

Let E be an expression defining the subtrees of a tree using x to denote nodes, and $($ and $)$ to denote the beginning and ending of subtrees.

G is a grammar for generating such expressions.

$G = (V_N, V_T, S, P)$

where

$V_N = \{ S, A, B, C \}$

$V_T = \{ x, (,) \}$

x is any character that is not $($ or $)$

λ is the empty string

$P = \{ S \rightarrow x \mid x(xA), A \rightarrow \lambda \mid xA \mid (B)C, C \rightarrow \lambda \mid x \mid xA, B \rightarrow x \mid xA \}$

$L(G) = \{ E : E \text{ is a valid tree-defining expression} \}$

We check the correctness of the tree-defining expressions by checking that the following conditions hold

- i. $E[i]$ is in $\{x, (,)\}$
- ii. $E[1] = x$
- iii. Number of $($ = Number of $)$
- iv. $)$ is not in E
- v. $()$ is not in E
- vi. $(($ is not in E