COMPUTATION AND VISUALIZATION OF CONCEPT LATTICES

SURYAPALLY VAMSHI KRISHNA

OUTLINE

- Introduction
  - Formal concept analysis
- Terminology
  - Partial order and pre-order
  - Formal context
  - Formal concept and Principal formal concept
  - Intent and extent
  - Concept lattice
- Proposed approach
- Applications
- Conclusion
- Bibliography

FORMAL CONCEPT ANALYSIS

- Formal concept analysis (FCA) arose around 1980 in Darmstadt as a mathematical theory and it has found many uses in informatics, e.g. for
  - Data Analysis
  - Knowledge Discovery
  - Software Engineering
- FCA is a framework for data analysis based on a notion of a “concept” i.e. the arrangement of data into encapsulating units.
- FCA helps to understand unknown, hidden and meaningful connections between group of objects and group of attributes.
- Formal concept analysis has its roots in lattice theory.
- There are two ways to describe a concept (Extension and Intension).
- Concept hierarchies can be created and visualized based on the given data sets.

PARTIAL ORDER AND PRE-ORDER

- On a set X, a relation R is said to be a partial order on X if R is:
  - Reflexive, i.e., $(x,x) \in R$, for all $x \in X$,
  - Antisymmetric, i.e., if $(x,y) \in R$ and $(y,x) \in R$ then $x = y$, for all $x,y \in X$,
  - Transitive, i.e., if $(x,y) \in R$ and $(y,z) \in R$ then $(x,z) \in R$, for all $x,y,z \in X$.
- If R is only reflexive and transitive, then R is called a pre-order. Any pre-order induces a partial-order on the set of equivalence classes induced by R.

FORMAL CONTEXT

- A formal context $K = (G,M,I)$ consists of
  - A set G of Objects
  - A set M of Attributes *and*
  - a binary relation between $I \subset G \times M$
- The relation I is called as the incidence relation of the context.
- To express that an object g is in a relation I with an attribute m, we use gIm or $(g , m) \in I$ and is read as object g has attribute m.

Attributes

Objects

	Student	Immigrant	Rich	Married	Employee
John		X		X	
Smith		X			X
Nick			X		
Mike	X	X	X		
Jack				X	
Paul		X	X		
Ben					X

INTENT AND EXTENT

- Given a set A of objects, i.e., $A \subseteq G$ we define its intent $A' \subseteq M$ as
  - $A' = \{ \mathrm{m} \mid oIm \text{ for all } \mathrm{o} \in A \}$.
- Analogously, the extent B' of a set of attribute is defined as
  - $B' = \{ \mathrm{o} \mid \mathrm{o}I\mathrm{m} \text{ for all } \mathrm{m} \in B \}$.
- Intent is a set of attributes that objects have in common.
- Extent is a set of objects that attributes have in common.

INTENT AND EXTENT – CONT'D

- Example of an extent:
  - {Immigrant}′= {John, Smith, Mike, Paul}

Attributes

Objects

	Student	Immigrant	Rich	Married	Employee
John		X		X	
Smith		X			X
Nick			X		
Mike	X	X	X		
Jack				X	
Paul		X	X		
Ben					X

INTENT AND EXTENT – CONT'D

- Example of an intent:
  - $\{\text{Mike}\}' = \{\text{Student, Immigrant, Rich}\}$
- Intersection of arbitrarily many extents is an extent.
- Intersection of arbitrarily many intents is an intent.

Attributes

Objects

	Student	Immigrant	Rich	Married	Employee
John		X		X	
Smith		X			X
Nick			X		
Mike	X	X	X		
Jack				X	
Paul		X	X		
Ben					X

