The word problem in semiconcept algebras

Philippe Balbiani

Institut de recherche en informatique de Toulouse

Aims of this talk

To present a short introduction to **formal concept analysis**

To define and to study **concept algebras**

To sketch a proof that **the word problem in semiconcept algebras** is *PSPACE*-complete

Formal concept analysis

Formal concept analysis

A context for the planets

	small	medium	large	near	far	yes	no
Mercury	×			×			×
Venus	×			×			×
Earth	×			×		×	
Mars	×			×		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Formal concept analysis

A context for the planets

Objects: the nine planets (*Mercury*, *Venus*, etc)

Attributes: the seven properties (*small*, *medium*, etc)

Concepts: ordered pairs (A, B) where

- A is a set of planets
- B is a set of properties
- A should contain just those planets sharing all the properties in B
- B should contain just those properties shared by all the planets in A

Formal concept analysis

A context for the planets

	small	medium	large	near	far	yes	no
Mercury	⊗			⊗			×
Venus	⊗			⊗			×
Earth	⊗			⊗		×	
Mars	⊗			⊗		×	
Jupiter			×		×	×	
Saturn			×		×	×	
Uranus		×			×	×	
Neptune		×			×	×	
Pluto	×				×	×	

Contexts, objects and attributes

Contexts: triples $\mathcal{S} = (Obj, Att, I)$ where *Obj* and *Att* are nonempty sets and $I \subseteq Obj \times Att$

$\mathcal{S}$-objects: elements of *Obj* (X, Y, etc)

$\mathcal{S}$-attributes: elements of *Att* (x, y, etc)

Polars and concepts

$\mathcal{S}$-polars: for $A \subseteq Obj$ and $B \subseteq Att$, define

- $A' = \{x \in Att\colon \text{for all } X \in A,\ I(X, x)\}$
- $B' = \{X \in Obj\colon \text{for all } x \in B,\ I(X, x)\}$

$\mathcal{S}$-concepts: pairs (A, B) where $A \subseteq Obj$ — the extent — and $B \subseteq Att$ — the intent — are such that

- $B' = A$
- $A' = B$

Concept lattice of a context $\mathcal{S} = (Obj, Att, I)$

$\mathcal{C}(\mathcal{S})$: set of all $\mathcal{S}$-concepts

$\leq$: $(A_1, B_1) \leq (A_2, B_2) \iff A_1 \subseteq A_2$ and $B_1 \supseteq B_2$

Formal concept analysis

Representation theorem

Let

- $\mathcal{S} = (Obj, Att, I)$ be a **context**

Then

- $(\mathcal{C}(\mathcal{S}), \leq)$ is a **complete lattice** in which join and meet are given by
  - $\bigvee_{j \in J}(A_j, B_j) = ((\bigcup_{j \in J} A_j)'', \bigcap_{j \in J} B_j)$
  - $\bigwedge_{j \in J}(A_j, B_j) = (\bigcap_{j \in J} A_j, (\bigcup_{j \in J} B_j)'')$

Let

- L be a **complete lattice**

Then

- there exists a **context** $\mathcal{S} = (Obj, Att, I)$ such that L is isomorphic to $(\mathcal{C}(\mathcal{S}), \leq)$

Formal concept analysis

Returning to the planets

	small	*medium*	*large*	*near*	*far*	*yes*	*no*
1 : *Mercury*	×			×			×
2 : *Venus*	×			×			×
3 : *Earth*	×			×		×	
4 : *Mars*	×			×		×	
5 : *Jupiter*			×		×	×	
6 : *Saturn*			×		×	×	
7 : *Uranus*		×			×	×	
8 : *Neptune*		×			×	×	
9 : *Pluto*	×				×	×	

The concept ($\emptyset$, {*small*, *medium*, *large*, *near*, *far*, *yes*, *no*})

Formal concept analysis

Returning to the planets

	small	*medium*	*large*	*near*	*far*	*yes*	*no*
1 : *Mercury*	⊗			⊗			⊗
2 : *Venus*	⊗			⊗			⊗
3 : *Earth*	×			×		×	
4 : *Mars*	×			×		×	
5 : *Jupiter*			×		×	×	
6 : *Saturn*			×		×	×	
7 : *Uranus*		×			×	×	
8 : *Neptune*		×			×	×	
9 : *Pluto*	×				×	×	

