Formal Concept Analysis

III Knowledge Discovery

Sebastian Rudolph

Computational Logic Group
Technische Universität Dresden
Agenda

7 Triadic Formal Concept Analysis
  • Motivation
    • Folksonomies
    • Motivation
    • Triadic Formal Concept
    • Concept-Tri-Lattice
    • Visualization of Tri-Lattices
    • Iceberg Tri-Lattices
    • Computing Tri-Concepts
    • Qualitative Evaluation
    • Neighborhoods
Motivation: Collaborative Tagging Systems

Entdecken / Tags / nature

Sortieren nach:
Neueste • Interessanteste

Cluster mit dem Tag nature
Entdecken und filtern Sie diese Liste mit dem Tag nature mit unserem tollen Cluster-Feature!

Dazu passende Tags:
macro, flower, green, landscapes, trees, sky, water, insect, flowers, leaves

Ähnliche Inhalte mit:
der Yahoo! Bildersuche suchen

Sponsoren-Links
PureNature Versand
Hier finden Sie alles für ein gesundes, allergiefreies Leben!
www.PureNature.de

Natururlaub in Frankreich
Natur pur und nachhaltige Konzepte
de finanzagile camphorblätterweise

Von Tony Reilly 1959
Von fließ
Von Diane &
Von Claytona
Von Andy
Von Tony Reilly 1959
Von Gerdé
Von Carme M.V.
Motivation: Collaborative Tagging Systems
Motivation: Collaborative Tagging Systems
Motivation: Collaborative Tagging Systems

- manage your web bookmarks and publication references
- open for the public since beginning of 2006, > 5000 active users
- developed and operated at L3S Research Center
Agenda

7. Triadic Formal Concept Analysis
   - Motivation
   - Folksonomies
   - Motivation
   - Triadic Formal Concept
   - Concept-Tri-Lattice
   - Visualization of Tri-Lattices
   - Iceberg Tri-Lattices
   - Computing Tri-Concepts
   - Qualitative Evaluation
   - Neighborhoods
Folksonomies

- data structure of collaborative tagging systems
- connects users, tags, and resources
- conceptual structure created by the people
Folksonomies: Hypergraph, Tensor
Folksonomies: Hypergraph, Tensor

Sebastian Rudolph (TUD)
Folksonomies: Hypergraph, Tensor
Folksonomies

Definition (Folksonomy)

\[ F := (U, T, R, Y) \] with

- \( U, T, R \) finite sets of users, tags, and resources, resp.
- \( Y \subseteq U \times T \times R \) ternary relation

- tripartite hypergraph
- boolean 3-dimensional tensor
- triadic formal context
Motivation

- conceptual clustering of folksonomies
  - find interesting concepts/clusters
  - support browsing, community detection, recommendations
  - get an overview into the structure of a folksonomy
Motivation

- conceptual clustering of folksonomies
  - find interesting concepts/clusters
  - support browsing, community detection, recommendations
  - get an overview into the structure of a folksonomy
- **tri-concept** \((A, B, C) \subseteq U \times T \times R\): maximal cuboid in which every user from \(A\) has tagged every resource from \(C\) with all tags from \(B\)

→ shared conceptualization
Motivation

Triadic Concept Analysis (Lehmann, Wille 1995)

Frequent Tri-Concepts

Iceberg Concept Lattices / Closed Itemset Mining (Lakhal/Stumme/Zaki 1999)

Formal Concept Analysis (Wille 1982)

Association Rules (Agrawal, Srikant 1993)
Motivation

We regard \( F = (U, T, R, Y) \) as triadic formal context.
In general, the elements of \( U, T \) and \( R \) are then called objects, attributes and conditions and \( (u, t, r) \in Y \) is read as “object \( u \) has the attribute \( t \) under condition \( r \)”.
Triadic Formal Concept

Definition (tri-concept)

triple \((A, B, C)\) with \(A \subseteq U, B \subseteq T, C \subseteq R\) and \(A \times B \times C \subseteq Y\), such that none of the three components can be enlarged without violating the condition \(A \times B \times C \subseteq Y\). We call \(A\) the *extent*, \(B\) the *intent* and \(C\) the *modus* of the formal tri-concept.

