Chương 2. TÌM KIẾM & SẮP XẾP CÁC THUẬT TOÁN SẮP XẾP CƠ BẢN

ThS. Nguyễn Chí Hiếu

2017

NỘI DUNG

Giới thiệu bài toán sắp xếp

Sắp xếp chọn trực tiếp

Sắp xếp chèn trực tiếp

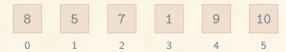
Sắp xếp đổi chỗ trực tiếp

Sắp xếp nổi bọt

Sắp xếp rung

Giới thiệu bài toán sắp xếp

- Sắp xếp là quá trình xử lý các phần tử của một danh sách, dãy số, dãy ký tự, ... theo đúng thứ tự (thỏa tiêu chuẩn nào đó).
- ► Trong bài toán sắp xếp, hai thao tác cơ bản là so sánh và gán giữa hai phần tử trong dãy.



Khái niệm nghịch thế

Định nghĩa

Xét dãy số a gồm n phần tử

$$a = \{a_0, a_1, ..., a_{n-1}\}$$

- Nếu i < j và $a_i > a_j$ thì gọi đó là một nghịch thế (trường hợp tăng dần).
- Mảng không thứ tự: chứa nghịch thế.
- Mảng có thứ tự: không chứa nghịch thế.

Ý tưởng

Cho dãy $a=\{a_0,a_1,...,a_{n-1}\}$ gồm n phần tử, giải thuật là một vòng lặp thực hiện n-1 lần.

Tại mỗi lần lặp thứ i=0,1,...,n-2

- ightharpoonup Chọn phần tử nhỏ nhất trong dãy $a_i, a_{i+1}, ..., a_{n-1}$.
- ightharpoonup Hoán vị phần tử được chọn với a_i .

```
Thuật toán 1: SelectionSort(a[], n)
- Đầu vào: mảng a gồm n phần tử.
- Đầu ra: mảng a có thứ tự tăng dần.

for i ← 0 to n - 2
    min ← i
    for j ← i + 1 to n - 1
        if a[j] < a[min]
        min ← j
    Swap(a[i], a[min])</pre>
```

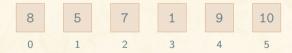
Giải thích

5

- Dòng 2: vị trí phần tử có giá trị nhỏ nhất trong mảng chưa có thứ tự.
- Dòng 6: gọi hàm hoán vị 2 giá trị.

Ví du 1

Cho dãy số a gồm 6 phần tử: 8,5,7,1,9,10. Áp dụng giải thuật sắp xếp chọn trực tiếp sắp dãy a theo thứ tự tăng dần.



Nguyễn Chí Hiếi

$$\begin{cases} i = 0, \\ j = 1, \\ min = 0 \end{cases}$$

$$\begin{cases} 8 \\ 5 \\ 7 \\ 1 \\ 9 \\ 10 \end{cases}$$

$$\begin{cases} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{cases}$$

$$\begin{cases} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{cases}$$

$$\begin{cases} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{cases}$$

$$\begin{cases} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{cases}$$

min

10

$$\begin{cases} i = 0, \\ j = 4, \\ min = 3 \end{cases}$$

8

5

5

1

2

3

9 10

min

$$\begin{cases} i = 0, \\ j = 6, \\ min = 3 \end{cases}$$
8 5 7 1 9 10

Nguyễn Chí Hiếu

min

$$\begin{cases} i = 1, \\ j = 2, \\ min = 1 \end{cases}$$
0
1
2
3
4
5

i)
min

..

$${i = 5}$$
 1 5 7 8 9 10

Đánh giá giải thuật

► Số phép so sánh

$$\sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} 1 = \sum_{i=0}^{n-2} ((n-1) - i) = \frac{n(n-1)}{2}.$$

► Mỗi lần hoán vị thực hiện 3 phép gán.

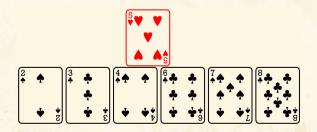
| Trường hợp | Số phép so sánh | Số hoán vị |
|-----------------------|-----------------------|------------|
| Tốt nhất | $\frac{n(n-1)}{2}$ | n-1 |
| Xấu nhất | $\frac{n(n-1)}{2}$ | n-1 |
| Độ phức tạp thời gian | $T\left(n\right) =O$ | (n^2) |

Ý tưởng

Cho dãy $a=\{a_0,a_1,...,a_{n-1}\}$ gồm n phần tử, giải thuật là một vòng lặp thực hiện n-1 lần.

