Compsci 201 Recitation 12

Professor Peck
Reem Mokhtar
11/22/2013

In this Recitation

- Huffman Coding
  - Brief intro
  - Practice
  - More Practice
- Submit via form: http://goo.gl/ivGZ97

Huffman Coding

- Problem:
  - Given a set of symbols and their frequencies (weights), find a way to encode these symbols using minimum encoding length on average
  - Construct a prefix-free binary code:
    - A binary code that satisfies that prefix property, where no codeword is a prefix of any other codeword
    - Example: {0,10,11} has the prefix property. However, {0,1,10,11} does not have that property, because “1” is a prefix of “10” and also “11”.
- How?
  - By constructing a binary tree with minimum average path length from the root

Huffman Coding

- Create the Huffman codes for the following characters.

a	s	p	r	t	o
12	8	2	6	4	10

- Do not forget the PSEUDO_EOF character.

1. First draw the Huffman tree
   a) When merging two trees, make the left child the smallest child.
   b) A '0' means 'left' and a '1' means right when generating codes from trees.
2. Use the tree to generate the binary (i.e. "1010..”) codes for each of the six characters and the PSEUDO_EOF.

Huffman Coding

- Create the Huffman codes for the following characters.

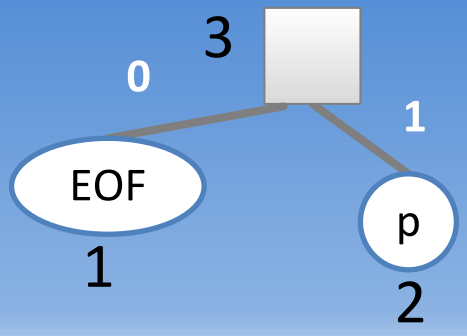
a	s	p	r	t	o
12	8	2	6	4	10

1. What are the codes for p, r, and o?
2. What letter/letters have the shortest codes? Why?
3. What is the encoding for “sport”?
4. One of Professor Peck's favorite cities is
   001 110 110 110 111 110. What city is it?

(Spacing does not denote specific letters. It is only there to make it easier for you to read.)

