Different Differences in Semirings

Dan Suciu
University of Washington, Seattle, WA, USA

To Val Tannen, my teacher, mentor, and friend

Abstract

Relational algebra operates over relations under either set semantics or bag semantics. In 2007 Val Tannen extended the semantics of relational algebra to K-relations, where each tuple is annotated with a value from a semiring. However, only the positive fragment of the relational algebra can be interpreted over K-relations. The reason is that a semiring contains only the operations *addition* and *multiplication*, and does not have a difference operation. This paper explores three ways of adding a difference operator to a semiring: as a freely generated algebra, by using the natural order, or by an explicit construction using products and quotients. The paper consists of both a survey of results from the literature, and of some novel results.

2012 ACM Subject Classification Information systems → Structured Query Language

Keywords and phrases Semirings, K-relations

Digital Object Identifier 10.4230/OASIcs.Tannen.2024.10

Funding NSF-BSF 2109922, NSF IIS 2314527, NSF SHF 2312195.
Dan Suciu: Partially Supported by NSF IIS 2314527.

1 Introduction

Val Tannen's seminal paper [13] extended the positive relational algebra to *K-relations*, where each tuple of the relation is associated with an element of the semiring. Yet this elegant generalization excluded one operator: set difference. The reason is that a semiring defines only the $\oplus$ and $\otimes$ operators and there is no canonical way to add a minus operation, although some semirings appear to admit a natural difference operator, see examples in Sec. 2. This lead several researchers to propose ways to define minus on K-relations. Tannen considered using the *ring* $\mathbb{Z}$ instead of a semiring in [12], and proved that relational algebra expressions admit a canonical form, as sum-of-product expressions. In recent work [8, 11] Tannen used dual-indeterminate polynomials, $\mathbb{N}[X, \bar{X}]$, where a positive and negative variable interact via the identity $x \cdot \bar{x} = 0$.

In this paper we explore three alternative ways to define a difference operator in a semiring; the paper consists both of a survey of related work, and some novel contributions. The first and most obvious approach to define difference algebraically. For any set of desired identities there is a unique way to extend freely a semiring with a difference operation that satisfies those identities. This is a standard technique in universal algebras and we review it in Sec. 3. The question is, what set of identities we should choose. For example if we ask for difference to convert the semiring into a ring, then the freely generate ring may collapse to a trivial ring. For example, the ring freely generated by the natural numbers $\mathbb{N}$ is $\mathbb{Z}$, but the ring freely generated by the Booleans $\mathbb{B}$ is the trivial ring $\{0\}$.

Therefore, in Section 4, we explore an alternative way to define difference: assuming that the semiring is naturally ordered, one can define the difference $b \ominus a$ as the smallest element z such that $a \oplus z \succeq b$. Bosbach [5] and Amer [3] considered naturally ordered semigroups, and monoids respectively, where such a difference operation exists, and proved that they form an equational class that can be described by a small set of axioms. This result is quite surprising, because the natural order does not appear to have a clear algebraic definition: we

The Provenance of Elegance in Computation – Essays Dedicated to Val Tannen.
Editors: Antoine Amarilli and Alin Deutsch; Article No. 10; pp. 10:1–10:20
OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

review this result and present a short, self-contained proof in Sec. 4.2. Geerts and Poggi [9] introduced m-semirings, which are naturally ordered semirings where the difference operation exists. Tannen [4] proved that m-semirings fail to satisfy an important axiom (called $(A5)$ in this paper) that is needed in query optimization: we review this in Sec. 4.3. Reference [4] ends by suggesting the addition of the axiom $(A5)$ to those of an m-semiring, in order to ensure that current optimizations performed by a query optimizer continue to hold over K-relations. However, adding $(A5)$ turns out to be insufficient. We show that, in order to preserve all identities valid under bag semantics, one must ensure that the semiring satisfies all identities that hold in $(\mathbb{N}, +, \cdot, 0, 1, \dot{-})$, where $\dot{-}$ is *monus*, an operation defined below in Eq. (3): we prove in Appendix A that this set is co-r.e. complete, and, thus, it is not finitely axiomatizable.

The definition of monus in Section 4 requires the semiring to be naturally ordered. An interesting question is whether the naturally ordered semirings form an equational class, i.e. whether they can be described by a set of identities. We answer this in Sec. 4.4: while natural order is not definable by a set of equations, it becomes definable if one allows one additional auxiliary operator.

