

Algoritmos de Ordenação Interna

Frank Coelho de Alcantara

Ordenação

Entre os algoritmos mais importantes destacam-se os algoritmos de ordenação, em inglês *sort*, isto porque a maior parte das tarefas diárias envolve alguma forma de ordenação.

Ordenamos registros em bancos de dados, as tarefas do dia, conhecimento em parágrafos, parágrafos em páginas e páginas em sites. Poderíamos estender esta lista para lembrar que ordenamos desde muito cedo, antes mesmo de aprender a contar. Esta tarefa, a ordenação está emaranhada com o desenvolvimento da humanidade. A própria civilização, como conhecemos, deve seu início à agricultura e esta à percepção da ordem dos ciclos solares. Ao longo da história os processos de ordenação tiveram impacto na ascensão e queda de impérios e empresas.

O *Google* deve seu sucesso a uma lista ordenada criada a partir de um algoritmo de classificação, o *PageRank*. Foi este algoritmo desenvolvido por Larry Page and Sergey Brin que permitiu que o *Google* retornasse uma lista de páginas na ordem que melhor atenda as demandas de busca na internet e com isso se tornasse uma empresa quase onipresente. A missão do *Google* era, ou é, indexar todo o conhecimento da humanidade mas, seu sucesso comercial se deve a ordem que este conhecimento é apresentado.

Com tal envolvimento na vida diária, uma parte importante das tarefas computacionais incluem alguma forma de ordenação. E, graças a isso, os estes algoritmos contínua e profundamente estudados. Em busca de mais eficiência para os algoritmos já conhecidos e na esperança de criação de algoritmos novos. Uma das primeiras regras de otimização, tarefa diária de todos os envolvidos com eficiência, diz que é melhor começar otimizar pelos algoritmos de ordenação. Não bastasse isso, como justificativa para o estudo dos algoritmos de ordenação ainda é necessário ressaltar que os algoritmos de ordenação servem como exemplo do uso das técnicas de desenvolvimento de algoritmos como: dividir-e-conquistar, estruturas de dados compostas e uso de aleatoriedade na criação de algoritmos.

Fora do mundo acadêmico é raro utilizarmos um algoritmo para ordenar uma lista de números, palavras ou frases. O comum é que o dado que precisa ser ordenado esteja representado por uma estrutura composta de vários campos, *record* com

um ou mais campos chaves, a partir dos quais realizaremos a ordenação. Este é o caso típico que encontramos em bancos de dados sejam eles relacionais, documentais ou gráficos.

Problema: dado um conjunto de dados qualquer, de que forma podemos ordenar estes dados em uma sequência determinada.

Conceito

A escolha do algoritmo para a solução de um determinado problema depende do problema, dos recursos disponíveis e da característica dos dados.

Nem sempre o algoritmo mais eficiente é a melhor escolha. Pode ser que o custo envolvido na implementação do algoritmo, e os recursos necessários para garantir que ele realmente seja o mais rápido não compense o benefício adquirido. Isto é mais perceptível quando a lista de itens que precisa ser ordenada é pequena ou quando esta lista tem um nível de entropia baixo.

Algoritmos diferentes terão eficiência diferente em dados diferentes. Isso quer dizer que um algoritmo pode ser muito eficiente em casos onde os dados estão arranjados em um conjunto com pouca entropia e não ter utilidade nenhuma em casos onde os dados estão na ordem inversa da desejada. A entropia, neste conceito se refere ao percentual de dados que está fora da ordem desejada. Dados que já se encontrem na ordem desejada terão entropia zero dados totalmente ordenados na ordem invertida, terão entropia 1.

Exemplo visual

Um dos benefícios da internet é poder ver animações de conceitos. Neste contexto Viktor Bohush publicou no YouTube o vídeo *Visualization and Comparison of Sorting Algorithms* que permite uma comparação visual de nove algoritmos de ordenação que pode ser visto na Figura 1.

Nesta visualização Bohoshu optou por testes diferentes: dados randômicamente organizados, onde, para determinarmos a entropia precisaríamos de acesso a estes dados (*random*); dados cujo conjunto é formado por repetições do mesmo valor, com pouca diferenciação (*few unique*); dados com entropia 1 (*reversed*), que estão na ordem reversa e, finalmente, dados que estão quase alinhados, entropia muito baixa (*almost sorted*). Em alguns casos, o shell sort se destaca, em outros o merge sort ou o quick sort, até o insertion sort teve seu momento de glória.

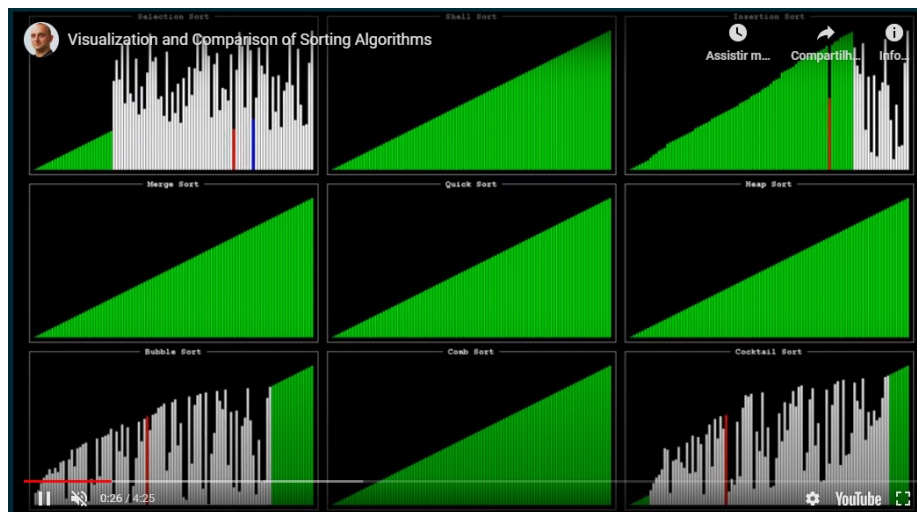


Figure 1: Fonte:(BOHUSH,2014) Animação de algoritmos de ordenção