Formal Concept Analysis

Bernhard Ganter

Institut für Algebra
TU Dresden
D-01062 Dresden
bernhard.ganter@tu-dresden.de

Description Logics Workshop Dresden 2008

Outline

1. Concept lattices
   - Data from a hospital
   - Formal definitions
   - More examples
2. Elements of structure theory
   - Context arithmetic
   - Context product
   - Order relation bond
3. Attribute logic
   - The base
4. Rough Sets
   - Traditional and generalised

Outline

Outline

1. Concept lattices
   - Data from a hospital
   - Formal definitions
   - More examples
2. Elements of structure theory
   - Context arithmetic
   - Context product
   - Order relation bond
3. Attribute logic
   - The base
4. Rough Sets
   - Traditional and generalised

Outline

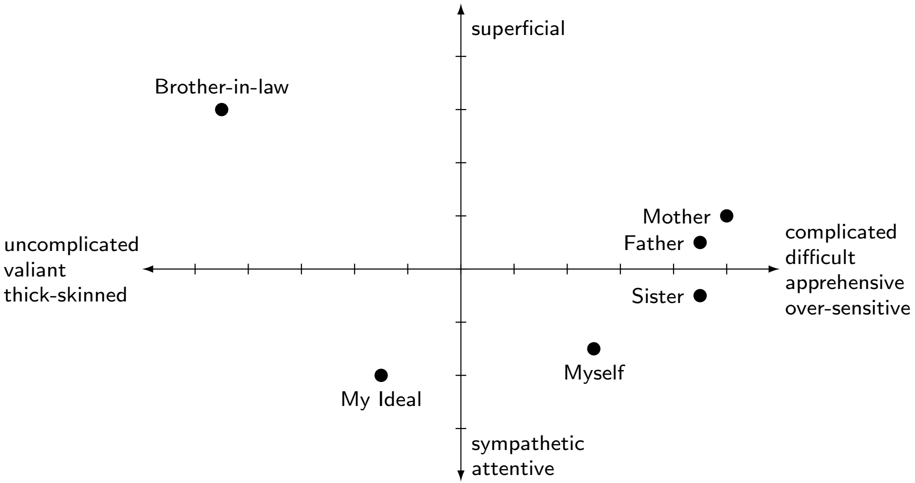
Outline

1. Concept lattices
   - Data from a hospital
   - Formal definitions
   - More examples
2. Elements of structure theory
   - Context arithmetic
   - Context product
   - Order relation bond
3. Attribute logic
   - The base
4. Rough Sets
   - Traditional and generalised

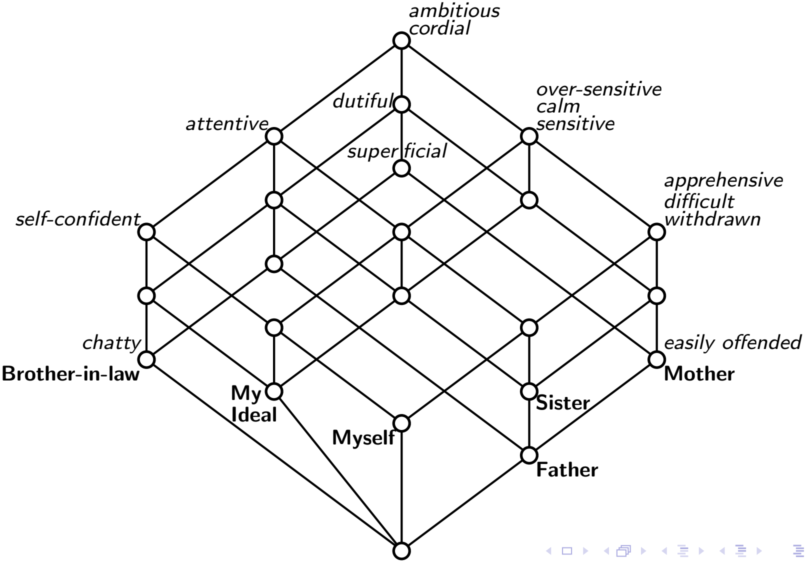
Interview data from a treatment of Anorexia nervosa

	over-sensitive	withdrawn	self-confident	dutiful	cordial	difficult	attentive	easily offended	calm	apprehensive	chatty	superficial	sensitive	ambitious
Myself	×	×	×		×	×	×		×	×			×	×
My Ideal	×		×	×	×		×		×				×	×
Father	×	×		×	×	×	×	×	×	×		×	×	×
Mother	×	×		×	×	×		×	×	×		×	×	×
Sister	×	×		×	×	×	×		×	×			×	×
Brother-in-law			×	×	×		×				×	×		×

