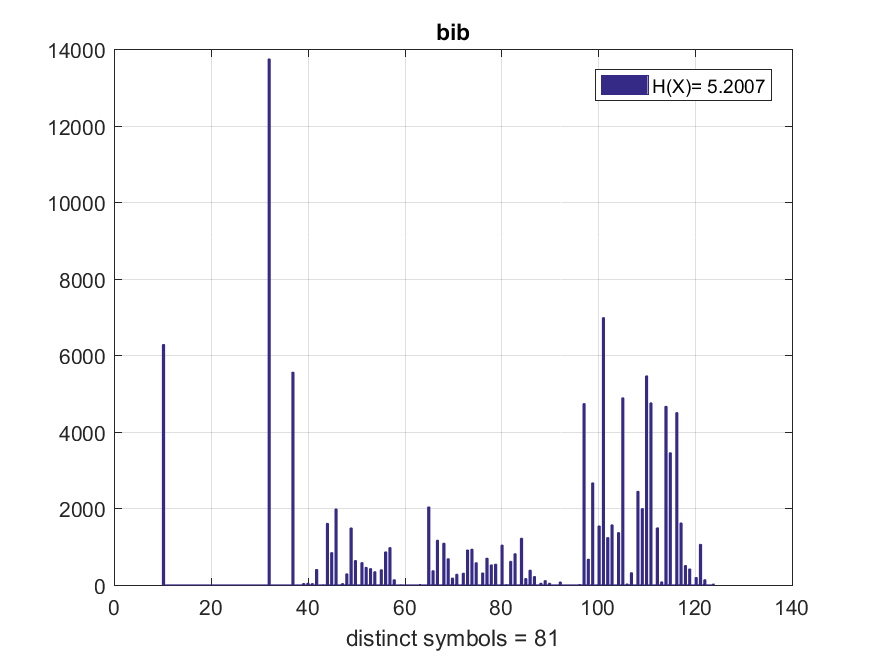
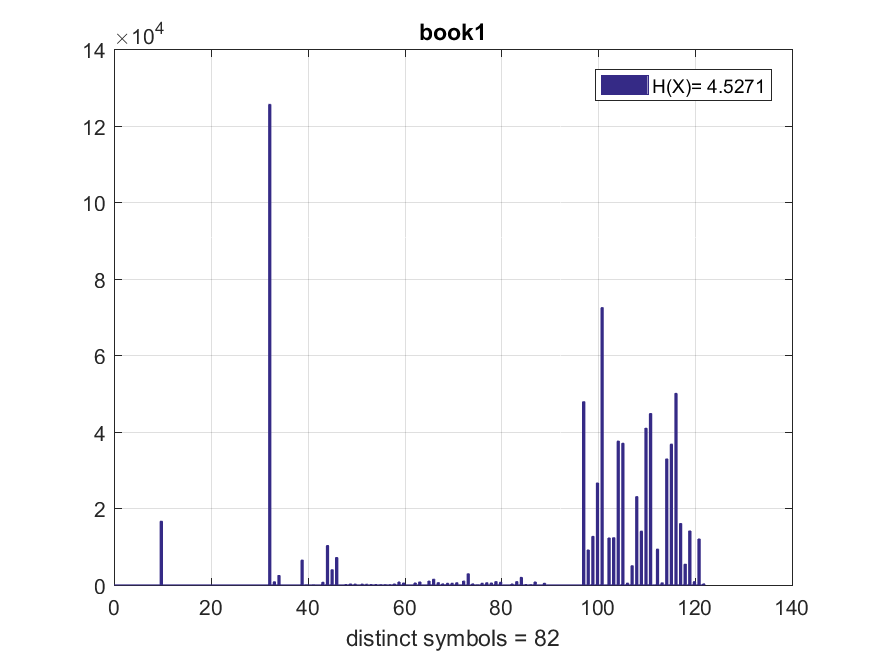
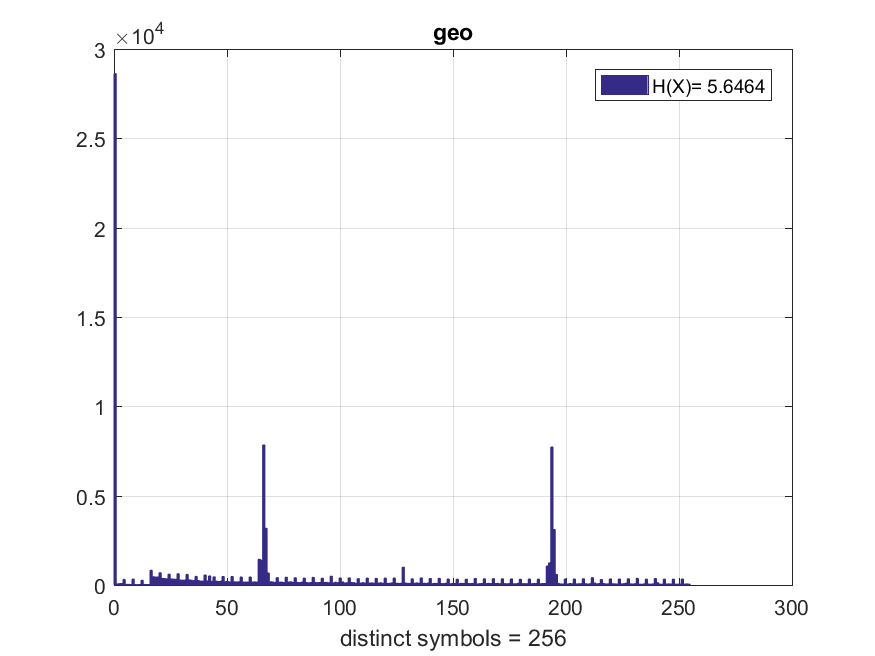
