INFORMATION AND CONTROL 12, 549 552 (1968)

A Note on Preservation of Languages by Transducers[1]

SEYMOUR GINSBURG AND GENE F. ROSE

System Development Corporation, Santa Monica, California 90406

An error in a previous paper is corrected. The rectified version involves pushdown transducers with accepting states instead of pushdown transducers. Some properties of pushdown transducers with accepting states are then noted.

Associated with any linear bounded automaton (pushdown automaton) is a class of linear bounded transducers, abbreviated lbt (pushdown transducers, abbreviated pdt), in which each move gives rise to an output. In a previous paper (Ginsburg and Rose, 1966) these transducers and their properties are studied. In Theorem 3.2 of Ginsburg and Rose (1966) it is claimed that[2] if $L = T(M)$ for a pda M and S is a pdt associated with M, then $S(L)$ is a context-free language.[3] Dr. Joseph Ullian has pointed out to us that the result is false. In rectifying the theorem we were led to consider transducers with accepting states, i.e., transducers in which the output is considered only when the state of the device is a final or accepting state. The latter part of this note shows that, with minor modifications, most of the results about pdt and lbt in Ginsburg and Rose (1966) hold for pdt with accepting states and lbt with accepting states, respectively.

We begin by showing how incorrect Theorem 3.2 is.[4] The argument is a variant of the counterexample furnished us by Dr. Ullian.

[1] Research sponsored in part by the Air Force Cambridge Research Laboratories, Office of Aerospace Research, USAF, under contract F1962867C0008, and by the Air Force Office of Scientific Research, Office of Aerospace Research, USAF, under AFOSR Grant No. AF-AFOSR-1203-67.

[2] We assume the reader is familiar with Ginsburg and Rose (1966) and we use the terminology and notation found therein.

[3] This statement also appears as Theorem 3.5.1 of Ginsburg (1966).

[4] A result numbered 3.1 or 3.2, etc., refers to that numbered result in Ginsburg and Rose (1966).

THEOREM 1. *For each recursively enumerable set Y, there is a pdt $\bar{S}$ and a pda $\bar{M}$ associated with $\bar{S}$ such that $\bar{S}(T(\bar{M})) = Y$.*[5]

Proof. By Theorem 3.1, there exist a context-free language X and a pdt $S = (K_S, \Sigma, \Gamma_S, \Delta, \mu_S, Z_S, q_S)$ such that $S(X) = Y$. Now $X = T(M)$ for some pda $M = (K_M, \Sigma, \Gamma_M, \delta_M, Z_M, q_M, F)$, with $K_S \cap K_M = \emptyset$. Let q_0 and Z_0 be new symbols, $K = \{q_0\} \cup K_S \cup K_M$, and $\Gamma = \{Z_0\} \cup \Gamma_S \cup \Gamma_M$. Let y_0 be a particular word in Y. (Obviously we may assume $Y \neq \emptyset$.) Let $\bar{M}$ be the pda $(K, \Sigma, \Gamma, \delta, Z_0, q_0, F)$ and $\bar{S}$ the pdt $(K, \Sigma, \Gamma, \Delta, \mu, Z_0, q_0)$, where δ and μ are defined as follows:

(1) $\delta(q_0, \epsilon, Z_0) = \{(q_S, Z_S), (q_M, Z_M)\}$ and

$$\mu(q_0, \epsilon, Z_0) = \{(q_S, Z_S, \epsilon), (q_M, Z_M, y_0)\}.$$

(2) If (p, x, Z) is in $K_S \times (\Sigma \cup \{\epsilon\}) \times \Gamma_S$, then

$$\delta(p, x, Z) = \{(q, \alpha)/(q, \alpha, y) \text{ in } \mu_S(p, x, Z), y \text{ in } \Delta^*\} \quad \text{and}$$

$$\mu(p, x, Z) = \mu_S(p, x, Z).$$

(3) If (p, x, Z) is in $K_M \times (\Sigma \cup \{\epsilon\}) \times \Gamma_M$, then

$$\delta(p, x, Z) = \delta_M(p, x, Z) \quad \text{and}$$

$$\mu(p, x, Z) = \{(q, \alpha, \epsilon)/(q, \alpha) \text{ in } \delta_M(p, x, Z)\}.$$

It is easy to verify that $\bar{S}$ and $\bar{M}$ have the required properties.

Remark. The proof of Theorem 3.2 breaks down on p. 170, lines 27–28, since $S(L)$ need not be a subset of $f[T(M')]$.

While Theorem 3.2 is false as given, it becomes true if the transducers have accepting states, i.e., are devices where the output is considered only along paths leading to final states. In the remainder of this note we examine such transducers and note how the results in Ginsburg and Rose (1966) fare for such devices.

DEFINITION. *A pushdown transducer with accepting states* (abbreviated *a-pdt*) is an 8-tuple $S = (K, \Sigma, \Gamma, \Delta, \mu, Z_0, q_0, F)$, where $K, \Sigma, \Gamma, \Delta, \mu, Z_0, q_0$, have the same significance as in a pdt and F (the set of *final* or *accepting* states) is a subset of K.

The a-pdt operates as follows.