Giả sử $a_0,a_1,...,a_{i-1}$ đã có thứ tự, tại mỗi lần lặp thứ i=1,2,...,n-1

▶ Chèn phần tử a_i vào đúng vị trí trong dãy $a_0,a_1,...,a_{i-1}$ để được dãy $a_0,a_1,...,a_{i-1},a_i$ có thứ tự.



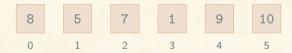
```
Thuât toán 2: InsertionSort(a[], n)
- Đầu vào: mảng a gồm n phần tử.
- Đầu ra: mảng a có thứ tự tăng dần.
  for i \leftarrow 1 to n - 1
     x \leftarrow a[i]
     pos \leftarrow i - 1
     while pos \geq 0 and a[pos] > x
       a[pos + 1] \leftarrow a[pos]
      pos \leftarrow pos - 1
     a[pos + 1] \leftarrow x
```

5

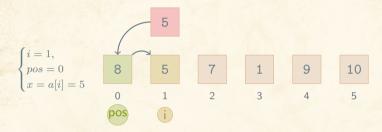
6

Ví du 2

Cho dãy số a gồm 6 phần tử: 8,5,7,1,9,10. Áp dụng giải thuật sắp xếp chèn trực tiếp sắp dãy a theo thứ tự tăng dần.



Nguyễn Chí Hiếi

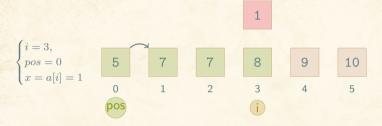




Nguyễn Chí Hiếu

Cấu trúc dữ liệu và Giải thuật 1

Nguyễn Chí Hiếu





Nguyễn Chí Hiếu

Cấu trúc dữ liệu và Giải thuật 1

9

$$\begin{cases} i = 4, \\ pos = 3 \\ x = a[i] = 9 \end{cases}$$

$$0$$

$$1$$

$$2$$

$$3$$

$$4$$

$$5$$

$$pos$$

$$i$$

10

Nguyễn Chí Hiếu

Đánh giá giải thuật

- Phép so sánh: a[pos] > x.
 - ► Trường hợp tốt nhất

$$\sum_{i=1}^{n-1} 1 = n - 1.$$

▶ Trường hợp xấu nhất

$$\sum_{i=1}^{n-1} \sum_{pos=0}^{i-1} 1 = \sum_{i=1}^{n-1} (i-1) = \frac{n(n-1)}{2}.$$

Đánh giá giải thuật

- ▶ Phép gán: x = a[i], a[pos + 1] = a[pos], a[pos + 1] = x.
 - ► Trường hợp tốt nhất

$$\sum_{i=1}^{n-1} 2 = 2(n-1).$$

► Trường hợp xấu nhất

$$\sum_{i=1}^{n-1} \left(\left(\sum_{nos=0}^{i-1} 1 \right) + 2 \right) = \frac{n(n-1)}{2} + 2(n-1) = \frac{(n-1)(n+4)}{2}.$$

Đánh giá giải thuật

| Trường hợp | Số phép so sánh | Số phép gán |
|-----------------------|---------------------------------------|------------------------|
| Tốt nhất | n-1 | 2(n-1) |
| Xấu nhất | $\frac{n(n-1)}{2}$ | $\frac{(n-1)(n+4)}{2}$ |
| Độ phức tạp thời gian | $T\left(n\right) = O\left(n^2\right)$ | |

Ý tưởng

Cho dãy $a=\{a_0,a_1,...,a_{n-1}\}$ gồm n phần tử, giải thuật là một vòng lặp thực hiện n-1 lần.

Tại mỗi lần lặp thứ i = 0, 2, ..., n-2

- ightharpoonup Tìm tất cả nghịch thế chứa phần tử a_i .
- ightharpoonup Đổi chỗ/hoán vị phần tử a_i và phần tử tương ứng trong nghịch thế.

```
Thuật toán 3: InterchangeSort(a[], n)
- Đầu vào: mảng a gồm n phần tử.
- Đầu ra: mảng a có thứ tự tăng dần.

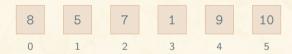
for i ← 0 to n - 2
  for j ← i + 1 to n - 1
   if a[j] < a[i]
        Swap(a[i], a[j])</pre>
```

Giải thích

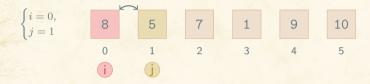
Dòng 3: kiểm tra nghịch thế giữa a[i] và a[j].

Ví du 3

Cho dãy số a gồm 6 phần tử: 8,5,7,1,9,10. Áp dụng giải thuật sắp xếp đổi chỗ trực tiếp sắp dãy a theo thứ tự tăng dần.