The concept $(\{1, 2\}, \{small, near, no\})$

Formal concept analysis

Returning to the planets

	small	*medium*	*large*	*near*	*far*	*yes*	*no*
1 : *Mercury*	×			×			×
2 : *Venus*	×			×			×
3 : *Earth*	⊗			⊗		⊗	
4 : *Mars*	⊗			⊗		⊗	
5 : *Jupiter*			×		×	×	
6 : *Saturn*			×		×	×	
7 : *Uranus*		×			×	×	
8 : *Neptune*		×			×	×	
9 : *Pluto*	×				×	×	

The concept $(\{3, 4\}, \{small, near, yes\})$

Formal concept analysis

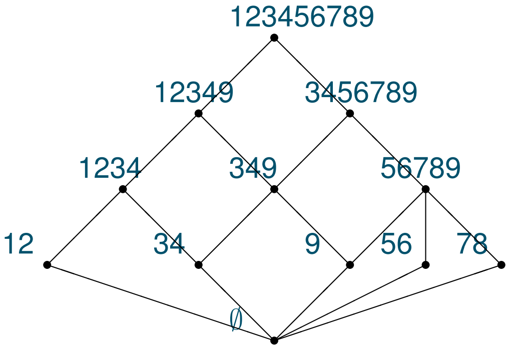
Returning to the planets

	small	*medium*	*large*	*near*	*far*	*yes*	*no*
1 : *Mercury*	×			×			×
2 : *Venus*	×			×			×
3 : *Earth*	×			×		×	
4 : *Mars*	×			×		×	
5 : *Jupiter*			×		×	×	
6 : *Saturn*			×		×	×	
7 : *Uranus*		×			×	×	
8 : *Neptune*		×			×	×	
9 : *Pluto*	⊗				⊗	⊗	

The concept $(\{9\}, \{small, far, yes\})$

Formal concept analysis

Returning to the planets



Concept algebras

Concept algebras

Join, meet and complement of concepts in context $\mathcal{S} = (Obj, Att, I)$

Join of concepts (A_1, B_1) and (A_2, B_2)

- $((A_1 \cup A_2)'', B_1 \cap B_2)$

Meet of concepts (A_1, B_1) and (A_2, B_2)

- $(A_1 \cap A_2, (B_1 \cup B_2)'')$

Complement of concept (A, B)

- $(Obj \setminus A, -)$? No since • is not always an extent
- $(-, Att \setminus B)$? No since • is not always an intent
- $((Obj \setminus A)'', (Obj \setminus A)')$? No since • may intersect A
- $((Att \setminus B)', (Att \setminus B)'')$? No since • may intersect B

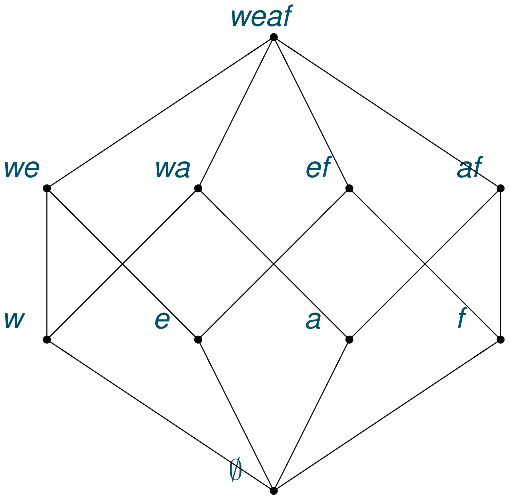
Concept algebras

Empedocle's conception of the four elements

	cold	*moist*	*dry*	*warm*
water	×	×		
earth	×		×	
air		×		×
fire			×	×

Concept algebras

Concept lattice of Empedocle's conception of the four elements



Concept algebras

Concepts, semiconcepts and protoconcepts

Contexts

- $\mathcal{S} = (Obj, Att, I)$ be a context
- $A \subseteq Obj$ be a set of objects
- $B \subseteq Att$ be a set of attributes

