Relational Lattices

Tadeusz Litak[1]*, Szabolcs Mikulás[2], and Jan Hidders[3]

[1] Informatik 8, Friedrich-Alexander-Universität Erlangen-Nürnberg
Martensstraße 3, 91058 Erlangen, Germany
tadeusz.litak@gmail.com
[2] School of Computer Science and Information Systems,
Birkbeck, University of London, WC1E 7HX London, UK
szabolcs@dcs.bbk.ac.uk
[3] Delft University of Technology, Elektrotechn., Wisk. and Inform.,
Mekelweg 4, 2628CD Delft, The Netherlands
A.J.H.Hidders@tudelft.nl

Abstract. *Relational lattices* are obtained by interpreting lattice connectives as *natural join* and *inner union* between database relations. Our study of their equational theory reveals that the variety generated by relational lattices has not been discussed in the existing literature. Furthermore, we show that addition of just the *header constant* to the lattice signature leads to undecidability of the quasiequational theory. Nevertheless, we also demonstrate that relational lattices 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 the structure of relational lattices via their standard contexts.

Keywords: relational lattices, relational algebra, database theory, algebraic logic, lattice theory, cylindric algebras, Formal Concept Analysis

1 Introduction

We study a class of lattices with a natural database interpretation [Tro, ST06, Tro05]. It does not seem to have attracted the attention of algebraists, even those investigating the connections between algebraic logic and relational databases (see, e.g., [IL84] or [DM01]).

The connective *natural join* (which we will interpret as lattice meet!) is one of the basic operations of Codd's *(named) relational algebra* [AHV95, Cod70].

* We would like to thank *Vadim Tropashko* and *Marshall Spight* for introducing the subject to the third author (who in turn introduced it to the other two) and discussing it in the usenet group `comp.databases.theory`, *Maarten Marx*, *Balder ten Cate*, *Jan Paredaens* for additional discussions and general support in an early phase of our cooperation and the *referees* for the comments. The first author would also like to acknowledge: a) *Peter Jipsen* for discussions in September 2013 at the Chapman University leading to recovery, rewrite and extension of the material (in particular for Sec. 5) and b) suggestions by participants of: TACL'09, ALCOP 2010 and the Birmingham TCS seminar (in particular for Sec. 2.1 and 6).

Incidentally, it is also one of its few genuine algebraic operations—i.e., defined for all arguments. Codd's "algebra", from a mathematical point of view, is only a *partial algebra*: some operations are defined only between relations with suitable headers, e.g., (set) union or the difference operator. Apart from the issues of mathematical elegance and generality, this partial nature of operations has also unpleasant practical consequences. For example, queries which do not observe constraints on headers can *crash* [VdBVGV07].

It turns out, however, that it is possible to generalize the union operation to *inner union* defined on all elements of the algebra and lattice-dual to natural join. This approach appears more natural and has several advantages over the embedding of relational "algebras" in cylindric algebras proposed in [IL84]. For example, we avoid an artificial uniformization of headers and hence queries formed with the use of proposed connectives enjoy the *domain independence property* (see, e.g., [AHV95, Ch. 5] for a discussion of its importance in databases).

We focus here on the (quasi)equational theory of natural join and inner union. Apart from an obvious mathematical interest, Birkhoff-style equational inference is the basis for certain query optimization techniques where algebraic expressions represent query evaluation plans and are rewritten by the optimizer into equivalent but more efficient expressions. As for *quasiequations*, i.e., definite Horn clauses over equalities, reasoning over many database constraints such as key constraints and foreign keys can be reduced to quasiequational reasoning. Note that an optimizer can consider more equivalent alternatives for the original expression if it can take the specified database constraints into account.

Strikingly, it turned out that relational lattices does not seem to fit anywhere into the rather well-investigated landscape of equational theories of lattices [JR92, JR98]. Nevertheless, there were some indications that the considered choice of connectives may lead to positive results concerning decidability/axiomatizability even for quasiequational theories. There is an elegant procedure known as *the chase* [AHV95, Ch. 8] applicable for certain classes of queries and database constraints similar to those that can be expressed with the natural join and inner union.

To our surprise, however, it turned out that when it comes to decidability, relational lattices seem to have a lot in common with other "untamed" structures from algebraic logic such as Tarski's relation algebras or cylindric algebras. As soon as an additional *header constant* H is added to the language, one can encode the word problem for semigroups in the quasiequational theory using a technique introduced by Maddux [Mad80]. This means that decidability of query equivalence under constraints for restricted positive database languages does not translate into decidability of corresponding quasiequational theories. However, our Theorem 4.5 and Corollary 4.6 do not rule out possible finite axiomatization results (except for quasiequational theory of *finite* structures) or decidability of equational theory.[4] And with H removed, i.e., in the pure lattice signature, the picture is completely open. Of course, such a language would be rather weak from a database point of view, but natural for an algebraist.

[4] Note, however, that an extension of our signature to a language with EDPC or a discriminator term would result in an undecidable *equational* theory.

We also obtained a number of positive results. First of all, representable relational lattices are pseudoelementary and hence their closure under subalgebras and products is a quasivariety—Theorem 4.1 and Corollary 4.2. The proof of pseudoelementarity shows how to encode the theory of concrete relational lattices in a sufficiently rich (many-sorted) first-order theory. This opens up the possibility of using generic proof assistants like Isabelle or Coq in future investigations—so far, we have only used Prover9/Mace4 to study interderivability of interesting (quasi)equations.[5] We have also used the tools of Formal Concept Analysis (Theorem 5.3) to investigate the dual structure of full concrete relational lattices and establish, e.g., their subdirect irreducibility (Corollary 5.4). Theorem 5.3 is likely to have further applications—see the discussion of Problem 6.1.

The structure of the paper is as follows. In Section 2, we provide basic definitions, establish that relational lattices are indeed lattices and note in passing a potential connection with category theory in Section 2.1. Section 3 reports our findings about the (quasi)equational theory of relational lattices: the failure of most standard properties such as weakening of distributivity (Theorem 3.2), those surprising equations and properties that still hold (Theorem 3.4) and dependencies between them (Theorem 3.5). In Section 4, we focus on quasiequations and prove some of most interesting results discussed above, both positive (Theorem 4.1 and Corollary 4.2) and negative ones (Theorem 4.5 and Corollary 4.6). Section 5 applies Formal Concept Analysis to relational lattices. Section 6 concludes and discusses future work, in particular possible extensions of the signature in Section 6.1.

2 Basic Definitions

Let $\mathcal{A}$ be a set of *attribute names* and $\mathcal{D}$ be a set of *domain values*. For $H \subseteq \mathcal{A}$, a *H-sequence from $\mathcal{D}$* or an *H-tuple over $\mathcal{D}$* is a function $x : H \to \mathcal{D}$, i.e., an element of ${}^{H}\mathcal{D}$. H is called the *header* of x and denoted as $h(x)$. The *restriction of x to H'* is denoted as $x[H']$ and defined as $x[H'] := \{(a, v) \in x \mid a \in H'\}$, in particular $x[H'] = \emptyset$ if $H' \cap h(x) = \emptyset$. We generalize this to the *projection of a set of H-sequences X to a header H'* which is $X[H'] := \{x[H'] \mid x \in X\}$. A *relation* is a pair $r = (H_r, B_r)$, where $H_r \subseteq \mathcal{A}$ is the *header* of r and $B_r \subseteq {}^{H_r}\mathcal{D}$ the *body* of r. The collections of all relations over $\mathcal{D}$ whose headers are contained in $\mathcal{A}$ will be denoted as $R(\mathcal{D}, \mathcal{A})$. For the relations r, s, we define the *natural join* $r \bowtie s$, and *inner union* $r \oplus s$:

$$r \bowtie s := (H_r \cup H_s, \{x \in {}^{H_r \cup H_s}\mathcal{D} \mid x[H_r] \in B_r \text{ and } x[H_s] \in B_s\})$$
$$r \oplus s := (H_r \cap H_s, \{x \in {}^{H_r \cap H_s}\mathcal{D} \mid x \in B_r[H_s] \text{ or } x \in B_s[H_r]\})$$

