

sequência pequena de texto dentro de uma sequência maior pode ser encontrada, isso não caracteriza um erro, mas para a comparação feita no algoritmo do AC poderia causar problemas (a comparação no sistema é feita termo a termo). Por exemplo, a palavra "ia" poderia ser encontrada na palavra "tecnologia". Para que esse problema não ocorra, o algoritmo verifica o tamanho das duas palavras a serem comparadas e tenta encontrar a maior palavra dentro da menor, assim dois termos serão correspondentes apenas se a diferença entre os dois respeitar o limite de erros permitidos.

A Figura 14 mostra a possibilidade de encontrar a palavra "tema" (termo menor) dentro da palavra "sistema" (termo maior). Nessa comparação o termo foi encontrado com nenhum erro.

Figura 14 - Comparando termo menor com termo maior

tema	→	sistemas	
sistem <u>a</u> s			Erros = 0

Fonte: Elaborado pelo autor

A Figura 15 mostra a comparação entre a palavra "sistemas" (termo maior) e a palavra "tema" (termo menor). Nessa comparação quatro erros foram encontrados. Se o limite de erros permitidos fosse menor que quatro, o casamento entre termos iria ocorrer, caso contrário não.

Figura 15 - Comparando termo maior com termo menor

sistemas	→	tema	
sistem <u>a</u> s			Erros = 4

Fonte: Elaborado pelo autor

O número máximo de erros definido no AC foi de vinte por cento para palavras com cinco ou mais caracteres e de um erro para palavras com até quatro caracteres. A porcentagem está relacionada à quantidade de caracteres de um termo.