

Introdução à Teoria da Informação

Prof. Derzu Omaia

Trabalho 2 - PPM

(Individual, dupla ou trio)

Implemente um compressor e descompressor baseado no modelo PPM-C (visto em sala de aula), com o mecanismo de exclusão, para fontes com alfabeto $A = \{0, 1, \dots, 255\}$. O contexto deve ter tamanho máximo K (parâmetro). O modelo PPM-C alimentará um codificador aritmético. Teste os compressores com um corpus de texto em português de 16MB. Faça um relatório e neste apresente as curvas de RC x K e de Tempo de Processamento x K , para $K = 0, 1, \dots, 8$.

Observações:

1. Os símbolos do arquivo de teste devem ser lidos na forma de byte (números) não devem ser interpretados como caracteres/strings.
2. Não é necessário implementar o codificador aritmético, utilize algum já existente.
3. O codificador deve receber como entrada um arquivo e gerar como saída o arquivo codificado. Testar com um corpus de 16MB.
4. O trabalho deve ser apresentado em sala de aula na data de entrega.
5. A execução dos experimentos é demorada, evite fazer os experimentos na véspera da entrega pois não dará tempo.
6. Relatório explicando o desenvolvimento.