A biplot of the interview data



The concept lattice of the interview data



Unfolding data in a concept lattice

The basic procedure of Formal Concept Analysis:

- Data is represented in a very basic data type, called a **formal context**.
- Each formal context is transformed into a mathematical structure called a **concept lattice**. The information contained in the formal context is preserved.
- The concept lattice is the basis for further data analysis. It may be represented graphically to support communication, or it may be investigated with algebraic methods to unravel its structure.

Unfolding data in a concept lattice

The basic procedure of Formal Concept Analysis:

- Data is represented in a very basic data type, called a **formal context**.
- Each formal context is transformed into a mathematical structure called a **concept lattice**. The information contained in the formal context is preserved.
- The concept lattice is the basis for further data analysis. It may be represented graphically to support communication, or it may be investigated with algebraic methods to unravel its structure.

Unfolding data in a concept lattice

The basic procedure of Formal Concept Analysis:

- Data is represented in a very basic data type, called a **formal context**.
- Each formal context is transformed into a mathematical structure called a **concept lattice**. The information contained in the formal context is preserved.
- The concept lattice is the basis for further data analysis. It may be represented graphically to support communication, or it may be investigated with algebraic methods to unravel its structure.

Unfolding data in a concept lattice

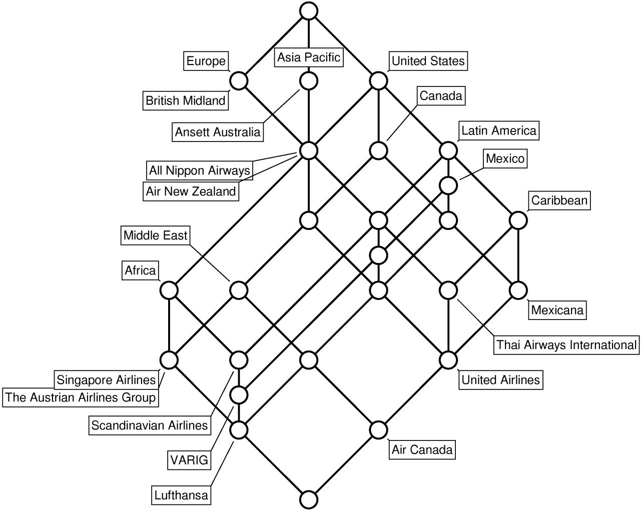
The basic procedure of Formal Concept Analysis:

- Data is represented in a very basic data type, called a **formal context**.
- Each formal context is transformed into a mathematical structure called a **concept lattice**. The information contained in the formal context is preserved.
- The concept lattice is the basis for further data analysis. It may be represented graphically to support communication, or it may be investigated with algebraic methods to unravel its structure.

An example about airlines . . .

	Latin America	Europe	Canada	Asia Pacific	Middle East	Africa	Mexico	Caribbean	United States
Air Canada	×	×	×	×	×		×	×	×
Air New Zealand		×		×					×
All Nippon Airways		×		×					×
Ansett Australia				×					
The Austrian Airlines Group		×	×	×	×	×			×
British Midland		×							
Lufthansa	×	×	×	×	×	×	×		×
Mexicana	×		×				×	×	×
Scandinavian Airlines	×	×		×		×			×
Singapore Airlines		×	×	×	×	×			×
Thai Airways International	×	×		×				×	×
United Airlines	×	×	×	×			×	×	×
VARIG	×	×		×		×	×		×

. . . and its concept lattice



Formal contexts

A **formal context** (G, M, I) consists of sets G, M and a binary relation $I \subseteq G \times M$.