→ natural extension of formal concepts
Concept-Tri-Lattice

- three quasi orders \( \preceq_1, \preceq_2, \preceq_3 \):
  \[(A_1, A_2, A_3) \preceq_i (B_1, B_2, B_3) \]
  \[\iff A_i \subseteq B_i, \text{ for } i = 1, 2, 3.\]

- not antisymmetric, i.e. from
  \[(A_1, A_2, A_3) \preceq_i (B_1, B_2, B_3) \]
  and
  \[(B_1, B_2, B_3) \preceq_i (A_1, A_2, A_3) \]
  does not follow \((A_1, A_2, A_3) = (B_1, B_2, B_3)\)

- concept tri-lattice \( \mathfrak{B}(K) \) of the triadic context \( K \)

- not a real (mathematical) lattice!
Visualization of Tri-Lattices

- Since it is not really a lattice, we cannot draw a lattice diagram
- Alternative:
  - every quasi-order is written along the edge of a virtual triangle
  - the tri-concepts are drawn into the triangle
- Example to the right: smallest non-trivial tri-lattice
  \[ \mathbb{B}_3 = \mathbb{B}(\{1\}, \{1\}, \{1\}, \emptyset) \]
- Visualization not always possible
  - satisfied tetrahedron condition
  - Violated Thomson condition
Visualization of Tri-Lattices
Iceberg Tri-Lattices

- Given support constraints \( \tau_u, \tau_t, \tau_r \): tri-concept \((A, B, C)\) frequent
  \[\iff |A| \geq \tau_u, |B| \geq \tau_t, \text{ and } |C| \geq \tau_r.\]

  \(\rightarrow\) iceberg tri-lattice
Computing Tri-Concepts

- **Given**
  - sets $U$, $T$, $R$
  - ternary relation $Y \subseteq U \times T \times R$
  - support constraints $\tau_u$, $\tau_t$, $\tau_r$

- **Find** $(A, B, C)$ with
  - $A \subseteq U$, $B \subseteq T$, $C \subseteq R$
  - $|A| \geq \tau_u$, $|B| \geq \tau_t$, $|C| \geq \tau_r$
  - $A \times B \times C \subseteq Y$
  - such that none of the sets $A$, $B$ or $C$ can be enlarged without violating the former condition
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

Algorithm
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

Algorithm

- Let $\tilde{Y} := \{(u, (t, r)) \mid (u, t, r) \in Y\}$
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

Algorithm

- Let $\tilde{Y} := \{(u, (t, r)) \mid (u, t, r) \in Y\}$
- Loop: Find (frequent) concepts $(A, I)$ in $(U, T \times R, \tilde{Y})$
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

Algorithm

- Let $\tilde{Y} := \{(u, (t, r)) \mid (u, t, r) \in Y\}$
- Loop: Find (frequent) concepts $(A, I)$ in $(U, T \times R, \tilde{Y})$

In the example:

$(A, I) = (\{u_2, u_3\}, \{(t_1, r_1), (t_1, r_2), (t_2, r_1)\})$
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

**Algorithm**

- Let $\tilde{Y} := \{(u, (t, r)) \mid (u, t, r) \in Y\}$
- Loop: Find (frequent) concepts $(A, I)$ in $(U, T \times R, \tilde{Y})$
  - Loop: Find (frequent) concepts $(B, C)$ in $(T, R, I)$

In the example:

$$(T, R, I) = (T, R, \{(t_1, r_1), (t_1, r_2), (t_2, r_1)\})$$
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

Algorithm

- Let $\tilde{Y} := \{(u, (t, r)) \mid (u, t, r) \in Y\}$
- Loop: Find (frequent) concepts $(A, I)$ in $(U, T \times R, \tilde{Y})$
  - Loop: Find (frequent) concepts $(B, C)$ in $(T, R, I)$

In the example:

$(B, C) = (\{t_1\}, \{r_1, r_2\})$
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

