COEN-352
Tutorial #7

June 6th, 2023

Radix Sort

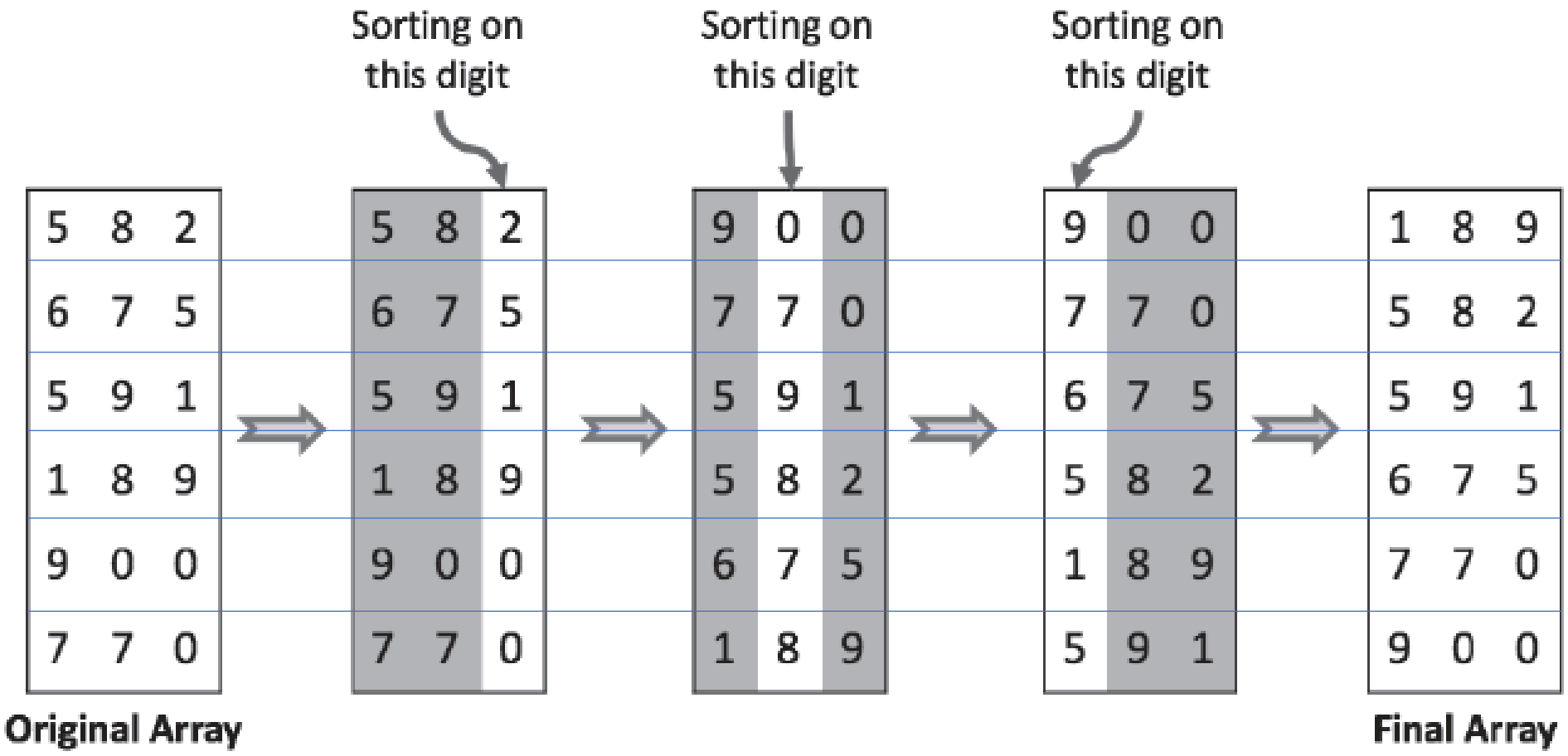
RadixSort: An almost linear sorting algorithm that does digit-by-digit sorting.

- It is not a comparative sorting algorithm.
- It uses Counting Sort as a subroutine to sort occurrences.

Time Complexity	
Best	O(n+k)
Worst	O(n+k)
Average	O(n+k)
Space Complexity	O(max)
Stability	O(n+k)

QUESTION: what is the lower bound of the algorithms we have seen already?

