

Pontifícia Universidade Católica do Rio Grande do Sul

Escola Politécnica

Algoritmos e Estruturas de Dados I

Professor(a): Leonardo Heredia

Trabalho 1 - Contador de Palavras

Participantes: Guilherme G. Santos

O Trabalho 1 da disciplina Algoritmos e Estruturas de Dados I (ALEST1) consistiu no desenvolvimento de um programa em linguagem Java para leitura e análise de arquivos de texto, com o objetivo de contabilizar o total de palavras, identificar as 20 palavras mais frequentes e listar os palíndromos encontrados. A implementação foi realizada de forma modular, com foco na organização e separação de responsabilidades entre os componentes do sistema. O código-fonte está disponível no GitHub, através do seguinte repositório: github.com/Quidoguim/ContadorDePalavras.

1. Implementação

O programa inicia com a leitura do arquivo de texto, seguida pela tokenização do conteúdo, etapa em que o texto é dividido em palavras com a remoção de pontuações e a conversão para letras minúsculas. Em seguida, cada palavra é armazenada em uma estrutura de lista - estática ou dinâmica - e sua frequência é incrementada sempre que ocorre uma repetição. Após o processamento, o programa realiza a identificação dos palíndromos, comparando os caracteres de cada palavra com sua versão invertida, e então ordena as palavras por frequência, da mais para a menos recorrente. As estruturas de dados foram implementadas manualmente, sem o uso de bibliotecas externas, permitindo a aplicação prática dos conceitos estudados em sala de aula. A distinção entre listas com alocação estática e dinâmica é refletida diretamente nas classes utilizadas, evidenciando suas características e limitações no contexto da aplicação.

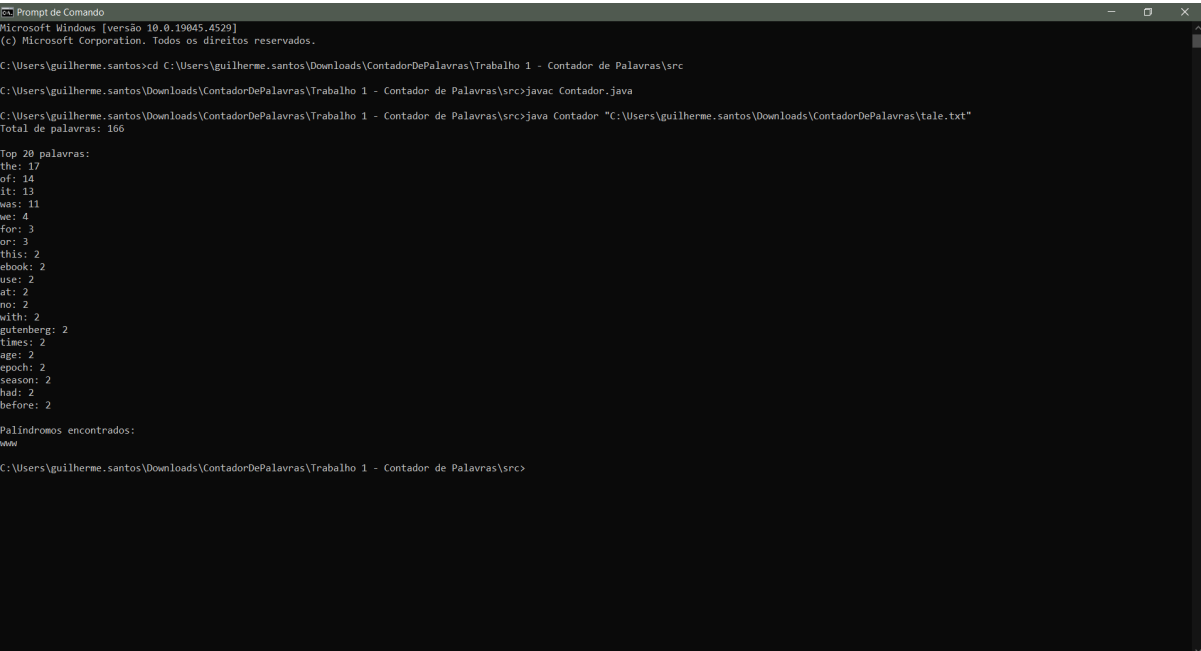
Por fim, ao ser executado, o programa exibe no terminal:

- A quantidade total de palavras processadas;
- As 20 palavras mais frequentes, acompanhadas de suas contagens;
- A lista completa de palavras identificadas como palíndromos.