For $A \subseteq G$ and $B \subseteq M$, define

$$\begin{aligned} A' &:= \{m \in M \mid g\ I\ m \text{ for all } g \in A\} \\ B' &:= \{g \in G \mid g\ I\ m \text{ for all } m \in B\}. \end{aligned}$$

The mappings

$$X \to X''$$

are closure operators.

Formal concepts

(A, B) is a **formal concept** of (G, M, I) iff

$$A \subseteq G, \quad B \subseteq M, \quad A' = B, \quad A = B'.$$

A is the **extent** and B is the **intent** of (A, B).

Concept lattice

Formal concepts can be ordered by

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

The set $\underline{\mathfrak{B}}(G, M, I)$ of all formal concepts of (G, M, I), with this order, is a complete lattice, called the **concept lattice** of (G, M, I).

The basic theorem

Theorem (The basic theorem of Formal Concept Analysis.) *The concept lattice of any formal context (G, M, I) is a complete lattice. For an arbitrary set $\{(A_i, B_i) \mid i \in I\} \subseteq \mathfrak{B}(G, M, I)$ of formal concepts, the supremum is given by*

$$\bigvee_{i\in I}(A_i, B_i) = \left((\bigcup_{i\in I} A_i)'', \bigcap_{i\in I} B_i \right)$$

and the infimum is given by

$$\bigwedge_{i\in I}(A_i, B_i) = \left(\bigcap_{i\in I} A_i, (\bigcup_{i\in I} B_i)'' \right).$$

A complete lattice L is isomorphic to $\underline{\mathfrak{B}}(G, M, I)$ iff there are mappings $\tilde{\gamma} : G \to L$ and $\tilde{\mu} : M \to L$ such that $\tilde{\gamma}(G)$ is supremum-dense and $\tilde{\mu}(M)$ is infimum-dense in L, and

$$g\ I\ m \iff \tilde{\gamma}(g) \leq \tilde{\mu}(m).$$

In particular, $L \cong \underline{\mathfrak{B}}(L, L, \leq)$.

Why complete lattices?

- Complete lattices are closely related to *closure operators* and thereby to *closure systems* (also called *Moore families*).
- Closure operators are closely related to propositional Horn theories. The formulas are *implications*, i.e., expressions of the form
$$A \to B,$$
where A and B are sets of attributes.

Combining this, we may say that

Complete lattices are the algebraic structures corresponding to propositional Horn theories.

Why algebraic structures?

- The mathematical approach is different from that of Logic. We ask different questions and thereby get other results.
- We admire the theoretical and algorithmic strength of DL. At the same time, we can build on a rich and well-developed mathematical theory of lattices and ordered sets.
- Structure theoretic aspects, such as decompositions, morphisms, representations and classifications can be combined with established visualisations.
- We try to systematically build a mathematical theory, the parts of which are as simple and as general as possible.

The boolean lattice

What are the formal concepts of $(G, G, \neq)$?

They are precisely the pairs

$$(A, G \setminus A) \quad A \subseteq G.$$

$(G, G, \neq)$ therefore is the standard context of the power set lattice of G.

	a	b	c	d
a		×	×	×
b	×		×	×
c	×	×		×
d	×	×	×	

Outline

Lattice constructions and context constructions

Many lattice properties and lattice constructions have nice counterparts on the formal context side. For example,

the dual lattice	corresponds to	the transposed formal context
complete sublattices	correspond to	closed subrelations
congruence relations	correspond to	compatible subcontexts
tolerance relations	correspond to	block relations
direct products	correspond to	context sums
tensor products	correspond to	context products
Galois connections	correspond to	bonds

and so on.

Closed relations and block relations

- A **closed relation** of (G, M, I) is a subrelation $C \subseteq I$ such that every formal concept of (G, M, C) also is a concept of (G, M, I).
  The closed relations correspond to the complete sublattices of $\underline{\mathfrak{B}}(G, M, I)$.
- A **block relation** of (G, M, I) is a superrelation $R \supseteq I$ such that every extent of (G, M, R) is an extent of (G, M, I) and every intent of (G, M, R) is an intent of (G, M, I).
  The block relations correspond to the complete tolerance relations of $\underline{\mathfrak{B}}(G, M, I)$.