NOTATION. Given an a-pdt $S = (K, \Sigma, \Gamma, \Delta, \mu, Z_0, q_0, F)$, let $\vdash$ and $\vdash^*$ be the relations on $K \times \Sigma^* \times \Gamma^* \times \Delta^*$ as defined for a pdt. For each x in Σ^*, let $S(x) = \{y | (q_0, x, Z_0, \epsilon) \vdash^* (q, \epsilon, \alpha, y)$ for some (q, α) in $F \times \Gamma^*\}$. For $X \subseteq \Sigma^*$, let $S(X) = \bigcup_{x \text{ in } X} S(x)$.

[5] Compare with Theorem 2.2 of Ginsburg and Rose (1966).

For a-pdt, the following counterparts of Theorems 3.1 to 3.6 and Lemma 3.1 hold.

THEOREM 3.1a. *Given any recursively enumerable set Y, there exists a context-free language X and an a-pdt S such that $S(X) = Y$.*

Theorem 3.1a is a corollary of Theorem 3.1. (Let $F = K$.)

DEFINITION. A pda $M = (K, \Sigma, \Gamma, \delta, Z_0, q_0, F)$ and an a-pdt $S = (K, \Sigma, \Gamma, \Delta, \mu, Z_0, q_0, F)$ are said to be *associated* if, for each (q, x, Z) in $K \times (\Sigma \cup \{\epsilon\}) \times \Gamma$,

$$\delta(q, x, Z) = \{(p, \alpha)/(p, \alpha, y) \quad \text{in} \quad \mu(q, x, Z) \quad \text{for some} \quad y \quad \text{in} \quad \Delta^*\}.$$

Note that an a-pdt has a unique associated pda, but each pda has an infinite number of associated a-pdt.

The amended form of Theorem 3.2 is now given.

THEOREM 3.2a. *If a pda M and an a-pdt S are associated, then $S(T(M))$ is a context-free language.*

THEOREM 3.3a. *$S(L)$ is context free for each a-pdt S and each regular set L.*

The proof is essentially the same as for Theorem 3.3. Given $S = (K, \Sigma, \Gamma, \Delta, \mu, Z_0, q_0, F)$ let $S' = (K \times K_A, \Sigma, \Gamma, \Delta, \mu', Z_0, (q_0, p_0), F \times F_A)$, with μ' defined as for Theorem 3.3. Let M' and M be the pda associated with S' and S, respectively.

THEOREM 3.4a. *Let $S = (K, \Sigma, \Gamma, \Delta, \mu, Z_0, q_0, F)$ be an ϵ-output free a-pdt. Then $S(L)$ is context sensitive for each context-sensitive language L.*

The proof is essentially the same as for Theorem 3.4. For productions of type (7), q is required to be in F.

LEMMA 3.1a. *Given a context-free language $Y \subseteq \Sigma^*$ and any word w, there exists an a-pdt S such that $S(w) = Y$.*

The proof is essentially the same as for Lemma 3.1. For $w \neq \epsilon$, a suitable a-pdt is $S = (K, \Sigma_1, V \cup \{Z_0\}, \Sigma, \mu, Z_0, q_0, \{q_{k+1}\})$, where K, Σ_1, V, Z_0, and μ are as in Lemma 3.1. For $w = \epsilon$, let $S = (\{q_0, q_1, q_2\}, \Sigma_1, V \cup \{Z_0\}, \Sigma, \mu, Z_0, q_0, \{q_2\})$, where $\mu(q_0, \epsilon, Z_0) = (q_1, Z_0\sigma, \epsilon)\}$, $\mu(q_1, \epsilon, x) = \{(q_1, \epsilon, x)\}$ for x in Σ, $\mu(q_1, \epsilon, \xi) = \{(q_1, \beta^R, \epsilon)/\xi \to \beta \text{ in } P\}$ for ξ in $V - \Sigma$, and $\mu(q_1, \epsilon, Z_0) = \{(q_2, \epsilon, \epsilon)\}$.

Thus we get

THEOREM 3.5a. *The following questions are recursively solvable for arbitrary context-free languages X and Y:*

(a) *Is there an a-pdt S such that $S(X) = Y$?*

(b) *Is there is a-pdt S such that $S(X) \subseteq Y$ and $S(X)$ is infinite if X is infinite?*

Note that it is no longer necessary, as in Theorem 3.5, to assume that X not contain ϵ. So the analog to Theorem 3.6, rather than declaring the unsolvability of mapping problems for context-free languages, is

THEOREM 3.6a. *Each of the following questions is recursively unsolvable for arbitrary context-sensitive languages X and Y:*

(a) *Is there an a-pdt S such that $S(X) = Y$?*

(b) *Is there an a-pdt S such that $S(X) \subseteq Y$ and $S(X)$ is infinite if X is infinite?*

The proof depends on Theorem 3.3a and Lemma 3.1a. The argument is essentially that in the last paragraph of Ginsburg and Rose (1966).

Remark. By Theorem 3.6, for pdt these questions are unsolvable for context-free X and Y.

Turning to the lbt, we find that, unlike the results about pdt, the results about lbt are essentially unaffected by the specification of accepting states. In particular, Theorems 2.2 to 2.5 of Ginsburg and Rose (1966) remain valid if "lbt" is replaced by "a-lbt," the given proofs carrying over with trivial modifications to the a-lbt case.

RECEIVED: June 6, 1967; revised May 15, 1968

REFERENCES

GINSBURG, S. (1966), "The Mathematical Theory of Context-free Languages." McGraw-Hill, New York.

GINSBURG, S. AND ROSE, G. F. (1966), Preservation of languages by transducers. *Inform. Control*, **9**, 153–176.