Concept lattices of incomplete data

Michal Krupka and Jan Laštovička (Palacký University)

DATA ANALYSIS AND MODELING LAB
Palacky University, Olomouc, Czech Republic

european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness
INVESTMENTS IN EDUCATION DEVELOPMENT

Introduction

Burmeiser and Holzer (2000)

- introduce incomplete context
- use Kleene logic ($L = \{0, 1, ?\}$)
- apply to attribute implications and attribute exploration

Drawback of Kleene logic

- In Kleene logic the implication satisfies $(? \to ?) = ?$
  The implication $a \to a$ should always hold, even if the value of a is unknown.

Obiedkov (2002)

- use Modal logic instead of Kleene logic
- overcome problem of Kleene logic

Our goal

- construct a concept lattice of an incomplete context
  Requirements: contains all information and small

Outline

Outline

Concept-Forming Operators and Complete Homomorphisms

Let

- $h\colon \mathbf{L} \to \mathbf{L}'$ be a complete homomorphism of complete residuated lattices
- $\langle X, Y, I \rangle$ be a formal $\mathbf{L}$-context

Lemma

For each $A \in \mathbf{L}^X$ and $B \in \mathbf{L}^Y$ it holds

$$(h \circ A)^{\uparrow_{h \circ I}} = h \circ A^{\uparrow_I}, \qquad (h \circ B)^{\downarrow_{h \circ I}} = h \circ B^{\downarrow_I}.$$

Relationship of $\mathscr{B}(X,Y,I)$ and $\mathscr{B}(X,Y,h\circ I)$

Let

- $h\colon \mathbf{L}\to\mathbf{L}'$ be a complete homomorphism of complete residuated lattices
- $\langle X,Y,I\rangle$ be a formal $\mathbf{L}$-context

Theorem

- *For each formal concept $\langle A,B\rangle \in \mathscr{B}(X,Y,I)$ it holds $\langle h\circ A, h\circ B\rangle \in \mathscr{B}(X,Y,h\circ I)$.*
- *The induced mapping $h^{\mathscr{B}(X,Y,I)}\colon \mathscr{B}(X,Y,I)\to\mathscr{B}(X,Y,h\circ I)$ is a complete homomorphism.*
- *For each $\langle A_1,B_1\rangle, \langle A_2,B_2\rangle \in \mathscr{B}(X,Y,I)$ it holds*

$$\left(h^{\mathscr{B}(X,Y,I)}(\langle A_1,B_1\rangle) \preceq h^{\mathscr{B}(X,Y,I)}(\langle A_2,B_2\rangle)\right) = h(\langle A_1,B_1\rangle \preceq \langle A_2,B_2\rangle).$$

- *If h is injective, then so is $h^{\mathscr{B}(X,Y,I)}$.*
- *If h is surjective, then so is $h^{\mathscr{B}(X,Y,I)}$.*

Structure of $\mathscr{B}(X,Y,I)$

Let

- $\langle X,Y,I\rangle$ be a formal $\mathbf{L}$-context
- $\mathbf{L}$ be isomorphic to the direct product $\mathbf{L}_1 \times \mathbf{L}_2$
- $p_1 : L \to L_1$ and $p_2 : L \to L_2$ be the respective projections

Theorem

- *$\mathscr{B}(X,Y,I)$ is isomorphic to the direct product $\mathscr{B}(X,Y,p_1 \circ I) \times \mathscr{B}(X,Y,p_2 \circ I)$.*
- *The mappings $p_1^{\mathscr{B}(X,Y,I)}$ and $p_2^{\mathscr{B}(X,Y,I)}$ correspond to the respective Cartesian projections.*

Outline

Boolean Algebras

Example

Two-element Boolean algebra $\mathbf{2}$

Example

Boolean algebra of all mappings from a set U to $\mathbf{2}$, denoted by $\mathbf{2}^U$

Boolean Algebras as Residuated Lattices

Let $\mathbf{L} = \langle L, \wedge, \vee, ', 0, 1 \rangle$ be a Boolean algebra

We set:

$$a \otimes b = a \wedge b, \qquad a \to b = a' \vee b,$$

$\langle L, \wedge, \vee, \otimes, \to, 0, 1 \rangle$ is a residuated lattice

Boolean Algebras with Variables and Assignments

Let $\mathbf{L}$ be a finite Boolean algebra
together with a set of *variables* $U = \{u_1, \ldots, u_k\}$
and mapping $\iota : U \to L$, such that $\mathbf{L}$ is generated by $\iota(U)$.

