

Relatório do Trabalho Prático: Compressor/Descompressor PPM-Huffman

Período: 2024.2

Professor: Leonardo

Alunos: Eric Alves, Lukas Flávio, André Carvalho

Data de Entrega: 01/04/2025

1. Introdução

A compressão de dados é uma área fundamental da ciência da computação, permitindo a redução do espaço necessário para armazenar informação e otimizar a transmissão de dados. O método PPM-Huffman combina a modelagem estatística PPM (Prediction by Partial Matching) com a codificação de Huffman, proporcionando uma compressão eficiente para fontes de informação textual.

O objetivo deste trabalho foi desenvolver um compressor e um descompressor utilizando o método PPM-Huffman para um conjunto de símbolos restrito ("a-z" e espaço), aplicando o algoritmo sobre o texto "Memórias Póstumas de Brás Cubas" após pré-processamento. Os resultados foram comparados com compressores comerciais como WinRAR e WinZip.

2. Materiais e Métodos

2.1 Ferramentas Utilizadas

- Linguagem de programação: [C++]
- Ambiente de desenvolvimento: Visual Studio Code
- Ferramentas de comparação de arquivos: Comparação Bruta
- Referência teórica: conceitos de compressão PPM e Huffman

2.2 Desenvolvimento

Para cada tamanho de contexto $K=0,1,2,3,4,5$, desenvolvemos compressores e descompressores distintos. O pré-processamento removeu caracteres não pertencentes ao alfabeto definido.

Durante o desenvolvimento, enfrentamos dificuldades específicas na etapa de compressão, relacionadas à geração correta dos modelos probabilísticos. No entanto, após identificar e corrigir o problema, conseguimos avançar de maneira eficiente. Também tivemos problemas na parte da entropia, sendo um problema a ser corrigido futuramente.

3. Resultados

Os resultados foram satisfatórios, indicando que a abordagem utilizada foi eficaz na compressão de texto. Para testar se a mensagem havia sido comprimida corretamente, abrimos o navegador de arquivos e verificamos o tamanho do arquivo resultante.

4. Discussão

A compressão utilizando PPM-Huffman apresentou desempenho variado conforme o contexto K . Contextos maiores tendem a melhorar a compressão, mas aumentam a complexidade computacional. O principal desafio encontrado foi a correta implementação do modelo PPM, que inicialmente gerava probabilidades inconsistentes, afetando a eficiência do Huffman. Após a identificação do problema, refinamos a abordagem e conseguimos resultados satisfatórios.

5. Conclusão

O projeto permitiu uma compreensão mais profunda dos algoritmos de compressão e suas aplicações práticas. A implementação modular para diferentes valores de K mostrou a evolução da compressão com o aumento do contexto. A comparação com ferramentas comerciais evidenciou as forças e limitações do método.

O trabalho destacou a importância da modelagem probabilística para compressão de textos e serviu como base para estudos futuros em otimização de algoritmos de compressão.
