Relational Lattices

Tadeusz Litak: Informatik 8, FAU

Szabolcs Mikulás: Birkbeck, University of London

Jan Hidders: TU Delft

29 Sept 2014

Contents

- ¿ We study an interpretation of lattice connectives proposed by Vadim Tropashko from Oracle: as natural join and inner union between DB relations ?
- ¿ We show that this interpretation yields a class of lattices which has not been considered in the existing lattice-theoretical literature ?
- ¿ We propose an equational axiomatization for a corresponding abstract algebraic class ?
- ¿ It turns out that addition of just the *header constant* to the lattice signature leads to undecidabilty of the quasiequational theory ?
- ¿ Relational lattices, however, are not as intangible as one may fear: for example, they do form a pseudoelementary class ?
- ¿ We also apply the tools of Formal Concept Analysis and investigate standard contexts of relational lattices. In particular, we'll see that finite relational lattices are subdirectly irreducible ?

We investigate the (named) relational model of database theory from an algebraic point of view

We investigate the (named) relational model of database theory from an algebraic point of view

Databases in relational model

Collections of relations over a fixed domain $\mathcal{D}omain$
with possibly different arities and different headers
taken from some fixed supply of attributes $\mathcal{A}ttr$

We investigate the (named) relational model of database theory from an algebraic point of view

Databases in relational model

Collections of relations over a fixed domain $\mathcal{D}omain$ with possibly different arities and different headers taken from some fixed supply of attributes $\mathcal{A}ttr$

A relation R

Consists of:

- Header (scheme): $H(R) \subseteq \mathcal{A}ttr$
- Body: a set of tuples that have the same attributes

$$Body(R) \subseteq \mathcal{D}omain^{H(R)}$$

An Example

Prefecture	Region
Mie	Kansai
Shiga	Kansai
Aomori	Tohoku
Iwate	Tohoku
Aichi	Chubu
Kagawa	Shikoku
Ishikawa	Chubu

LOCATION

```
Capital     Prefecture
```

Takamatsu	Kagawa
Kanazawa	Ishikawa
Sapporo	Hokkaido
Yokohama	Kanagawa
Naha	Okinawa

```
CAPITALS
```

Codd Relational algebra (not *relation algebra*!)

Partial algebra underlying the relational model

Codd Relational algebra (not *relation algebra*!)

Partial algebra underlying the relational model

Basic operations (*named* or *typed* case):

- selection $\sigma_{\phi} R$ where ϕ is a condition on attribute values
- projection $\pi_A R$ where A is a collection of attribute names
- (natural) join $R_1 \bowtie R_2$ (not to be confused with lattice join!)
- renaming $\rho_{a_1 \mapsto b_1, \ldots, a_n \mapsto b_n} R$ where $\overline{a}$ and $\overline{b}$ are attribute names
- union $R_1 \cup R_2$ (crashes for inputs with different types)
- difference (relative complementation) $R_1 - R_2$ (as above)

Codd Relational algebra (not *relation algebra*!)

Partial algebra underlying the relational model

Basic operations (*named* or *typed* case):

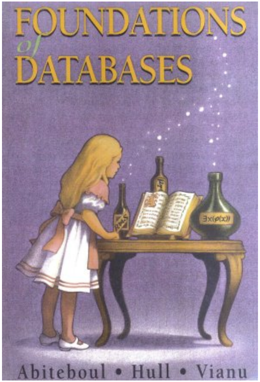
- selection $\sigma_\phi R$ where ϕ is a condition on attribute values
- projection $\pi_A R$ where A is a collection of attribute names
- (natural) join $R_1 \bowtie R_2$ (not to be confused with lattice join!)
- renaming $\rho_{a_1 \mapsto b_1, \ldots, a_n \mapsto b_n} R$ where $\overline{a}$ and $\overline{b}$ are attribute names
- union $R_1 \cup R_2$ (crashes for inputs with different types)
- difference (relative complementation) $R_1 - R_2$ (as above)

The goal: expressive completeness

ability to express domain-independent FO queries

(in an algebraic language)

. . . the Alice Book!



Available freely online: `http://webdam.inria.fr/Alice/`

Some of you may ask why Codd did not choose

- *cylindric algebras* (CA's)
- *polyadic algebras* (PA's)
- or *Tarski's relation algebras* (RA's)?

Reason I: History

Charles S. Peirce
pioneer of: algebra of relations,
FOL , SOL, computer science . . .

Arthur W. Burks
one of ENIAC designers
and cellular automata pioneers,
Peirce editor, one-time president
of Peirce Society

Edgar A. Codd
relational calculus, relational algebra . . .

SQL, Oracle, modern DB theory . . .

Ernst Schröder

Alfred Tarski

Maddux, Givant, Monk,
Jónsson, Pigozzi,
Hungarians starting with
Andreka & Németi . . .

Imieliński and Lipski in 1980's:
relational algebra can be embedded into cylindric algebra

Reason II: DB People Do Not Want Uniform Headers

Database reasons to be unhappy with CA's:

Reason II: DB People Do Not Want Uniform Headers

Database reasons to be unhappy with CA's:

- Tarskian uniformization:

  We artificially make all headers equal
  filling missing columns (axes, dimensions, attributes)
  with all possible values from the domain

Reason II: DB People Do Not Want Uniform Headers

Database reasons to be unhappy with CA's:

- Tarskian uniformization:

  We artificially make all headers equal
  filling missing columns (axes, dimensions, attributes)
  with all possible values from the domain

- Unrestricted nature of CA operations
  (cylindrification and unary negation)

Reason II: DB People Do Not Want Uniform Headers

Database reasons to be unhappy with CA's:

- Tarskian uniformization:

  We artificially make all headers equal
  filling missing columns (axes, dimensions, attributes)
  with all possible values from the domain

- Unrestricted nature of CA operations
  (cylindrification and unary negation)

Immediate problems with

- Domain Independence Property
- Boyce-Codd Normal Form . . .

