Realatório do trabalho 1 de CCD

4. Neste exercício pretende- se calcular a entropia de fonte, implementar um gerador com as mesmas caracteírsticas e obter o comprimento médio da arvore de huffman.

Os simbolos gerados por esta fonte vão estar compriendidos entre 3 e 18 pois se calhar 3 vezes 1 no dado lançado ou 3 vezes 6. Deste modo a fonte só vai ter 15 simbolos possiveis.

A análise do gerador é semelhante ao dos exercídios anteriores mas com uma particularidade para alem dos 15 simbolos possiveis.

Anteriormente os simbolos tinham no máximo 1 byte mas neste caso podem ter até 2 bytes pois se o valor tiver 2 caracters, por exemplo, o 18.

Para resolver este problema é necessário que no dicionário seja possivel representar este par de caracters na chave.

A árvore de Huffman está a ser gerada em função dos resultados estaísticos da fonte. Ordenando as probabilidades de forma crescente e gerar os nós parent apartir das folhas.

Em termos de simbolos gerados confirma-se que os simbolos que aparecem com mais frequência na geração da fonte tambem têm mais frequência no simulador por nós implementado.

7. Neste exercício pretende-se fazer compressão de ficheiros utilizando o algoritmo Deflate contido na biblioteca jzlib. Ao analisar a compressão obtida pelos diferentes niveis de compressão podemos concluir que:

* A taxa de compressão é maior quanto maior for o tamanho do ficheiro que pretendemos comprimir.
* A compressão sem compressão aumenta o tamanho final do ficheiro após compressão.
* A compressão best compression tem taxas de compressão muito semelhantes à default compression.
* A compressão default e a best compression são as compressões que comprimem em maior quantidade.
* A best speed compression apesar de ser mais rápida não comprime tanto como as restantes.
* Para casos de equiprobabilidade dos simbolos da fonte o desempenho da compressão é muito mau tendo mesmo o tamanho do output ser superior ao de input.

8. Neste exercício pretendemos usar o algoritmo de deflate para comparar o desempenho da compressão para grupos de ficheiros, sendo a compressão efetuada com a concatenação de todos os simbolos ou comprimir todos os ficheiros em separado.

Podemos concluir que a compressão em coletivo tem melhor desempenho que a compressão em separado.