Definition (Boolean algebras with variables)

- We call $\mathbf{L}$ a *Boolean algebra with variables* $u_1, \ldots, u_k$.
- If $\mathbf{L}$ is freely generated by $\iota(U)$, then the variables $u_1, \ldots, u_k$ are said to be *independent*.

Definition (assignments)

Mappings $v : U \to \mathbf{2}$ are called *assignments*.

Since $\mathbf{L}$ is generated by $\iota(U)$, then for each assignment v there exists at most one homomorphism $\bar{v} : \mathbf{L} \to \mathbf{2}$, such that $v = \bar{v} \circ \iota$.

Definition (admissible assignments)

If this homomorphism exists, then the assignment v is called *admissible*.

Admissible Assignments

- The variables $u_1, \ldots, u_k$ are independent if and only if each assignment is admissible.
- The other cases: dependencies between the unknown values

Example

If $\imath(u_1) \leq \imath(u_2)$, then there is no admissible assignment v such that $v(u_1) = 1$ and $v(u_2) = 0$.

Example

Let $\mathbf{L}$ be a Boolean algebra with variables u_1, u_2,
where assignment v with $v(u_1) = v(u_2) = 1$ is admissible.
Further, let $X = \{x_1, x_2, x_3\}$ be a set and $A = \{x_1, {}^{u_1 \vee u_2'}/x_2, {}^{u_1'}/x_3\}$ be an $\mathbf{L}$-set in X.

Then we obtain $\bar{v} \circ A = \{x_1, x_2\}$.

Construction of Boolean Algebras with Variables

Two natural requirements

1. Generality: for every possible dependency between variables there exists an appropriate Boolean algebra $\mathbf{L}$ (together with a mapping $\iota : U \to L$).
2. Efficiency: the Boolean algebra $\mathbf{L}$ is (up to isomorphism) the smallest element within the class of residuated lattices satisfying the generality requirement 1.

Theorem

The following holds for each subset $V \subseteq \mathbf{2}^U$.
1. Let $\mathbf{L} = \mathbf{2}^V$. Then there is a mapping $\iota : U \to L$ satisfying: For each $v \in V$ there is exactly one homomorphism $\bar{v} : \mathbf{L} \to \mathbf{2}$ such that $v = \bar{v} \circ \iota$.
2. Let $\mathbf{L}'$ be a residuated lattice and $\iota' : U \to L'$ a mapping such that for each $v \in V$ there is a homomorphism $\bar{v}' : \mathbf{L}' \to \mathbf{2}$ satisfying $v = \bar{v}' \circ \iota'$. Then there exists a surjective homomorphism of residuated lattices $s : \mathbf{L}' \to \mathbf{L}$ such that for each $v \in V$ it holds $\bar{v}' = \bar{v} \circ s$.

Example of Boolean Algebras with Variables

Let

- $U = \{u_1, u_2\}$
- $V = \{v_1, v_2, v_3\}$ with
  $v_1(u_1) = v_1(u_2) = 0$, $v_2(u_1) = 0$, $v_2(u_2) = 1$, and $v_3(u_1) = v_3(u_2) = 1$

$\mathbf{L} = \mathbf{2}^V$ is a Boolean algebra with variables u_1, u_2, where $\iota(u_1) \leq \iota(u_2)$

	u_1	u_2	0	$u_1' \wedge u_2$	u_2'	$u_1 \vee u_2'$	u_1'	1
$\bar{v}_1$	0	0	0	0	1	1	1	1
$\bar{v}_2$	0	1	0	1	0	0	1	1
$\bar{v}_3$	1	1	0	0	0	1	0	1

Elements of $\mathbf{L}$ are written in columns, their values in the mappings $\bar{v}_1$, $\bar{v}_2$, $\bar{v}_3$ are in rows.