Definition

A query expression $\phi(R_1, \ldots, R_n)$ is domain-independent if for any two domains of values $\mathcal{D}omain_1$, $\mathcal{D}omain_2$ and relations $\mathtt{R}_1$, ..., $\mathtt{R}_n$ s.t. for any $i \leq n$

$$Body(\mathtt{R}_i) \subseteq \mathcal{D}omain_1^{H(\mathtt{R}_i)} \cap \mathcal{D}omain_2^{H(\mathtt{R}_i)}$$

it holds that

$$[\phi(\mathtt{R}_1, \ldots, \mathtt{R}_n)]_{\mathcal{D}omain_1} = [\phi(\mathtt{R}_1, \ldots, \mathtt{R}_n)]_{\mathcal{D}omain_2}$$

In other words, the value of such a query depends only on its active domain

Queries which are not domain-independent are often called unsafe

The Effect of Tarskian Uniformization I

Capital	Prefecture	Region
Takamatsu	Mie	Kansai
Sapporo	Mie	Kansai
Yokohama	Mie	Kansai
Naha	Mie	Kansai
Takamatsu	Shiga	Kansai
Sapporo	Shiga	Kansai
Yokohama	Shiga	Kansai
Naha	Shiga	Kansai
Takamatsu	Aomori	Tohoku
Sapporo	Aomori	Tohoku
Yokohama	Aomori	Tohoku
Naha	Aomori	Tohoku
Takamatsu	Iwate	Tohoku
Sapporo	Iwate	Tohoku
Yokohama	Iwate	Tohoku
Naha	Iwate	Tohoku
	...	

$\texttt{UNIF\_LOCATION} = (\{\texttt{Capital}\}, \mathcal{D}omain^{\{\texttt{Capital}\}}) \bowtie \texttt{LOCATION}$

(assuming $\mathcal{A}ttr = \{\texttt{Capital}, \texttt{Prefecture}, \texttt{Region}\}$)

The Effect of Tarskian Uniformization II

Capital	Prefecture	Region
Takamatsu	Kagawa	Chubu
Takamatsu	Kagawa	Shikoku
Takamatsu	Kagawa	Kansai
Takamatsu	Kagawa	Tohoku
Kanazawa	Ishikawa	Chubu
Kanazawa	Ishikawa	Shikoku
Kanazawa	Ishikawa	Kansai
Kanazawa	Ishikawa	Tohoku
Sapporo	Hokkaido	Chubu
Sapporo	Hokkaido	Shikoku
Sapporo	Hokkaido	Kansai
Sapporo	Hokkaido	Tohoku
	...	

$\mathtt{UNIF\_CAPITALS} = \mathtt{CAPITALS} \bowtie (\{\mathtt{Region}\}, \mathcal{D}omain^{\{\mathtt{Region}\}})$

(assuming $\mathcal{A}ttr = \{\mathtt{Capital}, \mathtt{Prefecture}, \mathtt{Region}\}$)

Can We Do Any Better Then?

Thus, it seems of intrinsic interest to investigate an alternative,
heterogeneous algebraic setting
which models the database operations more faithfully

Can We Do Any Better Then?

Thus, it seems of intrinsic interest to investigate an alternative, heterogeneous algebraic setting
which models the database operations more faithfully

Perhaps one can obtain more positive results concerning equational/quasiequational theory?

On the algebraic side

Craig's finite axiomatization for

another heterogeneous and expressively complete
algebra of finite sequences

(his setting is somewhat different:

even *single relations* are not homogenous)

On the algebraic side

Craig's finite axiomatization for

another heterogeneous and expressively complete
algebra of finite sequences

(his setting is somewhat different:

even *single relations* are not homogenous)

On the DB side

Relational algebra operations without relative difference yield unions of conjunctive queries:
a paradigm example of well-behaved, decidable class of queries

Our goal

To provide a heterogeneous algebraic structure for relational queries

- which is a total algebra

  (unlike Codd's relational algebra)
- which preserves the Domain Independence Property

  (unlike CA's)
- whose primitive operations are natural from both DB and algebraic point of view

and investigate its equational/quasi-equational theory

First Basic Operation: Natural Join ⋈

Prefecture	Region
Mie	Kansai
Shiga	Kansai
Aomori	Tohoku
Iwate	Tohoku
Aichi	Chubu
Kagawa	Shikoku
Ishikawa	Chubu

LOCATION

First Basic Operation: Natural Join $\bowtie$

Capital	Prefecture
Takamatsu	Kagawa
Kanazawa	Ishikawa
Sapporo	Hokkaido
Yokohama	Kanagawa
Naha	Okinawa

CAPITALS

First Basic Operation: Natural Join $\bowtie$

Capital	Prefecture	Region
Takamatsu	Kagawa	Shikoku
Kanazawa	Ishikawa	Chubu

CAPITALS $\bowtie$ LOCATION

Formal Definition of $\bowtie$

$$(H(R), Body(R)) \bowtie (H(S), Body(S)) = (H(R) \cup H(S), Body(R) \bowtie Body(S))$$

where

$$Body(R) \bowtie Body(S) = \{t \in \mathcal{D}omain^{H(R) \cup H(S)} \,|\, t \upharpoonright_{H(R)} \in Body(R) \,\&$$
$$\& \, t \upharpoonright_{H(S)} \in Body(S)\}$$

$$(H(R), Body(R)) \bowtie (H(S), Body(S)) = (H(R) \cup H(S), Body(R) \bowtie Body(S))$$

where

$$Body(R) \bowtie Body(S) = \{t \in \mathcal{D}omain^{H(R) \cup H(S)} \mid t \restriction_{H(R)} \in Body(R) \, \& \\ \& \, t \restriction_{H(S)} \in Body(S)\}$$

