



Trabalho Prático 2

Mestrado Integrado em Engenharia de Telecomunicações e Informática

Tecnologias e Serviços Multimédia

Leandro Alves A82157

Estratégias adotadas

Neste trabalho prático, foi implementado um mecanismo de codificação prévia (LZW) e foi, também, integrado com o mecanismo desenvolvido no trabalho prático anterior (Shannon). Este mecanismo pode ser utilizado sozinho, ou então, caso o utilizador preferir, pode combinar os dois.

Compressão LZW

O mecanismo de codificação por dicionário é constituído por uma tabela com 4096 entradas (12 bits por símbolo).

Os passos de aplicação deste mecanismo são os seguintes:

1. Inicialização do dicionário (caracteres conhecidos, códigos de 0-255)
2. Leitura dos *bytes* do ficheiro, por blocos
3. Processamento desses *bytes* (codificação de um novo código)
4. Conversão do código obtido (12 bits) para 8 bits
5. Adicionar o código convertido no *buffer* final

Assim como no trabalho prático anterior, a leitura é feita por blocos de 1 MB.

Compressão LZW mais Shannon

Neste tipo de compressão, os passos são, ligeiramente, iguais à compressão LZW, apenas são adicionados os seguintes passos:

1. Construção da tabela de Shannon
2. Conversão dos bytes presentes em *buffer* final para *bits* recorrendo à tabela de Shannon
3. Escrita no ficheiro comprimido

Reparar que, nos passos 3 e 4 da compressão LZW, são feitas algumas alterações para que possa ser aplicado o algoritmo Shannon-Fanno, como, por exemplo, adicionar os códigos convertidos numa tabela de ocorrências de símbolos.

Descompressão LZW

Na descompressão são realizados os seguintes passos, por cada bloco:

1. Iniciar o dicionário com os padrões conhecidos
2. Leitura dos *bytes*
3. Conversão desses *bytes* (8 *bits*) para 12 *bits*
4. Aplicação do algoritmo LZW
5. Escrita dos *bytes* no ficheiro

Descompressão LZW mais Shannon

Bem como na compressão, a descompressão LZW mais Shannon é muito parecida à do LZW, só se tem de processar os *bits* contidos no *payload*, ou seja, converter de *bits* para *bytes*. Os restantes passos são exatamente os mesmos.

Testes e Tempos obtidos

Para testar a eficácia e a rapidez das ferramentas desenvolvidas, foi-nos disponibilizado, pelo docente, alguns ficheiros de teste. Nas seguintes figuras estão alguns desses ficheiros a serem comprimidos/descomprimidos usando as técnicas ditas anteriormente.

```

leandro@azkeryon:~/Documentos/TSM$ ./TP2 -c -l -v texto.txt
COMPRESSÃO LZW INICIADA...

      Bloco número 1 (1000000 bytes)
->Tempo de processamento do LZW: 43.77 ms
->Tamanho do bloco comprimido: 545984 bytes

      Bloco número 2 (1000000 bytes)
->Tempo de processamento do LZW: 27.91 ms
->Tamanho do bloco comprimido: 559110 bytes

      Bloco número 3 (1000000 bytes)
->Tempo de processamento do LZW: 28.71 ms
->Tamanho do bloco comprimido: 551517 bytes

      Bloco número 4 (1000000 bytes)
->Tempo de processamento do LZW: 26.92 ms
->Tamanho do bloco comprimido: 544008 bytes

      Bloco número 5 (1000000 bytes)
->Tempo de processamento do LZW: 34.26 ms
->Tamanho do bloco comprimido: 549828 bytes

      Bloco número 6 (776694 bytes)
->Tempo de processamento do LZW: 29.11 ms
->Tamanho do bloco comprimido: 435668 bytes

Nível de compressão: 1.81
Tempo de execução total: 200.14 ms

```

Figura 1 - Compressão LZW do ficheiro "texto.txt" com modo *verbose*

```

leandro@azkeryon:~/Documentos/TSM$ ./TP2 -d -l -v texto.txt
DESCOMPRESSÃO LZW INICIADA...

      Bloco número 1 (1000000 bytes)
-> Tempo de processamento do LZW e escrita no ficheiro: 19.50 ms

      Bloco número 2 (1000000 bytes)
-> Tempo de processamento do LZW e escrita no ficheiro: 14.30 ms
-> Tempo de processamento do LZW e escrita no ficheiro: 13.89 ms

      Bloco número 3 (1000000 bytes)
-> Tempo de processamento do LZW e escrita no ficheiro: 14.22 ms
-> Tempo de processamento do LZW e escrita no ficheiro: 13.24 ms

      Bloco número 4 (186115 bytes)
-> Tempo de processamento do LZW e escrita no ficheiro: 10.99 ms

Tempo de execução total: 90.84 ms
leandro@azkeryon:~/Documentos/TSM$ cmp NEW_texto.txt texto.txt
leandro@azkeryon:~/Documentos/TSM$ █

