A NOTE ON PUSHDOWN STORE AUTOMATA AND REGULAR SYSTEMS[1]

SHEILA A. GREIBACH

Recent work on pushdown store automata has focused attention on various sets of pushdown store tapes [8]. Certain sets of tapes associated with pushdown store automata can be proved regular. As a consequence we obtain a new proof of a theorem due to Büchi:[2] that regular canonical systems (i.e., productions of the form $\alpha Q \rightarrow \beta Q$) produce regular sets [2].[3]

In this paper we shall use a theorem of Bar-Hillel, Perles and Shamir [1] to show that the set of tapes left on the pushdown store by a regular set is regular,[4] and derive Büchi's theorem from that result.

First we shall need some definitions. We assume familiarity with the definition of production systems.[5]

DEFINITION. *A finite state grammar* is a quadruple $G=(I, T, X, P)$, where I and T are finite sets, $I \cap T = \varnothing$, $X \in I$ and P is a finite set of semi-Thue productions of the forms

$$Q_1 Z Q_2 \rightarrow Q_1 a Y Q_2, \qquad Q_1 Z Q_2 \rightarrow Q_1 a Q_2, \qquad Z,\ Y \in I,\ a \in T \cup \{\lambda\}.$$

A set L is *regular* iff $L = \{w \in T^* \mid X \Rightarrow^* w\}$ for some finite state grammar G.[6]

We must now define pushdown store automata and their actions.

Received by the editors January 17, 1966.

[1] The research reported in this paper was sponsored in part by the Air Force Cambridge Research Laboratories, Office of Aerospace Research, under Contract AF 19(682)-5166, CRL-Algorithmic Languages Program.

[2] Büchi's proof is more elementary; the proof in this paper is shorter and less complicated because it follows from other results.

[3] This theorem is not to be confused with Chomsky's observation that finite state grammars generate regular sets [3], or the theorem of Evey [6] and Matthews [13] that left generations of semi-Thue systems produce context-free sets; the systems involved are different in form.

[4] This result is part of the folklore on the subject, but, as far as this author is aware, has never appeared in print. Analogous theorems are proven in [8] and [9] by different methods; the present approach could have been used in [8].

[5] See [5], [14], [15] and [17] for further discussion of productions and combinatorial systems.

Notation. If $Q_1 \alpha Q_2 \rightarrow Q_1 \beta Q_2$ is a semi-Thue production, we write $w_1 \alpha w_2 \Rightarrow w_1 \beta w_2$ for any strings w_1, w_2. $\Rightarrow^*$ denotes the transitive closure of $\Rightarrow$.

[6] For any set R, the *closure of* R, denoted by R^*, is the free semigroup (with identity λ) generated by R.

Justification for this definition appears in [1] and [3]; different but equivalent characterizations of regular sets appear in [2], [12] and [16].

DEFINITION. A *pushdown store automaton* (pda) is a septuple $M = (K, \Sigma, \Gamma, \delta, q_0, \$, F)$, where (1) K, Σ, Γ are finite sets, $\$ \in \Gamma$, $F \subseteq K$, (2) δ is a function from $Kx(\Sigma \cup \{\lambda\})x\Gamma$ into the finite subsets of $Kx\Gamma^*$.

$$(q, ay, Aw') \vdash_M (q', y, ww') \text{ if } (q', w) \in \delta(q, a, A),$$
$$w' \in \Gamma^*, \; a \in \Sigma \cup \{\lambda\},$$

$A \in \Gamma$ and $y \in \Sigma^*$. $\vdash_M^*$ is the transitive closure of $\vdash_M$.[7] Null $(M) = \{w \mid \exists q \in F, (q_0, w, \$) \vdash_M^* (q, \lambda, \lambda)\}$.

Intuitively, Null (M) is the set of all input tapes that empty the pushdown store and cause the pda to enter a final state at the end.

DEFINITION. L is *context-free* iff $L = \text{Null}(M)$ for some pda M.[8]

In order to state the necessary results clearly, we give the following definitions.