$\bowtie$ intuitively: a hybrid of product and intersection

Can be considered a generalization of relational composition

Second Basic Operation: Inner Union ⊕

Prefecture	Region
Mie	Kansai
Shiga	Kansai
Aomori	Tohoku
Iwate	Tohoku
Aichi	Chubu
Kagawa	Shikoku
Ishikawa	Chubu

LOCATION

Second Basic Operation: Inner Union ⊕

Capital	Prefecture
Takamatsu	Kagawa
Kanazawa	Ishikawa
Sapporo	Hokkaido
Yokohama	Kanagawa
Naha	Okinawa

CAPITALS

Second Basic Operation: Inner Union $\oplus$

Prefecture
Mie
Shiga
Aomori
Iwate
Aichi
Kagawa
Ishikawa
Hokkaido
Kanagawa
Okinawa

CAPITALS $\oplus$ LOCATION

Formal Definition of $\oplus$

$$(H(R), Body(R)) \oplus (H(S), Body(S)) = (H(R) \cap H(S), Body(R) \oplus Body(S))$$

where

$$Body(R) \oplus Body(S) = Body(R) \restriction_{H(R) \cap H(S)} \cup Body(S) \restriction_{H(R) \cap H(S)}$$

Formal Definition of $\oplus$

$$(H(R), Body(R)) \oplus (H(S), Body(S)) = (H(R) \cap H(S), Body(R) \oplus Body(S))$$

where

$$Body(R) \oplus Body(S) = Body(R) \restriction_{H(R) \cap H(S)} \cup Body(S) \restriction_{H(R) \cap H(S)}$$

$\oplus$ intuitively: a hybrid of union and projection

As opposed to union in relational algebras,
does not require that both arguments have the same header

Extensions of the Signature (Sec. 6.1)

Just by adding constants to the basic ($\oplus$, $\bowtie$)–signature,
it is possible to define/express/simulate
the following operations of relational algebra:

- projections to concrete headers (inner unions with attribute constants)
- constant-based selection queries (natural joins with unary singleton constants)
- equality-based selection queries (natural join with equality constants)

  not respecting d.i.p., we need an unary operator to do better

However, in this paper we work either

- with pure lattice signature $\mathcal{L}$ or
- extension $\mathcal{L}_{\mathrm{H}}$ with only one constant H:
  relation with empty body and empty header
  which together with $\oplus$ allows to define relative projection:

  $$(\mathrm{H} \bowtie R_1) \oplus R_2$$

  projection of R_2 to $H(R_1)$, i.e., to the header of R_1

Lemma (Tropashko)

For any fixed domain $\mathcal{D}omain$ *and any set of attributes* $\mathcal{A}ttr$
the family of all relations over $\mathcal{D}omain$
with headers contained in $\mathcal{A}ttr$

forms a lattice with connectives $\oplus$ *and* $\bowtie$

Lemma (Tropashko)

For any fixed domain $\mathcal{Domain}$ and any set of attributes $\mathcal{Attr}$ the family of all relations over $\mathcal{Domain}$ with headers contained in $\mathcal{Attr}$

forms a lattice with connectives $\oplus$ and $\bowtie$

We call such lattices full relational lattices
Their $\mathbb{S}$-closure are concrete lattices
Their $(\mathbb{I})\mathbb{S}\mathbb{P}$-closure are representable lattices

Lemma (Tropashko)

For any fixed domain $\mathcal{D}omain$ and any set of attributes $\mathcal{A}ttr$ the family of all relations over $\mathcal{D}omain$ with headers contained in $\mathcal{A}ttr$

> *forms a lattice with connectives $\oplus$ and $\bowtie$*

We call such lattices full relational lattices
Their $\mathbb{S}$-closure are concrete lattices
Their $(\mathbb{I})\mathbb{S}\mathbb{P}$-closure are representable lattices

This is standard algebraic logic terminology

Lemma (Tropashko)

For any fixed domain $\mathcal{D}omain$ and any set of attributes $\mathcal{A}ttr$ the family of all relations over $\mathcal{D}omain$ with headers contained in $\mathcal{A}ttr$

forms a lattice with connectives $\oplus$ and $\bowtie$

We call such lattices full relational lattices
Their $\mathbb{S}$-closure are concrete lattices
Their $(\mathbb{I})\mathbb{SP}$-closure are representable lattices

This is standard algebraic logic terminology

The full ones will be also called Tropashko lattices
to honour the Founding Father

Proof.

- Define $Dom = \mathcal{A} \cup {}^{\mathcal{A}}\mathcal{D}$ and for any $X \subseteq Dom$ let

$$Cl(X) = X \cup \{x \in {}^{\mathcal{A}}\mathcal{D} \mid \exists y \in (X \cap {}^{\mathcal{A}}\mathcal{D}).\forall a \in (\mathcal{A} - X).x(a) = y(a)\}$$

  In other words, $Cl(X)$ is the sum of $X \cap \mathcal{A}$ (the set of attributes contained in X) with the *cylindrification* of $X \cap {}^{\mathcal{A}}\mathcal{D}$ along the attributes in $X \cap \mathcal{A}$.
  This means attributes in $X \cap \mathcal{A}$ are irrelevant!
- Cl is a closure operator and hence Cl-closed sets form a lattice, with the order being obviously $\subseteq$ inherited from the powerset of Dom.
- $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ is isomorphic to this lattice and the isomorphism is given by

$$(H, B) \mapsto (\mathcal{A} - H) \cup \{x \in {}^{\mathcal{A}}\mathcal{D} \mid x[H] \in B\}.$$

□

We take	$\bowtie$ (natural join)	to be lattice meet	$\wedge$
and	$\oplus$ (inner union)	to be lattice join	$\vee$

