SELF-REPRODUCING PUSHDOWN TRANSLATION

Ing. Lubos Lorenc, Doctoral Degree Programme (3)
Dept. of Information Systems, FIT, BUT
E-mail: lorenc@fit.vutbr.cz

Supervised by: Dr. Alexander Meduna

ABSTRACT

After a translation of an input string, x, to an output string, y, a self-reproducing pushdown transducer can make a self-reproducing step during which it moves y to its input tape and translates it. In this self-reproducing way, it can repeat the translation n-times for any $n \geq 1$. This paper demonstrates that every recursively enumerable language can be characterized by the domain or the range of the translation obtained from a self-reproducing pushdown transducer that repeats its translation no more than three times.

1 INTRODUCTION

Self-reproducing pushdown transducer represents a natural modified version of an ordinary pushdown transducer. The characterization described in the abstract is of some interest because it does not hold in terms of ordinary pushdown transducers because the domain or range obtained from any ordinary pushdown transducer is a context-free language.

2 DEFINITIONS

A *self-reproducing pushdown transducer* is a 8-tuple $M = (Q, \Gamma, \Sigma, \Omega, R, s, S, O)$, where Q is a finite set of states, Γ is a total alphabet such that $Q \cap \Gamma = \varnothing$, $\Sigma \subseteq \Gamma$ is an input alphabet, $\Omega \subseteq \Gamma$ is an output alphabet, R is a finite set of *translation rules* of the form $u_1qw \rightarrow u_2pv$ with $u_1, u_2, w, v \in \Gamma^*$ and $q, p \in Q$, $s \in Q$ is the *start state*, $S \in \Gamma$ is the *start pushdown symbol*, $O \subseteq Q$ is the set of *self-reproducing states*. A *configuration of* M is any string of the form $\$zqy\$x$, where $x, y, z \in \Gamma^*$, $q \in Q$, and \$ is a special *bounding symbol* ($\$ \notin Q \cup \Gamma$). If $u_1qw \rightarrow u_2pv \in R$, $y = \$hu_1qwz\$t$, and $x = \$hu_2pz\$tv$, where $h, u_1, u_2, w, t, v, z \in \Gamma^*$, $q, p \in Q$, then M makes a *translation step* from y to x in M, symbolically written as $y \ {}_t\!\Rightarrow x \ [u_1qw \rightarrow u_2pv]$ or, simply $y \ {}_t\!\Rightarrow x$ in M. If $y = \$hq\$t$, and $x = \$hqt\$$, where $t, h \in \Gamma^*$, $q \in O$, then M makes a *self-reproducing step* from y to x in M, symbolically written as $y \ {}_r\!\Rightarrow x$. Write $y \Rightarrow x$ if $y \ {}_t\!\Rightarrow x$ or $y \ {}_r\!\Rightarrow x$. In The standard manner, extend $\Rightarrow$ to $\Rightarrow^n$, where $n \geq 0$; then based on $\Rightarrow^n$ define $\Rightarrow^+$ and $\Rightarrow^*$. Let

$w, v \in \Gamma^*$; M *translates* w to v if $\$Ssw\$ \Rightarrow^* \$q\$v$ in M. *The translation obtained from* M, $T(M)$, is defined as $T(M) = \{(w, v) : \$Ssw\$ \Rightarrow^* \$q\$v$ with $w \in \Sigma^*,\ v \in \Omega^*,\ q \in Q\}$. Set *Domain*$(T(M)) = \{w : (w, x) \in T(M)\}$ and *Range*$(T(M)) = \{x : (w, x) \in T(M)\}$.