CALCULATING EXTENTS FOR THE GIVEN FORMAL CONTEXT

- **Extents:**
  - $E_1 = \{\text{Student}\}' = \{Mike\}$
  - $E_2 = \{\text{Immigrant}\}' = \{\text{John, Smith, Mike, Paul}\}$
  - $E_3 = \{\text{Rich}\}' = \{\text{Nick, Mike, Paul}\}$
  - $E_4 = \{\text{Married}\}' = \{John, Jack\}$
  - $E_5 = \{\text{Student}\}' = \{Smith, Ben\}$
  - $E_6 = E_1 \cap E_4 = \{\phi\}$
  - $E_7 = E_2 \cap E_3 = \{Mike, Paul\}$
  - $E_8 = E_2 \cap E_4 = \{John\}$
  - $E_9 = E_2 \cap E_5 = \{Smith\}$
  - $E_{10} = G = \{\text{John, Smith, Nick, Mike, Jack, Paul, Ben}\}$

Attributes

Objects

	Student	Immigrant	Rich	Married	Employee
John		X		X	
Smith		X			X
Nick			X		
Mike	X	X	X		
Jack				X	
Paul		X	X		
Ben					X

CALCULATING INTENTS FOR THE GIVEN FORMAL CONTEXT

- **Intents**:
  - I_1 = {Student, Immigrant, Rich}
  - I_2 ={Immigrant}
  - I_3 ={Rich}
  - I_4 = {Married}
  - I_5 = {Employee}
  - $I_6 = M$ = {Student, Immigrant, Rich, Married, Employee}
  - I_7 = {Immigrant, Rich}
  - I_8 = {Immigrant, Married}
  - I_9 = {Immigrant, Employee}
  - I_{10} = {ϕ}

Attributes

Objects

	Student	Immigrant	Rich	Married	Employee
John		X		X	
Smith		X			X
Nick			X		
Mike	X	X	X		
Jack				X	
Paul		X	X		
Ben					X

CONCEPTS FOR THE GIVEN FORMAL CONTEXT

Concepts	**Extents**	**Intents**
1	$\{Mike\}$	{Student, Immigrant, Rich}
2	{John, Smith, Mike, Paul}	{Immigrant}
3	{Nick, Mike, Paul}	{Rich}
4	$\{John, Jack\}$	{Married}
5	$\{Smith, Ben\}$	{Employee}
6	{ϕ}	{Student, Immigrant, Rich, Married, Employee}
7	$\{Mike, Paul\}$	{Immigrant, Rich}
8	$\{John\}$	{Immigrant, Married}
9	$\{Smith\}$	{Immigrant, Employee}
10	{John, Smith, Nick, Mike, Jack, Paul, Ben}	{ϕ}

FORMAL CONCEPT AND PRINCIPAL FORMAL CONCEPT

- A formal concept is a pair *(A,B)* with $A \subseteq G$ and $B \subseteq M$ so that $A' = B$ and $B' = A$
- The set *A* is called the extent (set of objects) and the set *B* is called the intent (set of attributes) of the formal concept *(A,B)*.
- If m ∈ M, then ({m}",{m}') is a formal concept. It is called the principal concept generated by m and denoted by $< m >$. Analogously, if g ∈ G, then ({g}',{g}") is a formal concept. It is called the principal concept generated by g and denoted by $< g >$. A formal concept is called a principal formal concept if it is equal to principal formal context generated by an element from *G* or *M*. The set of principal formal contexts form a partial order that is a suborder of the concept lattice

