

1º Exercício Escolar de Introdução à Programação
Prof Alexandre Mota – 03/10/2016

Questão. Na maioria dos sistemas baseados em texto atualmente (Email, SMS, Whatsapp, etc.), quando o usuário digita uma palavra errada, os sistemas dão sugestões das palavras mais próximas baseadas em um dicionário de palavras previamente cadastradas. Portanto, se um usuário digita: **escelente**, ou **scelente**, ou ainda **escelete**, o sistema sugere **excelente** como alternativa mais próxima. Normalmente mais de uma palavra que poderia ser equivalente é oferecida. Nota-se então que a sugestão é em função da maioria dos caracteres que aparecem na mesma ordem que a palavra presente no dicionário, onde a palavra errada se torna igual à sugerida por um certo número de substituições, inserções ou remoções de caracteres. Pede-se que se faça um programa em C que contenha uma matriz LxC (por exemplo, 1000x15) de caracteres como dicionário onde cada palavra a ser armazenada no dicionário tem a característica de que o último caractere de cada palavra termina com um '.'. Como exemplo de um tal dicionário poderíamos ter:

e	x	c	e	l	e	n	t	e	.					
p	r	a	z	e	r	.								
...														
i	m	p	r	e	s	s	i	o	n	a	n	t	e	.

Seu programa deve oferecer ao usuário todas as palavras que possam substituir a errada que o usuário escreveu inicialmente (esta palavra deve ser armazenada em vetor de tamanho C (por exemplo, 15) onde o último caractere deve ser um '.'), com similaridade de pelo menos 60% e que contenham pelo menos 3 caracteres, sem contar com o '.'. Para tanto, seu programa deve calcular o grau de similaridade da palavra digitada em relação às armazenadas no dicionário, levando em conta quantos caracteres são iguais e na mesma ordem, quantos caracteres teriam que ser substituídos, inseridos ou removidos. Seu programa só deve parar de ler novas palavras se apenas um '.' for digitado. Seu programa também não deve aceitar palavras com mais de C caracteres. Se isto ocorrer, uma mensagem de erro deve ser apresentada ao usuário.

Como exemplos, no caso de **escelente**, teríamos uma similaridade de $8/9=88,89\%$ e de **escelete**, uma similaridade de $7/9=77,78\%$. O denominador é o tamanho da palavra do dicionário, sem contar com o '.' e o numerador é formado pelos caracteres iguais e que obedecem a ordem. As palavras do dicionário têm que ser apresentadas do maior grau de semelhança para o de menor grau.

Boa sorte!!!

OBS.: Proibido o uso de strings tradicionais de C pois o assunto ainda não foi apresentado. Usem o que foi exposto na questão.