Correctness of Quicksort

```
Quicksort (A, l, r)
   if l < r
      m := Partition (A, l, r)
      Quicksort (A, l, m-1)
      Quicksort (A, m, r)
```

Claim: For all numbers l, r, where $l \leq r$, the array A is sorted when Quicksort(A, l, r) returns.

We prove the claim under the following assumption, which we will justify later on.

Assumption: Partition is correct, that is, if Partition(A, l, r) returns the value m, then

- $l < m \leq r$
- if $l \leq i \leq m-1$ and $m \leq j \leq r$, then $A[i] \leq A[j]$

Proof: We prove the claim by induction over the length of the array segment, that is, over $k = r - l + 1$.

Base case $k=1$: In this case $l = r$ and Quicksort returns without doing anything. Clearly, as a result the subarray $A[l..r]$ is sorted, since it contains only a single element.

Inductive step: Assume that the array segment $A[l..r]$ is sorted whenever Quicksort(A, l, r) returns, provided that $r - l + 1 \leq k$. Consider the call Quicksort(A, l, r) where $r - l + 1 = k+1$, that is, $r - l \leq k$. Since $l < m \leq r$, we have that

- $\underbrace{m-1}_{r'} - \underbrace{l}_{l'} + 1 = m - l \leq r - l \leq k$;
  according to the induction hypothesis, Quicksort$(A, l, m-1)$ sorts the subarray from l to $m-1$, that is $A[l..m-1]$
- $r - m + 1 < r - l + 1 = k+1$, that is, $r - m + 1 \leq k$; according to the induction hypothesis, Quicksort(A, m, r) sorts the subarray $A[m..r]$.

Now, $A[l, m-1]$ and $A[m, r]$ are each sorted and any element in $A[l, m-1]$ is less or equal than any element in $A[m, r]$. Hence, $A[l, r]$ is sorted.

Correctness of Partition

```
Partition(A, l, r)
    x := A[r]
    i := l - 1
    j := r + 1

    while true
        repeat j := j - 1
            until A[j] ≤ x
        repeat i := i + 1
            until A[i] ≥ x
        if i < j
            then swap A[i] ⟺ A[j]
            else return i
```

Claim 1: If $m = \text{Partition}(A, l, r)$ for $l < r$, then $m > l$.

Proof: The first repeat loop terminates the first time with "$j = r$", due to the choice of x.
The second repeat loop terminates the first time with "$i \geq l$".

Case "$i > l$": we are done, since i can only grow during the execution

Case "$i = \ell$": Then $i = \ell < r = j$, hence the if-clause is true and $A[i] \leftrightarrow A[j]$ are swapped.
Then the while loop is executed once more and i is incremented.

Claim 2: When Partition(A, ℓ, r) terminates, then

$A[i] \leq A[j]$ for $\ell \leq i \leq m-1$ and $m \leq j \leq r$.

Proof: We show that the following invariant holds during the while loop:

$$j' \geq j \Rightarrow A[j'] \geq x$$
$$i' \leq i \Rightarrow A[i'] \leq x$$

Initialization: The invariant holds initially, because initially $j > r$ and $i < \ell$, and there are no elements in $A[\ell..r]$ with index $j' \geq j$ or $i' \leq i$.

Maintenance: If j moves downward and encounters an element $A[j] < x$, then i moves upward until it encounters an element $A[i] \geq x$. The two elements are swapped and the invariant is maintained.

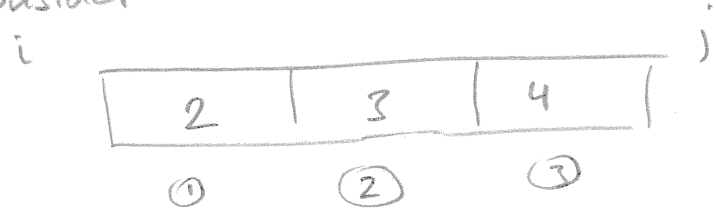
Termination: The loop terminates because eventually $i \geq j$, since i is continuously incremented and j is continously decremented.

