Module 3: Sorting and Randomized Algorithms

Mark Petrick
Based on lecture notes by many previous cs240 instructors

David R. Cheriton School of Computer Science, University of Waterloo

Fall 2020

References: Sedgewick 6.10, 7.1, 7.2, 7.8, 10.3, 10.5
Goodrich & Tamassia 8.3
Outline

1. Sorting and Randomized Algorithms
   - QuickSelect
   - Randomized Algorithms
   - QuickSort
   - Lower Bound for Comparison-Based Sorting
   - Non-Comparison-Based Sorting
Outline

1. Sorting and Randomized Algorithms
   - QuickSelect
   - Randomized Algorithms
   - QuickSort
   - Lower Bound for Comparison-Based Sorting
   - Non-Comparison-Based Sorting
Selection vs. Sorting

The selection problem: Given an array $A$ of $n$ numbers, and $0 \leq k < n$, find the element that would be at position $k$ of the sorted array.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>80</td>
<td>90</td>
<td>10</td>
<td>40</td>
<td>70</td>
</tr>
</tbody>
</table>

`select(3)` should return 30.

Special case: median finding = selection with $k = \lfloor \frac{n}{2} \rfloor$.

Selection can be done with heaps in time $\Theta(n + k \log n)$. Median-finding with this takes time $\Theta(n \log n)$.

This is the same cost as our best sorting algorithms.

Question: Can we do selection in linear time?
The quick-select algorithm answers this question in the affirmative.

The encountered sub-routines will also be useful otherwise.
Crucial Subroutines

*quick-select* and the related algorithm *quick-sort* rely on two subroutines:

- **choose-pivot**\( (A) \): Return an index \( p \) in \( A \). We will use the pivot-value \( v \leftarrow A[p] \) to rearrange the array.

  Simplest idea: Always select rightmost element in array

  \[
  \text{choose-pivot1}(A) \\
  1. \quad \text{return } A.\text{size} - 1
  \]

  We will consider more sophisticated ideas later on.

- **partition**\( (A, p) \): Rearrange \( A \) and return pivot-index \( i \) so that
  - the pivot-value \( v \) is in \( A[i] \),
  - all items in \( A[0, \ldots, i-1] \) are \( \leq v \), and
  - all items in \( A[i+1, \ldots, n-1] \) are \( \geq v \).

  \[
  \begin{array}{cc}
  \leq v & v \\
  i & \geq v
  \end{array}
  \]
Partition Algorithm

Conceptually easy linear-time implementation:

\[
\text{partition}(A, p)
\]
\[
A: \text{array of size } n, \quad p: \text{integer s.t. } 0 \leq p < n
\]
1. Create empty lists \textit{smaller}, \textit{equal} and \textit{larger}.
2. \(v \leftarrow A[p]\)
3. \textbf{for} each element \(x\) in \(A\)
4. \hspace{1em} \textbf{if} \(x < v\) \textbf{then} \(\text{smaller}.\text{append}(x)\)
5. \hspace{1em} \textbf{else if} \(x > v\) \textbf{then} \(\text{larger}.\text{append}(x)\)
6. \hspace{1em} \textbf{else} \(\text{equal}.\text{append}(x)\).
7. \(i \leftarrow \text{smaller}.\text{size}\)
8. \(j \leftarrow \text{equal}.\text{size}\)
9. Overwrite \(A[0 \ldots i-1]\) by elements in \textit{smaller}
10. Overwrite \(A[i \ldots i+j-1]\) by elements in \textit{equal}
11. Overwrite \(A[i+j \ldots n-1]\) by elements in \textit{larger}
12. \textbf{return} \(i\)

More challenging: partition \textbf{in place} (with \(O(1)\) auxiliary space).
### Efficient In-Place partition (Hoare)

