

Ordenação: Shellsort

Algoritmos e Estruturas de Dados II

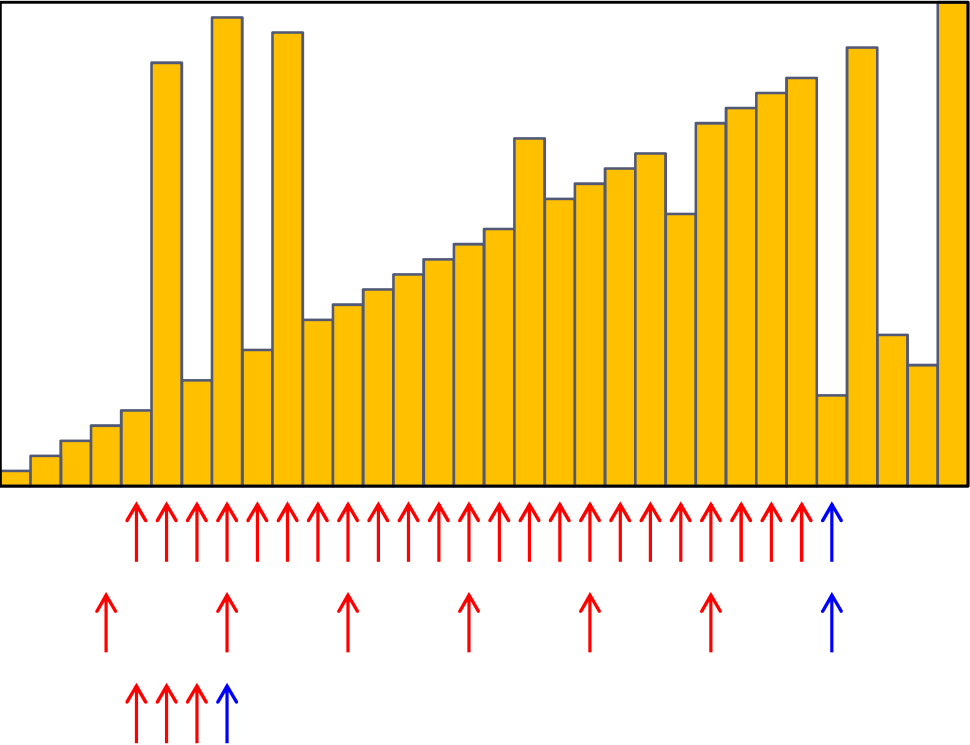
Introdução

- ▶ Proposto por Donald Shell em 1959
- ▶ Extensão do algoritmo de ordenação por inserção

Motivação

- ▶ Ordenação por inserção só troca itens adjacentes para determinar o ponto de inserção
 - ▶ São efetuadas $n-1$ comparações e movimentações quando o menor item está na última posição
- ▶ O método de Shell contorna este problema **permitindo trocas de registros distantes**

Exemplo



Algoritmo

- ▶ Os conjuntos de itens separados de h posições são ordenados
 - ▶ O elemento na posição x é comparado (e trocado) com o elemento na posição $x-h$
 - ▶ O vetor resultante é composto de h arquivos ordenados e entrelaçados
- ▶ O vetor (sequência) é dito estar h -ordenado
- ▶ Quando $h = 1$, o algoritmo é equivalente ao algoritmo de inserção

Exemplo: Shellsort

E	X	E	M	P	L	O
E	X	E	M	P	L	O
E	L	E	M	P	X	O
E	L	E	M	P	X	O

$h = 4$

E	L	E	M	P	X	O
E	L	E	M	P	X	O
E	L	E	M	P	X	O
E	L	E	M	P	X	O
E	L	E	M	P	X	O
E	L	E	M	O	X	P

$h = 2$

$h = 1$ (inserção)