Outline

Incomplete Contexts

Let $\mathbf{L}$ be a Boolean algebra with variables $u_1, \ldots, u_k$, $U = \{u_1, \ldots, u_k\}$

Suppose: $U \subseteq L$, $\imath$ is the inclusion $U \to L$.

Definition (incomplete context)

An *incomplete* $\mathbf{L}$*-context* is a triple $\langle X, Y, I \rangle$, where X and Y are sets and $I\colon X \times Y \to L$ is an $\mathbf{L}$-relation such that $I(X \times Y) \subseteq U \cup \{0, 1\}$.

Example

Boolean algebra $\mathbf{L}$ with variables u_1, u_2 is from the example.

	y_1	y_2	y_3	y_4	y_5
x_1			×	×	
x_2	u_1	×	u_2	×	
x_3	×	×	×		
x_4		×			

Completions

Definition (completion)

An ordinary formal context $\langle X, Y, J\rangle$ is a *completion* of $\langle X, Y, I\rangle$, if $J = \bar{v} \circ I$ for an admissible assignment $v : U \to \mathbf{2}$.

Example

$v_1(u_1) = v_1(u_2) = 0$

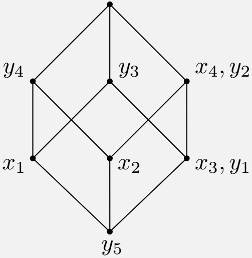
$\langle X, Y, I\rangle$

	y_1	y_2	y_3	y_4	y_5
x_1			×	×	
x_2	u_1	×	u_2	×	
x_3	×	×	×		
x_4		×			

$\langle X, Y, \bar{v}_1 \circ I\rangle$

	y_1	y_2	y_3	y_4	y_5
x_1			×	×	
x_2		×		×	
x_3	×	×	×		
x_4		×			

$\mathscr{B}(X, Y, \bar{v}_1 \circ I)$



Outline

$\mathbf{L}$-concept Lattice of an Incomplete $\mathbf{L}$-context

Let $\langle X, Y, I\rangle$ be incomplete $\mathbf{L}$-context.
$\mathscr{B}(X,Y,I)$ is the $\mathbf{L}$-concept lattice of $\mathbf{L}$-context $\langle X, Y, I\rangle$
Let v be an admissible assignment. $\bar{v}^{\mathscr{B}(X,Y,I)}$ is surjective.

Each $\langle A_0, B_0\rangle \in \mathscr{B}(X,Y,\bar{v}\circ I)$ can be obtained as the image of a concept $\langle A, B\rangle \in \mathscr{B}(X,Y,I)$ in the mapping $\bar{v}^{\mathscr{B}(X,Y,I)}$.

$\mathbf{L}$ is isomorphic to $\mathbf{2}^V$

Theorem

$\mathscr{B}(X,Y,I)$ is isomorphic to the direct product $\prod_{v\in V}\mathscr{B}(X,Y,\bar{v}\circ I)$. The mappings $\bar{v}^{\mathscr{B}(X,Y,I)}: \mathscr{B}(X,Y,I)\to\mathscr{B}(X,Y,\bar{v}\circ I)$ correspond to the respective Cartesian projections.

Example

The $\mathbf{L}$-concept lattice $\mathscr{B}(X,Y,I)$ of the incomplete context from the example is isomorphic to the direct product $\mathscr{B}(X,Y,\bar{v}_1\circ I)\times\mathscr{B}(X,Y,\bar{v}_2\circ I)\times\mathscr{B}(X,Y,\bar{v}_3\circ I)$ and has $8\cdot 8\cdot 7 = 448$ elements.

Crisply Generated Concepts of an Incomplete $\mathbf{L}$-context

Definition (Bělohlávek, Sklenář, Zacpal)

A formal $\mathbf{L}$-concept $\langle A,B\rangle \in \mathscr{B}(X,Y,I)$ is called *crisply generated*, if there is a crisp set $B_0 \subseteq Y$ such that $A = (B_0)^{\downarrow}$ (and hence $B = (B_0)^{\downarrow\uparrow}$). The set of all crisply generated $\mathbf{L}$-concepts is denoted by $\mathscr{B}_c(X,Y,I)$.

