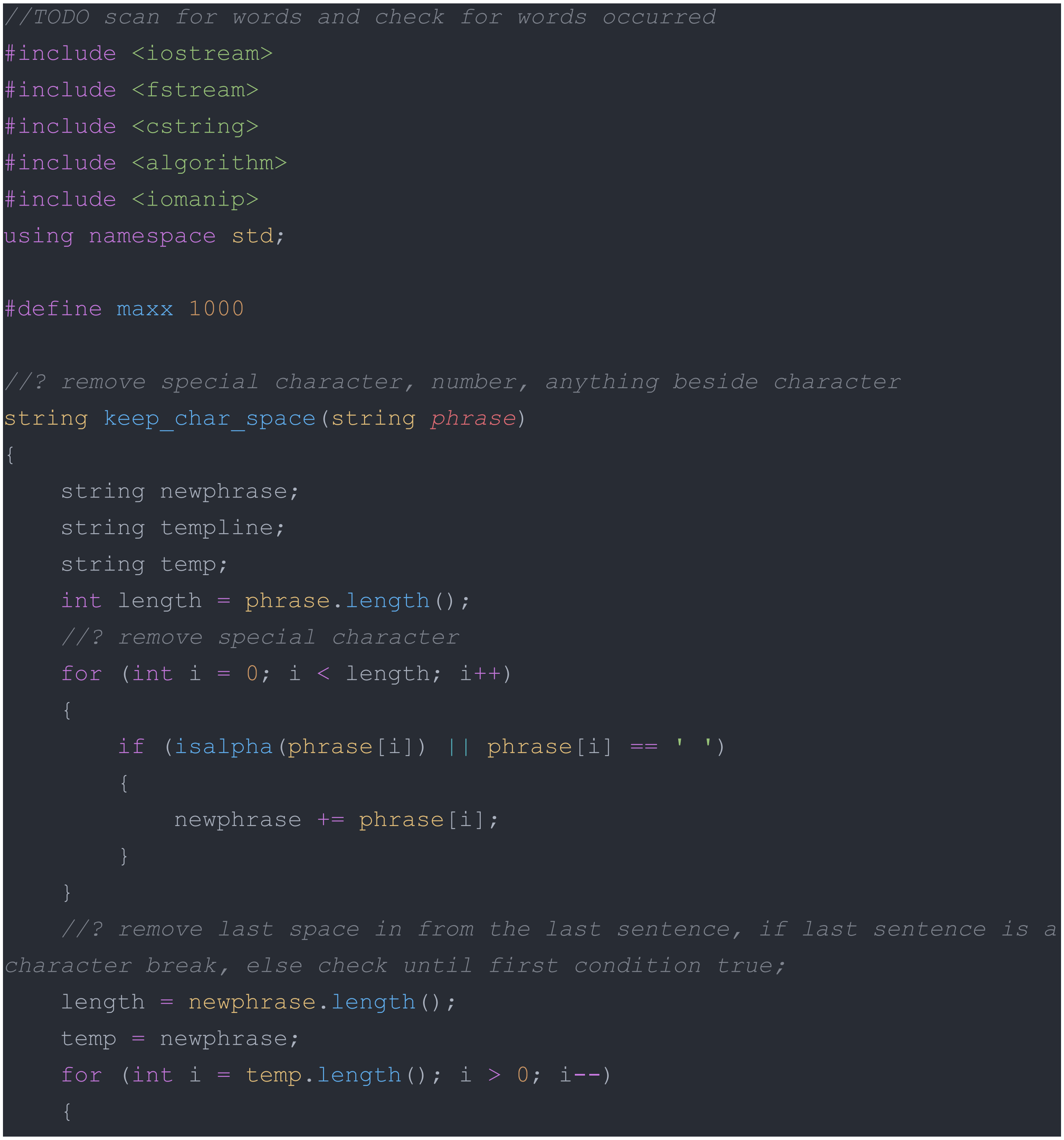
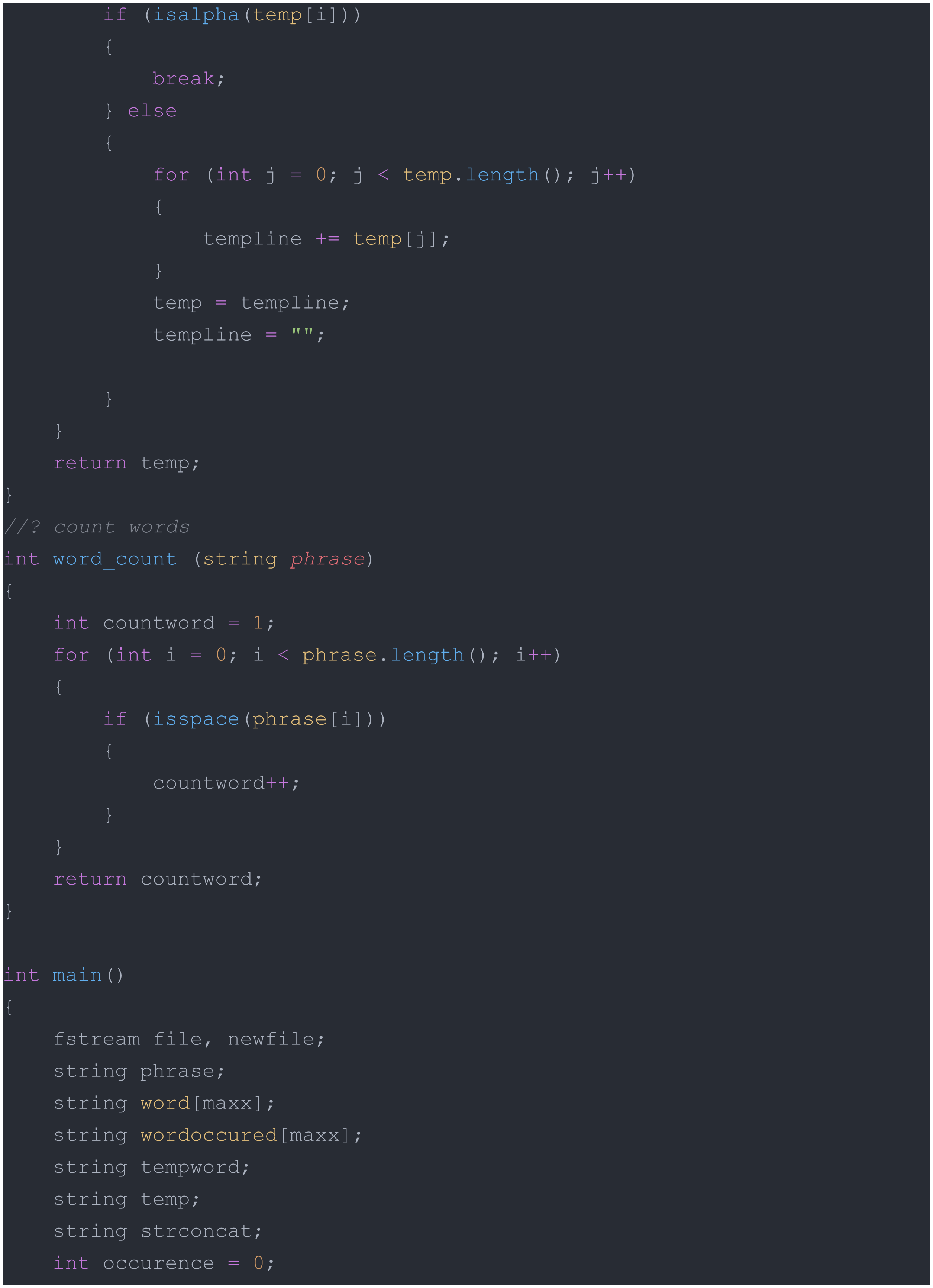
LAB 7

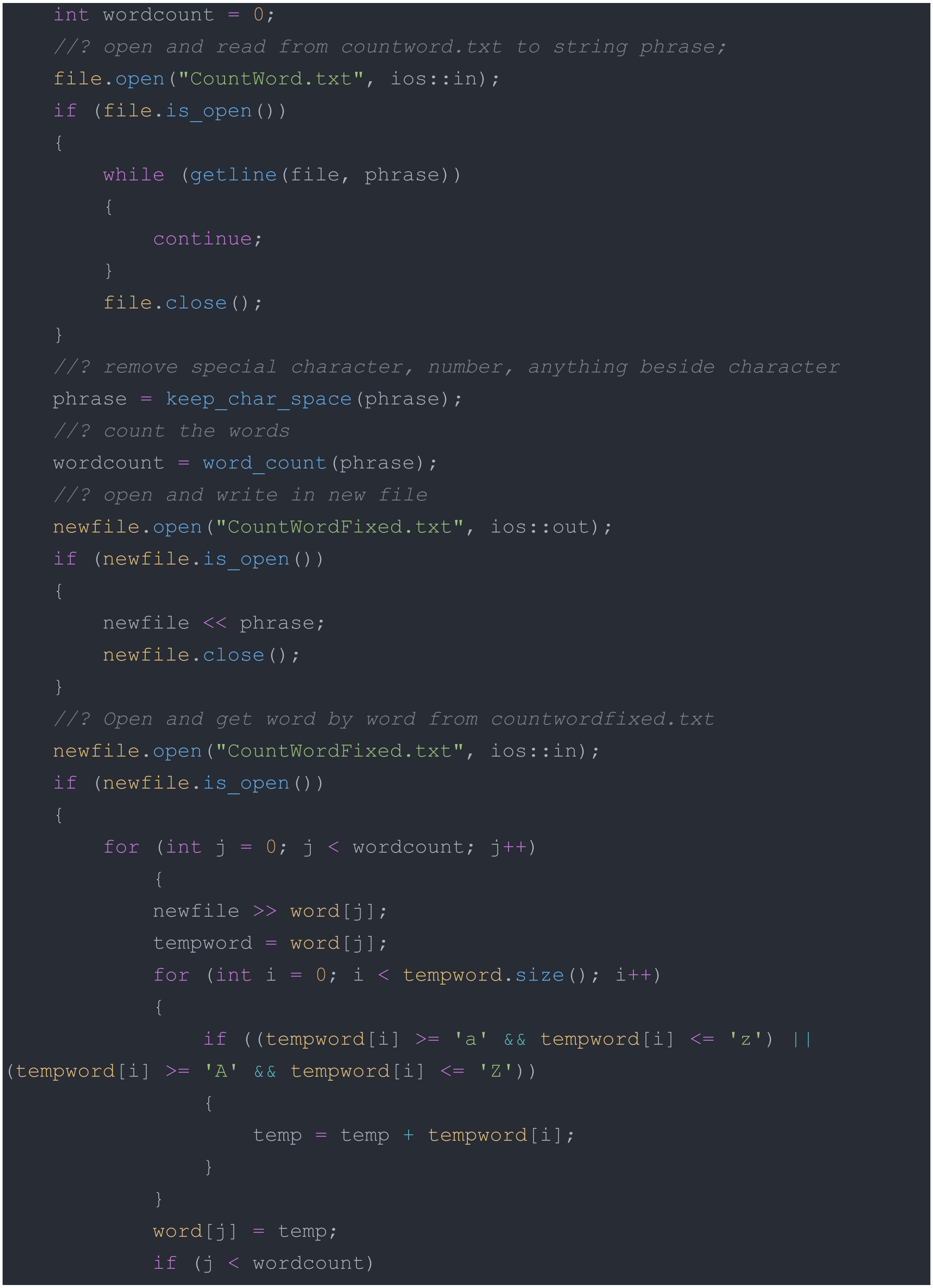
Seakmeng Hor
CS 126 Section 2



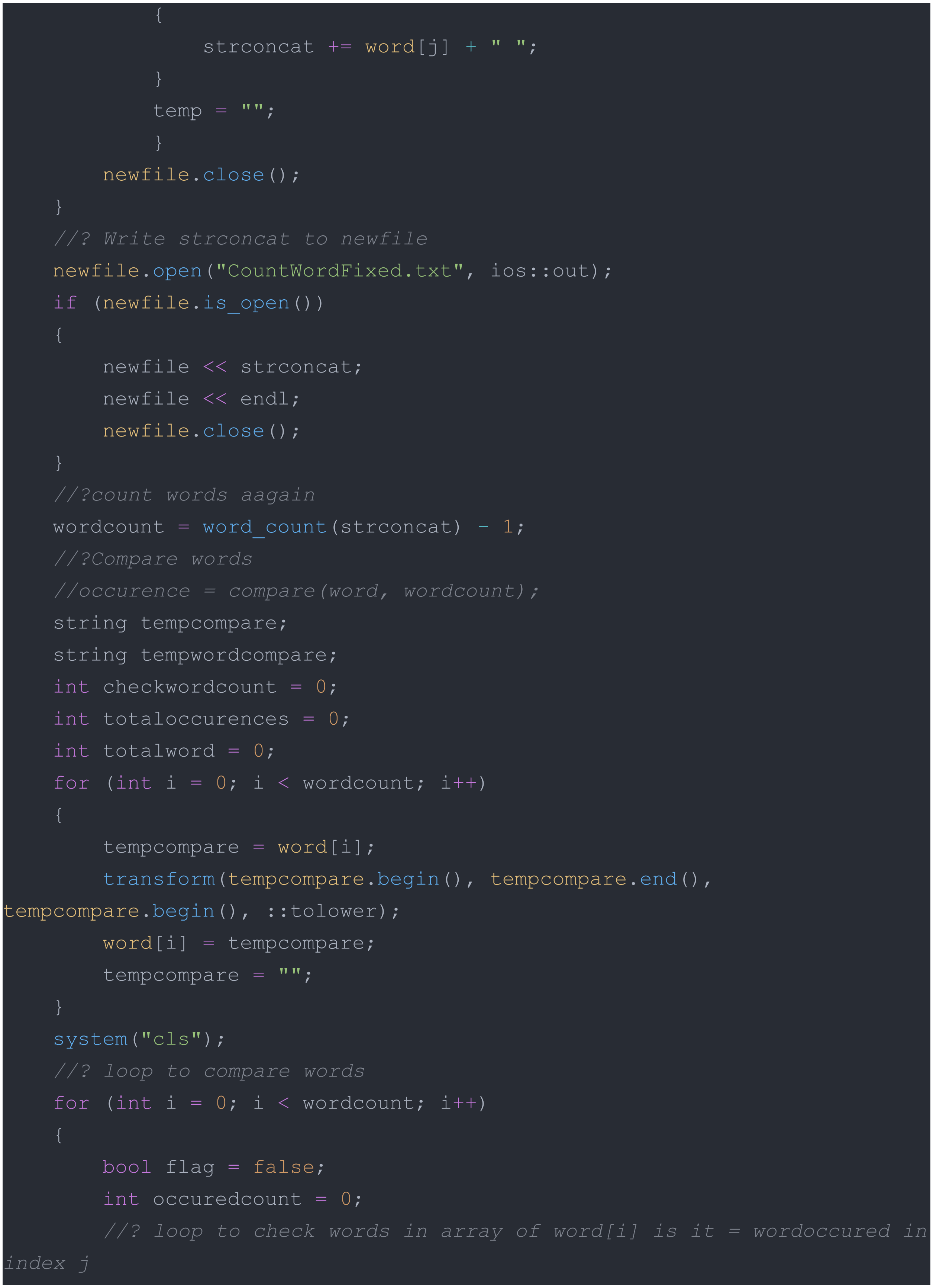
```
//TODO scan for words and check for words occurred
#include <iostream>
#include <fstream>
#include <cstring>
#include <algorithm>
#include <iomanip>
using namespace std;

#define maxx 1000

//? remove special character, number, anything beside character
string keep_char_space(string phrase)
{
    string newphrase;
    string templine;
    string temp;
    int length = phrase.length();
    //? remove special character
    for (int i = 0; i < length; i++)
    {
        if (isalpha(phrase[i]) || phrase[i] == ' ')
        {
            newphrase += phrase[i];
        }
    }
    //? remove last space in from the last sentence, if last sentence is a character break, else check until first condition true;
    length = newphrase.length();