Finally, in Sec. 5 we discuss a third, constructive method for adding difference, by following the same methodology as in the construction of integers $\mathbb{Z}$ from natural numbers $\mathbb{N}$. The traditional construction consists of equivalence classes of pairs (x, y) of natural numbers, where the equivalence relation is given by $(x, y) \equiv (u, v)$ when $x + v = y + u$. This set is isomorphic to $\mathbb{Z}$, and is often taken as the definition of $\mathbb{Z}$. We study whether this definition can be generalized from $\mathbb{N}$ to semirings. If we use the congruence $\equiv$ above, then the quotient semiring is often a trivial semiring, so we look for weaker notions of $\equiv$. We review the concept of an ideal I in a semiring in Sec. 5.2. In a ring, any ideal I defines a congruence relation $x \equiv_I y$, as $x - y \in I$. We describe two alternative ways to define $\equiv_I$ in a semiring, and use them to generalize the $\mathbb{N}$-to-$\mathbb{Z}$ construction. When applied to Booleans $\mathbb{B}$, or to a Boolean algebra, or to the tropical semiring Trop, this produces an extended semiring with a difference operation, which contains *positive elements* x, *negative elements* $\bar{x}$, and *over determined elements* of the form $x + \bar{x}$. These semirings resemble somewhat the dual-indeterminate polynomials $\mathbb{N}[X, \bar{X}]$ introduced in [8, 11], yet they are quite different, for example they do not satisfy the identity $x\bar{x} = 0$.

Finally, we conclude with a short discussion in Sec. 6.

2 Problem Definition and Examples

A *semiring* is a tuple $\boldsymbol{S} = (S, \oplus, \otimes, \mathbf{0}, \mathbf{1})$, where $(S, \oplus, \mathbf{0})$ is a commutative monoid, $(S, \otimes, \mathbf{1})$ is a monoid, $\otimes$ distributes over $\oplus$, and $x \otimes \mathbf{0} = \mathbf{0} \otimes x = \mathbf{0}$. When $\otimes$ is also commutative then we say that the semiring is commutative. We only consider commutative semirings in this paper. When no confusion arises we will denote the operators with $+, \cdot$, and the identities with $0, 1$, without boldface. The semiring is *trivial* when $0 = 1$: in that case $S = \{0\}$, because $x = x \cdot 1 = x \cdot 0 = 0$ for all x.

We denote by Σ_m and Σ_s the signature[1] of monoids and semirings, and by Σ_{mm}, Σ_{sm} their extension with a minus operator, thus:

$$\text{Monoids:} \qquad \Sigma_m \stackrel{\text{def}}{=} \{+, 0\} \qquad \Sigma_{mm} \stackrel{\text{def}}{=} \Sigma_m \cup \{-\} \tag{1}$$

$$\text{Semirings:} \qquad \Sigma_s \stackrel{\text{def}}{=} \{+, \cdot, 0, 1\} \qquad \Sigma_{sm} \stackrel{\text{def}}{=} \Sigma_s \cup \{-\} \tag{2}$$

[1] A signature is also called a vocabulary, or a type.

where $+, \cdot, -$ have arity 2, and $0, 1$ have arity 0. Thus, every monoid is an Σ_m-algebra, and every semiring is a Σ_s-algebra. The problem discussed in this paper is to extend an arbitrary semiring from an Σ_s-algebra to an Σ_{sm}-algebra; some of the discussion will be focused on how to extend the additive monoid to an Σ_{mm}-algebra.

Many semirings already admit a natural difference operator. We illustrate with some examples.

- Every *ring* is a semiring where, for each element x, there exists some $-x$, called the *inverse*, such that $x + (-x) = 0$. The inverse is unique, and $a \cdot (-x) = -(a \cdot x)$. Any ring can be naturally extended to an Σ_{sm}-algebra by defining minus as $y - x \stackrel{\text{def}}{=} y + (-x)$.
- The semiring of natural numbers $(\mathbb{N}, +, \cdot, 0, 1)$ can be also extended to an Σ_{sm}-algebra, by defining the *monus* operation as:

$$y \mathbin{\dot{-}} x \stackrel{\text{def}}{=} \begin{cases} y - x & \text{when } y \geq x \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

