

Corretor Ortográfico

Conceção e Análise de Algoritmos

1 de junho de 2015

Turma 2/Grupo A

Francisco Rodrigues - up201305627

Luís Oliveira - up201304515

Miguel Pereira - up201305998

Índice

1. Introdução	2
2. Definição do problema	3
2.1. Descrição do problema	3
3. Solução implementada	4
4. Lista de casos de utilização	5
5. Conclusão	6
5.1 Dificuldades	6
5.2 Contribuições	6

1. Introdução

No âmbito da unidade curricular de Conceção e Análise de Algoritmos foi-nos proposto um problema onde seria necessário desenvolver uma ferramenta de correção ortográfica de modo a que fosse possível detetar e apresentar possíveis soluções para erros ortográficos num determinado ficheiro fornecido pelo utilizador.

Neste relatório, será explicado o contexto do problema, e será feita uma descrição da solução, indicando os principais algoritmos implementados.

2. Definição do problema

2.1. Descrição do problema

É pretendido a partir de um ficheiro fornecido pelo utilizador e tendo em conta um dicionário escolhido pelo utilizador a partir de uma lista fornecida, encontrar os erros ortográficos presentes no ficheiro e apresentar sugestões de correção baseadas na sua proximidade à palavra por corrigir.

3. Solução implementada

De maneira a resolver o problema proposto, era necessário para cada palavra do ficheiro de texto tentar encontrar uma correspondência no dicionário, e caso esta tarefa não fosse bem sucedida assume-se que a palavra não existe, ou seja, contém erros ortográficos. A partir daqui esta palavra seria comparada com as palavras do dicionário utilizando um algoritmo de pesquisa aproximada de *strings*, através do qual conseguiríamos obter a distância de edição entre a palavra com erro e as palavras do dicionário. As palavras cuja distância de edição fosse menor, seriam as mais aproximadas à palavra com erro e consequentemente possíveis correções.

O programa segue a seguinte estrutura:

- Carrega um ficheiro contendo o dicionário escolhido pelo utilizador a partir dos que são disponibilizados.
- Pede o nome do ficheiro a verificar ao utilizador.
- Carrega o ficheiro e para cada palavra verifica se esta existe no dicionário, e caso esta exista não é considerada como errada.
- Caso a palavra não exista no dicionário é executado o algoritmo de pesquisa que devolve a distância de edição entre a palavra a considerar e cada palavra do dicionário (são consideradas edições a remoção/adição de letras, a substituição de uma letra ou a permuta entre duas letras adjacentes da palavra) guardando as possíveis soluções. O algoritmo de pesquisa aproximada entre duas palavras tem a complexidade temporal e espacial de:

$$O(|palavra1|. |palavra2|)$$

- São apresentadas as palavras erradas (identificando-se também a linha onde são apresentadas) e as respetivas sugestões de correção. Só são apresentadas como correções as palavras que tenham distância de edição inferior ou igual a $Tamanho\ da\ palavra\ errada \times 0.45$ – se este valor for superior a 5 edições, as 5 edições serão consideradas o novo máximo.

4. Lista de casos de utilização

- Verificar se uma palavra existe no dicionário
- Verificar se um texto está escrito sem erros ortográficos e obter possíveis correções para os eventuais erros.

5. Conclusão

5.1 Dificuldades

Não foram encontradas dificuldades de fazer notar na realização deste projeto.

5.2 Contribuições

- Miguel Pereira – 40%
- Luís Oliveira – 20%
- Francisco Rodrigues – 40%