

Algoritmos Avançados

2020/2021 — 1º Semestre

2º Trabalho — Algoritmos Probabilísticos

Data limite de entrega: 10 de janeiro de 2020

Hipótese A – Problemas Combinatórios

Na sequência do trabalho anterior, em que foram desenvolvidos algoritmos exaustivos para a resolução de problemas combinatórios, pode ser agora desenvolvida uma **abordagem probabilística para os mesmos problemas**.

Assim, para o mesmo problema que analisou no trabalho anterior, desenvolva e teste um algoritmo probabilístico para o resolver.

De seguida, analise o esforço computacional realizado pelo algoritmo desenvolvido. Para isso deve:

- Analisar a complexidade do algoritmo.
- Realizar uma sequência de testes com instâncias sucessivamente maiores do problema. E registar e analisar: (1) o tempo de execução e (2) o número soluções / configurações testadas.
- Comparar os resultados obtidos nos dois passos anteriores.
- Comparar o desempenho das estratégias exaustiva e probabilística.
- Escrever um relatório sucinto (máx. 6 págs.).

Hipótese B – Contagem Aproximada de Ocorrências – Cadeias de Caracteres Aleatórios

Pretende-se que sejam criados métodos para contar o **número de ocorrência de caracteres** em cadeias de caracteres aleatórios.

Os métodos a criar devem permitir contar os caracteres com **três tipos de contadores: exatos, aproximados com probabilidade fixa e aproximados com probabilidade decrescente logarítmica**, ou usando o **contador de vírgula flutuante de Csuros**.

Atenção:

- As bases dos contadores a usar deverão ser escolhidas com o docente da unidade curricular.
- As cadeias de teste usadas deverão ser geradas escolhendo aleatoriamente caracteres da cadeia de caracteres que representa o seu nome – em que todos os caracteres são **minúsculos e sem espaços**. Por exemplo: “antóniomanueldasilvaferreira”.

- **Deve ser assegurado que alguns dos caracteres têm um número de ocorrências significativamente maior que outros.**
- Crie cadeias de teste com 100, 1000, 10000, 1000000, etc. caracteres.

Deverá ser feita uma análise da eficiência computacional e das limitações dos métodos desenvolvidos. Para isso deve:

- a) Realizar um conjunto de testes.
- b) Comparar o desempenho dos métodos de contagem.
- c) Escrever um relatório sucinto (máx. 6 págs.).

O que fazer:

Obter as frequências de ocorrência de todas os **caracteres distintos** que surgem em cada uma das cadeias de teste.

Os resultados dos contadores aproximados devem ser comparados com os resultados dos contadores exatos – por exemplo: máximo erro relativo, mínimo erro relativo, média do erro relativo, etc.

Hipótese C – Contagem Aproximada de Ocorrências – Palavras em Ficheiros de Texto

Pretende-se que sejam criados métodos para contar o **número de ocorrência de palavras** em cadeias ficheiros de texto.

Como ficheiros de teste serão usados ficheiro de texto da mesma obra literária e comparada uma edição na língua original com uma ou mais traduções.

Sugestão, obtenha os ficheiros de texto a partir do Projeto Gutenberg: <http://www.gutenberg.org>

Os métodos a criar devem permitir contar as palavras com **três tipos de contadores: exatos, aproximados com probabilidade fixa e aproximados com probabilidade decrescente logarítmica**, ou usando o **contador de vírgula flutuante de Csuros**.

Atenção:

- **As bases dos contadores a usar deverão ser escolhidas com o docente da unidade curricular.**
- Remova os cabeçalhos dos ficheiros de texto.
- Não conte sinais de pontuação nem palavras simples (artigos definidos e indefinidos, etc.).

Deverá ser feita uma análise da eficiência computacional e das limitações dos métodos desenvolvidos. Para isso deve:

- a) Realizar um conjunto de testes.

- b) Comparar o desempenho dos métodos de contagem.
- c) Escrever um relatório sucinto (máx. 6 págs.).

O que fazer:

Obter as frequências de ocorrência de todas as **palavras distintas** que surgem em cada um dos ficheiros de teste.

Os resultados dos contadores aproximados devem ser comparados com os resultados dos contadores exatos – por exemplo: máximo erro relativo, mínimo erro relativo, média do erro relativo, etc.

Obtenha algumas conclusões por comparação das palavras mais frequentes em cada uma das versões da mesma obra.