- The semiring of Booleans $\mathbb{B} = (\{0, 1\}, \vee, \wedge, 0, 1)$ can be extended with the minus operator $y - x \stackrel{\text{def}}{=} y \wedge (\neg x)$. This extends to any Boolean algebra $(2^\Omega, \cup, \cap, \emptyset, \Omega)$ by defining difference as the standard set difference $y \setminus x$.
- An interesting example is the tropical semiring, $\mathsf{Trop} = ([0, \infty], \min, +, \infty, 0)$, where difference can be defined as:

$$b \ominus a \stackrel{\text{def}}{=} \begin{cases} \infty & \text{when } b \geq a \\ b & \text{otherwise} \end{cases} \tag{4}$$

  This operator is used for semi-naive evaluation of datalog programs over the tropical semiring [1]. For example, consider the APSP (All Pairs Shortest Path) problem. If the current shortest distance between two nodes is $d[x, y] = a$, and the algorithm discovers a new path of length b, then it updates $d[x, y] := \min(a, b)$. The semi-naive algorithm optimizes this step by first computing the difference $\delta \stackrel{\text{def}}{=} b \ominus a$ (using (4)), then updating $d[x, y] := \min(a, \delta)$, which the reader can verify is equal to $\min(a, b)$. The advantage is that, when $b \geq a$ then $\delta = \infty$ and no update is necessary: the algorithm simply ignores all edges where $\delta = \infty$, resulting in a smaller join between the edge relation and the δ relation.

These simple examples don't seem to have a unifying theme. Given an arbitrary semiring $\boldsymbol{S}$, what is the natural way to define difference? We discuss in this paper three approaches to define difference.

3 Difference by Equations

The first approach is to choose a set of identities E that we want the difference operator to satisfy, then consider the Σ_{sm}, E-algebra *freely generated* by $\boldsymbol{S}$. To explain this, we need a quick review universal algebras; there are many good textbooks, for example [6] is available online.

Given a signature Σ, a Σ-algebra is a pair $\boldsymbol{A} = (A, F)$ where F is a set of functions $f^A : A^n \to A$, one for each symbol $f \in \Sigma$ of arity n. Homomorphisms between Σ-algebras, $h : \boldsymbol{A} \to \boldsymbol{B}$, are defined in a straightforward way. The *free algebra* generated by a set X, denoted $T_\Sigma(X)$, is the set of all terms that can be formed from variables in X and function symbols in Σ, see e.g. [6].

An *identity* is a pair $(e_1, e_2) \in T_\Sigma(X)$. If E is a set of identities, then an Σ, E-algebra is an algebra that satisfies[2] all identities in E. The class of all Σ, E-algebras is called an *equational class*, or a *variety*.

A powerful tool for defining difference is the following theorem:

▶ **Theorem 1.** *Let $\Sigma_0 \subseteq \Sigma$, and let E be a set of Σ-identities. Then for any Σ_0-algebra $\boldsymbol{A}$, there exists a pair $(T_{\Sigma,E}(\boldsymbol{A}), \eta)$, where $T_{\Sigma,E}(\boldsymbol{A})$ is a Σ-algebra, $\eta : \boldsymbol{A} \to T_{\Sigma,E}(\boldsymbol{A})$ is a Σ_0-homomorphism, and the following property holds. For every Σ, E-algebra $\boldsymbol{B}$, and any Σ_0-homomorphism $h : \boldsymbol{A} \to \boldsymbol{B}$, there exists a unique Σ-homomorphism $\bar{h}$ such that the following diagram commutes:*

$$\begin{tikzcd} \boldsymbol{A} \arrow[r, "\eta"] \arrow[dr, "h"'] & T_{\Sigma,E}(\boldsymbol{A}) \arrow[d, dotted, "\bar{h}"] \\ & \boldsymbol{B} \end{tikzcd} \tag{5}$$

$T_{\Sigma,E}(\boldsymbol{A})$ is unique up to homomorphism, is called the Σ, E-algebra freely generated *by $\boldsymbol{A}$, and the diagram above is called the* universality property *of $T_{\Sigma,E}(\boldsymbol{A})$.*

