Principles of Computer Science II

Sorting Algorithms

Ioannis Chatzigiannakis

Sapienza University of Rome

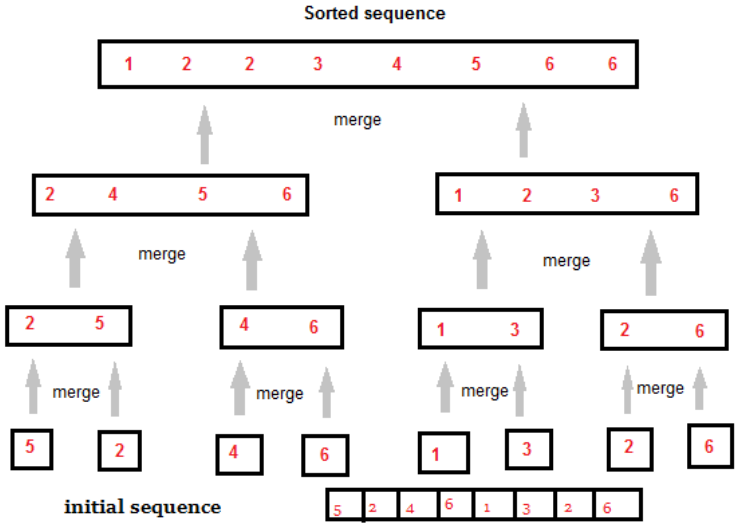
Lecture 9

Merge Sort Algorithm

In Merge Sort the unsorted list is divided into N sublists, each having one element, because a list consisting of one element is always sorted. Then, it repeatedly merges these sublists, to produce new sorted sublists, and in the end, only one sorted list is produced.

- Divide and Conquer algorithm
- Performance always same for Worst, Average, Best case

Merge Sort: Example



Merge Sort Code

```
a = [25, 52, 37, 63, 14, 17, 8, 6]

def mergesort(list):
    if len(list) == 1:
        return list

    left = list[0: len(list) // 2]
    right = list[len(list) // 2:]

    left = mergesort(left)
    right = mergesort(right)

    return merge(left, right)
```

Merge Sort Code

```
def merge(left, right):
    result = []
    while len(left) > 0 and len(right) > 0:
        if left[0] <= right[0]:
            result.append(left.pop(0))
        else:
            result.append(right.pop(0))

    while len(left) > 0:
        result.append(left.pop(0))

    while len(right) > 0:
        result.append(right.pop(0))

    return result

print("Before: ", a)
r = mergesort(a)
print("After: ", r)
```

How good is Merge Sort?

- How many comparisons are required until the list is sorted?
  - 1^{st} loop: two lists $\frac{n}{2}$ each
  - 2^{nd} loop: four lists $\frac{n}{4}$ each
  - ...
  - $\log n$ steps
  - For each partition we do n comparisons
  - In total $n \log n$ comparisons
- How much memory is needed ?
  - 1 additional slot.

Quick Sort Algorithm

Quick sort is very fast and requires very less additional space. It is based on the rule of Divide and Conquer. This algorithm divides the list into three main parts :

- Elements less than the Pivot element
- Pivot element(Central element)
- Elements greater than the pivot element

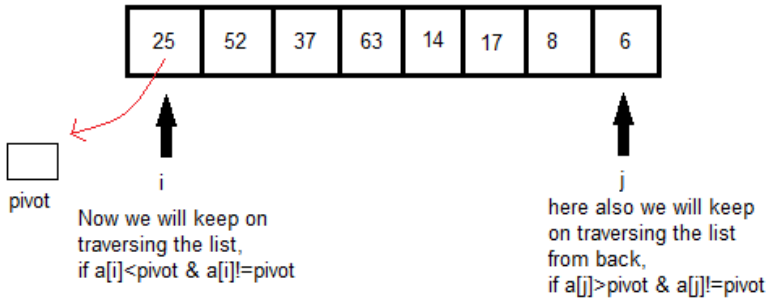
- Sorts any list very quickly
- Performance depends on the selection of the Pivot element

Quick Sort: Example

List: 25 52 37 63 14 17 8 6

- We pick 25 as the pivot.
- All the elements smaller to it on its left,
- All the elements larger than to its right.
- After the first pass the list looks like:
  6 8 17 14 25 63 37 52
- Now we sort two separate lists:
  6 8 17 14 and 63 37 52
- We apply the same logic, and we keep doing this until the complete list is sorted.

