AARYAN BAIRAGI

BE-47004

ISR LAB 1:

```
        import java.io.*;
import java.util.*;


public class Conflation {
   static List<String> stopWords = Arrays.asList("is", "am", "are", "the", "and", "a", "an", "in", "on",
"at", "of", "for", "to", "with", "by");


   public static void main(String[] args) throws IOException {
      Scanner sc = new Scanner(System.in);
      int ch;


      do {
         System.out.println("\n--- Conflation Algorithm Menu ---");
         System.out.println("1. Display the file");
         System.out.println("2. Remove Stop Words");
         System.out.println("3. Suffix Stripping (Stemming)");
         System.out.println("4. Count Frequency");
         System.out.println("5. Exit");
         System.out.print("Enter your choice: ");
         ch = sc.nextInt();


         switch (ch) {
            case 1:
               displayFile();
               break;
            case 2:
               removeStopWords();
               break;
            case 3:
               suffixStripping();
```

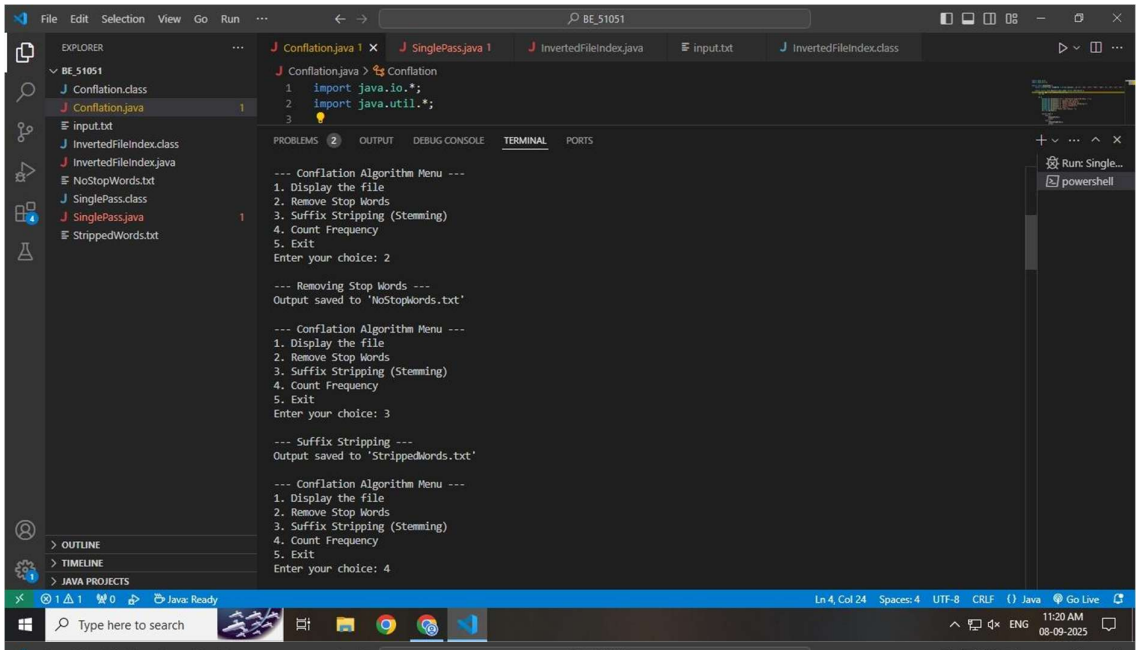
```
            break;
        case 4:
            countFrequency();
            break;
        case 5:
            System.out.println("Exiting program...");
            break;
        default:
            System.out.println("Invalid choice! Try again.");
      }

   } while (ch != 5);
}

// 1. Display file content
public static void displayFile() throws IOException {
   BufferedReader br = new BufferedReader(new FileReader("Input.txt"));
   String line;
   System.out.println("\n--- File Content ---");
   while ((line = br.readLine()) != null) {
      System.out.println(line);
   }
   br.close();
}

// 2. Remove stop words
public static void removeStopWords() throws IOException {
   BufferedReader br = new BufferedReader(new FileReader("Input.txt"));
   BufferedWriter bw = new BufferedWriter(new FileWriter("NoStopWords.txt"));
   String line;
   System.out.println("\n--- Removing Stop Words ---");
```

