



Informação e Codificação

Projeto 3

Raquel Pinto (92948), Alexandre Oliveira (93289), Pedro Loureiro (92953)

Link do repositório - https://github.com/AlexOliZ/IC_P3.git

Na pasta html do nosso projeto pode-se observar a documentação doxygen

Para testar este projeto, tem que se correr o ficheiro fcm.cpp uma vez para cada um dos ficheiros de texto de cada língua, para isso foi feito um script chamado build_tables.sh que compila os programas fcm.cpp e test_fcm.cpp e executa o ficheiro test_fcm.cpp para as várias línguas com $k = 3$ e $\alpha = 0$. Antes de correr este script é preciso fazer unzip IC_P3.zip pois aqui escreve-se as tabelas obtidas e lê-se os ficheiros de texto.

Para compilar os programas fcm.cpp e test_fcm.cpp sem ser pelo script pode-se executar make fcm.o, ficando com todas as tabelas dos ficheiros de teste para determinado k. Assim pode-se testar a parte B deste projeto para o programa lang e para o find_lang.

Para compilar o programa lang faz-se make lang.o (que compila o ficheiro lang.cpp e test_lang.cpp).

Para compilar o programa find_lang faz-se make find_lang.o (que compila o ficheiro test_find_lang.cpp e lang.cpp). Após isto, chamou-se várias vezes o programa alterando o valor de k e de alfa.

- Quando o ficheiro é português (txt2.txt) e para $k = 3$ e $\alpha = 0.01$;

```
./en/table_3.txt
no more info in file
compression: 168342
./pt/table_3.txt
no more info in file
compression: 82158
./french/table_3.txt
no more info in file
compression: 158428
./spanish/table_3.txt
no more info in file
compression: 132325
./german/table_3.txt
no more info in file
compression: 173213
./danish/table_3.txt
no more info in file
compression: 164725

./greek/table_3.txt
no more info in file
compression: 144138
./italian/table_3.txt
no more info in file
compression: 149497
./finnish/table_3.txt
no more info in file
compression: 170734
./dutch/table_3.txt
no more info in file
compression: 164836
min_size: 82158
language: pt
language -> pt
```

Na Figura 1 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt2.txt é português como esperado.

Figura 1 - Resultados para este caso.

- Quando o ficheiro é português (txt3.txt) e para $k = 3$ e $\alpha = 0.04$;

```
./en/table_3.txt
no more info in file
compression: 129665
./pt/table_3.txt
no more info in file
compression: 67070
./french/table_3.txt
no more info in file
compression: 125243
./spanish/table_3.txt
no more info in file
compression: 106125
./german/table_3.txt
no more info in file
compression: 132748
./danish/table_3.txt
no more info in file
compression: 126852

./greek/table_3.txt
no more info in file
compression: 120691
./italian/table_3.txt
no more info in file
compression: 118706
./finnish/table_3.txt
no more info in file
compression: 133979
./dutch/table_3.txt
no more info in file
compression: 126665
min_size: 67070
language: pt
language -> pt
```

Na Figura 2 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt3.txt é português como esperado.

Figura 2 - Resultados para este caso.

- Quando o ficheiro é português (txt3.txt) e para $k = 3$ e $\alpha = 0.5$;

```
./en/table_3.txt
no more info in file
compression: 115214
./pt/table_3.txt
no more info in file
compression: 70881
./french/table_3.txt
no more info in file
compression: 111754
./spanish/table_3.txt
no more info in file
compression: 98847
./german/table_3.txt
no more info in file
compression: 117900
./danish/table_3.txt
no more info in file
compression: 115313

./greek/table_3.txt
no more info in file
compression: 120691
./italian/table_3.txt
no more info in file
compression: 108587
./finnish/table_3.txt
no more info in file
compression: 118824
./dutch/table_3.txt
no more info in file
compression: 114831
min_size: 70881
language: pt
language -> pt
```

Na Figura 3 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt3.txt é português como esperado.

Figura 3 - Resultados para este caso.

- Quando o ficheiro é inglês (txt1.txt) e para $k = 3$ e $\alpha = 0.02$;

```
./en/table_3.txt
no more info in file
compression: 4639
./pt/table_3.txt
no more info in file
compression: 10036
./french/table_3.txt
no more info in file
compression: 9870
./spanish/table_3.txt
no more info in file
compression: 10110
./german/table_3.txt
no more info in file
compression: 10287
./danish/table_3.txt
no more info in file
compression: 9671

./greek/table_3.txt
no more info in file
compression: 10409
./italian/table_3.txt
no more info in file
compression: 10015
./finnish/table_3.txt
no more info in file
compression: 11082
./dutch/table_3.txt
no more info in file
compression: 9204
min_size: 4639
language: en
language -> en
```

Na Figura 4 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt1.txt é inglês como esperado.