Closed relations and block relations

- A **closed relation** of (G, M, I) is a subrelation $C \subseteq I$ such that every formal concept of (G, M, C) also is a concept of (G, M, I).
  The closed relations correspond to the complete sublattices of $\underline{\mathfrak{B}}(G, M, I)$.
- A **block relation** of (G, M, I) is a superrelation $R \supseteq I$ such that every extent of (G, M, R) is an extent of (G, M, I) and every intent of (G, M, R) is an intent of (G, M, I).
  The block relations correspond to the complete tolerance relations of $\underline{\mathfrak{B}}(G, M, I)$.

The context sum

The concept lattice of the context sum

$$(G, M, I) + (H, N, J)$$

is isomorphic to the direct product

$$\underline{\mathfrak{B}}(G, M, I) \times \underline{\mathfrak{B}}(H, N, J).$$

	M	N
G	I	$\times$
H	$\times$	J

Context product and tensor product

The **product** of formal contexts (G, M, I) and (H, N, J) is defined to be the context

$$(G, M, I) \times (H, N, J) := (G \times H, M \times N, \nabla),$$

where

$$(g, h) \nabla (m, n) :\iff g\ I\ m \text{ or } h\ J\ n.$$

The concept lattice

$$\underline{\mathfrak{B}}((G, M, I) \times (H, N, J))$$

of the product is the **tensor product**

$$\underline{\mathfrak{B}}(G, M, I) \otimes \underline{\mathfrak{B}}(H, N, J).$$

Context product and tensor product

The **product** of formal contexts (G, M, I) and (H, N, J) is defined to be the context

$$(G, M, I) \times (H, N, J) := (G \times H, M \times N, \nabla),$$

where

$$(g, h) \nabla (m, n) :\iff g\ I\ m \text{ or } h\ J\ n.$$

The concept lattice

$$\underline{\mathfrak{B}}((G, M, I) \times (H, N, J))$$

of the product is the **tensor product**

$$\underline{\mathfrak{B}}(G, M, I) \otimes \underline{\mathfrak{B}}(H, N, J).$$

An example of a context product

×		
	×	×
	×	

×

×	
	×

=

×	×	×		×	
×	×		×		×
×		×	×	×	×
	×	×	×	×	×
×		×	×	×	
	×	×	×		×

Bonds and closed relations of the sum

$R \subseteq G \times N$ is a bond from (G, M, I) to (H, N, J) if and only if

$$I \cup R \cup J \cup (H \times M)$$

is a closed relation of the context sum

$$(G, M, I) + (H, N, J)$$

(assuming $G \cap H = \emptyset = M \cap N$).

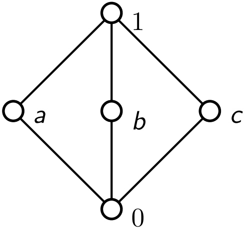
	M	N
G	I	R
H	×	J

The sublattice corresponding to the identity bond

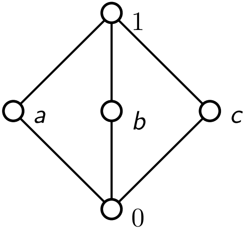
The concept lattice of the formal context made from the identity bond is isomorphic to the order relation of $\underline{\mathfrak{B}}(G, M, I)$, with the componentwise order.

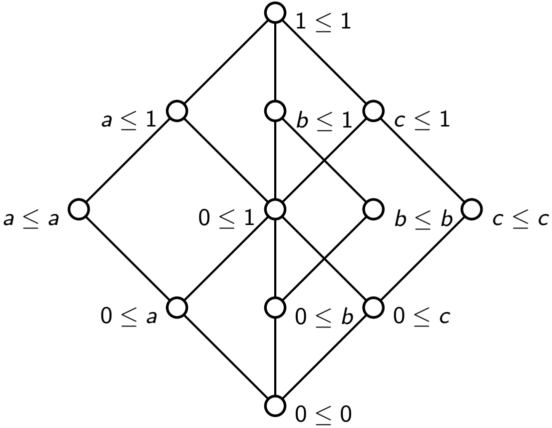
	M	M
G	I	I
G	×	I

The order lattice of M_3



The order lattice of M_3





The identity bond as a context product

Note that

$$\begin{array}{c|c|c|} & M & M \\ \hline G & I & I \\ \hline G & \times & I \\ \hline \end{array} \;=\; \begin{array}{|c|c|}\hline & \\ \hline \times & \\ \hline \end{array} \;\times\; (G, M, I)$$

In other words: The order relation of a lattice is isomorphic to the tensor product of the lattice with a three-element chain.

