Galois Sub-hierarchy and Orderings

SYLVIA ENCHEVA
Stord/Haugesund University College
Faculty of Technology, Business and Maritime Sciences
Bjørnsonsg. 45, 5528 Haugesund
NORWAY
sbe@hsh.no

Abstract: In this paper we investigate stabilities of concepts resulting from extending partial ordering of three elements to partial ordering of four elements. Iceberg lattices are further employed for the representation of reduced association rules.

Key Words: Decision support services, uncertainty management, partial orderings

1 Introduction

The regulation on waste delivered to shorere enforced by 2004 requires vessels entering ports within the European Economic Community (EEC) to report current status of waste onboard. This includes the amount of waste being produced, delivered in port, and planned to be delivered in next port of call. Prior to arrival all the required data has to be delivered to port authorities. Such information is of special interest regarding environmental reporting.

Cruise ships often generate waste that prevails their maximum storage capacity long before they have access to shore waste disposal facilities. Daily they also large amounts of black and grey water. According to MARPOL 73/78 discharge of treated waste water is allowed 4 Mi off shore, while some regional port policies require discharge treated waste water 12 Mi if offshore.

In this paper we investigate stabilities of concepts resulting from extending partial ordering of three elements to partial ordering of four elements. The obtained results are applied for facilitating the process of decision making related to ship generated waste management.

The rest of the paper is organised as follows. Related work and supporting theory may be found in Section 2. The main results are presented in Section 3 Conclusions and future work can be found in Section 4.

2 Posets

Let $P$ be a non-empty ordered set. If $\sup\{x, y\}$ and $\inf\{x, y\}$ exist for all $x, y \in P$, then $P$ is called a lattice [2]. In a lattice illustrating partial ordering of knowledge values, the logical conjunction is identified with the meet operation and the logical disjunction with the join operation.

A context is a triple $(G, M, I)$ where $G$ and $M$ are sets and $I \subseteq G \times M$. The elements of $G$ and $M$ are called objects and attributes respectively [2], [11]. For $A \subseteq G$ and $B \subseteq M$, define

$A' = \{m \in M \mid (\forall g \in A) \ gIm\}$,

$B' = \{g \in G \mid (\forall m \in B) \ gIm\}$

where $A'$ is the set of attributes common to all the objects in $A$ and $B'$ is the set of objects possessing the attributes in $B$.

A concept of the context $(G, M, I)$ is defined to be a pair $(A, B)$ where $A \subseteq G$, $B \subseteq M$, $A' = B$ and $B' = A$. The extent of the concept $(A, B)$ is $A$ while its intent is $B$. A subset $A$ of $G$ is the extent of some concept if and only if $A'' = A$ in which case the unique concept of the which $A$ is an extent is $(A, A')$. The corresponding statement applies to those subsets $B \in M$ which is the intent of some concepts.

The set of all concepts of the context $(G, M, I)$ is denoted by $\mathfrak{B}(G, M, I)$. $(\mathfrak{B}(G, M, I); \subseteq)$ is a complete lattice and it is known as the concept lattice of the context $(G, M, I)$.

2.1 Ordered Sets

Determining a consensus from a group of orderings and making statistically significant statements about orderings have been discussed in [1].

A relation $I$ is an indifference relation when given $AIB$ neither $A > B$ nor $A < B$ has place in the componentwise ordering. A partial ordering whose indifference relation is transitive is called a weak ordering.
Let \( w_1, w_2, w_3 \) be weak orderings. Then \( w_2 \) is between \( w_1 \) and \( w_3 \) if each decision made by \( w_2 \) is made by either \( w_1 \) or \( w_3 \) and any decision made by both \( w_1 \) and \( w_3 \) is made by \( w_2 \), i.e. \( w_1 \cap w_3 \subseteq w_2 \subseteq w_1 \cup w_3 \).

The distance \( d(w_1, w_3) \) is defined as \( d(w_1, w_2) + d(w_2, w_3) = d(w_1, w_3) \). The distance is a metric in the usual sense, it is invariant under permutation of alternatives, and the minimum positive distance is 1.

In [3], an introducer of an object is called an Object Concept and an introducer of a property is called a Property Concept.