We take	$\bowtie$ (natural join)	to be lattice meet	$\wedge$
and	$\oplus$ (inner union)	to be lattice join	$\vee$

$$(H(R), Body(R)) \sqsubseteq (H(S), Body(S))$$

iff

$$H(R) \supseteq H(S) \qquad \& \qquad Body(R) \restriction_{H(S)} \subseteq Body(S)$$

$$F^{\mathcal{A}}_{\mathcal{D}}(H) : \mathcal{P}^{\supseteq}(\mathcal{A}) \ni H \longrightarrow \mathcal{P}({}^{H}\mathcal{D})$$
$$F^{\mathcal{A}}_{\mathcal{D}}(H \supseteq H') = \quad ({}^{H}\mathcal{D} \supseteq B \mapsto B[H'] \subseteq {}^{H'}\mathcal{D})$$

defines a *quasifunctor*

- $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ is an instance of the (covariant) *Grothendieck construction/completion* $\int^{\mathcal{P}^{\supseteq}(\mathcal{A})} F^{\mathcal{A}}_{\mathcal{D}}$
- Note that to preserve the lattice structure we cannot consider $F^{\mathcal{A}}_{\mathcal{D}}$ as a functor into **Set**
- Note also that we chose the covariant definition on $\mathcal{P}^{\supseteq}(\mathcal{A})$ rather than the contravariant definition on $\mathcal{P}(\mathcal{A})$ to ensure the order $\sqsubseteq$ does not get reversed inside each slice $\mathcal{P}({}^{H}\mathcal{D})$

A related recent categorical approach by Samson Abramsky

Relational Databases and Bell Theorem

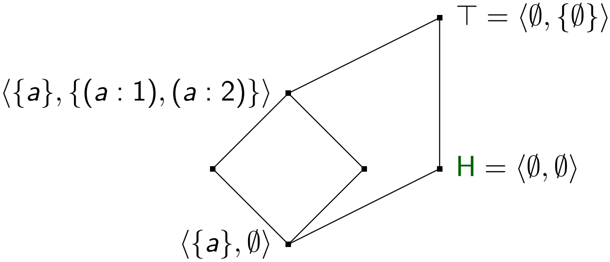
On Equations and Quasiequations

Question

Are the Tropasko lattices always distributive?
Or perhaps all lattices are $\mathbb{HSP}$-images of relational lattices?

NEITHER IS TRUE!

$$\mathcal{A}ttr = \{a\}, \quad \mathcal{D}omain = \{1, 2\}$$



Properties Relational Lattices Do NOT Have

- *semidistributivity* (and hence also *almost distributivity* and *neardistributivity*)

  $SD_\vee$: $a \vee b = a \vee c$ implies $a \vee b = a \vee (b \wedge c)$

- *upper-* or *lower- semimodularity* (and hence also modularity)

  if $a \wedge b$ covers a and b, then $a \vee b$ is covered by a and b

- *local distributivity/local modularity*
- *the Jordan-Dedekind chain condition*

  The cardinalities of two maximal chains between common end points are equal

- *supersolvability*

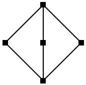
  for finite lattices:

  $\exists$ a maximal chain generating a distributive lattice with any other chain

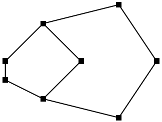
- ...

... But There Are Non-trivial Lattice Identities Valid

(AxRL1) $$x \bowtie y \oplus x \bowtie z = x \bowtie (y \bowtie (x \oplus z) \oplus z \bowtie (x \oplus y)),$$



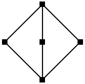
(AxRL2) $$t \bowtie ((x \oplus y) \bowtie (x \oplus z) \oplus (u \oplus w) \bowtie (u \oplus v)) = t \bowtie ((x \oplus y) \bowtie (x \oplus z) \oplus u \oplus w \bowtie v) \oplus t \bowtie ((u \oplus w) \bowtie (u \oplus v) \oplus x \oplus y \bowtie z).$$



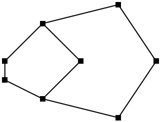
Both equalities are independent!

… But There Are Non-trivial Lattice Identities Valid

(AxRL1) $$x \bowtie y \oplus x \bowtie z = x \bowtie (y \bowtie (x \oplus z) \oplus z \bowtie (x \oplus y)),$$



(AxRL2) $$t \bowtie ((x \oplus y) \bowtie (x \oplus z) \oplus (u \oplus w) \bowtie (u \oplus v)) = t \bowtie ((x \oplus y) \bowtie (x \oplus z) \oplus u \oplus w \bowtie v) \oplus t \bowtie ((u \oplus w) \bowtie (u \oplus v) \oplus x \oplus y \bowtie z).$$



Both equalities are independent!

Lemma (Padmanabhan/McCune/Veroff)

Lattices satisfying

(AxRL1) $$x \bowtie y \oplus x \bowtie z = x \bowtie (y \bowtie (x \oplus z) \oplus z \bowtie (x \oplus y)),$$

satisfy also

($CD_\vee$) $$x \oplus y = x \oplus z \quad \textit{implies} \quad (x \bowtie y) \oplus (x \bowtie z) = x \bowtie (y \oplus z)$$

and this in turn implies the Huntington property.

A class of lattices K has the Huntington property
iff
all *uniquely complemented* lattices from K are distributive

Proposed Axioms for Abstract Relational Lattices

With constant H

all lattice axioms plus

AxRH1	$\mathsf{H} \bowtie x \bowtie (y \oplus z) \oplus y \bowtie z$	$=$	$(\mathsf{H} \bowtie x \bowtie y \oplus z) \bowtie (\mathsf{H} \bowtie x \bowtie z \oplus y)$
AxRH2	$x \bowtie (y \oplus z)$	$=$	$x \bowtie (z \oplus \mathsf{H} \bowtie y) \oplus x \bowtie (y \oplus \mathsf{H} \bowtie z)$

Without H

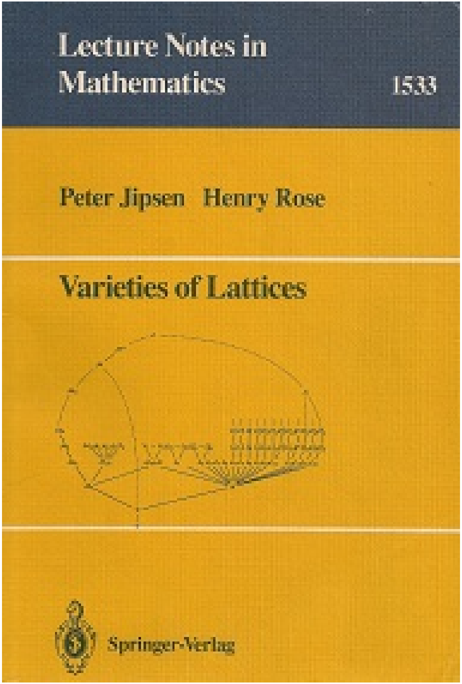
all lattice axioms plus

AxRL1	$x \bowtie y \oplus x \bowtie z$	$=$	$x \bowtie (y \bowtie (x \oplus z) \oplus z \bowtie (x \oplus y))$
AxRL2	$t \bowtie ((x \oplus y) \bowtie (x \oplus z) \oplus (u \oplus w) \bowtie (u \oplus v))$	$=$	$t \bowtie ((x \oplus y) \bowtie (x \oplus z) \oplus u \oplus w \bowtie v) \oplus$ $\oplus\, t \bowtie ((u \oplus w) \bowtie (u \oplus v) \oplus x \oplus y \bowtie z)$

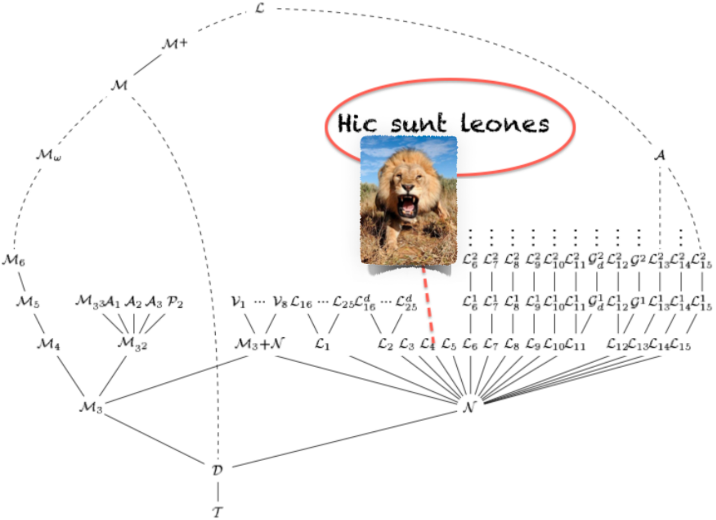
(derivable in the full signature)

Additional (quasi-)equations derivable in $\underline{R}^{\mathbf{H}}$ and $\underline{R}$:

$$
\begin{array}{lrcl}
\text{Qu1} & x \oplus y = x \oplus z & \Rightarrow & x \bowtie (y \oplus z) = x \bowtie y \oplus x \bowtie z. \\
\text{Qu2} & \mathsf{H} \bowtie (x \oplus y) = \mathsf{H} \bowtie (x \oplus z) & \Rightarrow & x \bowtie (y \oplus z) = x \bowtie y \oplus x \bowtie z. \\
\text{Eq1} & \mathsf{H} \bowtie x \bowtie (y \oplus z) & = & \mathsf{H} \bowtie x \bowtie y \oplus \mathsf{H} \bowtie x \bowtie z \\
\text{Der1} & \mathsf{H} \bowtie x \oplus x \bowtie y & = & x \bowtie (y \oplus \mathsf{H} \bowtie x)
\end{array}
$$


Lecture Notes in Mathematics
1533
Peter Jipsen Henry Rose
Varieties of Lattices
Springer-Verlag

Jipsen & Rose: The Lattice of Varieties of Lattices



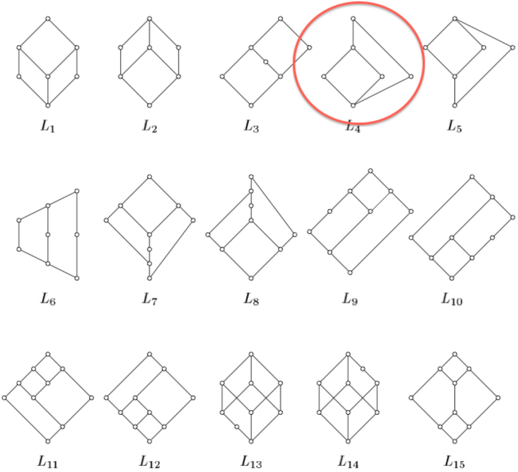


Figure 2.2

As you see, the equational theory is intriguing

As you see, the equational theory is intriguing

Nevertheless, there are are both database
and algebraic reasons to be
even more interested in the quasi-equational theory

The Algebraic Reason for Importance of Quasi-Equations

Theorem

The class of relational lattices is pseudoelementary

The Algebraic Reason for Importance of Quasi-Equations

Theorem

The class of relational lattices is pseudoelementary

hence closed under ultraproducts

The Algebraic Reason for Importance of Quasi-Equations

Theorem

The class of relational lattices is pseudoelementary

hence closed under ultraproducts

hence its $\mathbb{SP}$-closure is a quasivariety

The Algebraic Reason for Importance of Quasi-Equations

Theorem

The class of relational lattices is pseudoelementary

hence closed under ultraproducts

hence its $\mathbb{SP}$-closure is a quasivariety

Corollary

The quasi-equational theory of relational lattices is

recursively enumerable

The Algebraic Reason for Importance of Quasi-Equations

Theorem

The class of relational lattices is pseudoelementary

hence closed under ultraproducts

hence its $\mathbb{SP}$-closure is a quasivariety

Corollary

The quasi-equational theory of relational lattices is

recursively enumerable

Proof.