```

Figura 2 - Descompressão LZW do ficheiro original "texto.txt" com modo *verbose*

```
leandro@azkeryon:~/Documentos/TSM$ ./TP2 -c -ls -v music.mp3
COMPRESSÃO LZW + SHANNON INICIADA...

      Bloco número 1 (1000000 bytes)
-> Tempo de processamento do LZW: 64.65 ms
-> Entropia: 8.79 bits/símbolo
-> Tempo de tratamento e cálculo do Shannon: 0.16 ms
-> Comprimento médio: 9.63
-> Tempo de construção do bloco comprimido: 161.86 ms
-> Tamanho do bloco comprimido: 1204003 bytes

      Bloco número 2 (1000000 bytes)
-> Tempo de processamento do LZW: 44.14 ms
-> Entropia: 8.84 bits/símbolo
-> Tempo de tratamento e cálculo do Shannon: 0.08 ms
-> Comprimento médio: 9.69
-> Tempo de construção do bloco comprimido: 162.48 ms
-> Tamanho do bloco comprimido: 1211498 bytes

      Bloco número 3 (1000000 bytes)
-> Tempo de processamento do LZW: 46.29 ms
-> Entropia: 8.86 bits/símbolo
-> Tempo de tratamento e cálculo do Shannon: 0.08 ms
-> Comprimento médio: 9.74
-> Tempo de construção do bloco comprimido: 163.72 ms
-> Tamanho do bloco comprimido: 1217429 bytes

      Bloco número 4 (1000000 bytes)
-> Tempo de processamento do LZW: 47.85 ms
-> Entropia: 8.87 bits/símbolo
-> Tempo de tratamento e cálculo do Shannon: 0.09 ms
```

Figura 3 - Compressão LZW + Shannon do ficheiro "music.mp3" com modo *verbose*

```

leandro@azkeryon:~/Documentos/TSM$ ./TP2 -d -ls -v music.mp3
DESCOMPRESSÃO LZW + SHANNON INICIADA...

Bloco número 1 (1500000 bytes)
-> Tempo de processamento do cabeçalho e construção da tabela shannon: 0.12 ms
-> Tempo de processamento do payload: 376.81 ms
-> Tempo de processamento do LZW: 27.24 ms

Bloco número 2 (1500000 bytes)
-> Tempo de processamento do cabeçalho e construção da tabela shannon: 0.08 ms
-> Tempo de processamento do payload: 364.39 ms
-> Tempo de processamento do LZW: 25.58 ms

Bloco número 3 (1000000 bytes)
-> Tempo de processamento do cabeçalho e construção da tabela shannon: 0.08 ms
-> Tempo de processamento do payload: 373.70 ms
-> Tempo de processamento do LZW: 28.05 ms

Bloco número 4 (1500000 bytes)
-> Tempo de processamento do cabeçalho e construção da tabela shannon: 0.07 ms
-> Tempo de processamento do payload: 374.24 ms
-> Tempo de processamento do LZW: 27.28 ms

Bloco número 5 (1000000 bytes)
-> Tempo de processamento do cabeçalho e construção da tabela shannon: 0.08 ms
-> Tempo de processamento do payload: 376.58 ms
-> Tempo de processamento do LZW: 25.82 ms

Bloco número 6 (1500000 bytes)
-> Tempo de processamento do cabeçalho e construção da tabela shannon: 0.12 ms

```

Figura 4 - Descompressão LZW + Shannon do ficheiro original "music.mp3" com modo *verbose*

Os tempos obtidos para os diferentes tipos de técnicas foram os presentes nas tabelas seguintes.

Técnica LZW mais Shannon				
Ficheiro	Tamanho (bytes)	Tamanho comprimido (bytes)	Tempo de compressão	Tempo de descompressão
Texto.txt	5776694	3102087	~620ms	~960ms
Music.mp3	21058294	25628062	~4.4s	~8.5s
Music.flac	53380401	66197939	~11.4s	~20.9s
Music.wav	92847804	109824712	~19.5s	~34.5s

Técnica Shannon				
Ficheiro	Tamanho (bytes)	Tamanho comprimido (bytes)	Tempo de compressão	Tempo de descompressão
Texto.txt	5776694	3530242	~515ms	~836ms
Music.mp3	21058294	21072348	~3.2s	~6.1s
Music.flac	53380401	53582350	~8.1s	~15.2s
Music.wav	92847804	90511967	~13.3s	~27.1s

Técnica LZW				
Ficheiro	Tamanho (bytes)	Tamanho comprimido (bytes)	Tempo de compressão	Tempo de descompressão
Texto.txt	5776694	3186115	~200ms	~90ms
Music.mp3	21058294	30399509	~950ms	~650ms
Music.flac	53380401	77731857	~2.4s	~1.6s
Music.wav	92847804	131925078	~4.0s	~2.8s

Limitações

A principal limitação deste código é no processamento dos códigos de *Shannon*, pois este é trabalhado *bit a bit*, fazendo com que o programa se torne lento, quer na compressão, quer na descompressão.

Manual

Se correr o programa sem introduzir nenhum argumento, é mostrado um manual de utilização, no qual, está indicado o que cada argumento deverá ser. Como podemos ver na seguinte figura.

```
leandro@azkeryon:~/Documentos/TSM$ ./TP2
Usage: TP2 [MODE] [TYPE] [OPTION] [FILENAME]

MODE required
  -c,  Compression mode
  -d,  Decompression mode

TYPE required
  -ls, Apply LZW and Shannon compression
  -s,  Apply only Shannon compression
  -l,  Apply only LZW compression

OPTION
  -v,  Apply debug mode
```