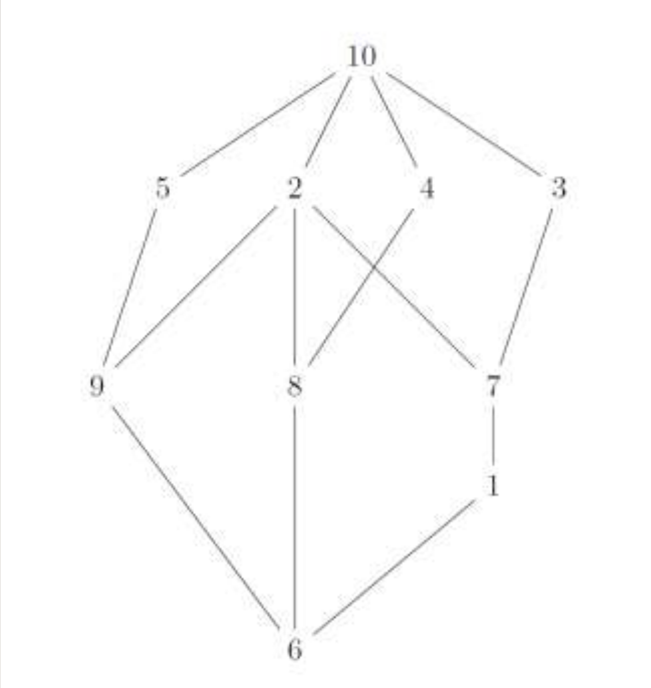
CONCEPT LATTICE

- The concept lattice of a formal context *(G,M,I)* is the set of all formal concepts together with the partial order $(A_1, B_1) \leq (A_2, B_2) :\Leftrightarrow A_1 \subseteq A_2 \ (\Leftrightarrow B_1 \supseteq B_2)$.
- Between two concepts (A_1, B_1) and (A_2, B_2) there is always a greatest common sub-concept and a least common super-concept.
- Concept lattice diagram is a mathematical structure based on the generality of the concept over the other concepts present in the set of all the concepts in a formal context. It can be visualized by its Hasse diagram.

HASSE DIAGRAM

- Graphically we can visualize a partial ordered set using a Hasse diagram. The Hasse diagram is characterized by:
    - There is a line between x and y if $x < y$ and there is no element z with $x < z < y$.
    - y is placed above x if $x < y$, i.e., smaller elements are placed lower than bigger elements.
    - Reflexivity and Transitivity edges are no shown.

CONCEPT LATTICE DIAGRAM



This program uses the LatDraw library, © 2002 Ralph Freese.

Concepts	Extents	Intents
1	$\{Mike\}$	{Student, Immigrant, Rich}
2	{John, Smith, Mike, Paul}	{Immigrant}
3	{Nick, Mike, Paul}	{Rich}
4	$\{John, Jack\}$	{Married}
5	$\{Smith, Ben\}$	{Employee}
6	{ϕ}	{Student, Immigrant, Rich, Married, Employee}
7	$\{Mike, Paul\}$	{Immigrant, Rich}
8	$\{John\}$	{Immigrant, Married}
9	$\{Smith\}$	{Immigrant, Employee}
10	{John, Smith, Nick, Mike,	{ϕ}

PROPOSED APPROACH

- Computing a concept lattice is very time consuming as we need to check every set of objects (resp. every set of attributes) whether it leads to a concept or not.
- Instead of computing it completely we use the following approach which needs to only compute $|G| + |M|$ elements and hence only takes polynomial time.
- We consider the $G + M$, i.e., the disjoint union of G and M. Now we define the relation on $G+M$ by distinguishing the four cases i.e., by defining a relation on pairs of objects, a relation on pairs of attributes, a relation between objects and attributes, and a relation between attributes and objects.

PROPOSED APPROACH – CONT'D

- Let *I* be the incident relation of a formal context. Then we define:
  - Relation on pairs of objects
    - $o_1(I/I)\ o_2 \Leftrightarrow \forall \mathrm{a} : o_2 I \mathrm{a} \ \rightarrow o_1 I \mathrm{a}$
  - Relation on pairs of attributes
    - $a_1(I \backslash I)\ a_2 \Leftrightarrow \forall \mathrm{o} : \mathrm{o} I a_1 \ \rightarrow \mathrm{o} I a_2$
  - Relation between attributes and objects
    - $\mathrm{a}\ bir(I)\ \mathrm{o} \Leftrightarrow \forall \mathrm{o}' \mathrm{a}'\ o' I \mathrm{a} \ \wedge\ \mathrm{o}\ I\ a' \rightarrow \mathrm{o}' I a'$
  - Relation between objects and attributes
    - Incidence relation *I*

PROPOSED APPROACH – CONT'D

- By taking the disjoint union of objects and attributes i.e. by merging *I* and the three relations above we obtain a new relation R as shown below:

$$\mathrm{R} = \begin{array}{cc} & \begin{array}{cc} G & M \end{array} \\ \begin{array}{c} G \\ M \end{array} & \begin{pmatrix} I/I & I \\ bir(I) & I\backslash \mathrm{I} \end{pmatrix} \end{array}$$

- R is defined in a matrix form and it is shown in [1] that R is a pre-order.

