Algorithms and Data Structures

A14. Sorting: Counting Sort & Radix Sort

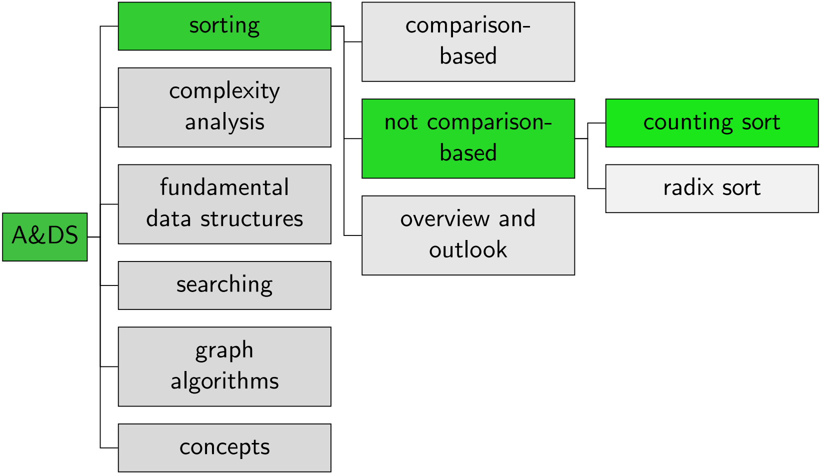
Gabriele Röger

University of Basel

March 21, 2024

Counting Sort

Content of the Course


A&DS

sorting

complexity analysis

fundamental data structures

searching

graph algorithms

concepts

comparison-based

not comparison-based

overview and outlook

counting sort

radix sort

Counting Sort: Idea

„Sort by counting“

- Assumption: Elements are from the range $0, \ldots, k-1$.
- Iterate once over the input array and count the number of occurrences of each element.
- Let $\#i$ be the number of occurrences of element i.
- For $i = 0, \ldots, k-1$ write $\#i$ times element i into the sequence.

Counting Sort: Algorithm

```python
def sort(array, k):
    counts = [0] * k  # list of k zeros
    for elem in array:
        counts[elem] += 1

    pos = 0
    for i in range(k):
        occurrences_of_i = counts[i]
        for j in range(occurrences_of_i):
            array[pos + j] = i
        pos += occurrences_of_i
```

Counting Sort: Algorithm

```
def sort(array, k):
    counts = [0] * k  # list of k zeros
    for elem in array:
        counts[elem] += 1

    pos = 0
    for i in range(k):
        occurrences_of_i = counts[i]
        for j in range(occurrences_of_i):
            array[pos + j] = i
        pos += occurrences_of_i
```

Running time: $O(n + k)$ (n size of input sequence)

Counting Sort: Algorithm

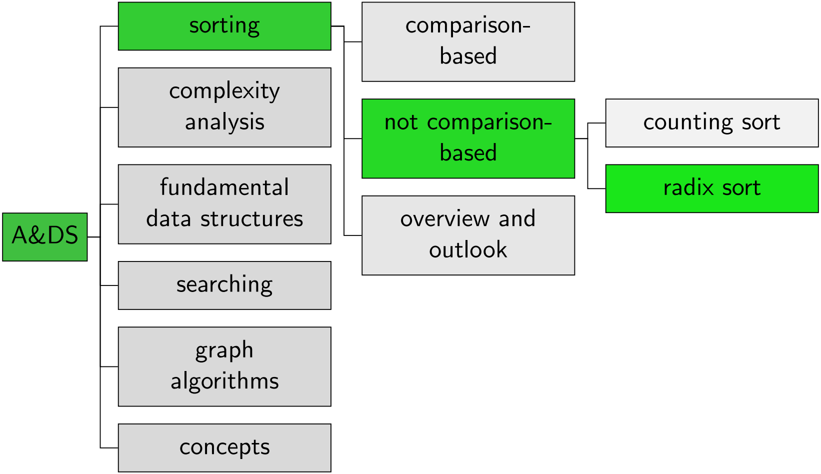
```python
def sort(array, k):
    counts = [0] * k  # list of k zeros
    for elem in array:
        counts[elem] += 1

    pos = 0
    for i in range(k):
        occurrences_of_i = counts[i]
        for j in range(occurrences_of_i):
            array[pos + j] = i
        pos += occurrences_of_i
```

Running time: $O(n + k)$ (n size of input sequence)

→ For fixed k or $k \in O(n)$ linear.