Concepts

- (A, B) is a $\mathcal{S}$-concept iff $B' = A$ and $A' = B$

Semiconcepts

- (A, B) is a $\mathcal{S}$-semiconcept iff $B' = A$ or $A' = B$

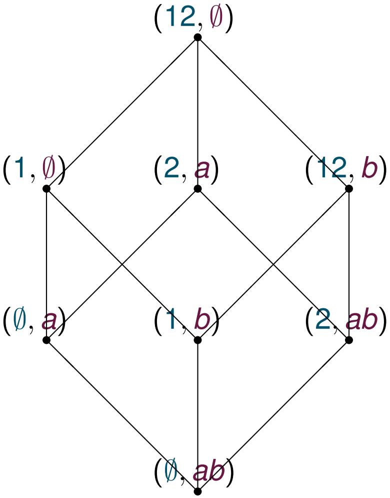
Protoconcepts

- (A, B) is a $\mathcal{S}$-protoconcept iff $B' = A''$ or $A' = B''$

Concept algebras

Example

	a	b
1		$\times$
2	$\times$	$\times$



Concept algebras

Protoconcept algebra of context $\mathcal{S} = (Obj, Att, I)$

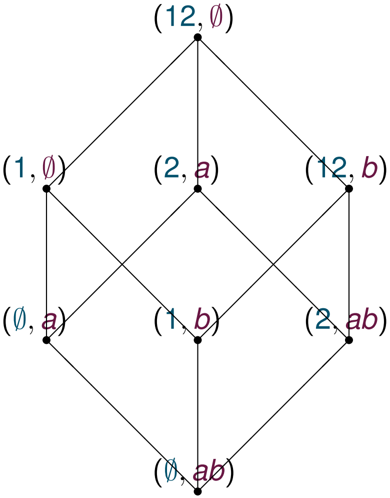
Structure $\mathcal{A}(\mathcal{S}) = (A^{\mathcal{S}}, \bot_l^{\mathcal{S}}, \top_r^{\mathcal{S}}, \top_l^{\mathcal{S}}, \bot_r^{\mathcal{S}}, \neg_l^{\mathcal{S}}, \neg_r^{\mathcal{S}}, \vee_l^{\mathcal{S}}, \wedge_r^{\mathcal{S}}, \wedge_l^{\mathcal{S}}, \vee_r^{\mathcal{S}})$ where $A^{\mathcal{S}}$ is the set of all $\mathcal{S}$'s protoconcepts and

- $\bot_l^{\mathcal{S}} = (\emptyset, Att)$
- $\top_r^{\mathcal{S}} = (Obj, \emptyset)$
- $\top_l^{\mathcal{S}} = (Obj, Obj')$
- $\bot_r^{\mathcal{S}} = (Att', Att)$
- $\neg_l^{\mathcal{S}}(A, B) = (Obj \setminus A, (Obj \setminus A)')$
- $\neg_r^{\mathcal{S}}(A, B) = ((Att \setminus B)', Att \setminus B)$
- $(A_1, B_1) \vee_l^{\mathcal{S}} (A_2, B_2) = (A_1 \cup A_2, (A_1 \cup A_2)')$
- $(A_1, B_1) \wedge_r^{\mathcal{S}} (A_2, B_2) = ((B_1 \cup B_2)', B_1 \cup B_2)$
- $(A_1, B_1) \wedge_l^{\mathcal{S}} (A_2, B_2) = (A_1 \cap A_2, (A_1 \cap A_2)')$
- $(A_1, B_1) \vee_r^{\mathcal{S}} (A_2, B_2) = ((B_1 \cap B_2)', B_1 \cap B_2)$

Concept algebras

Protoconcept algebra: example

	a	b
1		×
2	×	×



Concept algebras

Protoconcept algebra: equational theory

- $\wedge_l$ is AC — $\vee_r$ is AC
- $\wedge_l$ distributes over $\vee_l$ — $\vee_r$ distributes over $\wedge_r$
- $\neg_l(x \wedge_l x) = \neg_l x$ — $\neg_r(x \vee_r x) = \neg_r x$
- $x \wedge_l (y \wedge_l y) = x \wedge_l y$ — $x \vee_r (y \vee_r y) = x \vee_r y$
- $x \wedge_l (x \vee_l y) = x \wedge_l x$ — $x \vee_r (x \wedge_r y) = x \vee_r x$
- $x \wedge_l (x \vee_r y) = x \wedge_l x$ — $x \vee_r (x \wedge_l y) = x \vee_r x$
- $\neg_l(\neg_l x \wedge_l \neg_l y) = x \vee_l y$ — $\neg_r(\neg_r x \vee_r \neg_r y) = x \wedge_r y$
- $\neg_l \bot_l = \top_l$ — $\neg_r \top_r = \bot_r$
- $\neg_l \top_r = \bot_l$ — $\neg_r \bot_l = \top_r$
- $\top_r \wedge_l \top_r = \top_l$ — $\bot_l \vee_r \bot_l = \bot_r$
- $x \wedge_l \neg_l x = \bot_l$ — $x \vee_r \neg_r x = \top_r$
- $\neg_l \neg_l (x \wedge_l y) = x \wedge_l y$ — $\neg_r \neg_r (x \vee_r y) = x \vee_r y$
- $(x \vee_r x) \wedge_l (x \vee_r x) = (x \wedge_l x) \vee_r (x \wedge_l x)$

