AD 633194

TM-738/022/00

# PUSHDOWN AUTOMATA WITH BOUNDED BACKTRACK

## SCIENTIFIC REPORT NO. 7

Jeffrey Ullman

3 December 1965



Prepared for

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES

OFFICE OF AEROSPACE RESEARCH

UNITED STATES AIR FORCE

Supported by

SDC Independent Research Program

and Contract AF 19(628)-5166

CRL - Algorithmic Languages Program

Project No. 5632

AFCRL 66-550

SDC Document No.

TM-738/022/00

Technical Memorandum

SDC
TM SERIES

The research reported herein was supported by SDC's Independent Research Program and Contract AF 19(628)-5166, CRL - Algorithmic Languages Program, Project No. 5632, Task No. 563205.

SCIENTIFIC REPORT NO. 7

# PUSHDOWN AUTOMATA WITH BOUNDED BACKTRACK

by
Jeffrey Ullman

3 December 1965

Prepared for

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS

SYSTEM
DEVELOPMENT
CORPORATION
2500 COLORADO AVE.
SANTA MONICA
CALIFORNIA
90406

ABSTRACT

This report considers two classes of pushdown automata (pda), and the languages accepted by them. These pda accept their languages rapidly because they reread the input word a limited number of times. Hence, such languages are particularly useful as programming languages.

The first class, strong bounded backtrack pda, read input words from left to right, and jump from right to left (backtrack) at most k times for some integer k. The languages accepted by such automata will be shown to be equivalent to the finite unions of deterministic languages.

The second class, weak bounded backtrack pda, read each letter of the input word at most k times, although they may backtrack an arbitrary number of times. An alternative model envisioned for this device is one that has the storage space for k states and k pushdown tapes, but no more. The device reads a word from left to right, simulating the action of the pda. Every time the pda reaches a total configuration (state and pushdown tape) in which it is possible to read another input letter, that configuration is stored. If no move at all is possible in a given configuration, it is erased from storage. Thus one can accept the language with no backtrack without having to keep track of an arbitrary number of possible configurations of the pda.

Several results will be shown about each of these classes of pda, including operations that preserve the properties. While these properties are not preserved by all gsm mappings, it will be shown that information lossless gsm's preserve the weak bounded backtrack property, and information lossless gsm's of finite order preserve the strong bounded backtrack property.

# PUSHDOWN AUTOMATA WITH BOUNDED BACKTRACK*

## INTRODUCTION

The purpose of this report is to consider restrictions on pushdown automata causing them to accept or reject words in a simpler manner than the most general pushdown automaton. We will make two such related definitions and study their closure properties under various operations, such as intersection with regular sets. One restriction will be shown to be equivalent to the restriction that a language be the finite union of deterministic languages, and some results about deterministic languages will be shown.

Definition. A pushdown automaton (pda) [2,4], M, is a 7-tuple $(K, \Sigma, \Gamma, \delta, Z_0, q_0, F)$, where K is the finite nonempty set of states, $\Sigma$ is the finite nonempty set of pushdown tape symbols, $\delta$ is a mapping from $K \times (\Sigma \cup \{\epsilon\}) \times \Gamma$[(1)] to the finite subsets of $K \times \Gamma^*$, $Z_0$ in $\Gamma$ is the initial tape symbol, $q_0$ in K is the initial state, and $F \subseteq K$ is the set of acceptable final states.

Definition. For a pda, M, let $\vdash_{\overline{M}}$ (or $\vdash$ when M is understood) be the relation on $K \times \Sigma^* \times \Gamma^*$ such that $(q_1, w_1, \gamma_1) \vdash_{\overline{M}} (q_2, w_2, \gamma_2)$ if and only if there is some a in $\Sigma \cup \{\epsilon\}$, Z in $\Gamma$, $\gamma_1'$ and $\gamma$ in $\Gamma^*$ such that $w_1 = aw_2$, $\gamma_1 = \gamma_1' Z$, $\gamma_2 = \gamma_1' \gamma$, and $(q_2, \gamma)$ is in $\delta(q_1, a, Z)$.

---

*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(628)-5166, CRL - Algorithmic Languages Program.

(1) $\epsilon$ is the string of zero length. Also, for a set X, we will use $X^*$ as the closure of X under concatenation, $X^* = \bigcup_{i \geq 0} X^i$, and $\emptyset$ to stand for the null set.

Let $\vert\!\frac{*}{M}$ (or $\vert\!\frac{*}{}$ when M is understood) be the relation on $K \times \Sigma^* \times \Gamma^*$ such that for x and y in $K \times \Sigma^* \times \Gamma^*$, $x \vdash^{*} y$ if and only if for some z, $x \vdash^{*} Z$ and $Z \vdash y$, or $x = y$. Let $\vert\!\frac{d}{M}$ or $\vdash^{d}$ be the relation on $K \times \Sigma^* \times \Gamma^*$ such that $x \vdash^{d} y$ if and only if $x \vdash y$, $y = (q, w, \gamma Z)$, Z in $\Gamma$, and there is some a in $\Sigma$ such that $\delta(q,a,Z) \neq \emptyset$. Finally, let $\vert\!\frac{d*}{M}$ or $\vdash^{d*}$ be defined as $\vdash^{d}$ but with $\vdash^{*}$ in place of $\vdash$.(2)

Intuitively, we suppose that if $(q,\gamma)$ is in $\delta(p,a,Z)$, then the pda has the option, if Z is the symbol at the top of the pushdown tape, and a is the first input symbol, or $\epsilon$, of moving from state q to state p, replacing Z by the (possibly empty) string $\gamma$, and erasing a from the input tape. The set of words, w, such that for some p in F, and $\gamma$ in $\Gamma^*$, $(q_o, w, Z_o) \vert\!\frac{*}{M} (p, \epsilon, \gamma)$ is the set of words "accepted" by M, often denoted T(M). The sets of words accepted by some pda are exactly those sets that are context free languages [1].(3)

We shall define two different, related restrictions on pda implying that, in some sense, the languages accepted by pda, meeting the restriction, are easily compiled. The first, called "strong bounded backtrack," implies that

---

(2)Note that if M is deterministic, and $x \vdash^{d*} y$, then $x \vdash^{d*} y$ in the sense of [6,7], but the converse is not necessarily true.

(3)A context free language is a set generated by a grammar $G = (V, \Sigma, P, \sigma)$, where V is a finite vocabulary, $\Sigma \subseteq V$ is the set of terminal symbols, P a set of ordered pairs $(\tau, w)$ where $\tau$ is in $V-\Sigma$ and w in $V^*$, and $\sigma$ in $V-\Sigma$. The strings generated by G are defined: $\sigma$ is a string, and if $u\tau v$ is a string, u and v in $V^*$, and $\tau$ in $V-\Sigma$, and $(\tau, w)$ is in P, then uwv is a string generated by G. The language generated by G is the set of strings generated G intersected with $\Sigma^*$.

for some integer, k, we may determine whether or not a word, w, is accepted by a process that scans w from left to right and "backtracks" (jumps from right to left and restores the state and pushdown tape to the condition that prevailed when the new point of scan was last reached) at most k times.

The second restriction, called "weak bounded backtrack", implies that for some integer, k, and any word, w, we may determine whether or not w is accepted, by a process that scans each symbol of w at most k times. Alternatively, if we have space in memory for k copies of the pda, we may scan w only once, assigning various actions of the pda to the several copies. A more formal discussion follows.

<u>Notation</u>. For a set X, we will let <u>#X</u> represent the cardinality of X.

For x and y, words over some common alphabet, the relation $x \leq y$ shall hold if and only if there is some word, v, such that $xv = y$. Note that v may be $\varepsilon$.

For a given pda, $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$, x and y strings in $(K\Sigma^*\Gamma^*)^*$ let the relation xJy hold if and only if, for some w in $\Sigma^*$, $x = q_o w_o \gamma_o q_1 w_1 \gamma_1 \cdots q_n w_n \gamma_n$ and $y = q_o w_o w \gamma_o q_1 w_1 w \gamma_1 \cdots q_n w_n w \gamma_n q_{n+1} w_{n+1} \gamma_{n+1} \cdots q_m w_m \gamma_m$, where $0 \leq n \leq m$, and for $0 \leq i \leq m$, $q_i$ is in K, $w_i$ is in $\Sigma^*$ and $\gamma_i$ is in $\Gamma^*$.

We will here, and throughout the paper, assume that the set of states, set of tape symbols, and set of input symbols for any pda, are mutually disjoint.

<u>Definition</u>. For pda, $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$, let $P_M(w)$ be the set $\{q_o w_o \gamma_o q_1 w_1 \gamma_1 \cdots q_n w_n \gamma_n /$ for $0 \leq i \leq n$, $q_i$ is in K, $w_i$ in $\Sigma^*$, and $\gamma_i$ in $\Gamma^*$, $w_o = w, \gamma_o = Z_o$, $w_n = \varepsilon$, for $0 \leq i < n$, $(q_i, w_i, \gamma_i) \vdash (q_{i+1}, w_{i+1}, \gamma_{i+1})$ and for no q and $\gamma$ is $(q_n, \varepsilon, \gamma_n) \vdash (q, \varepsilon, \gamma)$ true$\}$.

Let $Q_M(x,y)$ be the binary function from $\Sigma^* \times \Sigma^*$ to the subsets of $(K\Sigma^*\Gamma^*)^*$ defined as follows:

$$Q_M(x,x) = P_M(x)$$

If it is false that $x \leq y$, $Q_M(x,y) = \emptyset$.

For $a$ in $\Sigma$ and $w$ in $\Sigma^*$, let $Q_M(x,xaw) = \{y/y$ is in $P_M(x)$ and for all $z$ such that $yJz$, $z$ is not in $P_M(xa)\}$.

Definition. A pda M is said to be a strong bounded backtrack (sbb) pda with bound k if and only if M has no infinite loops with input $\epsilon$[4], and for some integer k, and any x in $\Sigma^*$, there are no more than k elements, y, in $(K\Sigma^*\Gamma^*)^*$ such that for some $z \leq x$, y is in $Q_M(z,x)$.

Intuitively, $P_M(x)$ represents the set of responses of M to word x, for which the entire word x is input to M, and M cannot operate further without a non-$\epsilon$ input. $Q_M(y,x)$ represents the responses to x for which, after reading the proper initial subword y, of x, M can make no move that will enable it to read the next symbol of x. Thus the sbb property bounds the number of distinct paths a pda may traverse while reading any word, a path, of course, being any succession of triple s consisting of state, remaining input word, and pushdown tape contents.

Definition. For the pda $M = (K,\Sigma,\Gamma,\delta,Z_o,q_o,F)$, let $R_M(x) = \{(q,\gamma)/(q_o,x,Z_o) \overset{d*}{\underset{M}{\vdash}} (q,\epsilon,\gamma)\}$.

---

[4] M has an infinite loop with input $\epsilon$ if for some w in $\Sigma^*$ and $q_1,\gamma_1,q_2,\gamma_2,\ldots,q_i,\gamma_i,\ldots$ there exists an infinite sequence, $(q_ow,Z_o) \overset{*}{\vdash} (q_1,\epsilon,\gamma_1) \vdash (q_2,\epsilon,\gamma_2) \vdash \ldots \vdash (q_i,\epsilon,\gamma_i) \vdash (q_{i+1},\epsilon,\gamma_{i+1}) \vdash \ldots$

A pda will be said to be of <u>weak bounded backtrack</u> (wbb) with bound R if for some integer k and any x in $\Sigma^*$, $\#R_M(x) \leq k$.

A language will be said to be of sbb (wbb) if it is accepted by some sbb (wbb) pushdown automaton.

<u>Definition</u>. The pda M as above will be said to be <u>deterministic</u> [6,8,10] if for all q, a and Z in $K, \Sigma \cup \{\epsilon\}$ and $\Gamma$ respectively, $\#\delta(q,a,Z) \leq 1$, and if $\delta(q,\epsilon,Z) \neq \emptyset$, then for all a in $\Sigma$, $\delta(q,a,Z) = \emptyset$.

Intuitively, a deterministic pda is one for which, for any input word, at most one response is possible.

## <u>Section 1. Strong Bounded Backtrack and Deterministic Languages</u>

The primary result of this section is that every sbb language is the union of a finite number of deterministic languages. For this result we need several lemmas.

<u>Lemma 1.1</u>. A deterministic language is an sbb language.