2. Execução

A imagem mostra a execução do programa Contador de Palavras no Prompt de Comando do Windows. Primeiro, o usuário acessa a pasta do projeto e compila o arquivo Contador.java. Em seguida, executa o programa informando o caminho do arquivo de texto tale.txt como entrada. O programa processa o conteúdo do arquivo, contabiliza as palavras e identifica os palíndromos. Como resultado, são exibidos no terminal: o total de palavras encontradas (166), as 20 palavras mais frequentes com suas respectivas quantidades e a lista de palíndromos identificados, que nesse caso contém apenas a palavra "www".

A execução confirma que o programa funciona corretamente e atende aos objetivos propostos no Trabalho 1.



```
Microsoft Windows [versão 10.0.19045.4529]
(c) Microsoft Corporation. Todos os direitos reservados.

C:\Users\guilherme.santos>cd C:\Users\guilherme.santos\Downloads\ContadorDePalavras\Trabalho 1 - Contador de Palavras\src
C:\Users\guilherme.santos\Downloads\ContadorDePalavras\Trabalho 1 - Contador de Palavras\src>javac Contador.java
C:\Users\guilherme.santos\Downloads\ContadorDePalavras\Trabalho 1 - Contador de Palavras\src>java Contador "C:\Users\guilherme.santos\Downloads\ContadorDePalavras\tale.txt"
Total de palavras: 166

Top 20 palavras:
the: 17
of: 14
it: 13
was: 11
we: 4
for: 3
pr: 3
this: 2
ebook: 2
use: 2
is: 2
no: 2
with: 2
gutenberg: 2
times: 2
age: 2
epoch: 2
season: 2
had: 2
before: 2

Palíndromos encontrados:
www

C:\Users\guilherme.santos\Downloads\ContadorDePalavras\Trabalho 1 - Contador de Palavras\src>
```

A imagem mostra uma nova execução do programa Contador de Palavras, agora com o arquivo dracula.txt como entrada. O terminal apresenta os resultados processados pelo programa, confirmando que o funcionamento segue conforme esperado. O programa contabilizou um total de 124.487 palavras. Em seguida, listou as 20 palavras mais frequentes, entre elas: "the", "and", "i", "to", "of", "that", e "you", com suas respectivas quantidades. Após a contagem, o programa exibiu a lista de palíndromos identificados no texto, com destaque para palavras como "deed", "madam", "noon", "level", "pop", "eye", "wow", "did", "aha" e "xxx".

Essa saída comprova que o programa foi capaz de lidar com um arquivo extenso, mantendo o desempenho e retornando corretamente os dados solicitados: total de palavras, palavras mais frequentes e os palíndromos encontrados.

```
C:\Users\guilherme.santos\Downloads\ContadorDePalavras\Trabalho 1 - Contador de Palavras\src>java Contador "C:\Users\guilherme.santos\Downloads\ContadorDePalavras\dracula.txt"
Total de palavras: 124487

Top 20 palavras:
the: 8164
and: 5974
i: 4844
to: 4746
of: 3754
that: 3615
me: 2583
in: 2571
that: 2503
it: 2198
was: 1983
as: 1692
for: 1564
we: 1501
is: 1528
you: 1489
his: 1472
me: 1457
not: 1429
with: 1334

Palíndromos encontrados:
deed
noon
ere
ere
deed
ere
madam
ere
ere
noon
ii
ii
ere
ere
pop
ere
ere
aha
level
eye
eye
wow
did
ere
xx
ix
iii
ii
wow

C:\Users\guilherme.santos\Downloads\ContadorDePalavras\Trabalho 1 - Contador de Palavras\src>
```