h^i(d)| > i$. But then $L(d)$ is infinite, hence $L(s_0)$ is infinite, a contradiction. So for every $k > j$, s'_k does not contain d. Let $S'' = \{s'_{j+1}, s'_{j+2}, \ldots, s'_n\}$. Then $Jumps(A, S') = 1 + Jumps(A - d, S'')$. So by the induction hypothesis, $Jumps(A, S') \leq 1 + (|A| - 1 + 1)! \leq 1 + |A|!$. Now for each c in s_0, there will be some S' constructed in this way. For any two occurrences of the same c, S' will be the same. Then there are at most $|A|$ distinct S's. Therefore $Jumps(A, S) \leq |A| \cdot (1 + |A|!) \leq (|A| + 1)!$, completing the induction.

Now take any $h_1, \ldots, h_n \in H$ such that $n > m$. For each $0 \leq i \leq n$, let $s_i = h_i \cdots h_1(w)$. Then $s_0 = w$. So $S = \{s_0, s_1, s_2, \ldots, s_n\}$ is relevant to A.

Hence $Jumps(A, S) \leq (|A| + 1)!$. Then since $n > m$, there are j', k' such that $|s_{j'}| = |s_{k'}|$ and $j' + |A|^{|A|} \leq k$. Then every s_i between $s_{j'}$ and $s_{k'}$ has the same length. But at most $|A|^{|A|}$ such s_i are distinct, since each c in $s_{j'}$ has only one descendant in each s_i, and any two occurrences of the same c have the same descendant. So there are j, k such that $0 \leq j' \leq j < k \leq k' \leq n$ and $s_j = s_k$, which was to be shown. □

Theorem 18. *For every alphabet A, there is a $b \geq 0$ such that for every finite DT0L system $G = (A, H, w)$, for any $h_1, \ldots, h_n \in H$ such that $n > b$, there are p, q such that $1 \leq p \leq q \leq n$ and $h_n \cdots h_{q+1} h_{p-1} \cdots h_1(w) = h_n \cdots h_1(w)$.*

Proof. Take any alphabet A. Take any m meeting the conditions of Lemma 17 for A. Take any $b > m \cdot 2^{|A|}$. Take any finite DT0L system $G = (A, H, w)$. Take any $h_1, \ldots, h_n \in H$ such that $n > b$. Let $s = h_n \cdots h_1(w)$. For each $0 \leq i \leq n$, let $s_i = h_i \cdots h_1(w)$. Then $s_0 = w$ and $s_n = s$. For each $0 \leq i < n$, let $f_i = h_n \cdots h_{i+1}$ and let $Stay_i = \{c \in A \mid f_i(c) \neq \lambda\}$. Each $Stay_i$ is one of $2^{|A|}$ possible sets. Then since $n > m \cdot 2^{|A|}$, there is a subset $Stay$ of A and $0 \leq z_0 < z_1 < z_2 < \cdots < z_m < n$ such that for every $0 \leq i \leq m$, $Stay_{z_i} = Stay$. Let $Gone = A - Stay$. Notice that for every $c \in Gone$ and $1 \leq i \leq n$, $h_i(c)$ is in $Gone^*$. For each $x \in A^*$, let $Core(x)$ be the string obtained from x by erasing all occurrences of symbols in $Gone$.

We now construct a finite PDT0L system G'. Let $A' = Stay$ and $w' = Core(s_{z_0})$. For each $1 \leq i \leq m$ and $c \in Stay$, let $h'_i(c) = Core(h_{z_i} \cdots h_{z_{i-1}+1}(c))$. Let $H' = \{h'_1, \ldots, h'_m\}$. Now, for any $c \in Stay$ and any $h'_i \in H'$, clearly $h'_i(c)$ is in $Stay^*$. Further, since $f_{z_{i-1}}(c) \neq \lambda$, $h'_i(c) \neq \lambda$. Therefore $G' = (A', H', w')$ is a PDT0L system. Further, since $L(G)$ is finite, and w' was obtained by erasing letters from a string in $L(G)$, and each h'_i was obtained by composing tables from H and erasing letters from the result, $L(G')$ is finite.

Now for each $0 \leq i \leq m$, let $s'_i = t'_i \cdots t'_1(w)$. Then by Lemma 17, there are j, k such that $0 \leq j < k \leq m$ and $s'_j = s'_k$. Notice that for each $0 \leq i \leq m$, $s'_i = Core(s_{z_i})$. Then $Core(s_{z_j}) = Core(s_{z_k}) = s'_j = s'_k$. Now $f_{z_k}(s_{z_k}) = s$. Therefore $f_{z_k}(Core(s_{z_k})) = s$. Then $f_{z_k}(Core(s_{z_j})) = s$. So then $f_{z_k}(s_{z_j}) = s$. So set $p = z_j + 1$ and $q = z_k$. Then $h_n \cdots h_{q+1} h_{p-1} \cdots h_1(w) = h_n \cdots h_{q+1}(s_{z_j})$ $= f_{z_k}(s_{z_j}) = s$, as desired. □

