CMPSC 465 Data Structures & Algorithms
Fall 2025 Ke Chen and Yana Safonova

Worksheet 9

Monday, Nov 03, 2025

1. **Huffman Properties.** Prove the following: if some character occurs with frequency more than 2/5, then there is guaranteed to be a codeword of length 1.

2. **Huffman Encoding.** Let $n \geq 2$ and label symbols $s_1, \ldots, s_n$ with frequencies

$$f_i = 2^{n-i} \qquad (i = 1, \ldots, n),$$

so the frequencies (in descending order) are $2^{n-1}, 2^{n-2}, \ldots, 2, 1$. Construct the Huffman code for these frequencies and determine the codeword lengths $L(s_i)$ for all i.

3. **Worst Case for Greedy Set Cover.** Let n be a power of 2. Show that there exists an instance of the set cover problem such that: 1) there are n elements in the base set; 2) the optimal solution uses only two sets; and 3) the greedy algorithm picks at least $\log n$ sets.

4. **Weighted Set Cover and the Greedy Algorithm**

Let us consider the weighted version of the Set Cover problem.

Suppose the universe is

$$U = \{1, 2, 3, 4, 5, 6, 7, 8\}.$$

The following sets and weights are given:

$$\begin{aligned} T_1 &= \{1, 2, 3, 4\}, \quad w(T_1) = 9, \\ T_2 &= \{5, 6\}, \quad w(T_2) = 7, \\ T_3 &= \{7\}, \quad w(T_3) = 5, \\ A &= \{1, 3, 5, 7\}, \quad w(A) = 10, \\ B &= \{2, 4, 6, 8\}, \quad w(B) = 10. \end{aligned}$$

(a) What can be a possible greedy strategy for solving this weighted set cover problem?

(b) Apply the proposed greedy strategy step by step to find the sets selected by the greedy algorithm and the total cost of the cover it produces.

(c) Determine the optimal solution and its cost. Compare it with the greedy solution.