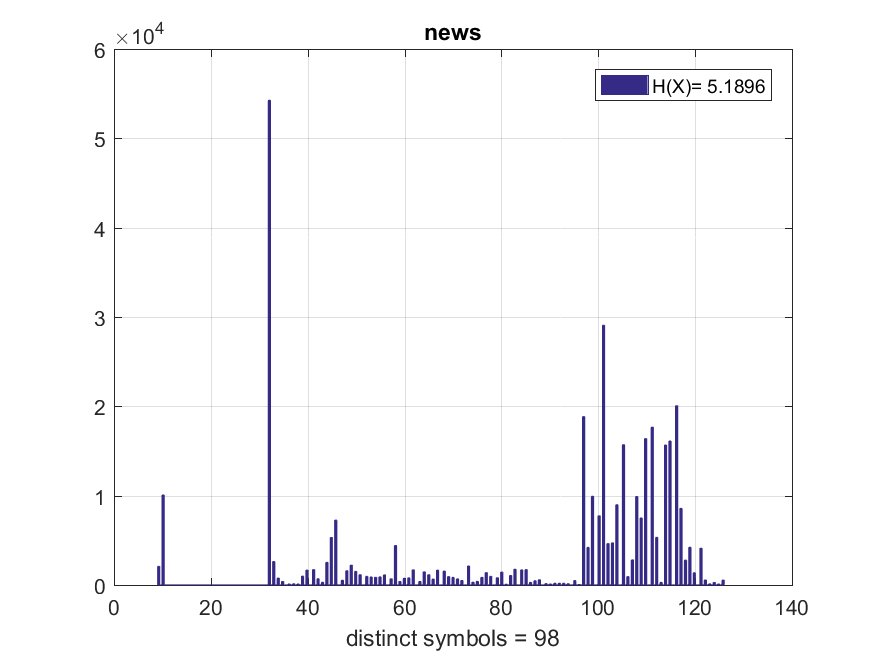
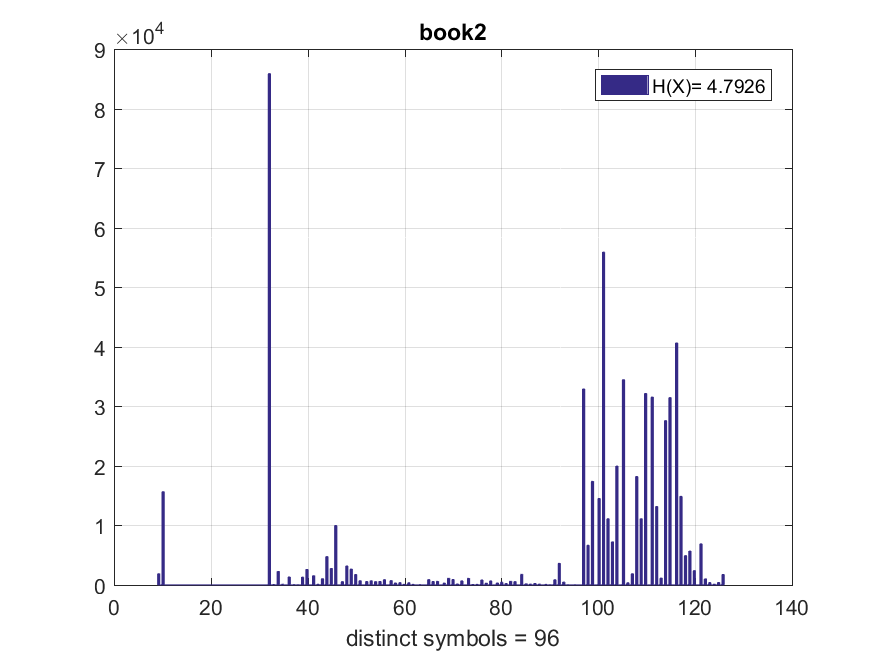