DEFINITION. L is a *ucv*-language iff for some finite vocabulary T and some $c \notin T$, $L \subseteq T^*cT^*$. If L is a *ucv*-language, let

$$f_L(u) = \{v \mid ucv \in L\}, \qquad g_L(v) = \{u \mid ucv \in L\},$$
$$U(L) = \{u \mid f_L(u) \neq \emptyset\}, \quad V(L) = \{v \mid g_L(v) \neq \emptyset\}.$$

We can now state the relevant theorem of Bar-Hillel, Perles and Shamir [1] as:

THEOREM 1. *Let L be a context-free ucv-language. If for every u, $f_L(u)$ is finite, then $V(L)$ is regular. If for every v, $g_L(v)$ is finite, then $U(L)$ is regular.*

We now focus attention on the tapes left on the pushdown store when reading any member of a given regular set and ending in a given state.

THEOREM 2. *Let M: $(K, \Sigma, \Gamma, \delta, q_0, \$, F)$ be a pda. Let $q \in K$, and let $R \subseteq \Sigma^*$ be regular. Then*

$$V_q = \{u \mid \exists w \in R, (q_0, w, \$) \vdash_M^* (q, \lambda, u)\}$$

is regular.

PROOF. Let c be a new symbol. First we shall see that

[7] The notation used here is the reverse of that employed by the author elsewhere [8]; here we read pushdown store tapes from left to right for convenience; if we considered productions $Q\alpha \rightarrow Q\beta$, the other notation would be preferable.

[8] Context-free languages are usually defined by special semi-Thue systems; see [1] or [3]. The equivalence of the present definition to the standard one can be easily obtained as a corollary to results in [4] and [6]; a particularly clean proof appears in [7].

$$L_1 = \{wcu \mid (q_0, w, \$) \vdash_M^* (q, \lambda, u)\}$$

is context-free. We modify the pda M to produce a pda M_1, which imitates M unless and until it sees c in state q. Then it empties the pushdown store, checking against the input tape. Clearly M_1 can be constructed, Null $(M_1) = L_1$ and $V_q = V(L_1 \cap Rc\Gamma^*)$. If $f_{L_1}(w)$ is finite for all $w \in R$, we are done. But the λ-rules may allow M infinitely many actions on one input tape and hence one input might leave infinitely many distinct tapes on the pushdown store. So, instead of λ-rules, we use a dummy symbol (or "clock pulse") $d \notin \Sigma \cup \Gamma \cup \{c\}$. M_1 is modified to produce a pda M_2 which behaves like M_1, except that where, for s in K and A in Γ,

$$\delta(s, \lambda, A) = \delta_1(s, \lambda, A) \neq \varnothing$$

we have in M_2:

$$\delta_2(s, \lambda, A) = \varnothing \quad \text{and} \quad \delta_2(s, d, A) = \delta(s, \lambda, A).$$

Let $L_2 = \mathrm{Null}(M_2)$. Let $\psi(a) = a$ for $a \neq d$ and $\psi(d) = \lambda$. Let $L_3 = L_2 \cap (\psi^{-1}(Rc\Gamma^*))$. Clearly, $V_q = V(L)_3$, and for each w, $f_{L_3}(w)$ is finite. L_2 is context-free. The inverse of a homomorphism preserves regularity [**10**], and the intersection of a context-free language and a regular language is context-free [**1**]. Hence L_3 is context-free and, by Theorem 1, $V_q = V(L_3)$ is regular.

Now we must define regular canonical systems to derive the desired results.

DEFINITION. A *regular canonical system* is a quintuple $R = (I, T, U, V, P)$, where

(1) I and T are finite sets and $I \cap T = \varnothing$,

(2) $U, V \subseteq (I \cup T)^*$, and

(3) P is a finite set of regular productions of the form

$$\alpha Q \to \beta Q \qquad \alpha, \beta \in (I \cup T)^*.$$

Notation. When treating a regular canonical system R we shall write R-deductions as $u \Rightarrow_p v$ or $u \Rightarrow_p^* v$.[9]

DEFINITION. Let $R = (I, T, U, V, P)$ be a regular canonical system. The set $\tau(U, P, V)$ of words *produced* by R is

$$\tau(U, P, V) = \{x \in T^* \mid \exists\, u \in U,\ v \in V,\ u \Rightarrow_p^* vx\}.$$

The set $\beta(U, P, V)$ of words *accepted* by R is

[9] If $\alpha Q \to \beta Q$ is in P, then $\alpha u \Rightarrow_p \beta u$ for any $u \in (I \cup T)^*$. If $w_i \Rightarrow_p w_{i+1}$ for $1 \leq i < n$, then $w_1 \Rightarrow_p^* w_n$.

$$\beta(U, P, V) = \{x \in T^* \mid \exists u \in U, v \in V, ux \Rightarrow_p^* v\}.$$

$\tau(U, P, \{\lambda\})$ is the set of theorems of R.