<u>Proof</u>. We will show that every loop-free deterministic pda satisfies the definition of strong bounded backtrack, with a bound of k = 1. It has been shown [6] that every language accepted by a deterministic pda is accepted by a loop-free deterministic pda. Let $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$ be such a pda. Suppose that there exist y and y' such that either y and y' are in $Q_M(v,x)$ for some v and x such that $v \leq x$, or y is in $Q_M(v,x)$ and y' is in $Q_M(v',x)$ for distinct v and v' such that $v \leq v' \leq x$.

In the first case, suppose that $y = q_o v Z_o q_1 v_1 \gamma_1 \cdots q_n v_n \gamma_n$ and $y' = q_o v Z_o q_1' v_1' \gamma_1' \cdots q_m' v_m' \gamma_m'$. Without loss of generality, we may assume that $n \leq m$.

Since M is deterministic, for $1 \leq i \leq n$, $q_i = q_i'$, $v_i = v_i'$ and $\gamma_i = \gamma_i'$. If $m = n$, then $y = y'$, so we may assume $n < m$. But then $(q_o, v, Z_o) \overset{*}{\vdash} (q_n, \epsilon, \gamma_n) \overset{*}{\vdash} (q_m, \epsilon, \gamma_m)$, contradicting the assumption that y was in $P_M(v)$.

In the second case, suppose $y = q_o v Z_o q_1 v_1 \gamma_1 \cdots q_n v_n \gamma_n$ and $y' = q_o v' Z_o q_1' v_1' \gamma_1' \cdots q_m' v_m' \gamma_m'$. Then since M is deterministic, and v is a proper initial subword of v', for $1 \leq i \leq n$, it must be the case that $q_i = q_i'$, $\gamma_i = \gamma_i'$ and $v_i' = v_i v_n'$. But then there must be some a in $\Sigma$ such that $va \leq v'$ and a largest s such that $y'' = q_o vaZ_o q_1 v_1 a \gamma_1 \cdots q_n v_n a \gamma_n \cdots q_s' \epsilon \gamma_s'$ and $y''Jy'$. Since M is loop free, either $y''$ is in $P_M(va)$, or there will be some finite sequence $(q_s', \epsilon, \gamma_s') \vdash (q_1'', \epsilon, \gamma_1'') \vdash (q_2'', \epsilon, \gamma_2'') \vdash (q_t'', \epsilon, \gamma_t'')$, such that $y''' = y'' q_1'' \gamma_1'' q_2'' \gamma_2'' \cdots q_t'' \gamma_t''$ is in $P_M(va)$ and in either case, $P_M(va)$ is not empty, so y could not be in $Q_M(v,x)$. We thus conclude that M is a sbb machine with a bound of one.

<u>Lemma 1.2</u>. The union of sbb languages is a sbb language. Also, the union of wbb languages is a wbb language.

<u>Proof</u>. Let $M_1 = (K_1, \Sigma, \Gamma, \delta_1, Z_o, q_o, F_1)$ and $M_2 = (K_2, \Sigma, \Gamma, \delta_2, Z_o, q_o', F_2)$ and assume that $K_1 \cap K_2 = \emptyset$. Let $q_o''$ not be in $K_1 \cup K_2$, $K = K_1 \cup K_2 \cup \{q_o''\}$ and $F = F_1 \cup F_2$. Define $\delta$ as follows:

$$\delta(q,a,Z) = \delta_1(q,a,Z) \text{ for } q \text{ in } K_1,\ a \text{ in } \Sigma \cup \{\epsilon\} \text{ and } Z \text{ in } \Gamma$$

$$\delta(q,a,Z) = \delta_2(q,a,Z) \text{ for } q \text{ in } K_2,\ a \text{ in } \Sigma \cup \{\epsilon\} \text{ and } Z \text{ in } \Gamma$$

$$\delta(q_o'', \epsilon, Z_o) = \{(q_o, Z_o), (q_o', Z_o)\}$$

$$\delta(q_o'', a, Z_o) = \emptyset \text{ for } a \text{ in } \Sigma$$

Consider the pda $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o'', F)$. Surely, $T(M) = T(M_1) \cup T(M_2)$. Also, it is easy to see for any x in $\Sigma^*$, that y is in $R_{M_1}(x) \cup R_{M_2}(x)$ if and only if $q_o'' x Z_o y$ is in $R_M(x)$. Hence, $\# R_M(x) \leq \# R_{M_1}(x) + \# R_{M_2}(x)$, so that if $M_1$ and $M_2$ are wbb machines, M is a wbb machine.

Likewise, y is in $P_{M_1}(x) \cup P_{M_2}(x)$ if and only if $q_o''xZ_oy$ is in $P_M(x)$. Suppose y is in $Q_M(v,x)$ for some v, a proper initial subword of x. Then $y = q_o''vZ_oy'$ for some y' in $P_{M_1}(v) \cup P_{M_2}(v)$. Say y' is in $P_{M_1}(v)$ without loss of generality. Suppose there is an a in $\Sigma$ and z in $P_{M_1}(va)$ such that y'Jz. Then $yJq_o''vaZ_oz$, which contradicts the assumption that y is in $Q_M(v,x)$. Thus $\#Q_M(v,x) \leq \#Q_{M_1}(v,x) + \#Q_{M_2}(v,x)$ for all $v \leq x$, and if $M_1$ and $M_2$ are sbb machines, M will be likewise.

<u>Definition</u>. Given a pda $M = (K,\Sigma,\Gamma,\delta,Z_o,q_o,F)$, a <u>branch point</u> is an element $(q,a,Z)$ of $K \times (\Sigma \cup \{\epsilon\}) \times \Gamma$ such that if a is in $\Sigma$, $\#\delta(q,a,Z) + \#\delta(q,\epsilon,Z) \geq 2$ and $\delta(q,a,Z) \neq \emptyset$. If $a = \epsilon$, then either $\#\delta(q,\epsilon,Z) \geq 2$ or $\delta(q,\epsilon,Z) \neq \emptyset$ and for some b in $\Sigma$, $\delta(q,b,Z) \neq \emptyset$.

<u>Lemma 1.3.</u> If $M = (K,\Sigma,\Gamma,\delta,Z_o,q_o,F)$ is sbb with bound k, $(q_o,x_o,\gamma_o) \vdash (q_1,x_1,\gamma_1) \underset{M}{\vdash} \cdots \underset{M}{\vdash} (q_n,x_n,\gamma_n)$, where $\gamma_o = Z_o$ and $x_n = \epsilon$, and for $0 \leq i < n$, $a_i$ in $\Sigma \cup \{\epsilon\}$ is defined by $a_ix_{i+1} = x_i$ and $Z_i$ in $\Gamma$ is defined by $\gamma_i = \gamma_i'Z_i$ for some $\gamma_i'$ in $\Gamma^*$, then $\#\{(q_i,a_i,Z_i)/(q_i,a_i,Z_i)$ is a branch point and $i < n\} \leq k-1$.

<u>Proof</u>. Suppose $(q_{b(1)},a_{b(1)},Z_{b(1)}),(q_{b(2)},a_{b(2)},Z_{b(2)}),\ldots,(q_{b(k)},a_{b(k)},Z_{b(k)})$ are k branch points, not necessarily distinct, but with $b(i) < b(j)$ for $i < j$, and $b(k) < n$. Then for $1 \leq i \leq k$, $(q_{b(i)},x_{b(i)},\gamma_{b(i)}) \vdash (q_i',x_i',\gamma_i')$, where either $q_i' \neq q_{b(i)+1}$ or $x_i' \neq x_{b(i)+1}$ or $\gamma_i' \neq \gamma_{b(i)+1}$. Then, since M has no infinite loops, either for some $q_i''$ and $\gamma_i''$, $(q_o,x_o,\gamma_o) \overset{*}{\vdash} (q_{b(i)},x_{b(i)},\gamma_{b(i)}) \vdash (q_i',x_i',\gamma_i') \overset{*}{\vdash} (q_i'',\epsilon,\gamma_i'')$ and for no q in K and $\gamma$ in $\Gamma^*$ does $(q_i'',\epsilon,\gamma_i'') \vdash (q,\epsilon,\gamma)$, or for some $q_i''$, $x_i''$ and $\gamma_i''$, $(q_o,x_o,\gamma_o) \overset{*}{\vdash} (q_{b(i)},x_{b(i)},\gamma_{b(i)}) \vdash (q_i',x_i',\gamma_i') \overset{*}{\vdash} (q_i'',x_i'',\gamma_i'')$ and if a is in $\Sigma \cup \{\epsilon\}$, and $x_i'' = ax_i'''$, for $x_i'''$ in $\Sigma^*$, then for no q and $\gamma$ does $(q_i,x_i,\gamma_i) \vdash (q,x_i''',\gamma)$.

In the first case, there is a $y_i$ in $Q_M(x_o,x_o)$, and in the second, a $y_i$ in $Q_M(x,x_o)$, where $xx_i'' = x_o$. Also, there must be a y in $Q_M(x_o,x_o)$ such that

$q_o x_o \gamma_o q_1 x_1 \gamma_1 \cdots q_n x_n \gamma_n Jy$. It should be easy to see that $y \neq y_i$ for any i, and that for distinct i and j, $y_i \neq y_j$. Thus, there would be at least k + 1 elements in $\bigcup_{u \leq x_o} Q_M(u, x_o)$, which violates the assumption that M is sbb with bound k.

Lemma 1.4. An sbb language is the finite union of deterministic languages.

Proof. Let $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$ be an sbb pushdown automaton with bound k on backtrack. Let M have branch points $B_1, B_2, \ldots, B_e$, certainly a finite number. If $B_i = (q_i, a_i, Z_i)$, $q_i$ in K, $a_i$ in $\Sigma \cup \{\epsilon\}$, and $Z_i$ in $\Gamma$, has $f_i$ members in $\delta(q_i, a_i, Z_i)$, let these be known as $C_{i1}, C_{i2}, \ldots, C_{if_i}$. Consider the set S = $\{(C_{i_1 j_1}, C_{i_2 j_2}, \ldots, C_{i_{k-1} j_{k-1}}) / 1 \leq i_n \leq e \text{ and } 1 \leq j_n \leq f_{i_n} \text{ for } 1 \leq n \leq k-1\}$. Also, define $K' = \{q^{(i)} / q \text{ in } K \text{ and } 1 \leq i \leq k\}$ and $F' = \{q^{(i)} / q \text{ in } F \text{ and } 1 \leq i \leq k\}$. Then for each s in S, define $M_s = (K', \Sigma, \Gamma, \delta_s, Z_o, q_o^{(1)}, F')$, where $\delta_s$ is defined as follows:

If (q,a,Z) in $K \times (\Sigma \cup \{\epsilon\}) \times \Gamma$ is not a branch point, $\delta_s(q^{(i)}, a, Z) = \{(p^{(i)}, \gamma)\}$ if $\delta(q, a, Z) = \{(p, \gamma)\}$ and $\delta_s(q^{(i)}, a, Z) = \emptyset$ if $\delta(q, a, Z) = \emptyset$, for all i, $1 \leq i \leq k$.

If for some m, $1 \leq m \leq k-1$, the $m^{th}$ coordinate of s is $C_{i_m j_m} = (p, \gamma)$, and (q,a,Z) is the $i_m^{th}$ branch point, let $\delta_s(q^{(m)}, a, Z) = \{(p^{(m+1)}, \gamma)\}$. Otherwise, let $\delta_s(q^{(m)}, a, Z) = \emptyset$.

It is easy to see that each $M_s$ so defined is deterministic. If $(q_o^{(1)}, x, Z_o) \vdash_{M_s} (q_1^{(i_1)}, x_1, \gamma_1) \vdash_{M_s} \cdots \vdash_{M_s} (q_n^{(i_n)}, \epsilon, \gamma_n)$, then certainly, $(q_o, x, Z_o) \vdash_M (q_1, x_1, \gamma_1) \vdash_M \cdots \vdash_M (q_n, \epsilon, \gamma_n)$, so that $T(M_s) \subseteq T(M)$. Hence, $\bigcup_{s \text{ in } S} T(M_s) \subseteq T(M)$.