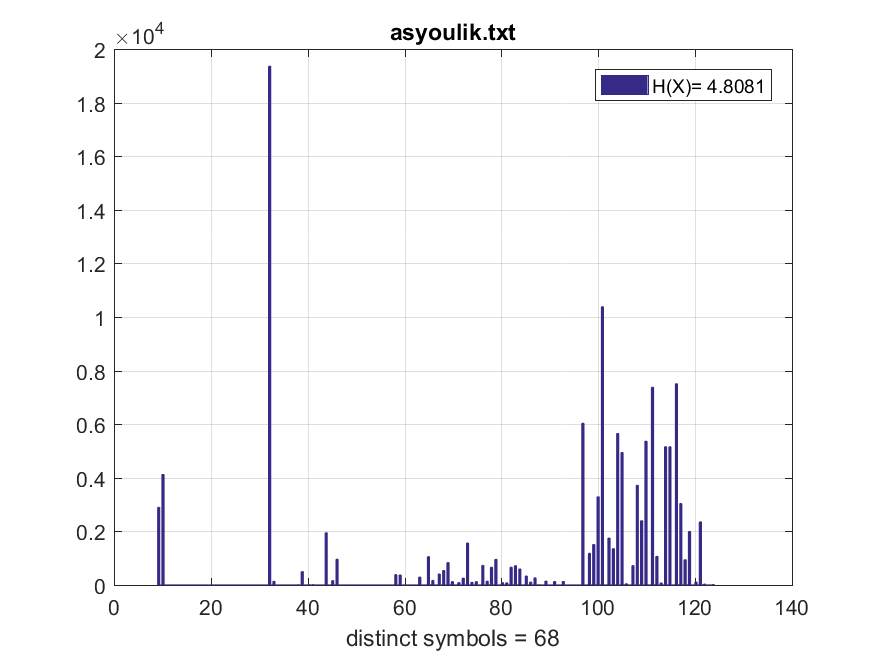
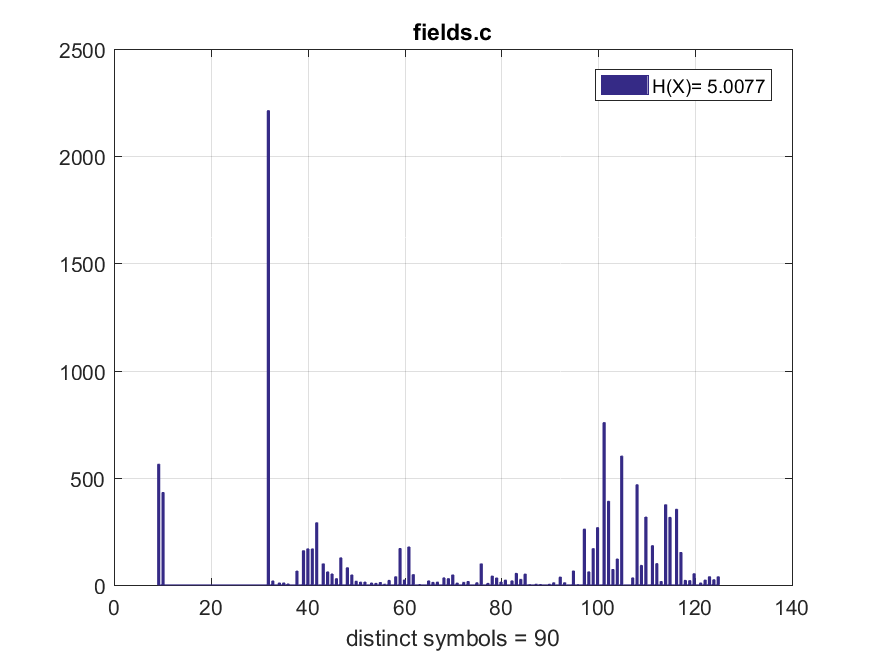
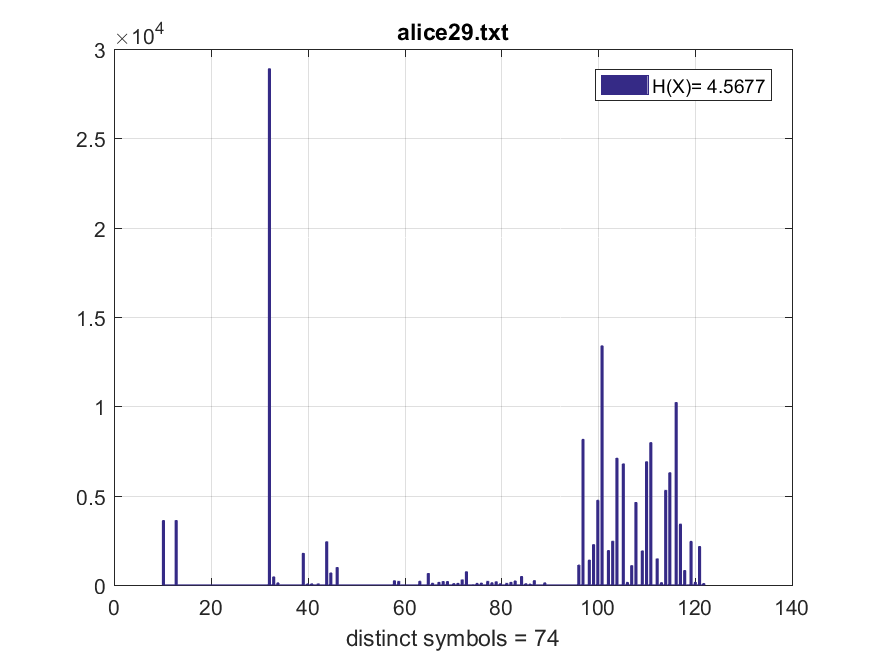
