Exercise 1 (frequent concept intents and closure systems)

Definition (frequent concept intent). Let $\mathbb{K} = (G, M, I)$ be a formal context.

(a) The support of a set $B \subseteq M$ of attributes in $\mathbb{K}$ is given by

$$\text{supp}(B) := \frac{|B'|}{|G|}.$$ 

(b) For a given minimal support $\text{minsupp}$ the set of frequent concept intents is given by

$$\{ B \subseteq M \mid \exists A \subseteq G : (A, B) \in \mathbb{B}(G, M, I) \land \text{supp}(B) \geq \text{minsupp} \}.$$ 

Show that the set of frequent concept intents together with the set $M$ forms a closure system.

Solution:

Proof by contradiction:

We have to show that the intersection of any two subsets of the set of frequent concept intents is a frequent concept intent, too. We already know that the set of all concept intents forms a closure system and hence the intersection of two concept intents is again a concept intent. It remains to show that the intersection of two frequent concept intents is again a frequent concept intent.

Without loss of generality, let us consider any two fixed but arbitrary frequent concept intents, e.g., $S_1 = \{a_1, a_2, \ldots, a_k, a_{k+1}\}$ and $S_2 = \{a_1, a_2, \ldots, a_k, a_{k+2}, \ldots, a_n\}$. The intersection of these two sets is given by $S_1 \cap S_2 = \{a_1, a_2, \ldots, a_k\}$. Let’s say the intersection is not a frequent concept intent, which implies that there is at least one element in both $S_1$ and $S_2$ that is not frequent. This is a contradiction because a frequent concept intent, in our case $S_1$ or $S_2$ can’t be frequent if any of its elements is not frequent. Thus, $S_1 \cap S_2$ is frequent, too, and the above set forms a closure system.

Exercise 2 (support)

Show the validity of the properties of the support function that are employed by the TITANIC algorithm:

Let $(G, M, I)$ be a formal context $X, Y \subseteq M$. Then it holds:

1) $X \subseteq Y \implies \text{supp}(X) \geq \text{supp}(Y)$
2) $X'' = Y'' \implies \text{supp}(X) = \text{supp}(Y)$
3) $X \subseteq Y \land \text{supp}(X) = \text{supp}(Y) \implies X'' = Y''$

Solution:
1. Let $X \subseteq Y$, then $Y' \subseteq X'$ holds as we saw in Exercise Sheet 1. This implies, 
$\text{supp}(Y) = \frac{|Y'|}{|G|} \leq \frac{|X'|}{|G|} = \text{supp}(X)$

2. $X'' = Y'' \implies \text{supp}(X) = \text{supp}(Y)$
   
   $X'' = Y'' \iff X''' = Y''' \iff X' = Y' \implies \text{supp}(X) = \frac{|X'|}{|G|} = \frac{|Y'|}{|G|} = \text{supp}(|Y|)$.

3. $X \subseteq Y \land \text{supp}(X) = \text{supp}(Y) \implies X'' = Y''$
   
   $\text{supp}(X) = \text{supp}(Y) \implies |X'| = |Y'|$ and $X \subseteq Y \implies X' \supseteq Y'$. Hence $X' = Y'$, since $X'$ and $Y'$ are finite. It follows, $X'' = Y'''$.

**Exercise 3** (computing concept intents with TITANIC)

The following context contains transactions in a supermarket. Compute the closure system of all concept intents using the TITANIC algorithm. (hint: use the table structure from the example computation in the lecture slides)

<table>
<thead>
<tr>
<th></th>
<th>apples (a)</th>
<th>beer (b)</th>
<th>chips (c)</th>
<th>tv magazine (d)</th>
<th>toothpaste (e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td></td>
</tr>
<tr>
<td>$t_2$</td>
<td></td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td></td>
</tr>
<tr>
<td>$t_3$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td></td>
</tr>
<tr>
<td>$t_4$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_5$</td>
<td></td>
<td>$\times$</td>
<td>$\times$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_6$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_7$</td>
<td></td>
<td></td>
<td>$\times$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t_8$</td>
<td></td>
<td>$\times$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Solution:**

In the first pass, the algorithm deals with the empty set and singletons, all 1-sets. It returns the results for $k = 0$ and $k = 1$:

<table>
<thead>
<tr>
<th>step 1</th>
<th>step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$x.s$</td>
</tr>
<tr>
<td>$x \in k_k$?</td>
<td></td>
</tr>
<tr>
<td>$\emptyset$</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>steps 4+5</th>
<th>step 7</th>
<th>step 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>$X.p_s$</td>
<td>$X.s$</td>
</tr>
<tr>
<td>${a}$</td>
<td>1</td>
<td>$\frac{3}{8}$</td>
</tr>
<tr>
<td>${b}$</td>
<td>1</td>
<td>$\frac{5}{8}$</td>
</tr>
<tr>
<td>${c}$</td>
<td>1</td>
<td>$\frac{6}{8}$</td>
</tr>
<tr>
<td>${d}$</td>
<td>1</td>
<td>$\frac{4}{8}$</td>
</tr>
<tr>
<td>${e}$</td>
<td>1</td>
<td>$\frac{2}{8}$</td>
</tr>
</tbody>
</table>

Step 8 returns: $\emptyset$. closure $\leftarrow \emptyset$
Exercise 4 (optimizing TITANIC for iceberg concept lattices)

In the lecture the following steps to optimize TITANIC for the computation of iceberg concept lattices have been mentioned:

1. Stop, as soon as only non-frequent minimal generators are computed.
2. Return only the closures of frequent minimal generators.
3. Generate candidates only from the frequent minimal generators.
4. All subsets of candidates with $k - 1$ elements must be frequent.

Implement the corresponding modifications in TITANIC. Utilize the fact that for a formal context $\mathbb{K} = (G, M, I)$ and a minimal support constraint $\minsupp$ the set of frequent concept intents together with $M$ form a closure system. The corresponding closure operator $h$ and the support function $\text{support}$ are given by

$$h(X) := \begin{cases} X'' & \text{if } \text{supp}(X) \geq \minsupp \\ M & \text{otherwise} \end{cases}$$

$$\text{support}(X) := \begin{cases} \text{supp}(X) & \text{if } \text{supp}(X) \geq \minsupp \\ -1 & \text{otherwise} \end{cases}$$
Insert the corresponding changes directly into the algorithms attached to this exercise sheet.

**Solution:**
Changes are highlighted in blue in the modified Algorithms 1, 2, and 3.

---

**Algorithm 1 TITANIC**

1) $0.s ← 1$;
2) $\mathcal{K}_0 ← \{\emptyset\}$;
3) $k ← 1$;
4) forall $m ∈ M$ do \{$m\}.p_\_s ← 0.s;
5) $C ← \{\{m\} | m ∈ M\}$;
6) loop begin
7) Support($C$);
8) forall $X ∈ \mathcal{K}_{k−1}$ do $X.\text{closure} ← \text{Closure}(X)$;
9) $\mathcal{K}_k ← \{X ∈ C | X.s ≠ X.p_\_s\}$;
10) if \{$X ∈ \mathcal{K}_k | X.s = −1\} = \emptyset$ then exit loop;
11) $k ← k+1$;
12) $C ← \text{TITANIC-GEN}(\mathcal{K}_{k−1})$;
13) end loop;
14) return $\bigcup_{i=0}^{k−1} \{X.\text{closure} | X ∈ \mathcal{K}_i, X.s ≠ −1\}$.

---

**Algorithm 2 TITANIC-GEN**

Input: $\mathcal{K}_{k−1}$, the set of key $(k−1)$-sets $K$ with their weight $K.s$.

Output: $C$, the set of candidate $k$-sets $C$ with the values $C.p_\_s := \min\{s(C \setminus \{m\} | m ∈ C\}$.

The variables $p_\_s$ assigned to the sets \{$m_1, \ldots, m_k\}$ which are generated in step 1 are initialized by \{$m_1, \ldots, m_k\}.p_\_s ← s_{\text{max}}$.

1) $C ← \{\{m_1 < m_2 < \cdots < m_k\} | \{m_1, \ldots, m_{k−2}, m_{k−1}\}, \{m_1, \ldots, m_{k−2}, m_k\} ∈ \{K ∈ \mathcal{K}_{k−1} | K.s ≠ −1\}$;
2) forall $X ∈ C$ do begin;
3) forall $(k−1)$-subsets $S$ of $X$ do begin
4) if $S ∉ \mathcal{K}_{k−1}$ or $S.s = −1$ then begin $C ← C \setminus \{X\}$; exit forall ; end;
5) $X.p_\_s ← \min(X.p_\_s, S.s)$;
6) end;
7) end;
8) return $C$.  