This is a very powerful theorem. It says that we can always add new operators to Σ_0, and enforce new identities, in a canonical way. While the proof of the theorem is constructive[3], it is not practical. $T_{\Sigma,E}(\boldsymbol{A})$ may be a superset of $\boldsymbol{A}$, or may be a homomorphic image, or may simply collapse to a trivial algebra with a single element. The theorem only tells us that $T_{\Sigma,E}(\boldsymbol{A})$ exists and is unique. We can use the theorem to add a difference operation to any semiring: all we need is to choose what identities we want difference to satisfy. For example, assume we choose the ring identities:[4]

$$x - x = \mathbf{0} \tag{6}$$

$$x + (y - z) = (x + y) - z \tag{7}$$

Then $T_{\Sigma_s,E}(\boldsymbol{S})$ is the ring freely generated by the semiring $\boldsymbol{S}$. For example, if we apply this construction to the natural numbers $(\mathbb{N}, +, \cdot, 0, 1)$, then $T_{\Sigma_{sm},E}(\mathbb{N})$ is isomorphic to $\mathbb{Z}$. But the freely generated ring can sometimes be trivial, as can seen from this lemma.

▶ **Lemma 2.** *Let $\boldsymbol{S}$ be a semiring where addition is idempotent, $x + x = x$. Then, if $\boldsymbol{R}$ is any ring such that there exists a homomorphism $h : \boldsymbol{S} \to \boldsymbol{R}$, then $\boldsymbol{R}$ is the trivial ring. In particular, the ring freely generated by $\boldsymbol{S}$ is trivial.*

Proof. In $\boldsymbol{S}$ it holds that $1 + 1 = 1$, therefore $1 + 1 = h(1) + h(1) = h(1) = 1$ holds in $\boldsymbol{R}$. By adding -1 to both sides we obtain $0 = 1$ in $\boldsymbol{R}$, hence $\boldsymbol{R}$ is trivial. ◀

The take-away is that, in order to define a difference operation "freely", we need to choose carefully what identities we want it to satisfy. If we insist on the ring identities, then we may end up with the trivial ring. Yet, Sec. 2 showed useful examples of difference operations that were not rings. We consider next an alternative way to defined difference, by using the natural order.

---

[2] For a formal definition of what it means for an algebra $\boldsymbol{A}$ to satisfy (e_1, e_2) we refer to [6].

[3] $T_{\Sigma,E}(\boldsymbol{A})$ is defined as $T_{\Sigma \cup A} / \equiv_{E \cup E_{\boldsymbol{A}}}$, where $\Sigma \cup A$ extends Σ with one nulary operator a for every constant $a \in A$, the set $E_{\boldsymbol{A}}$ consists of all grounded identities of the form $(f(a_1, \ldots, a_m), b)$ where $b = f^A(a_1, \ldots, a_m)$, and $\equiv_{E \cup E_{\boldsymbol{A}}}$ is the smallest congruence relation that contains $E_{\boldsymbol{A}}$ and all groundings of E

[4] Equations (6) and (7) imply that the operation $-x \stackrel{\text{def}}{=} \mathbf{0} - x$ is the additive inverse, because $x + (-x) = x + (\mathbf{0} - x) = (x + \mathbf{0}) - x = x - x = \mathbf{0}$.

4 Difference by Natural Order

Given a commutative monoid $\boldsymbol{M} = (M, +, 0)$ the *natural preorder*, $\preceq$, is defined as follows:

$$a \preceq b \quad \text{if} \quad \exists z : a + z = b \tag{8}$$

Then, $\preceq$ is transitive and reflexive, thus a preorder. When it is antisymmetric then it is called the *natural order* of $\boldsymbol{M}$, and $\boldsymbol{M}$ is called a naturally ordered monoid. In a naturally ordered monoid 0 is the smallest element: $0 \preceq x$ for all $x \in M$. For example, $(\mathbb{N}, +, 0)$ is naturally ordered, while $(\mathbb{Z}, +, 0)$ is not. Similarly, a *naturally ordered semiring* is a semiring $(S, +, \cdot, 0, 1)$ where the additive monoid $(S, +, 0)$ is naturally ordered. Many semirings are naturally ordered, so it makes sense to try to use the natural order to define difference.