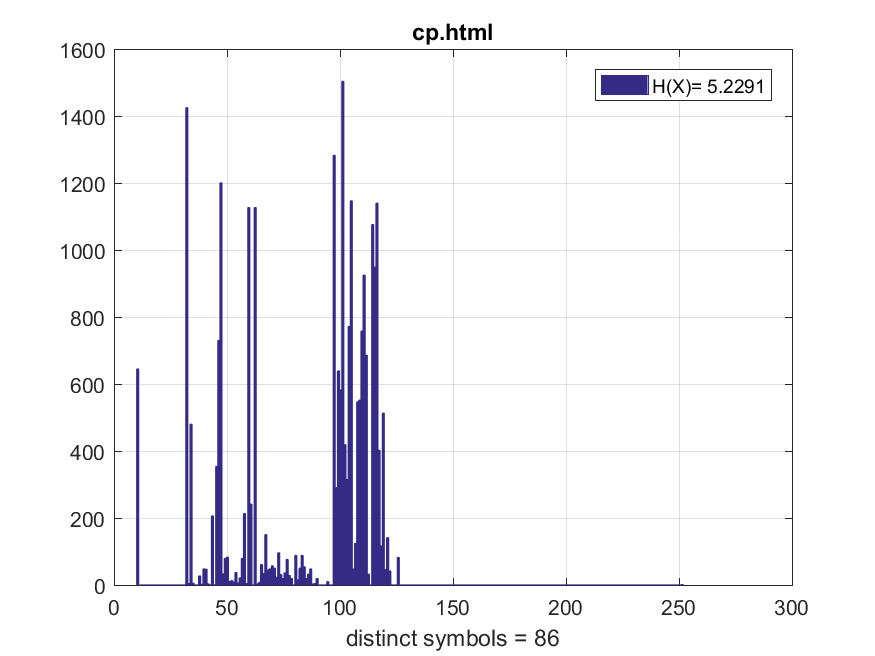
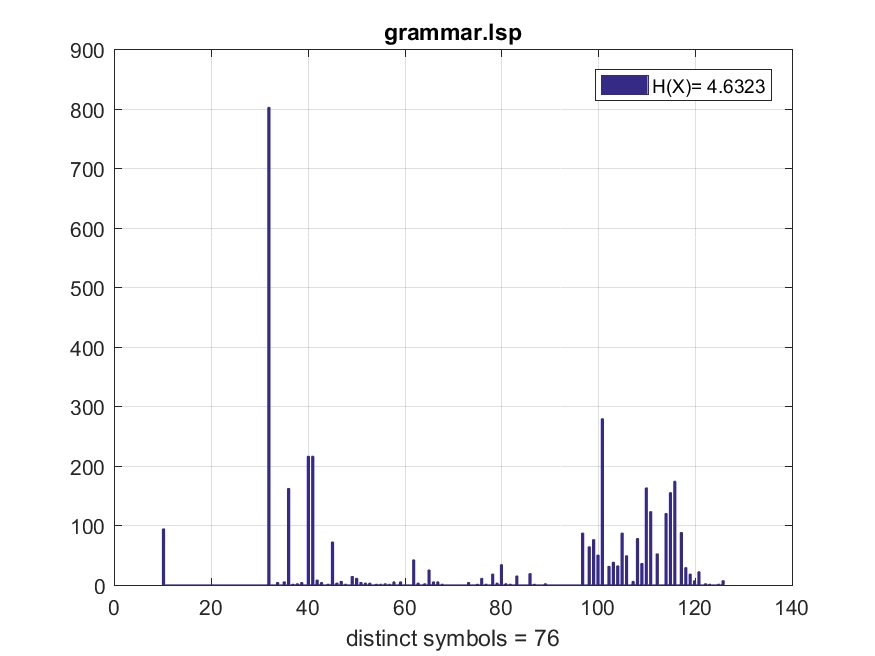
**2**