The elementary axiomatization in extended multi-sorted language used in the pseudoelementarity proof is finite □

Proof.

- Recall $Dom = \mathcal{A} \cup {}^{\mathcal{A}}\mathcal{D}$ and for any $X \subseteq Dom$ let

$$Cl(X) = X \cup \{x \in {}^{\mathcal{A}}\mathcal{D} \mid \exists y \in (X \cap {}^{\mathcal{A}}\mathcal{D}).\forall a \in (\mathcal{A} - X).x(a) = y(a)\}$$

  In other words, $Cl(X)$ is the sum of $X \cap \mathcal{A}$ (the set of attributes contained in X) with the *cylindrification* of $X \cap {}^{\mathcal{A}}\mathcal{D}$ along the axes in $X \cap \mathcal{A}$.
- Take sorts $A = \mathcal{A}$, $F = {}^{\mathcal{A}}\mathcal{D}$, $D = \mathcal{D}$ and R for Cl-closed subsets of Dom
- $\bowtie, \oplus : R \times R \to R$, $\mathsf{H} : R$
- $assign : F \times A \mapsto D$ (value of F on A)
- $inR \subseteq (A \cup F) \times R$ (membership of attribute/sequence in a closed subset)

(to be continued) □

Proof.

(continued) FO axioms forcing the correctness of this interpretation:

- extensionality for F and R (via axioms on inR),
- each element of R is closed/cylindrified
- $\bowtie$ and $\oplus$ are, respectively, genuine infimum and supremum operations on R.
- inR assigns no elements of R and all elements of A to H.

□

Aside: Automatization and Proof Assistants

- The above proof shows how an encoding into a first-order theory in a richer language looks like
- We can (and should ...) use it in a proof assistant
- So far, only used equational theorem provers/countermodel finders (Prover9/Mace4) to investigate the equational theory
- It is worth mentioning though that the inventor of relational lattices, Vadim Tropashko, has developed in the meantime a dedicated tool QBQL: `https://code.google.com/p/qbql/`

The Database Reason for Importance of Quasi-Equations

Reasoning over database constraints

(key constraints, foreign keys)

can be reduced to quasi-equational reasoning.

Also relevant for rule-based query optimization

The Database Reason for Importance of Quasi-Equations

Reasoning over database constraints

(key constraints, foreign keys)

can be reduced to quasi-equational reasoning.

Also relevant for rule-based query optimization

Note: in order to formulate
simplest constraints or just type infomation:

"R_1 and R_2 have the same header"

it is necessary to add H:

$$\mathrm{H} \bowtie R_1 = \mathrm{H} \bowtie R_2$$

Theorem 4.7

Unfortunately, the addition of identity AxRH1 to lattice axioms:

$$\mathsf{H} \bowtie x \bowtie (y \oplus z) \oplus y \bowtie z = (\mathsf{H} \bowtie x \bowtie y \oplus z) \bowtie (\mathsf{H} \bowtie x \bowtie z \oplus y)$$

allows to imitate Maddux' technique for CA_3 and

embed the word problem for semigroups

into quasi-equational theory . . .

Corollary (4.8)

Any quasiequational theory contained between

- *the eq. theory of lattices with AxRH1 and*
- *the quasiequational theory of finite relational lattices*

is undecidable

Corollary (4.8)

Any quasiequational theory contained between

- *the eq. theory of lattices with AxRH1 and*
- *the quasiequational theory of finite relational lattices*

is undecidable

Corollary (4.9)

The quasiequational theory of finite relational lattices

is not finitely axiomatizable

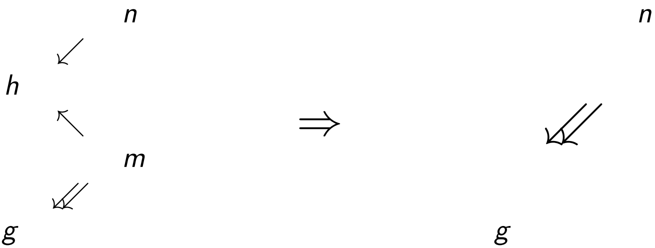
Undecidability knocks on our door earlier than expected . . .

Finite Duality and Formal Concept Analysis

- Let $\mathcal{L}$ be a finite lattice
- $\mathfrak{J}(\mathcal{L})$ the set of its join-irreducibles and $\mathfrak{M}(\mathcal{L})$ the set of its meet-irreducibles
- The standard context $\mathrm{con}(\mathcal{L}) := (\mathfrak{J}(\mathcal{L}), \mathfrak{M}(\mathcal{L}), \mathrm{I}_\leq)$, where $\mathrm{I}_\leq := \leq \cap (\mathfrak{J}(\mathcal{L}) \times \mathfrak{M}(\mathcal{L}))$
- Define:

$$
\begin{aligned}
g \swarrow m &: \quad g \text{ is } \leq\text{-minimal in } \{h \in \mathfrak{J}(\mathcal{L}) \mid \text{NOT } h\,\mathrm{I}_\leq\, m\} \\
g \nwarrow m &: \quad m \text{ is } \leq\text{-maximal in } \{n \in \mathfrak{M}(\mathcal{L}) \mid \text{NOT } g\,\mathrm{I}_\leq\, n\} \\
g \swarrow\!\!\!\!\nwarrow m &: \quad g \swarrow m \;\&\; g \nwarrow m \quad \text{(modified notation!)}
\end{aligned}
$$

- $\swarrow\!\!\!\swarrow$ is the smallest relation containing $\swarrow$ and closed s.t. $g \swarrow\!\!\!\swarrow m$, $h \nwarrow m$ and $h \swarrow n$ imply $g \swarrow\!\!\!\swarrow n$


n
h
m
g
n
g

Theorem (Ganter, Wille)