4.1 Monus in Naturally Ordered Monoids

▶ **Definition 3.** *Let $(M, +, 0)$ be a naturally ordered monoid. Given two elements $a, b \in M$, consider the set of all elements $z \in M$ s.t. $a + z \succeq b$. If this set has a minimal element c, then we define:*

$$b \dot{-} a \stackrel{\text{def}}{=} \min\{z \mid a + z \succeq b\} \tag{9}$$

Amer [3] called $(M, +, 0)$ a *Commutative Monoid with Monus*, or CMM, if it is naturally ordered and $b \dot{-} a$ exists for all $a, b \in M$. The monoid of natural numbers $(\mathbb{N}, +, 0)$ is a CMM, and its monus operation given by (9) is the same as monus in equation (3). We give two examples of classes of CMMs.

▶ **Definition 4.** *Call a naturally ordered monoid $(M, +, 0)$* complete and distributive *if $\preceq$ forms a complete, distributive lattice, and $+$ distributes over with $\bigwedge$, in other words, $x + \bigwedge\{z \mid z \in A\} = \bigwedge\{x + z \mid z \in A\}$, for any set $A \subseteq M$.*

Every complete, distributive monoid is a CMM, and its monus operation is:

$$b \dot{-} a \stackrel{\text{def}}{=} \bigwedge\{z \mid a + z \succeq b\} \tag{10}$$

We check that (10) satisfies Definition 3, and for that we need to show that $\bigwedge\{z \mid a + z \succeq b\}$ is the minimum element of the set $\{z \mid a + z \succeq b\}$, in other words we need to show that $a + \bigwedge\{z \mid a + z \succeq b\} \succeq b$. This follows from the fact that $+$ distributes over $\bigwedge$: $a + \bigwedge\{z \mid a + z \succeq b\} = \bigwedge\{a + z \mid a + z \succeq b\}$ and the latter is obviously $\succeq b$.

One example of a complete, distributive monoid is $(\mathbb{N}, +, 0)$, where monus (10) is the same as Eq. (3). Another example is a Boolean algebra $(2^{\Omega}, \cup, \emptyset)$, where (10) is set difference $b \setminus a$.

Let $(M, \preceq)$ be a complete total order, meaning that $x \preceq y$ or $y \preceq x$ for all $x, y \in M$, and $\bigwedge A$ exists for all $A \subseteq M$. Then $(M, \preceq)$ is a distributive lattice[5], and $(M, \vee, \perp)$ is a CMM, where $x \vee y \stackrel{\text{def}}{=} \bigwedge\{z \mid x \preceq z, y \preceq z\}$, $\perp \stackrel{\text{def}}{=} \bigwedge M$ is the smallest element of M. The monus operation in (10) further simplifies to:

$$b \dot{-} a = \begin{cases} 0 & \text{when } b \preceq a \\ b & \text{when } b \succ a \end{cases} \tag{11}$$

[5] We prove the identity $x \vee \bigwedge\{z \mid z \in A\} = \bigwedge\{x \vee z \mid z \in A\}$ by considering two cases. If there exists $y \in A$, $x \succeq y$, then both sides are equal to x. Assuming $x \preceq z$ for all $z \in A$ we have $\bigwedge\{x \vee z \mid z \in A\} = \bigwedge\{z \mid z \in A\}$, and the identity follows immediately.

An example of such a CMM is $([0, \infty], \min, \infty)$, where monus (11) is the same as Eq. (4). [6]

We caution that not every naturally ordered monoid a CMM. For a counterexample, consider the non-distributive lattice M_3 with elements $0 < a, b, c < 1$, where $a \vee b = a \vee c = b \vee c = 1$ and $a \wedge b = a \wedge c = b \wedge c = 0$. Then $(M_3, \vee, 0)$ is naturally ordered, but it is not a CMM because the set of z's for which $a \vee z \geq 1$ is $\{b, c, 1\}$ and it has no smallest element.

4.2 Monus as an Equational Class

In 1965 [5] Bosbach proved a remarkable result, which implies that CMMs form an equational class.[7] Amer [3] presented a simplified statement of Bosbach' result, and claimed (without proof) that CMMs are precisely the equational class defined by the axioms $(A1 - A4)$ below. We will show here Amer's identities, and give a simplified proof of Bosbach's result.