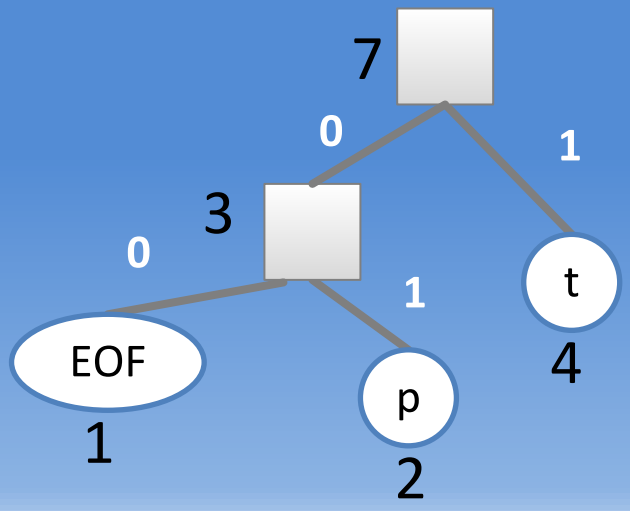
Huffman Coding – Solution

Character	Frequency
EOF	1
p	2
t	4
r	6
s	8
o	10
a	12



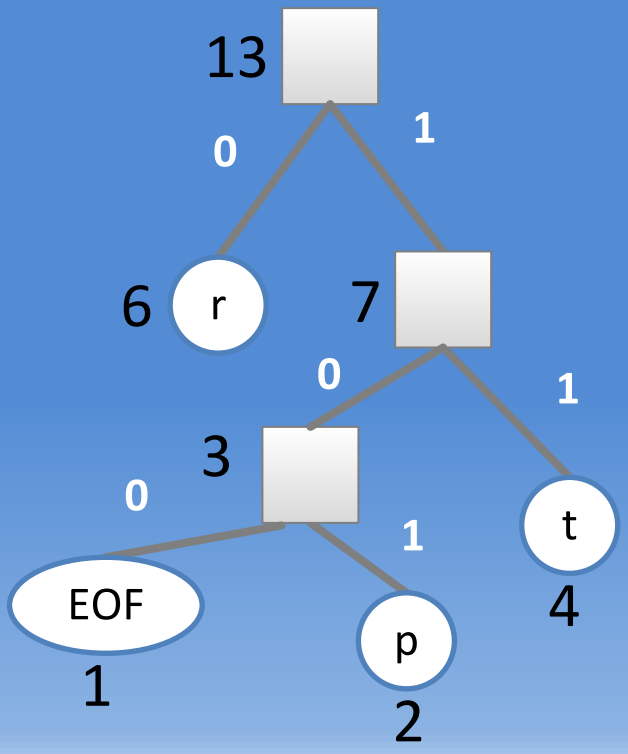
Huffman Coding – Solution

Character	Frequency
EOF, p	3
t	4
r	6
s	8
o	10
a	12



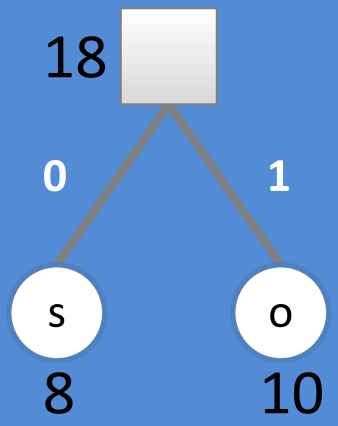
Huffman Coding – Solution

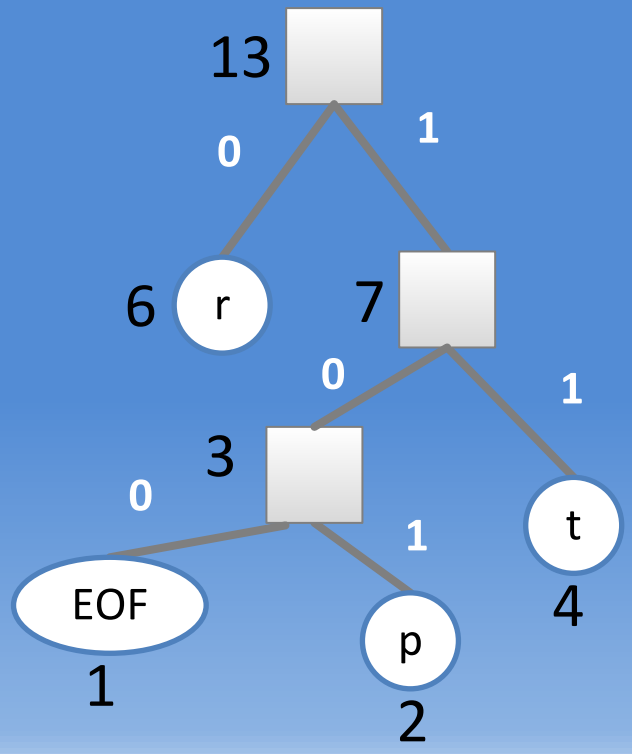
Character	Frequency
r	6
EOF, p, t	7
s	8
o	10
a	12



Huffman Coding – Solution

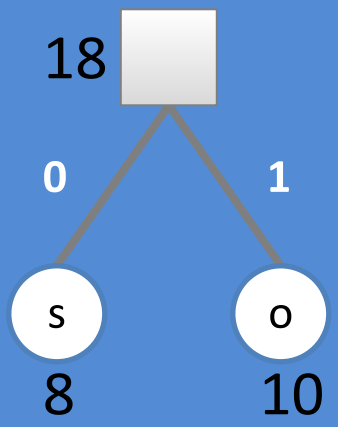
Character	Frequency
s	8
o	10
a	12
EOF, p, t, r	13