Figura 4 - Resultados para este caso.

- Quando o ficheiro é inglês (txt3.txt) e para $k = 3$ e $\alpha = 0.04$;

```

./en/table_3.txt
no more info in file
compression: 4668
./pt/table_3.txt
no more info in file
compression: 9716
./french/table_3.txt
no more info in file
compression: 9576
./spanish/table_3.txt
no more info in file
compression: 9801
./german/table_3.txt
no more info in file
compression: 9972
./danish/table_3.txt
no more info in file
compression: 9400

./greek/table_3.txt
no more info in file
compression: 10409
./italian/table_3.txt
no more info in file
compression: 9703
./finnish/table_3.txt
no more info in file
compression: 10687
./dutch/table_3.txt
no more info in file
compression: 8970
min_size: 4668
language: en
language -> en

```

Na Figura 5 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt3.txt é inglês como esperado.

Figura 5 - Resultados para este caso.

- Quando o ficheiro é inglês (txt6.txt) e para $k = 3$ e $\alpha = 1$;

```

./en/table_3.txt
no more info in file
compression: 6808
./pt/table_3.txt
no more info in file
compression: 11942
./french/table_3.txt
no more info in file
compression: 11457
./spanish/table_3.txt
no more info in file
compression: 12137
./german/table_3.txt
no more info in file
compression: 12477
./danish/table_3.txt
no more info in file
compression: 12103

./greek/table_3.txt
no more info in file
compression: 15150
./italian/table_3.txt
no more info in file
compression: 12107
./finnish/table_3.txt
no more info in file
compression: 13749
./dutch/table_3.txt
no more info in file
compression: 11525
min_size: 6808
language: en
language -> en

```

Na Figura 6 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt6.txt é inglês como esperado.

Figura 6 - Resultados para este caso.

- Quando o ficheiro é espanhol (txt2.txt) e para $k = 3$ e $\alpha = 0.02$;

```

./en/table_3.txt
no more info in file
compression: 91299
./pt/table_3.txt
no more info in file
compression: 69224
./french/table_3.txt
no more info in file
compression: 87853
./spanish/table_3.txt
no more info in file
compression: 45144
./german/table_3.txt
no more info in file
compression: 95985
./danish/table_3.txt
no more info in file
compression: 91417

./greek/table_3.txt
no more info in file
compression: 86419
./italian/table_3.txt
no more info in file
compression: 81501
./finnish/table_3.txt
no more info in file
compression: 96073
./dutch/table_3.txt
no more info in file
compression: 91324
min_size: 45144
language: spanish
language -> spanish

```

Na Figura 7 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt2.txt é espanhol como esperado.

Figura 7 - Resultados para este caso.

- Quando o ficheiro é espanhol (txt3.txt) e para $k = 3$ e $\alpha = 0.3$;

```
./en/table_3.txt
no more info in file
compression: 140579
./pt/table_3.txt
no more info in file
compression: 121362
./french/table_3.txt
no more info in file
compression: 133996
./spanish/table_3.txt
no more info in file
compression: 88853
./german/table_3.txt
no more info in file
compression: 143608
./danish/table_3.txt
no more info in file
compression: 140281

./greek/table_3.txt
no more info in file
compression: 144831
./italian/table_3.txt
no more info in file
compression: 129659
./finnish/table_3.txt
no more info in file
compression: 144690
./dutch/table_3.txt
no more info in file
compression: 140089
min_size: 88853
language: spanish
language -> spanish
```

Na Figura 8 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt3.txt é espanhol como esperado.

Figura 8 - Resultados para este caso.

- Quando o ficheiro é espanhol (txt4.txt) e para $k = 3$ e $\alpha = 4$;

```
./en/table_3.txt
no more info in file
compression: 125089
./pt/table_3.txt
no more info in file
compression: 112016
./french/table_3.txt
no more info in file
compression: 122820
./spanish/table_3.txt
no more info in file
compression: 95834
./german/table_3.txt
no more info in file
compression: 130230
./danish/table_3.txt
no more info in file
compression: 129743

./greek/table_3.txt
no more info in file
compression: 138593
./italian/table_3.txt
no more info in file
compression: 120714
./finnish/table_3.txt
no more info in file
compression: 130449
./dutch/table_3.txt
no more info in file
compression: 129364
min_size: 95834
language: spanish
language -> spanish
```

Na Figura 9 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt4.txt é espanhol como esperado.

Figura 9 - Resultados para este caso.