3 RESULTS

Lemma 1. *For every recursively enumerable language, L, there exists a left-extended queue grammar, Q, satisfying $L(Q) = L$.*

Proof. Recall that every recursively enumerable language is generated by queue grammar (see [2]). Clearly, for every queue grammar, there exists an equivalent left-extended queue grammar. Thus, this lemma holds. □

Lemma 2. *Let Q' be an left-extended queue grammar. Then there exists a left-extended queue grammar, $Q = (V, T, W, F, s, R)$, such that $L(Q') = L(Q), W = X \cup Y \cup \{1\}$, where $X, Y, \{1\}$ are pairwise disjoint, and every $(a, b, x, c) \in R$ satisfies either $a \in V - T,\ b \in X,\ x \in (V - T)^*,\ c \in X \cup \{1\}$ or $a \in V - T,\ b \in Y \cup \{1\},\ x \in T^*,\ c \in Y$. Q generates every $h \in L(Q)$ in this way*

$$
\begin{array}{ll}
\#a_0q_0 & \\
\Rightarrow a_0\#x_0q_1 & [(a_0, q_0, z_0, q_1)] \\
\Rightarrow a_0a_1\#x_1q_2 & [(a_1, q_1, z_1, q_2)] \\
\vdots & \\
\Rightarrow a_0a_1 \dots a_k\#x_kq_{k+1} & [(a_k, q_k, z_k, q_{k+1})] \\
\Rightarrow a_0a_1 \dots a_ka_{k+1}\#x_{k+1}y_1q_{k+2} & [(a_{k+1}, q_{k+1}, y_1, q_{k+2})] \\
\vdots & \\
\Rightarrow a_0a_1 \dots a_ka_{k+1} \dots a_{k+m-1}\#x_{k+m-1}y_1 \dots y_{m-1}q_{k+m} & [(a_{k+m-1}, q_{k+m-1}, y_{m-1}, q_{k+m})] \\
\Rightarrow a_0a_1 \dots a_ka_{k+1} \dots a_{k+m}\#y_1 \dots y_mq_{k+m+1} & [(a_{k+m}, q_{k+m}, y_m, q_{k+m+1})]
\end{array}
$$

where $k, m \geq 1,\ a_i \in V - T$ *for* $i = 0, \dots, k + m,\ x_j \in (V - T)^*$ *for* $j = 1, \dots, k + m - 1,\ s = a_0q_0,\ a_jx_j = x_{j-1}z_j$ *for* $j = 1, \dots, k,\ a_1 \dots a_kx_k = z_0 \dots z_k,\ a_{k+1} \dots a_{k+m} = x_k,\ q_0, q_1, \dots q_{k+m} \in W - F$ *and* $q_{k+m+1} \in F,\ z_1, \dots, z_k \in (V - T)^*,\ y_1, \dots, y_m \in T^*,\ h = y_1y_2 \dots y_{m-1}y_m,\ q_{k+1} \in \{1\}$.

Proof. See Lemma 1 in [4]. □

For brevity, the following proofs are only sketches, because the full proofs are too long to fit in this paper.

Lemma 3. *Let Q be a left-extended queue grammar satisfying the properties given in lemma 2. Then, there exists a 2-self-reproducing pushdown transducer, M, such that Domain$(T(M)) = L(Q)$ and Range$(T(M)) = \{\varepsilon\}$.*

Proof. Let $G = (V, T, W, F, s, P)$ be a left-extended queue grammar satisfying the properties given in lemma 2. Without any loss of generality, assume that $\{0, 1\} \cap (V \cup W) = \varnothing$. For some positive integer, n, define an injection, ι, from P to $(\{0, 1\}^n - \{1\}^n)$ so that ι is an injective homomorphism when its domain is extended to $(VW)^*$; after this extension,

ι thus represents an injective homomorphism from $(VW)^*$ to $(\{0,1\}^n - \{1\}^n)^*$; a proof that such an injection necessarily exists is simple and left to the reader. Based on ι, define the substitution, ν, from V to $(\{0,1\}^n - \{1\}^n)$ so that for every $a \in V,\ \nu(a) = \{\iota(p) : p \in P,\ p = (a,b,x,c)$ for some $x \in V^*;\ b,c \in W\}$. Extend the domain of ν to V^*. Furthermore, define the substitution, μ, from W to $(\{0,1\}^n - \{1\}^n)$ so that for every $q \in W,\ \mu(q) = \{\iota(p) : p \in P,\ p = (a,b,x,c)$ for some $a \in V,\ x \in V^*;\ b,c \in W\}$. Extend the domain of μ to W^*.

