Check Plagiarism

ANALYSIS OF ALGORITHMS

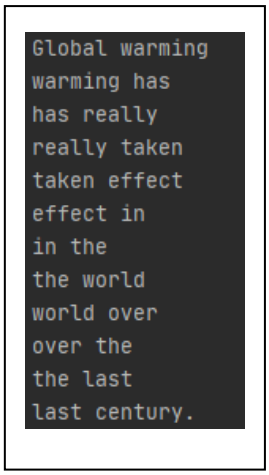
PROJECT REPORT

BERKE AYAN | 1729662
KÜBRA BAL | 1729922
MELİHA BUSE İRİCAN | 1731212

-IMPLEMENTED IN JAVA LANGUAGE-

METHODS

In our program we create pairs of subsequent words of the main source text file as shown in the example screenshot, and we compare all of these pairs with each sentences of the input files.Therefore after the comparison we define a similarity rate of texts. We benefit from the Boyer Moore-Horspool Algorithm to check similarity between texts.We used array,arraylists,linked hash map and a hashtable for storing.



-Horspool Algorithm Shift Table

For Horspool algorithm we created Shift Table.ShiftMyTable method takes pattern and text file as a paramater, according to these paramaters shift table will be created .For store the values of shift table we decided to use hash table. The type of ShiftMyTable method is Hashtable and it returns the table itself.A boolean type flag is also created which checks the last letter of the pattern . A for loop has been created so that the file can be viewed from beginning to end.If the value of the letter in the text file has not yet been added to the pattern table; a variable named 's' is created in the string type, and this variable takes the value of character in the i-th position of the text file. After that, the possibilities are in order;

1)If the letter in text file does not exist in the pattern

For that letter we put the corresponding value into the shift table For that value since the letter does not exist in the pattern the corresponding value will be length of the pattern.For this we used length() fuction to take length of the pattern and we put the corresponding character and the value into shift table.For inserting these two values we used put() function.

2)If the letter in text file exist in the pattern

If the letter in the text file is in pattern, it is followed by the for loop, starting from the letter that is before the last letter to the first letter of the word.If the letter is found,the flag value becomes true,since we don't check the last letter in the for loop we use the flag to check that letter .The put() function is used to add the letter and the shift value into the shift table .The corresponding shift value is : length of the pattern - the position in which it is located-i .After we calculete these value we put it with letter into the shift table. The loop is broken with the break keyword so that it does not run indefinitely. If flag value is false, that is, if the letter in the file is the last letter in the pattern, the length of the pattern is added to the shift table('table'). In addition, the flag created with the false value is again left with the false value so that other words can be checked.Shift table is returned so that we can access it in the HorspoolSearch method.

-Horspool Algorithm Search

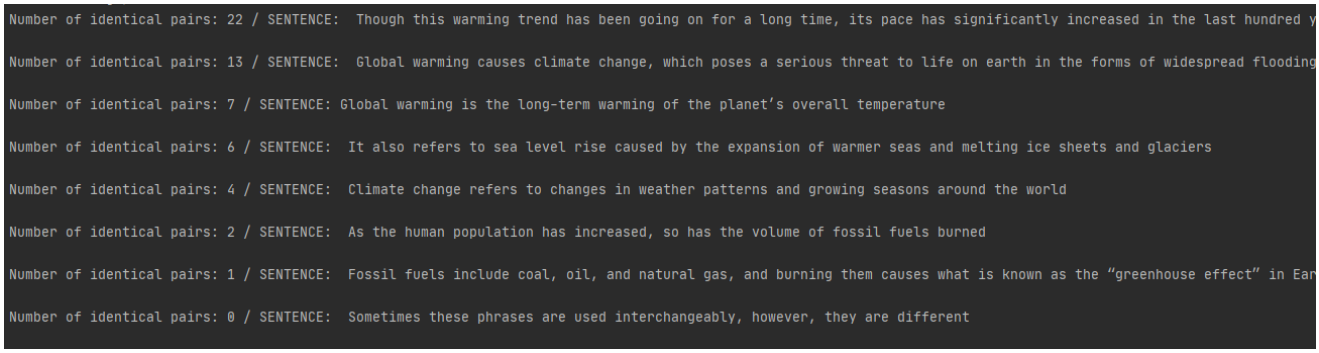
This method allows the created pattern table to be shifted according to the values placed. An i value of integer type was created to keep the length of the pattern.A while loop has been created that will continue as long as the length of the pattern is less than or equal to the length of the file.A k value of integer type was created to see how many letters i have in common in this loop.So in while condition we look at if the letters are matched and the k value is less than pattern length.If these conditions satisfies which means there is a match and the value k is increased.So if k value is equal to length of the pattern we can say that we found the pattern and we return the position of the match.If there are no matches we shift the index i according to shift table. A value of -1 is returned if there is no letter match.