Radix Sort Illustration



Huffman Coding

Def: Huffman coding is a lossless data compression algorithm that uses priority binary trees.

- Each leaf represents an encoding
- Internal nodes also have a frequency weight
- Usually, left branches represent a `'0'` bit and right is `'1'` bit
- There could be more than one possible encoding

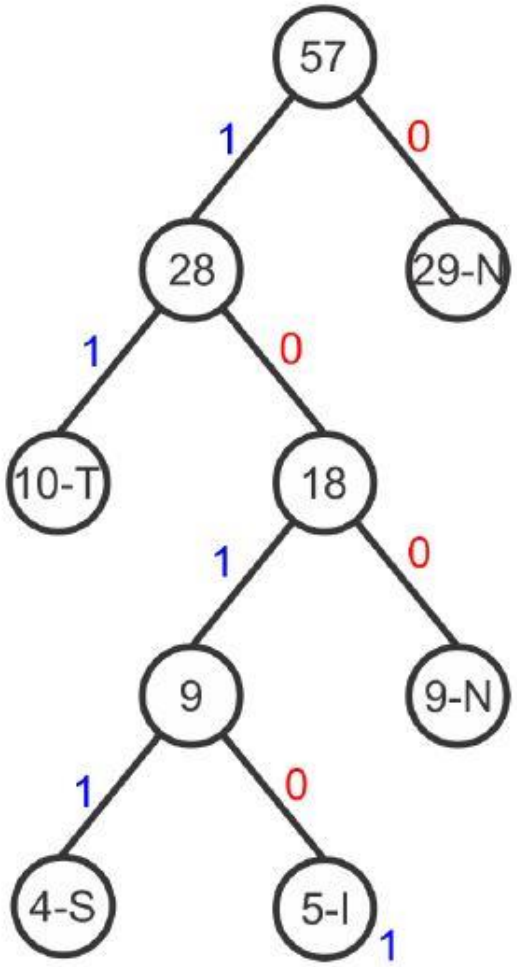
Huffman Coding is heavily used in data compression without losing any of the details.

- Compared to an ASCII encoding, in Huffman encoding the number of bits is dynamic not constant.