$\mathscr{B}_c(X,Y,I)$ is a complete lattice

Theorem

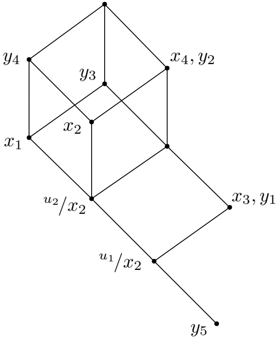
For each $v \in V$, the restriction $\bar{v}^{\mathscr{B}_c(X,Y,I)} : \mathscr{B}_c(X,Y,I) \to \mathscr{B}(X,Y,v \circ I)$ of $\bar{v}^{\mathscr{B}(X,Y,I)}$ is

- *a surjective,*
- *$\bigwedge$-preserving mapping, such that*
- *for each $\langle A_1,B_1\rangle, \langle A_2,B_2\rangle \in \mathscr{B}_c(X,Y,I)$ it holds $\bar{v}^{\mathscr{B}_c(X,Y,I)}(\langle A_1,B_1\rangle) \leq \bar{v}^{\mathscr{B}_c(X,Y,I)}(\langle A_2,B_2\rangle)$ iff $\bar{v}(\langle A_1,B_1\rangle \preceq \langle A_2,B_2\rangle) = 1$.*

Example of $\mathscr{B}_c(X,Y,I)$

$\langle X,Y,I\rangle$	$\mathscr{B}_c(X,Y,I)$

	y_1	y_2	y_3	y_4	y_5
x_1			$\times$	$\times$	
x_2	u_1	$\times$	u_2	$\times$	
x_3	$\times$	$\times$	$\times$		
x_4		$\times$			



Elements of $\mathscr{B}_c(X,Y,I)$:

$$\langle\{x_1,x_2,x_3,x_4\},\emptyset\rangle, \langle\{x_1,x_2\},\{{}^{u_2}/y_3,y_4\}\rangle, \langle\{x_1,{}^{u_2}/x_2,x_3\},\{y_3\}\rangle,$$

$$\langle\{x_2,x_3,x_4\},\{y_2\}\rangle, \langle\{x_1,{}^{u_2}/x_2\},\{y_3,y_4\}\rangle, \langle\{x_2\},\{{}^{u_1}/y_1,y_2,{}^{u_2}/y_3,y_4\}\rangle,$$

$$\langle\{{}^{u_2}/x_2,x_3\},\{{}^{u_1\vee u_2'}/y_1,y_2,y_3\}\rangle, \langle\{{}^{u_2}/x_2\},\{{}^{u_1\vee u_2'}/y_1,y_2,y_3,y_4,{}^{u_2'}/y_5\}\rangle,$$

$$\langle\{{}^{u_1}/x_2,x_3\},\{y_1,y_2,y_3\}\rangle, \langle\{{}^{u_1}/x_2\},\{y_1,y_2,y_3,y_4,{}^{u_1'}/y_5\}\rangle, \langle\emptyset,\{y_1,y_2,y_3,y_4,y_5\}\rangle$$

Size of $\mathscr{B}_c(X,Y,I)$

- in the worst case - exponentially
- experiments - linear (for small number of variables)

Results of experiments

- The Digits Context (48 concepts)

No. of variables	1	2	3	4	5	6	7	8	9	10
Avg. no. of concepts	50.31	54.31	56.26	59.38	63.49	65.84	68.21	71.01	73.05	77.39
Std. deviation	3.16	6.34	6.51	8.00	8.19	8.67	10.37	11.71	11.31	11.67

- The Tea Ladies Context (65 concepts)

No. of variables	1	2	3	4	5	6	7
Avg. no. of concepts	67.88	71.54	74.73	78.3	82.12	86.09	89.67
Std. deviation	2.56	3.79	4.66	5.0	7.53	7.02	8.44

Outline

Conclusions and Future Work

Conclusions

- Our approach could find practical applications.

Future work

- Attribute implications in our framework.
- More results on the size of concept lattices of incomplete contexts.

Remark

- Generalizing the theory to FCA in fuzzy setting was done. Article was sent to IPMU 2012.