**Definition 1** [4] The Galois sub-hierarchy (GSH) of a concept lattice is the partially ordered set of elements \( X \times Y, X \cup Y \neq \emptyset \), such that there exists a concept where \( X \) is the set of objects introduced by this concept and \( Y \) is the set of properties introduced by this concept. The ordering of the elements in the GSH is the same as in the lattice.

Let \( E_1 \) and \( E_2 \) be two elements of the GSH. We will denote \( E_1 < E_2 \) if \( E_1 \) is an ancestor (i.e. represented below in the GSH) of \( E_2 \) and \( E_1 \leq E_2 \) if \( E_1 = E_2 \) or \( E_1 < E_2 \).

Iceberg lattices are represented below as in [8].

**Definition 2** Let \( B \subseteq M \). The support count of the attribute set \( B \) in \( \mathbb{K} \) is

\[
\sigma(B) = \frac{|B^*|}{|G|}
\]

Let \( \text{minsupp} \) be a threshold \( \in [0, 1] \), then \( B \) is said to be a frequent itemset if \( \sigma(B) \geq \text{minsupp} \). A concept is called frequent concept if its intent is frequent.

**Definition 3** The set of all frequent concepts of a context \( \mathbb{K} \) is called iceberg lattice of the context \( \mathbb{K} \).

Iceberg Lattices can be used to discover and visualize association rules. Within a formal context \( K = (G, M, I) \), the task of mining association rules is to determine all pairs \( X \rightarrow Y \) of \( M \) such that \( \sigma(X \rightarrow Y) = \sigma(X \cup Y) \geq \text{minsupp} \), and the confidence \( \text{conf}(X \rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)} \) is above a given threshold \( \text{minconf} \in [0, 1] \).

## 3 Posets of Elements

In this section we consider changes within various concepts when an object is added or removed and when an attribute is added or removed.

1) Suppose the system offers three elements. The following orderings are to be considered: the three elements are ordered consecutively, one of the elements is placed higher than the other two elements and no preference is shown with respect to these two elements, one of the elements is placed lower than the other two elements and no preference is shown with respect to these two elements, only two of the elements are considered.

Representatives of these posets are summarized in Fig. 1.

2) The following orderings are to be considered in the case of four elements: the four elements are not ranked, only two elements are ranked, two couples of elements are ranked, three couples of elements are ranked, four couples of elements are ranked, one element is ranked higher than the other three, three elements are ranked higher than the forth element, three elements are ranked linearly and the forth element is not ranked, three elements are ranked linearly and the forth element is ranked higher than the last element, three elements are ranked linearly and the forth element is ranked below than the top element, one element is compared to the other three elements and another element ranked higher than the rest, one element is compared to the other three elements and two elements are ranked higher than the rest, every element is compared to two of the other elements, and all elements are ranked linearly.

Representatives of these posets are summarized in Fig. 2.
Table 1: Relationships between an ordered set of 3-elements and an ordered set of 4-elements

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td></td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G31</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Orderings of four elements

Figure 3: Complete lattice
Figure 4: A lattice without object G2r

Galois sub-hierarchy (GSH) of a concept lattice in Fig. 3 is shown in Fig. 4.

Let \( c = (X, Y) \) a concept in \( L_1 \). The frequency of \( c \), denoted \( \text{freq}(c) \), is defined as the ratio of its extent and the size of the object set: \( \text{freq}(c) = \frac{|X|}{|O|} \). Given \( \alpha \) a minimal threshold of support defined by the user, the concept \( c \) is frequent if \( \text{freq}(c) \geq \alpha \). The iceberg concept lattice generated by \( \alpha L_1 \), is made of all frequent concepts. An iceberg is thus a join-semi-lattice, a sub-semi-lattice of the complete concept lattice [10].

For example, the iceberg \( L_{0.4} \) obtained from the complete lattice of Fig. 3 with \( \alpha \geq 0.4 \) is shown in Fig. 5.

Example 4 In terms of ship generated waste the elements \( a, b, c, d \) can be taken as greywater, blackwater, oily bilge water and solid waste.

4 Conclusion

There is a need for providing information about waste management to ship owners, ship operator, authorities and ports. Different users however have needs on level of detail and regularity of information. Application of iceberg lattices in the process of environmental reporting can considerably improve the process of selecting the most important information without inclusion of unnecessary details.

References:


