lecture 15 — Oct 07

- Onboarding Exam ?
- paper & pen(cil)
- UR Courses chat with instructor
- closed-book exam

Insertion Sort

```
for (i=1; i<N; i++)
  {
     tmp = a[i];
     for (j=i; j>0 && tmp < a[j-1]; j--)
          a[j] = a[j-1];
     a[j] = tmp;
  }
```

redo Example 32.

input	51	7	32	11	27	21	29	# comp.
~~after~~ i=1	7	51 |	32	11	27	21	29	1
i=2	7	32	51 |	11	27	21	29	2
i=3	7	11	32	51 |	27	21	29	3
i=4	7	11	27	32	51 |	21	29	3
i=5	7	11	21	27	32	51 |	29	4
i=6	7	11	21	27	29	32	51 |	3 // 16 comparisons

analysis

- worst case $T(N) = O(1 + 2 + 3 + 4 + 5 + \ldots + N-1)$
$$= O(N^2)$$

if array initially in reverse order, each pass i exactly i comparisons $\Rightarrow O(N^2)$ is attained

$\Rightarrow T_{worst}(N) = \Theta(N^2)$.

- best case if array initially sorted, then exactly one comparison per pass is made, i.e., a total of $N-1$ comparisons

(there is no better case, because outer loop runs $N-1$ times, always making at least one comparison).

$\Rightarrow T_{best}(N) = \Theta(N)$.

- average case goal: prove that $T_{avg}(N) = \Theta(N^2)$, assuming that

the array contains no duplicates and every permutation of list elements is equally likely (*)

Since $T_{worst}(N) = O(N^2)$, we also know $T_{avg}(N) = O(N^2)$.

Hence it suffices to prove

$$T_{avg}(N) = \Omega(N^2) \quad \text{under condition } (*)$$

We prove even more:

Theorem 8. <u>Any</u> sorting algorithm that exchanges only adjacent elements needs $\Omega(N^2)$ time on average, under assumption (*).

Lemma 1. The avg. number of pairs (k, ℓ) with

$$k < \ell \quad \text{and} \quad a[k] > a[\ell]$$

(called "inversions") in an array with N elements is

$$\frac{N(N-1)}{4}$$

Proof of Theorem 8.

Since the algorithm swaps only adjacent elements, each swap remove at most inversion.

$$\Rightarrow \text{avg } \# \text{swaps} \geq \text{avg } \# \text{inversions} \overset{\text{Lemma 1}}{=} \frac{N(N-1)}{4} = \Omega(N^2)$$

$$\Rightarrow T_{avg}(N) = \Omega(N^2)$$

$\square$

Proof of Lemma 1.

We group arrays into two groups A and B of the same size.

For each list in A its reversed list is in B and vice versa

A	B
1, 3, 2, 4	4, 2, 3, 1
1, 2, 3, 4	4, 3, 2, 1
⋮	⋮

Each list L in A has $\frac{N(N-1)}{2}$ pairs of indices (k,l) with $k<l$.

Each pair is an inversion either in L, or the reversed list of L in B.

$\Rightarrow$ Each of these $\frac{N(N-1)}{2}$ pairs is an inversion in exactly half of the arrays.

$\Rightarrow$ A list has on average $\frac{N(N-1)}{2} \cdot \frac{1}{2} = \frac{N(N-1)}{4}$ inversions. $\square$

$\Rightarrow T_{avg}(N) = \Theta(N^2)$ for insertion sort!