**Algorithm**

- Let \( \tilde{Y} := \{(u, (t, r)) \mid (u, t, r) \in Y\} \)
- Loop: Find (frequent) concepts \((A, I)\) in \((U, T \times R, \tilde{Y})\)
  - Loop: Find (frequent) concepts \((B, C)\) in \((T, R, I)\)
    - If \( A = (B \times C)\tilde{Y} \), then output \((A, B, C)\)

In the example:

\[(B \times C)\tilde{Y} = (\{t_1\} \times \{r_1, r_2\})\tilde{Y}\]
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

Algorithm

- Let \( \tilde{Y} := \{(u, (t, r)) \mid (u, t, r) \in Y\} \)
- Loop: Find (frequent) concepts \((A, I)\) in \((U, T \times R, \tilde{Y})\)
  - Loop: Find (frequent) concepts \((B, C)\) in \((T, R, I)\)
    - If \( A = (B \times C) \tilde{Y} \), then output \((A, B, C)\)

In the example:

\[
(B \times C) \tilde{Y} = (\{t_1\} \times \{r_1, r_2\}) \tilde{Y} = \{u_2, u_3\} = A
\]
Computing Tri-Concepts

computes the iceberg tri-lattice of a triadic formal context

Algorithm

- Let \( \tilde{Y} := \{ (u, (t, r)) \mid (u, t, r) \in Y \} \)
- Loop: Find (frequent) concepts \((A, I)\) in \((U, T \times R, \tilde{Y})\)
  - Loop: Find (frequent) concepts \((B, C)\) in \((T, R, I)\)
    - If \(A = (B \times C)\tilde{Y}\), then output \((A, B, C)\)

In the example:
\( (A, B, C) = (\{u_2, u_3\}, \{t_1\}, \{r_1, r_2\}) \)
Computing Tri-Concepts

Require: $U, T, R, Y, \tau_u, \tau_t, \tau_r$

1. $\tilde{Y} := \{(u, (t, r)) \mid (u, t, r) \in Y\}$
2. $(A, I) := \text{FirstFrequentConcept}((U, T \times R, \tilde{Y}), \tau_u)$
3. repeat
4. if $|I| \geq \tau_t \cdot \tau_r$ then
5. $(B, C) := \text{FirstFrequentConcept}((T, R, I), \tau_t)$
6. repeat
7. if $|C| \geq \tau_r$ then
8. if $A = (B \times C)^{\tilde{Y}}$ then
9. print $A, B, C$
10. end if
11. end if
12. until not $\text{NextFrequentConcept}((B, C), (T, R, I), \tau_t)$
13. end if
14. until not $\text{NextFrequentConcept}((A, I), (U, T \times R, \tilde{Y}), \tau_u)$
Computing Tri-Concepts

The *FirstFrequentConcept* method:

**Require:** \((G, M, I), \tau\)

1. \(A := \emptyset^I\)
2. \(B := A^I\)
3. **if** \(|A| < \tau\) **then**
4. \(\text{NextFrequentConcept}((A, B), (G, M, I), \tau)\)
5. **end if**
6. **return** \((A, B)\)
Computing Tri-Concepts

the \textit{NextFrequentConcept} method:

\begin{algorithm}
\begin{algorithmic}
\Require $(A, B), (G, M, I), \tau$
\State $i := \max(M)$
\While {defined($i$)}
\State $A := (B \bullet i)^I$
\If {$|A| \geq \tau$}
\State $D := A^I$
\If {$B <_i D$}
\State $B := D$
\State \text{return true}
\EndIf
\EndIf
\State $i := \max(M \setminus B \cap \{1, \ldots, i - 1\})$
\EndWhile
\State \text{return false}
\end{algorithmic}
\end{algorithm}
Qualitative Evaluation

**BibSonomy Dataset:**
- all publication records until November 23rd, 2006
- removed: DBLP, posts with the tag “imported”
- $|U| = 262$, $|T| = 5954$, $|R| = 11101$, $|Y| = 44944$