Questions:

- What if we do

  $x := A[\ell]$ (instead of $A[r]$)?

Consider

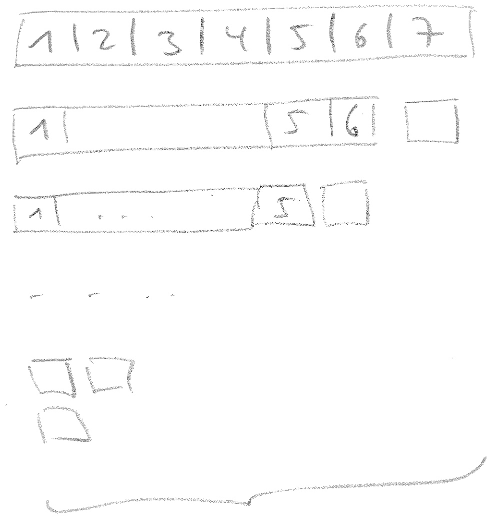


Running Time

- Worst case

A is sorted

n call levels



$\Rightarrow m = 7$

$\Rightarrow m = 6$

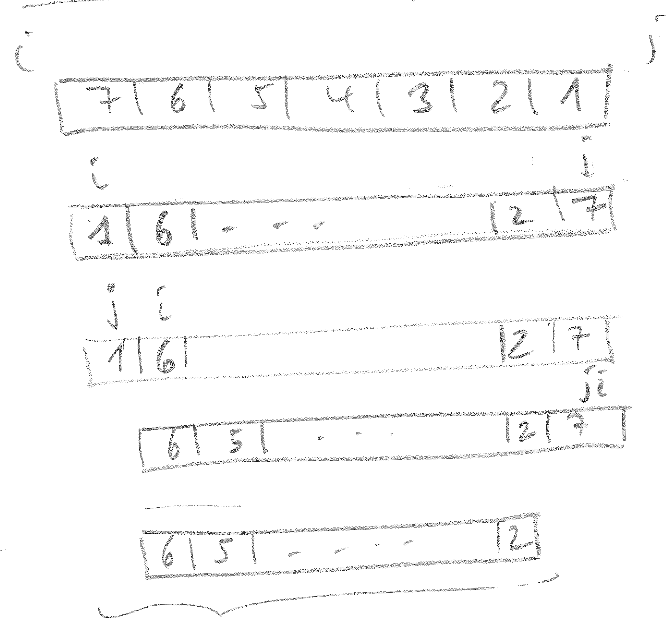
$\Rightarrow m = 5$

segments to partition have length k, $1 \leq k \leq n$

$$\Rightarrow T(n) = \sum_{k=1}^{n} \Theta(k) = \Theta\left(\frac{n(n+1)}{2}\right) = \Theta(n^2)$$