In our notation, $\bowtie$ always binds stronger than $\oplus$. The *header constant* $\mathsf{H} := (\emptyset, \emptyset)$ plays a special role, as $(H, B) \bowtie \mathsf{H} = (H, \emptyset)$. Hence, r_1 and r_2 have the same headers iff $\mathsf{H} \bowtie r_1 = \mathsf{H} \bowtie r_2$. Note also that the projection of r_1 to H_{r_2} can be defined

[5] It is worth mentioning that the database inventor of relational lattices has in the meantime developed a dedicated tool [Tro].

a	b
1	1
2	2
3	2
3	3

$\bowtie$

b	c
1	1
2	2
2	3
4	4

$=$

a	b	c
1	1	1
2	2	2
2	2	3
3	2	2
3	2	3

a	b
1	1
2	2
3	2
3	3

$\oplus$

b	c
1	1
2	2
2	3
4	4

$=$

b
1
2
3
4

Fig. 1. Natural join and inner union. In this example, $\mathcal{A} = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}, D = \{1, 2, 3, 4\}$.

as $r_1 \oplus (\mathsf{H} \bowtie r_2)$. In fact, we can identify $\mathsf{H} \bowtie r$ and H_r. We denote $(R(\mathcal{D}, \mathcal{A}), \bowtie, \oplus, \mathsf{H})$ as $\mathfrak{R}^{\mathsf{H}}(\mathcal{D}, \mathcal{A})$, with $\mathcal{L}_{\mathsf{H}}$ denoting the corresponding algebraic signature. $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ is its reduct to the signature $\mathcal{L} = \{\bowtie, \oplus\}$.

Lemma 2.1. *For any $\mathcal{D}$ and $\mathcal{A}$, $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ is a lattice.*

Proof. This result is due to Tropashko [Tro, ST06, Tro05], but let us provide an alternative proof. Define $Dom = \mathcal{A} \cup {}^{\mathcal{A}}\mathcal{D}$ and for any $X \subseteq Dom$ set

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

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 $X \cap \mathcal{A}$. It is straightforward to verify 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. It remains to observe $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ is isomorphic to this lattice and the isomorphism is given by

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

$\square$

The lattice order given by these operations is

$$(H_r, B_r) \sqsubseteq (H_s, B_s) \text{ iff } H_s \subseteq H_r \text{ and } B_r[H_s] \subseteq B_s.$$

Therefore, we call $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ *the (full) relational lattice over ($\mathcal{D}$,$\mathcal{A}$)* and $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ its *lattice reduct*. We also use the alternative name *Tropashko lattices* to honor the inventor of these structures.

For classes of algebras, we use $\mathbb{H}, \mathbb{S}, \mathbb{P}$ to denote closures under, respectively, homomorphisms, (isomorphic copies of) subalgebras and products. Let $\mathcal{R}^{\mathsf{H}}_{\text{fin}} := \mathbb{S}\{\mathfrak{R}(\mathcal{D}, \mathcal{A}, \mathsf{H}) \mid \mathcal{D}, \mathcal{A} \text{ finite}\}$, $\mathcal{R}^{\mathsf{H}}_{\text{unr}} := \mathbb{S}\{\mathfrak{R}(\mathcal{D}, \mathcal{A}, \mathsf{H}) \mid \mathcal{D}, \mathcal{A} \text{ unrestricted}\}$ and let $\mathcal{R}_{\text{fin}}$ and $\mathcal{R}_{\text{unr}}$ denote the $\mathcal{L}$-reducts of respective classes.

2.1 Relational Lattice as the Grothendieck Construction

Given $\mathcal{D}$ and $\mathcal{A}$, a category theorist may note that

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

defines a *quasifunctor* assigning to an element of the powerset $\mathcal{P}^{\supseteq}(\mathcal{A})$ (considered as a poset with reverse inclusion order) the poset $\mathcal{P}({}^{H}\mathcal{D})$ considered as a small category. Then one readily notes that $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ is an instance of what is known as the (covariant) *Grothendieck construction/completion*[6] of $F^{\mathcal{A}}_{\mathcal{D}}$ [Jac99, Definition 1.10.1] denoted as $\int^{\mathcal{P}^{\supseteq}(\mathcal{A})} F^{\mathcal{A}}_{\mathcal{D}}$. As such considerations are irrelevant for the rest of our paper, for the time being we just note this category-theoretical connection as a curiosity, but it might lead to an interesting future study.

3 Towards the Equational Theory of Relational Lattices

Let us begin the section with an open

Problem 3.1. Are $\mathbb{SP}(\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}) = \mathbb{HSP}(\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}})$ and $\mathbb{SP}(\mathcal{R}_{\mathrm{unr}}) = \mathbb{HSP}(\mathcal{R}_{\mathrm{unr}})$?

If the answer is "no", it would mean that relational lattices should be considered a quasiequational rather than equational class (cf. Corollary 4.2 below). Note also that the decidability of equational theories seems of less importance from a database-theoretical point of view than decidability of quasiequational theories. Nevertheless, relating to already investigated varieties of lattices seems a good first step. It turns out that weak forms of distributivity and similar properties (see [JR92, JR98, Ste99]) tend to fail dramatically:

Theorem 3.2. $\mathcal{R}_{\mathrm{fin}}$ *(and hence* $\mathcal{R}_{\mathrm{unr}}$*) does not have any of the following properties (see the above references or the proof below for definitions):*

1. *upper- and lower-semidistributivity,*
2. *almost distributivity and neardistributivity,*
3. *upper- or lower-semimodularity (and hence also modularity),*
4. *local distributivity/local modularity,*
5. *the Jordan–Dedekind chain condition,*
6. *supersolvability.*

Proof. For most clauses, it is enough to observe that $\mathcal{R}(\{0,1\}, \{0\}))$ is isomorphic to L_4, one of the covers of the non-modular lattice N_5 in [McK72] (see also [JR98]): a routine counterexample in such cases. In more detail:
Clause 1: Recall that *semidistributivity* is the property
$SD_{\oplus}$: $a \oplus b = a \oplus c$ implies $a \oplus b = a \oplus (b \bowtie c)$.
Now take a to be H and b and c to be the atoms with the header $\{0\}$.
Clause 2: This is a corollary of Clause 1, see [JR92, Th 4.2 and Sec 4.3].
Clause 3: Recall that *semimodularity* is the property
if $a \bowtie b$ covers a and b, then $a \oplus b$ is covered by a and b
Again, take a to be H and b to be either of the atoms with the header $\{0\}$.
Clause 4: This is a corollary of Clause 3, see [Mae74].

[6] Note that to preserve the lattice structure of $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ we *cannot* consider $F^{\mathcal{A}}_{\mathcal{D}}$ as a functor into **Set**, which would yield a special case of the Grothendieck construction known as the *category of elements*. 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 $\mathcal{P}({}^{H}\mathcal{D})$.

Clause 5: Recall that *the Jordan-Dedekind chain condition* is the property that the cardinalities of two maximal chains between common end points are equal. This obviously fails in N_5.
Clause 6: Recall that for finite lattices, *supersolvability* [Sta72] boils down to the existence of a maximal chain generating a distributive lattice with any other chain. Again, this fails in N_5.

□

Remark 3.3. Theorem 3.2 has an additional consequence regarding the notion called by some lattice theorists rather misleadingly *boundedness* (see e.g., [JR92, p. 27]): being an image of a freely generated lattice by a *bounded morphism.* We use the term *McKenzie-bounded*, as McKenzie showed that for finite subdirectly irreducible lattices, this property amounts to splitting the lattice of varieties of lattices [JR92, Theorem 2.25]. We will see that finite full relational lattices are subdirectly irreducible (Corollary 5.4 below) but already the first item of Theorem 3.2 means they are not McKenzie-bounded by [JR92, Lemma 2.30].

Nevertheless, Tropashko lattices do not generate the variety of all lattices. The results of our investigations so far on valid (quasi)equations are summarized by the following theorems:

Theorem 3.4. *Axioms of* $\underline{R}^{\mathsf{H}}$ *in Table 1 are valid in* $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ *(and consequently in* $\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}}$*). Similarly, axioms of* $\underline{R}$ *are valid in* $\mathcal{R}_{\mathrm{unr}}$ *(and consequently* $\mathcal{R}_{\mathrm{fin}}$*).*

Table 1. (Quasi)equations Valid in Tropashko Lattices

Class $\underline{R}^{\mathsf{H}}$ in the signature $\mathcal{L}_{\mathsf{H}}$:

all lattice axioms

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	$x\bowtie(y\oplus z) = x\bowtie(z\oplus \mathsf{H}\bowtie y)\oplus x\bowtie(y\oplus \mathsf{H}\bowtie z)$
AxRL1	$x\bowtie y\oplus x\bowtie z = x\bowtie(y\bowtie(x\oplus z)\oplus z\bowtie(x\oplus y))$

Class $\underline{R}$ in the signature $\mathcal{L}$ (without H):

all lattice axioms **together with AxRL1** and

AxRL2 $t\bowtie((x\oplus y)\bowtie(x\oplus z)\oplus(u\oplus w)\bowtie(u\oplus v)) =$
$= t\bowtie((x\oplus y)\bowtie(x\oplus z)\oplus u\oplus w\bowtie v)\oplus t\bowtie((u\oplus w)\bowtie(u\oplus v)\oplus x\oplus y\bowtie z)$
(in $\mathcal{L}_{\mathsf{H}}$, AxRL2 is derivable from AxRH1 and AxRH2 above)

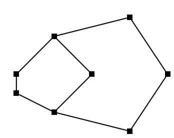
Additional (quasi)equations derivable in $\underline{R}^{\mathsf{H}}$ and $\underline{R}$:

Qu1	$x\oplus y = x\oplus z$	$\Rightarrow$	$x\bowtie(y\oplus z) = x\bowtie y\oplus x\bowtie z.$
Qu2	$\mathsf{H}\bowtie(x\oplus y) = \mathsf{H}\bowtie(x\oplus z)$	$\Rightarrow$	$x\bowtie(y\oplus z) = x\bowtie y\oplus x\bowtie z.$
Eq1	$\mathsf{H}\bowtie x\bowtie(y\oplus z)$	$=$	$\mathsf{H}\bowtie x\bowtie y\oplus \mathsf{H}\bowtie x\bowtie z$
Der1	$\mathsf{H}\bowtie x\oplus x\bowtie y$	$=$	$x\bowtie(y\oplus \mathsf{H}\bowtie x)$

Theorem 3.5. ***Assuming all lattice axioms***, *the following dependencies hold:*

1. *Axioms of* $\underline{R}$ *are mutually independent.*
2. *Each of the axioms of* $\underline{R}^{\mathsf{H}}$ *is independent from the remaining ones, with a possible exceptions of AxRL1.*
3. *[PMV07] AxRL1 forces Qu1.*
4. *Qu2 together with Eq1 imply AxRL2.*
5. *Eq1 is implied by AxRH1. The converse implication does not hold even in presence of AxRL1.*
6. *AxRH1 and AxRH2 jointly imply Qu2, although each of the two equations separately is too weak to entail Qu2. In the converse direction, Qu2 implies AxRH2 but not AxRH1.*
7. *AxRH1 implies Der1.*

Proof. Clause 1: The example showing that the validity of AxRL2 does not imply the validity of AxRL1 is the non-distributive diamond lattice M_3, while the reverse implication can be disproved with an eight-element model:



Clause 2: Counterexamples can be obtained by appropriate choices of the interpretation of H in the pentagon lattice.
Clause 4: Direct computation.
Clause 5: The first part has been proved with the help of Prover9 (66 lines of proof). The counterexample for the converse is obtained by choosing H to be the top element of the pentagon lattice.
Clause 6: Prover9 was able to prove the first statement both in presence and in absence of AxRL1, although there was a significant difference in the length of both proofs (38 lines vs. 195 lines). The implication from Qu2 to AxRH2 is straightforward. All the necessary counterexamples can be found by appropriate choices of the interpretation of H in the pentagon lattice.
Clause 7: Substitue x for z and use the absorption law.

□

AxRL1 comes from [PMV07] as an example of an equation which forces *the Huntington property* (distributivity under unique complementation). Qu1 is a form of weak distributivity, denoted as $CD_{\vee}$ in [PMV07] and $WD_{\wedge}$ in [JR98].

Problem 3.6. Are the equational theories of $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ and $\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}}$ equal?

Problem 3.7. Is the equational theory of $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ ($\mathcal{R}_{\mathrm{unr}}$) equal to $\underline{R}^{\mathsf{H}}$ ($\underline{R}$, respectively)? If not, is it finitely axiomatizable at all?

If the answer to the last question is in the negative, one can perhaps attempt a rainbow-style argument from algebraic logic [HH02].

4 Relational Lattices as a Quasiequational Class

We have already hinted at the database-theoretic reasons why a quasiequational axiomatization would be of even more interest than an equational one. Now it is time for an algebraic reason: the class of representable Tropashko lattices (i.e., the $\mathbb{SP}$-closure of $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ or $\mathcal{R}_{\mathrm{unr}}$) is a quasivariety.

Theorem 4.1. *$\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ and $\mathcal{R}_{\mathrm{unr}}$ are pseudoelementary classes and hence are closed under ultraproducts.*

Proof. (sketch) Assume a multi-sorted language extending $\mathcal{L}_{\mathsf{H}}$ with sorts A, F, D and R, the last used to interpret all connectives of $\mathcal{L}_{\mathsf{H}}$. In addition, we assume a relation $inR \subseteq (A \cup F) \times R$ and a function $assign : (F \times A) \mapsto D$. The interpretation of these sorts, relations and functions is as suggested by the closure system used in the proof of Lemma 2.1. That is, A corresponds to $\mathcal{A}$, F corresponds to ${}^{\mathcal{A}}\mathcal{D}$, D corresponds to $\mathcal{D}$ and R—to the family of Cl-closed subsets of Dom. Moreover, $assign(f, a)$ represents the value assigned by the $\mathcal{A}$-sequence denoted by f to attribute denoted by a and $inR(x, r)$ (for x of either A or F sort)—the membership in the subset of Dom denoted by r. Then the following axioms force the correctness of this interpretation: extensionality for F (i.e., injectivity of $assign$) and R (via axioms on inR), the axiom forcing that each element of R is equal to its own closure as specified in the proof of Lemma 2.1 and an axiom forcing that $\bowtie$ and $\oplus$ are, respectively, genuine infimum and supremum operations on R. For $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$, we add an axiom forcing that inR assigns no elements of R and all elements of A (the latter means all attributes are *irrelevant* for the element under consideration!) to the interpretation of H. □

Corollary 4.2. *The $\mathbb{SP}$-closures of $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ and $\mathcal{R}_{\mathrm{unr}}$ are quasiequational classes.*

Corollary 4.3. *The quasiequational theories of $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ and $\mathcal{R}_{\mathrm{unr}}$ are recursively enumerable (the same applies to universal and elementary theories of these classes).*

Proof. The axiomatizations in the extended language proposed in the proof of Theorem 4.1 are finite. □

Note that we would not be able to prove the Theorem 4.1 if we demanded that headers are finite subsets of $\mathcal{A}$. This would require including an axiom forcing finiteness of $A - r$ for every r of sort R in the proof of Theorem 4.1 and this condition cannot be forced by first-order sentences. However, concrete database instances always belong to $\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}}$ and, as it turns out, the decidability status of the quasiequational theory of $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ and $\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}}$ is the same.