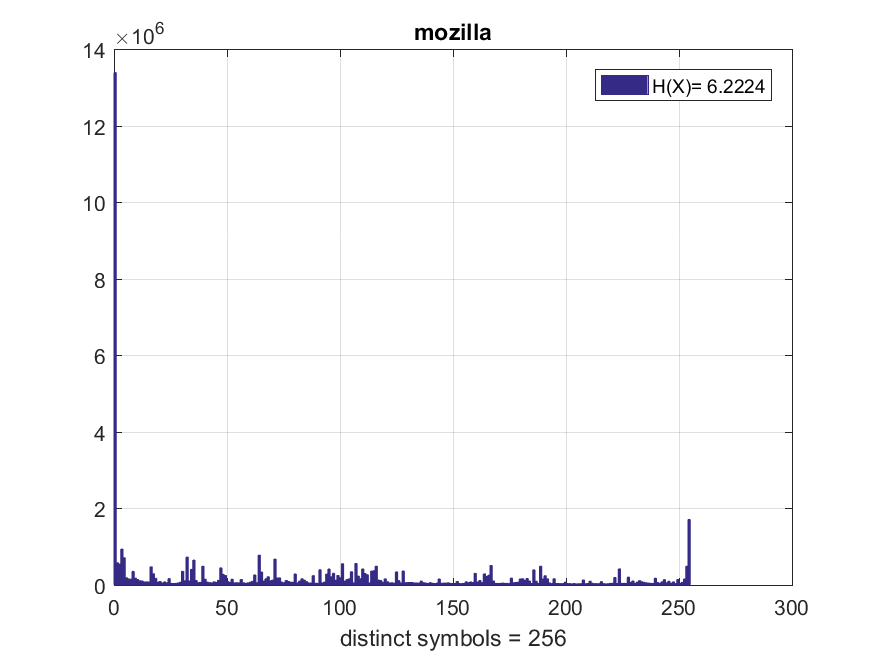
**a)** Do conjunto Calgary Corpus foram escolhidos os ficheiros: bib, book1, book2, geo, news.

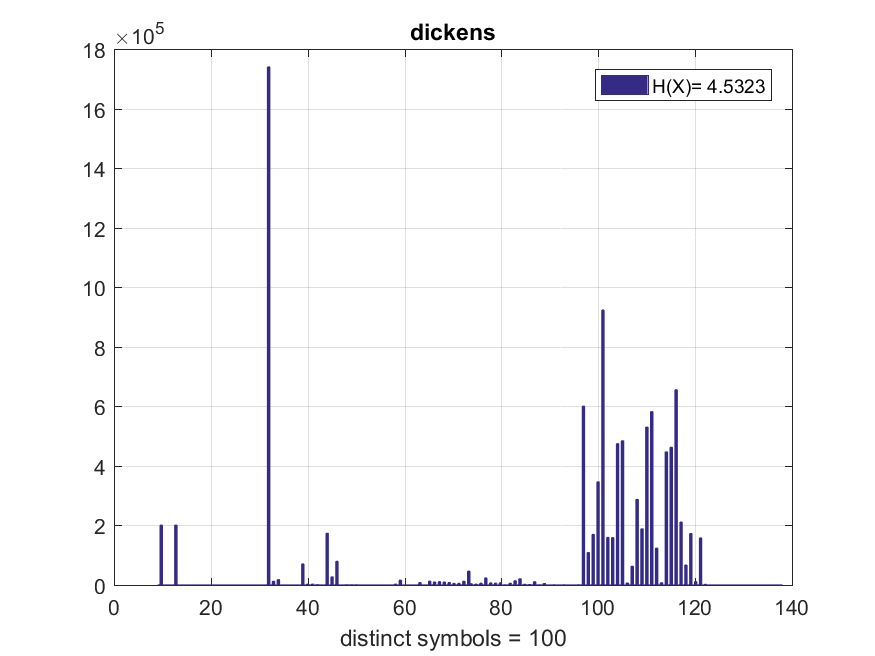
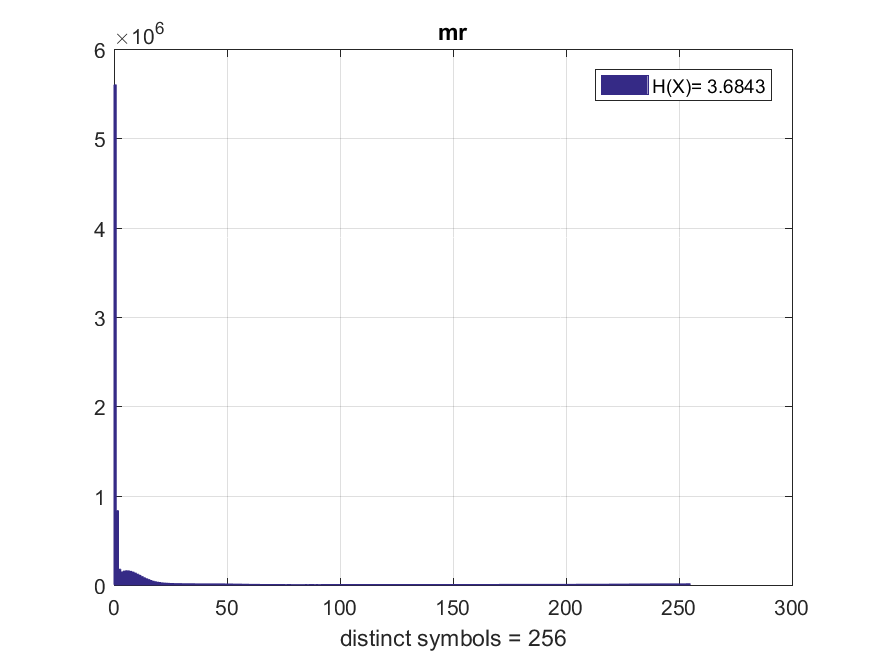
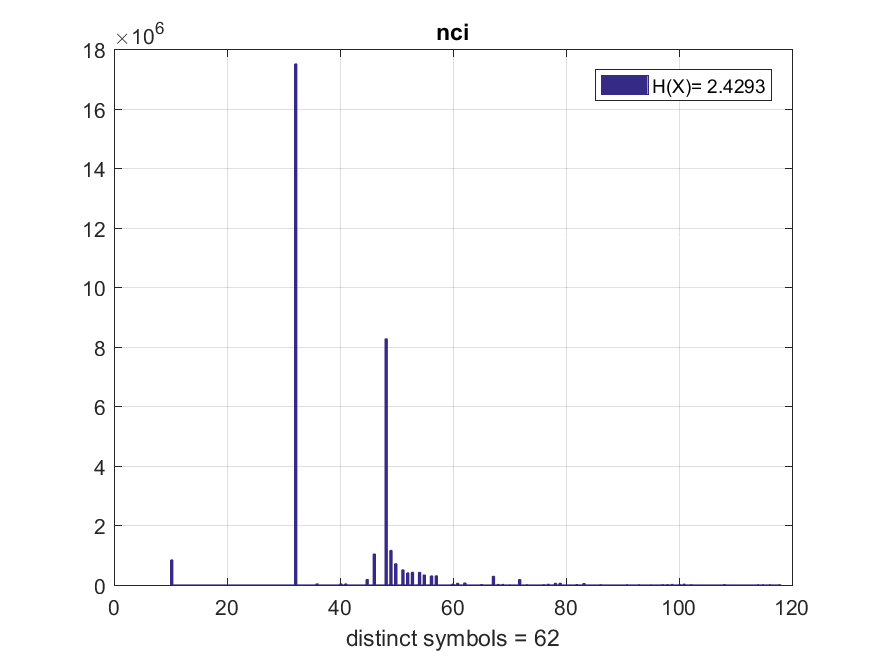


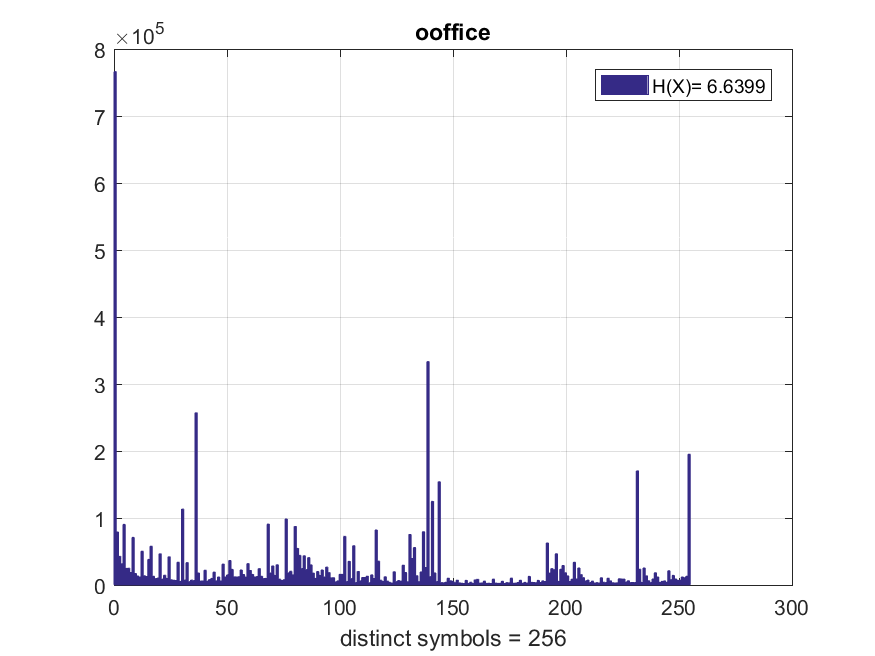


Do conjunto Canterbury foram escolhidos os ficheiros: alice29.txt, asyoulik.txt, cp.html, fields.c, gramar.lsp

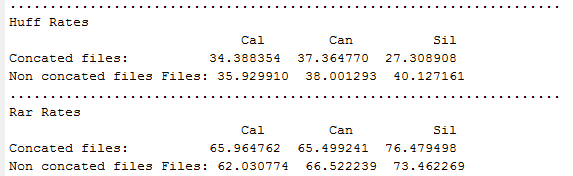


Do conjunto Silesia Corpus foram escolhidos os ficheiros: Dickens, Mozilla, mr, nci, ooffice.





**2 c)**



O codificador WinRar obtém taxas de compressão superiores ao codificador Huffman semi-adaptativo.

A concatenação dos cinco ficheiros faz com que na utilização do codificador WinRar, obtenha-se melhor taxas de compressão para o conjunto Calgary Corpus e Silesia Corpus. Com o Codificador Huffman semi-adaptativo, a concatenação dos ficheiros obtém taxas de compressão mais reduzidas. A maior diferença é no conjunto Silesia Corpus onde a taxa de compressão tem uma redução de 13%. A diferença explica-se porque os ficheiros já têm entropia elevada?