

# Sorting Algorithms

---

## Sorting Algorithms

---

Các thuật toán sắp xếp gồm hai loại: sắp xếp nội (internal sorting), đưa toàn bộ dữ liệu vào bộ nhớ trong. Sắp xếp nội có kích thước dữ liệu không lớn và thời gian thực hiện thuật toán nhanh. Loại thứ hai là sắp xếp ngoại (external sorting), chỉ có một phần dữ liệu được đưa vào bộ nhớ trong, phần còn lại lưu ở bộ nhớ ngoài. Sắp xếp ngoại có kích thước dữ liệu rất lớn và thời gian thực hiện chậm.

**Internal memory** – bộ nhớ trong, chính xác là RAM. Sẽ mất dữ liệu khi tắt máy. Truy xuất dữ liệu ngẫu nhiên.

**External memory** – bộ nhớ ngoài, ngoài RAM hoặc Register. Điển hình là HDD hoặc SSD. Không mất dữ liệu khi tắt máy. Truy xuất dữ liệu trực tiếp phụ thuộc vào vị trí vật lý của dữ liệu.

**Inplace** : Không dùng bộ nhớ trung gian.

**Stable** : Một thuật toán được gọi là stable khi chúng giữ nguyên thứ tự của các phần tử tương tự tính chất với nhau. Chẳng hạn như ta có dãy

1, 7, 4, 5a, 5b, 5c, 7, 6

Sau khi sắp xếp xong, thuật toán Stable sẽ không đảo thứ tự giữa 5a, 5b và 5c. Còn các thuật toán Unstable có thể đảo thành 5a, 5c, 5b hoặc 5c, 5b, 5a gì đấy.

Một ví dụ khác là:

1, 9, 82, 2, 19, 27

Một thuật toán Stable, sau khi sắp xếp dãy trên xong, thứ tự giữa 82 và 2 là không đổi, do chúng có cùng tính chất là giống nhau hàng đơn vị (82 và 02).

Ngoại trừ Radix Sort thì còn lại đều là **Comparison-based Sorting Algorithm**. Tất cả các thuật toán Comparison-based mang tính tổng quát (General Purpose) thì lower bound là  $O(n \log(n))$ . Còn Radix Sort, mang tính đặc dụng hơn (Specific Purpose), thì có lower bound là  $O(n)$ .

Chúng ta phải xem xét các thuật toán dựa trên một ngữ cảnh nhất định.

# Internal Sorting

Có bốn loại internal sorting đó là

- Đổi chỗ (Exchange Sort) gồm Bubble Sort, Shaker Sort, Interchange Sort, Quick Sort,...
- Chèn (Insertion Sort) gồm Simple Insertion Sort, Binary Insertion Sort, Shell Sort,...
- Trộn (Merge Sort) gồm Merge Sort, Natural Merge Sort,...
- Lựa chọn (Selection Sort) gồm Simple Selection Sort, Heap Sort,...

Trong đó Shell Sort, Heap Sort, Quick Sort, Merge Sort và Radix Sort là các thuật toán phức tạp nhưng chi phí thấp. Các thuật toán còn lại đơn giản nhưng chi phí cao.

Idea của Bubble, Selection, Insertion và Heap Sort đều chia mảng thành hai phần đã sắp xếp và chưa sắp xếp. Thuật toán nào không nêu cụ thể đầu vào đầu ra ngầm hiểu có đầu vào đầu ra là một mảng chưa sắp xếp.

# Comparision

Comparison of Sorting Algorithms		
	Worst case	Average case
Selection sort	$n^2$	$n^2$
Bubble sort	$n^2$	$n^2$
Insertion sort	$n^2$	$n^2$
Mergesort	$n * \log n$	$n * \log n$
Quicksort	$n^2$	$n * \log n$
Radix sort	$n$	$n$
Treesort	$n^2$	$n * \log n$
Heapsort	$n * \log n$	$n * \log n$