Construction 1 (of M). Introduce the self-reproducing pushdown transducer

$$M = (Q, T \cup \{0,1,S\}, T, \varnothing, R, z, S, O)$$

where $Q = \{o,f,z\} \cup \{\langle p,i\rangle : p \in W \text{ and } i \in \{1,2\}\},\ O = \{o,f\}$, and R is constructed by performing the following steps 1 through 6.

1. if $a_0q_0 = s$, where $a \in V - T$ and $q \in W - F$,
   then add $Sz \to uS\langle q_0,1\rangle w$ to R, for all $w \in \mu(q_0)$ and all $u \in \nu(a_0)$;
2. if $(a,q,y,p) \in P$, where $a \in V - T,\ p,q \in W - F$, and $y \in (V-T)^*$,
   then add $S\langle q,1\rangle \to uS\langle p,1\rangle w$ to R, for all $w \in \mu(p)$ and $u \in \nu(y)$;
3. for every $q \in W - F$, add $S\langle q,1\rangle \to S\langle q,2\rangle$ to R
4. if $(a,q,y,p) \in P$, where $a \in V - T,\ p,q \in W - F$, and $y \in T^*$,
   then add $S\langle q,2\rangle y \to S\langle p,2\rangle w$ to R, for all $w \in \mu(p)$;
5. if $(a,q,y,p) \in P$, where $a \in V - T,\ q \in W - F,\ y \in T^*$, and $p \in F$,
   then add $S\langle q,2\rangle y \to SoS$ to R;
6. add $o0 \to 0o,\ o1 \to 1o,\ oS \to c,\ 0c \to c0,\ 1c \to c1,\ Sc \to f,\ 0f0 \to f,\ 1f1 \to f$ to R.

Claim 1. *M accepts every $h \in L(M)$ in this way*

$$\begin{aligned}
& \$Szy_1y_2\dots y_{m-1}y_m\$ \\
\Rightarrow\ & \$g_0\langle q_0,1\rangle y_1y_2\dots y_{m-1}y_m\$t_0 \\
\Rightarrow\ & \$g_1\langle q_1,1\rangle y_1y_2\dots y_{m-1}y_m\$t_1 \\
& \vdots \\
\Rightarrow\ & \$g_k\langle q_k,1\rangle y_1y_2\dots y_{m-1}y_m\$t_k \\
\Rightarrow\ & \$g_k\langle q_k,2\rangle y_1y_2\dots y_{m-1}y_m\$t_k \\
\Rightarrow\ & \$g_k\langle q_{k+1},2\rangle y_1y_2\dots y_{m-1}y_m\$t_{k+1} \\
\Rightarrow\ & \$g_k\langle q_{k+2},2\rangle y_2\dots y_{m-1}y_m\$t_{k+2} \\
& \vdots \\
{}_t\Rightarrow\ & \$g_k\langle q_{k+m},2\rangle y_m\$t_{k+m} \\
{}_t\Rightarrow\ & \$g_kSo\$t_{k+m}S \\
{}_r\Rightarrow\ & \$g_kSot_{k+m}S\$ \\
{}_t\Rightarrow^{\iota}\ & \$g_kSt_{k+m}oS\$ \\
{}_t\Rightarrow\ & \$g_kSt_{k+m}c\$ \\
{}_t\Rightarrow^{\iota}\ & \$u_1Sc\$v_1 \\
{}_t\Rightarrow\ & \$u_1f\$v_1 \\
{}_r\Rightarrow\ & \$u_1fv_1\$ \\
\Rightarrow\ & \$u_2fv_2\$ \\
& \vdots \\
\Rightarrow\ & \$u_\varpi fv_\varpi\$ \\
\Rightarrow\ & \$f\$
\end{aligned}$$