Theorem 19. *The class of finite DT0L systems is alphabet step-bounded.*

Proof. Take any alphabet A. Take any b meeting the conditions of Theorem 18 for A. Take any finite DT0L system $G = (A, H, w)$. Then by Theorem 18, any derivation in G with more than b steps can be shortened. So G is step-bounded by b. Hence the class of finite DT0L systems is alphabet step-bounded. □

4.3 T0L

Theorem 20. *The class of finite T0L systems is not alphabet step-bounded.*

Proof. Let $A = \{\mathtt{a}, \mathtt{b}, \mathtt{x}\}$ and take any $n \geq 0$. Let $w = (\mathtt{ax})^{n+1}$. For every $1 \leq i \leq n+1$, let σ_i be a finite substitution on A such that $\sigma_i(\mathtt{a}) = \{\mathtt{a}, \mathtt{b}^i\}$, $\sigma_i(\mathtt{b}) = \{\mathtt{b}\}$, and $\sigma_i(\mathtt{x}) = \{\mathtt{x}\}$. Let $T = \{\sigma_1, \ldots, \sigma_{n+1}\}$. Let $G = (A, T, w)$. Clearly G is finite. Let $s = \mathtt{bxbbxbbbx} \cdots \mathtt{b}^{n+1}\mathtt{x}$. Then s can be derived from w in $n+1$ steps, by applying each table in turn to replace an $\mathtt{a}$ by $\mathtt{b}$s. In any derivation of s from w, at each step, at most one $\mathtt{a}$ can be replaced by $\mathtt{b}$s, otherwise s would become unreachable. So at least $n+1$ steps are needed to derive s. Hence G is not step-bounded by n. So the class of finite T0L systems is not alphabet step-bounded. $\square$

Corollary 21. *The class of finite T0L systems is not alphabet size-bounded.*

5 D0L Subsystems

By Corollary 6, every infinite T0L language has an infinite D0L subset. In this section, we consider a related notion, that of a D0L subsystem of a 0L or DT0L system. Such a subsystem not only generates a subset of the original language, but also shares structural characteristics with the original system.

5.1 0L

Let $G = (A, \sigma, w)$ be a 0L system. A **D0L subsystem** of G is a D0L system $G' = (A, h, w)$ such that for every $c \in A$, $h(c)$ is in $\sigma(c)$. Notice that $L(G') \subseteq L(G)$.

Lemma 22. *Take any 0L system $G = (A, \sigma, w)$ with mortal symbols M and vital symbols V. Take any D0L subsystem $G' = (A, h, w)$ of G such that for every $c \in V$, $h(c)$ contains some $d \in V$. Let G' have mortal symbols M' and vital symbols V'. Then $M' = M$ and $V' = V$.*

Proof. Take any $c \in M$. Then $\sigma^i(c) = \{\lambda\}$ for some $i \geq 0$. Then since G' is a D0L subsystem of G, $h^i(c) = \lambda$. So c is in M'. Now take any $c \in V$. We will show that c is in V' by induction on the number n of symbols which are reachable from c under h.

For the base case, suppose $n = 1$. Then only c is reachable from c under h. Then since c is in V, $h(c)$ must contain c. Then c is recursive under h, so c is in V'.

So say $n \geq 1$. Suppose for induction that for every c' in V and $n' < n$ such that n' is the number of symbols reachable from c' under h, c' is in V'. Since c is in V, $h(c)$ contains some $d \in V$. Suppose c is reachable from d. Then c is recursive under h, so c is in V'. So say c is not reachable from d. Then since every symbol reachable from d is reachable from c, the number of symbols reachable from d is at most $n-1$. So by the induction hypothesis, d is in V'. Then c is in V', completing the induction.

So $M \subseteq M'$ and $V \subseteq V'$. Then since $M \cup V = M' \cup V' = A$ and $M' \cap V' = \{\}$, $M' = M$ and $V' = V$. $\square$

Theorem 23. *Every infinite 0L system has an infinite D0L subsystem.*

Proof. Take any infinite 0L system $G = (A, \sigma, w)$. By Corollary 4, there is a derivation $D : s_0 \to \cdots \to s_k \to \cdots \to s_n$ such that $0 \leq k < n$, $s_0 = w$, s_k contains a $c \in A$ whose descendant string in s_n contains distinct occurrences of c and a vital symbol d, and no shorter derivation has these properties. The c in s_n has an ancestor symbol c_i in each s_i, and each c_i generates a string x_i contained in s_{i+1}, for $0 \leq i < n$. Similarly, the d in s_n has an ancestor symbol d_i in each s_i, for $0 \leq i < n$. Let m be the highest i such that $k \leq i < n$ and the d in s_n is descended from c_i. Intuitively, m designates the string containing the last common ancestor of the c and d in s_n. Let h be a morphism on A constructed as follows. First, for i from $n-1$ down to 0, set $h(c_i) = x_i$ unless $h(c_i)$ has already been set. Then for all $e \in A$ for which $h(e)$ has not been set, if some $s \in \sigma(e)$ contains a vital symbol, set $h(e)$ to any such s, otherwise set $h(e)$ to any $s \in \sigma(e)$. Then $G' = (A, h, w)$ is a D0L subsystem of G. We will show that G' is infinite.