An undecidability result is also provable for the corresponding abstract class, much like for relation algebras and cylindric algebras—in fact, we build on a proof of Maddux [Mad80] for CA_3. Moreover, we do not even need all the axioms of $\underline{R}^{\mathsf{H}}$. Let $\underline{RH1}$ be the variety of $\mathcal{L}_{\mathsf{H}}$-algebras axiomatized by the lattice axioms and AxRH1. It is straightforward to verify that $\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}} \subset \mathcal{R}^{\mathsf{H}}_{\mathrm{unr}} \subset \mathbb{SP}(\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}) \subseteq \underline{R}^{\mathsf{H}} \subset \underline{RH1}$ and by clause 7 of Theorem 3.5, Der1 holds in $\underline{RH1}$. Note that AxRH1 would hold if we take H to be a bottom element if a given lattice has one and it would also hold for an arbitrary choice of H in a distributive lattice.

Definition 4.4. *Let $\overline{e} = (u_0, u_1, u_2, e_0, e_1)$ be an arbitrary 5-tuple of variables. We abbreviate $u_0 \bowtie u_1 \bowtie u_2$ as u. For arbitrary L-terms s, t define*

$$\begin{aligned}
\mathbf{c}_0^{\overline{e}}\langle t\rangle &= u \bowtie (\mathsf{H} \bowtie u_1 \bowtie u_2 \oplus u \bowtie t)\\
\mathbf{c}_1^{\overline{e}}\langle t\rangle &= u \bowtie (\mathsf{H} \bowtie u_0 \bowtie u_2 \oplus u \bowtie t)\\
\mathbf{c}_2^{\overline{e}}\langle t\rangle &= u \bowtie (\mathsf{H} \bowtie u_0 \bowtie u_1 \oplus u \bowtie t)\\
s \circ^{\overline{e}} t &= \mathbf{c}_2^{\overline{e}}\left\langle \mathbf{c}_1^{\overline{e}}\left\langle e_0 \bowtie \mathbf{c}_2^{\overline{e}}\langle s\rangle\right\rangle \bowtie \mathbf{c}_0^{\overline{e}}\left\langle e_1 \bowtie \mathbf{c}_2^{\overline{e}}\langle s\rangle\right\rangle\right\rangle
\end{aligned}$$

Let $T_n(x_1, \ldots, x_n)$ be the collection of all semigroup terms in n variables. Whenever $\overline{e} = (x_{n+1}, \ldots, x_{n+5})$ define the translation $\tau^{\overline{e}}$ of semigroup terms as follows: $\tau^{\overline{e}}(x_i) = x_i$ for $i \leq n$ and $\tau^{\overline{e}}(s \circ t) = s \circ^{\overline{e}} t$ for any $s, t \in T_n(x_1, \ldots, x_n)$.

Whenever $\overline{e}$ is clear from the context, we will drop it to ensure readability.

Theorem 4.5. *For any $p_0, \ldots, p_m, r_0, \ldots, r_m, s, t \in T_n(x_1, \ldots, x_n)$, the following conditions are equivalent :*

(I) The quasiequation

$$\text{(Qu3)} \quad \forall x_1, \ldots, x_n. (p_0 = r_0 \,\&\, \ldots \,\&\, p_m = r_m \Rightarrow s = t)$$

holds in all semigroups (finite semigroups)

(II) For $\overline{e} = (x_{n+1}, \ldots, x_{n+5})$ as in Definition 4.4, the quasiequation

$$\text{(Qu4)} \quad \begin{aligned}
&\forall x_0, x_1, \ldots, x_{n+5}. (\tau^{\overline{e}}(p_0) = \tau^{\overline{e}}(r_0) \,\&\, \ldots \tau^{\overline{e}}(p_m) = \tau^{\overline{e}}(r_m) \,\&\\
&\qquad \&\, x_{n+4} = \mathbf{c}_0^{\overline{e}}\langle x_{n+4}\rangle \,\&\, x_{n+5} = \mathbf{c}_1^{\overline{e}}\langle x_{n+5}\rangle) \Rightarrow\\
&\qquad\qquad \Rightarrow \tau^{\overline{e}}(s) \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}}\langle x_0\rangle = \tau^{\overline{e}}(t) \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}}\langle x_0\rangle))
\end{aligned}$$

holds in every member of $\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ (every member of $\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}}$).

(III) Qu4 above holds in every member of $\underline{RH1}$ (finite member of $\underline{RH1}$).

Proof. (I) $\Rightarrow$ (III). By contraposition:

Take any $\mathfrak{A} \in \underline{RH1}$ and arbitrarily chosen elements $u_0, u_1, u_2 \in \mathfrak{A}$. In order to use Maddux's technique, we have to prove that for any $a, b \in \mathfrak{A}$ and $k, l < 3$

(b) $\mathbf{c}_k\langle \mathbf{c}_k\langle a\rangle\rangle = \mathbf{c}_k\langle a\rangle$
(c) $\mathbf{c}_k\langle a \bowtie \mathbf{c}_k\langle b\rangle\rangle = \mathbf{c}_k\langle a\rangle \bowtie \mathbf{c}_k\langle b\rangle$
(d) $\mathbf{c}_k\langle \mathbf{c}_l\langle a\rangle\rangle = \mathbf{c}_l\langle \mathbf{c}_k\langle a\rangle\rangle$

(we deliberately keep the same labels as in the quoted paper), where $\mathbf{c}_k\langle a\rangle$ is defined in the same way as in Definition 4.4 above. We will denote by $u_{\hat{k}}$ the product of u_i's such that $i \in \{0, 1, 2\} - \{k\}$. For example, $u_{\hat{0}} = u_1 \bowtie u_2$.

For (b):

$$\begin{aligned}
L &= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a))\\
&= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \bowtie (u \oplus \mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a) \oplus u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a)) && \text{by lattice laws}\\
&= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \bowtie u \oplus \mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a) \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a) \oplus u) && \text{by AxRH1}\\
&= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a) \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u) && \text{by lattice laws}\\
&= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a) && \text{by lattice laws}\\
&= R
\end{aligned}$$

(c) is proved using a similar trick:

$$\begin{aligned}
L &= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie b)) \\
&= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \bowtie (u \bowtie a \oplus \mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie b) \oplus u \bowtie a \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie b)) && \text{by lattice laws} \\
&= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \bowtie u \bowtie a \oplus \mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie b) \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie b) \oplus u \bowtie a) && \text{by AxRH1} \\
&= u \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie b) \bowtie (\mathsf{H} \bowtie u_{\hat{k}} \oplus u \bowtie a) && \text{by lattice laws} \\
&= R
\end{aligned}$$

(d) is obviously true for $k = l$, hence we can restrict attention to $k \neq l$. Let j be the remaining element of $\{0, 1, 2\}$. Thus,

$$\begin{aligned}
L &= u \bowtie (\mathsf{H} \bowtie u_l \bowtie u_j \oplus u \bowtie (\mathsf{H} \bowtie u_k \bowtie u_j \oplus u \bowtie a)) \\
&= u \bowtie (\mathsf{H} \bowtie u_l \bowtie u_j \oplus u_l \bowtie (\mathsf{H} \bowtie u_k \bowtie u_j \oplus u \bowtie a)) && \text{by Der1} \\
&= u \bowtie (\mathsf{H} \bowtie u_l \bowtie u_j \bowtie (u_l \oplus \mathsf{H} \bowtie u_k \bowtie u_j \oplus u \bowtie a) \oplus u_l \bowtie (\mathsf{H} \bowtie u_k \bowtie u_j \oplus u \bowtie a)) && \text{by lattice laws} \\
&= u \bowtie (\mathsf{H} \bowtie u_l \bowtie u_j \oplus \mathsf{H} \bowtie u_k \bowtie u_j \oplus u \bowtie a) \bowtie (\mathsf{H} \bowtie u_l \bowtie u_j \bowtie (\mathsf{H} \bowtie u_k \bowtie u_j \oplus u \bowtie a) \oplus u_l) && \text{by AxRH1} \\
&= u \bowtie (\mathsf{H} \bowtie u_l \bowtie u_j \oplus \mathsf{H} \bowtie u_k \bowtie u_j \oplus u \bowtie a) \bowtie u_l && \text{by lattice laws} \\
&= u \bowtie (\mathsf{H} \bowtie u_l \bowtie u_j \oplus \mathsf{H} \bowtie u_k \bowtie u_j \oplus u \bowtie a) && \text{by lattice laws}
\end{aligned}$$

In the last term, u_l and u_k may be permuted by commutativity. Now doing analogous sequence of transformations in the reverse direction with the roles of u_k and u_l replaced we obtain the right side of the equation.