- Quando o ficheiro é dinamarquês (txt1.utf8) e para $k = 3$ e $\alpha = 0.32$;

```
./en/table_3.txt
no more info in file
compression: 320629
./pt/table_3.txt
no more info in file
compression: 328934
./french/table_3.txt
no more info in file
compression: 327971
./spanish/table_3.txt
no more info in file
compression: 328290
./german/table_3.txt
no more info in file
compression: 313654
./danish/table_3.txt
no more info in file
compression: 245643

./greek/table_3.txt
no more info in file
compression: 320317
./italian/table_3.txt
no more info in file
compression: 325753
./finnish/table_3.txt
no more info in file
compression: 321503
./dutch/table_3.txt
no more info in file
compression: 311246
min_size: 245643
language: danish
language -> danish
```

Na Figura 10 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt1.utf8 é dinamarquês como esperado.

Figura 10 - Resultados para este caso.

- Quando o ficheiro é dinamarquês (txt1.utf8) e para $k = 3$ e $\alpha = 0.1$;

```

./en/table_3.txt
no more info in file
compression: 329911
./pt/table_3.txt
no more info in file
compression: 339208
./french/table_3.txt
no more info in file
compression: 338055
./spanish/table_3.txt
no more info in file
compression: 337874
./german/table_3.txt
no more info in file
compression: 321572
./danish/table_3.txt
no more info in file
compression: 236108

./greek/table_3.txt
no more info in file
compression: 320317
./italian/table_3.txt
no more info in file
compression: 334128
./finnish/table_3.txt
no more info in file
compression: 330560
./dutch/table_3.txt
no more info in file
compression: 316992
min_size: 236108
language: danish
language -> danish

```

Figura 11 - Resultados para este caso.

Na Figura 11 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt1.utf8 é dinamarquês como esperado.

- Quando o ficheiro é dinamarquês (txt1.utf8) e para $k = 3$ e $\alpha = 2.1$;

```

./en/table_3.txt
no more info in file
compression: 307109
./pt/table_3.txt
no more info in file
compression: 314091
./french/table_3.txt
no more info in file
compression: 313501
./spanish/table_3.txt
no more info in file
compression: 314411
./german/table_3.txt
no more info in file
compression: 304016
./danish/table_3.txt
no more info in file
compression: 271691

./greek/table_3.txt
no more info in file
compression: 320317
./italian/table_3.txt
no more info in file
compression: 313994
./finnish/table_3.txt
no more info in file
compression: 310030
./dutch/table_3.txt
no more info in file
compression: 305597
min_size: 271691
language: danish
language -> danish

```

Figura 12 - Resultados para este caso.

Na Figura 12 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt1.utf8 é dinamarquês como esperado.

- Quando o ficheiro é português (txt2.txt) e para $k = 2$ e $\alpha = 0.01$;

```

./en/table_2.txt
no more info in file
compression: 48088
./pt/table_2.txt
no more info in file
compression: 31097
./french/table_2.txt
no more info in file
compression: 43646
./spanish/table_2.txt
no more info in file
compression: 39918
./german/table_2.txt
no more info in file
compression: 47932
./danish/table_2.txt
no more info in file
compression: 47219

./greek/table_2.txt
no more info in file
compression: 49101
./italian/table_2.txt
no more info in file
compression: 44786
./finnish/table_2.txt
no more info in file
compression: 47180
./dutch/table_2.txt
no more info in file
compression: 47192
min_size: 31097
language: pt
language -> pt

```

Na Figura 13 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt2.txt é português como esperado.

Figura 13 - Resultados para este caso.

- Quando o ficheiro é português (txt3.txt) e para $k = 2$ e $\alpha = 0.04$;

```

./en/table_2.txt
no more info in file
compression: 38849
./pt/table_2.txt
no more info in file
compression: 26498
./french/table_2.txt
no more info in file
compression: 36569
./spanish/table_2.txt
no more info in file
compression: 33870
./german/table_2.txt
no more info in file
compression: 38740
./danish/table_2.txt
no more info in file
compression: 38133

./greek/table_2.txt
no more info in file
compression: 37876
./italian/table_2.txt
no more info in file
compression: 37265
./finnish/table_2.txt
no more info in file
compression: 38917
./dutch/table_2.txt
no more info in file
compression: 37943
min_size: 26498
language: pt
language -> pt

```

Na Figura 14 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt3.txt é português como esperado.

Figura 14 - Resultados para este caso.