Now we must show that $T(M) \subseteq \bigcup_{s \text{ in } S} T(M_s)$. Suppose that x is in T(M). Then $(q_o, a_o x_o, \gamma_o Z_o) \vdash_M (q_1, a_1 x_1, \gamma_1 Z_1) \vdash_M \cdots \vdash_M (q_n, a_n x_n, \gamma_n Z_n)$ where $a_o x_o = x$,

$a_{i+1}x_{i+1} = x_i$ for $0 \le i < n$ and $\gamma_o = a_n = x_n = \epsilon$. From Lemma 1.3, at most k-1 of the triplets $(q_i, a_i, Z_i)$, $0 \le i < n$ are branch points. Let these be $(q_{b(1)}, a_{b(1)}, Z_{b(1)})$, $(q_{b(2)}, a_{b(2)}, Z_{b(2)}), \ldots, (q_{b(k')}, a_{b(k')}, Z_{b(k')})$ for some $k' \le k-1$. Define $\gamma_i'$ by $\gamma_{b(i)+1}Z_{b(i)+1} = \gamma_{b(i)}\gamma_i'$, $1 \le j_i \le k'$. Then $(q_{b(i)}, a_{b(i)}, Z_{b(i)})$ is $B_{g_i}$ and $(q_{b(i)+1}, \gamma_i')$ is some $C_{g_i j_i}$, for $1 \le i \le k'$, $1 \le g_i \le e$ and $1 \le j_i \le f_{g_i}$. Let $s = (C_{g_1 j_1}, C_{g_2 j_2}, \ldots, C_{g_{k'} j_{k'}}, C_{11}, C_{11}, \ldots, C_{11})$ have exactly k-1 components, the right-hand $C_{11}$'s being used only if $k' < k-1$. Then $M_s$ accepts x. For $(q_o^{(1)}, a_o x_o, \gamma_o Z_o) \overset{*}{\vdash}_{M_s} (q_{b(1)}^{(1)}, a_{b(1)}x_{b(1)}, \gamma_{b(1)}Z_{b(1)}) \vdash_{M_s}$ $(q_{b(1)+1}^{(2)}, a_{b(1)+1}x_{b(1)+1}, \gamma_{b(1)+1}Z_{b(1)+1}) \overset{*}{\vdash}_{M_s} \cdots \overset{*}{\vdash}_{M_s} (q_{b(k')}^{(k')}, a_{b(k')}x_{b(k')}, \gamma_{b(k')}Z_{b(k')})$ $\vdash_{M_s} (q_{b(k')+1}^{(k'+1)}, a_{b(k')+1}x_{b(k')+1}, \gamma_{b(k')+1}Z_{b(k')+1}) \overset{*}{\vdash}_{M_s} (q_n^{(k'+1)}, a_n x_n, \gamma_n Z_n)$, where $a_o x_o = x$, $a_{i+1}x_{i+1} = x_i$, for $0 \le i < n$, $\gamma_o = a_n = x_n = \epsilon$, and $q_n^{(k'+1)}$ is in F'.

Theorem 1.1. A language is sbb if and only if it is the finite union of deterministic languages.

Proof. Immediate from Lemmas 1.1, 1.2, and 1.4.

Definition. An operation, f, that maps sets to sets is said to be additive if for any set S, $f(S) = \bigcup_{x \text{ in } S} f(\{x\})$.[5]

Lemma 1.5. If f is an additive operation that preserves deterministic languages, f preserves sbb languages.

Proof. If L is an sbb language, by Theorem 1.1, for some integer k and deterministic languages $L_1, L_2, \ldots, L_k$, we have $L = \bigcup_{i=1}^{k} L_i$. Then $f(L) = f(\bigcup_{i=1}^{k} L_i) =$ $f(\bigcup_{i=1}^{k} \bigcup_{w \text{ in } L_i} w) = \bigcup_{i=1}^{k} [\bigcup_{w \text{ in } L_i} f(w)] = \bigcup_{i=1}^{k} f(L_i)$. Since $f(L_i)$ is deterministic,

---

[5] From here, we will use f(x) for f({x}). No confusion will arise.

$f(L)$ is the finite union of deterministic languages.

Theorem 1.2. The following operations preserve sbb languages ($L$ is an sbb language, and $R$ a regular set[6]): a) $L \cap R$ b) $L-R$ c) $L/R = \{u/$ for some $v$ in $R$, $uv$ is in $L\}$ d) $LR = \{uv/u$ in $L$ and $v$ in $R\}$ e) $G^{-1}(L)$ for any generalized sequential machine mapping, $G$[7] f) $\mathrm{Init}(L) = \{u/$for some string $v$, $uv$ is in $L\}$ g) $f_c(L) = \{u/ucv$ is in $L$ for some string $v$ and $u$ contains no occurrence of $c$ $\}$.

Proof. Each of the above operations is easily seen to be additive, and each was shown to preserve deterministic languages in [6].

## Section 2. Operations on Weak Bounded Backtrack Languages

Having disposed of several important questions regarding operations that preserve the sbb property, we will now proceed to consider the same questions for wbb languages. The properties of sbb and wbb languages turn out to be quite similar, and in fact, all the operations in Theorem 1.2 can be shown to preserve wbb languages. To begin, we prove the following simple but important theorem:

---

(6) The regular sets form the smallest class of sets containing the finite sets and closed under union, concatenation and closure (*) [9]. Alternatively, a regular set is a set accepted by some finite automaton [11].

(7) A generalized sequential machine [5] (gsm) is a six-tuple $(K, \Sigma, \Delta, \delta, \lambda, q_o)$ where $K$ is the finite nonempty set of states, $\Sigma$ the finite nonempty set of input symbols, and $\Delta$ the finite nonempty set of output symbols. $\delta$ is a mapping from $K \times \Sigma$ to $K$, and $\lambda$, a mapping from $K \times \Sigma$ to $\Delta^*$. $q_o$ is the initial state, a member of $K$. We may extend $\delta$ and $\lambda$ to $K \times \Sigma^*$ by $\delta(q, \epsilon) = q, \lambda(q, \epsilon) = \epsilon$ and for $w$ in $\Sigma^*$ and $a$ in $\Sigma$, $\delta(q,wa) = \delta(\delta(q,w),a)$ and $\lambda(q,wa) = \lambda(q,w)\ \lambda(\delta(q,w),a)$. The corresponding gsm mapping takes any word $w$ in $\Sigma^*$ to $\lambda(q_o,w)$. The inverse gsm mapping is, of course, the mapping which takes $u$ in $\Delta^*$ to the set of $w$ in $\Sigma^*$ for which $u = \lambda(q_o,w)$.

<u>Theorem 2.1</u>. The intersection or difference of a wbb language and a regular set is a wbb language.

<u>Proof</u>. Let $L = T(M)$ for wbb pda $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$, and $R = T(A)$ for finite automaton $A = (K_A, \Sigma, \delta_A, p_o, F_A)$.(8) Consider the pda $N = (K \times K_A, \Sigma, \Gamma, \delta', Z_o, (q_o, p_o), F')$, where $\delta'((q,p),a,Z) = \{((q',p'),\gamma)/\delta_A(p,a) = p'$ and $\delta(q,a,Z)$ contains $(q',\gamma)\}$, for all $a$ in $\Sigma$, $q$ and $q'$ in $K$, $p$ and $p'$ in $K_A$, and $Z$ in $\Gamma\}$. Also, $\delta'((q,p),\epsilon,Z) = \{((q',p),\gamma)/\delta(q,\epsilon,Z)$ contains $(q',\gamma)$, for $q$ and $q'$ in $K$, $p$ in $K_A$ and $Z$ in $\Gamma\}$.

It is clear that $((q_o,p_o),w,Z_o) \overset{*}{\underset{N}{\vdash}} ((q,p),\epsilon,\gamma)$ if and only if $\delta_A(p_o,w) = p$ and $(q_o,w,Z_o) \overset{*}{\underset{M}{\vdash}} (q,\epsilon,\gamma)$. Hence, if we define $F' = \{(q,p)/q$ in $F$ and $p$ in $F_A\}$, $T(N) = L \cap R$. If $F' = \{(q,p)/q$ in $F$ and $p$ not in $F_A\}$, $T(N) = L-R$. Also, for $w$ in $\Sigma^*$, and $((q,p),\gamma)$ in $R_N(w)$, $p = \delta_A(p_o,w)$ and $(q,\gamma)$ is in $R_M(w)$. Hence $\#R_N(w) \leq \#R_M(w)$, so $N$ is a wbb pda.

<u>Notation</u>. For a given set, $X$, let $X^{(n)} = 0 \leq = \bigcup_{i \leq n} X^i$.

<u>Theorem 2.2</u>. If $G$ is a gsm mapping and $L$ a wbb language, then $G^{-1}(L)$ is a wbb language.

---

(8) A finite automaton [11] is a five-tuple consisting of a finite nonempty set of states, $K_A$, a finite nonempty set of input symbols, $\Sigma$, a mapping from $K_A \times \Sigma$ to $K_A$, $\delta_A$, a start state in $K_A$, $p_o$, and a set of acceptable final states $F_A \subseteq K_A$. $\delta_A$ is extended to domain $K_A \times \Sigma^*$, as for the mapping $\delta$ for the gsm in footnote (7). The set of words accepted by the finite automaton, $T(A)$ is $\{w/\delta_A(p_o,w)$ is in $F_A\}$. As noted, each such set is regular, and every regular set is $T(A)$ for some finite automaton, $A$.

Proof. Let $M = (K_M, \Delta, \Gamma_M, \delta_M, Z_o, q_o, F_M)$ be a wbb pda accepting L, and $S = (K_S, \Sigma, \Delta, \delta_S, \lambda, p_o)$ be a gsm with $\lambda(p_o, w) = G(w)$ for all w in $\Sigma^*$. Let $r = \max\{|\lambda(p,a)|/p \text{ in } K_S, \text{ a in } \Sigma\}$. Define $K = K_M \times K_S \times \Delta^{(r)} \cup K_S \cup \{q_o'\}$, where $q_o'$ is in neither $K_S$ nor $K_M \times K_S \times \Delta^{(r)}$. Let $X_o$ not be in $\Gamma_M$, and $\Gamma = \Gamma_M \cup \{X_o\}$. Let $F = K_S \cup \{(q,p,\epsilon)/q \text{ in } F, \text{ p in } K_S\}$. Then consider the pda $N = (K, \Sigma, \Gamma, \delta, X_o, q_o', F)$, where $\delta$ is defined as follows:

For q in $K_M$, p in $K_S$, and Z in $\Gamma_M$, if and only if there is a b in $\Delta$ such that $\delta_M(q,b,Z) \neq \emptyset$, for all a in $\Sigma$ with $\lambda(p,a) \neq \epsilon$, let $((q,p',w),Z)$ be in $\delta((q,p,\epsilon),a,Z)$, where $p' = \delta_S(p,a)$ and $w = \lambda(p,a)$.

If $(q,\gamma)$ is in $\delta_M(q',\epsilon,Z)$, for q and q' in $K_M$, Z in $\Gamma_M$, then for every p in $K_S$, let $(q,p,\epsilon),\gamma)$ be in $\delta((q',p,\epsilon),\epsilon,Z)$.

For b in $\Delta \cup \{\epsilon\}$, u in $\Delta^*$, q and q' in $K_M$, p in $K_S$, Z in $\Gamma_M$, and $\gamma$ in $\Gamma_M^*$, if $(q',\gamma)$ is in $\delta_M(q,b,Z)$, let $((q',p,u),\gamma)$ be in $\delta((q,p,bu),\epsilon,Z)$.

For fixed q in $F_M$, Z in $\Gamma_M$, and any p in $K_S$, if for no b in $\Delta$ is $\delta_M(q,b,Z) \neq \emptyset$, let $(p,\epsilon)$ be in $\delta((q,p,\epsilon),\epsilon,Z)$.

Let $\delta(p,\epsilon,Z) = \{(p,\epsilon)\}$ for all Z in $\Gamma_M$ and p in $K_S$.

For a in $\Sigma$, and p in $K_S$, $\lambda(p,a) = \epsilon$, let $\delta(p,a,X_o) = \{(\delta_S(p,a),X_o)\}$. If $\lambda(p,a) \neq \epsilon$, let $\delta(p,a,X_o) = \emptyset$.

$\delta(q_o',\epsilon,X_o) = \{((q_o,p_o,\epsilon),X_oZ_o)\}$.