    temp = newphrase;
    for (int i = temp.length(); i > 0; i--)
    {
```



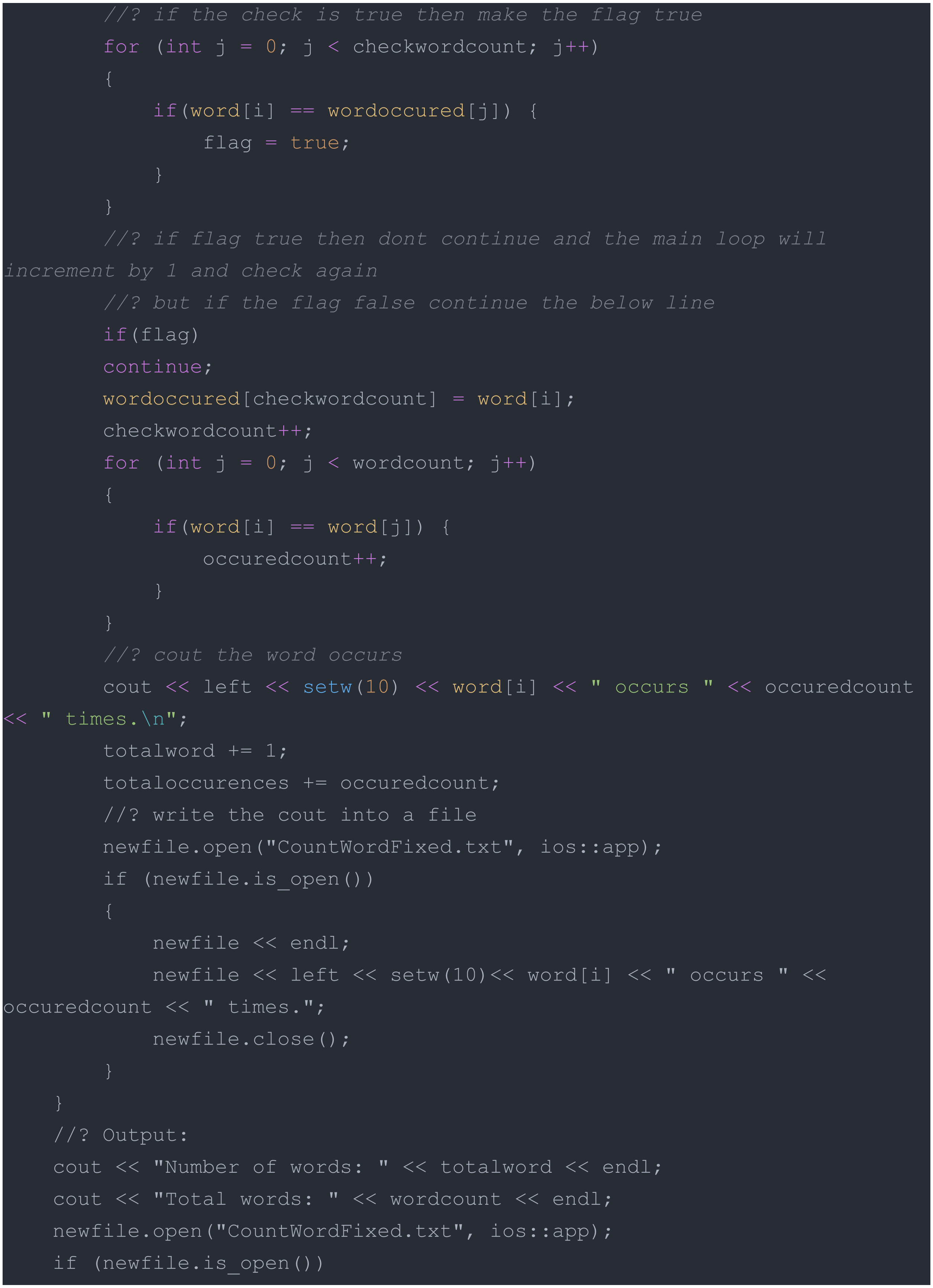
```
        if (isalpha(temp[i]))
        {
            break;
        } else
        {
            for (int j = 0; j < temp.length(); j++)
            {
                templine += temp[j];
            }
            temp = templine;
            templine = "";

        }
    }
    return temp;
}
//? count words
int word_count (string phrase)
{
    int countword = 1;
    for (int i = 0; i < phrase.length(); i++)
    {
        if (isspace(phrase[i]))
        {
            countword++;
        }
    }
    return countword;
}

int main()
{
    fstream file, newfile;
    string phrase;
    string word[maxx];
    string wordoccured[maxx];
    string tempword;
    string temp;
    string strconcat;
    int occurence = 0;
```



```
    int wordcount = 0;
    //? open and read from countword.txt to string phrase;