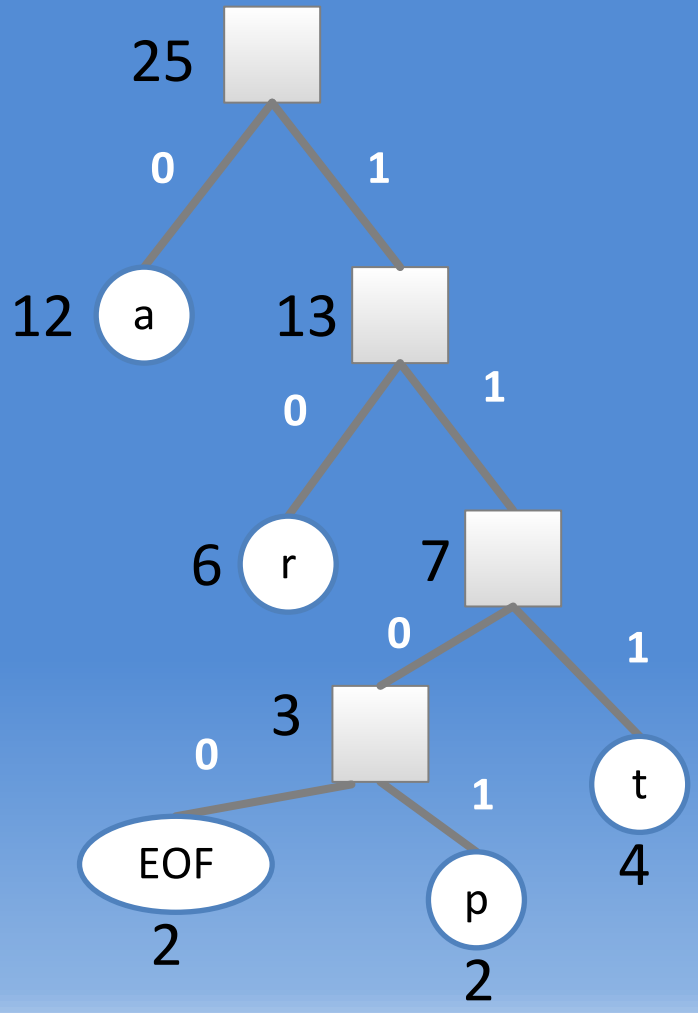




Huffman Coding – Solution

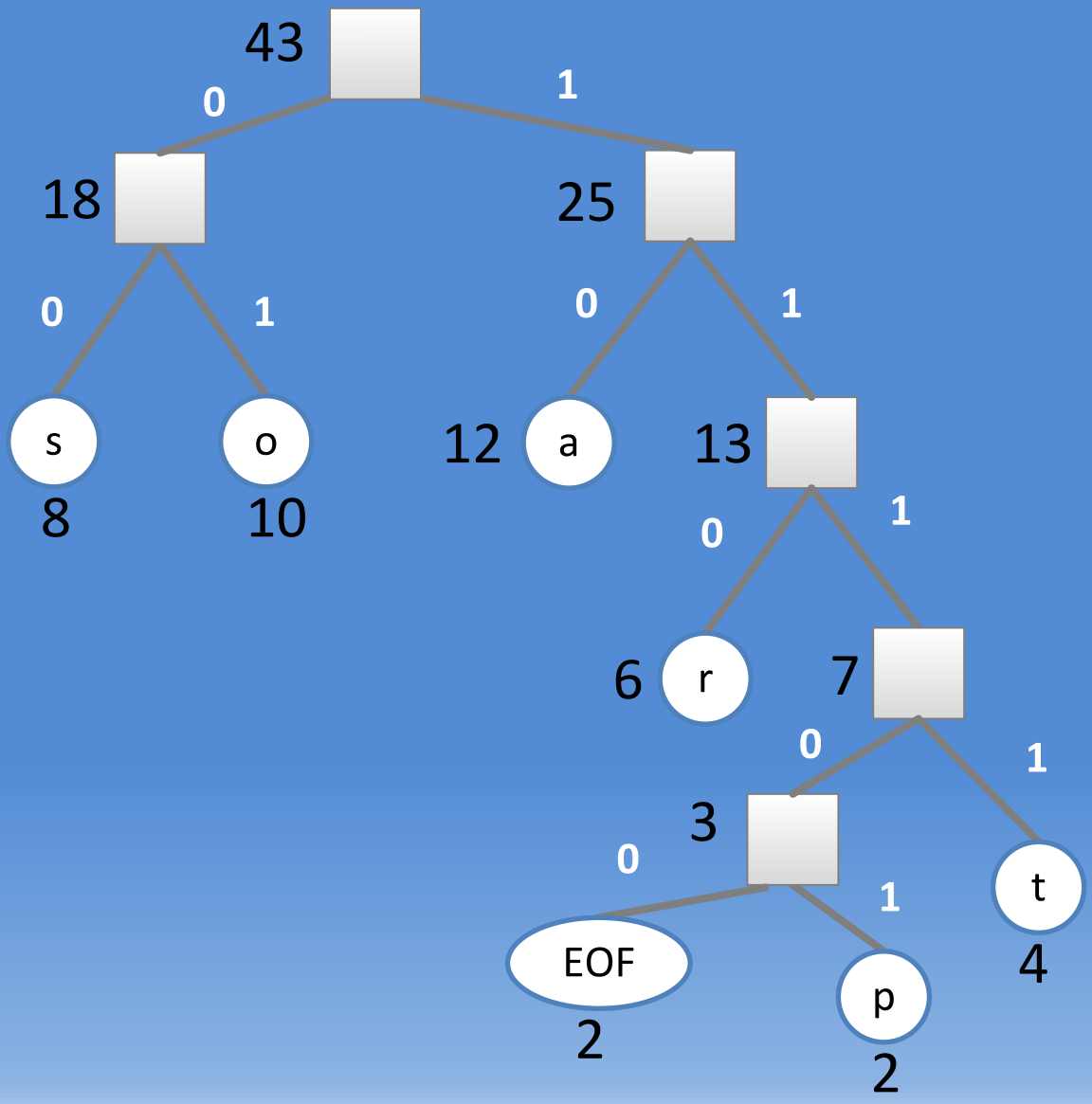
Character	Frequency
a	12
EOF, p, t, r	13
s, o	18