Suppose $\lambda(p_o,w) = u$ in L, for given w in $\Sigma^*$, and $(q_o,u,Z_o) \overset{*}{\underset{M}{\vdash}} (q,\epsilon,\gamma)$ for q in F and in $\Gamma_M^*$. Then either $(q_o',w,X_o) \overset{*}{\underset{N}{\vdash}} (Q,\epsilon,X_o\gamma)$, where $Q = (q,\delta_S(p_o,w),\epsilon)$, or there is some initial subword, w', of w, with w'w" = w, such that $(q_o',w,X_o) \overset{*}{\underset{N}{\vdash}} (Q,w'',X_o\gamma) \overset{*}{\underset{N}{\vdash}} (p,\epsilon,X_o)$, where $Q = (q,\delta_S(p_o,w'),\epsilon)$, $p = \delta_S(p_o,w)$ and $\lambda(\delta_S(p_o,w'),w'') = \epsilon$. Hence $G^{-1}(L) \subseteq T(N)$. But it is easy to see that $T(N) \subseteq G^{-1}(L)$, so N accepts G(L).

Now suppose $(q_o', w, X_o) \overset{d*}{\underset{N}{\vdash}} (Q, \epsilon, \gamma)$ for $Q = (q, p, u)$ in $K_M \times K_S \times \Delta^{(r)}$ and $\gamma$ in $\Gamma^*$. Then, $u = \epsilon$ and $p = \delta_S(p_o, w)$, and $(q, \gamma)$ is in $R_M(w)$. If $(q_o', w, X_o) \overset{d*}{\underset{N}{\vdash}} (p, \epsilon, \gamma)$ for $p$ in $K_S$ and $\gamma$ in $\Gamma^*$, then $\gamma = X_o$. Hence $\#R_N(w) \leq \#R_M(w) + \#K_S$, and N is wbb.

One operation that preserves wbb languages, which will be used to show that other operations do likewise, is the operation $f_c$. We first show the following lemmas.

<u>Lemma 2.1</u>. If L is a wbb language, then L = Null(N)[9] for some wbb pda, N.

<u>Proof</u>. Let L = T(M), where $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$ is a wbb pda with bound k. Let $p_1$ and $p_2$ not be symbols of K, and X not in $\Gamma$. Let $K_N = K \cup \{p_1, p_2\}$, $\Gamma' = \{Z'/Z \text{ in } \Gamma\}$, where Z' is an abstract symbol, and $\Gamma_N = \Gamma \cup \Gamma' \cup \{X\}$. Define $\delta_N$ as follows:

If $(p, \gamma)$ is in $\delta(q, a, Z)$, let $(p, \gamma)$ be in $\delta_N(q, a, Z)$ for q in K, a in $\Sigma \cup \{\epsilon\}$ and Z in $\Gamma$.

If $(p, Y\gamma)$ is in $\delta(q, a, Z)$, let $(p, Y'\gamma)$ be in $\delta_N(q, a, Z')$, for Y in $\Gamma$, $\gamma$ in $\Gamma^*$.

If $(p, \epsilon)$ is in $\delta(q, a, Z)$, let $(p_1, X)$ be in $\delta_N(q, a, Z')$.

If q is in F, let $(p_2, \epsilon)$ be in $\delta_N(q, \epsilon, Z)$ for all Z in $\Gamma \cup \Gamma'$.

For all a in $\Sigma$ and Z in $\Gamma \cup \Gamma'$, let $\delta_N(p_1, a, X) = \{(p_1, X)\}$ and $\delta_N(p_2, \epsilon, Z) = \{(p_2, \epsilon)\}$.

Let N be the pda $(K_N, \Sigma, \Gamma_N, \delta_N, Z_o', q_o, \emptyset)$. It is known [7] that T(M) = Null(N). We note that if, for $q \neq p_2$, $(q_o, w, Z_o') \overset{*}{\underset{N}{\vdash}} (q, \epsilon, \gamma)$, then either $\gamma = X$ or $\gamma = Z'\gamma_1$ for Z' in $\Gamma'$ and $\gamma_1$ in $\Gamma^*$. Hence, $\epsilon$ never occurs on the pushdown

---

(9) For pda $N = (K_N, \Sigma, \Gamma_N, \delta_N, Z_o', q_o, F)$, $\text{Null}(N) = \{w/(q_o, w, Z_o') \overset{*}{\underset{N}{\vdash}} (q, \epsilon, \epsilon)$ for some q in $K_N\}$.

tape unless N is in state $p_2$. Also, it is never true that $(q_o,w,Z_o') \overset{d*}{\underset{N}{\vdash}} (p_2, \epsilon, \gamma)$ for any w in $\Sigma^*$, $\gamma$ in $\Gamma_N{}^*$, and if $(q_o,w,Z_o') \overset{d*}{\underset{N}{\vdash}} (p_1, \epsilon, \gamma)$, then $\gamma = X$. Finally, if $(q_o,w,Z_o') \overset{d*}{\underset{N}{\vdash}} (q, \epsilon, Z'\gamma)$ for any q in K, Z' in $\Gamma'$ and $\gamma$ in $\Gamma^*$, then $(q_o,w,Z_o) \overset{d*}{\underset{M}{\vdash}} (q, \epsilon, Z\gamma)$. We thus conclude that $\#R_N(w) \leq \#R_M(w) + 1$, hence that N is a wbb pda.

__Note__. The standard proof [7] also shows that if L = Null(M) for a wbb pda M, then L is a wbb language. Also, if M is sbb, then L is an sbb language. However, it is not true that every sbb language is Null(M) for some sbb pda, M. In fact, it is not hard to show that in an sbb language, L is Null(M) for some sbb pda, M, if and only if there exists an integer $k_1$ such that for any word, w, in L, $\#\{u/u \leq w \text{ and } u \text{ in } L\} \leq k_1$.

__Lemma 2.2__. It is decidable for a given pda $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$ and $q_1$ in K, whether or not there exists w in $\Sigma^*$ such that $(q_o,w,Z_o) \overset{*}{\vdash} (q_1, \epsilon, \epsilon)$.

__Proof__. Let $\Gamma' = \{Z'/Z \text{ in } \Gamma\}$, and $\Gamma_N = \Gamma \cup \Gamma'$. Define $\delta_N$ as follows:

$\delta_N(q,a,Z) = \delta(q,a,Z)$ for all q in K, a in $\Sigma \cup \{\epsilon\}$ and Z in $\Gamma$.

If $(q',Y\gamma)$ is in $\delta(q,a,Z)$, let $(q',Y'\gamma)$ be in $\delta_N(q,a,Z')$.

If $(q',\epsilon)$ is in $\delta(q,a,Z)$, let $(q',\epsilon)$ be in $\delta_N(q,a,Z')$ if and only if $q' = q_1$.

Then it is easy to see that if $N = (K_N, \Sigma, \Gamma_N, \delta_N, Z_o', q_o, \emptyset)$, then Null(N) = $\{w/(q_o,w,Z_o) \overset{*}{\underset{M}{\vdash}} (q_1, \epsilon, \epsilon)\}$. Since Null(N) is a language for any pda, N [2,4,7], it is decidable whether or not Null(N) = $\emptyset$.

__Corollary__. For c in $\Sigma$ and $q_1$ in K, it is decidable for pda M, whether or not there exists w in $\Sigma^*$ such that $(q_o,cw,Z_o) \overset{*}{\underset{M}{\vdash}} (q_1, \epsilon, \epsilon)$.

__Proof__. Null(N) $\cap\ c\Sigma^*$ is a language.

<u>Theorem 2.3</u>. Let L be a wbb language $\subseteq \Sigma^*$, and c an element of $\Sigma$. Then $f_c(L)$ is a wbb language.

<u>Proof</u>. By Lemma 2.1, L = Null(M) for some wbb pda $M = (K, \Sigma, \Gamma, \delta, Z_0, q_0, \emptyset)$ with bound k. Define $\Gamma_N = \{[Z,S,T]/Z \text{ in } \Gamma,\ S \text{ and } T \text{ subsets of } K\}$. For (q,Z) in $K \times \Gamma$, define the function $\theta$ by $\theta(q,Z) = \{p/p \text{ in } K \text{ and for some } w \text{ in } \Sigma^*,\ (q, cw, Z) \overset{*}{\underset{M}{\vdash}} (p, \epsilon, \epsilon)\}$. Define $\psi(q,Z) = \{p/p \text{ in } K \text{ and for some } w \text{ in } \Sigma^*,\ (q,w,Z) \overset{*}{\underset{M}{\vdash}} (p, \epsilon, \epsilon)\}$. Define $\chi(q,Z) = \{p/p \text{ in } K \text{ and } (q, \epsilon, Z) \overset{*}{\underset{M}{\vdash}} (p, \epsilon, \epsilon)\}$. $\theta$ and $\psi$ are effectively calculable by Lemma 2.2, and it is easy to see that $\chi$ is likewise.[10] Let $K_N = K \cup \{q'/q \text{ in } K\}$ and $F_N = \{q'/q \text{ in } K\}$. Define $\delta_N$ as follows:

If $(p, \epsilon)$ is in $\delta(q,a,Z)$ for q in K, a in $\Sigma \cup \{\epsilon\}$ and Z in $\Gamma$, let $(p, \epsilon)$ be in $\delta_N(q, a, [Z,S,T])$ for all S, $T \subseteq K$.

If $(p, Z_1 Z_2 \ldots Z_s)$ is in $\delta(q,a,Z)$, $s \geq 1$, let $(p, [Z_1,S_1,T_1][Z_2,S_2,T_2]\ldots[Z_s,S_s,T_s])$ be in $\delta_N(q,a,[Z,S,T])$ for each S and T, each contained in K, where $S_1 = S$, $T_1 = T$, and for $2 \leq i \leq s$, $S_i = \{q/\psi(q,Z_{i-1}) \cap S_{i-1} \neq \emptyset\}$, $T_i = \{q/\theta(q,Z_{i-1}) \cap S_{i-1} \neq \emptyset\} \cup \{q/\chi(q,Z_{i-1}) \cap T_{i-1} = \emptyset\}$.

If, for q in K and [Z,S,T] in $\Gamma_N$, $\theta(q,Z) \cap S \neq \emptyset$ or $\chi(q,Z) \cap T \neq \emptyset$, let $(q', [Z,S,T])$ be in $\delta_N(q, \epsilon, [Z,S,T])$.

Then consider the pda $N = (K_N, \Sigma, \Gamma_N, \delta_N, X_0, q_0, F_N)$, where $X_0 = [Z_0, K, \emptyset]$. It is easy to see by induction on the length of the pushdown tape, $[Z_1,S_1,T_1]$

---

[10] It is not actually necessary to have $\theta$, $\psi$, and $\chi$ calculable. We could consider all possible pda defined as the pda N is defined in this proof, but with arbitrary functions $\theta$, $\psi$, and $\chi$ from $K \times \Gamma$ to the subsets of K, and know that one of them accepted the desired set. However, the proof is somewhat simpler because the computability is used to advantage.

$[Z_2,S_2,T_2]\ldots[Z_r,S_r,T_r]$, that q is in $S_i$, $1 \le i \le r$, if and only if for some w in $\Sigma^*$ and p in K, $(q,w,Z_1Z_2\ldots Z_{i-1}) \vdash_M^* (p,\epsilon,\epsilon)$. Also, q is in $T_i$ if and only if for some w in $\Sigma^*$ and p in K, $(q,cw,Z_1Z_2\ldots Z_{i-1}) \vdash_M^* (p,\epsilon,\epsilon)$. Let h be the homomorphism from $\Gamma_N^*$ to $\Gamma^*$ that sends $[Z,S,T]$ to Z for all S and T. A word, w, is in T(N) if and only if for some q in K and $\gamma$ in $\Gamma_N^*$, $(q_0,w,X_0) \vdash_N^* (q,\epsilon,\gamma) \vdash_N (q',\epsilon,\gamma)$. But then and only then will there be a word w' in $\Sigma^*$ and p in K such that $(q,cw',h(\gamma)) \vdash_M^* (p,\epsilon,\epsilon)$, so that wcw' is in Null (M). Thus $T(N) = \{u/ucv$ is in L for u and v in $\Sigma^*\}$, and if N is wbb, $T(N) \cap (\Sigma-\{c\})^* = f_c(L)$ will be a wbb language. Hence it is sufficient to show that N is a wbb pda.