in M, where $k, m \geq 1$; $q_0, q_1, \ldots, q_{k+m} \in W - F$; $y_1, \ldots, y_m \in T^*$; $t_i \in \mu(q_0 q_1 \ldots q_i)$ for $i = 0, 1, \ldots, k+m$; $g_j \in \nu(d_0 d_1 \ldots d_j)$ with $d_1, \ldots, d_j \in (V-T)^*$ for $j = 0, 1, \ldots, k$; $d_0 d_1 \ldots d_k = a_0 a_1 \ldots a_{k+m}$ where $a_1, \ldots, a_{k+m} \in V - T$, $d_0 = a_0$, and $s = a_0 q_0$; $g_k = t_{k+m}$ (that is, $\nu(a_0 a_1 \ldots a_{k+m})$ and $\mu(q_0 q_1 \ldots q_{k+m})$ are identical); $v_i \in$ Prefix$(\mu(q_0 q_1 \ldots q_{k+m}), |\mu(q_0 q_1 \ldots q_{k+m})| - i)$ for $i = 1, \ldots, \upsilon$ with $\upsilon = |\mu(q_0 q_1 \ldots q_{k+m})|$; $u_j \in$ Suffix$(\nu(a_0 a_1 \ldots a_{k+m}), |\nu(a_0 a_1 \ldots a_{k+m})| - j)$ for $j = 1, \ldots, \varpi$ with $\varpi = |\nu(a_0 a_1 \ldots a_{k+m})|$; $h = y_1 y_2 \ldots y_{m-1} y_m$.

Proof of the claim. Examine steps 1 through 6 of the construction of R. Notice that in every successful computation, M uses the rules introduced in step i before it uses the rules introduced in step $i+1$, for $i = 1, \ldots, 5$. During $\$ g_k S o \$ t_{k+m} S \Rightarrow^* \$ f \$$ only the rules of 6 are used. Recall these rules: $o0 \rightarrow 0o$, $o1 \rightarrow 1o$, $oS \rightarrow c$, $0c \rightarrow c0$, $1c \rightarrow c1$, $Sc \rightarrow f$, $0f0 \rightarrow f$, $1f1 \rightarrow f$. This computation implies $g_k = t_{k+m}$. As a result, the claim holds. □

Let M accepts $h \in L(M)$ in the way described in the above claim. Examine the construction of R to see that at this point P contains (a_0, q_0, z_0, q_1), $\ldots$, (a_k, q_k, z_k, q_{k+1}), $(a_{k+1}, q_{k+1}, y_1, q_{k+2})$, $\ldots$, $(a_{k+m-1}, q_{k+m-1}, y_{m-1}, q_{k+m})$, $(a_{k+m}, q_{k+m}, y_m, q_{k+m+1})$, where $z_1, \ldots, z_k \in (V-T)^*$, so G makes the generation of h in the way described in lemma 2. Thus $h \in L(G)$. Consequently $L(M) \subseteq L(G)$. Let G generates $h \in L(G)$ in the way described in lemma 2. Then, M accepts h in the way described in the above claim, so $L(G) \subseteq L(M)$. As $L(M) \subseteq L(G)$ and $L(G) \subseteq L(M)$, $L(G) = L(M)$. From the above claim, it follows that M is a 2-self-reproducing pushdown transducer. Thus, lemma 3 holds. □

Lemma 4. *Let Q be a left-extended queue grammar satisfying the properties given in lemma 2. Then, there exists a 2-self-reproducing pushdown transducer, M, such that Domain$(T(M)) = \{\varepsilon\}$ and Range$(T(M)) = L(Q)$.*

Proof. Let $G = (V, T, W, F, s, P)$ be a left-extended queue grammar satisfying the properties given in lemma 2.
Construction of M. Introduce the self-reproducing pushdown transducer

$$M = (Q, V \cup \{S\}, \varnothing, T, R, z, S, O)$$

where $Q = \{z \cup \{\langle p, i\rangle : p \in W, i \in \{1, 11\}\} \cup \{\langle 1, i\rangle : i \in \{1, \ldots, 11\}\}\}$, $O = \{\langle 1, 3\rangle, \langle 1, 8\rangle\}$ and R is constructed by performing the following steps 1 through 6.

