

Introdução à Teoria da Informação

Prof. Derzu Omaia

Trabalho 1 – PPM/LZW

(Individual, dupla ou trio)

Implemente um compressor e descompressor utilizando o PPM-C (visto em sala de aula) com o codificador aritmético, ou utilizando o algoritmo LZW. Considere que as mensagens são geradas por fontes com alfabeto $A = \{0, 1, \dots, 255\}$. Teste o compressor/descompressor com um [corpus de texto](#) em português de 16MB e com um [arquivo binário](#) de vídeo.

PPM-C + Aritmético:

O contexto deve ter tamanho máximo K (parâmetro). O modelo PPM-C alimentará um codificador aritmético. Utilize o mecanismo de exclusão quando necessário. No relatório apresente as curvas de RC x K e de Tempo de Processamento x K , para $K = 0, 1, \dots, 8$. Não é necessário implementar o codificador aritmético, utilize algum já existente.

LZW:

O índice do dicionário deve ser testado com diferentes tamanhos K bits (parâmetro). Exemplo: $K=9$ bits tamanho do dicionário: $2^9=512$, $K=10$ bits tamanho do dicionário $2^9=1024$. No relatório apresente as curvas de RC x K e de Tempo de Processamento x K , para $K = 9, 10, \dots, 16$ bits.

Observações:

1. Os símbolos do arquivo de teste devem ser lidos no modo binário (números) e não no modo texto (caracteres/strings).
2. O codificador deve receber como entrada um arquivo e gerar como saída o arquivo codificado..
3. O trabalho deve ser apresentado em sala de aula na data de entrega.
4. A execução dos experimentos é demorada, evite fazer os experimentos na véspera da entrega pois não dará tempo.
5. Relatório explicando o desenvolvimento.

Links:

<https://www.dropbox.com/s/akjcu9rb9lhxqx0/corpus16MB.zip>

<https://www.dropbox.com/s/hh3kqnwvyr5aamz/mapa.mp4>