Concept algebras

Semiconcept algebra of context $\mathcal{S} = (Obj, Att, I)$

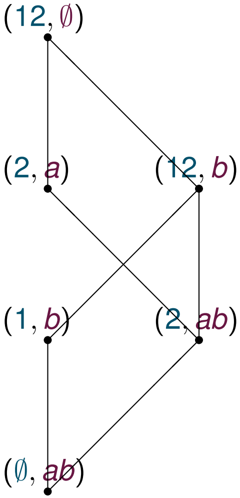
Structure $\mathcal{A}(\mathcal{S}) = (A^{\mathcal{S}}, \bot_l^{\mathcal{S}}, \top_r^{\mathcal{S}}, \top_l^{\mathcal{S}}, \bot_r^{\mathcal{S}}, \neg_l^{\mathcal{S}}, \neg_r^{\mathcal{S}}, \vee_l^{\mathcal{S}}, \wedge_r^{\mathcal{S}}, \wedge_l^{\mathcal{S}}, \vee_r^{\mathcal{S}})$ where $A^{\mathcal{S}}$ is the set of all $\mathcal{S}$'s semiconcepts and

- $\bot_l^{\mathcal{S}} = (\emptyset, Att)$
- $\top_r^{\mathcal{S}} = (Obj, \emptyset)$
- $\top_l^{\mathcal{S}} = (Obj, Obj')$
- $\bot_r^{\mathcal{S}} = (Att', Att)$
- $\neg_l^{\mathcal{S}}(A, B) = (Obj \setminus A, (Obj \setminus A)')$
- $\neg_r^{\mathcal{S}}(A, B) = ((Att \setminus B)', Att \setminus B)$
- $(A_1, B_1)\vee_l^{\mathcal{S}}(A_2, B_2) = (A_1 \cup A_2, (A_1 \cup A_2)')$
- $(A_1, B_1)\wedge_r^{\mathcal{S}}(A_2, B_2) = ((B_1 \cup B_2)', B_1 \cup B_2)$
- $(A_1, B_1)\wedge_l^{\mathcal{S}}(A_2, B_2) = (A_1 \cap A_2, (A_1 \cap A_2)')$
- $(A_1, B_1)\vee_r^{\mathcal{S}}(A_2, B_2) = ((B_1 \cap B_2)', B_1 \cap B_2)$

Concept algebras

Semiconcept algebra: example

	a	b
1		×
2	×	×



Concept algebras

Semiconcept algebra: equational theory

- $\wedge_l$ is AC — $\vee_r$ is AC
- $\wedge_l$ distributes over $\vee_l$ — $\vee_r$ distributes over $\wedge_r$
- $\neg_l(x \wedge_l x) = \neg_l x$ — $\neg_r(x \vee_r x) = \neg_r x$
- $x \wedge_l (y \wedge_l y) = x \wedge_l y$ — $x \vee_r (y \vee_r y) = x \vee_r y$
- $x \wedge_l (x \vee_l y) = x \wedge_l x$ — $x \vee_r (x \wedge_r y) = x \vee_r x$
- $x \wedge_l (x \vee_r y) = x \wedge_l x$ — $x \vee_r (x \wedge_l y) = x \vee_r x$
- $\neg_l(\neg_l x \wedge_l \neg_l y) = x \vee_l y$ — $\neg_r(\neg_r x \vee_r \neg_r y) = x \wedge_r y$
- $\neg_l \bot_l = \top_l$ — $\neg_r \top_r = \bot_r$
- $\neg_l \top_r = \bot_l$ — $\neg_r \bot_l = \top_r$
- $\top_r \wedge_l \top_r = \top_l$ — $\bot_l \vee_r \bot_l = \bot_r$
- $x \wedge_l \neg_l x = \bot_l$ — $x \vee_r \neg_r x = \top_r$
- $\neg_l \neg_l(x \wedge_l y) = x \wedge_l y$ — $\neg_r \neg_r(x \vee_r y) = x \vee_r y$
- $(x \vee_r x) \wedge_l (x \vee_r x) = (x \wedge_l x) \vee_r (x \wedge_l x)$
- $x \wedge_l x = x$ or $x \vee_r x = x$