<table>
<thead>
<tr>
<th>i=1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>j=9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>80</td>
<td>90</td>
<td>20</td>
<td>40</td>
<td>v=70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>i=5</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>i=5</th>
<th>6</th>
<th>7</th>
<th>j=8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>80</td>
<td>90</td>
<td>20</td>
<td>40</td>
<td>v=70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>i=5</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>i=5</th>
<th>6</th>
<th>7</th>
<th>j=8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>40</td>
<td>90</td>
<td>20</td>
<td>80</td>
<td>v=70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>i=6</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>i=6</th>
<th>j=7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>40</td>
<td>90</td>
<td>20</td>
<td>80</td>
<td>v=70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>i=6</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>i=6</th>
<th>j=7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>40</td>
<td>20</td>
<td>90</td>
<td>80</td>
<td>v=70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>i=7</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>j=6</th>
<th>i=7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>40</td>
<td>20</td>
<td>90</td>
<td>80</td>
<td>v=70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>i=7</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>j=6</th>
<th>i=7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30</td>
<td>60</td>
<td>10</td>
<td>0</td>
<td>50</td>
<td>40</td>
<td>20</td>
<td>70</td>
<td>80</td>
<td>90</td>
</tr>
</tbody>
</table>
Efficient In-Place partition (Hoare)

Idea: Keep swapping the outer-most wrongly-positioned pairs.

\[
\begin{array}{cccc}
A & \leq v & ? & \geq v \\
i & j & n-1
\end{array}
\]

\[
\text{partition}(A, p)
\]

\[A: \text{array of size } n, \quad p: \text{integer s.t. } 0 \leq p < n\]

1. \(\text{swap}(A[n-1], A[p])\)
2. \(i \leftarrow -1, \quad j \leftarrow n-1, \quad v \leftarrow A[n-1]\)
3. \text{loop}
4. \quad \text{do } i \leftarrow i + 1 \text{ while } i < n \text{ and } A[i] < v
5. \quad \text{do } j \leftarrow j - 1 \text{ while } j > 0 \text{ and } A[j] > v
6. \quad \text{if } i \geq j \text{ then break} \quad \text{(goto 9)}
7. \quad \text{else } \text{swap}(A[i], A[j])
8. \text{end loop}
9. \text{swap}(A[n-1], A[i])
10. \text{return } i

Running time: \(\Theta(n)\).
QuickSelect Algorithm

quick-select1(A, k)
A: array of size n,  k: integer s.t. 0 ≤ k < n
1.  \( p \leftarrow \text{choose-pivot1}(A) \)
2.  \( i \leftarrow \text{partition}(A, p) \)
3.  \( \text{if } i = k \text{ then} \)
4.      \( \text{return } A[i] \)
5.  \( \text{else if } i > k \text{ then} \)
6.      \( \text{return } \text{quick-select1}(A[0, 1, \ldots, i - 1], k) \)
7.  \( \text{else if } i < k \text{ then} \)
8.      \( \text{return } \text{quick-select1}(A[i + 1, i + 2, \ldots, n - 1], k - i - 1) \)
Analysis of *quick-select*1

**Worst-case analysis:** Recursive call could always have size $n - 1$. 

Recurrence given by $T(n) = \begin{cases} 
T(n - 1) + cn, & n \geq 2 \\
c, & n = 1 
\end{cases}$

Solution: $T(n) = cn + c(n - 1) + c(n - 2) + \cdots + c \cdot 2 + c \in \Theta(n^2)$

**Best-case analysis:** First chosen pivot could be the $k$th element 

No recursive calls; total cost is $\Theta(n)$.

**Average case analysis?**
Sorting Permutations

- Need to take average running time over all inputs.
- How to characterize input of size $n$?
  (There are infinitely many sets of $n$ numbers.)
- **Simplifying assumption:** All input numbers are distinct.
- Observe: quick-select1 would act the same on inputs
  14, 2, 3, 6, 1, 11, 7 and
  14, 2, 4, 6, 1, 12, 8
- The actual numbers do not matter, only their relative order.
- Characterize input via **sorting permutation**: the permutation that
  would put the input in order.
- Assume all $n!$ permutations are equally likely.

Average cost is sum of costs for all permutations, divided by $n!$
Average-Case Analysis of *quick-select*1

Define $T(n, k)$ as average cost for selecting $k$th item from size-$n$ array. Then $T(1, k) = c$ and

$$T(n, k) = cn + \frac{1}{n} \left( \sum_{i=0}^{k-1} T(n - i - 1, k - i - 1) + \sum_{i=k+1}^{n-1} T(i, k) \right)$$