A is inversely sorted

n call levels



$\rightarrow m = 2$

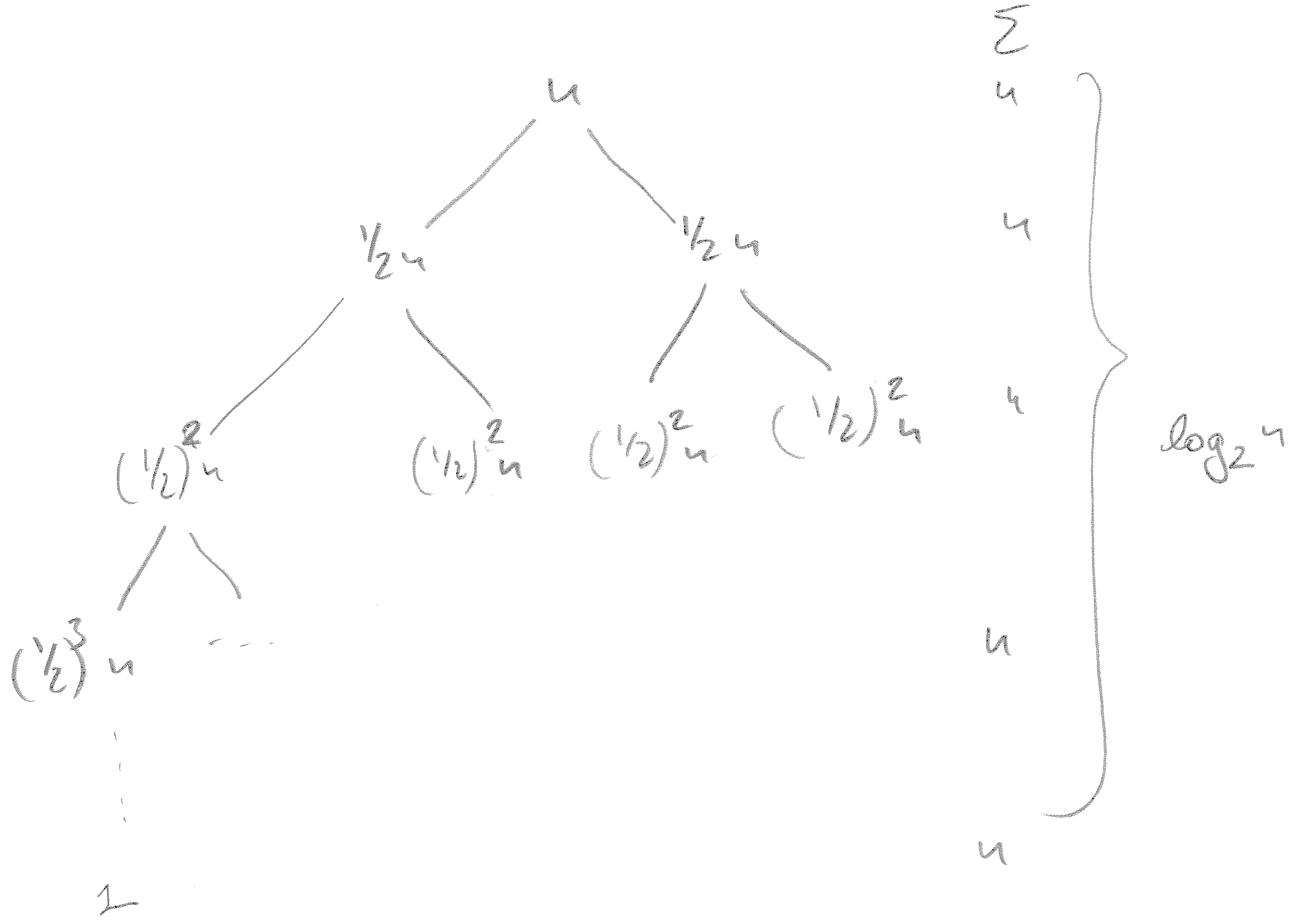
$\rightarrow m = 7$

$\rightarrow m = 3$

$\rightarrow m = 6$

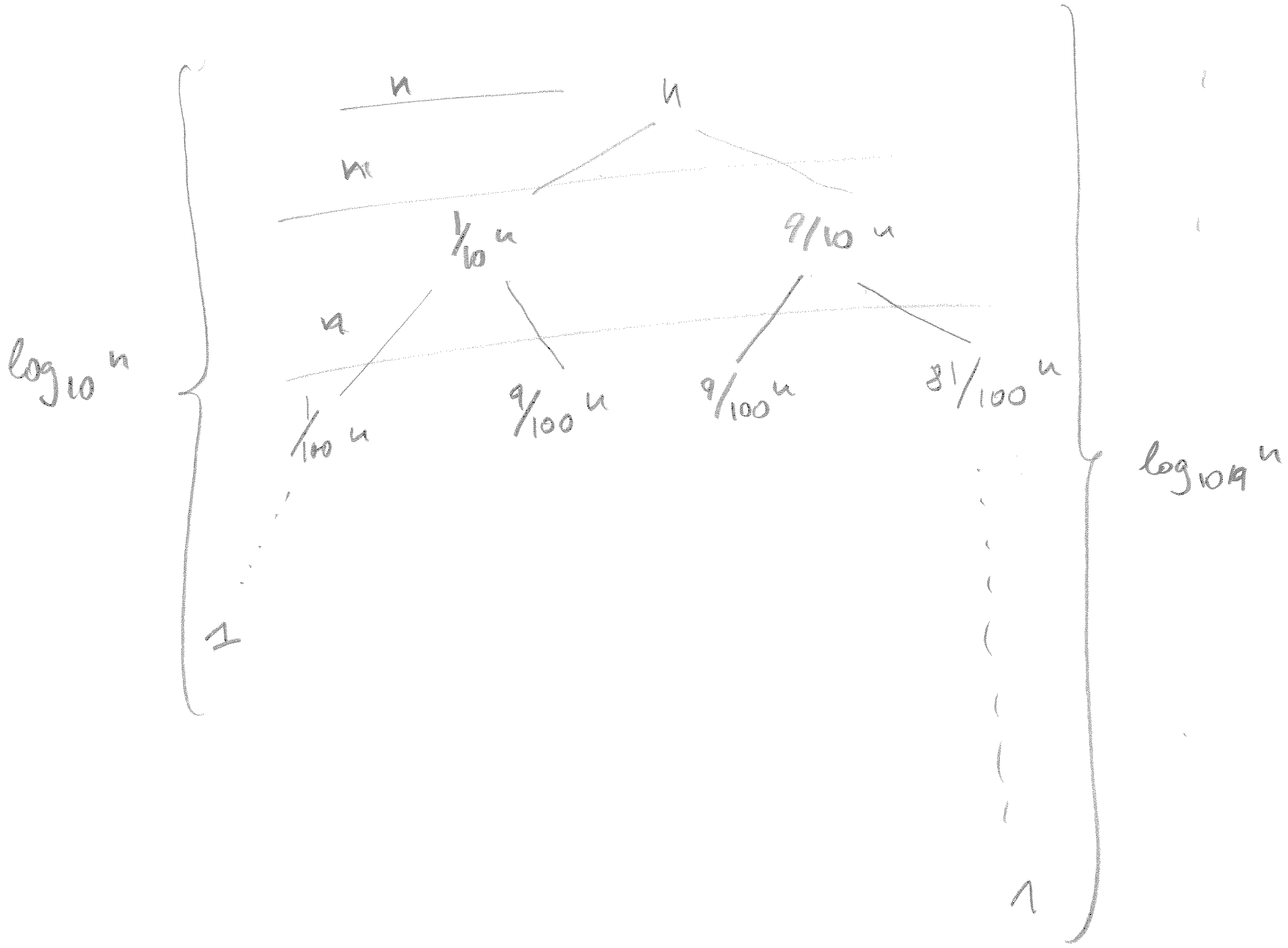
segments of length k where $1 \leq k \leq n$.

Best Case



$$T(n) = \Theta(n \cdot \log_2 n)$$

1/10 : 9/10 Split



$$T(n) \leq \Theta(n \log_{10/9} n)$$

$$= \Theta(n \log n)$$

Lucky vs. Unlucky Splits

$$L(n) = 2\,U(n/2) + \Theta(n)$$

$$U(n) = L(n-1) + \Theta(n)$$

$$\Rightarrow L(n) = 2\left(L(n/2 - 1) + \Theta(n) \right) + \Theta(n)$$
$$= 2\,L(n/2 - 1) + \Theta(n)$$

$$\Rightarrow L(n) \leq \Theta(n \cdot \log n)$$