

### Observações:

- Data de entrega: **23 de Outubro de 2017.**

## 1 Problema: as $k$ palavras mais proximas

Pretende-se desenvolver uma aplicação que permita determinar, dado uma palavra  $p$  e um inteiro positivo  $k$ , as  $k$  palavras distintas com o prefixo mais próximo da palavra  $p$  que ocorrem em pelo menos um dos  $n$  ficheiros ordenados lexicograficamente, em que o número de palavras total em todos os ficheiros é  $m$ . Assuma que  $m \geq 1$ .

O problema é descrito por:

- Um conjunto  $F = f_1, \dots, f_n$  de  $n$  ficheiros de texto, em que  $n > 0$ ;
- Cada linha de um ficheiro  $f_n$  tem uma palavra;
- Uma palavra  $p$ ;
- Um inteiro  $k$ .

O objetivo da aplicação a desenvolver é a produção de um novo ficheiro de texto contendo as  $k$  palavras distintas com o prefixo mais próximo da palavra  $p$  que ocorram em pelo menos um dos  $n$  ficheiros considerados.

### Parâmetros de Execução

A aplicação a desenvolver terá de suportar as seguintes opções:

- `java -Xmx32m MaisProximas k p outputFile sourceFiles ...` Esta opção corresponde à produção de um ficheiro de texto designado por `outputFile` contendo as  $k$  palavras distintas com o prefixo mais próximo da palavra  $p$  que ocorrem pelo menos num dos ficheiros de texto `sourceFiles`. A opção `-Xmx32m` estabelece para o Java um *heap size* máximo de 32 MB.

Por exemplo, dados os ficheiros de texto `f1`, `f2` e `f3`, ordenados lexicograficamente, se executarmos:

```
java -Xmx32m MaisProximas 4 aed output.txt 10 f1.txt f2.txt f3.txt
```

iremos produzir um ficheiro de texto designado por `output.txt`, contendo as 4 palavras distintas com o prefixo mais próximo da palavra `aed` que ocorrem pelo menos num dos ficheiros de texto `f1`, `f2` e `f3`.

### Avaliação Experimental

Realize uma avaliação experimental do(s) algoritmo(s) desenvolvido(s) para a resolução deste problema. Como exemplo, poderá utilizar os ficheiros que se encontram disponíveis juntamente com este enunciado. Apresente os resultados graficamente, utilizando uma escala adequada.