Quick Sort: Example



if both sides we find the element not satisfying their respective conditions, we swap them. And keep repeating this.

DIVIDE AND CONQUER - QUICK SORT

Quick Sort Code

```
a = [25, 52, 37, 63, 14, 17, 8, 6]

def partition(list, p, r):
    pivot = list[p]
    i = p
    j = r
    while(1):
        while(list[i] < pivot and list[i] != pivot):
            i += 1

        while(list[j] > pivot and list[j] != pivot):
            j -= 1

        if(i < j):
            temp = list[i]
            list[i] = list[j]
            list[j] = temp
        else:
            return j
```

Quick Sort Code

```
def quicksort(list, p, r):
    if (p < r):
        q = partition(list, p, r)
        quicksort(list, p, q);
        quicksort(list, q + 1, r);

print("Before: ", a)
quicksort(a, 0, len(a) - 1)
print("After: ", a)
```

How good is Quick Sort?

- How many comparisons are required until the list is sorted?

How good is Quick Sort?

- How many comparisons are required until the list is sorted?
- What if we choose the smallest or the largest item as pivot?

How good is Quick Sort?

- How many comparisons are required until the list is sorted?
- What if we choose the smallest or the largest item as pivot?
  - 1^{st} loop: n - 1
  - 2^{nd} loop: n - 2
  - ...
  - (n-1)+(n-2)+(n-3)+ ...+3+2+1 comparisons are required
  - $\sum \frac{n(n-1)}{2}$ comparisons are required

How good is Quick Sort?

- How many comparisons are required until the list is sorted?
- What if we choose the smallest or the largest item as pivot?
  - 1^{st} loop: n - 1
  - 2^{nd} loop: n - 2
  - ...
  - (n-1)+(n-2)+(n-3)+ ...+3+2+1 comparisons are required
  - $\sum \frac{n(n-1)}{2}$ comparisons are required
- What if we choose the median item as pivot?

How good is Quick Sort?

- How many comparisons are required until the list is sorted?
- What if we choose the smallest or the largest item as pivot?
  - 1^{st} loop: n - 1
  - 2^{nd} loop: n - 2
  - ...
  - (n-1)+(n-2)+(n-3)+ ...+3+2+1 comparisons are required
  - $\sum \frac{n(n-1)}{2}$ comparisons are required
- What if we choose the median item as pivot?
  - 1^{st} loop: two lists $\frac{n}{2}$ each
  - 2^{nd} loop: four lists $\frac{n}{4}$ each
  - ...
  - $\log n$ steps
  - For each partition we do n comparisons
  - In total $n \log n$ comparisons

How good is Quick Sort?

- How many comparisons are required until the list is sorted?
- What if we choose the smallest or the largest item as pivot?
  - 1^{st} loop: n - 1
  - 2^{nd} loop: n - 2
  - ...
  - (n-1)+(n-2)+(n-3)+ ...+3+2+1 comparisons are required
  - $\Sigma \frac{n(n-1)}{2}$ comparisons are required
- What if we choose the median item as pivot?
  - 1^{st} loop: two lists $\frac{n}{2}$ each
  - 2^{nd} loop: four lists $\frac{n}{4}$ each
  - ...
  - log n steps
  - For each partition we do n comparisons
  - In total $n \log n$ comparisons
- How much memory is needed ?

How good is Quick Sort?

- How many comparisons are required until the list is sorted?
- What if we choose the smallest or the largest item as pivot?
  - 1^{st} loop: n - 1
  - 2^{nd} loop: n - 2
  - ...
  - (n-1)+(n-2)+(n-3)+ ...+3+2+1 comparisons are required
  - $\Sigma \frac{n(n-1)}{2}$ comparisons are required
- What if we choose the median item as pivot?
  - 1^{st} loop: two lists $\frac{n}{2}$ each
  - 2^{nd} loop: four lists $\frac{n}{4}$ each
  - ...
  - log n steps
  - For each partition we do n comparisons
  - In total $n \log n$ comparisons

- How much memory is needed ?
  - 1 additional slot.

3^{rd} Assignment

- https://www.rosalind.info/
  - Complete the following challenges:
    fib hamm fibd mrna prtm lcsm perm revp lexf lgis
  - http://rosalind.info/problems/{challenge}
- Create a GitHub repository and upload the code for each exercise.
- Email `ichatz@diag.uniroma1.it`
  Subject: [PCS2] Homework 3
  Your GitHub repository with your solutions, for all challenges.
  Also send your account user account link:
  http://rosalind.info/users/{username}