    file.open("CountWord.txt", ios::in);
    if (file.is_open())
    {
        while (getline(file, phrase))
        {
            continue;
        }
        file.close();
    }
    //? remove special character, number, anything beside character
    phrase = keep_char_space(phrase);
    //? count the words
    wordcount = word_count(phrase);
    //? open and write in new file
    newfile.open("CountWordFixed.txt", ios::out);
    if (newfile.is_open())
    {
        newfile << phrase;
        newfile.close();
    }
    //? Open and get word by word from countwordfixed.txt
    newfile.open("CountWordFixed.txt", ios::in);
    if (newfile.is_open())
    {
        for (int j = 0; j < wordcount; j++)
            {
            newfile >> word[j];
            tempword = word[j];
            for (int i = 0; i < tempword.size(); i++)
            {
                if ((tempword[i] >= 'a' && tempword[i] <= 'z') ||
(tempword[i] >= 'A' && tempword[i] <= 'Z'))
                {
                    temp = temp + tempword[i];
                }
            }
            word[j] = temp;
            if (j < wordcount)
```

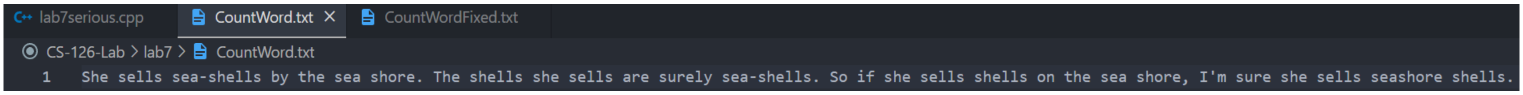


```
            {
                strconcat += word[j] + " ";
            }
            temp = "";
            }
        newfile.close();
    }
    //? Write strconcat to newfile
    newfile.open("CountWordFixed.txt", ios::out);
    if (newfile.is_open())
    {
        newfile << strconcat;
        newfile << endl;
        newfile.close();
    }
    //?count words aagain
    wordcount = word_count(strconcat) - 1;
    //?Compare words
    //occurence = compare(word, wordcount);
    string tempcompare;
    string tempwordcompare;
    int checkwordcount = 0;
    int totaloccurences = 0;
    int totalword = 0;
    for (int i = 0; i < wordcount; i++)
    {
        tempcompare = word[i];
        transform(tempcompare.begin(), tempcompare.end(),