The rest of the proof mimics the one in [Mad80]. In some detail: assume there is $\overline{e} = (u_0, u_1, u_2, e_0, e_1) \in \mathfrak{A}$ such that

(a) $\mathbf{c}_0^{\overline{e}} \langle e_0 \rangle = e_0$, $\mathbf{c}_1^{\overline{e}} \langle e_1 \rangle = e_1$

holds. Using (a)–(d) we prove that for every $a, b \in \mathfrak{A}$ the following hold:

(i) $\mathbf{c}_1^{\overline{e}} \left\langle a \circ^{\overline{e}} b \right\rangle = a \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle b \rangle$
(ii) $a \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle b \rangle = \mathbf{c}_1^{\overline{e}} \left\langle \mathbf{c}_2^{\overline{e}} \langle a \rangle \bowtie \mathbf{c}_0^{\overline{e}} \left\langle \mathbf{c}_2^{\overline{e}} \left\langle e_0 \bowtie e_1 \bowtie \mathbf{c}_2^{\overline{e}} \left\langle \mathbf{c}_1^{\overline{e}} \langle b \rangle \right\rangle \right\rangle \right\rangle \right\rangle$
(iii) $(a \circ^{\overline{e}} b) \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle c \rangle = a \circ^{\overline{e}} (b \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle c \rangle)$
(iv) $((a \circ^{\overline{e}} b) \circ^{\overline{e}} c) \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle d \rangle = (a \circ^{\overline{e}} (b \circ^{\overline{e}} c)) \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle d \rangle$

Now, a failure of Qu4 implies the existence of $\mathfrak{A} \in \mathcal{R}^{\mathsf{H}}_{\mathrm{unr}}$ (or $\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}}$) and $\overline{e} = (u_0, u_1, u_2, e_0, e_1)$ such that elements of $\overline{e}$ interpret variables $(x_{n+1}, \ldots, x_{n+5})$ in Qu4. This means (a) is satisfied, hence (i)–(iv) hold for every element of $\mathfrak{A}$. We define an equivalence relation $\equiv$ on $\mathfrak{A}$:

$a \equiv b$ iff for all $c \in \mathfrak{A}$, $a \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle c \rangle = b \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle c \rangle$.

We take $\circ^{\overline{e}}$ to be the semigroup operation on $\mathfrak{A}/\equiv$. Following [Mad80], we use (i)–(iv) to prove that this operation is well-defined (i.e., independent of the choice of representatives) and satisfies semigroup axioms. It follows from the assumptions that the semigroup thus defined fails Qu3.

(III) $\Rightarrow$ (II). Immediate.

(II) $\Rightarrow$ (I). As in [Mad80], define the full relational lattice with at least three attributes over a semigroup $\mathfrak{B} = (B, \circ, \mathsf{u})$ (with an identity element u) failing Qu3 with some valuation v witnessing this failure. That is, let $\mathcal{D}$ be the domain of $\mathfrak{B}$ and $\mathcal{A}$ be $\{0, 1, 2\}$; note that for finite $\mathfrak{B}$ this yields a finite relational lattice. Define a valuation in this lattice as follows:

$$\begin{aligned}
w(x_0) =&(\{0,1,2\}, \{\{(0, v(r)), (1, a), (2, b)\} \mid a, b \in \mathfrak{B}\}) \\
w(x_i) =&(\{0,1,2\}, \{\{(0, a), (1, a \circ v(x_i)), (2, b)\} \mid a, b \in \mathfrak{B}\}) && i \leq n \\
w(x_{n+i}) =&(\{i\}, \{\{(i, b)\} \mid b \in \mathfrak{B}\}) && (0 < i \leq 3) \\
w(x_{n+4}) =&(\{0,1,2\}, \{\{(0, a), (1, b), (2, b)\} \mid a, b \in \mathfrak{B}\}) \\
w(x_{n+5}) =&(\{0,1,2\}, \{\{(0, b), (1, a), (2, b)\} \mid a, b \in \mathfrak{B}\})
\end{aligned}$$

It is proved by induction that

$$w(\tau^{\overline{e}}(t)) = (\{0,1,2\}, \{\{(0, a), (1, a \circ v(t)), (2, b)\} \mid a, b \in \mathfrak{B}\})$$

(where $e = (x_{n+1}, \ldots, x_{n+5})$) for every $t \in T(x_1, \ldots x_n)$ and also

$$\begin{aligned}
w(\tau^{\overline{e}}(s) \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle x_0 \rangle) =&(\{0,1,2\}, \{\{(0, a), (1, b), (2, c)\} \mid a, b, c \in \mathfrak{B}, v(r) \circ a = v(s)\}) \\
w(\tau^{\overline{e}}(r) \circ^{\overline{e}} \mathbf{c}_1^{\overline{e}} \langle x_0 \rangle) =&(\{0,1,2\}, \{\{(0, a), (1, b), (2, c)\} \mid a, b, c \in \mathfrak{B}, v(r) \circ a = v(r)\})
\end{aligned}$$

Any tuple whose value for attribute 0 is u belongs to the first relation, but not to the second. Thus w is a valuation refuting Qu4.

□

Corollary 4.6. *The quasiequational theory of any class of algebras between* $\mathcal{R}_{\text{fin}}^{\mathsf{H}}$ *and* $\underline{RH1}$ *is undecidable.*

Proof. Follows from Theorem 4.5 and theorems of Gurevič [Gur66] (see also [GL84]) and Post [Pos47] (for finite and arbitrary semigroups, respectively).

□

Corollary 4.7. *The quasiequational theory of* $\mathcal{R}_{\text{fin}}^{\mathsf{H}}$ *is not finitely axiomatizable.*

Proof. Follows from Theorem 4.5 and the Harrop criterion [Har58]. □

Problem 4.8. Are the quasiequational theories of $\mathcal{R}_{\text{unr}}$ and $\mathcal{R}_{\text{fin}}$ (i.e., of lattice reducts) decidable?

