8.4 Bucket-Sort

- Sorts an array of values in the interval $[0,1)$

Bucket-Sort (A)

let $B[0,1,\dots,n-1]$ be a new array

for $i = 0$ to $n-1$ || $\Theta(n)$

make $B[i]$ an empty list

for $i = 1$ to n || $\Theta(n)$

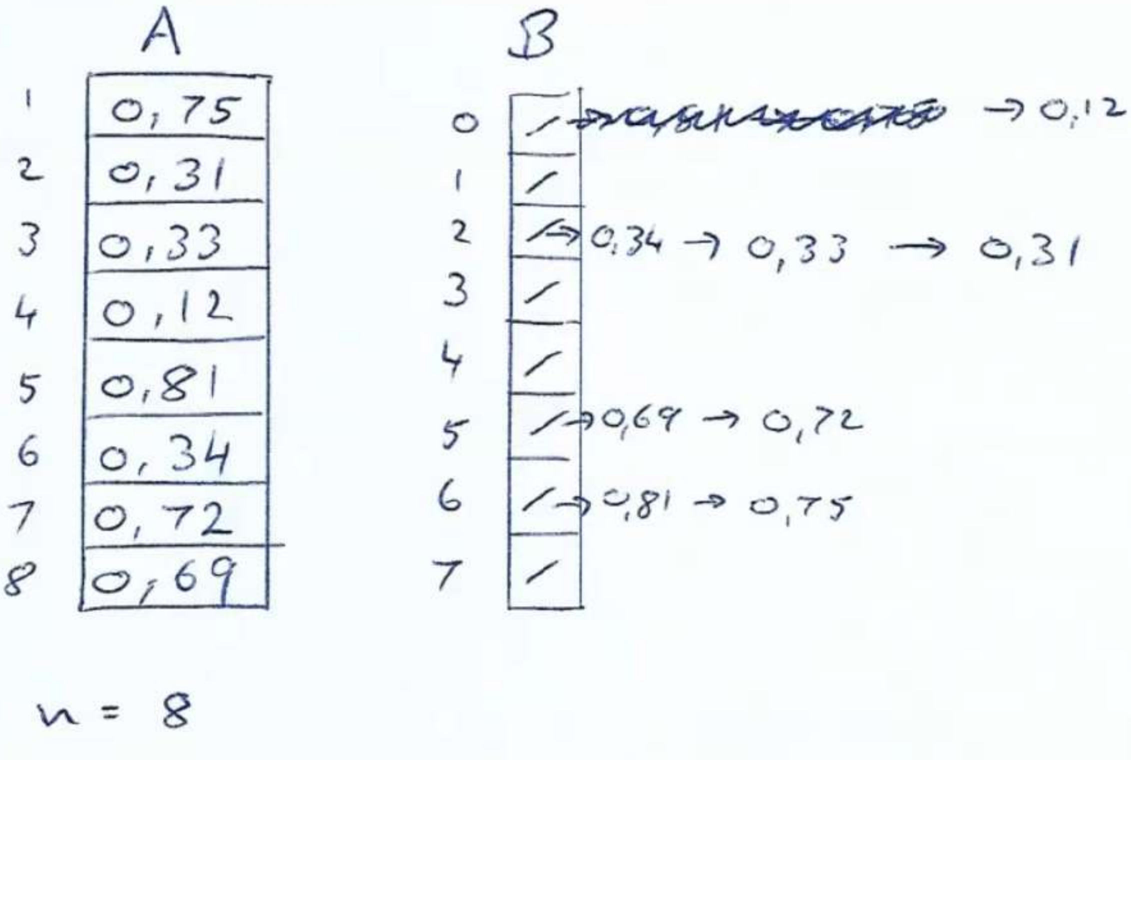
insert $A[i]$ into list $B[\lfloor nA[i] \rfloor]$

for $i = 0$ to $n-1$

for $i = 0$ to $n-1$

sort list $B[i]$ using InsertionSort $\quad \| \; O(n^2)$

concatenate lists $B[0], \ldots, B[n-1]$.


A
B
1 0,75
2 0,31
3 0,33
4 0,12
5 0,81
6 0,34
7 0,72
8 0,69
0 → 0,12
1
2 → 0,34 → 0,33 → 0,31
3
4
5 → 0,69 → 0,72
6 → 0,81 → 0,75
7

n = 8

$$A[i] \rightarrow B[\lfloor n A[i] \rfloor]$$

$$A[1] \rightarrow B[\lfloor 8A[1] \rfloor] = B[\lfloor 8 \cdot 0{,}75 \rfloor] = B[6]$$

$$A[2] \rightarrow B[\lfloor 8 \cdot 0{,}31 \rfloor] = B[\lfloor 2{,}48 \rfloor] = B[2]$$

$$\vdots \quad \vdots$$

$$A[8] \rightarrow B[\lfloor 8 \cdot 0{,}69 \rfloor] = B[\lfloor 5{,}52 \rfloor] = B[5]$$

sort each list in B:

B

0	/	0,12
1	/	~~[illegible]~~
2	/	0,31 → 0,33 → 0,34
3	/	
4	/	
5	/	0,69 → 0,72
6	/	0,75 → 0,81
7	/	

concatenate lists

B

0,12
0,31
0,33
0,34
0,69
0,72
0,75
0,81

If all values, or even most, fall into one bucket, then Insertion sort will take $O(n^2)$. So worst case is $O(n^2)$.

But we expect that the values will be distributed over all buckets, so that Insertion Sort operates on small lists.

We make the assumption that the values in A are uniformly distributed.

We show that the average case running time of Bucket-Sort is $\Theta(n)$

running time of Bucket-Sort is :

$$T(n) = \Theta(n) + \sum_{i=0}^{n-1} O(n_i^2)$$

where n_i is number of values in the bucket at index i.

$$E[T(n)] = E\left[\Theta(n) + \sum_{i=0}^{n-1} O(n_i^2)\right]$$

$$= E[\Theta(n)] + \sum_{i=0}^{n-1} E[O(n_i^2)]$$

$$= \Theta(n) + \sum_{i=0}^{n-1} O(E[n_i^2])$$

$i = 0$

$$\begin{array}{c|c|} \hline n_1 & 1 \\ \hline n_2 & 0 \\ \hline n_3 & 2 \\ \hline n_4 & 0 \\ \hline n_5 & 1 \\ \hline n_6 & 2 \\ \hline n_7 & 1 \\ \hline \end{array} \longrightarrow \left.\begin{array}{c|c|} \hline n_1^2 & 1 \\ \hline n_2^2 & 0 \\ \hline n_3^2 & 4 \\ \hline n_4^2 & 0 \\ \hline n_5^2 & 1 \\ \hline n_6^2 & 4 \\ \hline n_7^2 & 1 \\ \hline \end{array}\right\}$$

average value of n_i^2 is $\frac{11}{7}$.

We use the fact that $E[n_i^2] = 2 - \frac{1}{n}$

Then $E[T(n)] = \Theta(n) + \sum_{i=0}^{n-1} O(2 - \frac{1}{n})$

$= \Theta(n) + O(n(2 - \frac{1}{n}))$

$= \Theta(n) + O(2n - 1)$

$= \Theta(n)$