**Result:**
- 13,992 tri-concepts (75 minutes on a 2 GHz PC)
- with support constraints $\tau_u = 3$, $\tau_t = 2$, $\tau_r = 2$:
  - 21 tri-concepts
  - contain 41 publications, 15 users and 36 tags
Qualitative Evaluation

visualisation of the iceberg tri-lattice for $\tau_u = 3$, $\tau_t = 2$, $\tau_r = 2$
Qualitative Evaluation

two topical groups:
- semantic
- social

semantic further divided:
- wiki
- web
- folksonomy
Qualitative Evaluation

<table>
<thead>
<tr>
<th></th>
<th>t</th>
<th>r</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>A Finite-State Model for On-Line Analytical Processing in Triadic Contexts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Annotation and Navigation in Semantic Wikis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>A Semantic Wiki for Mathematical Knowledge Management</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>BibSonomy: A Social Bookmark and Publication Sharing System</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Bringing the &quot;Wiki-Way&quot; to the Semantic Web with Rhizome</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Building and Using the Semantic Web</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Conceptual Clustering of Text Clusters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Content Aggregation on Knowledge Bases using Graph Clustering</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Creating and using Semantic Web information with Makna</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Emergent Semantics in BibSonomy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Explaining Text Clustering Results using Semantic Structures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Harvesting Wiki Consensus - Using Wikipedia Entries as Ontology Elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Information Retrieval in Folksonomies: Search and Ranking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>KAON – Towards a Large Scale Semantic Web</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Kaukolu: Hub of the Semantic Corporate Intranet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Kollaboratives Wissensmanagement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Learning with Semantic Wikis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Mining Association Rules in Folksonomies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>On Self-Regulated Swarms, Societal Memory, Speed and Dynamics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Ontologies improve text document clustering</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Proceedings of the First Workshop on Semantic Wikis – From Wiki To Semantics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Proc. of the European Web Mining Forum 2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Semantic Network Analysis of Ontologies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Semantic Web Mining</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Semantic Web Mining and the Representation, Analysis, and Evolution of Web Space</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Semantic Web Mining for Building Information Portals (Position Paper)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Social Bookmarking Tools (I): A General Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Social Bookmarking Tools (II). A Case Study – Connotea</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Social Cognitive Maps, Swarm Collective Perception and Distributed Search on Dynamic Landscapes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>SweetWiki : Semantic Web Enabled Technologies in Wiki</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Text Clustering Based on Background Knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>The ABCDE Format Enabling Semantic Conference Proceedings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>The Courseware Watchdog: an Ontology-based tool for Finding and Organizing Learning Material</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Towards a Wiki Interchange Format (WIF) – Opening Semantic Wiki Content and Metadata</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Towards Semantic Web Mining</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>TRIAS - An Algorithm for Mining Iceberg Tri-Lattices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Usage Mining for and on the Semantic Web (Book)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Usage Mining for and on the Semantic Web (Workshop)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Wege zur Entdeckung von Communities in Folksonomies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>WordNet improves text document clustering</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Qualitative Evaluation
Qualitative Evaluation

Formal Concept Analysis
Qualitative Evaluation
Qualitative Evaluation

publication 33: "‘The ABCDE Format Enabling Semantic Conference Proceedings’"
The visualization of tri-lattices is . . .

- at the moment manual work,
- time-intensive and pretty complicated,
- or even impossible (cf. *tetrahedron condition* and *Thomson condition*).

Thus: easier visualization option desireable
Neighborhoods

Idea:

- We regard tri-concepts as nodes in a graph.
- We connect two tri-concepts with an edge, when they contain the same tags, users, or resources.

More formally:

- Two tri-concepts \((A_1, A_2, A_3)\) and \((B_1, B_2, B_3)\) are neighbors, if for an \(i \in \{1, 2, 3\}\) it holds \(A_i = B_i\).
- neighbor relation \(\sim \subseteq (\mathcal{B}(F) \times \mathcal{B}(F))\)
- The *neighborhood graph* then is \((\mathcal{B}(F), \sim)\).
Neighborhoods

neighborhood graph for the tri-concept
\(\{\text{jaeschke, schmitz, stumme}\}, \{\text{fca, triadic}\}, \{1, 37\}\)