5 The Concept Structure of Tropashko Lattices

In their foundational work on Formal Concept analysis, Ganter and Wille [GW96] show how to use the *incidence relation* between the join- and the meet-irreducibles of a finite lattice (i.e., its *standard context*) to investigate its structure. Recall that if $\mathcal{L}$ is a finite lattice, $\mathfrak{J}(\mathcal{L})$ is the set of its join-irreducibles and $\mathfrak{M}(\mathcal{L})$ is the set of its meet-irreducibles, then the standard context of $\mathcal{L}$ is defined as $\mathsf{con}(\mathcal{L}) := (\mathfrak{J}(\mathcal{L}), \mathfrak{M}(\mathcal{L}), \mathsf{I}_{\leq})$, where $\mathsf{I}_{\leq} := \leq \cap (\mathfrak{J}(\mathcal{L}) \times \mathfrak{M}(\mathcal{L}))$. Set

$$\begin{array}{ll} g \swarrow m: & g \text{ is } \leq\text{-minimal in } \{h \in \mathfrak{J}(\mathcal{L}) \mid \text{not } h \,\mathsf{I}_{\leq}\, m\} \\ g \nearrow m: & m \text{ is } \leq\text{-maximal in } \{n \in \mathfrak{M}(\mathcal{L}) \mid \text{not } g \,\mathsf{I}_{\leq}\, n\} \\ g \swarrow\!\!\!\nearrow m: & g \swarrow m \,\&\, g \nearrow m \end{array}$$

Let also $\swarrow\!\!\swarrow$ be the smallest relation containing $\swarrow$ and satisfying the condition

$$g \swarrow\!\!\swarrow m, \; h \nearrow m \text{ and } h \swarrow n \text{ imply } g \swarrow\!\!\swarrow n;$$

in a more compact notation, $\swarrow\!\!\swarrow \circ \nearrow \circ \swarrow \subseteq \swarrow\!\!\swarrow$. We have the following

Proposition 5.1. *[GW96, Theorem 17] A finite lattice is*

- *subdirectly irreducible iff there is $m \in \mathfrak{M}(\mathcal{L})$ such that $\swarrow\!\!\swarrow \supseteq \mathfrak{J}(\mathcal{L}) \times \{m\}$*
- *simple iff $\swarrow\!\!\swarrow = \mathfrak{J}(\mathcal{L}) \times \mathfrak{M}(\mathcal{L})$*

Let us describe $\mathfrak{J}(\mathfrak{R}(\mathcal{D}, \mathcal{A}))$ and $\mathfrak{M}(\mathfrak{R}(\mathcal{D}, \mathcal{A}))$ for finite $\mathcal{D}$ and $\mathcal{A}$. Set

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

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

It is worth noting that $\mathfrak{R}(\mathcal{D}, \mathcal{A})$ naturally divides into what we may call *boolean H-slices*—i.e., the powerset algebras of ${}^{H}\mathcal{D}$ for each $H \subseteq \mathcal{A}$. Furthermore, the projection mapping from H-slice to H'-slice where $H' \subseteq H$ is a join-homomorphism. Lastly, note that the bottom elements of H-slices—i.e., elements of the form $(H, \emptyset)$—and top elements of the form $(H, {}^{H}\mathcal{D})$ form two additional boolean slices, which we may call the *lower attribute slice* and *the upper attribute slice*, respectively. Both are obviously isomorphic copies of the powerset algebra of $\mathcal{A}$. The intention of our definition should be clear then:

- The join-irreducibles are only the atoms of the $\mathcal{A}$-slice (i.e., the slice with the longest tuples) plus the atoms of the lower attribute slice
- The meet-irreducibles are much richer: they consists of the coatoms of *all H-slices* (note $\mathcal{M}_{\mathcal{D},\mathcal{A}}$ includes H as the sole element of $\mathcal{C}oDom_{\mathcal{D},\emptyset}$) plus all coatoms of the *upper* attribute slice.

Let us formalize these two itemized points as

Theorem 5.2. *For any finite $\mathcal{A}$ and $\mathcal{D}$ such that $|\mathcal{D}| \geq 2$, we have*

$$\begin{array}{ll} \mathcal{J}_{\mathcal{D},\mathcal{A}} = \mathfrak{J}(\mathfrak{R}(\mathcal{D}, \mathcal{A})) & \textit{(join-irreducibles)} \\ \mathcal{M}_{\mathcal{D},\mathcal{A}} = \mathfrak{M}(\mathfrak{R}(\mathcal{D}, \mathcal{A})) & \textit{(meet-irreducibles)} \end{array}$$

Proof. (join-irreducibles): To prove the $\subseteq$-direction, simply observe that the elements of $\mathcal{J}_{\mathcal{D},\mathcal{A}}$ are exactly the atoms of $\mathfrak{R}(\mathcal{D},\mathcal{A})$. For the converse, note that

- every element in a H-slice is a join of the atoms of this slice, as each H-slice has a boolean structure and in the boolean case atomic = atomistic,
- the header elements $(H,\emptyset)$ are joins of elements of $\mathcal{AA}tt_{\mathcal{D},\mathcal{A}}$
- the atoms of H-slices are joins of header elements with elements of $\mathcal{AA}tt_{\mathcal{D},\mathcal{A}}$. Hence, no element of $\mathfrak{R}(\mathcal{D},\mathcal{A})$ outside $\mathcal{AA}tt_{\mathcal{D},\mathcal{A}}$ can be join-irreducible.

(meet-irreducibles): This time, the $\supseteq$-direction is easier to show: $\mathcal{M}_{\mathcal{D},\mathcal{A}}$ includes the coatoms of the H-slices and the upper attribute slices. Hence, the basic properties of finite boolean algebras imply all meet-irreducibles must be contained in $\mathcal{M}_{\mathcal{D},\mathcal{A}}$: every element of $\mathfrak{R}(\mathcal{D},\mathcal{A})$ can be obtained as an intersection of elements of $\mathcal{M}_{\mathcal{D},\mathcal{A}}$. For the $\subseteq$-direction, it is clear that elements of $\mathcal{C}o\mathcal{A}tt_{\mathcal{D},\mathcal{A}}$ are meet-irreducible, as they are coatoms of the whole $\mathfrak{R}(\mathcal{D},\mathcal{A})$. This also applies to $\mathsf{H} \in \mathcal{C}o\mathcal{D}om_{\mathcal{D},\emptyset}$. Now take $\mathsf{codom}^H(x) = (H, {}^H\mathcal{D} - \{x\})$ for a non-empty $H = \{1,\dots,h\}$ and $x = (x_1,\dots x_h) \in {}^H\mathcal{D}$ and assume $\mathsf{codom}^H(x) = r \bowtie s$ for $r,s \neq \mathsf{codom}^H(x)$. That is, $H = H_r \cup H_s$ and

$$ {}^H\mathcal{D} - \{x\} = \{y \in {}^{H_r\cup H_s}\mathcal{D} \mid y[H_r] \in B_r \quad \text{and} \quad y[H_s] \in B_s\}. $$

Note that wlog $H_r \subsetneq H$ and $r \subseteq \mathsf{codom}^{H_r}(z)$ for some $z \in {}^{H_r}\mathcal{D}$; otherwise, if both r and s were top elements of their respective slices, their meet would be $(H, {}^H\mathcal{D})$. Thus ${}^H\mathcal{D} - \{x\} \subseteq \{y \in {}^H\mathcal{D} \mid y[H_r] \neq z\}$ and by contraposition

$$ \{y \in {}^H\mathcal{D} \mid y[H_r] = z\} \subseteq \{x\}. \quad (1) $$

This means that $z = x[H_r]$. But now take any $i \in H - H_r$, pick any $d \neq x_i$ (here is where we use the assumption that $|\mathcal{D}| \geq 2$) and set

$$ x' := (x_1,\dots,x_{i-1},d,x_{i+1},\dots,x_h). $$

Clearly, $x'[H_r] = x[H_r] = z$, contradicting (1).

□

Theorem 5.3. *Assume $\mathcal{D},\mathcal{A}$ are finite sets such that $|\mathcal{D}| \geq 2$ and $\mathcal{A} \neq \emptyset$. Then $\mathrel{\mathsf{I}_{\leq}}$, $\swarrow$, $\nearrow$ and $\swarrow\!\!\!\nearrow$ look for $\mathfrak{R}(\mathcal{D},\mathcal{A})$ as follows:*

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

Proof (Sketch).