-Array Lists

We store main source text's pairs and texts which will be checked in array lists.We used array lists instead of arrays since arrays are fixed size but array lists are more flexible than arrays.Array lists are much better advantages over arrays such as array lists are resizable and has many methods for modifications.

-Linked Hash Map

Firstly, we store sentences of texts which is checked and their similar pair values in hashmap then in order to sort sentences in descending order,we pass these to the linked hash map thanks to linked hash map's comparator.reverseorder() method we successfully sort the sentences according to their similar pair values and get a sequence of sentences as shown in above.



- Hash Table

In the ShiftMyTable method in the Search Class, a hashtable was created to hold both the pattern and the text file together.A shift table is an object named 'table' of hash type.The created 'Table' object was also used in the HorspoolSearch method.In this way, shifting was done by looking at the table created in the hash type.In addition, the object 'a' of the hastable type was used to hold the pairs created in the main class.ShiftMyTable and HorspoolSearch with the help of this object ,its methods have been called.

LIBRARIES

```
import java.io.BufferedReader;
```

-Its used for read texts and buffers characters.It provides an efficient reading.

```
import java.io.FileReader;
```

- FileReader is class that reads character files and contains necessary info about the file, a BufferReader takes FileReader as an input with the FileReader ,the same BufferReader class can read multiple files.

```
import java.io.IOException;
```

- Library that is written so that the error occurs when there is an error while reading, writing, and searching file operations.

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

-It contains basic useful collection types and their basic utility classes.We implemented for using arraylists,hashlinkedlist.

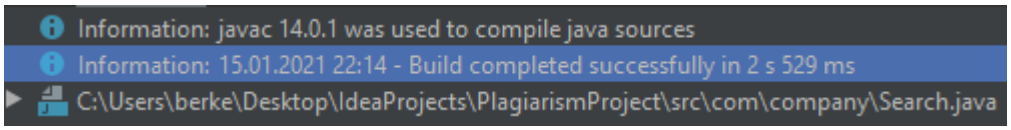
```
import java.util.Comparator;
```

- It is a member of the java collection framework interface ,where we implemented it for using sorted() function to sort sentences in descending order in the LinkedHashMap.

```
import java.util.Hashtable;
```

-In order to use hash table we implement its library.

AVERAGE SPEED OF THE PROGRAM



MAIN TEXT

Global warming has really taken effect in the world over the last century. It is the unusually rapid increase in the Earth's average surface temperature over the past century primarily due to the greenhouse gases released as people burn fossil fuels. Global warming is due to the enhancing greenhouse gases emission and build-up in the Earth's environment. The gases that have an influence on the atmosphere are water vapor, carbon dioxide, dinitrogen-oxide, and methane. Almost 30 percent of incoming sunlight is reflected back into space by bright surfaces like clouds and ice. In the other 70 percent, most is absorbed by the land and ocean, and the rest is absorbed by the atmosphere. The absorbed solar energy heats our planet. This absorption and radiation of heat by the atmosphere is beneficial for life on Earth. Today, the atmosphere contains more greenhouse gas molecules, so more of the infrared energy emitted by the surface ends up being absorbed by the atmosphere. By increasing the concentration of greenhouse gases, we are making Earth's atmosphere a more efficient greenhouse. Climate has cooled and warmed throughout the Earth history for various reasons. Rapid warming like we see today is unusual in the history of our planet. Some of the factors that have an effect on climate, like volcanic eruptions and changes in the amount of solar energy, are natural. Climate can change if there is a change in the amount of solar energy that gets to the Earth. Volcano eruptions can really affect climate, because when it erupts it spews out more than just lava and ash. Volcanoes release tiny particles made of sulfur dioxide into the atmosphere. These particles get into the stratosphere and reflect solar radiation back out to space. Snow and ice also have a great effect on climate. When snow and ice melts Earth's climate warms, less energy is reflected and this causes even more warming. There are many different ways that plants, animals, and other life on our planet can affect climate. Some can produce greenhouse gases that trap heat and aid global warming through the greenhouse effect. Carbon dioxide is taken out of the atmosphere by plants as they make their food by photosynthesis. During the night, plants release some carbon dioxide back into the atmosphere. Methane is made while farm animals, such as cattle and sheep digest their food. Cars and trucks can effect climate by releasing carbon dioxide when fossil fuels are burned to power them. When wildfires occur, carbon dioxide is released into the atmosphere. However, if a forest of similar size grows again, about the same amount of carbon that was added to the atmosphere during the fire will be removed. Some effects that scientists have predicted in the past would result when global change was occurring: loss of sea ice, accelerated sea level rise, and more intense heat waves. Scientists have confidence that global temperatures will continue to rise for decades to come, largely due to greenhouse gases produced by human activities. The Intergovernmental Panel on Climate Change (IPCC) stated that the extent climate change effects on individual regions will vary over time and with ability of different societal and environmental systems mitigate or adapt to change (The Intergovernmental Panel on Climate Change). This has been the warmest decade since 1880. According to the National Oceanic and Atmospheric Administration, 2010 and 2005 has been the warmest years on record. The earth could warm by an additional 7 degrees Fahrenheit during the 21st century if we fail to reduce emissions from burning fossil fuels (The National Oceanic and Atmospheric Administration). The rising of temperature will have great effects on the earth's climate patterns and on all living things. Industrial activities that our modern civilization depends upon have raised atmospheric carbon dioxide from 280 parts per million to 379 parts per million in the last 150 years (The Intergovernmental Panel on Climate Change).