Nguyễn Chí Hiếu





$$\begin{cases} i = 1, \\ j = 2 \end{cases}$$
0
1
8
7
5
9
10
0
1
2
3
4
5

Đánh giá giải thuật

► Số phép so sánh

$$\sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} 1 = \sum_{i=0}^{n-2} ((n-1) - i) = \frac{n(n-1)}{2}.$$

► Mỗi lần hoán vị thực hiện 3 phép gán.

| Trường hợp | Số phép so sánh | Số hoán vị |
|-----------------------|-----------------------|--------------------|
| Tốt nhất | $\frac{n(n-1)}{2}$ | 0 |
| Xấu nhất | $\frac{n(n-1)}{2}$ | $\frac{n(n-1)}{2}$ |
| Độ phức tạp thời gian | $T\left(n\right) =O$ | (n^2) |

Sắp xếp nổi bọt (bubble sort)

Ý tưởng

Cho dãy $a = \{a_0, a_1, ..., a_{n-1}\}$ gồm n phần tử, giải thuật là một vòng lặp thực hiện n-1 lần.

Tại mỗi lần lặp thứ i=0,1,...,n-2

- ▶ Bắt đầu từ cuối/đầu dãy, tìm nghịch thế giữa a_j và a_{j-1} , với j=n-1,n-2,...,1.
- ► Hoán vị hai phần tử trong nghịch thế này.

A Phần tử nhỏ (nhẹ) sẽ nổi lên trên và phần tử lớn (nặng) sẽ chìm xuống đáy.

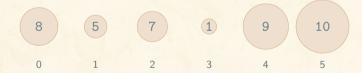
Sắp xếp nổi bọt (bubble sort)

```
Thuật toán 4: BubbleSort(a[], n)
- Đầu vào: mảng a gồm n phần tử.
- Đầu ra: mảng a có thứ tự tăng dần.
for i ← 0 to n - 2
  for j ← n - 1 downto i
    if a[j] < a[j - 1]
        Swap(a[j], a[j - 1])</pre>
```

Sắp xếp nổi bọt (bubble sort)

Ví du 4

Cho dãy số a gồm 6 phần tử: 8,5,7,1,9,10. Áp dụng giải thuật sắp xếp nổi bọt sắp dãy a theo thứ tự tăng dần.



 $\begin{cases} i = 0, \\ j = 5 \end{cases}$

$$\begin{cases} i = 0, \\ j = 4 \end{cases}$$

j 0

$$\begin{cases} i = 0, \\ j = 3 \end{cases}$$

- j 0 (8
 - 1 5
 - 2 7
- j 3 1
 - 4 9
 - 5 10

$$\begin{cases} i = 0, \\ j = 2 \end{cases}$$

- i 0 8
 - 1 5
 - 3 7
 - 4 9
 - 5 10

$$\begin{cases} i = 0, \\ j = 1 \end{cases}$$

- i 0 8
- j 1 1
 - 2 5
 - 3 7
 - 4 9
 - 5 10

- $\begin{cases} i = 1, \\ j = 5 \end{cases}$
 - 0 1
- j 1 (8)
 - 2 5
 - 3 7
 - 4 9
- j 5 10

$$\begin{cases} i = 1, \\ j = 4 \end{cases}$$

j 1

$$\begin{cases} i = 1, \\ j = 3 \end{cases}$$

$$\begin{cases} i = 1, \\ j = 2 \end{cases}$$



$$\begin{cases} i = 2, \\ j = 5 \end{cases}$$

 $\begin{cases} i = 2, \\ j = 4 \end{cases}$

39/55

$$\begin{cases} i = 3, \\ j = 5 \end{cases}$$

0

1

1 5

2 (7)

j 3 (8)

4 9

j 5 (10

$$\begin{cases}i=5\end{cases}$$

0 1

1 5

2 7

3 8

4 9

5 10

Sắp xếp nổi bọt (bubble sort)

M Nhân xét

- Không nhận biết được dãy đã có thứ tự hay có thứ tự từng phần.
- Các phần tử nhỏ được đưa về đúng vị trí rất nhanh, nhưng các phần tử lớn thì rất châm.