**Proof:**

- For $i = 0, \ldots, n - 1$, a fraction of $1/n$ of all permutations has pivot index $i$.
- The average runtime for these permutations is

\[
\begin{align*}
    cn + T(n - i - 1, k - i - 1) & \quad \text{if } i < k \\
    cn & \quad \text{if } i = k \\
    cn + T(i, k) & \quad \text{if } k < i.
\end{align*}
\]
Average-Case Analysis of \textit{quick-select1}

\[
T(n, k) = cn + \frac{1}{n} \left( \sum_{i=0}^{k-1} T(n-i-1, k-i-1) + \sum_{i=k+1}^{n-1} T(i, k) \right)
\]

\textbf{Theorem:} \(T(n, k) \leq 4cn.\)

\textbf{Proof:} By induction on \(n\)
Outline

1. Sorting and Randomized Algorithms
   - QuickSelect
   - Randomized Algorithms
   - QuickSort
   - Lower Bound for Comparison-Based Sorting
   - Non-Comparison-Based Sorting
A randomized algorithm is one which relies on some random numbers in addition to the input.

Computers cannot generate randomness. We assume that there exists a pseudo-random number generator (PRNG), a deterministic program that uses an initial value or seed to generate a sequence of seemingly random numbers. The quality of randomized algorithms depends on the quality of the PRNG!

- The run-time will depend on the input and the random numbers used.
- **Goal:** Shift the dependency of run-time from what we can’t control (the input) to what we *can* control (the random numbers).

*No more bad instances, just unlucky numbers.*
Expected running time

Define $T(I, R)$ to be the running time of a randomized algorithm $\mathcal{A}$ for an instance $I$ and the sequence of random numbers $R$.

The expected running time $T^{\text{exp}}(I)$ for instance $I$ is the expected value for $T(I, R)$:

$$T^{\text{exp}}(I) = E[T(I, R)] = \sum_R T(I, R) \cdot \Pr[R]$$

- We could now take the maximum or the average over all instances of size $n$ to define the expected running time of $\mathcal{A}$.
- But we usually design $\mathcal{A}$ such that all instances of size $n$ have the same expected run-time.
- Then maximum and average are the same, so we have

$$T^{\text{exp}}(n) := \max_{\{I : \text{size}(I) = n\}} T^{\text{exp}}(I) = \frac{\sum_{\{I : \text{size}(I) = n\}} T^{\text{exp}}(I)}{|\{I : \text{size}(I) = n\}|}$$
Randomized QuickSelect: Shuffle

**Goal:** Create a randomized version of *QuickSelect* for which all input has the same expected run-time.

**First idea:** Randomly permute the input first using *shuffle*:

```plaintext
shuffle(A)
A: array of size n
1. for i ← 0 to n − 2 do
2. swap( A[i], A[i + random(n − i)] )
```

We assume the existence of a function `random(n)` that returns an integer uniformly from \{0, 1, 2, \ldots, n − 1\}.

*Expected cost* becomes the same as the average cost: $\Theta(n)$. 

Randomized QuickSelect: Random Pivot

Second idea: Change the pivot selection.

\[
\text{choose-pivot2}(A)
\]
1. \( \text{return } \text{random}(n) \)

\[
\text{quick-select2}(A, k)
\]
1. \( p \leftarrow \text{choose-pivot2}(A) \)
2. \( \ldots \)

With probability \( \frac{1}{n} \) the random pivot has index \( i \), so the analysis is just like that for the average-case. The expected running time is again \( \Theta(n) \).

This is generally the fastest quick-select implementation.

There exists a variation that has worst-case running time \( O(n) \), but it uses double recursion and is slower in practice. (\( \rightsquigarrow \text{cs341} \))
Outline

1. Sorting and Randomized Algorithms
   - QuickSelect
   - Randomized Algorithms
   - QuickSort
   - Lower Bound for Comparison-Based Sorting
   - Non-Comparison-Based Sorting
QuickSort

Hoare developed *partition* and *quick-select* in 1960. He also used them to *sort* based on partitioning:

```
quick-sort1(A)
A: array of size n
1. if n ≤ 1 then return
2. p ← choose-pivot1(A)
3. i ← partition(A, p)
4. quick-sort1(A[0, 1, ..., i − 1])
5. quick-sort1(A[i + 1, ..., n − 1])
```
QuickSort analysis

Define $T(n)$ to be the run-time for quick-sort1 in a size-$n$ array.

- $T(n)$ depends again on the pivot-index $i$.
- If we know $i$: $T(n) = \Theta(n) + T(i) + T(n - i - 1)$.