We note that if $(q_0,w,X_0) \vdash_N^{d*} (q,\epsilon,\gamma)$, then q is a member of K and $(q_0,w,Z_0) \vdash_M^{d*} (q,\epsilon,h(\gamma))$. Also, it is easy to see by induction on the length of the push-down dape, $[Z_1,S_1,T_1][Z_2,S_2,T_2]\ldots[Z_r,S_r,T_r]$, that the sequence $Z_1,Z_2,\ldots,Z_r$, together with the functions $\theta$, $\psi$, and $\chi$, uniquely determines $S_i$ and $T_i$ for all i, $1 \le i \le r$. Thus if $(q_0,w,X_0) \vdash_N^{d*} (q,\epsilon,\gamma_1)$ and $(q_0,w,X_0) \vdash_N^{d*} (q,\epsilon,\gamma_2)$ and $h(\gamma_1) = h(\gamma_2)$, then $\gamma_1 = \gamma_2$. We may conclude that $\#R_N(w) = \#R_M(w)$, hence that N is wbb.

<u>Theorem 2.4</u>. If L is a wbb language, and R a regular set, (a) L/R is a wbb language, (b) Init(L) is a wbb language.

<u>Proof</u>. (a) As in [6], we may define a gsm mapping, G, that takes a to a for all a in $\Sigma$, the joint vocabulary of L and R, and c to $\epsilon$ for some c not in $\Sigma$. Then $L/R = f_c(G^{-1}(L) \cap \Sigma^*cR)$. Since $f_c$, $G^{-1}$ and intersection with a regular set preserve wbb languages, by Theorems 2.1, 2.2 and 2.3, L/R is a wbb language.

(b) $Init(L) = L/\Sigma^*$, hence Init also preserves wbb languages.

<u>Theorem 2.5</u>. If L is a wbb language and R a regular set, then LR is a wbb language.

Proof. Let $L = T(M)$, where $M = (K_M, \Sigma, \Gamma_M, \delta_M, Z_o, q_o, F_M)$ is a wbb machine with bound k. Let $R = T(A)$, where $A = (K_A, \Sigma, \delta_A, p_o, F_A)$ is a finite automaton. We will assume $K_A \cap K_M = \emptyset$, and that $q_o'$ is a member of neither $K_A$ nor $K_M$. Let $K = K_A \cup K_M \cup \{q_o'\}$, $X_o$ not be an element of $\Gamma_M$, and $\Gamma = \Gamma_M \cup \{X_o\}$. Define $\delta$ as follows:

$\delta(q_o', \epsilon, X_o) = \{(q_o, X_oZ_o)\}$

If q is in $K_M - F_M$, a in $\Sigma \cup \{\epsilon\}$, and Z in $\Gamma_M$, let $\delta(q,a,Z) = \delta_M(q,a,Z)$. For q in $F_M$, Z in $\Gamma_M$, let $\delta(q,a,Z) = \delta_M(q,a,Z) \cup \{(p_o, \epsilon)\}$.

$\delta(p_o, \epsilon, Z) = \{(p_o, \epsilon)\}$ for all Z in $\Gamma_M$.

$\delta(p, a, X_o) = \{(\delta_A(p,a), X_o)\}$, for all p in $K_A$ and a in $\Sigma$.

$\delta(q,a,Z) = \emptyset$ otherwise.

Define pda $N = (K, \Sigma, \Gamma, \delta, X_o, q_o', F_A)$. Evidently, if $(q_o', w, X_o) \overset{d*}{\underset{N}{\vdash}} (q, \epsilon, \gamma)$, and q is in $K_A$, then $\gamma = X_o$. If q is in $K_M$, $(q_o, w, Z_o) \overset{d*}{\underset{M}{\vdash}} (q, \epsilon, \gamma')$, where $\gamma = X_o\gamma'$. Furthermore, it is impossible that q could be $q_o'$. Thus $\#R_N(w) \leq \#R_M(w) + \#K_A$, and if M is wbb, N is wbb.

## Section 3. Information Lossless Gsm Mappings

We will now consider two restricted classes of gsm's, and show that one preserves deterministic, sbb, and wbb languages, the other wbb languages only. These results are significant, for we will show later that not all gsm mappings preserve sbb languages, and it is strongly suspected that the same applies to wbb languages.

Definition. A gsm, $S = (K, \Sigma, \Delta, \delta, \lambda, p_o)$, is said to be information lossless (IL) if for $x_1$ and $x_2$ in $\Sigma^*$, $x_1 \neq x_2$ implies either $\delta(p_o, x_1) \neq \delta(p_o, x_2)$ or $\lambda(p_o, x_1) \neq \lambda(p_o, x_2)$ [3].

A gsm is said to be <u>information lossless of order k</u> (IL-k) if for $x_1 = ax_1'$ and $x_2 = bx_2'$, a and b in $\Sigma$, $x_1'$ and $x_2'$ in $\Sigma^*$, $|x_1| > k$, $|x_2| > k$, and any p in K, $\lambda(p,x_1) \leq \lambda(p,x_2)$ implies a = b.

Note that an IL-k gsm is IL. An IL gsm not IL-k for any k is said to be <u>information lossless of infinite order</u>.

We will show that IL-k gsm's preserve deterministic and sbb languages, but first will need a simple lemma.

<u>Lemma 3.1</u>. If L is a deterministic language, then there exists a deterministic pda, M, with T(M) = L and Null (M) = $\emptyset$.

<u>Proof</u>. Let L = T(N), where N = $(K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$ is a deterministic pda. Let $X_o$ not be in $\Gamma$, and $\Gamma_M = \Gamma \cup \{X_o\}$. Let $q_o'$ not be in K, and $K_M = K \cup \{q_o'\}$. Define $\delta_M(q,a,Z) = \delta(q,a,Z)$ for all q in K, a in $\Sigma \cup \{\epsilon\}$ and Z in $\Gamma$, and $\delta_M(q_o', \epsilon, X_o) = \{(q_o, X_oZ_o)\}$. $\delta_M$ has value $\emptyset$ elsewhere. Then it is trivial to show that M = $(K_M, \Sigma, \Gamma_M, \delta_M, X_o, q_o', F)$ satisfies the requirements of the lemma.

<u>Lemma 3.2.</u> If L is a deterministic language and G an IL-k gsm mapping, then G(L) is a deterministic language.

<u>Proof</u>. Let M = $(K_M, \Sigma, \Gamma_M, \delta_M, Z_o, q_o, F)$ be a deterministic pda accepting L. Let S = $(K_S, \Sigma, \Delta, \delta_S, \lambda, p_o)$ be an IL-k gsm, with $G(w) = \lambda(p_o, w)$ for all w in $\Sigma^*$. We assume, without loss of generality, by Lemma 3.1, that Null(M) = $\emptyset$. Let $r_1$ = $\max\{|\lambda(p,a)|/a \text{ in } \Sigma,\ p \text{ in } K_S\}$, and $r = r_1(k+1)$. Define $K' = K_M \times K_S \times \Delta^{(r)}$, $K'' = \{x'/x \text{ in } K'\}$, and $K = K' \cup K''$. Also, let $\Gamma = \{(Z,\tau)/Z \text{ is in } \Gamma_M$ and $\tau$ is a mapping from $K_M$ to the subsets of $\Sigma^{(k)}\}$. Intuitively, $\tau(q)$ will represent the words in $\Sigma^{(k)}$ that would leave the pda, M, in a state in F, beginning in state q, but with the top symbol of the pushdown tape erased.

Define a function, $\mu$, from $K_S \times \Delta^{(r)}$ to $\Sigma \cup \{\emptyset\}$, such that $\mu(p,w) = a$ if $a$ is the unique element of $\Sigma$ such that $\lambda(p,a) \leq w$, and $\mu(p,w) = \emptyset$ otherwise.

For $q$ and $q'$ in $K_M$, and $Z$ in $\Gamma_M$, define $\psi(q,q',Z) = \{w/(q,w,Z) \vdash_M^* (q', \epsilon, \epsilon)\} \cap \Sigma^{(k)}$.

Finally, define $\theta(q,Z) = \{w/(q,w,Z) \vdash_M^* (q', \epsilon, \gamma)$ for $q'$ in $F$ and $\gamma$ in $\Gamma^*\} \cap \Sigma^{(k)}$.

Let $N$ be the pda $(K, \Delta, \Gamma, \delta, X_0, Q_0, K'')$, where $X_0 = (Z_0, \tau_0)$ and for all $q$ in $K_M$, $\tau_0(q) = \emptyset$, and $Q_0 = (q_0, p_0, \epsilon)$. Let $\delta$ be defined as follows:

(1) If $(Z,\tau)$ is in $\Gamma$, $(q,p,w)$ in $K'$, and for some word, $x$, in $\theta(q,Z) \cup \bigcup_{q' \text{ in } K_M} \psi(q,q',Z)\tau(q') \cap \Sigma^{(k)}$, $\lambda(p,x) = w$, let $\delta((q,p,w), \epsilon, (Z,\tau)) = \{((q,p,w)', (Z,\tau))\}$.

(2) If $\delta_M(q,\epsilon,Z) = \{(t,\epsilon)\}$, let $\delta(Q,\epsilon,X) = \{(T,\epsilon)\}$, for all $X = (Z,\tau)$, in $\Gamma$, $\tau$ arbitrary, $w$ in $\Delta^{(r)}$, and $p$ in $K_S$, where $Q = (q,p,w)'$ if $\delta((q,p,w),\epsilon,X)$ is defined by rule (1), $Q = (q,p,w)$ if not, and $T = (t,p,w)$.

(3) If $\delta_M(q,\epsilon,Z) = \{(t, Z_1 Z_2 \ldots Z_s)\}$, $s \geq 1$, let $\delta(Q,\epsilon,X) = \{(T, X_1 X_2 \ldots X_s)\}$ for all $X$ in $\Gamma$, $X = (Z,\tau)$, arbitrary $\tau$, $w$ in $\Delta^{(r)}$, and $p$ in $K_S$, where $T$ and $Q$ are defined as in (2), $X_1 = (Z_1, \tau_1)$, $\tau_1 = \tau$, and for $2 \leq i \leq s$, $X_i = (Z_i, \tau_i)$ where for $q'$ in $K_M$, $\tau_i(q') = \theta(q', Z_{i-1}) \cup \bigcup_{q'' \text{ in } K_M} \psi(q', q'', Z_{i-1})\tau_{i-1}(q'') \cap \Sigma^{(k)}$.

(4) Let $q$ be in $K_M$, $p$ in $K_S$, $w$ in $\Delta^{(r)}$, $Z$ in $\Gamma_M$, $\delta_M(q,\epsilon,Z) = \emptyset$, and $\mu(p,w) = a$ in $\Sigma$. Further, define $w'$ by $\lambda(p,a)w' = w$, $p'$ by $\delta_S(p,a) = p'$, and assume $\delta_M(q,a,Z) = \{(t,\epsilon)\}$. Then let $\delta(Q,\epsilon,X) = \{(T,\epsilon)\}$, where $Q = (q,p,w)'$ if $\delta((q,p,w),\epsilon,X)$ is defined by rule (1), $Q = (q,p,w)$ if not, $X = (Z,\tau)$, arbitrary $\tau$, and $T = (t,p',w')$.

(5) Under the same conditions as (4), except $\delta_M(q,a,Z) = \{(t,Z_1Z_2\cdots Z_s)\}$, $s \geq 1$, let $\delta(Q,\epsilon,X) = \{(T,X_1X_2\cdots X_s)\}$, Q, X, and T as in (4), $X_1,X_2,\ldots,X_s$ as in (3).

(6) For q in $K_M$, p in $K_S$, w in $\Delta^{(r-1)}$, X in $\Gamma$, $X = (Z,\tau)$, $\delta_M(q,\epsilon,Z) = \emptyset$ and $\mu(p,a) = \emptyset$, for all b in $\Delta$, let $\delta(Q,b,X) = \{((q,p,wb),X)\}$, where $Q = (q,p,w)'$ if $\delta((q,p,w),\epsilon,X)$ is defined by rule (1), $Q = (q,p,w)$ if not.

(7) $\delta(Q,b,X) = \emptyset$ for all Q in K, b in $\Delta \cup \{\epsilon\}$ and X in $\Gamma$ unless explicitly defined otherwise by rules (1) - (6).