The identity bond as a context product

Note that

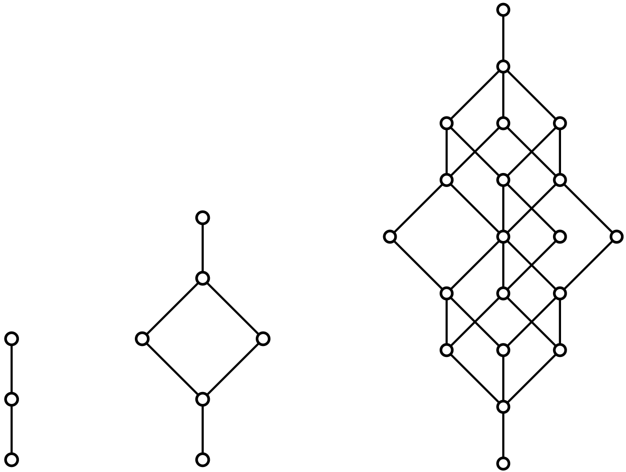
$$\begin{array}{c|c|c|} & M & M \\ \hline G & I & I \\ \hline G & \times & I \\ \hline \end{array} \quad = \quad \begin{array}{|c|c|} \hline & \\ \hline \times & \\ \hline \end{array} \quad \times \quad (G, M, I)$$

In other words: The order relation of a lattice is isomorphic to the tensor product of the lattice with a three-element chain.

Free distributive lattices

- The tensor product of n three-element chains is the **free distributive lattice** on n generators, $FCD(n)$, also known as the lattice of all **monotone boolean functions** in n variables.
- The order relation of $FCD(n)$ is isomorphic to the tensor product of $FCD(n)$ and a three-element chain. Therefore
- The free distributive lattice on $n+1$ generators is isomorphic to the order relation of the free distributive lattice on n generators.

Free distributive lattices $FCD(1)$, $FCD(2)$, $FCD(3)$



Outline

The stem base

Duquenne and Gigues have discovered that each implicational theory has a canonical base (with respect to the Armstrong derivation rules). We call this the **stem base**.

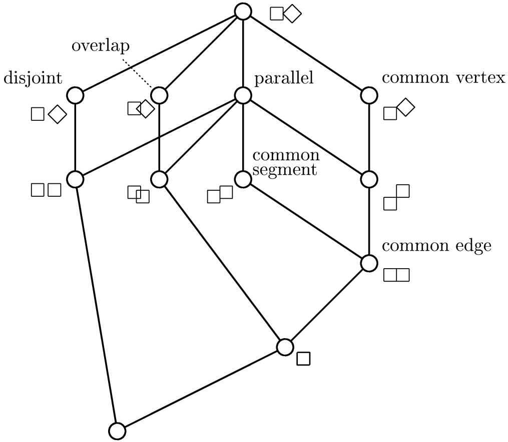
To find the stem base for a given formal context (G, M, I) requires the notion of a pseudo-intent: $P \subseteq M$ is a **pseudo-intent** iff

- P is not an intent and
- P contains the closure of every pseudo-intent that is properly contained in P.

The stem base then is

$$\{P \to P'' \mid P \text{ is a pseudo-intent}\}.$$

Pairs of squares



But did we consider all possible cases?

How can we decide if our selection of examples is complete?

A possible strategy is to prove that every implication that holds for these examples, holds in general.

- Compute the stem base of the context of examples, and
- prove that these implications hold in general,
- or find counterexamples and extend the example set.

This can nicely be organised in an algorithm, called **attribute exploration**.

But did we consider all possible cases?

