Greibach Normal Form

Definition Greibach Normal Form (GNF)

A CFG is in **Greibach Normal Form** if all productions are of the form

$A \rightarrow a X$

with $a \in T$ and $X \in V^*$.

This means, all productions start with a terminal on the PRODUCTION and have only variables following.

Theorem:

For every CFG G with $\lambda \notin L(G)$ exists an equivalent grammar G' in Greibach NF.

Proof and algorithm

Firstly, we convert the given grammar into Chomsky NF:

Start with a grammar G = (V, T, P, S)

Eliminate useless variables that cannot become terminals.

Eliminate useless variables that cannot be reached.

Eliminate λ-productions.

Eliminate unit productions.

Convert grammar to Chomsky Normal Form.

Then, we convert this grammar into an equivalent grammar in Greibach NF.

Convert a CNF grammar into Greibach Normal Form:

1. Re-label all variables such that the names are A_1, A_2, ... , A_m.

2. We want to order the productions, which are not terminal but contain variables. We use for this purpose the indexing of the variables, so that

   $A_i \rightarrow A_j\, \alpha$ with i<j for all i = 1,...,m and j = 2, ..., m

   We perform the ordering process by substitution of the first variable on the PRODUCTION, if the production violates the condition above (see 2 and 3).

3. We start the ordering with A_1.

   A_1-productions can have only a higher numbered variable as first variable on the PRODUCTION, or a single terminal.

4. We assume now that all rules are okay up to A_{k-1}. The next rule we encounter, with A_k on the production, is the first one, which is not okay:

   $A_k \rightarrow A_l\, \alpha$ with k>l

   We resolve this problem by substituting A_l. Since l<k, the A_l-rules have already gone through the sorting process and are in the proper format:

   $A_l \rightarrow A_j\, \alpha$ with l<j

   Now, we substitute the PRODUCTIONs of A_l in the A_k-rule, and come up with:

   $A_k \rightarrow A_l\, \alpha$ or $A_k \rightarrow a$ for some a

If l is still less than k, we substitute again. And again and again, until we get <u>at least</u> A_k on the PRODUCTION - all rules up to A_{k-1} are already sorted and in proper form, and the first variable on the PRODUCTION of the A_{k-1}-production must be therefore at least A_k.

<u>EXAMPLE 1: TRANSFORM A CFG INTO GREIBACH NORMAL FORM</u>

Bring the grammar G with V={S, A, B}, T={a, b} and productions P

$S \rightarrow A$

$A \rightarrow a\ B\ a \mid a$

$B \rightarrow b\ A\ b \mid b$

into GNF.

Solution:

1. Simplify G:

No useless variables or productions, no λ-productions.

Remove unit-production $S \rightarrow A$:

Replace A with PRODUCTION of A (after calculating transitive closure of unit-productions - but there is only one unit-dependency here $A \Rightarrow B$):

$S \rightarrow a\ B\ a \mid a$ new rule

2. Transform G into an equivalent grammar G' in <u>Chomsky NF</u>:

Substitute terminals on PRODUCTION with variables:

$S \rightarrow C_1\ B\ C_1 \mid a \quad C_1 \rightarrow a$

$A \rightarrow C_1\ B\ C_1 \mid a \quad C_2 \rightarrow b$

$B \to C_2\ A\ C_2 \mid b$

Break down the rules:

$S \to C_1\ D_1 \mid a \quad D_1 \to B\ C_1 \quad C_1 \to a$

$A \to C_1\ D_1 \mid a$

$B \to C_2\ D_2 \mid b \quad D_2 \to A\ C_2 \quad C_2 \to b$

3. Transform G' into equivalent grammar G' in <u>Greibach NF</u>:

a. Rename the variables to V_1, V_2, ... in the productions P':

$S=V_1$; $A=V_2$; $B=V_3$; $C_1=V_4$; $C_2=V_5$; $D_1=V_6$; $D_2=V_7$

$V_1 \to V_4\ V_6 \mid a \quad V_6 \to V_3\ V_4 \quad V_4 \to a$

$V_2 \to V_4\ V_6 \mid a$

$V_3 \to V_5\ V_7 \mid b \quad V_7 \to V_2\ V_5 \quad V_5 \to b$

b. Order the productions:

$V_1 \to V_4\ V_6 \mid a$

$V_2 \to V_4\ V_6 \mid a$

$V_3 \to V_5\ V_7 \mid b$

$V_4 \to a$

$V_5 \to b$

$V_6 \to V_3\ V_4$

$V_7 \to V_2\ V_5$

Rules for V_1, V_2, V_3, V_4, V_5 are okay.

Modify V_6-rule $V_6 \rightarrow V_3\ V_4$:

Substitute PRODUCTIONs of V_3 in V_6-rule:

$V_6 \rightarrow V_5\ V_7\ V_4 \mid b\ V_4$

Substitute PRODUCTIONs of V_5 in modified V_6-rule:

$V_6 \rightarrow b\ V_7\ V_4 \mid b\ V_4$ final V_6-rule

Modify V_7-rule $V_7 \rightarrow V_2\ V_5$:

Substitute PRODUCTIONs of V_2 in V_7-rule:

$V_7 \rightarrow V_4\ V_6\ V_5 \mid a\ V_5$

Substitute PRODUCTIONs of V_4 in modified V_7-rule:

$V_7 \rightarrow a\ V_6\ V_5 \mid a\ V_5$ final V_7-rule

c. Substitute to achieve Greibach Normal Form:

Everything already done in this case. Grammar is in GNF:

$V_1 \rightarrow V_4\ V_6 \mid a$

$V_2 \rightarrow V_4\ V_6 \mid a$

$V_3 \rightarrow V_5\ V_7 \mid b$

$V_4 \rightarrow a$

$V_5 \rightarrow b$

$V_6 \rightarrow b\ V_7\ V_4 \mid b\ V_4$

$V_7 \rightarrow a\ V_6\ V_5 \mid a\ V$