

MAC0323 Algoritmos e Estruturas de Dados II

Eduardo Pinheiro(8936798)

12-03-2016

3.1.52 Word Frequencies:

Write a program (or several programs and use piping) that reads in a text file and prints out a list of the words in decreasing order of frequency. Consider breaking it up into 5 pieces and use piping: read in text and print the words one per line in lowercase, sort to bring identical words together, remove duplicates and print count, sort by count.

Considerações:

- O programa uso o método *StdIn.readLine()* para leitura do texto, portanto o texto deve estar em apenas uma linha, isso ocorre pois os métodos *StdIn.readAll()* *StdIn.readAllLines()* *StdIn.readAllStrings()* por algum motivo não terminam sua execução, assim o programa não prosegue a execução.
- Como pedido no enunciado++ o programa é compilado por executado por(respectivamente):

```
$ javac -algs4 WordFrequencies.java  
$ java -algs4 WordFrequencies
```

- Também há junto ao código um *Makefile* que em minha opinião facilita esse processo, para compilar e executar desta outra forma(respectivamente):

```
$ make  
$ make run
```

Word Frequencies implementação:

Para resolver o problema criei um tipo de dados chamado Word, que possui os seguintes atributos:

1. *String Word*
2. *int frequency*

*Cada objeto possui uma palavra e sua frequência no texto.

A classe Word tem a seguinte declaração:

```
public class Word implements Comparable < Word >
```

A implementação da interface Comparable, é utilizada para sobrescrever o método:

```
public int compareTo(Word o)
```

Isso é utilizado pois o método de ordenação: *Collections.sort(ArrayList)*; utiliza *compareTo* para fazer sua ordenação, e este método de ordenação tem complexidade de tempo de execução $\theta(N \log N)$ garantida[1].

API de Word:

- *public Word(String word)*
- *public String getWord()*
- *public void setWord(String word)*
- *public int getFrequency()*
- *public void setFrequency(int frequency)*
- *public String toString()*
- *public int compareTo(Word o)*

Word Frequencies Teste:

Para a seguinte entrada:

Você já se perguntou o que é, de fato, um texto? Geralmente, entendemos o texto como um conjunto de frases, ou seja, algo que foi feito para ser lido. Mas a definição de texto não é tão simples quanto parece.

Temos a seguinte saída:

```
de 3
texto 3
o 2
que 2
é 2
um 2
você 1
já 1
se 1
perguntou 1
fato 1
geralmente 1
entendemos 1
como 1
conjunto 1
frases 1
ou 1
seja 1
algo 1
foi 1
feito 1
para 1
ser 1
lido 1
mas 1
a 1
definição 1
não 1
tão 1
simples 1
quanto 1
parece 1
```

*Funciona com palavras com acento.

Referências:

[1]<http://docs.oracle.com/javase/6/docs/api/java/util/Collections.html#sort%28java.util.List%29>