Concept algebras

Semiconcept algebra: representation theorem

Let $\mathcal{S} = (Obj, Att, I)$ be a **context**

- If $A^{\mathcal{S}}$ is the set of all $\mathcal{S}$'s semiconcepts then the structure $\mathcal{A}(\mathcal{S}) = (A^{\mathcal{S}}, \bot_l^{\mathcal{S}}, \top_r^{\mathcal{S}}, \neg_l^{\mathcal{S}}, \neg_r^{\mathcal{S}}, \vee_l^{\mathcal{S}}, \wedge_r^{\mathcal{S}})$ is a **semiconcept algebra**

Let $\mathcal{A} = (A, \bot_l, \top_r, \neg_l, \neg_r, \vee_l, \wedge_r)$ be a **semiconcept algebra**

- There exists a **context** $\mathcal{S}(\mathcal{A}) = (Obj^{\mathcal{A}}, Att^{\mathcal{A}}, I^{\mathcal{A}})$ such that $\mathcal{A}$ is embeddable into the structure $\mathcal{A}(\mathcal{S}(\mathcal{A})) = (A^{\mathcal{S}(\mathcal{A})}, \bot_l^{\mathcal{S}(\mathcal{A})}, \top_r^{\mathcal{S}(\mathcal{A})}, \neg_l^{\mathcal{S}(\mathcal{A})}, \neg_r^{\mathcal{S}(\mathcal{A})}, \vee_l^{\mathcal{S}(\mathcal{A})}, \wedge_r^{\mathcal{S}(\mathcal{A})})$

The word problem in semiconcept algebras

The word problem in semiconcept algebras
Syntax

We define terms as follows

- $s ::= x \mid 0_l \mid 1_r \mid -_l s \mid -_r s \mid (s \sqcup_l t) \mid (s \sqcap_r t)$

We define the following abbreviations

- $1_l ::= -_l 0_l$
- $0_r ::= -_r 1_r$
- $(s \sqcap_l t) ::= -_l(-_l s \sqcup_l -_l t)$
- $(s \sqcup_r t) ::= -_r(-_r s \sqcap_r -_r t)$

The word problem in semiconcept algebras

Semantics

A valuation based on a semiconcept algebra $\mathcal{A} = (A, \bot_l, \top_r, \neg_l, \neg_r, \vee_l, \wedge_r)$ is a function

- $\theta \colon x \mapsto \theta(x) \in A$

θ induces a function $\bar{\theta} \colon s \mapsto \bar{\theta}(s) \in A$ as follows:

- $\bar{\theta}(x) = \theta(x)$
- $\bar{\theta}(0_l) = \bot_l$
- $\bar{\theta}(1_r) = \top_r$
- $\bar{\theta}(-_l s) = \neg_l \bar{\theta}(s)$
- $\bar{\theta}(-_r s) = \neg_r \bar{\theta}(s)$
- $\bar{\theta}(s \sqcup_l t) = \bar{\theta}(s) \vee_l \bar{\theta}(t)$
- $\bar{\theta}(s \sqcap_r t) = \bar{\theta}(s) \wedge_r \bar{\theta}(t)$

The word problem in semiconcept algebras

The word problem

Input:

- Terms s, t

Output:

- Decide whether $s \not\approx t$, i.e. whether there exists a valuation θ based on a semiconcept algebra $\mathcal{A} = (A, \bot_l, \top_r, \neg_l, \neg_r, \vee_l, \wedge_r)$ such that $\bar{\theta}(s) \neq \bar{\theta}(t)$

Computational complexity:

- The word problem in semiconcept algebras is *PSPACE*-complete

K^2: a basic 2-sorted modal logic:

- Syntax:
  - $F ::= P \mid \bot \mid \neg F \mid (F \vee G) \mid \Box f$
  - $f ::= p \mid \bot \mid \neg f \mid (f \vee g) \mid \Box F$
- Semantics:
  - $\mathcal{M} = (\mathcal{S}, V)$ where $\mathcal{S} = (Obj, Att, I)$ is a context and:
    - V: $P \mapsto V(P) \subseteq Obj$
    - v: $p \mapsto v(p) \subseteq Att$
  - $\mathcal{M}, X \models P$ iff $X \in V(P)$
  - $\mathcal{M}, x \models p$ iff $x \in v(p)$
  - $\mathcal{M}, X \models \Box f$ iff for all $x \in Att$, if $\mathcal{M}, x \models f$ then XIx
  - $\mathcal{M}, x \models \Box F$ iff for all $X \in Obj$, if $\mathcal{M}, X \models F$ then XIx
- The satisfiability problem for K^2 is *PSPACE*-complete

Restriction of the syntax of K^2:

- $F ::= P \mid \bot \mid \neg F \mid (F \vee G) \mid \Box f$
- $f ::= \bot \mid \neg f \mid (f \vee g) \mid \Box F$

Syntax of the word problem:

- $s ::= x \mid 0_l \mid 1_r \mid -_l s \mid -_r s \mid (s \sqcup_l t) \mid (s \sqcap_r t)$

A reduction from K^2 to the word problem:

- $T(P_i) = x_i$
- $T(\bot) = 0_l$ $\quad t(\bot) = 1_r$
- $T(\neg F) = -_l T(F)$ $\quad t(\neg f) = -_r t(f)$
- $T(F \vee G) = T(F) \sqcup_l T(G)$ $\quad t(f \vee g) = t(f) \sqcap_r t(g)$
- $T(\Box f) = -_l -_l -_r -_r t(f)$ $\quad t(\Box F) = -_r -_r -_l -_l T(F)$

F is satisfiable iff $T(F) \not\approx 0_l$

The word problem in semiconcept algebras

Solving the word problem

Syntax of the word problem:

- $s ::= x \mid 0_l \mid 1_r \mid -_l s \mid -_r s \mid (s \sqcup_l t) \mid (s \sqcap_r t)$

Syntax of K^2:

- $F ::= P \mid \bot \mid \neg F \mid (F \vee G) \mid \Box f$
- $f ::= p \mid \bot \mid \neg f \mid (f \vee g) \mid \Box F$

A reduction from the word problem to K^2:

- $F(x_i) = P_i$ — $f(x_i) = p_i$
- $F(0_l) = \bot$ — $f(0_l) = \Box\bot$
- $F(1_r) = \Box\bot$ — $f(1_r) = \bot$
- $F(-_l s) = \neg F(s)$ — $f(-_l s) = \Box\neg F(s)$
- $F(-_r s) = \Box\neg f(s)$ — $f(-_r s) = \neg f(s)$
- $F(s \sqcup_l t) = F(s) \vee F(t)$ — $f(s \sqcup_l t) = \Box(F(s) \vee F(t))$
- $F(s \sqcap_r t) = \Box(f(s) \vee f(t))$ — $f(s \sqcap_r t) = f(s) \vee f(t)$

$s \not\approx t$ iff $\neg(F(s) \leftrightarrow F(t))$ is p-satisfiable or $\neg(f(s) \leftrightarrow f(t))$ is p-satisfiable

Conclusion

What we have done

- *PSPACE*-completeness of the word problem in semiconcept algebras

Open problems

- *PSPACE*-completeness of the word problem in protoconcept algebras
- Algorithm for deciding the word problem in semiconcept/protoconcept algebras
- Algorithm for drawing semiconcept/protoconcept algebras

Bibliography

- Davey, B., Priestley, H.: *Introduction to Lattices and Order.* Cambridge University Press (2002, Second Edition).
- Ganter, B., Wille, R.: *Formal Concept Analysis. Mathematical Foundations.* Springer-Verlag (1999).
- Herrmann, C., Luksch, P., Skorsky, M., Wille, R.: *Algebras of semiconcepts and double Boolean algebras.* Technische Universität Darmstadt (2000).
- Vormbrock, B.: *A solution of the word problem for free double Boolean algebras.* In Kuznetsov, S., Schmidt, S. (Editors): *ICFCA 2007.* Springer-Verlag (2007) 240–270.
- Wille, R.: *Boolean concept logic.* In Eklund, P. (Editor): *ICFCA 2004.* Springer-Verlag (2004) 1–13.
- Wille, R.: *Preconcept algebras and generalized double Boolean algebras.* In Ganter, B., Mineau, G. (Editors): *ICCS 2000.* Springer-Verlag (2000) 317–331.