---


Algorithm 3 \textsc{Closure}(X) for \(X \in K_{k-1}\)

1) \textbf{if} \(X.s = -1\) \textbf{then return} \(M\);
2) \(Y \leftarrow X\);
3) \textbf{forall} \(m \in X\) \textbf{do} \(Y \leftarrow Y \cup (X \setminus \{m\}).\text{closure}\);
4) \textbf{forall} \(m \in M \setminus Y\) \textbf{do begin}
5) \textbf{if} \(X \cup \{m\} \in \mathcal{C}\) \textbf{then} \(s \leftarrow (X \cup \{m\}).s\)
6) \textbf{else} \(s \leftarrow \min\{K.s \mid K \in \mathcal{K}, K \subseteq X \cup \{m\}\}\);
7) \textbf{if} \(s = X.s\) \textbf{then} \(Y \leftarrow Y \cup \{m\}\)
8) \textbf{end};
9) \textbf{return} \(Y\).

Exercise 5 (computing iceberg concept lattices)

We are regarding the following excerpt from the mushroom database:

<table>
<thead>
<tr>
<th></th>
<th>edible (e)</th>
<th>poisonous (p)</th>
<th>cap shape: convex (c)</th>
<th>cap shape: flat (f)</th>
<th>cap surface: fibrous (i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mushroom 1</td>
<td>Mushroom 2</td>
<td>Mushroom 3</td>
<td>Mushroom 4</td>
<td>Mushroom 5</td>
</tr>
<tr>
<td></td>
<td>Mushroom 6</td>
<td>Mushroom 7</td>
<td>Mushroom 8</td>
<td>Mushroom 9</td>
<td>Mushroom 10</td>
</tr>
</tbody>
</table>

a) Compute the corresponding iceberg concept lattice for \(\text{minsupp} = 30\%\) using the modified algorithm from the previous exercise.

b) Compute the corresponding iceberg concept lattices and label each (frequent) concept with its corresponding support value.

Solution:

In the first pass, the algorithm deals with the empty set and singletons, all 1-sets. It returns the results for \(k = 0\) and \(k = 1\):

<table>
<thead>
<tr>
<th></th>
<th>step 1</th>
<th>step 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k = 0):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\emptyset)</td>
<td>1</td>
<td>yes</td>
</tr>
</tbody>
</table>
### k = 1:

<table>
<thead>
<tr>
<th>X</th>
<th>X.p.s</th>
<th>X.s</th>
<th>X \in K_k?</th>
</tr>
</thead>
<tbody>
<tr>
<td>{e}</td>
<td>1</td>
<td>6/10</td>
<td>yes</td>
</tr>
<tr>
<td>{p}</td>
<td>1</td>
<td>14/10</td>
<td>yes</td>
</tr>
<tr>
<td>{c}</td>
<td>1</td>
<td>14/10</td>
<td>yes</td>
</tr>
<tr>
<td>{l}</td>
<td>1</td>
<td>6/10</td>
<td>yes</td>
</tr>
<tr>
<td>{i}</td>
<td>1</td>
<td>17/10</td>
<td>yes</td>
</tr>
</tbody>
</table>

Step 8 returns: ∅.closure ← ∅

### k = 2:

<table>
<thead>
<tr>
<th>X</th>
<th>X.p.s</th>
<th>X.s</th>
<th>X \in K_k?</th>
</tr>
</thead>
<tbody>
<tr>
<td>{e, p}</td>
<td>4/10</td>
<td>-1</td>
<td>no</td>
</tr>
<tr>
<td>{e, c}</td>
<td>4/10</td>
<td>4/10</td>
<td>yes</td>
</tr>
<tr>
<td>{e, l}</td>
<td>6/10</td>
<td>-1</td>
<td>yes</td>
</tr>
<tr>
<td>{e, i}</td>
<td>6/10</td>
<td>4/10</td>
<td>yes</td>
</tr>
<tr>
<td>{p, c}</td>
<td>4/10</td>
<td>-1</td>
<td>yes</td>
</tr>
<tr>
<td>{p, l}</td>
<td>4/10</td>
<td>4/10</td>
<td>yes</td>
</tr>
<tr>
<td>{p, i}</td>
<td>4/10</td>
<td>3/10</td>
<td>yes</td>
</tr>
<tr>
<td>{c, l}</td>
<td>4/10</td>
<td>-1</td>
<td>no</td>
</tr>
<tr>
<td>{c, i}</td>
<td>4/10</td>
<td>-1</td>
<td>yes</td>
</tr>
<tr>
<td>{l, i}</td>
<td>6/10</td>
<td>5/10</td>
<td>yes</td>
</tr>
</tbody>
</table>