tempcompare.begin(), ::tolower);
        word[i] = tempcompare;
        tempcompare = "";
    }
    system("cls");
    //? loop to compare words
    for (int i = 0; i < wordcount; i++)
    {
        bool flag = false;
        int occuredcount = 0;
        //? loop to check words in array of word[i] is it = wordoccured in
index j
```



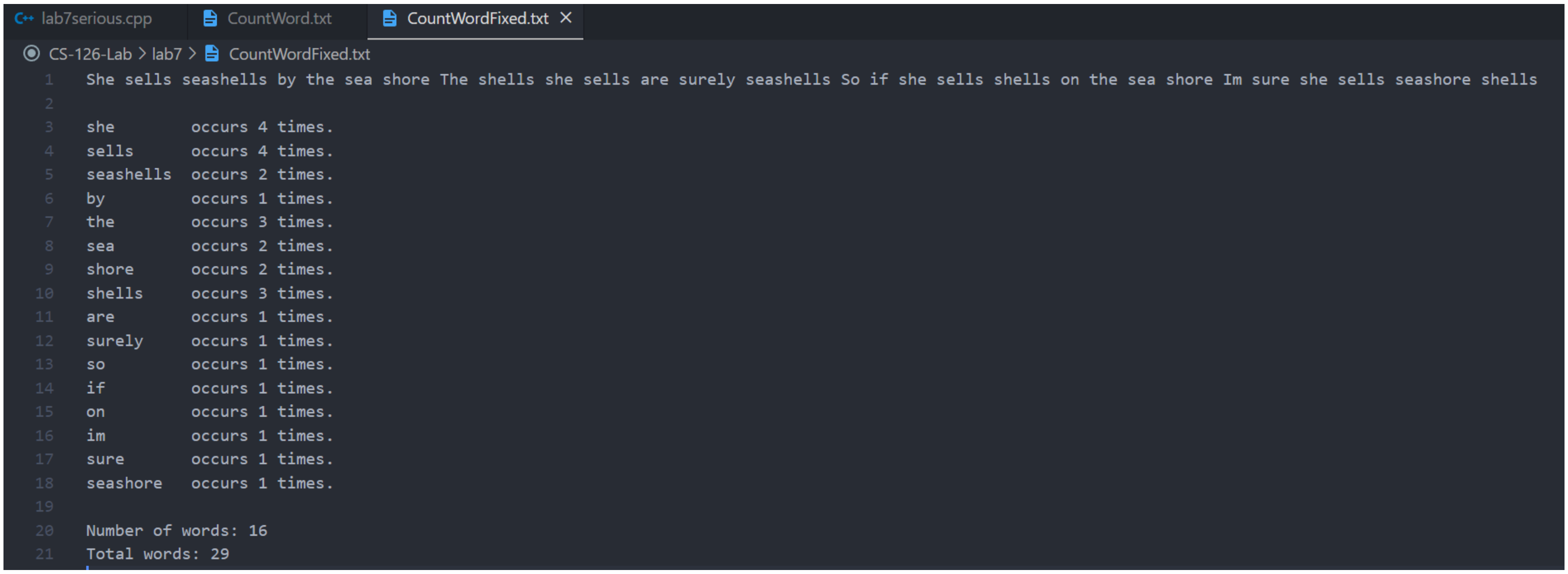
```
        //? if the check is true then make the flag true
        for (int j = 0; j < checkwordcount; j++)
        {
            if(word[i] == wordoccured[j]) {
                flag = true;
            }
        }
        //? if flag true then dont continue and the main loop will
increment by 1 and check again
        //? but if the flag false continue the below line
        if(flag)
        continue;
        wordoccured[checkwordcount] = word[i];
        checkwordcount++;
        for (int j = 0; j < wordcount; j++)
        {