Sắp xếp nổi bọt (bubble sort)

Đánh giá giải thuật

► Số phép so sánh

$$\sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} 1 = \sum_{i=0}^{n-2} ((n-1) - i) = \frac{n(n-1)}{2}.$$

► Mỗi lần hoán vị thực hiện 3 phép gán.

| Trường hợp | Số phép so sánh | Số hoán vị |
|-----------------------|-----------------------|--------------------|
| Tốt nhất | $\frac{n(n-1)}{2}$ | 0 |
| Xấu nhất | $\frac{n(n-1)}{2}$ | $\frac{n(n-1)}{2}$ |
| Độ phức tạp thời gian | $T\left(n\right) =O$ | (n^2) |

Ý tưởng

Tương tự như sắp xếp nổi bọt, nhưng khắc phục được khuyết điểm của giải thuật này bằng cách

- Lượt đi: xuất phát từ cuối dãy không thứ tự, đẩy phần tử nhỏ nhất về đầu dãy.
- Lượt về: xuất phát từ đầu dãy không thứ tự, đẩy phần tử lớn nhất về cuối dãy.
- ► Ghi nhận lại vị trí những đoạn đã có thứ tự.



```
Thuât toán 5: ShakerSort(a[], n)
   - Đầu vào: mảng a gồm n phần tử.
   - Đầu ra: mảng a có thứ tư tăng dần.
     left \leftarrow 0, right \leftarrow n - 1
    k \leftarrow right
     while left < right
4
        for j ← right downto left - 1
          if a[j] < a[j - 1]
             Swap(a[j], a[j - 1])
7
       k \leftarrow i
       left \leftarrow k
        for i \leftarrow left to right - 1
          if a[i] > a[i + 1]
           Swap(a[i], a[i + 1])
            k \leftarrow i
    \texttt{right} \leftarrow \texttt{k}
```

5

6

8

10

12

13

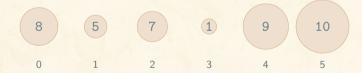
14

Giải thích

- Dòng 2: ghi nhận vị trí xảy ra hoán vị cuối cùng.
- ightharpoonup Dòng 4 ightharpoonup 7: đưa phần tử nhỏ về đầu dãy.
- Dòng 8: loại các phần tử có thứ tự ở đầu dãy.
- ightharpoonup Dòng 10
 ightharpoonup 13: đưa phần tử lớn về cuối dãy.
- Dòng 14: loại các phần tử có thứ tự ở cuối dãy.

Ví du 5

Cho dãy số a gồm 6 phần tử: 8, 5, 7, 1, 9, 10. Áp dụng giải thuật sắp xếp rung sắp dãy a theo thứ tự tăng dần.



$$\begin{cases} left = 0, \\ right = 5, \\ j = 5 \end{cases}$$

$$\begin{cases} 8 \\ 5 \\ \end{cases}$$

$$\begin{cases} 7 \\ 1 \\ \end{cases}$$

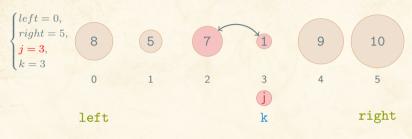
left

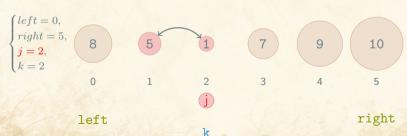
right

10

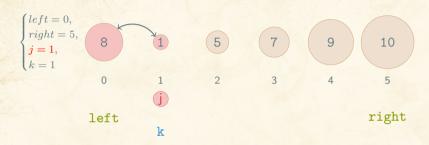
9

Nguyễn Chí Hiếu





Nguyễn Chí Hiểu





Nguyễn Chí Hiếu

$$\begin{cases} left = 1, \\ right = 5, \\ \frac{i = 2}{k}, \\ k = 2 \end{cases}$$
0
1
2
3
4
5
left
k



Nguyễn Chí Hiếu

$$\begin{cases} left = 1, \\ right = 5, \\ \frac{i = 4,}{k = 2} \end{cases}$$

$$0 \qquad 1 \qquad 2 \qquad 3 \qquad 4 \qquad 5$$

$$left \qquad \qquad \qquad right$$

$$k$$



Nguyễn Chí Hiếu

$$\begin{cases} left = 2, \\ right = 2, \\ \frac{i = 2}{k}, \\ k = 2 \end{cases}$$
0
1
2
3
4
5
left = right
k

Đánh giá giải thuật

- ► Trong trường hợp dãy số có thứ tự bộ phận thuật toán sẽ thực hiện số phép so sánh ít hơn Bubble Sort.
- ► Số phép hoán vị tương tự giải thuật Bubble Sort.

| Trường hợp | Số phép so sánh | Số hoán vị |
|-----------------------|---------------------------------------|--------------------|
| Tốt nhất | n-1 | 0 |
| Xấu nhất | $\frac{n(n-1)}{2}$ | $\frac{n(n-1)}{2}$ |
| Độ phức tạp thời gian | $T\left(n\right) = O\left(n^2\right)$ | |