**Worst-case analysis**: $i = 0$ or $n–1$ always. Then as for quick-select

$$T(n) = \begin{cases} 
T(n - 1) + cn, & n \geq 2 \\
c, & n = 1
\end{cases}$$

for some constant $c > 0$. This resolves to $\Theta(n^2)$.

**Best-case analysis**: $i = \lfloor \frac{n}{2} \rfloor$ or $\lceil \frac{n}{2} \rceil$ always. Then

$$T(n) = \begin{cases} 
T(\lfloor \frac{n-1}{2} \rfloor) + T(\lceil \frac{n-1}{2} \rceil) + cn, & n \geq 2 \\
c, & n = 1
\end{cases}$$

Similar to merge-sort: This resolves to $\Theta(n \log n)$. 
Average-case analysis of *quick-sort1*

Now let $T(n)$ to be the *average-case* run-time for *quick-sort1* in a size-$n$ array.

- As before, $(n - 1)!$ permutations have pivot-index $i$.
- So average running time is

$$T(n) = \frac{1}{n!} \sum_{i=0}^{n-1} \sum_{l : \text{size}(l) = n \atop l \text{ has pivot-index } i} \text{running time for instance } l$$

$$\leq \frac{1}{n!} \sum_{i=0}^{n-1} (n - 1)! \left( c \cdot n + T(i) + T(n - i - 1) \right)$$

$$= c \cdot n + \frac{1}{n} \sum_{i=0}^{n-1} \left( T(i) + T(n - i - 1) \right)$$

**Theorem:** $T(n) \in \Theta(n \log n)$.

**Proof:** Can prove that $T(n) \leq 2cn \log(n)$ by induction on $n$.
Improvement ideas for QuickSort

- We can randomize by using \textit{choose-pivot2}, giving $\Theta(n \log n)$ expected time for \textit{quick-sort2}.

- The auxiliary space is $\Omega$(recursion depth).
  - This is $\Theta(n)$ in the worst-case.
  - It can be reduced to $\Theta(\log n)$ worst-case by recursing in smaller sub-array first and replacing the other recursion by a while-loop.

- One should stop recursing when $n \leq 10$.
  One run of InsertionSort at the end then sorts everything in $O(n)$ time since all items are within 10 units of their required position.

- Arrays with many duplicates can be sorted faster by changing \textit{partition} to produce three subsets $\leq v, = v, \geq v$.

- Two programming tricks that apply in many situations:
  - Instead of passing full arrays, pass only the range of indices.
  - Avoid recursion altogether by keeping an explicit stack.
QuickSort with tricks

\[
\text{quick-sort3}(A, n)
\]
1. Initialize a stack \( S \) of index-pairs with \( \{ (0, n-1) \} \)
2. \textbf{while} \( S \) is not empty
3. \( (\ell, r) \leftarrow S.p\text{op}() \)
4. \textbf{while} \( (r-\ell+1 > 10) \) do
5. \( p \leftarrow \text{choose-pivot2}(A, \ell, r) \)
6. \( i \leftarrow \text{partition}(A, \ell, r, p) \)
7. \textbf{if} \( (i-\ell > r-i) \) do
8. \( S.p\text{ush}( (\ell, i-1)) \)
9. \( \ell \leftarrow i+1 \)
10. \textbf{else}
11. \( S.p\text{ush}( (i+1, r)) \)
12. \( r \leftarrow i-1 \)
13. \( \text{InsertionSort}(A) \)

This is often the most efficient sorting algorithm in practice.
Outline

1. Sorting and Randomized Algorithms
   - QuickSelect
   - Randomized Algorithms
   - QuickSort
   - Lower Bound for Comparison-Based Sorting
   - Non-Comparison-Based Sorting
Lower bounds for sorting

We have seen many sorting algorithms:

<table>
<thead>
<tr>
<th>Sort</th>
<th>Running time</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection Sort</td>
<td>$\Theta(n^2)$</td>
<td>worst-case</td>
</tr>
<tr>
<td>Insertion Sort</td>
<td>$\Theta(n^2)$</td>
<td>worst-case</td>
</tr>
<tr>
<td>Merge Sort</td>
<td>$\Theta(n \log n)$</td>
<td>worst-case</td>
</tr>
<tr>
<td>Heap Sort</td>
<td>$\Theta(n \log n)$</td>
<td>worst-case</td>
</tr>
<tr>
<td>quick-sort1</td>
<td>$\Theta(n \log n)$</td>
<td>average-case</td>
</tr>
<tr>
<td>quick-sort2</td>
<td>$\Theta(n \log n)$</td>
<td>expected</td>
</tr>
</tbody>
</table>

