

**Relatório de Projeto CheckFile**

|  |  |
| --- | --- |
| Gabriel Madeira Vieira, nº2200661 |  |
|  |  |
| Diogo dos Anjos Barbeiro, nºNUMERO |  |

SUBSTITUIR IMAGEM

“Gabriel Madeira Vieira (2200661) e Diogo dos Anjos Barbeiro (NUMERO) declaram sob compromisso de honra que o presente trabalho (código, relatórios e afins) foi integralmente realizado por nós, sendo que as contribuições externas se encontram claramente e inequivocamente identificadas no próprio código. Mais se declara que os estudantes acima identificados não disponibilizaram o código ou partes dele a terceiros”.

Leiria, Outubro de 2021

# Funcionalidades

|  |  |
| --- | --- |
| Funcionalidade | Estado |
| Modo “c” | Totalmente operacional |
| Modo “C” | Totalmente operacional |
| Modo “p” | Totalmente operacional |
| Modo “P” | Totalmente operacional |
| Modo “t” | Totalmente operacional |
| Modo “T” | Totalmente operacional |

# Metodologia de contagem

Para a contagem de caracteres, foram considerados números como palavras e palavras com hífen/ponto foram consideradas como uma só palavra. Por exemplo, “sentar-se” e “ficheiro.txt” é contado como uma palavra. Inicialmente, concatenaram-se todas as palavras de um determinado ficheiro com o comando *cat*.

De seguida, foram utilizados os comandos *sed* e *tr* encadeados para remover pontuações e caracteres especiais. Primeiramente, os pontos e os hífens foram substituídos por outros caracteres (“«” e “»” por ordem) pois o comando *tr -d [:punct:]* não os remove. Depois de removidas as pontuações voltou-se a trocar os pontos e hífens com o comando *tr*. Por fim, filtram-se as palavras que correspondem aos tipos “sentar-se”, “ficheiro.txt” e “palavra” obtendo-se uma lista limpa.

Exemplo de resultado:

* Sentar-se – uma palavra
* Sentar- se – duas palavras
* Ficheiro.txt – uma palavra
* Ajuda. – uma palavra

Excerto de código abaixo:

local punct="sed -r 's/(\-)/«/g' | sed -r 's/(\.)/»/g' | tr -d \"[:punct:]\""

punct="$punct | sed -r 's/(«)/-/g' | sed -r 's/(»)/./g'"

punct="$punct | sed -r 's/(^(\.)|(\.)$|(\ \.)|(\ \. \ )|(\.\ ))/\n/g'"

punct="$punct | sed -r 's/((\ )|(\ \-\ )|(\ \-)|(\-\ )|^(\-)|(\-)$)/\n/g'"

punct="$punct | tr -d \"'«»•–\" | sed -r 's/(^(\.)|(\.)$)/ /g'"

O próximo comando é modular, ou seja, este é diferente de acordo com o resultado desejado. É usado o comando *sort –ignore-case* para ordenar por ordem alfabética ignorando as maiúsculas e minúsculas em ambos os casos. O comando *grep -w -v -i -f* é apenas executado para a remoção de palavras irrelevantes caso a letra do modo introduzido seja maiúscula (C, T e P).

O seguinte comando foi o *awk NF*. Este comando foi implementado com o objetivo de eliminar eventuais linhas vazias que possam existir após o pré-processamento das palavras.

Para a contagem de palavras únicas foi utilizado o comando *uniq -c -i*. Este comando conta as ocorrências de cada palavra ignorando letras maiúsculas e minúsculas (*case insensitive*).

De seguida, organizaram-se os resultados com o comando *sort -rn* para ordenar a lista das palavras por ordem decrescente e seu valor numérico (opção *-n*), obtendo assim as palavras numeradas com maior ocorrência no topo da lista.

Para finalizar, foi utilizado o comando *cat -n* para mostrar a numeração das linhas do resultado final e o comando *tr -d* para remover um *tab* do resultado final, apenas para melhorar o resultado visual.

# Extras

O Tamanho da imagem do gráfico no modo *plot*, ajusta-se consoante o número de palavras a serem apresentadas. Utilizando a razão de crescimento dada nas variáveis globais no início do programa. Caso o tamanho inicial seja 1366\*768 e o grau de crescimento introduzido for de 0.8, então será multiplicado o número de resultados com o fator também multiplicado por 10, assim obtém-se a largura. Para obter o comprimento divide-se a altura inicial com a largura e multiplica-se com a nova largura.

# Web grafia

https://tldp.org/LDP/Bash-Beginners-Guide/html/sect\_07\_01.html

https://opensource.com/article/18/5/you-dont-know-bash-intro-bash-arrays

https://stackoverflow.com/questions/14370133/is-there-a-way-to-create-key-value-pairs-in-bash-script

https://linuxize.com/post/bash-functions/

https://stackoverflow.com/questions/965053/extract-filename-and-extension-in-bash

https://stackoverflow.com/questions/15028567/get-the-index-of-a-value-in-a-bash-array

https://stackoverflow.com/questions/13219634/easiest-way-to-check-for-an-index-or-a-key-in-an-array

https://linuxize.com/post/bash-case-statement/

https://www.baeldung.com/linux/use-command-line-arguments-in-bash-script

https://stackoverflow.com/questions/10931915/how-can-i-read-words-instead-of-lines-from-a-file

https://linuxize.com/post/bash-write-to-file/

https://stackoverflow.com/questions/12916352/shell-script-read-missing-last-line

https://askubuntu.com/questions/674333/how-to-pass-an-array-as-function-argument

https://stackoverflow.com/questions/4069188/how-to-pass-an-associative-array-as-argument-to-a-function-in-bash

https://stackoverflow.com/questions/19417015/how-to-copy-an-array-in-bash

https://stackoverflow.com/questions/8217049/bash-associative-array-sorting-by-value

https://lists.gnu.org/archive/html/bug-bash/2017-06/msg00005.html

https://linuxize.com/post/how-to-compare-strings-in-bash/

http://man.he.net/

https://linux.die.net/man/1/wc

https://linuxize.com/post/how-to-use-grep-command-to-search-files-in-linux/

https://misc.flogisoft.com/bash/tip\_colors\_and\_formatting

https://stackoverflow.com/questions/1728683/case-insensitive-comparison-of-strings-in-shell-script

https://stackoverflow.com/questions/806906/how-do-i-test-if-a-variable-is-a-number-in-bash

https://unix.stackexchange.com/questions/35333/what-is-the-fastest-way-to-view-images-from-the-terminal

http://gnuplot.sourceforge.net/docs\_4.2/

https://stackoverflow.com/questions/40294902/bash-script-check-string-for-uppercase-letter

https://stackoverflow.com/questions/6744006/can-i-use-sed-to-manipulate-a-variable-in-bash

http://gnuplot-surprising.blogspot.com/2011/10/add-value-labels-to-top-of-bars-in-bar.html

https://stackoverflow.com/questions/9084257/bash-array-with-spaces-in-elements

https://stackoverflow.com/questions/806906/how-do-i-test-if-a-variable-is-a-number-in-bash

https://stackoverflow.com/questions/3685970/check-if-a-bash-array-contains-a-value

https://www.gnu.org/software/bash/manual/html\_node/Bash-Conditional-Expressions.html

https://unix.stackexchange.com/questions/40786/how-to-do-integer-float-calculations-in-bash-or-other-languages-frameworks