E	L	E	M	O	X	P
E	L	E	M	O	X	P
E	L	E	M	O	X	P
E	E	L	M	O	X	P
E	E	L	M	O	X	P
E	L	E	M	O	X	P
E	L	E	M	O	X	P
E	L	E	M	O	P	X

Escolha da distância de salto h

- ▶ Qualquer sequência terminando com $h = 1$ garante ordenação correta ($h = 1$ é ordenação por inserção)
- ▶ Forte impacto no desempenho do algoritmo

- ▶ Sequência para h :

$$h(s) = 1, \text{ para } s = 1$$

$$h(s) = 3h(s-1) + 1, \text{ para } s > 1$$

- ▶ A sequência corresponde a 1, 4, 13, 40, 121, 364, 1093, 3280, ...
- ▶ Knuth (1973, p. 95) mostrou experimentalmente que esta sequência é difícil de ser batida por mais de 20% em eficiência
 - ▶ Outras escolhas são possíveis

Escolha da distância de salto h

- ▶ Qualquer sequência terminando com $h = 1$ garante ordenação correta
 - ▶ $h = 1$ é ordenação por inserção
- ▶ Forte impacto no desempenho do algoritmo
- ▶ Exemplo de sequência ruim: 1, 2, 4, 8, 16, ...
 - ▶ Não compara elementos em posições pares com elementos em posições ímpares até a última iteração

Escolha da distância de salto h

- ▶ Sequência para h :

$$h(s) = 1, \text{ para } s = 1$$

$$h(s) = 3h(s-1) + 1, \text{ para } s > 1$$

- ▶ A sequência corresponde a 1, 4, 13, 40, 121, 364, 1093, 3280, ...
- ▶ Knuth (1973, p. 95) mostrou experimentalmente que esta sequência é difícil de ser batida por mais de 20% em eficiência
 - ▶ Outras sequências têm desempenho similar

Shellsort

```
void shellsort(struct item *v, int n) {
    int i, j, h;
    struct item aux;

    for(h = 1; h < n; h = 3*h+1); /* calcula o h inicial. */

    while(h > 0) {
        h = (h-1)/3; /* atualiza o valor de h. */
        for(i = h; i < n; i++) {
            aux = v[i];
            j = i;
            /* efetua comparações entre elementos com distância h: */
            while(v[j - h].chave > aux.chave) {
                v[j] = v[j - h];
                j -= h;
                if(j < h) break;
            }
            v[j] = aux;
        }
    }
}
```



Shellsort – Análise

- ❑ A complexidade do algoritmo ainda não é conhecida
- ❑ Ninguém ainda foi capaz de encontrar uma fórmula fechada para sua função de complexidade
 - ❑ A sua análise contém alguns problemas matemáticos muito difíceis
 - ❑ Exemplo: escolher a sequência de incrementos
 - ❑ O que se sabe é que cada incremento não deve ser múltiplo do anterior



Shellsort – Análise

- Conjecturas referentes ao número de comparações para a seqüência de Knuth:

Conjectura 1: $C(n) = O(n^{1,25})$

Conjectura 2: $C(n) = O(n (\ln n)^2)$



Vantagens e Desvantagens

■ Vantagens:

- Shellsort é uma ótima opção para arquivos de tamanho moderado
- Sua implementação é simples e requer uma quantidade de código pequena

■ Desvantagens:

- O tempo de execução do algoritmo é sensível à ordem inicial do arquivo
- O método não é estável



Exercícios

1. Dê um exemplo de um vetor com N elementos que maximiza o número de vezes que o mínimo é atualizado no método de ordenação seleção.
2. Mostre um exemplo de entrada que demonstra que o método de ordenação seleção não é estável.
3. Mostre um exemplo que demonstra que o Shellsort é instável para sequencia $h=1,2$
4. O método da bolha não é adaptável, altere o código para que ele se torne adaptável.
5. Qual dos métodos: bolha, inserção e seleção executa menos comparações para um vetor de entrada contendo valores idênticos.