First we show that c is reachable from w under h. Take any i such that $0 \leq i < n$. We will show by induction that c is reachable from c_i under h. For the base case, suppose $i = n-1$. Clearly c is reachable from c_{n-1} under h, since $h(c_{n-1}) = x_{n-1}$, which contains c. So say $i < n-1$. Suppose for induction that for all j such that $i < j < n$, c is reachable from c_j under h. Suppose there is a j such that $i < j < n$ and $c_j = c_i$. Then by the induction hypothesis, c is reachable from $c_j = c_i$ under h. So say there is no such j. Then $h(c_i) = x_i$, which contains c_{i+1}. By the induction hypothesis, c is reachable from c_{i+1} under h, hence c is reachable from c_i under h, completing the induction. So c is reachable from c_0 under h, hence c is reachable from w under h.

Next we show that there are no i, j such that $k \leq i < j < n$ and $c_i = c_j$. Suppose there are such i, j. Suppose $k \leq i < j \leq m$. Then c_m could be derived from c_i in $m-j$ steps instead of $m-i$ steps, so D is not minimal, a contradiction. So suppose $m < i < j < n$. Then the steps from i to j could be skipped, so that at step $n-(j-i)$, c_i would reach c and d_i would reach $d_{n-(j-i)}$, which is vital. But then D is not minimal. So suppose $k \leq i \leq m < j < n$. Then the descendant string in s_j of the c_i in s_i contains distinct occurrences of c_i and d_j. But then D could be shortened from length n to length j. So there are no such i, j.

So for all $k \leq i < n$, $h(c_i) = x_i$. Hence c_m is reachable from c under h and c is reachable from c_{m+1} under h. Now $h(c_m) = x_m$, which contains distinct occurrences of c_{m+1} and d_{m+1}. By Lemma 22, d_{m+1} is vital under h. Then some string s containing distinct occurrences of c and a vital symbol can be derived from x_m under h. Then some string containing s can be derived from c under h. Hence c is recursive under h, and not monorecursive under h.

Then since c is reachable, recursive, and not monorecursive in G', G' is infinite by Lemma 9. Therefore every infinite 0L system has an infinite D0L subsystem. ◻

5.2 DT0L

Let $G = (A, H, w)$ be a DT0L system. A **D0L subsystem** of G is a D0L system $G' = (A, h, w)$ such that h is in H. Notice that $L(G') \subseteq L(G)$.

Theorem 24. *There is an infinite DT0L system with no infinite D0L subsystem.*

Proof. Let $A = \{\mathtt{a}, \mathtt{b}\}$. Let h_1 and h_2 be morphisms on A such that $h_1(\mathtt{a}) = \mathtt{ab}$, $h_1(\mathtt{b}) = \lambda$, $h_2(\mathtt{a}) = \lambda$, and $h_2(\mathtt{b}) = \mathtt{bb}$. Let $H = \{h_1, h_2\}$. Let $w = \mathtt{a}$. Then the DT0L system (A, H, w) is infinite but has no infinite D0L subsystem. $\square$

6 Conclusion

In this paper we have extended to 0L, DT0L, and T0L systems the work of Vitányi [7] on infiniteness and boundedness of D0L systems. In doing so, we relaxed the condition of alphabet size-boundedness (which holds for the class of finite D0L systems) to one of alphabet step-boundedness (which holds also for the classes of finite 0L and DT0L systems). One direction for further work would be to find a related boundedness condition which holds for the class of finite T0L systems. We have also shown that every infinite T0L language has an infinite D0L subset, and that every infinite 0L system has an infinite D0L subsystem. It would be interesting to see whether in classes of L systems beyond the ones studied here, infiniteness is similarly characterized by the presence of notable infinite subsets and subsystems.

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

References

1. Jungers, R.M., Protasov, V., Blondel, V.D.: Efficient algorithms for deciding the type of growth of products of integer matrices. Linear Algebra and its Applications 428, 2296–2311 (2008)
2. 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)
3. Nishida, T.: Quasi-deterministic 0L systems. In: Kuich, W. (ed.) Automata, Languages and Programming, Lecture Notes in Computer Science, vol. 623, pp. 65–76. Springer Berlin / Heidelberg (1992)
4. Nishida, T.Y., Salomaa, A.: Slender 0L languages. Theoretical Computer Science 158(1–2), 161–176 (1996)
5. Rabkin, M.: Ogden's lemma for ET0L languages. In: Proceedings of the 6th International Conference on Language and Automata Theory and Applications. pp. 458–467. LATA'12, Springer-Verlag, Berlin, Heidelberg (2012)
6. Rozenberg, G., Salomaa, A.: Mathematical Theory of L Systems. Academic Press, Inc., Orlando, FL, USA (1980)
7. Vitányi, P.: On the size of D0L languages. In: Rozenberg, G., Salomaa, A. (eds.) L Systems, Lecture Notes in Computer Science, vol. 15, pp. 78–92. Springer Berlin / Heidelberg (1974)