RELATION ON PAIRS OF OBJECTS

$o_1 (I/I)\ o_2 \Leftrightarrow \forall a : o_2 I a \ \rightarrow o_1 I a$

Objects

Objects

	John	Smith	Nick	Mike	Jack	Paul	Ben
John	X				X		
Smith		X					X
Nick			X				
Mike			X	X		X	
Jack					X		
Paul			X			X	
Ben							X

RELATION ON PAIRS OF ATTRIBUTES

$a_1(I\backslash I)\ a_2 \Leftrightarrow \forall \mathrm{o} : \mathrm{o}I a_1 \ \rightarrow \mathrm{o}I a_2$

Attributes

Attributes

	Student	Immigrant	Rich	Married	Employee
Student	X	X	X		
Immigrant		X			
Rich			X		
Married				X	
Employee					X

RELATION BETWEEN ATTRIBUTES AND OBJECTS

$a\ bir(I)\ o \Leftrightarrow \forall o'a'\ o'Ia \land o I a' \rightarrow o'Ia'$

Objects

Attributes

	John	Smith	Nick	Mike	Jack	Paul	Ben
Student			X	X		X	
Immigrant							
Rich			X				
Married					X		
Employee							X

DISJOINT UNION OF OBJECTS AND ATTRIBUTES

$G+M$

$G+\mathrm{M}$

	John	Smith	Nick	Mike	Jack	Paul	Ben	Student	Immigrant	Rich	Married	Employee
John	X				X				X		X	
Smith		X					X		X			X
Nick			X							X		
Mike			X	X		X		X	X	X		
Jack					X						X	
Paul			X			X			X	X		
Ben							X					X
Student			X	X		X		X	X	X		
Immigrant									X			
Rich			X							X		
Married					X						X	
Employee							X					X

EQUIVALENCE RELATION AND EQUIVALENCE CLASSES

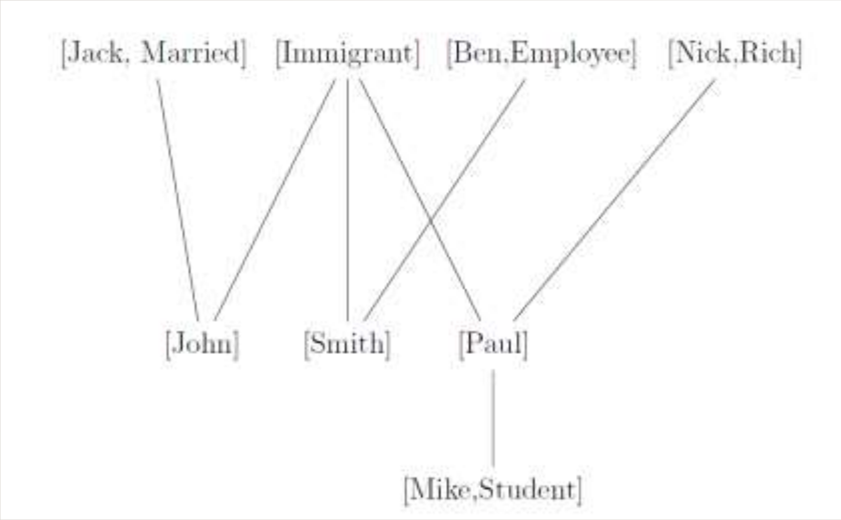
- An equivalence relation is a relation that is reflexive, symmetric, and transitive.
- If two elements are related by some equivalence relation, we will say that they are equivalent.
- Two elements x and y belong to equivalence class if and only if they are equivalent.
- In the given formal context of the previous slide we can observe that
  - Nick and Rich are equivalent.
  - Mike and Student are equivalent.
  - Jack and Married are equivalent.
  - Ben and Employee are equivalent.