Amer's identities are the following:

$$\begin{aligned}
&(A1) && a + (b \dot{-} a) = b + (a \dot{-} b)\\
&(A2) && (a \dot{-} b) \dot{-} c = a \dot{-} (b + c)\\
&(A3) && a \dot{-} a = 0\\
&(A4) && 0 \dot{-} a = 0
\end{aligned}$$

▶ **Theorem 5.** *[3, 5] A commutative monoid $\boldsymbol{M} = (M, +, 0)$ is a CMM iff there exists an operation $\dot{-}$ that satisfies $(A1) - (A4)$. In particular, the class of CMMs is the restriction to the signature Σ_m of an equational class of $\Sigma_m \cup \{\dot{-}\}$ algebras.*

Proof. Assume first that $\boldsymbol{M}$ is a CMM, and let $b \dot{-} a$ be given as in Definition 3. We prove that $\boldsymbol{M}$ satisfies $(A1)$ and $(A2)$, and leave it up to the reader to check $(A3 - A4)$. By assumption $\boldsymbol{M}$ is naturally ordered, with partial order $\preceq$. Identity $(A1)$ follows from these implications:

$$\begin{aligned}
a + (b \dot{-} a) &\succeq b && \text{Definition of } b \dot{-} a\\
\exists z, a + (b \dot{-} a) &= b + z && \text{Definition of } \succeq && (12)\\
a &\preceq b + z && \text{From } a \preceq a + (b \dot{-} a)\\
a \dot{-} b &\preceq z && \text{Definition of } a \dot{-} b && (13)\\
a + (b \dot{-} a) &\succeq b + (a \dot{-} b) && \text{From (12) and (13)}
\end{aligned}$$

The opposite inequality $a + (b \dot{-} a) \preceq b + (a \dot{-} b)$ is proven similarly, and this implies $(A1)$. For $(A2)$, we start by proving $(a \dot{-} b) \dot{-} c \succeq a \dot{-} (b + c)$. By definition $a \dot{-} (b + c)$ is the smallest z satisfying the condition $z + (b + c) \succeq a$, hence it suffices to prove that $(a \dot{-} b) \dot{-} c$ also satisfies this condition. This follows from:

$$((a \dot{-} b) \dot{-} c) + b + c = (((a \dot{-} b) \dot{-} c) + c) + b \succeq (a \dot{-} b) + b \succeq a$$

[6] A small variation is the monoid $(\mathbb{R} \cup \{\infty\}, \min, \infty)$. This is also a CMM, with monus defined by the same Eq. (4). In both these CMM's the natural order $\preceq$ is the reverse of the standard order, i.e. $x \preceq y$ if $x \geq y$. This probably confused the authors of [9], who claimed incorrectly in Example 4 that $(\mathbb{R} \cup \{\infty\}, \min, \infty)$ is not a CMM.

[7] Bosbach considered naturally ordered semigroups $(M, +)$ (i.e. monoids without 0, and not necessarily commutative), which he called *holoids*, and defined a *complemented holoid* (komplementäres Holoid) to be a holoid where $b \dot{-} a$ given as in Definition 3 exists for all $a, b \in M$. Then he proved that the set of complemented holoids is an equational class, defined by just four identities over the signature $\{+, \dot{-}\}$. He further proved that every complemented holoid $(M, +, \dot{-})$ is also a commutative monoid, meaning that $+$ is commutative and has an identity, namely $0 \stackrel{\text{def}}{=} x \dot{-} x$, which, he showed, is independent on the choice of x.

Similarly, for the opposite inequality $a \dot{-} (b + c) \succeq (a \dot{-} b) \dot{-} c$, it suffices to prove that $(a \dot{-} (b + c)) + c \succeq a \dot{-} b$, and, for that, it suffices to prove that $((a \dot{-} (b + c)) + c) + b \succeq a$. This follows immediately by writing the inequality as $(a \dot{-} (b + c)) + (b + c) \succeq a$.

We now prove the interesting part: if $(M, +, 0)$ is a commutative monoid and admits a difference operation $\dot{-}$ that satisfies $(A1 - A4)$, then M is a CMM. Recall that $a \preceq b$ is defined as: $\exists x, a + x = b$. We first establish a simple property:

$$(P1): \qquad a \preceq b \text{ iff } a \dot{-} b = 0$$

In one direction, if $a \preceq b$ then $a \dot{-} b = a \dot{-} (a + x) = (a \dot{-} a) \dot{-} x$ (by $(A2)$) $= 0 \dot{-} x = 0$ (by $(A3), (A4)$). In the other direction, we have $b = b + 0 = b + (a \dot{-} b) = a + (b \dot{-} a)$ (by $(A1)$) which implies $b \succeq a$ by definition.

