Announcements

- MidTerm Exam 1: October 16 in class
- MidTerm Exam 2: Last day of class
- Final: NO FINAL EXAM

QuickSelect: a variant of QuickSort

QUICKSELECT(*array* A, *int* k, *int* p, *int* r)

$\triangleright$ Select k-th largest in subarray $A[p..r]$

if $(p = r)$

  then return $A[p]$

$q \leftarrow$ PARTITION(A, p, r)

$i \leftarrow q - p + 1$ $\triangleright$ Compute rank of pivot

if $(i = k)$

  then return $A[q]$

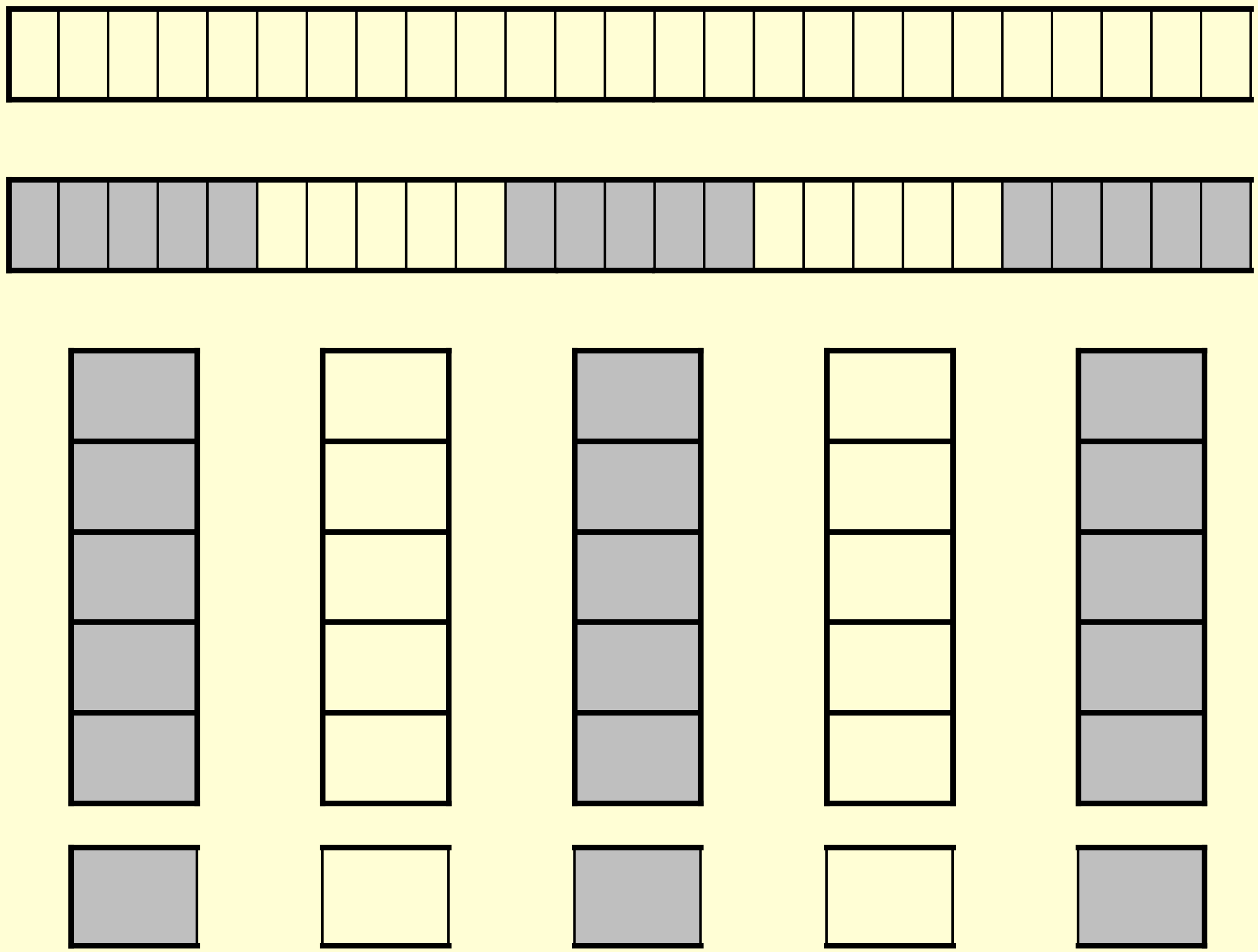
if $(i > k)$

  then return QUICKSELECT$(A, k, p, q - 1)$

  else return QUICKSELECT$(A, k - i, q + 1, r)$

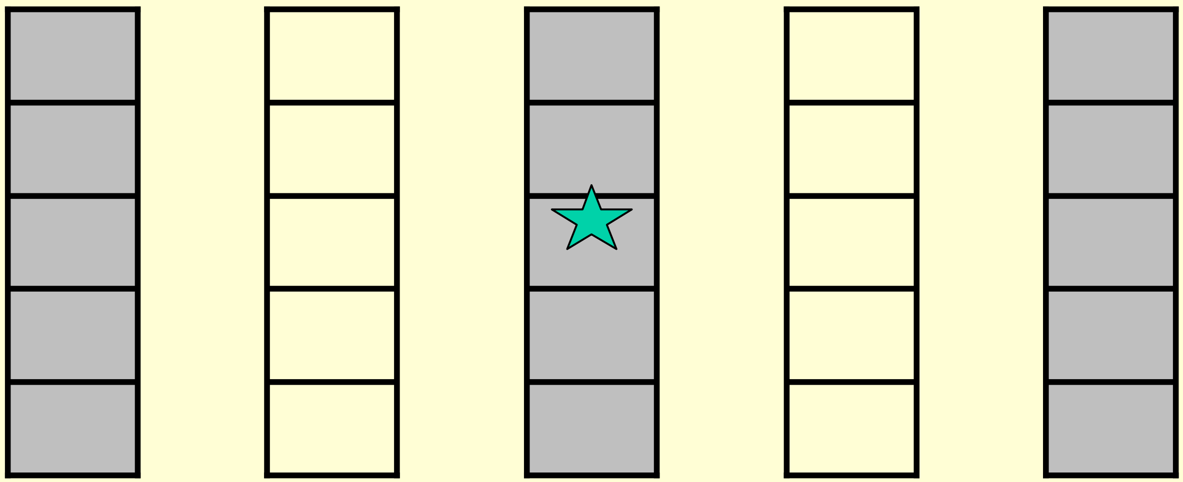
k-Selection & Median: Improved Algorithm

- Start with initial array



k-Selection & Median: Improved Algorithm(Cont'd)

- Use median of medians as pivot



- $T(n) < O(n) + T(n/5) + T(3n/4)$

ImprovedSelect

ImprovedSelect($array\ A, int\ k, int\ p, int\ r$)

    ▷ Select k-th largest in subarray $A[p..r]$

if $(p = r)$

    then return $A[p]$

    else $N \leftarrow r - p + 1$

Partition $A[p..r]$ into subsets of 5 elements and collect all medians of subsets in $B[1..\lceil N/5 \rceil]$.

$Pivot \leftarrow$ ImprovedSelect$(B, 1, \lceil N/5 \rceil, \lceil N/10 \rceil)$

$q \leftarrow$ PivotPartition$(A, p, r, Pivot)$

$i \leftarrow q - p + 1$ ▷ Compute rank of pivot

if $(i = k)$

    then return $A[q]$

if $(i > k)$