A finite lattice is

- *subdirrectly irreducible iff there is* $m \in \mathfrak{M}(\mathcal{L})$ *s.t.* $\swarrow\!\!\!\swarrow \supseteq \mathfrak{J}(\mathcal{L}) \times \{m\}$
- *simple iff* $\swarrow\!\!\!\swarrow = \mathfrak{J}(\mathcal{L}) \times \mathfrak{M}(\mathcal{L})$

Hence, in order to investigate the structure of finite Tropashko lattices, it would be useful to have a description of $\mathfrak{J}(\mathfrak{R}(\mathcal{D}, \mathcal{A}))$ and $\mathfrak{M}(\mathfrak{R}(\mathcal{D}, \mathcal{A}))$ ...

$$
\begin{array}{llll}
\mathcal{AD}om_{\mathcal{D},\mathcal{A}} := \{\mathrm{adom}(x) \mid x \in {}^{\mathcal{A}}\mathcal{D}\} & \text{where} & \mathrm{adom}(x) & := (\mathcal{A}, \{x\}) \\
\mathcal{AA}tt_{\mathcal{D},\mathcal{A}} := \{\mathrm{aatt}(a) \mid a \in \mathcal{A}\} & \text{where} & \mathrm{aatt}(a) & := (\mathcal{A} - \{a\}, \emptyset) \\
\mathcal{C}oDom_{\mathcal{D},H} := \{\mathrm{codom}^H(x) \mid x \in {}^{H}\mathcal{D}\} & \text{where} & \mathrm{codom}^H(x) & := (H, {}^{H}\mathcal{D} - \{x\}) \\
\mathcal{C}oAtt_{\mathcal{D},\mathcal{A}} := \{\mathrm{coatt}(a) \mid a \in \mathcal{A}\} & \text{where} & \mathrm{coatt}(a) & := (\{a\}, {}^{\{a\}}\mathcal{D})
\end{array}
$$

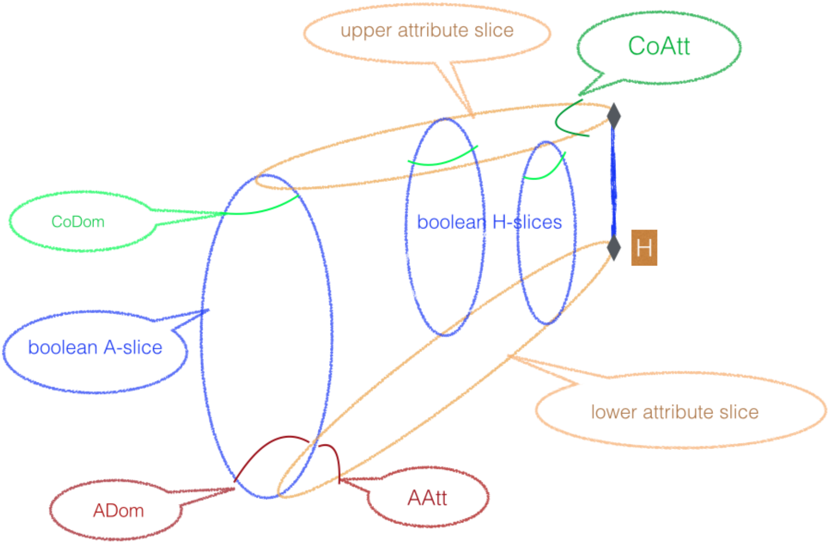
$$
\begin{array}{lll}
\mathcal{J}_{\mathcal{D},\mathcal{A}} & := & \mathcal{AD}om_{\mathcal{D},\mathcal{A}} \cup \mathcal{AA}tt_{\mathcal{D},\mathcal{A}} \\
\mathcal{M}_{\mathcal{D},\mathcal{A}} & := & \mathcal{C}oAtt_{\mathcal{D},\mathcal{A}} \cup \bigcup_{H \subseteq \mathcal{A}} \mathcal{C}oDom_{\mathcal{D},H}
\end{array}
$$

Theorem (5.2)

For any finite $\mathcal{A}$ and $\mathcal{D}$ such that $|\mathcal{D}| \geq 2$, we have

$$\mathcal{J}_{\mathcal{D},\mathcal{A}} = \mathfrak{J}(\mathfrak{R}(\mathcal{D},\mathcal{A})) \qquad \textit{(join-irreducibles)}$$

$$\mathcal{M}_{\mathcal{D},\mathcal{A}} = \mathfrak{M}(\mathfrak{R}(\mathcal{D},\mathcal{A})) \qquad \textit{(meet-irreducibles)}$$


upper attribute slice
CoAtt
CoDom
boolean H-slices
H
boolean A-slice
lower attribute slice
ADom
AAtt

Theorem (5.3)

Assume $\mathcal{D}, \mathcal{A}$ are finite sets s.t. $|\mathcal{D}| \geq 2$ and $\mathcal{A} \neq \emptyset$. Then $\mathbf{I}_{\leq}$, $\swarrow$, $\nwarrow$ and $\swarrow\!\!\!\swarrow$ look for $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ as follows:

$r =$ $s =$	$\mathrm{adom}(x)$ $\mathrm{coatt}(a)$	$\mathrm{aatt}(a)$ $\mathrm{coatt}(b)$	$\mathrm{adom}(x)$ $\mathrm{codom}^H(y)$	$\mathrm{aatt}(a)$ $\mathrm{codom}^H(y)$
$r \mathbf{I}_{\leq} s$	*always*	$a \neq b$	$x[H] \neq y$	$a \notin H$
$r \swarrow s$	*never*	$a = b$	$x[H] = y$	$a \in H$
$r \nwarrow s$	*never*	$a = b$	$x[H] = y$	*never*
$r \swarrow\!\!\!\swarrow s$	*never*	$a = b$	*always*	*always*

Corollary (5.4)

Whenever $\mathcal{D}, \mathcal{A}$ are finite sets s.t. $|\mathcal{D}| \geq 2$ and $\mathcal{A} \neq \emptyset$, then