$(P1)$ implies that $(M, +, 0)$ is naturally ordered. Indeed, if both $a \dot{-} b = 0$ and $b \dot{-} a = 0$ hold, then, by $(A1)$: $a = a + 0 = a + (b \dot{-} a) = b + (a \dot{-} b) = b + 0 = b$.

It remains to prove that $(M, +, 0)$ is a CMM. For this purpose we prove a second property:

$$(P2): \qquad b \preceq a + z \text{ iff } b \dot{-} a \preceq z$$

In one direction, we use $(P1)$ and $b \preceq a + z$ to derive $0 = b \dot{-} (a + z) = (b \dot{-} a) \dot{-} z$ (by $(A2)$) and we use again $(P1)$ to conclude $b \dot{-} a \preceq z$. In the other direction, we add a to both sides of $z \succeq b \dot{-} a$ and derive $a + z \succeq a + (b \dot{-} a) = b + (a \dot{-} b)$ (by $(A1)$) $\succeq b$ (by definition of $\succeq$). We prove now that, for any two elements $a, b \in M$, the operation $b \dot{-} a$ satisfies the condition in Definition 3. On one hand $a + (b \dot{-} a) = b + (a \dot{-} b) \succeq b$. On the other hand, if z also satisfies $a + z \succeq b$, then by $(P2)$ we have $b \dot{-} a \preceq z$. Thus, $b \dot{-} a$ is the smallest element z with this property, as required. ◀

4.3 Monus in Naturally Ordered Semirings

Geerts and Poggi [9] extended CMMs from monoids to semirings. They defined an *m-semiring* to be a semiring $\boldsymbol{S} = (S, +, \cdot, 0, 1)$ where the monoid $(S, +, 0)$ is a CMM; i.e. $\boldsymbol{S}$ is naturally ordered and $b \dot{-} a \stackrel{\text{def}}{=} \min\{z \mid a + z \succeq b\}$ exists for elements $a, b \in S$. In particular, monus satisfies identities $(A1 - A4)$. One drawback of m-semirings is that monus is defined using only the additive monoid $(S, +, 0)$, ignoring the multiplicative operator: this creates problems, as we see next.

Tannen [4] considered the use of m-semirings as annotations of relations, and asked whether the identities $(A1 - A4)$ are sufficient to capture identities of the relational algebra. In particular, they considered the following relational algebra identity:

$$(R - S) \bowtie T = R \bowtie T - S \bowtie T \tag{14}$$

In order for (14) to hold when the relations R, S, T are annotated with values from an m-semiring, the semiring must satisfy the following identity:

$$(A5) \qquad (b \dot{-} a) \cdot c = b \cdot c \dot{-} a \cdot c$$

However, $(A5)$ does not hold in general.[8] For example, it holds in $(\mathbb{N}, +, \cdot, 0, 1, \dot{-})$, and, by extension, in the m-semiring of polynomials[9] $(\mathbb{N}[X], +, \cdot, 0, 1, \dot{-})$, but it fails in the following semiring (adapted from [4]): $\boldsymbol{S} = (\{0, \frac{1}{2}, 1\}, \vee, \wedge, 0, 1)$, where monus is given by equation (11). Here $(1 \dot{-} \frac{1}{2}) \wedge 0 = 1 \wedge 0 = 0$ while $(1 \wedge \frac{1}{2}) \dot{-} (\frac{1}{2} \wedge \frac{1}{2}) = \frac{1}{2} \dot{-} \frac{1}{2} = 0$, thus $(A5)$ fails.

[8] Inequality does hold in one direction, namely $(b \dot{-} a) \cdot c \succeq b \cdot c \dot{-} a \cdot c$, because $a + (b \dot{-} a) \succeq b$ implies $a \cdot c + (b \dot{-} a) \cdot c \succeq b \cdot c$, and, by property $(P2)$, we obtain $b \dot{-} a \succeq b \cdot c \dot{-} a \cdot c$.

[9] Monus on polynomials is defined by applying it to each monomial. For example $5x \dot{-} 2x = 3x$, and $2x \dot{-} 5x = 0$, and $x \dot{-} y = x$.

