Elimination of unit productions

Given a CFG G we can construct a CFG G_1 such that $L(G_1) = L(G)$ and G_1 has no unit productions.

For this we have to first find the unit pairs.

A pair of variables (A,B) is a unit pair if either A = B or $A \rightarrow B_1$, $B_1 \rightarrow B_2$, ...

$B_{k-1} \rightarrow B_k = B$.

All unit pairs are found as follows.

1) For every A in N (A,A) is a unit pair.

2) If (A,B) is a unit pair and B -> C where C is a variable then (A,C) is also a unit pair.

All unit pairs are found this way.

After finding unit pairs, the unit productions can be eliminated by starting from an empty set of productions and adding for every unit pair (A,B), A -> all the non-unit production bodies of B. Since (A,A) is a unit pair all the original non-unit productions are retained.

Example : S -> 0A0 | 00 |1B1 |11 | BB | B

A -> C

B -> S | A

C -> S

Unit pairs : (S,S), (A,A), (B,B), (C,C) are unit pairs.

Since (S,S) is a unit pair and S -> B, (S,B) is a unit pair.

.........(S,B) ....B-> A, (S,A) is a unit pair.

In this way (S,C), (A,C), (A,S), (A,B) are unit pairs.

Also (B,A), (B,C) and (B,S) are unit pairs.

Finally (C,S), (C,B) and (C,A) are unit pairs. Thus all pairs are unit pairs.

Now to get rid of unit productions, we note that S -> 0A0 | 00 | 1B1| 11 | BB

are the non-unit productions of S. Hence since (S,S), (A,S), (B,S), (C,S) are unit pairs, we get after getting rid of the unit productions the Grammar,

S -> 0A0 | 00 | 1B1 | 11 | BB

A -> 0A0 | 00 | 1B1 | 11 | BB

B -> 0A0 | 00 | 1B1 | 11 | BB

C -> 0A0 | 00 | 1B1 | 11 | BB

which is equivalent to the given Grammar.