            if(word[i] == word[j]) {
                occuredcount++;
            }
        }
        //? cout the word occurs
        cout << left << setw(10) << word[i] << " occurs " << occuredcount
<< " times.\n";
        totalword += 1;
        totaloccurences += occuredcount;
        //? write the cout into a file
        newfile.open("CountWordFixed.txt", ios::app);
        if (newfile.is_open())
        {
            newfile << endl;
            newfile << left << setw(10)<< word[i] << " occurs " <<
occuredcount << " times.";
            newfile.close();
        }
    }
    //? Output:
    cout << "Number of words: " << totalword << endl;
    cout << "Total words: " << wordcount << endl;
    newfile.open("CountWordFixed.txt", ios::app);
    if (newfile.is_open())
    {
        newfile << endl << endl;
        newfile << "Number of words: " << totalword << endl;
        newfile << "Total words: " << wordcount << endl;
        //newfile << "Number of words: " << totaloccurences;
        newfile.close();
    }
    return 0;
}
```

Original text



```
She sells sea-shells by the sea shore. The shells she sells are surely sea-shells. So if she sells shells on the sea shore, I'm sure she sells seashore shells.
```

After running the code



```
She sells seashells by the sea shore The shells she sells are surely seashells So if she sells shells on the sea shore Im sure she sells seashore shells

she        occurs 4 times.
sells      occurs 4 times.
seashells  occurs 2 times.
by         occurs 1 times.
the        occurs 3 times.
sea        occurs 2 times.
shore      occurs 2 times.
shells     occurs 3 times.
are        occurs 1 times.
surely     occurs 1 times.
so         occurs 1 times.
if         occurs 1 times.
on         occurs 1 times.
im         occurs 1 times.
sure       occurs 1 times.
seashore   occurs 1 times.

Number of words: 16
Total words: 29
```