- $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ *is subdirectly irreducible*
- $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ *is not simple*

Recap

- We study an interpretation of lattice connectives proposed by Vadim Tropashko from Oracle: as natural join and inner union between DB relations
- We show that this interpretation yields a class of lattices which has not been considered in the existing lattice-theoretical literature
- We propose an equational axiomatization for a corresponding abstract algebraic class
- It turns out that addition of just the *header constant* to the lattice signature leads to undecidabilty of the quasiequational theory
- Relational lattices, however, are not as intangible as one may fear: for example, they do form a pseudoelementary class
- We also apply the tools of Formal Concept Analysis and investigate standard contexts of relational lattices. In particular, we'll see that finite relational lattices, while not "bounded" in the McKenzie sense, are subdirectly irreducible

Recap

- ✓ We study an interpretation of lattice connectives proposed by Vadim Tropashko from Oracle: as natural join and inner union between DB relations
- ? We show that this interpretation yields a class of lattices which has not been considered in the existing lattice-theoretical literature
- ? We propose an equational axiomatization for a corresponding abstract algebraic class
- ? It turns out that addition of just the *header constant* to the lattice signature leads to undecidabilty of the quasiequational theory
- ? Relational lattices, however, are not as intangible as one may fear: for example, they do form a pseudoelementary class
- ? We also apply the tools of Formal Concept Analysis and investigate standard contexts of relational lattices. In particular, we'll see that finite relational lattices, while not "bounded" in the McKenzie sense, are subdirectly irreducible

Recap

- ✓ We study an interpretation of lattice connectives proposed by Vadim Tropashko from Oracle: as natural join and inner union between DB relations
- ✓ We show that this interpretation yields a class of lattices which has not been considered in the existing lattice-theoretical literature
- ? We propose an equational axiomatization for a corresponding abstract algebraic class
- ? It turns out that addition of just the *header constant* to the lattice signature leads to undecidabilty of the quasiequational theory
- ? Relational lattices, however, are not as intangible as one may fear: for example, they do form a pseudoelementary class
- ? We also apply the tools of Formal Concept Analysis and investigate standard contexts of relational lattices. In particular, we'll see that finite relational lattices, while not "bounded" in the McKenzie sense, are subdirectly irreducible

Recap

- ✓ We study an interpretation of lattice connectives proposed by Vadim Tropashko from Oracle: as natural join and inner union between DB relations
- ✓ We show that this interpretation yields a class of lattices which has not been considered in the existing lattice-theoretical literature
- ✓ We propose an equational axiomatization for a corresponding abstract algebraic class
- ? It turns out that addition of just the *header constant* to the lattice signature leads to undecidabilty of the quasiequational theory
- ? Relational lattices, however, are not as intangible as one may fear: for example, they do form a pseudoelementary class
- ? We also apply the tools of Formal Concept Analysis and investigate standard contexts of relational lattices. In particular, we'll see that finite relational lattices, while not "bounded" in the McKenzie sense, are subdirectly irreducible

Recap

- ✓ We study an interpretation of lattice connectives proposed by Vadim Tropashko from Oracle: as natural join and inner union between DB relations
- ✓ We show that this interpretation yields a class of lattices which has not been considered in the existing lattice-theoretical literature
- ✓ We propose an equational axiomatization for a corresponding abstract algebraic class
- ✓ It turns out that addition of just the *header constant* to the lattice signature leads to undecidabilty of the quasiequational theory
- **?** Relational lattices, however, are not as intangible as one may fear: for example, they do form a pseudoelementary class
- **?** We also apply the tools of Formal Concept Analysis and investigate standard contexts of relational lattices. In particular, we'll see that finite relational lattices, while not "bounded" in the McKenzie sense, are subdirectly irreducible

Recap

- ✓ We study an interpretation of lattice connectives proposed by Vadim Tropashko from Oracle: as natural join and inner union between DB relations
- ✓ We show that this interpretation yields a class of lattices which has not been considered in the existing lattice-theoretical literature
- ✓ We propose an equational axiomatization for a corresponding abstract algebraic class
- ✓ It turns out that addition of just the *header constant* to the lattice signature leads to undecidabilty of the quasiequational theory
- ✓ Relational lattices, however, are not as intangible as one may fear: for example, they do form a pseudoelementary class
- ? We also apply the tools of Formal Concept Analysis and investigate standard contexts of relational lattices. In particular, we'll see that finite relational lattices, while not "bounded" in the McKenzie sense, are subdirectly irreducible

Recap

- ✓ We study an interpretation of lattice connectives proposed by Vadim Tropashko from Oracle: as natural join and inner union between DB relations
- ✓ We show that this interpretation yields a class of lattices which has not been considered in the existing lattice-theoretical literature
- ✓ We propose an equational axiomatization for a corresponding abstract algebraic class
- ✓ It turns out that addition of just the *header constant* to the lattice signature leads to undecidabilty of the quasiequational theory
- ✓ Relational lattices, however, are not as intangible as one may fear: for example, they do form a pseudoelementary class
- ✓ We also apply the tools of Formal Concept Analysis and investigate standard contexts of relational lattices. In particular, we'll see that finite relational lattices, while not "bounded" in the McKenzie sense, are subdirectly irreducible

List of questions

- Is quasi-equational theory of arbitrary relational lattices axiomatizable?
- How about decidability/axiomatizability of quasi-equational and equational theories of lattice reducts (i.e., without H)?
- Prove that representable relational lattices

  do not form an equational class (or do they?)
- Investigate the connections with
  - boolean algebras of finite sequences

    (Craig and Quine/Kuhn)
  - multi-sorted cylindric algebras

    (Bernays/Schwartz/Börner)
- (Venema) Duality theory?
  - more or less done for full relational lattices via our FCA
  - generalize to: concrete, representable, abstract ones
- (Hirsch) Representability? Also should use the FCA results . . .