EXEMPLARY OUTPUT FOR 'CHECK1.TXT'

```
MOST SIMILAR 1th SENTENCE
 Though this warming trend has been going on for a long time, its pace has significantly increased in the last hundred years due to the burning of fossil fuels

MOST SIMILAR 2th SENTENCE
 Global warming causes climate change, which poses a serious threat to life on earth in the forms of widespread flooding and extreme weather

MOST SIMILAR 3th SENTENCE
Global warming is the long-term warming of the planet's overall temperature

MOST SIMILAR 4th SENTENCE
 It also refers to sea level rise caused by the expansion of warmer seas and melting ice sheets and glaciers

MOST SIMILAR 5th SENTENCE
 Climate change refers to changes in weather patterns and growing seasons around the world

Similarity rate between texts = %8
```

EXEMPLARY OUTPUT FOR 'CHECK2.TXT'

```
MOST SIMILAR 1th SENTENCE
 For our sake and for existence of all life on our heavenly planet earth, it is the need of the hour that we make sincere efforts and attempts to contain the mons

MOST SIMILAR 2th SENTENCE
 But increase in greenhouse gases has led to more absorbation of infrared radiation which in turn has led to the increase of earth's surface temperature

MOST SIMILAR 3th SENTENCE
 These human activities and trends have led to increased emission of greenhouse gases like carbon dioxide, methane, nitrous oxide, etc

MOST SIMILAR 4th SENTENCE
 It is a wake-up call for us to amend our ways of living to limit the use of products that contribute to the increase in global warming

MOST SIMILAR 5th SENTENCE
 Global warming deplets the ozone layer and might lead us to the doom's day

Similarity rate between texts = %10
```

EXEMPLARY OUTPUT FOR 'CHECK3.TXT'

```
MOST SIMILAR 1th SENTENCE
 They also said the odds of similar droughts happening in the future had roughly doubled over the past century

MOST SIMILAR 2th SENTENCE
Scientists agree that the earth's rising temperatures are fueling longer and hotter heat waves, more frequent droughts, heavier rainfall, and more powerful hurri

MOST SIMILAR 3th SENTENCE
 In fact, scientists have found that the frequency of North Atlantic hurricanes has increased since the early 1980s, as well as the number of storms that reach c

MOST SIMILAR 4th SENTENCE
 The earth's ocean temperatures are getting warmer, too—which means that tropical storms can pick up more energy

MOST SIMILAR 5th SENTENCE
 history-struck New Orleans; the second-costliest, Hurricane Sandy, hit the East Coast in 2012

Similarity rate between texts = %3
```

WORST CASE BIG-OH COMPLEXITY

```
for (int i=0;i<checklist.size();i++){                                    //FO
    for (int j=0;j<pairslist.size();j++) {                               //S
    Search shift=new Search();
    Hashtable a=shift.ShiftMyTable(pairslist.get(j),checklist.get(i));
    int match =shift.HorspoolSearch(pairslist.get(j),checklist.get(i),a);
```

Basic operation of the program is in the for loops of the above screenshot.

Worst case of Horspool Algorithm is O(mn).

For our input file, we can consider as it can have 'n' characters.

For our main file, we can consider as it can have 'm' characters in it.

We can say average length of the each words is 'k'.

Each pair has 2 words, so we can say each of the pair size is '2k'.

For one pair which is '2k', we search these pairs in the input file which has 'n' characters.This makes '2nk'.This process is repeat for every pairs.We have to multiply '2nk' with number of pairs.

In order to find number of pairs, we should find number of the words in source file.

Number of the words in source file = m/k

Number of pairs = (m/k) – 1

BigOH= (2nk) * ((m/k)-1)

BigOH = O(mn)