Semantics of Context-Free Languages: Correction*

by

DONALD E. KNUTH

Stanford University

Page 137, delete "when $\mathscr{T}_j$ is of the type $q(j)$" at the end of the proof of the theorem.

Page 135, line 14, change T to $\mathscr{T}$.

In the last formula of (2.4), replace "$\{(l(X_{po}), X_{po}) | X_{po} \neq S\}$" by "$\{(l(X_{pj}), X_{pj}) | 1 \leq j \leq n_p\}$".

Page 140, line 1, change "p" to "P"; line 13, change "S" to "Σ".

Page 141, rules 3.1 and 3.2, change "S" to "Σ".

Page 141, rule 6.2, delete "follow (L_2) = newsymbol;" and also delete "include follow (L_2) in Q;".

Finally, replace lines 12 and following of p. 136, up to and including the statement of the theorem on p. 137, by:

The following algorithm may now be used (the algorithm originally published in *Math. Systems Theory* is incorrect): For each X in V let $S(X)$ be a set of directed graphs on the vertices $A(X)$. Initially $S(X)$ is empty for all $X \in N$; and $S(X)$ is the single directed graph with vertices $A(X)$ and no arcs, for all $X \notin N$. Now we add further directed graphs to the sets $S(X)$ by the following procedure until no further graphs can be added: Choose an integer p, with $1 \leq p \leq m$, and for $1 \leq j \leq n_p$ choose a directed graph D_j' in $S(X_{pj})$. Then include in $S(X_{po})$ the directed graph whose vertices are $A(X_{po})$ and whose arcs run from α to α' if and only if there is an oriented path from (X_{po}, α) to (X_{po}, α') in the directed graph

$$D_p[D_1', \cdots, D_{n_p}']. \tag{3.5}$$

It is clear that this process must ultimately terminate with no more directed graphs created, since only finitely many directed graphs are possible in all.

In the case of grammar (1.5), this algorithm constructs the sets

$$S(N) = \{\overset{v}{\cdot}\}, \quad S(L) = \{\overset{v}{\cdot}\ \overset{1}{\cdot}\ \overset{s}{\cdot},\ \overset{v}{\cdot}\ \overset{1}{\cdot}\ \overset{s}{\cdot}\}, \quad S(B) = \{\overset{v}{\cdot}\ \overset{s}{\cdot},\ \overset{v}{\cdot}\ \overset{s}{\cdot}\}, \quad S(0) = S(1) = \{\ \}.$$

If $\mathscr{T}$ is a derivation tree with root X, let $D'(\mathscr{T})$ be the directed graph with vertices $A(X)$ whose arcs run from α to α' if and only if there is an oriented

*The original paper appeared in *Math. Systems Theory* **2** (1968), 127–145.

path from (X, α) to (X, α') in $D(\mathscr{T})$. After the above algorithm terminates, we can show for all $X \in V$ that $S(X)$ *is the set of all* $D'(\mathscr{T})$, *where* $\mathscr{T}$ *is a derivation tree with root* X. For the construction does not add any directed graph to $S(X)$ unless it is such a $D'(\mathscr{T})$; the algorithm could readily be extended so that it would in fact print out an appropriate derivation tree $\mathscr{T}$ for each graph in $S(X)$. Conversely if $\mathscr{T}$ is any derivation tree, we can prove by induction on the number of nodes of $\mathscr{T}$ that $D'(\mathscr{T})$ is in the relevant $S(X)$. Otherwise $\mathscr{T}$ must have the form (3.3), and $D(\mathscr{T})$ is "pasted together" from D_p, $D(\mathscr{T}_1), \cdots, D(\mathscr{T}_{n_p})$. By induction and the fact that no arcs run from $D(\mathscr{T}_j)$ to $D(\mathscr{T}_{j'})$ for $j \neq j'$, any arcs of the assumed path which appear in $D(\mathscr{T}_1), \cdots, D(\mathscr{T}_{n_p})$ may be replaced by appropriate arcs in $D_p[D_1', \cdots, D_{n_p}']$, where D_j' is a member of $S(X_{pj})$ for $1 \leq j \leq n_p$. This directed graph is therefore in $S(X_{po})$, and it is equal to $D'(\mathscr{T})$.

The above algorithm now affords a solution to the problem posed in this section:

THEOREM. *Semantic roles added to a grammar as described in Section 2 are well defined for all derivation trees in the language if and only if none of the directed graphs (3.5), for any choice of* p *and* $D_1' \in S(X_{p1}), \cdots, D_{n_p}' \in S(X_{n_p})$, *contains an oriented cycle.*