

# Huffman Coding Trees

ASCII codes: 8 bits per character.

- Fixed-length coding.

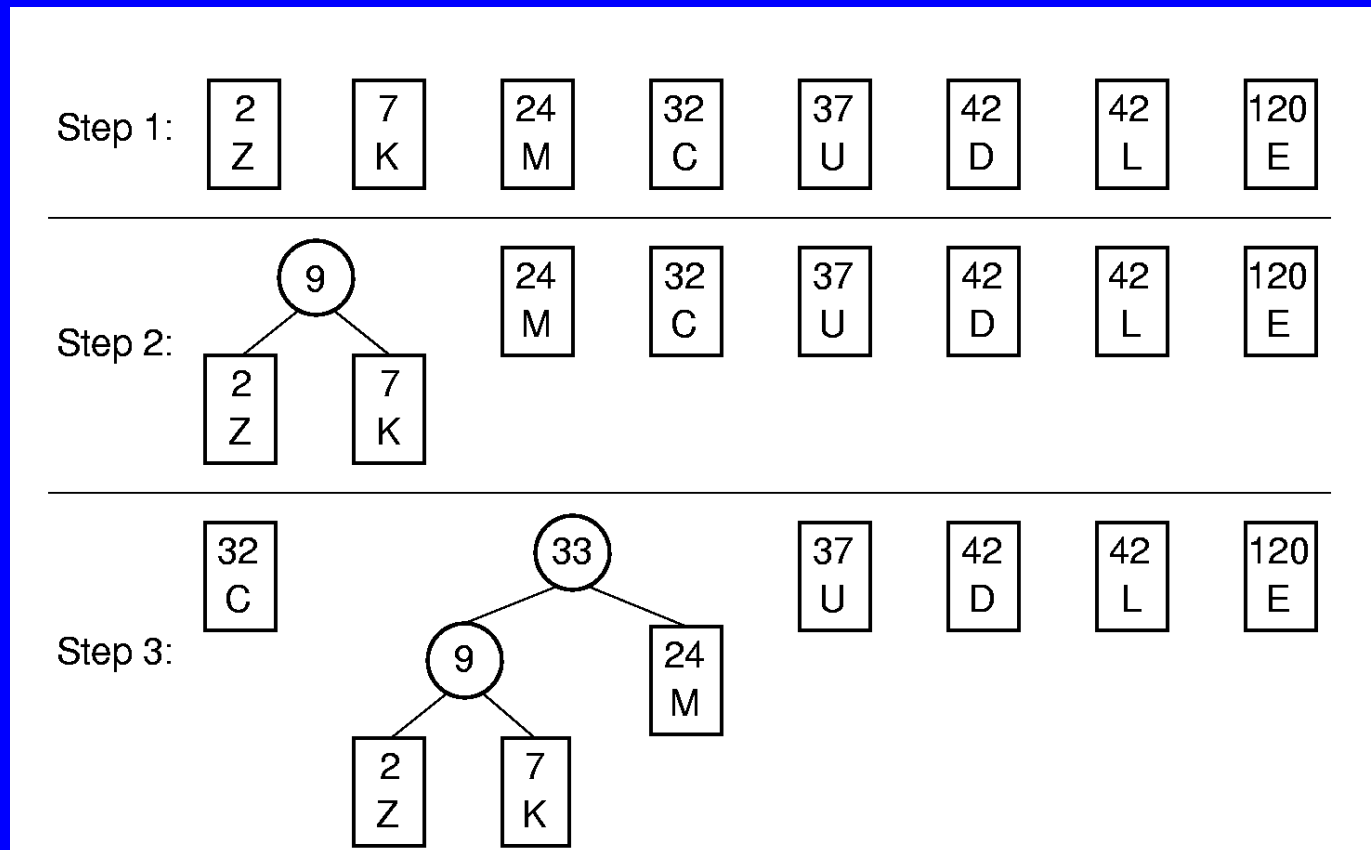
Can take advantage of relative frequency of letters to save space.

