Elimination of epsilon productions

We have to first identify the nullable variables.

A variable A is called nullable if A $\rightarrow^*$ ϵ . Note that a in ∑ cannot be nullable.

Identification of nullable variables

1. E <- { A Ɛ N | A -> ϵ }
2. if A -> A_1 A_2 .... A_k A_i Ɛ E then insert A in E.
3. Repeat 2 until no more insertions in E.

E is the set of nullable variables.

Epsilon-productions can now be eliminated by

1. If A -> X_1 X_2 ..... X_k, X_i in E add a new production

A -> X_1 X_2 ... $X_{i-1}X_{i+1}$ ...... X_k.

2. Repeat 1 until no more new productions can be added,

3. Eliminate all productions of the form A -> ϵ.

In this way we get a Grammar G_1 such that $L(G_1) = L(G) – \{ϵ\}$. To check whether L(G) contains ϵ, we can check whether S is nullable.

Example : S -> A B C

A -> B C | a

B -> b A C | ϵ

C -> d A B | ϵ

Initially E = {B, C}

Since A -> BC , A gets inserted.

Since S -> ABC, S gets inserted.

Thus E = {S, A, B, C}

Since A, B are in E from C -> d A B, we add C -> dA | dB

Since A is in E from C -> d A we add C -> d

Similarly from B -> b A C, we add B -> b A | b C | b

From A -> B C, we add A -> B | C | ϵ

From S -> A B C, we add S -> A B | A C | B C | A | B | C | ϵ

Finally getting rid of epsilon-productions we get,

S -> A B C | A B | A C | B C | A | B | C

A -> B C | B | C

B -> b A C | b A | b C | b

C -> d A B | d A | d B | d.

Since S is nullable (is in E) ϵ is in the language as we can check

S -> A B C -> B C B C -> C B C -> B C -> C -> ϵ .