It is straightforward to see that N is a deterministic pda, since under no circumstances may two rules be applied simultaneously. Let $\gamma$ be in $\Gamma^*$, $\gamma = (Z_1,\tau_1)(Z_2,\tau_2)\cdots(Z_s,\tau_s)$, $s \geq 1$. Then it is easy to see by induction on s, that for any q in $K_M$, and $2 \leq i \leq s$, $\tau_i(q) = \{w/(q,w,Z_1Z_2\cdots Z_{i-1}) \vdash_M^* (q',\epsilon,\gamma')$ for some q' in F and $\gamma'$ in $\Gamma_M^*\} \cap \Sigma^{(k)}$. Also, $\tau_1(q) = \emptyset$.

Let h be the homomorphism from $\Gamma^*$ to $\Gamma_M^*$ which takes $(Z,\tau)$ to Z for all $\tau$. Suppose for some u in $\Delta^*$, Q in K' and $\gamma$ in $\Gamma^*$, $(Q_o,u,X_o) \vdash_N^* (Q,\epsilon,\gamma)$. Let $Q = (q,p,u')$. Then for some u" in $\Delta^*$, u"u' = u. Also, there is some w in $\Sigma^*$ such that $\lambda(p_o,w) = u''$, $\delta_S(p_o,w) = p$, and $(q_o,w,Z_o) \vdash_M^* (q,\epsilon,h(\gamma))$. We note that $h(\gamma)$ is not $\epsilon$, hence $\gamma \neq \epsilon$. But $(Q,\epsilon,\gamma) \vdash_N (Q',\epsilon,\gamma)$ if and only if there is some word w' in $\Sigma^{(k)}$ such that ww' is in L and $\lambda(p_o,ww') = u$. Thus $T(N) \subseteq G(L)$.

<u>Theorem 3.1</u>. Information lossless gsm mappings of finite order, k, preserve sbb languages.

<u>Proof</u>. All gsm mappings are additive, hence sbb languages are preserved by IL-k mappings according to Lemmas 1.5 and 3.2.

<u>Theorem 3.2</u>. If L is a wbb language, and G an IL gsm mapping, then G(L) is a wbb language.

<u>Proof</u>. Let $M = (K_M, \Sigma, \Gamma, \delta_M, Z_o, q_o, F_M)$, $L = T(M)$, and M be a wbb pda with bound k. Let $S = (K_S, \Sigma, \Delta, \delta_S, \lambda, p_o)$ be an IL gsm realizing the mapping, G. Suppose the maximum of $|\lambda(p,a)|$, over all a in $\Sigma$ and p in $K_S$, is r. Let $K = K_M \times K_S \times \Delta^{(r-1)}$. Define $\delta$ as follows:

If (q,p,w) is in K, b in $\Delta$, Z in $\Gamma$, and $|w| = r-1$, let $\delta((q,p,w),b,Z) = \bigcup_{\substack{a \text{ in } \Sigma \\ \lambda(p,a)=wb}} \{((q', \delta_S(p,a), \epsilon), \gamma)/(q',\gamma) \text{ is in } \delta_M(q,a,Z)\}$.

If $|w| < r-1$, and for some a in $\Sigma$, $wb \le \lambda(p,a)$ and $\delta_M(q,a,Z) \ne \emptyset$, let $\delta((q,p,w),b,Z) = \bigcup_{\substack{a \text{ in } \Sigma \\ \lambda(p,a)=wb}} \{((q', \delta_S(p,a), \epsilon), \gamma)/(q',\gamma) \text{ is in } \delta_M(q,a,Z)\} \cup \{((q,p,wb),Z)\}$.

For q in $K_M$, p in $K_S$ and Z in $\Gamma$, let $\delta((q,p,\epsilon),\epsilon,Z) = \{((q',p,\epsilon),\gamma)/(q',\gamma)$ is in $\delta_M(q,\epsilon,Z)\} \cup \bigcup_{\substack{a \text{ in } \Sigma \\ \lambda(p,a)=\epsilon}} \{((q', \delta_S(p,a), \epsilon), \gamma)/(q',\gamma) \text{ is in } \delta_M(q,a,Z)\}$.

Let $F = \{(q,p,\epsilon)/q \text{ in } F_M \text{ and } p \text{ in } K_S\}$. Define the pda, $N = (K, \Delta, \Gamma, \delta, Z_o, Q_o, F)$, where $Q_o = (q_o, p_o, \epsilon)$.

Suppose $w = a_1 a_2 \cdots a_s$, $s \ge 0$, and w is in L. Then for some $q_i$, $q_i'$, $\gamma_i$, $\gamma_i'$, $1 \le i \le s$, $q_{s+1}$ and $\gamma_{s+1}$, $(q_o, w, Z_o) \overset{*}{\underset{M}{\vdash}} (q_1, w, \gamma_1) \underset{M}{\vdash} (q_1', a_2 a_3 \cdots a_s, \gamma_1') \overset{*}{\underset{M}{\vdash}} (q_2, a_2 a_3 \cdots a_s, \gamma_2) \underset{M}{\vdash} (q_2', a_3 a_4 \cdots a_s, \gamma_2') \overset{*}{\underset{M}{\vdash}} \cdots \overset{*}{\underset{M}{\vdash}} (q_s, a_s, \gamma_s) \underset{M}{\vdash} (q_s', \epsilon, \gamma_s') \overset{*}{\underset{M}{\vdash}} (q_{s+1}, \epsilon, \gamma_{s+1})$, where $q_{s+1}$ is in $F_M$. For $1 \le i \le s$, let $p_i = \delta_S(p_o, a_1 a_2 \cdots a_i)$, $p_i' = \delta_S(p_o, a_1 a_2 \cdots a_{i-1})$, $Q_i' = (q_i', p_i, \epsilon)$, and $Q_i = (q_i, p_i', u_i)$, where if $\lambda(p_i', a_i) = \epsilon$, then $u_i = \epsilon$ and $c_i = \epsilon$, and $\lambda(p_i', a_i) = u_i c_i$ for $u_i$ in $\Delta^*$ and $c_i$ in $\Delta$, otherwise. Then it should be evident that $(Q_o, \lambda(p_o, w), Z_o) \overset{*}{\underset{N}{\vdash}} (Q_1, c_1 \lambda(p_1, a_2 a_3 \cdots a_s), \gamma_1) \underset{N}{\vdash} (Q_1', \lambda(p_1, a_2 a_3 \cdots a_s), \gamma_1') \overset{*}{\underset{N}{\vdash}} (Q_2, c_2 \lambda(p_2, a_3 a_4 \cdots a_s), \gamma_2) \underset{N}{\vdash} (Q_2', \lambda(p_2, a_3 a_4 \cdots a_s), \gamma_2') \overset{*}{\underset{N}{\vdash}} \cdots \overset{*}{\underset{N}{\vdash}} (Q_s, c_s, \gamma_s) \underset{N}{\vdash} (Q_s', \epsilon, \gamma_s') \overset{*}{\underset{N}{\vdash}} ((q_{s+1}, p_s, \epsilon), \epsilon, \gamma_{s+1})$. Thus, if w is in T(M), $\lambda(p_o, w)$ is in T(N).

Now suppose $(Q_o,u,Z_o) \vdash_N^* (Q,\epsilon,\gamma)$, for Q in K, of the form $(q,p,\epsilon)$, $\gamma$ in $\Gamma^*$, and u in $\Delta^*$. We will prove by induction on $|u|$, that there must be some w in $\Sigma^*$ such that $\lambda(p_o,w) = u$ and $(q_o,w,Z_o) \vdash_M^* (q,\epsilon,\gamma)$. Suppose $u = \epsilon$. Then $(Q_o,\epsilon,\gamma_o) \vdash_N (Q_1,\epsilon,\gamma_1) \vdash_N \cdots \vdash_N (Q_s,\epsilon,\gamma_s)$, where $Q_s = Q, \gamma_s = \gamma$ and $\gamma_o = Z_o$. Let $Q_i = (q_i,p_i,\epsilon)$, $0 \le i \le s$. Then for each i between 1 and s, either $(q_{i-1},\epsilon,\gamma_{i-1}) \vdash_M (q_i,\epsilon,\gamma_i)$ or $(q_{i-1},a,\gamma_{i-1}) \vdash_M (q_i,\epsilon,\gamma_i)$ for some a in $\Sigma$. Thus it is obvious that a w exists with $\lambda(p_o,w) = \epsilon$ and $(q_o,w,Z_o) \vdash_M^* (q,\epsilon,\gamma)$.

Now, let $|u| = t$, $t \ge 1$, and assume the inductive hypothesis for all $t' < t$. Then $u = u'b$ for some $u'$ in $\Delta^*$ and b in $\Delta$. There are some $Q', Q'', \gamma'$, and $\gamma''$ such that $(Q_o,u,Z_o) \vdash_N^* (Q',b,\gamma') \vdash_N (Q'',\epsilon,\gamma'') \vdash_N^* (Q,\epsilon,\gamma)$, where $Q'$ is of the form $(q',p',u'')$ and $Q''$ is of the form $(q'',p'',\epsilon)$. $u''$ is a terminal subword of $u'$ of length r-1 or less, so let $u'''u'' = u'$. Then for some $Q'''$ of the form $(q''',p''',\epsilon)$, and $\gamma'''$ in $\Gamma^*$, $(Q_o,u''',Z_o) \vdash_N^* (Q''',\epsilon,\gamma''')$ and $\lambda(p''',a) = u''b$ for some a in $\Sigma$. By the inductive hypothesis, there is some $w'$ such that $\lambda(p_o,w') = u'''$ and $(q_o,w',Z_o) \vdash_M^* (q''',\epsilon,\gamma''')$. Then $\lambda(p_o,wa) = u$ and $(q_o,wa,Z_o) \vdash_M^* (q'',\epsilon,\gamma'')$. By the argument used in the case $u = \epsilon$, there will be a $w''$ in $\Sigma^*$ such that $\lambda(p'',w'') = \epsilon$ and $(q'',w'',\gamma'') \vdash_M^* (q,\epsilon,\gamma)$. Since $(q,p,\epsilon)$ is in F if and only if q is in $F_M$, we conclude that if u is in T(N), then there is a w in T(M) such that $\lambda(p_o,w) = u$. Thus T(N) = G(L).

Now suppose $(Q_o,v,Z_o) \vdash_N^{d*} ((q,p,u),\epsilon,\gamma)$ for v in $\Delta^*$, q in $K_M$, p in $K_S$, u in $\Delta^{(r-1)}$, and $\gamma$ in $\Gamma^*$. There exists b in $\Delta$ such that $((q,p,u),b,\gamma) \vdash_N (Q,\epsilon,\gamma')$, Q in K, $\gamma'$ in $\Gamma^*$. Let $v = u'u$. Then there is some a in $\Sigma$ and exactly one w such that $(q_o,w,Z_o) \vdash_M^* (q,\epsilon,\gamma)$, $\lambda(p_o,w) = u'$, $\delta_S(p_o,w) = p$, $ub \le \lambda(p,a)$, and $(q,a,\gamma) \vdash_M (q'',\epsilon,\gamma'')$ for some $q''$ in $K_M$ and $\gamma''$ in $\Gamma^*$. Hence $(q_o,w,Z_o) \vdash_M^{d*} (q,\epsilon,\gamma)$. Therefore, if $\#K_S = m$, there are at most kmr pairs, $((q,p,u),\gamma)$, such that $(Q_0,v,Z_0) \overset{d*}{\underset{N}{\vdash}} ((q,p,u),\epsilon,\gamma)$. Hence, N is a wbb machine, and the theorem is proven.

One question has been left unanswered. Do IL gsm's of infinite order preserve sbb languages? The answer is no, but this result will be left for a later section.

## Section 4. The Relation Between the Sbb and Wbb Properties

The main result here is that the sbb languages form a proper subset of the wbb languages.

Lemma 4.1. Every sbb pda is a wbb pda.

