Exercise 4: Hash Tree, FP-Growth, Association Rules

Exercise 4-1 Hash-Tree

(a) **Construction.** Using the hash function

\[ h(x) = x \mod 3 \]  \hspace{1cm} (1)

construct a hash tree with maximum number of itemsets in inner nodes equal to 4 given the following set of candidates:

\begin{center}
\begin{tabular}{cccccccc}
(1, 9, 11) & (2, 5, 10) & (3, 6, 8) & (4, 7, 9) & (6, 12, 13) & (9, 12, 14) \\
(1, 10, 12) & (2, 5, 12) & (3, 7, 10) & (4, 7, 13) & (6, 12, 14) & (10, 11, 15) \\
(2, 4, 7) & (2, 9, 10) & (3, 12, 14) & (5, 7, 9) & (8, 11, 11) & (12, 12, 15) \\
(2, 5, 8) & (3, 3, 5) & (4, 5, 8) & (5, 7, 13) & (8, 11, 15) & (14, 14, 15)
\end{tabular}
\end{center}

(b) **Counting.** Given the transaction \( t = (t_1, \ldots, t_5) = (1, 3, 7, 9, 12) \), find all candidates of length \( k = 3 \) in the previously constructed tree from exercise (a). In absolute and relative numbers: How many candidates need to be refined? How many nodes are visited?

Exercise 4-2 FP-Tree and FP-Growth Algorithm

Given a set of items \( \{a, b, c, d, e, f, g, h\} \) and a set of transactions \( T \) according to the following table, construct the FP-tree and use the FP-Growth algorithm to compute all frequent itemsets for \( minSup = 0.1 \) (i.e. 2 transactions are needed for an itemset to be frequent).

\begin{tabular}{ll}
TID & Items \\
1 & ag \\
2 & cg \\
3 & eg \\
4 & dg \\
5 & bdfg \\
6 & dg \\
7 & ag \\
8 & ag \\
9 & ae \\
10 & ag \\
11 & afh \\
12 & af \\
13 & ade \\
14 & bdfg
\end{tabular}
Exercise 4-3 Association Rules

Given the following frequent itemsets extract all strong association rules with a minimum confidence of $\text{minConf} = 80\%$. Which candidates can be pruned based on anti-monotonicity?

<table>
<thead>
<tr>
<th>Itemset</th>
<th>Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.00</td>
</tr>
<tr>
<td>B</td>
<td>1.00</td>
</tr>
<tr>
<td>D</td>
<td>0.75</td>
</tr>
<tr>
<td>AB</td>
<td>1.00</td>
</tr>
<tr>
<td>AD</td>
<td>0.75</td>
</tr>
<tr>
<td>BD</td>
<td>0.75</td>
</tr>
<tr>
<td>ABD</td>
<td>0.75</td>
</tr>
</tbody>
</table>