Chương 2. TÌM KIẾM & SẮP XẾP Sắp xếp rung

Sắp xếp rung (shaker sort)

Sắp xếp rung (shaker sort)

- in gia giải thuật. Trong trường hợp dẫy số có thứ tự bộ phận thuật toán sẽ thực hiện số phép sc
- Số phép hoán vị tương tự giải thuật Bubble Sort.

| Trường hợp | Số phép so sánh | Số hoán vị |
|-----------------------|-----------------|-------------------|
| Tốt nhất | n-1 | 0 |
| Xấu nhất | <u>s(n-1)</u> | n(n-1) |
| Độ phức tạp thời gian | T(n) = O | (n ²) |

- Hỏi & đáp
 - Tại sao cần phải có nhiều giải thuật sắp xếp?
 Mỗi giải thuật có một số ưu điểm đối với từng trường hợp riêng biệt.
 - Giải thuật nào ổn định nhất? Giải thuật là ổn định (stable) nếu sau khi thực hiện sắp xếp, thứ tự tương đối của các mẫu tin có khóa bằng nhau không đổi. Bubble Sort kiểm tra nghịch thế giữa 2 phần tử kề nhau a_i và a_{i-1} .

Bài tập

- Cho dãy ký tự X, I, N, C, H, A, O. Áp dụng các giải thuật sắp xếp đã học để sắp theo đúng thứ tự bảng chữ cái Latinh.
- Cho ví dụ minh họa ưu điểm của Shaker Sort so với Bubble Sort khi sắp xếp một dãy số.
- 3. Bài toán di chuyển đĩa
 - ► Cho một dãy đĩa gồm 2 màu sáng và tối



Xây dựng giải thuật di chuyển các đĩa sáng về bên trái và đĩa tối về bên phải. Mỗi bước chỉ được hoán vị hai đĩa kề nhau.



Dêm số hoán vị hai đĩa khi thực hiện.

Gợi ý: xem dãy đĩa như một dãy số, đĩa màu sáng có giá trị là 0 và đĩa tối có giá trị là 1.

Chương 2. TÌM KIẾM & SẮP XẾP Bài tập

-Bài tâp

- Bài tập

 1. Coa dy, y tr 2, 1, 1, y, v, 1, 1, 4, 10. Áp dung các giải thuật sắp sắp the ped là chiếu chiếu phá y tràng thát (i.e. thuật chiếu ch
- Xét dãy số 2,3,4,5,6,7,8,9,1, sắp xếp dãy theo thứ tự tăng dần.
 - Shaker sort: chỉ cần 1 lần duyệt để đưa phần tử 1 về đầu dãy.
 - Bubble sort: cần n-1=8 lần duyệt (không nhận biết một phần dãy đã có thứ tự).

```
Chương 2. TÌM KIẾM & SẮP XẾP

Bài tập
```

```
EN tip

1. Cho dip by is X, T, B, C, K, A, C, Ap drang các giữ thuất sắp dọ đã học đi
dực thu được thủ tru hiểng chủ cái Linho.

2. Cho ví dực minh họu sự điện cái Shake Sern sự Vũ Bhábh Sort Mi dại sắp nột ti

1. Bi ti thiết chuyển định các thuyện có thể

1. Cho nhỏ dực định chuyển có thể chuyển có thuyến có thủ có thuyến có thuyến có thủ có thuyến có thủ có thuyến có thuyến có thủ có thuyến có thuyến có thủ có thuyến có
```

Áp dụng giải thuật Bubble Sort và một cờ để đánh dấu sau khi một cặp đĩa được hoán vị.
 BubbleSortEx(a[], n)

```
for i \leftarrow 0 downto n - 2

swapped = false

for j \leftarrow n - 1 to i

if a[j] < a[j - 1]

Swap(a[j], a[j - 1])

swapped \leftarrow true

if swapped = false

break
```

• Lưu ý, số hoán vị sẽ được giảm rất nhiều lần so với Bubble Sort.

Tài liệu tham khảo



Dương Anh Đức, Trần Hạnh Nhi.

Nhập môn Cấu trúc dữ liệu và Thuật toán. Đại học Khoa học tự nhiên TP Hồ Chí Minh, 2003.



Donald E. Knuth.

The Art of Computer Programming, Volume 3. Addison-Wesley, 1998.



Niklaus Wirth.

Algorithms + Data Structures = Programs. Prentice-Hall, 1976.



Robert Sedgewick.

Algorithms in C. Addison-Wesley, 1990.