Proof. Suppose $M = (K,\Sigma,\Gamma,\delta,Z_0,q_0,F)$ were an sbb pda with bound k, but not wbb. Let $n = \#\Sigma$. Then there is some x in $\Sigma^*$ for which $\#R_M(x) > nk$, hence some a in $\Sigma$ and distinct $(q_i,\gamma_i)$, $1 \le i \le k+1$, in $K \times \Gamma^*$ for which $(q_0,a,Z_0) \overset{*}{\vdash} (q_i,a,\gamma_i) \vdash (q_i',\epsilon,\gamma_i')$ for some $q_i'$ in K and $\gamma_i'$ in $\Gamma^*$. Since M may have no infinite loops with input $\epsilon$, for each i, $1 \le i \le k+1$, there will be some $q_i''$ and $\gamma_i''$ in K and $\Gamma^*$, respectively, such that $(q_i',\epsilon,\gamma_i') \overset{*}{\vdash} (q_i'',\epsilon,\gamma_i'')$ and for no q in K, $\gamma$ in $\Gamma^*$ is $(q_i'',\epsilon,\gamma_i'') \vdash (q,\epsilon,\gamma)$ true. Thus, for each i, $1 \le i \le k+1$, there will be an element in $P_M(xa)$, and it is not hard to see that these elements must be distinct. But then the assumption that k was a bound on $\bigcup_{y=xa} Q_M(y,xa)$ has been violated, so we conclude that if M is sbb, it is wbb.

We will now exhibit a language that is wbb, but not sbb. Let $L_1 = \{a^n b^n / n \ge 1\}$, $L_2 = \{a^n b^n c / n \ge 1\}$, $L_3 = \{a^n b^{2n} cc / n \ge 1\}$, and $L_4 = (L_1 \cup L_2 \cup L_3)^*$.(11)

(11)The language $L_4$ may seem more complicated than necessary, but this language will be needed in the next section to prove an additional result.

We will prove $L_4$ to be wbb, but not sbb. These languages will be as defined above for the remainder of the paper.

Lemma 4.2. Let $M = (K, \Sigma, \Gamma, \delta, Z_o, q_o, F)$ be a loop-free, deterministic pda, with $\Sigma = \{a,b,c\}$, and $T(M) \subseteq L_4$. Let w be in $L_4$, and for an infinity of n, let $wa^n b^n$ be in T(M). Then there must be some m such that $wa^m b^m$ is in T(M), but for no w' in $\Sigma^*$ is $wa^m b^{2m} ccw'$ in T(M).

Proof. Let $S = \{n / wa^n b^n$ is in $T(M)\}$. Assume for some integer, r, and an infinity of arbitrary integers, n, $(q_o, wa^n, Z_o) \stackrel{*}{\vdash} (q_n, \epsilon, \gamma_n)$ for $q_n$ in K, $\gamma_n$ in $\Gamma^*$ and $|\gamma_n| \le r$. Then there exist q and $\gamma$ such that $q = q_n$ and $\gamma = \gamma_n$ for an infinity of n, since K and $\Gamma^{(r)}$ are finite sets. Let $n_1$ be the smallest n with $q_{n_1} = q$ and $\gamma_{n_1} = \gamma$, and $n_2$ the smallest element of S with $n_2 \ge n_1$. We can find $n_3$ with $q_{n_3} = q$ and $\gamma_{n_3} = \gamma$ and $n_3 > n_2$, else $\{n / q_n = q$ and $\gamma_n = \gamma\}$ would be bounded, hence finite, But then, $(q_o, wa^{n_3+n_2-n_1} b^{n_2}, Z_o) \stackrel{*}{\vdash} (q, a^{n_2-n_1} b^{n_2}, \gamma)$ $\stackrel{*}{\vdash} (q', \epsilon, \gamma')$ for some q' in F and $\gamma'$ in $\Gamma^*$. But $n_3 + n_2 - n_1 > n_2$, so $wa^{n_3+n_2-n_1} b^{n_2}$ cannot be a word of $L_4$.

We thus conclude that for every integer, r, there is a smallest integer, $m_r$, such that if $n \ge m_r$, and $(q_o, wa^n, Z_o) \stackrel{*}{\vdash} (q, \epsilon, \gamma)$, then $|\gamma| \ge r$. Since M is loop free, if $(q_o, wa^n, Z_o) \stackrel{*}{\vdash} (q, \epsilon, \gamma)$, the set $\{m / (q, a, \gamma) \stackrel{*}{\vdash} (q', \epsilon, \gamma')$ and $|\gamma'| = m\}$ is bounded. Hence, no m may be $m_r$ for more than a finite number of r, and $\{m / m = m_r$ for some $r\}$ is infinite. We may therefore find an infinite, monotonically increasing sequence of integers, $n_1, n_2, \ldots, n_i, \ldots$, such that for $i \ge 1$, $(q_o, wa^{n_i}, Z_o) \stackrel{*}{\vdash} (q_i, \epsilon, \gamma_i Z_i)$ where $q_i$ is in K, $\gamma_i$ in $\Gamma^*$, and $Z_i$ in $\Gamma$, and if, for $j \ge 1$, $(q_i, a^j, \gamma_i Z_i) \stackrel{*}{\vdash} (q', \epsilon, \gamma')$, then $\gamma_i \le \gamma'$. (Note that it is sufficient that $n_i + 1$ be $m_r$ for some r, and $\delta(q_i, a, Z_i) \ne \emptyset$.) We may then find some

$q$ and $Z$ such that for an infinity of $i$, $q = q_i$ and $Z = Z_i$. Renumbering, if necessary, we may find a new infinite, monotonically increasing sequence, $n_1, n_2, \ldots, n_i, \ldots$ with the property that for all $i \geq 1$, $(q_o, wa^{n_i}, Z_o) \overset{*}{\vdash} (q, \epsilon, \gamma_i Z)$ for some $\gamma_i$ in $\Gamma^*$, and if for $j \geq 1$, $(q, a^j, \gamma_i Z) \overset{*}{\vdash} (q', \epsilon, \gamma')$, then $\gamma_i \leq \gamma'$.

Now suppose there is some $n$ in $S$, $n_i \leq n < n_{i+1}$, for some $i > 2$, such that $(q_o, wa^n b^n, Z_o) \overset{*}{\vdash} (q, a^{n-n_i} b^n, \gamma_i Z) \vdash (q_1', x_1, \gamma_1') \vdash (q_2', x_2, \gamma_2') \vdash \cdots \vdash (q_s', x_s, \gamma_s')$, where $x_s = \epsilon$, $q_s'$ is in $F$, and for $1 \leq j \leq s$, $x_j$ is a terminal subsequence of $a^{n-n_i} b^n$ and $\gamma_i \leq \gamma_j'$. Then $(q_o, wa^{n+n_2-n_1} b^n, Z_o) \overset{*}{\vdash} (q, a^{n-n_i} b^n, \gamma_2 \gamma_i'' Z) \overset{*}{\vdash} (q_s', \epsilon, \gamma_2 \gamma_s''')$, where $\gamma_i = \gamma_1 \gamma_i''$ and $\gamma_s' = \gamma_1 \gamma_s'''$. But $n_2 - n_1 > 0$, so we again have a word in $L_4$ that does not belong.

We have thus shown the existence, for each $n$ in $S$, $n_i \leq n < n_{i+1}$, $i > 2$, of a smallest integer, $p_n \leq n$, such that $(q_o, wa^n b^{p_n}, Z_o) \overset{*}{\vdash} (q, a^{n-n_i} b^{p_n}, \gamma_i Z) \overset{*}{\vdash} (q_n', b^{p_n}, \gamma_i \gamma_n') \overset{*}{\vdash} (q_n'', \epsilon, \gamma_2)$, for some $q_n'$ and $q_n''$ in $K$, and $\gamma_n'$ in $\Gamma^*$. We note that there will be some $q'$ such that $q' = q_n''$ for an infinity of $n$ in $S$.

At this point, let us introduce the assumption contrary to the statement of the lemma we are trying to prove, namely that for each $n$ in $S$, there is a $w'$ in $\Sigma^*$ such that $wa^n b^{2n} ccw'$ is in $T(M)$. Then we may find $s$ and $t$ in $S$, with $q_s'' = q_t'' = q'$, $s > 2t$, and $(q_o, wa^t b^{2s-p_s+p_t} ccw', Z_o) \overset{*}{\vdash} (q', b^{2s-p_s} ccw', \gamma_2) \overset{*}{\vdash} (q'', \epsilon, \gamma'')$, for some $w'$ in $\Sigma^*$, $q''$ in $F$, and $\gamma''$ in $\Gamma^*$. However, $s \geq p_s$, so $2s - p_s + p_t > 2t$, which implies, again, that $T(M)$ is not contained in $L_4$. We may thus conclude the lemma.

<u>Lemma 4.3</u>. $L_4$ is not the finite union of deterministic languages.

<u>Proof</u>. Suppose $L_4 = \bigcup_{i=1}^{k} T(M_i)$, where $M_i$ is deterministic for each $i$. We may assume $M_i$ is loop-free deterministic without loss of generality. Let $w_o = \epsilon$.

For each $j > 0$, assume that $w_{j-1}$ is in $L_4$. Then for all n, $w_{j-1}a^nb^n$ is in $L_4$, and we can find some $i_j$ such that for an infinity of n, $w_{j-1}a^nb^n$ is in $T(M_{i_j})$. Then we can find, by Lemma 4.2, an m such that for no w' in $\{a,b,c\}^*$ is $w_{j-1}a^mb^{2m}ccw'$ in $T(M_{i_j})$. Define $w_j = w_{j-1}a^mb^{2m}cc$, certainly a word in $L_4$. Since $w_o$ is in $L_4$, we see that $w_j$ is in $L_4$ for all j, when defined in the manner above. However, consider $w_k$. No integer may be $i_j$ for two different j, thus, for no i, $1 \leq i \leq k$, does $M_i$ accept a word of the form $w_kw'$, for w' in $\{a,b,c\}^*$. But $w_kab$, for example, is in $L_4$, thus $L_4$ could not be the finite union of deterministic languages.

<u>Lemma 4.4</u>. $L_4$ is a wbb language.

<u>Proof</u>. Consider the pda $M = (\{q_o,q_1,q_2,q_3,q_4\},\{a,b,c\},\{Z_o,Z_1\},\delta,Z_o,q_o,\{q_o\})$ where $\delta$ is defined in table 4.1.

| q | Z | $\delta(q,a,Z)$ | $\delta(q,b,Z)$ | $\delta(q,c,Z)$ | $\delta(q,\epsilon,Z)$ |
|---|---|---|---|---|---|
| $q_o$ | $Z_o$ | $\{(q_o,Z_oZ_1)\}$ | $\emptyset$ | $\emptyset$ | $\emptyset$ |
| $q_o$ | $Z_1$ | $\{(q_o,Z_1Z_1)\}$ | $\{(q_1,\epsilon),(q_2,Z_1)\}$ | $\emptyset$ | $\emptyset$ |
| $q_1$ | $Z_o$ | $\emptyset$ | $\emptyset$ | $\{(q_o,Z_o)\}$ | $\{(q_o,Z_o)\}$ |
| $q_1$ | $Z_1$ | $\emptyset$ | $\{(q_1,\epsilon)\}$ | $\emptyset$ | $\emptyset$ |
| $q_2$ | $Z_o$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ |
| $q_2$ | $Z_1$ | $\emptyset$ | $\{(q_3,\epsilon)\}$ | $\emptyset$ | $\emptyset$ |
| $q_3$ | $Z_o$ | $\emptyset$ | $\emptyset$ | $\{(q_4,Z_o)\}$ | $\emptyset$ |
| $q_3$ | $Z_1$ | $\emptyset$ | $\{(q_2,Z_1)\}$ | $\emptyset$ | $\emptyset$ |
| $q_4$ | $Z_o$ | $\emptyset$ | $\emptyset$ | $\{(q_o,Z_o)\}$ | $\emptyset$ |
| $q_4$ | $Z_1$ | $\emptyset$ | $\emptyset$ | $\emptyset$ | $\emptyset$ |

Table 4.1

It is not hard to see that $T(M) = L_4$ and that m is a wbb pda with a bound of three. The verification will be left to the reader.

Theorem 4.1. The sbb languages are properly included in the wbb languages.

Proof. Immediate from Lemmas 4.1, 4.3 and 4.4.

Section 5. Some Operations that Fail to Preserve the Sbb Property

The common operations of complementation and arbitrary gsm mapping do not preserve sbb languages. We strongly suspect that the same is true of the wbb languages, but do not have a proof, due to the difficulty in exhibiting that a specific language is not wbb. We will also show that IL gsm mappings not of finite order do not necessarily preserve sbb languages.

Lemma 5.1. The language $L_5 = \{a^n b^m / n \le m \le 2n\}$ is not the finite union of deterministic languages.