### k = 3:

Step 12: returns the empty set. Hence there is nothing to **WEIGH** in Step 7.

Step 8 returns:
- {e}.closure ← {e}
- {p}.closure ← {p, l}
- {c}.closure ← {e, c}
- {l}.closure ← {l}
- {i}.closure ← {i}

Step 14: Collects all concept intents:

∅, {e}, {p, l}, {c, e}, {l}, {i}, {e, i}, {p, l, i}, {l, i}, M

---

**Diagram:**

- **Cape Shape:** Fibrous (100%)
  - Edible (60%)
  - Cape Shape: Convex (30%)
  - 40%

- **Cape Shape:** Flat (60%)
  - Poisonous (50%)
  - Cape Shape: Convex (30%)
  - 30%
Algorithm 4 TITANIC

1) $\text{Support}\left\{\emptyset\right\}$;
2) $\mathcal{K}_0 \leftarrow \{\emptyset\}$;
3) $k \leftarrow 1$;
4) $\forall m \in M \text{ do } \{m\}.p_s \leftarrow \emptyset.s$;
5) $C \leftarrow \{\{m\} \mid m \in M\}$;
6) loop begin
7) $\text{Support}(C)$;
8) $\forall X \in \mathcal{K}_{k-1} \text{ do } X.\text{closure} \leftarrow \text{CLOSURE}(X)$;
9) $\mathcal{K}_k \leftarrow \{X \in C \mid X.s \neq X.p_s\}$;
10) if $\mathcal{K}_k = \emptyset$ then exit loop ;
11) $k + +$;
12) $\mathcal{C} \leftarrow \text{TITANIC-GEN}(\mathcal{K}_{k-1})$;
13) end loop ;
14) return $\bigcup_{i=0}^{k-1}\{X.\text{closure} \mid X \in \mathcal{K}_i\}$.

Algorithm 5 TITANIC-GEN

Input: $\mathcal{K}_{k-1}$, the set of key $(k - 1)$-sets $K$ with their weight $K.s$.
Output: $\mathcal{C}$, the set of candidate $k$-sets $C$
with the values $C.p_s := \min\{s(C \setminus \{m\} \mid m \in C\}$.

The variables $p_s$ assigned to the sets $\{m_1, \ldots, m_k\}$ which are generated in step 1 are initialized by $\{m_1, \ldots, m_k\}.p_s \leftarrow s_{\text{max}}$.

1) $C \leftarrow \{\{m_1 < m_2 < \cdots < m_k\} \mid \{m_1, \ldots, m_{k-2}, m_{k-1}\}, \{m_1, \ldots, m_{k-2}, m_k\}\}$
2) forall $X \in C$ do begin
3) $\forall (k - 1)$-subsets $S$ of $X$ do begin
4) if $S \notin \mathcal{K}_{k-1}$ then begin $\mathcal{C} \leftarrow \mathcal{C} \setminus \{X\}$; exit forall ; end;
5) $X.p_s \leftarrow \min(X.p_s, S.s)$;
6) end;
7) end;
8) return $\mathcal{C}$.

Algorithm 6 CLOSURE$(X)$ for $X \in \mathcal{K}_{k-1}$

1) $Y \leftarrow X$;
2) forall $m \in X$ do $Y \leftarrow Y \cup (X \setminus \{m\}).\text{closure}$; forall $m \in M \setminus Y$ do begin
3) if $X \cup \{m\} \in \mathcal{C}$ then $s \leftarrow (X \cup \{m\}).s$
4) else $s \leftarrow \min\{K.s \mid K \in \mathcal{K}, K \subseteq X \cup \{m\}\}$;
5) if $s = X.s$ then $Y \leftarrow Y \cup \{m\}$
6) end;
7) return $Y$. 