**Question:** Can one do better than $\Theta(n \log n)$ running time?

**Answer:** Yes and no! *It depends on what we allow.*

- **No:** Comparison-based sorting lower bound is $\Omega(n \log n)$.
- **Yes:** Non-comparison-based sorting can achieve $O(n)$ (under restrictions!). → see below
The Comparison Model

In the **comparison model** data can only be accessed in two ways:

- comparing two elements
- moving elements around (e.g. copying, swapping)

This makes very few assumptions on the kind of things we are sorting. We count the number of above operations.

All sorting algorithms seen so far are in the comparison model.
Decision trees

Comparison-based algorithms can be expressed as decision tree. To sort \( \{x_0, x_1, x_2\} \):

(i.e., \( x_0 \leq x_1 \leq x_2 \))
Lower bound for sorting in the comparison model

**Theorem.** Any correct *comparison-based* sorting algorithm requires at least $\Omega(n \log n)$ comparison operations to sort $n$ distinct items.

**Proof.**
Outline

1. Sorting and Randomized Algorithms
   - QuickSelect
   - Randomized Algorithms
   - QuickSort
   - Lower Bound for Comparison-Based Sorting
   - Non-Comparison-Based Sorting
Non-Comparison-Based Sorting

- Assume keys are numbers in base $R$ (R: radix)
  - $R = 2, 10, 128, 256$ are the most common.

  Example ($R = 4$): 123 230 21 320 210 232 101

- Assume all keys have the same number $m$ of digits.
  - Can achieve after padding with leading 0s.

  Example ($R = 4$): 123 230 021 320 210 232 101

- Can sort based on individual digits.
  - How to sort 1-digit numbers?
  - How to sort multi-digit numbers based on this?
(Single-digit) Bucket Sort

Sort array $A$ by last digit:

$$
\begin{array}{c|c|c|c|c|c}
123 & 230 & 021 & 232 & 123 & \\
230 & 320 & 101 & & & \\
021 & & & & & \\
320 & & & & & \\
210 & & & & & \\
232 & & & & & \\
101 & & & & & \\
123 & & & & & \\
\end{array}
$$
(Single-digit) Bucket Sort

- Sorts numbers by a single digit.
- Create a “bucket” for each possible digit: Array $B[0..R-1]$ of lists
- Copy item with digit $i$ into bucket $B[i]$
- At the end copy buckets in order into $A$.

```
Bucket-sort(A, d)
A: array of size $n$, contains numbers with digits in \{0, \ldots, R - 1\}
d: index of digit by which we wish to sort

1. Initialize an array $B[0...R - 1]$ of empty lists
2. for $i \leftarrow 0$ to $n - 1$ do
3. \hspace{1cm} Append $A[i]$ at end of $B[d^{th} \text{ digit of } A[i]]$
4. \hspace{1cm} $i \leftarrow 0$
5. for $j \leftarrow 0$ to $R - 1$ do
6. \hspace{1cm} while $B[j]$ is non-empty do
7. \hspace{2cm} move first element of $B[j]$ to $A[i++]$
```

- This is stable: equal items stay in original order.
- Run-time $\Theta(n + R)$, auxiliary space $\Theta(n)$
Count Sort

- Bucket sort wastes space for linked lists.
- Observe: We know exactly where numbers in $B[j]$ go:
  - The first of them is at index $|B[0]| + |B[1]| + \cdots + |B[j-1]|$
  - The others follow.
- So we don’t need the lists; it’s enough to count how many there would be in it.
Count Sort Pseudocode

\[ \text{count-sort}(A, d) \]
\n\( A \): array of size \( n \), contains numbers with digits in \( \{0, \ldots, R - 1\} \)
\( d \): index of digit by which we wish to sort

// count how many of each kind there are
1. \( \text{count} \leftarrow \text{array of size } R, \text{ filled with zeros} \)
2. \( \text{for } i \leftarrow 0 \text{ to } n - 1 \text{ do} \)
3. \( \text{increment } \text{count}[d^{\text{th}} \text{ digit of } A[i]] \)