    then return ImprovedSelect$(A, k, p, q - 1)$

    else return ImprovedSelect$(A, k - i, q + 1, r)$

PivotPartition

```
PIVOTPARTITION(array A, int p, int r, item Pivot)
    ▷ Partition using provided Pivot
    i ← p − 1
    for j ← p to r
            do if (A[j] ≤ Pivot)
                    then i ← i + 1
                            exchange A[i] ↔ A[j]
    return i + 1
```

Analysis of ImprovedSelect

Number of elements greater than "median of medians" is at least

$$3\left(\left\lceil \frac{1}{2}\left\lceil \frac{n}{5} \right\rceil \right\rceil - 2\right) \geq \frac{3n}{10} - 6$$

Why?

Our recurrence is given by:

$$T(n) = O(n) + T(\lceil n/5 \rceil) + T(3n/4)$$

Thus there exists a positive constant a such that

$$T(n) \leq an + T(\lceil n/5 \rceil) + T(3n/4)$$

Using the substitution method, let's guess that $T(n) = O(n)$, i.e., $T(n) \leq cn$.
Then we need to show that

$$an + c\lceil n/5 \rceil + c(3n/4) \leq cn$$

What positive values of c and n_0 would enforce the above inequality?
When $n > 70$, and choosing $c \geq 20a$ will satisfy above inequality.

Data Structure Evolution

- Standard operations on data structures
  - **Search**
  - **Insert**
  - **Delete**
- Linear Lists
  - Implementation: **Arrays** **(Unsorted and Sorted)**
- **Dynamic** Linear Lists
  - Implementation: **Linked Lists**
- **Dynamic** Trees
  - Implementation: **Binary Search Trees**