Proof. The proof of this lemma has many details in common with the proof of Lemma 4.2. We will therefore not go into detail when steps can be filled in a manner analogous to that used in Lemma 4.2.

Assume that for some k, $L_5 = \bigcup_{i=1}^{k} T(M_i)$ for loop-free deterministic pda $M_1$, $M_2, \ldots, M_k$. For $n \ge 0$, let $S_n = \{M_i / M_i$ accepts two words $a^n b^{m_1}$ and $a^n b^{m_2}$ with $m_1 - m_2 \ge n/2k\}$. We observe that for $n > k$, $S_n$ is not empty, for otherwise there would be at most n words of the form $a^n b^m$, in $L_5$, when in fact there are $n + 1$. Also observe that for some i, $M_i$ is in $S_n$ for an infinity of n. Let us choose one such i, and let $M_i = M = (K, \Sigma, \Gamma, \delta, Z_0, q_0, F)$. Let $T = \{n/M$ is in $S_n\}$. By an argument quite similar to that employed in Lemma 4.2, we may show that it is not possible for some r and an infinity of integers, n, to have $(q_0, a^n, Z_0) \overset{*}{\underset{M}{\vdash}}$

$(q, \epsilon, \gamma)$ for some $q$ in $K$ and $\gamma$ in $\Gamma^*$, with $|\gamma| \le r$. Also, we can show that we must then have a $q$ in $K$ and $Z$ in $\Gamma$, and an infinite, monotonically increasing sequence of integers, $n_1, n_2, \ldots, n_i, \ldots$ such that for all $i \ge 1$, $(q_o, a^{n_i}, Z_o) \overset{*}{\underset{M}{\vdash}} (q, \epsilon, \gamma_i Z)$ for some $\gamma_i$ in $\Gamma^*$, and if for $j \ge 1$, $(q, a^j, \gamma_i Z) \overset{*}{\underset{M}{\vdash}} (q', \epsilon, \gamma')$, then $\gamma_i \le \gamma'$. Finally, as in Lemma 4.2, we may see that if $n$ is in $T$, $i > 2$, and $n_i \le n < n_{i+1}$, then it is not possible that $(q_o, a^n b^m, Z_o) \overset{*}{\underset{M}{\vdash}} (q, a^{n-n_i} b^m, \gamma_i Z) \underset{M}{\vdash} (q_1, x_1, \gamma_1') \underset{M}{\vdash} (q_2, x_2, \gamma_2') \underset{M}{\vdash} \cdots \underset{M}{\vdash} (q_s, x_s, \gamma_s')$, where $q_s$ is in $F$, $x_s = \epsilon$, and for $1 \le j \le s$, $x_j$ is a terminal subsequence of $a^{n-n_i} b^m$, and $\gamma_2$ is an initial subsequence of $\gamma_j'$. Hence, for each $n$ in $T$, $n \ge n_3$, there is a smallest integer, $p_n$, such that $(q_o, a^n b^{p_n}, Z_o) \overset{*}{\underset{M}{\vdash}} (q, a^{n-n_i} b^{p_n}, \gamma_i Z) \overset{*}{\underset{M}{\vdash}} (q_n', b^{p_n}, \gamma_i \gamma_n') \overset{*}{\underset{M}{\vdash}} (q_n'', \epsilon, \gamma_2)$, for $q_n'$ and $q_n''$ in $K$, $\gamma_n'$ in $\Gamma^*$, and $n_i \le n \le n_{i+1}$, and if $a^n b^m$ is in $T(M)$, for any $m$, then $m \ge p_n$. We may thus choose $q'$ in $K$ such that for an infinity of $n$ in $T$, $q_n'' = q'$.

Now let us choose $s$ and $t$ in $T$, such that $q_s'' = q_t'' = q'$, and $s > 4kt$. We may find $m_1$ and $m_2$, with $m_1 - m_2 \ge s/2k$, and $a^s b^{m_1}$ and $a^s b^{m_2}$ words of $T(M)$. But then, $(q_o, a^t b^{m_1 + p_t - p_s}, Z_o) \overset{*}{\underset{M}{\vdash}} (q', b^{m_1 - p_s}, \gamma_2) \overset{*}{\underset{M}{\vdash}} (q'', \epsilon, \gamma')$ for some $q''$ in $F$, and $\gamma'$ in $\Gamma^*$. However, $m_1 + p_t - p_s \ge m_1 - m_2 + p_t \ge s/2k > 2t$. Hence $a^t b^{m_1 + p_t - p_s}$ is not in $L_5$. We have thus contradicted the assumption that $L_5$ is the finite union of deterministic languages.

<u>Theorem 5.1</u>. The sbb languages are not closed under (a) complementation (b) arbitrary gsm mappings.

<u>Proof</u>. (a) Let $\Sigma = \{a,b\}$. Then $\Sigma^* - L_5 = \{a^n b^m / m < n\} \cup \{a^n b^m / n \ge 1, m > 2n\} \cup aa^*bb^*a\Sigma^* \cup b\Sigma^*$ is easily seen to be the finite union of deterministic languages. Its complement, $L_5$, of course, is not.

(b) Let G be the gsm mapping which always takes a to a, b to b, and c to $\epsilon$. Let $L_6 = \{a^i c a^j b^{i+2j} / i \geq 0,\ j \geq 0\}$.[12] It is easily seen that $L_6$ is deterministic. But $G(L_6) = L_5$, so not all gsm mappings preserve sbb languages.

Lastly, we would like to show that it is not generally true that an IL gsm mapping of infinite order will preserve the sbb property. We will use the languages $L_1$ and $L_4$ as defined in the previous section.

<u>Theorem 5.2</u>. IL gsm mappings of infinite order do not preserve sbb languages.

<u>Proof</u>. Consider the gsm $S = (\{p_o, p_1, p_2, p_3, p_4\}, \{a,b\}, \{a,b,c\}, \delta, \lambda, p_o)$, where $\delta$ and $\lambda$ are defined in table 5.1.

| $p$ | $\delta(p,a)$ | $\lambda(p,a)$ | | $\delta(p,b)$ | $\lambda(p,b)$ |
|---|---|---|---|---|---|
| $p_o$ | $p_1$ | a | | $p_2$ | b |
| $p_1$ | $p_o$ | $\epsilon$ | | $p_4$ | b |
| $p_2$ | $p_1$ | cca | | $p_2$ | b |
| $p_3$ | $p_1$ | ca | | $p_4$ | b |
| $p_4$ | $p_1$ | a | | $p_3$ | $\epsilon$ |

Table 5.1

Any input word to S may be written $a^{i_1}b^{j_1}a^{i_2}b^{j_2}\ldots a^{i_n}b^{j_n}, n \geq 1$, $i_1, j_1,$ and $j_n \geq 0$, $i_n \geq 1$, and $i_m$ and $j_m \geq 1$ for $2 \leq m < n$. An initial sequence of b's will appear, followed by cc at the output of S. $a^i b^j$, surrounded by either $\epsilon$ or b on the left and a on the right, can be considered to be transformed into

---

[12] The use of this language was first suggested by Joseph Ullian.

$a^{\frac{1}{2}i}b^{j}cc$ if i is even, $a^{\frac{1}{2}(i+1)}b^{\frac{1}{2}j}c$ if i is odd and j is even, and $a^{\frac{1}{2}(i+1)}b^{\frac{1}{2}(j+1)}$ if both i and j are odd. If $a^{i}b^{j}$ is a terminal sequence, the same thing will occur, but without the terminal c's. Note that when $a^{i}b^{j}(i, j \geq 1)$ is a terminal sequence, then the final state distinguishes between the cases, i even, i odd and j even, and i and j odd; and if $a^{i}(i \geq 1)$ is a terminal sequence, then the state likewise distinguishes between i even and odd. Thus S is an IL gsm.

Consider $S(L_1^*) = L_7$. Surely $L_1^*$ is a deterministic language. Suppose $L_7$ were the finite union of deterministic languages. Then $(L_7 \cap a\Sigma^*)/aa^*bb^*$ would be also, by Theorem 1.2, where $\Sigma = \{a,b,c\}$. But $(L_7 \cap a\Sigma^*)/aa^*bb^* = L_4$, hence $L_4$ is not the finite union of deterministic languages, by Lemma 4.3. Thus we have exhibited an IL gsm (obviously of infinite order) and a deterministic language mapped by the gsm into a language that is not sbb.

## REFERENCES

[1] N. Chomsky, "On Certain Formal Properties of Grammars," Information and Control, Vol. 2, 1959, pp. 137-167.

[2] N. Chomsky, "Context Free Grammars and Pushdown Storage," Quarterly Progress Report #65, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., 1962.

[3] S. Even, "Information Lossless Sequential Machines," Ph.D. Thesis, Harvard University, Cambridge, Mass., 1963.

[4] R. J. Evey, "The Theory and Application of Pushdown Store Machines," Report NSF-10, the Computation Laboratory of Harvard University, Cambridge, Mass., May 1963.

[5] S. Ginsburg, "Examples of Abstract Machines," IRE Transactions on Electronic Computers, Vol. EC11, 1962, pp. 132-135.

[6] S. Ginsburg and S. A. Greibach, "Deterministic Context Free Languages," System Development Corporation Report TM-738/014/00, May 1965.

[7] S. Ginsburg, "The Mathematical Theory of Context Free Languages," McGraw Hill Publishing Company, to be published.

[8] L. G. Haines, "Generation and Recognition of Formal Languages," Ph.D. Thesis, Massachusetts Institute of Technology, June 1965.

[9] S. C. Kleene, "Representation of Events in Nerve Nets," "Automata Studies," Edited by C. E. Shannon and J. McCarthy, Princeton University Press, Princeton, N. J., 1956, pp. 3-40.

[10] D. E. Knuth, "On the Translation of Languages from Left to Right," California Institute of Technology, 1965.

[11] M. O. Rabin and D. Scott, "Finite Automata and their Decision Problems," IBM Journal of Research and Development, Vol. 3, 1959, pp. 114-125.

Unclassified

Security Classification

**DOCUMENT CONTROL DATA - R&D**

*(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)*

1 ORIGINATING ACTIVITY *(Corporate author)*

System Development Corporation
Santa Monica, California

2a REPORT SECURITY CLASSIFICATION
Unclassified

2b GROUP

3 REPORT TITLE

Pushdown Automata With Bounded Backtrack

4 DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Scientific Report Interim

5 AUTHOR(S) *(Last name, first name, initial)*

Ullman, J.

| 6 REPORT DATE | 7a. TOTAL NO. OF PAGES | 7b. NO. OF REFS |
|---|---|---|
| 3 December 1965 | 33 | 11 |

8a CONTRACT OR GRANT NO. AF-19(628)-5166, CPL - Algorithmic Languages Program

b. PROJECT & Task No. 5632-05

c. DoD Element No. 61445014

d. DoD Subelement No. 681305

9a. ORIGINATOR'S REPORT NUMBER(S)

TM-738/022/00

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

AFCRL-66-550

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Hq. AFCRL, OAR (CRB)
United States Air Force
Bedford, Massachusetts

13 ABSTRACT This report considers two classes of pushdown automata (pda), and the languages accepted by them. These pda accept their languages rapidly because they reread the input word a limited number of times, Hence, such languages are particularly useful as programming languages. The first class, strong bounded backtrack pda, read input words from left to right, and jump from right to left (backtrack). The languages accepted by such automata will be shown to be equivalent to the finite unions of deterministic languages. The second class, weak bounded backtrack an arbitrary number of times. The device reads a word from left to right, simulating the action of the pda. Every time the pda reaches a total configuration (state and pushdown tape) in which it is possible to read another input letter, that configuration is stored. If no move at all is possible in a given configuration, it is erased from storage. Thus one can accept the language with no backtrack without having to keep track of an arbitrary number of possible configurations of the pda. Several results will be shown about each of these classes of pda, including operations that preserve the properties. While these properties are not preserved by all gsm mappings, it will be shown that information lossless gsm's preserve the weak bounded backtrack property, and information lossless gsm's of finite order preserve the strong bounded backtrack property.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

Unclassfied

Security Classification

| 14 KEY WORDS | LINK A | | LINK B | | LINK C | |
|---|---|---|---|---|---|---|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| Pushdown<br>Automata<br>Bounded<br>Backtrack<br>Programming<br>Mathematical<br>Languages<br>PDA | | | | | | |

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.

Unclassified

Security Classification