How can we decide if our selection of examples is complete?

A possible strategy is to prove that every implication that holds for these examples, holds in general.

- Compute the stem base of the context of examples, and
- prove that these implications hold in general,
- or find counterexamples and extend the example set.

This can nicely be organised in an algorithm, called **attribute exploration**.

But did we consider all possible cases?

How can we decide if our selection of examples is complete?

A possible strategy is to prove that every implication that holds for these examples, holds in general.

- Compute the stem base of the context of examples, and
- prove that these implications hold in general,
- or find counterexamples and extend the example set.

This can nicely be organised in an algorithm, called **attribute exploration**.

Stem base of the example set

- common edge $\longrightarrow$ parallel, common vertex, common segment
- common segment $\longrightarrow$ parallel
- parallel, common vertex, common segment $\longrightarrow$ common edge
- overlap, common vertex $\longrightarrow$ parallel, common segment, common edge
- overlap, parallel, common segment $\longrightarrow$ common edge, common vertex
- overlap, parallel, common vertex $\longrightarrow$ common segment, common edge
- disjoint, common vertex $\longrightarrow \perp$
- disjoint, parallel, common segment $\longrightarrow \perp$
- disjoint, overlap $\longrightarrow \perp$

Two of the implications do not hold in general

- common edge $\longrightarrow$ parallel, common vertex, common segment
- common segment $\longrightarrow$ parallel
- parallel, common vertex, common segment $\longrightarrow$ common edge
- overlap, common vertex $\longrightarrow$ parallel, common segment, common edge
- overlap, parallel, common segment $\longrightarrow$ common edge, common vertex
- overlap, parallel, common vertex $\longrightarrow$ common segment, common edge
- disjoint, common vertex $\longrightarrow \perp$
- disjoint, parallel, common segment $\longrightarrow \perp$
- disjoint, overlap $\longrightarrow \perp$

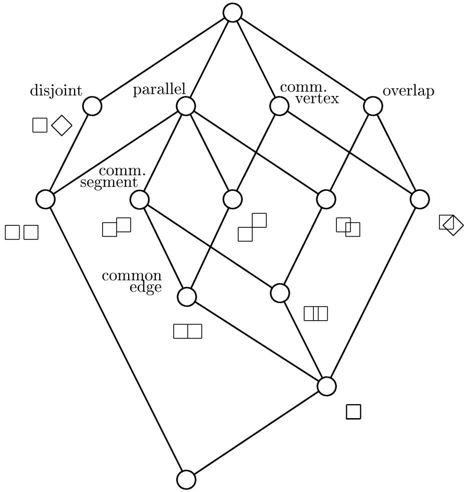
Conterexamples for the two implications

- overlap, common vertex $\longrightarrow$ parallel, common segment, common edge
- Counterexample: 
- overlap, parallel, common segment $\longrightarrow$ common edge, common vertex
- Counterexample: 

Conterexamples for the two implications

- overlap, common vertex → parallel, common segment, common edge
- Counterexample: 
- overlap, parallel, common segment → common edge, common vertex
- Counterexample: 

A better choice of examples

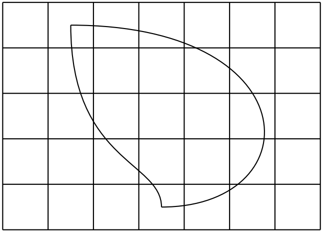


Outline

Zdzisław Pawlak's rough sets

An **approximation space** $(U, \sim)$ is a set U with an equivalence relation $\sim$ on U.

A subset $A \subseteq U$ is **rough** if it is not a union of equivalence classes.



Approximations

A rough set A is represented (approximately) by a pair $(\underline{R}(A), \overline{R}(A))$, giving the

lower approximation $\underline{R}(A)$

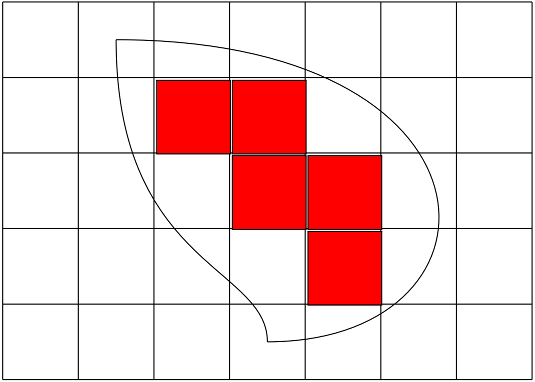
and the

upper approximation $\overline{R}(A)$

of A.