Radix Sort

Content of the Course



Radix Sort: Idea

- Assumption: Keys are decimal numbers
  z.B. 763, 983, 96, 286, 462

Radix Sort: Idea

- Assumption: Keys are decimal numbers
  z.B. 763, 983, 96, 286, 462
- Separate items by the least significant (= last) digit:

0	1	2	3	4	5	6	7	8	9
		462	763			96			
			983			286			

Radix Sort: Idea

- Assumption: Keys are decimal numbers
  z.B. 763, 983, 96, 286, 462
- Separate items by the least significant (= last) digit:

0	1	2	3	4	5	6	7	8	9
		462	763			96			
			983			286			

- Collect items from left to right/top to bottom:
  462, 763, 983, 96, 286

Radix Sort: Idea

- Assumption: Keys are decimal numbers
  z.B. 763, 983, 96, 286, 462
- Separate items by the least significant (= last) digit:

0	1	2	3	4	5	6	7	8	9
		462	763			96			
			983			286			

- Collect items from left to right/top to bottom:
  462, 763, 983, 96, 286
- Separate items by the second last digit and collect them.
- Separate items by the third last digit and collect them.
- . . . until you considered all positions of digits.

Radix Sort: Example

- Input: 263, 983, 96, 462, 286

Radix Sort: Example

- Input: 263, 983, 96, 462, 286
- Separation by last digit:

0	1	2	3	4	5	6	7	8	9
		462	263			96			
			983			286			

After collection: 462, 263, 983, 96, 286

Radix Sort: Example

- Input: 263, 983, 96, 462, 286
- Separation by last digit:

0	1	2	3	4	5	6	7	8	9
		462	263			96			
			983			286			

After collection: 462, 263, 983, 96, 286

- Separation by second last digit:

0	1	2	3	4	5	6	7	8	9
						462		983	96
						263		286	

After collection: 462, 263, 983, 286, 96

Radix Sort: Example

- Input: 263, 983, 96, 462, 286
- Separation by last digit:

0	1	2	3	4	5	6	7	8	9
		462	263			96			
			983			286			

After collection: 462, 263, 983, 96, 286

- Separation by second last digit:

0	1	2	3	4	5	6	7	8	9
						462		983	96
						263		286	

After collection: 462, 263, 983, 286, 96

- Separation by third last digit:

0	1	2	3	4	5	6	7	8	9
096		263		462					983
		286							

After collection: 96, 263, 286, 462, 983

Jupyter Notebook

Jupyter notebook: `radix_sort.ipynb`

Radix Sort: Algorithm (for arbitrary base)

```python
def sort(array, base=10):
    if not array:  # array is empty
        return
    iteration = 0
    max_val = max(array) # identify largest element
    while base ** iteration <= max_val:
        buckets = [[] for num in range(base)]
        for elem in array:
            digit = (elem // (base ** iteration)) % base
            buckets[digit].append(elem)
        pos = 0
        for bucket in buckets:
            for elem in bucket:
                array[pos] = elem
                pos += 1
        iteration += 1
```

Radix Sort: Running Time

- m: Maximal number of digits in representation with given base b.
- n: length of input sequence
- Running time $O(m \cdot (n + b))$

Radix Sort: Running Time

- m: Maximal number of digits in representation with given base b.
- n: length of input sequence
- Running time $O(m \cdot (n + b))$

For fixed m and b, radix sort has linear running time.

Radix Sort: High-level Perspective

All entries in the array have d digits, where the lowest-order digit is at position 0 and the highest-order digit at position d-1.

```
def radix_sort(array, d)
  for i in range(d):
    # use a stable sort to sort array on the digit at position i
```

Summary

Summary

- Counting sort and radix sort are not comparison-based and allow us (under certain restrictions) to sort in linear time.
- However, they place additional restrictions on the keys used.