For the $\mathsf{I}_{\leq}$-row: this is just spelling out the definition of $\leq$ on $\mathfrak{R}(\mathcal{D},\mathcal{A})$ as restricted to $\mathcal{J}_{\mathcal{D},\mathcal{A}} \times \mathcal{M}_{\mathcal{D},\mathcal{A}}$.

For the $\swarrow$-row: the set of join-irreducibles consists of only of the atoms of the whole lattice, hence $\swarrow$ is just the complement of $\leq$.

This observation already yields $\nearrow \subseteq \swarrow$ and $\swarrow\!\!\!\nearrow = \nearrow$. The last missing piece of information to define $\nearrow$ is provided by the analysis of restriction of $\leq$ to $\mathcal{M}_{\mathcal{D},\mathcal{A}} \times \mathcal{M}_{\mathcal{D},\mathcal{A}}$:

$$\text{for} \quad \begin{array}{ll} r = \mathsf{coatt}(a), & s = \mathsf{coatt}(b), \\ r = \mathsf{coatt}(a), & s = \mathsf{codom}^H(x), \\ r = \mathsf{codom}^H(x), & s = \mathsf{coatt}(a), \\ r = \mathsf{codom}^H(x), & s = \mathsf{codom}^H(y), \end{array} \quad r \leq s \quad \text{iff} \quad \begin{array}{l} \text{never} \\ \text{never} \\ a \in H \\ \text{never.} \end{array}$$

Finally, for $\swarrow\!\!\swarrow$ we need to observe that composing $\swarrow$ with $\nearrow \circ \swarrow$ does not allow to reach any new elements of $\mathcal{C}oAtt_{\mathcal{D},\mathcal{A}}$. As for elements of $\mathcal{M}_{\mathcal{D},\mathcal{A}}$ of the form $\mathsf{codom}^H(y)$, note that

$$\exists h.(h \nearrow \mathsf{coatt}(a) \,\&\, h \swarrow \mathsf{codom}^H(y)) \text{ if } a \in H \tag{2}$$

$$\exists h.(h \nearrow \mathsf{codom}^{H_x}(x) \,\&\, h \swarrow \mathsf{codom}^{H_y}(y)) \text{ if } x[H_x \cap H_y] = y[H_x \cap H_y]. \tag{3}$$

Furthermore, we have that

- for any $x \in {}^{\mathcal{A}}\mathcal{D}$ and any $H \subseteq \mathcal{A}$, $\mathsf{adom}(x) \swarrow \mathsf{codom}^H(x[H])$
- for any $a \in \mathcal{A}$ and any $x \in {}^{\mathcal{A}}\mathcal{D}$, $\mathsf{aatt}(a) \swarrow \mathsf{codom}^{\mathcal{D}}(x)$

Using (3), we obtain then that $\mathcal{J}_{\mathcal{D},\mathcal{A}} \times \{\mathsf{H}\} \subseteq \swarrow\!\!\swarrow$ and using (3) again—that $\mathcal{J}_{\mathcal{D},\mathcal{A}} \times \{\mathsf{codom}^H(y)\} \subseteq \swarrow\!\!\swarrow$ for any $y \in {}^{\mathcal{A}}\mathcal{D}$ and any $H \subseteq \mathcal{A}$.

□

Corollary 5.4. *If $\mathcal{D},\mathcal{A}$ are finite sets such that $|\mathcal{D}| \geq 2$ and $\mathcal{A} \neq \emptyset$, then $\mathfrak{R}(\mathcal{D},\mathcal{A})$ is subdirectly irreducible but not simple.*

Proof. Follows immediately from Proposition 5.1 and Theorem 5.3. □

6 Conclusions and Future Work

6.1 Possible Extensions of the Signature

Clearly, it is possible to define more operations on $\mathcal{R}^{\mathsf{H}}_{\text{unr}}$ than those present in $\mathcal{L}_{\mathsf{H}}$. Thus, our first proposal for future study, regardless of the negative result in Corollary 4.6, is a systematic investigation of extensions of the signature. Let us discuss several natural ones; see also [ST06, Tro].

The top element $\top = (\emptyset, \{\emptyset\})$. Its inclusion in the signature would be harmless, but at the same time does not appear to improve expressivity a lot.

The bottom element $\bot = (\mathcal{A}, \emptyset)$. Whenever $\mathcal{A}$ is infinite, including $\bot$ in the signature would exclude subalgebras consisting of relations with finite headers—i.e., exactly those arising from concrete database instances. Another undesirable feature is that the interpretation of $\bot$ depends on $\mathcal{A}$, i.e., the collection of all possible attributes, which is not explicitly supplied by a query expression.

The full relation $\mathsf{U} = (\mathcal{A}, {}^{\mathcal{A}}\mathcal{D})$ [Tro, ST06]. Its inclusion would destroy the *domain independence property (d.i.p.)* [AHV95, Ch. 5] mentioned above. Note that for non-empty $\mathcal{A}$ and $\mathcal{D}$, U is a complement of H.

Attribute constants $\underline{\mathsf{a}} = (\{\mathsf{a}\}, \emptyset\})$ *for* $\mathsf{a} \in \mathcal{A}$. We touch upon an important difference between our setting and that of both *named SPJR algebra* and *unnamed SPC algebra* in [AHV95, Ch. 4], which are *typed*: expressions come with an explicit information about their headers (*arities* in the unnamed case). Our expressions are untyped *query schemes*. On the one hand, $\mathcal{L}_{\mathsf{H}}$ allows, e.g., *projection of* r *to the header of* s: $r \oplus (s \bowtie \mathsf{H})$, which does not correspond to any *single* SPJR expression. On the other hand, only with attribute constants we can write the SPJR *projection of* r *to a* <u>*concrete*</u> *header* $\{\mathsf{a}_1, \ldots, \mathsf{a}_n\}$: $\pi_{\mathsf{a}_1,\ldots,\mathsf{a}_n}(r) = r \oplus \underline{\mathsf{a}_1} \bowtie \ldots \bowtie \underline{\mathsf{a}_n}$.

Unary singleton constants $(\underline{\mathsf{a}} : \mathsf{d}) = (\{\mathsf{a}\}, \{(\mathsf{a} : \mathsf{d})\})$ for $\mathsf{a} \in \mathcal{A}$, $\mathsf{d} \in \mathcal{D}$. These are among the *base SPJR queries* [AHV95, p. 58]. Note they add more expressivity than attribute constants: whenever the signature includes $(\underline{\mathsf{a}} : \mathsf{d})$ for some $d \in \mathcal{D}$, we have $\underline{\mathsf{a}} = (\underline{\mathsf{a}} : \mathsf{d}) \bowtie \mathsf{H}$. They also allow to define $\top$ as $\top = (\underline{\mathsf{a}} : \mathsf{d}) \oplus \mathsf{H}$ and, more importantly, the SPJR *constant-based selection queries* $\sigma_{\underline{\mathsf{a}}=\mathsf{d}}(r) = r \bowtie (\underline{\mathsf{a}} : \mathsf{d})$.

The equality constant $\mathbf{\Delta} = (\mathcal{A}, \{x \in {}^{\mathcal{A}}\mathcal{D} \mid \forall a, a'. x(a) = x(a')\})$. With it, we can express the *equality-based selection queries*: $\sigma_{\underline{\mathsf{a}}=\underline{\mathsf{b}}}(r) = r \bowtie (\mathbf{\Delta} \oplus \underline{\mathsf{a}} \bowtie \underline{\mathsf{b}})$. But the interpretation of $\mathbf{\Delta}$ violates d.i.p., hence we prefer *the di-equality operator*:

$$\overline{\overline{r}} = (H_r, \{x \in {}^{H_r}\mathcal{D} \mid \exists x' \in r. \exists a' \in H_r. \forall a \in H_r. x(a) = x'(a')\}),$$

which also allows to define $\sigma_{\underline{\mathsf{a}}=\underline{\mathsf{b}}}(r) = r \bowtie (\overline{\overline{r}} \oplus \underline{\mathsf{a}} \bowtie \underline{\mathsf{b}})$.