- Quando o ficheiro é português (txt3.txt) e para $k = 2$ e $\alpha = 0.5$;

```

./en/table_2.txt
no more info in file
compression: 35544
./pt/table_2.txt
no more info in file
compression: 27056
./french/table_2.txt
no more info in file
compression: 34211
./spanish/table_2.txt
no more info in file
compression: 32263
./german/table_2.txt
no more info in file
compression: 35615
./danish/table_2.txt
no more info in file
compression: 35287

./greek/table_2.txt
no more info in file
compression: 34146
./italian/table_2.txt
no more info in file
compression: 34800
./finnish/table_2.txt
no more info in file
compression: 35761
./dutch/table_2.txt
no more info in file
compression: 35133
min_size: 27056
language: pt
language -> pt

```

Na Figura 15 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt3.txt é português como esperado.

Figura 15 - Resultados para este caso.

- Quando o ficheiro é inglês (txt1.txt) e para $k = 2$ e $\alpha = 0.02$;

```

./en/table_2.txt
no more info in file
compression: 3655
./pt/table_2.txt
no more info in file
compression: 5318
./french/table_2.txt
no more info in file
compression: 5435
./spanish/table_2.txt
no more info in file
compression: 5414
./german/table_2.txt
no more info in file
compression: 5469
./danish/table_2.txt
no more info in file
compression: 5267

./greek/table_2.txt
no more info in file
compression: 6358
./italian/table_2.txt
no more info in file
compression: 5486
./finnish/table_2.txt
no more info in file
compression: 5901
./dutch/table_2.txt
no more info in file
compression: 4948
min_size: 3655
language: en
language -> en

```

Na Figura 16 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt1.txt é inglês como esperado.

Figura 16 - Resultados para este caso.

- Quando o ficheiro é inglês (txt3.txt) e para $k = 2$ e $\alpha = 0.04$;

```

./en/table_2.txt
no more info in file
compression: 3659
./pt/table_2.txt
no more info in file
compression: 5238
./french/table_2.txt
no more info in file
compression: 5349
./spanish/table_2.txt
no more info in file
compression: 5328
./german/table_2.txt
no more info in file
compression: 5397
./danish/table_2.txt
no more info in file
compression: 5193

./greek/table_2.txt
no more info in file
compression: 6089
./italian/table_2.txt
no more info in file
compression: 5393
./finnish/table_2.txt
no more info in file
compression: 5798
./dutch/table_2.txt
no more info in file
compression: 4891
min_size: 3659
language: en
language -> en

```

Na Figura 17 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt3.txt é inglês como esperado.

Figura 17 - Resultados para este caso.

- Quando o ficheiro é inglês (txt6.txt) e para $k = 2$ e $\alpha = 1$;

```

./en/table_2.txt
no more info in file
compression: 5034
./pt/table_2.txt
no more info in file
compression: 6904
./french/table_2.txt
no more info in file
compression: 6732
./spanish/table_2.txt
no more info in file
compression: 6963
./german/table_2.txt
no more info in file
compression: 7163
./danish/table_2.txt
no more info in file
compression: 6975

./greek/table_2.txt
no more info in file
compression: 7885
./italian/table_2.txt
no more info in file
compression: 7072
./finnish/table_2.txt
no more info in file
compression: 7781
./dutch/table_2.txt
no more info in file
compression: 6672
min_size: 5034
language: en
language -> en

```

Na Figura 18 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt6.txt é inglês como esperado.

Figura 18 - Resultados para este caso.

- Quando o ficheiro é espanhol (txt2.txt) e para $k = 2$ e $\alpha = 0.02$;

```

./en/table_2.txt
no more info in file
compression: 27726
./pt/table_2.txt
no more info in file
compression: 22591
./french/table_2.txt
no more info in file
compression: 25997
./spanish/table_2.txt
no more info in file
compression: 18759
./german/table_2.txt
no more info in file
compression: 28348
./danish/table_2.txt
no more info in file
compression: 27638

./greek/table_2.txt
no more info in file
compression: 29131
./italian/table_2.txt
no more info in file
compression: 26487
./finnish/table_2.txt
no more info in file
compression: 27844
./dutch/table_2.txt
no more info in file
compression: 27719
min_size: 18759
language: spanish
language -> spanish

```

Na Figura 19 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt2.txt é espanhol como esperado.

Figura 19 - Resultados para este caso.

