

INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA
LICENCIATURA EM ENGENHARIA INFORMÁTICA E DE COMPUTADORES
COMUNICAÇÃO DIGITAL

Verão 2021/2022 - Trabalho Prático (Módulo 1)

Data de publicação: 28 de março de 2022

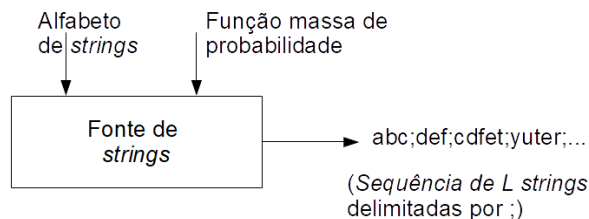
Data de entrega: 9 de maio de 2022

Objetivos:

- Desenvolvimento de programas e aplicações em linguagem 'C' e 'Python'.
- Estudo e aplicação de conceitos fundamentais sobre SCD, teoria de informação e codificação de fonte.

*O código desenvolvido e o respetivo relatório deverão ser entregues em formato eletrónico no sistema Moodle.
A apresentação da resolução dos vários exercícios decorrerá em aula a definir em cada turma.*

1. Exercícios 4 e 5 apresentados no guia da primeira aula prática.
 - (a) Entregue o código desenvolvido para a implementação de cada uma das funções e os respetivos programas de teste.
 - (b) Usando os programas de teste, apresente resultados experimentais que comprovem o correto funcionamento das funções desenvolvidas.
2. **(Python)** A figura apresenta a fonte de *strings* com alfabeto e função massa de probabilidade (FMP) genéricos.



As sequências de L strings produzidas pela fonte deverão ser escritas num ficheiro de texto, sendo delimitadas pelo carater ';'.

- (a) Implemente a fonte de *strings*, de acordo com as especificações indicadas acima. Considere ainda que a implementação deverá apresentar: o valor da entropia do ficheiro gerado; o histograma do ficheiro gerado.
- (b) Utilize convenientemente a implementação, para gerar as seguintes sequências de símbolos:
 - (i) **Geração automática de palavras-passe** com número variável de caracteres (entre o mínimo e o máximo estabelecidos). A palavra passe deverá ser alfanumérica, contendo obrigatoriamente caracteres minúsculos, maiúsculos, símbolos de pontuação e algarismos.
 - (ii) **Geração automática de sequências alfanuméricas**, ou seja, algarismos decimais e letras maiúsculas, com $L = 24$, semelhantes a chaves de ativação e registo de software (por exemplo: RTY9 GHUI 1JER 82TY SGJP IUDS).

Apresente resultados experimentais (cinco ficheiros com palavras-passe e cinco ficheiros com sequências alfanuméricas) obtidos na implementação destas funcionalidades. Proponha uma função para medir a robustez da palavra-passe e aplique-a aos ficheiros gerados. Comente os resultados.

- (c) Utilize convenientemente a implementação, para a **geração automática de conteúdos de tabelas a utilizar num sistema de informação** de um jogo de sorte. Pretende-se preencher duas tabelas com dados, gerados aleatoriamente, sobre indivíduos (apostadores) e as respetivas apostas. Por cada indivíduo, teremos a seguinte informação:

Número de Cidadão | Nome(s) Próprio(s) e Apelido(s) | Concelho de Residência | Profissão

A tabela com as apostas terá registos na forma

Número de Cidadão | Aposta | Data

O número de cidadão (8 algarismos decimais) é único e podem existir nomes iguais. Recorra aos ficheiros Nomes.txt, Apellidos.txt, Concelhos.txt e Profissões.txt para gerar os diferentes campos do registo. Para a geração do nome completo, considere: um nome próprio e um apelido, dois nomes próprios e um apelido, dois nomes próprios e dois apelidos. A aposta é formada por 5 números inteiros (diferentes) no intervalo $\{1, \dots, 50\}$ seguidos de outros 2 números inteiros (diferentes) no intervalo $\{1, \dots, 11\}$. A data tem a forma dd-mm-aaaa.

Apresente resultados experimentais em ficheiros de texto, com mais de 1000 registos.

3. (C) Considere o código unário (*comma code*).
- (a) Implemente o par codificador/descodificador de código unário, a funcionar em modo semi-adaptativo.
 - (b) Apresente resultados experimentais que comprovem o correto funcionamento do par codificador/descodificador, recorrendo a ficheiros do conjunto `CD_TestFiles.zip`.
 - (c) Para ficheiros produzidos pela fonte do exercício 2, alínea c), apresente a taxa de compressão obtida. Apresente resultados experimentais tais que comprovem o cumprimento do primeiro Teorema de Shannon.
4. (C/Python) A função `LZ77_Tokenizer` descreve o ficheiro de entrada através de *tokens* LZ77, obtidos com *search window* (SW) e *look-ahead-buffer* (LAB) de dimensões indicadas pelo utilizador. A aplicação:
- (i) Gera um ficheiro de texto com os *tokens* obtidos no processamento do ficheiro de entrada, sendo que cada linha contém um *token*.
 - (ii) Apresenta os histogramas e a entropia dos campos *position* e *length*.
- (a) Implemente a função `LZ77_Tokenizer`.
 - (b) Apresente resultados experimentais obtidos com ficheiros do conjunto `CD_TestFiles.zip` e gerados pela fonte do exercício 2, alínea c). Comente os resultados obtidos.