Problem 1 (10 points). Construct the Huffman code (i.e., the optimum prefix code) for the alphabet \{a, b, c, d, e, f, g\} with the following frequencies:

<table>
<thead>
<tr>
<th>Symbols</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequencies</td>
<td>50</td>
<td>10</td>
<td>38</td>
<td>25</td>
<td>55</td>
<td>90</td>
<td>5</td>
</tr>
</tbody>
</table>

What is the weighted length of the code (i.e., the sum over all symbols the frequency of the symbol times its encoding length)?

Problem 2 (20 points). Let $I = (k, n, T, (p_1, p_2, \cdots, p_k), (r_1, r_2, \cdots, r_T))$ be an offline caching instance with initial set of pages, where
- $k$ is the number of pages the cache can hold,
- $n$ is the number of different pages,
- $T$ is the length of the request sequence,
- $p_1, p_2, \cdots, p_k \in [n] := \{1, 2, 3, \cdots, n\}$ are the $k$ pages in the cache initially (for simplicity we assume the $k$ pages are all different and there are no empty pages),
- $r_t \in [n]$ for every $t \in [T]$ is the page requested at time $t$.

Let $I' = (k, n, T, (p_1, p_2, \cdots, p_k - 1, p_k'), (r_1, r_2, \cdots, r_T))$ be the instance obtained from $I$ by changing $p_k$ to $p_k'$. Prove the minimum number of misses we can achieve for the instance $I'$ is at most that for the instance $I$ plus 1.

Problem 3 (20 points). Given a set of $n$ points $X = \{x_1, x_2, \cdots, x_n\}$ on the real line, we want to use the smallest number of unit-length closed intervals to cover all the points in $X$. For example, the points $X$ in Figure 1 can be covered by 3 unit-length intervals.

Suppose our greedy strategy is to choose some unit-length interval, and include it in the optimal solution. Which unit-length interval do you want to choose? Give your strategy and prove that it is safe to include it in the solution.
Figure 1: Using 3 unit-length intervals (denoted by thick lines) to cover points in $X$ (denoted by the solid circles).