- Quando o ficheiro é espanhol (txt3.txt) e para $k = 2$ e $\alpha = 0.3$;

```

./en/table_2.txt
no more info in file
compression: 40917
./pt/table_2.txt
no more info in file
compression: 36854
./french/table_2.txt
no more info in file
compression: 38980
./spanish/table_2.txt
no more info in file
compression: 31553
./german/table_2.txt
no more info in file
compression: 41071
./danish/table_2.txt
no more info in file
compression: 40427

```

```

./greek/table_2.txt
no more info in file
compression: 38923
./italian/table_2.txt
no more info in file
compression: 38990
./finnish/table_2.txt
no more info in file
compression: 41178
./dutch/table_2.txt
no more info in file
compression: 40588
min_size: 31553
language: spanish
language -> spanish

```

Na Figura 20 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt3.txt é espanhol como esperado.

Figura 20 - Resultados para este caso.

- Quando o ficheiro é espanhol (txt4.txt) e para $k = 2$ e $\alpha = 4$;

```

./en/table_2.txt
no more info in file
compression: 36824
./pt/table_2.txt
no more info in file
compression: 34082
./french/table_2.txt
no more info in file
compression: 36217
./spanish/table_2.txt
no more info in file
compression: 31922
./german/table_2.txt
no more info in file
compression: 37193
./danish/table_2.txt
no more info in file
compression: 37140

```

```

./greek/table_2.txt
no more info in file
compression: 37942
./italian/table_2.txt
no more info in file
compression: 36570
./finnish/table_2.txt
no more info in file
compression: 37393
./dutch/table_2.txt
no more info in file
compression: 37046
min_size: 31922
language: spanish
language -> spanish

```

Na Figura 21 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt4.txt é espanhol como esperado.

Figura 21 - Resultados para este caso.

- Quando o ficheiro é dinamarquês (txt1.utf8) e para $k = 2$ e $\alpha = 0.32$;

```

./en/table_2.txt
no more info in file
compression: 81559
./pt/table_2.txt
no more info in file
compression: 82813
./french/table_2.txt
no more info in file
compression: 84202
./spanish/table_2.txt
no more info in file
compression: 83491
./german/table_2.txt
no more info in file
compression: 78526
./danish/table_2.txt
no more info in file
compression: 67203

```

```

./greek/table_2.txt
no more info in file
compression: 73019
./italian/table_2.txt
no more info in file
compression: 83624
./finnish/table_2.txt
no more info in file
compression: 81668
./dutch/table_2.txt
no more info in file
compression: 78682
min_size: 67203
language: danish
language -> danish

```

Na Figura 22 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt1.utf8 é dinamarquês como esperado.

Figura 22 - Resultados para este caso.

- Quando o ficheiro é dinamarquês (txt1.utf8) e para $k = 2$ e $\alpha = 0.1$;

```

./en/table_2.txt
no more info in file
compression: 85123
./pt/table_2.txt
no more info in file
compression: 86769
./french/table_2.txt
no more info in file
compression: 88360
./spanish/table_2.txt
no more info in file
compression: 87615
./german/table_2.txt
no more info in file
compression: 81226
./danish/table_2.txt
no more info in file
compression: 65460
./greek/table_2.txt
no more info in file
compression: 75506
./italian/table_2.txt
no more info in file
compression: 87495
./finnish/table_2.txt
no more info in file
compression: 85388
./dutch/table_2.txt
no more info in file
compression: 81300
min_size: 65460
language: danish
language -> danish

```

Na Figura 23 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt1.utf8 é dinamarquês como esperado.

Figura 23 - Resultados para este caso.

- Quando o ficheiro é dinamarquês (txt1.utf8) e para $k = 2$ e $\alpha = 2.1$;

```

./en/table_2.txt
no more info in file
compression: 77415
./pt/table_2.txt
no more info in file
compression: 78414
./french/table_2.txt
no more info in file
compression: 79314
./spanish/table_2.txt
no more info in file
compression: 79012
./german/table_2.txt
no more info in file
compression: 75469
./danish/table_2.txt
no more info in file
compression: 70088
./greek/table_2.txt
no more info in file
compression: 73365
./italian/table_2.txt
no more info in file
compression: 79618
./finnish/table_2.txt
no more info in file
compression: 77643
./dutch/table_2.txt
no more info in file
compression: 75922
min_size: 70088
language: danish
language -> danish

```

Na Figura 24 o programa percorreu todas as tabelas criadas, calculou o número estimado de bits necessários para comprimir o ficheiro e viu qual o mais pequeno. Chegando à conclusão que o ficheiro txt1.utf8 é dinamarquês como esperado.

Figura 24 - Resultados para este caso.

Conclusão

Como podemos ver é possível encontrar a linguagem de um certo ficheiro através da entropia usando um certo modelo para a linguagem, contudo é preciso ter cuidado e encontrar textos adequados para a construção do modelo para este ter resultados válidos. Após construir o modelo podemos comparar o texto com todos os modelos e usar aquele que usa o menor número de bits na sua compressão.