Büchi showed that if U is finite, then $\tau(U, P, \{\lambda\})$ is regular and, moreover, all regular sets can be obtained in this fashion [**2**]. We shall see that, if U, V are any regular sets over $I \cup T$, then $\tau(U, P, V)$ and $\beta(U, P, V)$ are regular.

Notation. If a is an individual symbol, $\check{a}=a$; $\check{\lambda}=\lambda$. If $x=A_1 \cdots A_m$ is a string, then $\check{x}=\check{A}_m \cdots \check{A}_1$. If S is a set, $\check{S}=\{x \mid x \in S\}$. S is regular iff $\check{S}$ is regular [**16**].

THEOREM 3. *Let $R=(I, T, U, V, P)$ be a regular canonical system. Let U and V be regular sets (finite or infinite). Then $\tau(U, P, V)$ is regular.*

PROOF. We now construct a special pda M that accepts members of $(I \cup T)^*$ as input, places them (reversed) on the pushdown store, and proceeds to imitate the deductions of R.

Let $n = \operatorname{Max}\{l(\alpha) \mid \exists \beta, \alpha \boldsymbol{Q} \rightarrow \beta \boldsymbol{Q} \in P\}$.[10] Let $M=(K, \Sigma, \Gamma, \delta, q_0, \$, F)$. We define:

$K=\{q(w) \mid w \in (I \cup T)^*, 0 \leqq l(w) \leqq n\} \cup \{q_0, q_f\}$, each $q(w)$ a new symbol,

$\Gamma = I \cup T \cup \{\$\}$, $\$$ a symbol not in $I \cup T$,

$\Sigma = I \cup T$,

$F = \{q_f\}$.

δ is defined in the following parts:

(I) For all $A \in I \cup T$, $B \in I \cup T \cup \{\$\}$, we set

$$\delta(q_0, A, B) = \{(q_0, AB), (q(\lambda), AB)\}.$$

$$\delta(q_0, \lambda, \$) = \{(q(\lambda), \$)\}.$$

(II) For each $A \in I \cup T$ and $w \in (I \cup T)^*$, with $0 \leqq l(w) < n$,

$$(q(wA), \lambda) \in \delta(q(w), \lambda, A).$$

(III) For each $\alpha \boldsymbol{Q} \rightarrow \beta \boldsymbol{Q}$ in P and each $A \in I \cup T \cup \{\$\}$,

$$(q(\lambda), \beta A) \in \delta(q(\alpha), \lambda, A).$$

(IV) For each $A \in I \cup T \cup \{\$\}$,

$$(q_f, A) \in \delta(q(\lambda), \lambda, A).$$

The parts of the pda work as follows:

---

[10] $l(\alpha)$ is the length of the string α, $l(\lambda)=0$.

(I) $(q_0, \check{u}, \$) \vdash_M^* (q(\lambda), \lambda, u\$)$ for $u \in (I \cup T)^*$.

(II) $(q(\lambda), \lambda, wy) \vdash_M^* (q(w), \lambda, y)$ iff $0 \leqq l(w) \leqq n$.

(III) $(q(\alpha), \lambda, y) \vdash_M (q(\lambda), \lambda, \beta y)$ iff $\alpha Q \to \beta Q$ is in P.

(IV) $(q(\lambda), \lambda, w\$) \vdash_M (q_f, \lambda, w\$)$.