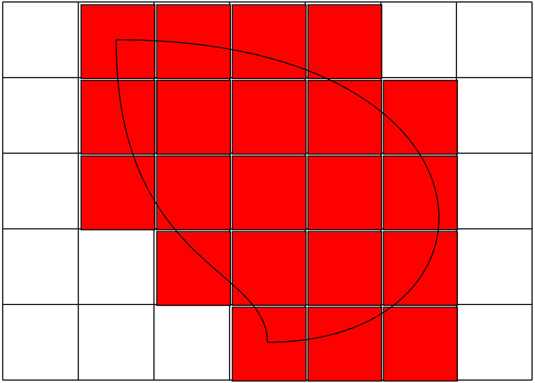
In this way, a *interval arithmetic* for sets is established.

The lower approximation



$\underline{R}(A)$

The upper approximation



$\overline{R}(A)$

The algebra of approximating pairs

The algebra of approximating pairs (also called rough sets) is well studied. It can be obtained from the formal context shown here.

	U	U
U	$\not\sim$	$\not\sim$
U	×	$\not\sim$

Note this formal context is the context product of the indiscernibility relation $(u, u, \not\sim)$ with the context of a three element chain. The lattice of rough sets therefore is the lattice of all intervals of the power set of indiscernibility equivalence classes. It is a Stone lattice.

Operator pairs for generalised approximation

Meanwhile, rough sets have been generalised. Approximations are considered that do not rely on equivalence relations.
Natural assumptions to keep are that

- $\underline{R}$ is a kernel operator, and
- $\overline{R}$ is a closure operator.

The family of approximating pairs

$$(\underline{R}(A), \overline{R}(A))$$

then generates (but not necessarily forms!) a sublattice of the direct product of the lattice of kernels and the lattice of closures.

The approximation lattice

The lattice generated by such generalised approximations can be characterised in general: With some technical cheating, the formal context is as given in the diagram, where

(G, M, I) is a formal context whose extent-complements are the kernel sets,

the extents of (G, N, J) are the closures, and

$m \perp n :\iff m' \cup n' = G$.

	G	N
M	I^{-1}	$\perp$
G	×	J

A natural generalisation

From the general characterisation it can be derived that there is a unique generalisation having most of the nice properties of the classical case. It occurs when the indiscernibility equivalence relation is generalised to an **indiscernibility quasi-order**.

For an indiscernibility quasi-order the algebra of approximations gets a strikingly simple form.

It is the context product of $(G, G, \ngeq)$ with the context of the three element chain.

The concept lattice therefore is isomorphic to the order relation of the lattice of quasi-order ideals.

	G	G
G	$\ngeq$	$\ngeq$
G	×	$\ngeq$

A non-symmetric indiscernibility?

We have arrived at a seemingly counter-intuitive conclusion: The mathematically most promising generalisation of the rough set approach is obtained by replacing the indiscernibility relation by a not necessarily symmetric quasi-order.
Does non-symmetric indiscernibility make any sense?

Yes, it does! The intuition that an attribute-based indiscernibility should be symmetric tacitely assumes that attributes are scaled *nominally*, that is, with mutually exclusive attribute values.
The generalised version is natural when subsumption between attribute values is permitted.

A non-symmetric indiscernibility?

We have arrived at a seemingly counter-intuitive conclusion: The mathematically most promising generalisation of the rough set approach is obtained by replacing the indiscernibility relation by a not necessarily symmetric quasi-order.
Does non-symmetric indiscernibility make any sense?

Yes, it does! The intuition that an attribute-based indiscernibility should be symmetric tacitely assumes that attributes are scaled *nominally*, that is, with mutually exclusive attribute values.
The generalised version is natural when subsumption between attribute values is permitted.

Thank you for your attention!