

Especificação do Trabalho 2 – Busca composta utilizando arquivos invertidos

Cada aluno deve desenvolver um programa na linguagem C que receba vários nomes de arquivo como parâmetro. Cada arquivo terá um texto qualquer, e todas as palavras contidas nestes textos possuem apenas os caracteres minúsculos [a-z] e maiúsculos [A-Z] sem acento. O texto ainda pode conter espaços, quebras de linha, tabs, traços e pontuações, que devem ser considerados separadores (a sentença “pode-se” deve ser considerada duas palavras: “pode” e “se”). Para efeitos de comparação, letras maiúsculas e suas respectivas versões minúsculas não devem ser consideradas diferentes. O exemplo abaixo mostra como o programa será testado.

Ex:

```
./a.out arq1.txt arq2.txt ... arqN.txt
```

O trabalho deverá ser enviado por e-mail (mpamplona2.ufba@gmail.com), com o assunto “TRABALHO 2 - EDAII.2014.2”. Este e-mail deve conter um único anexo de nome “trabalho2.tgz”. Dentro deste arquivo compactado deve haver uma pasta com o nome “trabalho2” que contém o código fonte e um arquivo Makefile que gere um executável chamado “a.out” no diretório “trabalho2”. Cada aluno deverá utilizar o mesmo arquivo de identificação do primeiro trabalho, que deve ser incluso inalterado no diretório “trabalho2”.

O programa deve então ler da entrada padrão uma sequência de frases (no mínimo uma palavra, e no máximo dez), e imprimir na saída padrão os nomes dos arquivos nos quais estas frases aparecem. O nomes dos arquivos devem estar separados por um espaço (não há espaço após o último nome de arquivo), e a linha deve terminar com uma quebra de linha. Se a sequência não ocorrer em nenhum arquivo, deve-se imprimir na saída padrão a frase “FRASE NAO ENCONTRADA!\n”. Deve-se utilizar o conceito de arquivos invertidos para este trabalho.

Segue abaixo alguns exemplos:

arq1.txt	arq2.txt	arq3.txt
O rato roeu a roupa do rei de roma.	Quem matou o rato? Roeu algo proibido?	O rato nanico roeu um pedacinho do queijo.

\$./a.out arq1.txt arq2.txt arq3.txt

Entrada padrão	Saída padrão
o rato roeu o rato proibido O RATO	arq1.txt arq2.txt FRASE NAO ENCONTRADA! arq1.txt arq2.txt arq3.txt