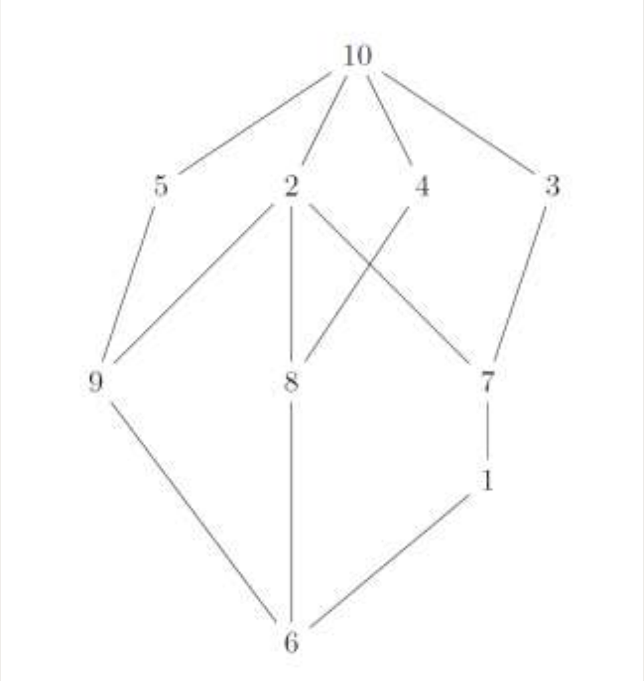
DISJOINT UNION OF *G+M* RELATION AFTER SETTING EQUIVALENCE CLASSES

	[John]	[Smith]	[Nick, Rich]	[Mike, Student]	[Jack, Married]	[Paul]	[Ben, Employee]	[Immigrant]
[John]	X				X			X
[Smith]		X					X	X
[Nick, Rich]			X					
[Mike, Student]			X	X		X		X
[Jack, Married]					X			
[Paul]			X			X		
[Ben, Employee]							X	
[Immigrant]								X

PARTIAL ORDER OF PRIMITIVE CONCEPTS DIAGRAM

- The completion of this partial order is isomorphic to the concept lattice of *I*.





APPLICATIONS

- Formal Concept Analysis or Concept Lattice Diagrams have applications in
  - Knowledge Representation and Discovery
  - Data Mining
  - Information Retrieval
  - Ontology
  - Machine Learning
  - Program Analysis

CONCLUSION

- Computing a concept lattice is very time consuming, so the disjoint union of $G + M$ when distinguished into four cases gives a pre-order relation which on completion is isomorphic to the concept lattice of I.

BIBLIOGRAPHY

- [1] Rudolf Berghammer and Michael Winter. Decomposition of relations and concept lattices. Fundamenta Informaticae, 126(1):37{82, 2013.
- [2] Bernhard Ganter, Gerd Stumme, and R Wille. Formal concept analysis: Methods and applications in computer science. TU Dresden, 2003.
- [3] Bernhard Ganter and Rudolf Wille. Formal concept analysis: mathematical foundations. Springer Science & Business Media, 2012.
- [4] Dmitry I Ignatov. Introduction to formal concept analysis and its applications in information retrieval and related fields. In Russian Summer School in Information Retrieval, pages 42–141. Springer, 2014.
- [5] Sergei O Kuznetsov. Machine learning on the basis of formal concept analysis. Automation and Remote Control, 62(10):1543–1564, 2001.

BIBLIOGRAPHY – CONT'D

- [6] Jirapond Muangprathub. A novel algorithm for building concept lattice. Applied Mathematical Sciences, 8:507–515, 01 2014.
- [7] Marek Obitko, Vaclav Snasel, Jan Smid, and V Snasel. Ontology design with formal concept analysis. In CLA, volume 128, pages 1377–90, 2004.
- [8] Jonas Poelmans, Paul Elzinga, Stijn Viaene, and Guido Dedene. Formal concept analysis in knowledge discovery: a survey. In International conference on conceptual structures, pages 139–153. Springer, 2010.
- [9] Gerd Stumme. Efficient data mining based on formal concept analysis. In International conference on database and expert systems applications, pages 534–546. Springer, 2002.
- [10] Uta Priss. Linguistic applications of formal concept analysis. In Formal Concept Analysis, pages 149–160. Springer, 2005.


thank
you!