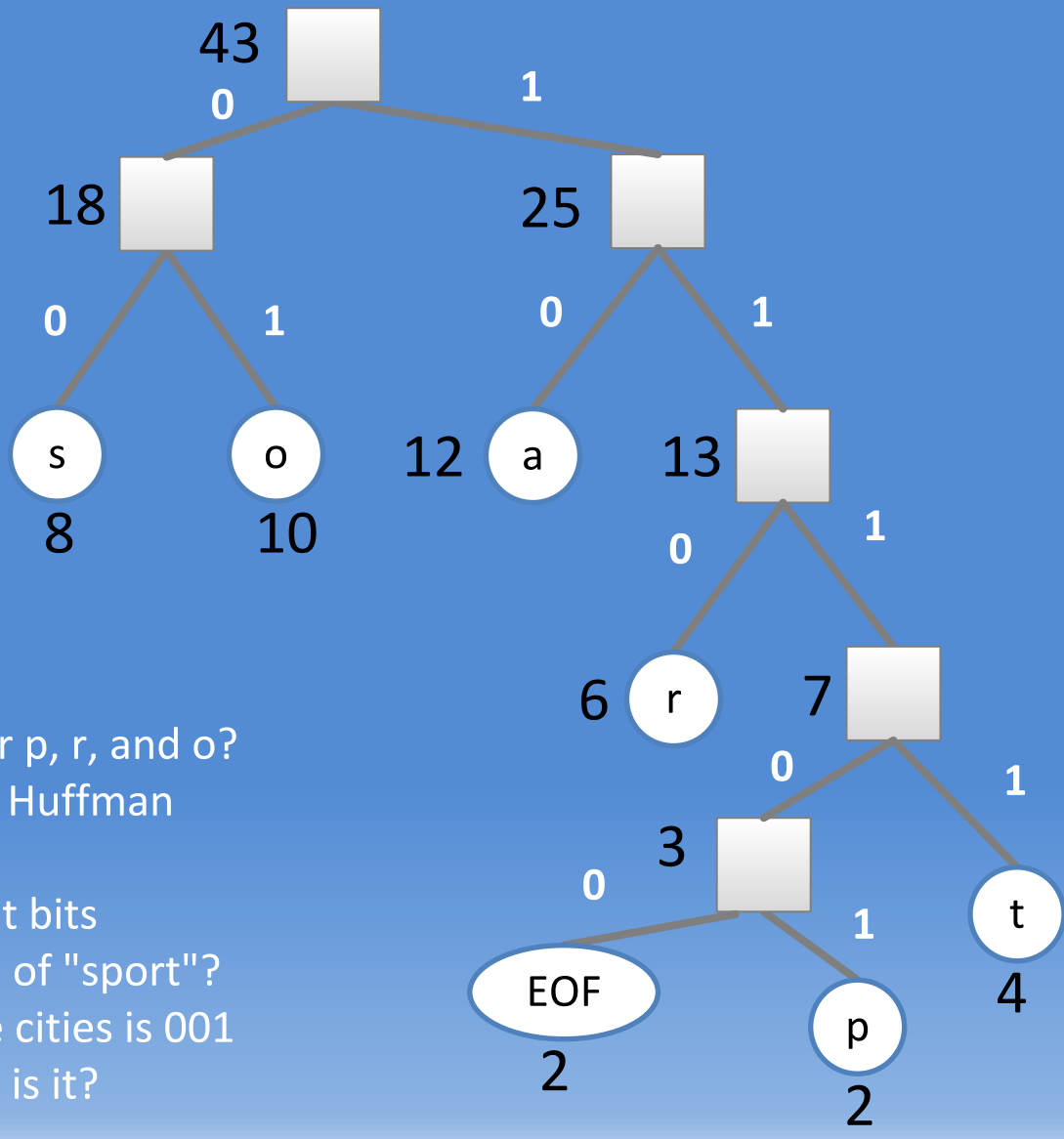
Huffman Coding – Solution

Character	Frequency
s, o	18
EOF, p, t, r, a	25



Huffman Coding – Solution

Character	Frequency	Codewords
EOF	1	11100
p	2	11101
t	3	1111
r	6	110
s	7	00
o	9	01
a	10	10



1. What are the Huffman codes for p, r, and o?
2. Which letters have the shortest Huffman codes? Why?
3. Using your Huffman codes, what bits represent the compressed form of "sport"?
4. One of Professor Peck's favorite cities is 001 110 110 110 111 110. What city is it?

More Questions

See the page linked from the course site on the recitations page.

Have a good weekend!

Don’t forget to submit!