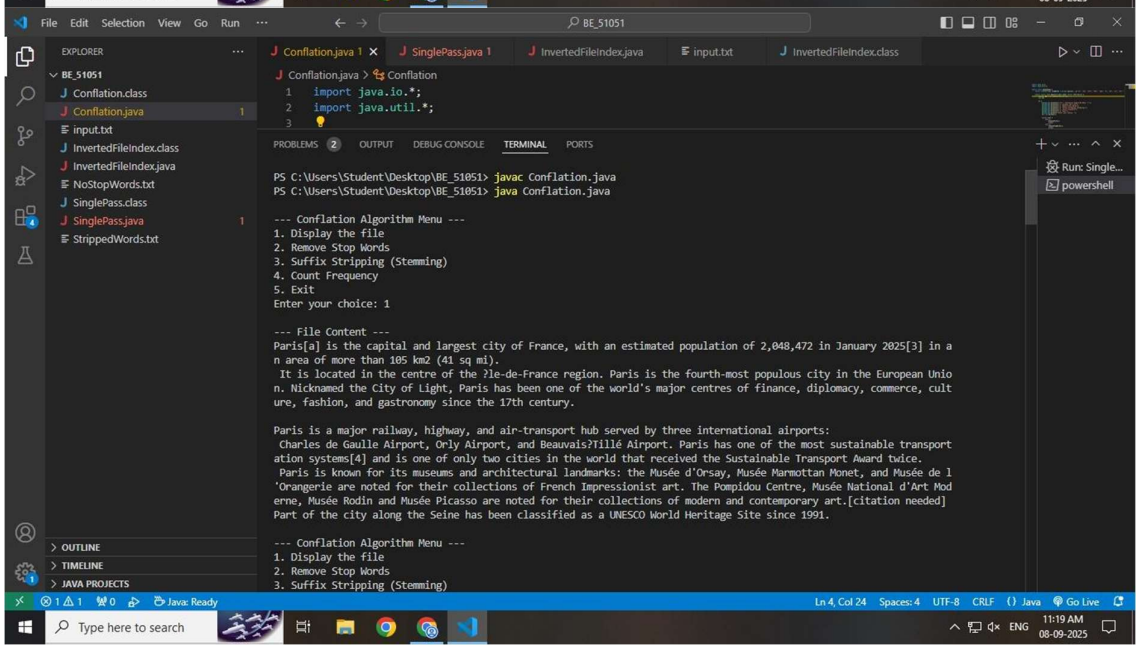
```
    while ((line = br.readLine()) != null) {
        String[] words = line.toLowerCase().split("\\W+");
        for (String word : words) {
            if (!stopWords.contains(word) && !word.trim().isEmpty()) {
                bw.write(word + " ");
            }
        }
        bw.newLine();
    }

    br.close();
    bw.close();
    System.out.println("Output saved to 'NoStopWords.txt'");
}

// 3. Suffix Stripping (very basic stemming)
public static void suffixStripping() throws IOException {
    BufferedReader br = new BufferedReader(new FileReader("NoStopWords.txt"));
    BufferedWriter bw = new BufferedWriter(new FileWriter("StrippedWords.txt"));
    String line;

    System.out.println("\n--- Suffix Stripping ---");
    while ((line = br.readLine()) != null) {
        String[] words = line.split("\\W+");
        for (String word : words) {
            String stemmed = word;
            if (word.endsWith("ing") && word.length() > 4)
                stemmed = word.substring(0, word.length() - 3);
            else if (word.endsWith("ed") && word.length() > 3)
                stemmed = word.substring(0, word.length() - 2);
```

```
            else if (word.endsWith("es") && word.length() > 3)
                stemmed = word.substring(0, word.length() - 2);
            else if (word.endsWith("s") && word.length() > 2)
                stemmed = word.substring(0, word.length() - 1);

            bw.write(stemmed + " ");
        }
        bw.newLine();
    }

    br.close();
    bw.close();
    System.out.println("Output saved to 'StrippedWords.txt'");
}

// 4. Word frequency count
public static void countFrequency() throws IOException {
    BufferedReader br = new BufferedReader(new FileReader("StrippedWords.txt"));
    HashMap<String, Integer> freqMap = new HashMap<>();
    String line;

    System.out.println("\n--- Word Frequency ---");
    while ((line = br.readLine()) != null) {
        String[] words = line.toLowerCase().split("\\W+");
        for (String word : words) {
            if (!word.trim().isEmpty()) {
                freqMap.put(word, freqMap.getOrDefault(word, 0) + 1);
            }
        }
    }
```

```
        for (Map.Entry<String, Integer> entry : freqMap.entrySet()) {
            System.out.println(entry.getKey() + " : " + entry.getValue());
        }

        br.close();
    }
}
```