Putting this together we get

$$(q_0, \check{u}, \$) \vdash_M^* (q(\lambda), \lambda, u\$) \vdash_M^* (q(\lambda), \lambda, v\$) \vdash_M (q_f, \lambda, v\$) \qquad \text{iff } u \Rightarrow_p^* v.$$

Since U is regular, so is $\check{U}$. By Theorem 2, the set

$$V_{q_f} = \{v \mid \exists u \in, \check{U}(q_0, u, \$) \vdash_M^* (q_f, \lambda, v)$$

is regular. This yields

$$V_{q_f} = \{x \mid \exists u \in U, u \Rightarrow_p^* x\}.$$

Since the quotient of regular sets is regular [**11**], then

$$\tau(U, P, V) = \{x \mid \exists v \in V, vx\$ \in V_{q_f}\} \cap T^*$$

is regular.

COROLLARY. *Let $R=(I, T, U, V, P)$ be a regular canonical system. Let U and V be regular sets. Then $\beta(U, P, V)$ is regular.*

PROOF. Let $\hat{P}$ be the set of regular productions defined as follows:

$$\hat{P} = \{\beta Q \to \alpha Q \mid \alpha Q \to \beta Q \in P\}$$

and let $\hat{R}$ be the regular canonical system, $R=(I, T, U, V, \hat{P})$. Then clearly $\beta(U, P, V)=\tau(V, \hat{P}, U)$, so that $\beta(U, P, V)$ is regular by Theorem 3.

REMARKS. Obviously we have the same results for reverse regular canonical systems whose productions have the form $Q\alpha \to Q\beta$.

Instead of appealing to the quotient theorem in the proof of Theorem 3, we could have had M delete some member of V from the pushdown store before going to q_f; from this we could have easily obtained another proof of the quotient theorem for regular sets.

REFERENCES

1. Y. Bar-Hillel, M. Perles and E. Shamir, *On formal properties of simple phrase structure grammars*, Z. Phonetik Sprachwiss. Kommunikat. **14** (1961), 143–172.

2. J. R. Büchi, *Regular canonical sysetms*, Arch. Math. Logik Grundlagenforsch **6** (1964), 91–111.

3. N. Chomsky, *On certain formal properties of grammars*, Information and Control **2** (1959), 137–167.

4. ———, *Context-free grammars and pushdown storage*, Quarterly Progress Report No. 65, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., 1962.

5. M. Davis, *Computability and unsolvability*, McGraw-Hill, New York, 1958.

6. J. Evey, *The theory and applications of pushdown store machines*, Ph.D. Thesis, Harvard University, Cambridge, Mass., 1963.

7. S. Ginsburg, *The mathematical theory of context-free languages*, McGraw-Hill, New York, 1966.

8. S. Ginsburg and S. Greibach, *Deterministic context-free languages*, Information and Control **9** (1966), 620–648.

9. S. Ginsburg, S. Greibach and M. Harrison, *Stack automata and compiling*, J. Assoc. Comput. Mach. **14** (1967), 172–201.

10. S. Ginsburg and G. F. Rose, *Operations which preserve definability in languages*, J. Assoc. Comput. Mach. **10** (1963), 175–195.

11. S. Ginsburg and E. Spanier, *Quotients of context-free languages*, J. Assoc. Comput. Mach. **10** (1963), 487–492.

12. S. C. Kleene, *Representation of events in nerve nets and finite automata*, Automata Studies, (C. E. Shannon and J. McCarthy, Eds.), pp. 3–40, Princeton Univ. Press, Princeton, N. J., 1956.

13. G. H. Matthews, *A note on asymmetry in phrase structure grammars*, Information and Control **7** (1964), 360–365.

14. E. L. Post, *Formal Reductions of the general combinatorial decision problem*, Amer. J. Math. **65** (1943), 197–215.

15. ———, *Recursive unsolvability of a problem of Thue*, J. Symbolic Logic **12** (1947), 1–11.

16. M. Rabin and D. Scott, *Finite automata and their decision problems*, IBM J. Res. Develop. **3** (1959), 114–125.

17. A. Thue, *Probleme über Veränderungen von Zeichenreihen nach gegebenen Regeln*, Skr. Vid.-Selskapet Kristiania. I, (1914), no. 10, 1–34.

SYSTEM DEVELOPMENT CORPORATION, SANTA MONICA, CALIFORNIA AND HARVARD UNIVERSITY