// find left boundary for each kind
4. \( \text{idx} \leftarrow \text{array of size } R, \text{ idx}[0] = 0 \)
5. \( \text{for } i \leftarrow 1 \text{ to } R - 1 \text{ do} \)
6. \( \text{idx}[i] \leftarrow \text{idx}[i - 1] + \text{count}[i - 1] \)

// move to new array in sorted order, then copy back
7. \( \text{aux} \leftarrow \text{array of size } n \)
8. \( \text{for } i \leftarrow 0 \text{ to } n - 1 \text{ do} \)
9. \( \text{aux}[\text{idx}[d^{\text{th}} \text{ digit of } A[i]]] \leftarrow A[i] \)
10. \( \text{increment } \text{idx}[d^{\text{th}} \text{ digit of } A[i]] \)
11. \( A \leftarrow \text{copy}(\text{aux}) \)
Example: Count Sort

<table>
<thead>
<tr>
<th>A</th>
<th>count</th>
<th>idx</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>12③</td>
<td>0, 1, 2, 3</td>
<td>0</td>
<td>230</td>
</tr>
<tr>
<td>23①</td>
<td>0, 1, 2</td>
<td>3</td>
<td>320</td>
</tr>
<tr>
<td>02①</td>
<td>0, 1</td>
<td>5</td>
<td>210</td>
</tr>
<tr>
<td>32①</td>
<td>0, 1</td>
<td>6</td>
<td>021</td>
</tr>
<tr>
<td>21①</td>
<td></td>
<td></td>
<td>101</td>
</tr>
<tr>
<td>23①</td>
<td></td>
<td></td>
<td>232</td>
</tr>
<tr>
<td>10①</td>
<td></td>
<td></td>
<td>123</td>
</tr>
</tbody>
</table>

$\Rightarrow$
MSD-Radix-Sort

Sorts array of \( m \)-digit radix-\( R \) numbers recursively:
sort by leading digit, then each group by next digit, etc.

\[
\text{MSD-Radix-sort}(A, \ell \leftarrow 0, \ r \leftarrow n-1, \ d \leftarrow 1)
\]

1. \( \text{if } \ell < r \)
2. \( \text{count-sort}(A[\ell..r], d) \)
3. \( \text{if } d < m \)
4. \( \text{for } i \leftarrow 0 \text{ to } R - 1 \text{ do} \)
5. \( \text{let } \ell_i \text{ and } r_i \text{ be boundaries of } i\text{th bin} \)
6. \( \text{(i.e., } A[\ell_i..r_i] \text{ all have } d\text{th digit } i) \)
7. \( \text{MSD-Radix-sort}(A, \ell_i, r_i, d+1) \)

- \( \ell_i \text{ and } r_i \) are automatically computed with count-sort
MSD-Radix-Sort Example

- **Drawback of MSD-Radix-Sort**: many recursions
- **Auxiliary space**: $\Theta(n + R + m)$ (for count-sort and recursion stack)
- **Run-time**: $\Theta(m(n + R))$ can be achieved if auxiliary arrays for count-sort are shared across subroutines.
LSD-Radix-Sort

\[ \text{LSD-radix-sort}(A) \]

\[ A: \text{array of size } n, \text{ contains } m\text{-digit radix-} R \text{ numbers} \]

1. for \( d \leftarrow m \) down to 1 do
2. \( \text{count-sort}(A, d) \)

Loop-invariant: \( A \) is sorted w.r.t. digits \( d, \ldots, m \) of each entry.

Time cost: \( \Theta(m(n + R)) \) \hspace{1cm} Auxiliary space: \( \Theta(n + R) \)
Summary

- Sorting is an important and very well-studied problem.
- Can be done in $\Theta(n \log n)$ time; faster is not possible for general input.
- **HeapSort** is the only $\Theta(n \log n)$-time algorithm we have seen with $O(1)$ auxiliary space.
- **MergeSort** is also $\Theta(n \log n)$, selection & insertion sorts are $\Theta(n^2)$.
- **QuickSort** is worst-case $\Theta(n^2)$, but often the fastest in practice.
- **CountSort** and **RadixSort** achieve $o(n \log n)$ if the input is special.

- Randomized algorithms can eliminate “bad cases”
- Best-case, worst-case, average-case, expected-case can all differ, but for well-designed randomizations of algorithms, the expected case is the same as the average-case of the non-randomized algorithm.