1. if $y_s q_0 = s$, where $y \in V - T$ and $q \in W - F$,
   then add $Sz \rightarrow SS\langle q_0, 1\rangle S y_s$ to R

2. if $(a, q, y, p) \in P$, where $a \in V - T$, $p, q \in W - F$, and $y \in (V-T)^*$,
   then add $S\langle q, 1\rangle \rightarrow aS\langle p, 1\rangle y$ to R

3. add following rules to R

   $S\langle 1,1\rangle \rightarrow \langle 1,2\rangle S$, $a\langle 1,2\rangle \rightarrow \langle 1,2\rangle a$, $S\langle 1,2\rangle \rightarrow \langle 1,3\rangle S$, $\langle 1,3\rangle S \rightarrow S\langle 1,4\rangle$,
   $\langle 1,4\rangle a \rightarrow a\langle 1,4\rangle$, $\langle 1,4\rangle S \rightarrow \langle 1,5\rangle S$, $a\langle 1,5\rangle \rightarrow \langle 1,5\rangle a$, $S\langle 1,5\rangle \rightarrow S\langle 1,6\rangle$,
   $\langle 1,6\rangle a \rightarrow a\langle 1,6\rangle$, $\langle 1,6\rangle S \rightarrow \langle 1,7\rangle S$, $a\langle 1,7\rangle \rightarrow \langle 1,7\rangle a$, $S\langle 1,7\rangle \rightarrow \langle 1,8\rangle S$,
   $\langle 1,8\rangle S \rightarrow S\langle 1,9\rangle$, $\langle 1,9\rangle a \rightarrow a\langle 1,9\rangle$, $\langle 1,9\rangle S \rightarrow \langle 1,10\rangle$, $a\langle 1,10\rangle a \rightarrow \langle 1,10\rangle$

4. if $(a, q, y, p) \in P$, where $a \in V - T,\ p = 1, q \in W - F$, and $y \in T^*$, then add $a\langle 1, 10\rangle S \to \langle p, 11\rangle y$
5. if $(a, q, y, p) \in P$, where $a \in V - T,\ p, q \in W - F$, and $y \in T^*$, then add $a\langle q, 11\rangle \to \langle p, 11\rangle y$ to R
6. if $(a, q, y, p) \in P$, where $a \in V - T,\ q \in W - F, y \in T^*$, and $p \in F$, then add $Sa\langle q, 11\rangle \to \langle p, 11\rangle y$ to R

Proof of the claim. Examine steps 1 through 6 of the construction of R. Notice that in every successful computation, M uses the rules introduced in step i before it uses the rules introduced in step $i + 1$, for $i = 1, \ldots, 5$. Thus, in greater detail, every successful computation can be splitted into three main steps. In these steps the self-reproducing pushdown transducer simulates the work of queue grammar. In the first step only the rules of 1 and 2 are used. There are generated two strings. To the output tape is generated the string of nonterminals and to the pushdown is generated the string of nonterminals that the queue grammar rewrites until its state is q_{k+1}. In the second step the generated strings are swapped, reverted and compared using rules of 3. These rules simulates rewriting of nonterminals in the queue grammar. After this simulation using one rule from 4 the self-reproducing transducer reaches the last part of simulation. In the last part only rules of 5 (or 6 for the last move) are used. Using these rules the remaining nonterminals are translated into resulting string of terminals h, where h is the string generated by G. Thus, the claim holds. A detailed proof is left to the reader. □

□

Theorem 1. *For every recursively enumerable language, L, there exists a 2-self-reproducing pushdown transducer, M, such that Domain$(T(M)) = L$ and Range$(T(M)) = \{\varepsilon\}$ or Domain$(T(M)) = \{\varepsilon\}$ and Range$(T(M)) = L$.*

Proof. This theorem follows from lemmas 1, 2, 3 and 4. □

REFERENCES

[1] Harrison, M. A.: Introduction to Formal Language Theory. Addison-Wesley, Reading, 1978.

[2] Kleijn, H. C. M., Rozenberg, G.: On the Generative Power of Regular Pattern Grammars. Acta Informatica, Vol. 20, pp. 391–411, 1983.

[3] Meduna, A.: Automata and Languages: Theory and Applications. Springer, London, 2000.

[4] Meduna, A.: Simultaneously One-Turn Two-Pushdown Automata. Inter. Computer Math., Vol. 80, pp. 679–687, 2003

[5] Meduna, A., Lorenc, L.: Self-Reproducing Pushdown Transducers. In: Proceedings of 7th International Conference ISIM'04 Information Systems Implementation and Modelling, Ostrava, CZ, MARQ, pp. 155–160, 2004