The header-narrowing operator $r \pitchfork s = (H_r - H_s, \{x[H_r - H_s] \mid x \in H_r\})$. This one is perhaps more surprising, but now we can define the *attribute renaming* operators [AHV95, p. 58] as $\rho_{\underline{\mathsf{a}} \mapsto \underline{\mathsf{b}}}(r) = (r \bowtie \overline{\overline{(r \oplus \underline{\mathsf{a}}) \bowtie (\underline{\mathsf{b}} : \mathsf{d})}}) \pitchfork \underline{\mathsf{a}}$, where $d \in \mathcal{D}$ is arbitrary. Instead of using $\pitchfork$, one could add constants for elements $\mathsf{aatt}(\mathsf{a})$ introduced in Section 5, but this would lead to the same criticism as $\bot$ above: indeed, such constants would make $\bot$ definable as $\bot = \mathsf{aatt}(\mathsf{a}) \bowtie \underline{\mathsf{a}}$.

Overall, one notices that just to express the operators discussed in [AHV95, Ch. 4], it would be sufficient to add special constants, but more care is needed in order to preserve the d.i.p. and similar relativization/finiteness properties.

The difference operator $r - s = (H_r, \{x \in B_r \mid x \notin B_s\})$. This is a very natural extension from the DB point of view [AHV95, Ch. 5], which leads us beyond the SPJRU setting towards the question of *relational completeness* [Cod70]. Here again we break with the partial character of Codd's original operator. Another option would be $(H_{r \cap s}, \{x \in B_r[H_s] \mid x \notin B_s[H_r]\})$, but this one can be defined using the operator above as $(r \oplus s) - (s \oplus (r \bowtie \mathsf{H}))$.

6.2 Summary and Other Directions for Future Research

We have seen that relational lattices form an interesting class with rather surprising properties. Unlike Codd's relational algebra, all operations are total and in contrast to the encoding of relational algebras in cylindric algebras, the domain independence property obtains automatically. We believe that with the extensions of the language proposed in Section 6.1, one can ultimately obtain most natural algebraic treatment of SPRJ(U) operators and relational query languages. Besides, given how well investigated the lattice of varieties of lattices is in general [JR92], it is intriguing to discover a class of lattices with a natural CS motivation which does not seem to fit anywhere in the existing picture.

To save space and reader's patience, we are not going to recall again all the conjectures and open questions posed above, but without settling them we cannot claim to have grasped how relational lattices behave as an algebraic class. None of them seems trivial, even with the rich supply of algebraic logic tools available in existing literature. A reference not mentioned so far and yet potentially relevant is [Cra74]. An interesting feature of Craig's setting from our point of view is that it allows tuples of varying arity.

We would also like to mention the natural question of *representability*:

Problem 6.1 (Hirsch). Given a finite algebra in the signature $\mathcal{L}_{\mathsf{H}}$ ($\mathcal{L}$), is it decidable whether it belongs to $\mathbb{SP}(\mathcal{R}^{\mathsf{H}}_{\mathrm{unr}})$, $\mathbb{SP}(\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}})$ ($\mathbb{SP}(\mathcal{R}_{\mathrm{unr}})$, $\mathbb{SP}(\mathcal{R}_{\mathrm{fin}})$)?

We believe that the analysis of the concept structure of finite relational lattices in Section 5 may lead to an algorithm recognizing whether the concept lattice of a given context belongs to $\mathbb{SP}(\mathcal{R}^{\mathsf{H}}_{\mathrm{fin}})$ (or $\mathbb{SP}(\mathcal{R}_{\mathrm{fin}})$). It also opens the door to a systematic investigation of a research problem suggested by Yde Venema: *duality theory of relational lattices*. See also Section 2.1 above for another category-theoretical connection.

References

[AHV95] Abiteboul, S., Hull, R., Vianu, V.: Foundations of Databases. Addison-Wesley (1995)

[Cod70] Codd, E. F.: A Relational Model of Data for Large Shared Data Banks. Commun. ACM 13, 377–387 (1970)

[Cra74] Craig, W.: Logic in Algebraic Form. Three Languages and Theories, Studies in Logic and the Foundations of Mathematics, vol. 72. North Holland (1974)

[DM01] Düntsch, I., Mikulás, S.: Cylindric structures and dependencies in relational databases. Theor. Comput. Sci. 269, 451–468 (2001)

[GW96] Ganter, B., Wille, R.: Applied Lattice Theory: Formal Concept Analysis In Grätzer, G. (ed.) General Lattice Theory, second edition. Birkhäuser (1996)

[Gur66] Gurevich, Y.: The word problem for certain classes of semigroups. Algebra and Logic 5, 25–35 (1966)

[GL84] Gurevich, Y., Lewis, H. R.: The Word Problem for Cancellation Semigroups with Zero. The Journal of Symbolic Logic 49, 184–191 (1984)

[Har58] Harrop, R.: On the existence of finite models and decision procedures for propositional calculi. Mathematical Proceedings of the Cambridge Philosophical Society 54, 1–13 (1958)

[HH02] Hirsch, R., Hodkinson, I.: Relation Algebras by Games, Studies in Logic and the Foundations of Mathematics, vol. 147. Elsevier (2002)

[IL84] Imieliński, T., Lipski, W.: The Relational Model of Data and Cylindric Algebras. J. Comput. Syst. Sci. 28, 80–102 (1984)

[Jac99] Jacobs, B.: Categorical Logic and Type Theory. Studies in Logic and the Foundations of Mathematics **141**. North Holland, Amsterdam (1999)

[JR92] Jipsen, P., Rose, H.: Varieties of Lattices, Lecture Notes in Mathematics, vol. 1533. Springer Verlag (1992)

[JR98] Jipsen, P., Rose, H.: Varieties of Lattices In Grätzer, G. (ed.) General Lattice Theory, pp. 555–574. Birkhäuser Verlag (1998). Appendix F to the second edition

[Mad80] Maddux, R.: The Equational Theory of CA_3 is Undecidable. The Journal of Symbolic Logic 45, 311–316 (1980)

[Mae74] Maeda, S.: Locally Modular Lattices and Locally Distributive Lattices. Proceedings of the American Mathematical Society 44, pp. 237–243 (1974)

[McK72] McKenzie, R.: Equational bases and non-modular lattice varieties. Trans. Amer. Math. Soc. 174, 1–43 (1972)

[PMV07] Padmanabhan, R., McCune, W., Veroff, R.: Lattice Laws Forcing Distributivity Under Unique Complementation. Houston Journal of Mathematics 33, 391–401 (2007)

[Pos47] Post, E. L.: Recursive Unsolvability of a Problem of Thue. The Journal of Symbolic Logic 12, 1–11 (1947)

[ST06] Spight, M., Tropashko, V.: First Steps in Relational Lattice. http://arxiv.org/abs/cs/0603044 (2006)

[Sta72] Stanley, R. P.: Supersolvable lattices. Algebra Universalis 2, 197–217 (1972)

[Ste99] Stern, M.: Semimodular Lattices, Encyclopedia of Mathematics and its Applications, vol. 73. Cambridge University Press (1999)

[Tro] Tropashko, V.: The website of QBQL: Prototype of relational lattice system. `https://code.google.com/p/qbql/`

[Tro05] Tropashko, V.: Relational Algebra as non-Distributive Lattice. http://arxiv.org/abs/cs/0501053 (2005)

[VdBVGV07] Van den Bussche, J., Van Gucht, D., Vansummeren, S.: A crash course on database queries. In: PODS '07: Proceedings of the twenty-sixth ACM SIGMOD-SIGACT-SIGART symposium on Principles of Database Systems, pp. 143–154, New York, NY, USA (2007). ACM