Output:


File Edit Selection View Go Run
BE_51051
EXPLORER
BE_51051
Conflation.class
Conflation.java 1
input.txt
InvertedFileIndex.class
InvertedFileIndex.java
NoStopWords.txt
SinglePass.class
SinglePass.java 1
StrippedWords.txt
Conflation.java 1
SinglePass.java 1
InvertedFileIndex.java
input.txt
InvertedFileIndex.class
Conflation.java > Conflation
1 import java.io.*;
2 import java.util.*;
3
PROBLEMS 2
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
Run: Single...
powershell
--- Conflation Algorithm Menu ---
1. Display the file
2. Remove Stop Words
3. Suffix Stripping (Stemming)
4. Count Frequency
5. Exit
Enter your choice: 2
--- Removing Stop Words ---
Output saved to 'NoStopWords.txt'
--- Conflation Algorithm Menu ---
1. Display the file
2. Remove Stop Words
3. Suffix Stripping (Stemming)
4. Count Frequency
5. Exit
Enter your choice: 3
--- Suffix Stripping ---
Output saved to 'StrippedWords.txt'
--- Conflation Algorithm Menu ---
1. Display the file
2. Remove Stop Words
3. Suffix Stripping (Stemming)
4. Count Frequency
5. Exit
Enter your choice: 4
OUTLINE
TIMELINE
JAVA PROJECTS
Java: Ready
Ln 4, Col 24
Spaces: 4
UTF-8
CRLF
Java
Go Live
Type here to search
ENG
11:20 AM
08-09-2025


File Edit Selection View Go Run
BE_51051
EXPLORER
BE_51051
Conflation.class
Conflation.java 1
input.txt
InvertedFileIndex.class
InvertedFileIndex.java
NoStopWords.txt
SinglePass.class
SinglePass.java 1
StrippedWords.txt
Conflation.java 1
SinglePass.java 1
InvertedFileIndex.java
input.txt
InvertedFileIndex.class
Conflation.java > Conflation
1 import java.io.*;
2 import java.util.*;
3
PROBLEMS 2
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
Run: Single...
powershell
PS C:\Users\Student\Desktop\BE_51051> javac Conflation.java
PS C:\Users\Student\Desktop\BE_51051> java Conflation.java
--- Conflation Algorithm Menu ---
1. Display the file
2. Remove Stop Words
3. Suffix Stripping (Stemming)
4. Count Frequency
5. Exit
Enter your choice: 1
--- File Content ---
Paris[a] is the capital and largest city of France, with an estimated population of 2,048,472 in January 2025[3] in a
n area of more than 105 km2 (41 sq mi).
It is located in the centre of the ?le-de-France region. Paris is the fourth-most populous city in the European Unio
n. Nicknamed the City of Light, Paris has been one of the world's major centres of finance, diplomacy, commerce, cult
ure, fashion, and gastronomy since the 17th century.
Paris is a major railway, highway, and air-transport hub served by three international airports:
Charles de Gaulle Airport, Orly Airport, and Beauvais?Tillé Airport. Paris has one of the most sustainable transport
ation systems[4] and is one of only two cities in the world that received the Sustainable Transport Award twice.
Paris is known for its museums and architectural landmarks: the Musée d'Orsay, Musée Marmottan Monet, and Musée de l
'Orangerie are noted for their collections of French Impressionist art. The Pompidou Centre, Musée National d'Art Mod
erne, Musée Rodin and Musée Picasso are noted for their collections of modern and contemporary art.[citation needed]
Part of the city along the Seine has been classified as a UNESCO World Heritage Site since 1991.
--- Conflation Algorithm Menu ---
1. Display the file
2. Remove Stop Words
3. Suffix Stripping (Stemming)
OUTLINE
TIMELINE
JAVA PROJECTS
Java: Ready
Ln 4, Col 24
Spaces: 4
UTF-8
CRLF
Java
Go Live
Type here to search
ENG
11:19 AM
08-09-2025