- Variable-length coding

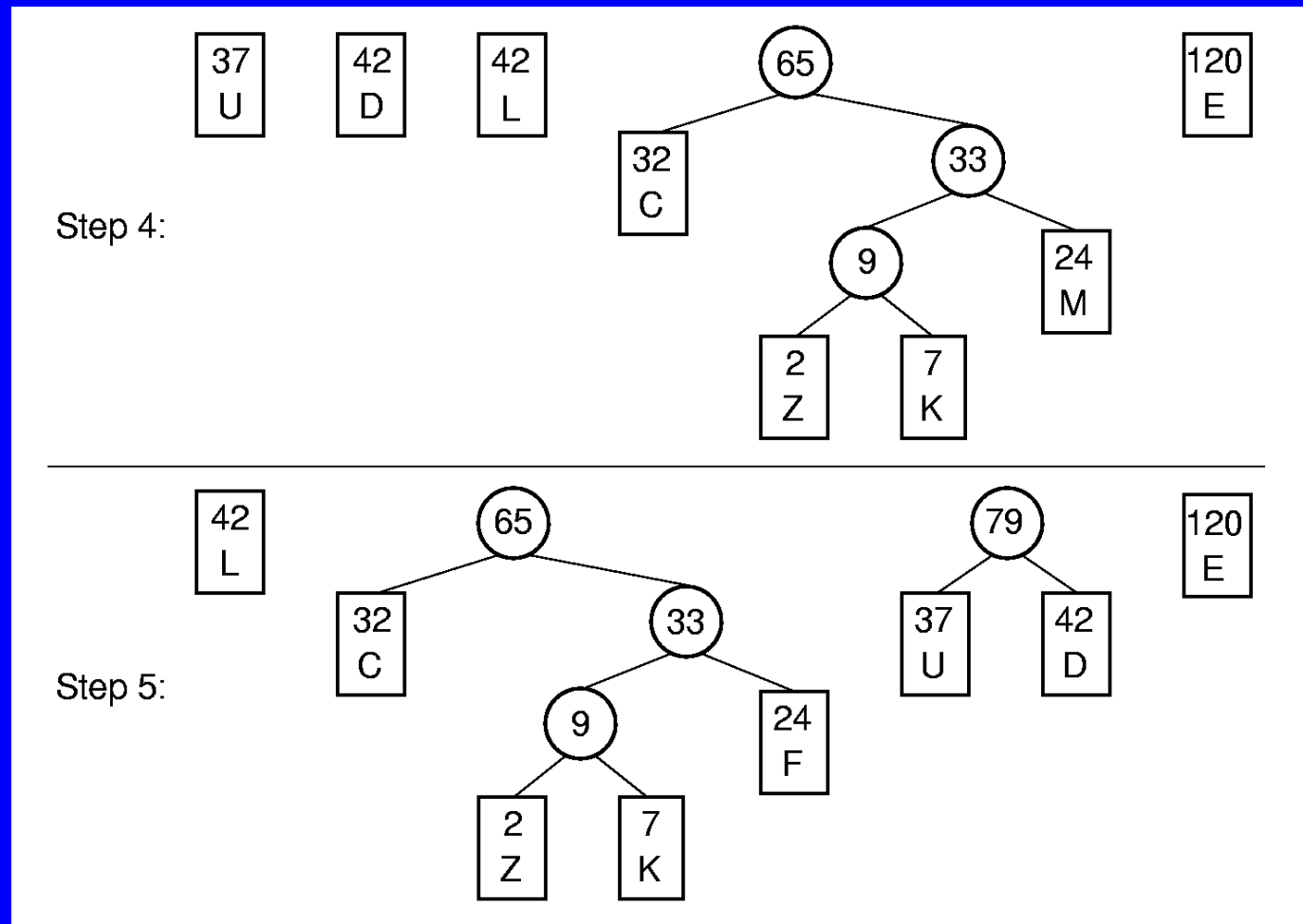
Z	K	M	C	U	D	L	E
2	7	24	32	37	42	42	120

Build the tree with minimum external path weight.

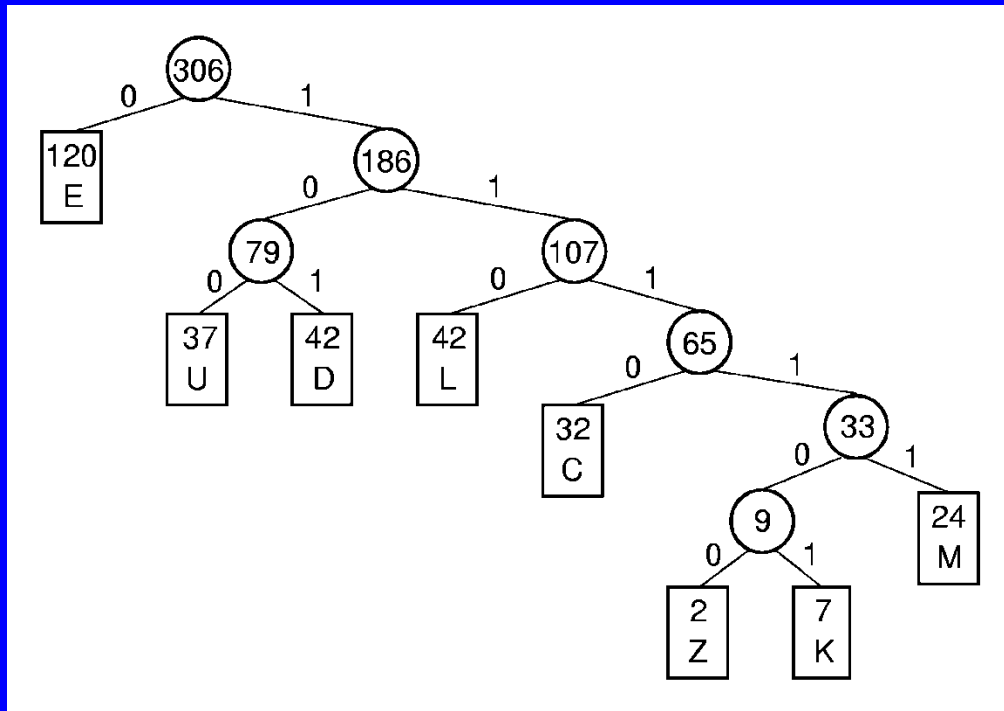
# Huffman Tree Construction (1)



# Huffman Tree Construction (2)



# Assigning Codes



Letter	Freq	Code	Bits
C	32		
D	42		
E	120		
M	24		
K	7		
L	42		
U	37		
Z	2		

# Coding and Decoding

A set of codes is said to meet the prefix property if no code in the set is the prefix of another.

Code for DEED:

Decode 1011001110111101:

Expected cost per letter:

# Search Tree vs. Trie

- In a BST, the root value splits the key range into everything less than or greater than the key
  - The split points are determined by the data values
- View Huffman tree as a search tree
  - All keys starting with 0 are in the left branch, all keys starting with 1 are in the right branch
  - The root splits the key range in half
  - The split points are determined by the data structure, not the data values
  - Such a structure is called a Trie