Example 1: A Huffman Tree

Chars	Frequency	Huff Code
E	29	0
T	10	11
N	9	100
I	5	1010
S	4	1011

Result



Example 2: Creation of the tree

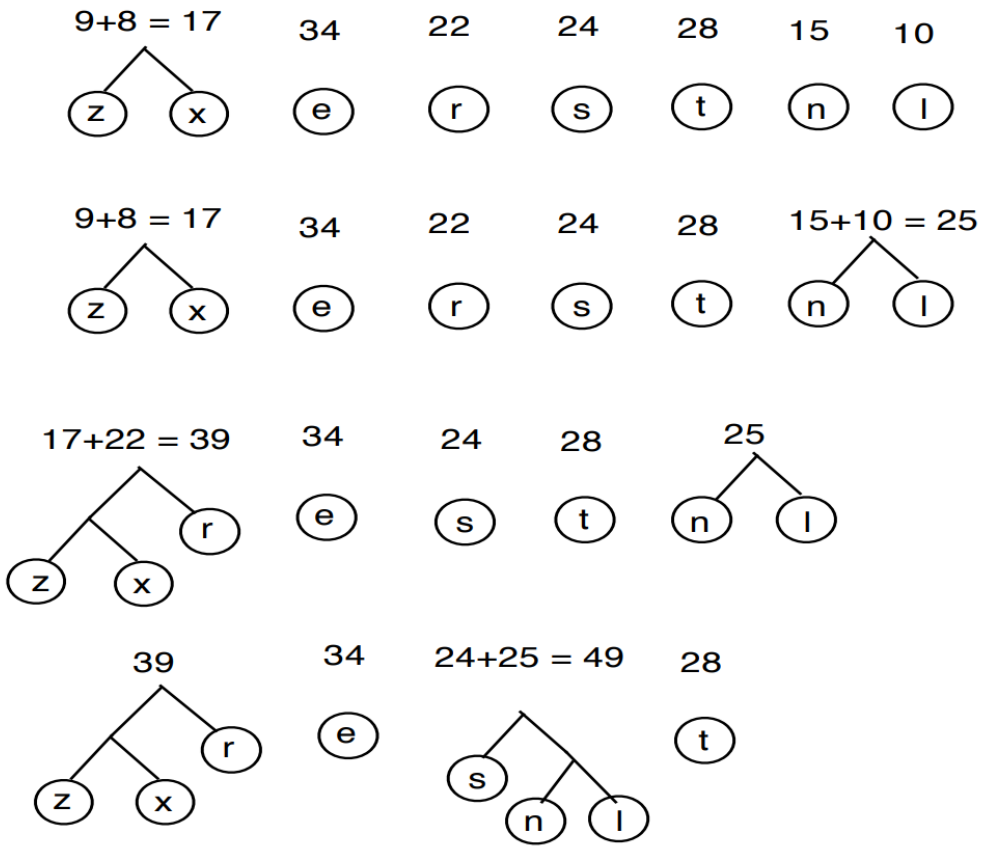
e	r	s	t	n	l	z	x
34	22	24	28	15	10	9	8

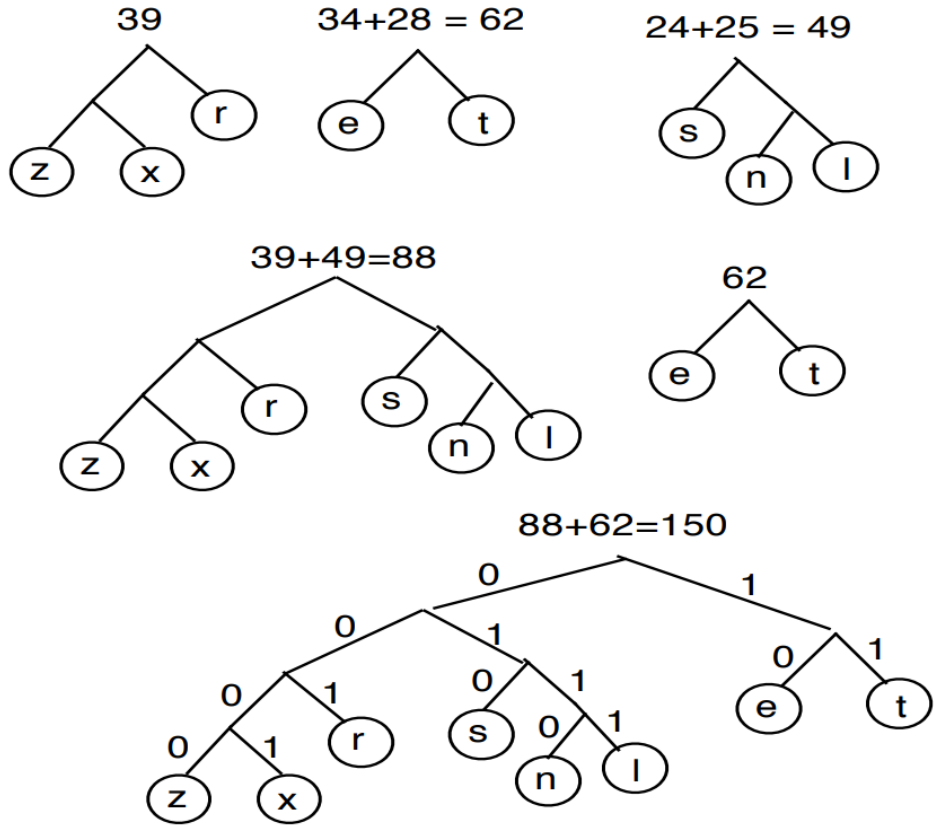
Frequency in an average sample of size 150 letters

The tree is shown in the next page. This leads to the following codes.

z = 0000	n = 0110
x = 0001	l = 0111
r = 001	e = 10
s = 010	t = 11

Example 2: Creation of the tree (contd)





EXERCISES

Exercise: Write an algorithm to return the most frequently repeated character in a Huffman tree.

- Use the implementation of Huffman Coding from the GitHub Repo.
- The algorithm should traverse to where that character lies in the Huffman tree.
- **Q:** What type of traversal are we doing here?

THANK YOU