BLG 311E Term Project - Phase 2

For this second phase of the project, you'll implement a recognizer for Context Free Grammars (CFGs) with CYK algorithm (https://en.wikipedia.org/wiki/CYK_algorithm). You should write your program either with **C** or **C++**. Given a grammar (in **grammar.txt** file) and a list of words (in **words.txt** file), your program will decide whether the given words are in the language defined by the grammar. Note that in this phase, the grammar will be in Chomsky Normal Form (CNF).

Specifications

Input:

A sample **grammar.txt** file is given below. First line states the terminals (alphabet) of the language. Second line gives the Non-terminals of the grammar. Note that first non-terminal is the **start symbol.** After these initial two lines, rules of the grammars are given (one rule on each line). Each line starts with the ID of the rule (an integer). Every term in the grammar file is separated by a single space character. For simplicity, you can assume that all the terminals and non-terminals will be just a single character.

```
a b c
S A B C
1 S -> AS
2 S -> CA
3 A -> AB
4 A -> a
5 C -> c
6 C -> CB
7 B -> b
```

The **words.txt** file will contain words (one word on each line). An example words.txt file is given below.

```
abbcbabb
ccba
aaaacabc
```

Output:

For each word in the words.txt file, if the word cannot be generated by the grammar you will only print NO. If the word is in the language, you will only print YES. Note that this is different from phase 1, you don't have to write the production steps. The results should be printed to **results.txt** file. For the grammar and words file given in the previous section, your results.txt file should be as follows:

```
YES
YES
NO
```

Bonus part

If you want to get extra marks, you can also give the production steps for a left-most derivation. For this bonus, part you have to output another file named **bonus.txt**. The format of the file should be just like in phase 1. An example bonus.txt file for the above **grammar.txt** and **words.txt** file is given below. Note that you will still output the **results.txt** file even though the same information is available in **bonus.txt**.

```
YES
1 3 3 4 7 7 2 6 5 7 3 3 4 7 7
YES
2 6 6 5 5 7 4
NO
```

Regulations

- You have to implement in C or C++.
- You will submit two files. One is your parser in file **parser.c (or parser.cpp)** and the other is the report in pdf format, **report.pdf** .
- You are not allowed to use any special parsing libraries.
- You should do the project with the same group members in phase 1.
- Do not share anything in the code level with your friends. Any unnecessary similarities between the codes will receive our special attention. If we find out that you have some piece of code that is not result of your own coding (from friends or from internet without any reference in your report), not only you'll get a 0 from this project but also will also lead to disciplinary action. Worst of all, you'll be stamped as a cheater and a fraud.