An intriguing observation by Geerts and Poggi [9] is that, although $(\mathbb{N}[X], +, \cdot, 0, 1, \dot{-})$ is an m-semiring, it is not the freely generated m-semiring.[10] The reason is that monus in $(\mathbb{N}[X], +, \cdot, 0, 1, \dot{-})$ is not defined "freely". For example, consider $\boldsymbol{B} = (\{0,1\}, \vee, \wedge, 0, 1, \dot{-})$, which is an m-semiring where $x \dot{-} y = x \wedge \neg y$. Given two variables, $X = \{x, y\}$, and the function $h : \{x, y\} \to \mathbb{B}$, $h(x) = h(y) = 1$, its unique extension $\bar{h} : \mathbb{N}[X] \to \mathbb{B}$ to a semiring homomorphism fails to be a homomorphism of m-semirings, because, on one hand, $\bar{h}(y \dot{-} x) = \bar{h}(y) = 1$, while $\bar{h}(y) \dot{-} \bar{h}(x) = 1 \dot{-} 1 = 0$.

In summary, there are two pieces of bad news for defining difference using the natural order. On one hand, m-semirings do not satisfy $(A5)$ in general, which implies that some optimizations performed by a traditional query optimizer may fail when the relations are interpreted over m-semirings. On the other hand, if we restrict only to m-semirings that satisfy $(A5)$, then we no longer have a familiar freely generated semiring. Tannen [4] suggested a deeper investigation of the freely generated m-semiring satisfying $(A5)$. However, it turns out that this is only a partial solution: such semirings will ensure that the optimization (14) is sound, but may fail other optimization rules. In fact, the only way to ensure that all relational algebra identities that are valid over bag semantics remain valid over m-semirings is to require the latter to satisfy all identities satisfied by $(\mathbb{N}, +, \cdot, 0, 1, \dot{-})$. This set of identities is co-r.e. complete, and, therefore not finitely axiomatizable. We defer the proof to Appendix A.

4.4 Natural Order and Equational Classes

Bosbach's result [5] that Commutative Monoid with Monus (CMMs) form an equational class is surprising, because it is not obvious how to define a natural order using only algebraic operations and identities. Here we investigate whether such a definition is possible. More precisely, we ask: do the naturally ordered monoids form an equational class? Similarly, do the naturally ordered semirings form an equational class? Bosbach's result does not answer this question, because it only concerns a subclass of naturally ordered monoids (semirings), namely those where monus exists. We answer the general question both negatively and positively!

First, the negative answer:

▶ **Lemma 6.** *The naturally ordered monoids are not an equational class.*
The naturally ordered semirings are not an equational class.

Proof. We will use the following known fact, which is also easy to check directly: for any signature Σ and surjective homomorphism $h : \boldsymbol{A} \to \boldsymbol{B}$, if $\boldsymbol{A}$ satisfies a set of identities E, then so does $\boldsymbol{B}$. We will fix Σ to be either $\Sigma_m = \{+, 0\}$ (the signature of monoids (1)) or $\Sigma_s = \{+, \cdot, 0, 1\}$ (the signature of semirings (2)). Assume that the set of ordered monoids (or semirings) are the equational class defined by a set of Σ_m-identities E. Consider the monoid $\boldsymbol{A} \stackrel{\text{def}}{=} \mathbb{N} \times \mathbb{N}$ where $+$ is defined component-wise, $(a, b) + (c, d) \stackrel{\text{def}}{=} (a + c, b + d)$. $\boldsymbol{A}$ is naturally ordered (this is easily verified), thus, by our assumption satisfies the identities E. Consider the homomorphism $h : \boldsymbol{A} \to \boldsymbol{Z}$, $h(a, b) \stackrel{\text{def}}{=} a - b$. Since h is surjective, we conclude that $\boldsymbol{Z}$ also satisfies E, thus, by our assumption, it is naturally ordered, which is a contradiction, proving the lemma for monoids. To prove the lemma for semirings it suffices to define the multiplication operator on $\mathbb{N} \times \mathbb{N}$ as $(a, b) \cdot (c, d) \stackrel{\text{def}}{=} (ac + bd, ad + bc)$; then h is also a homomorphism of semirings. ◀

[10] This follows immediately from the fact that $\mathbb{N}[X]$ satisfies $(A5)$, while some m-semirings don't.

However, naturally ordered monoids and semirings *are* an equational class, if we are allowed to use an additional operator, denote it $\vee$. More precisely, we extend the signature of monoids to $\Sigma_m \cup \{\vee\} = \{+, 0, \vee\}$ (Σ_m was defined in Equations (1)), and extend similarly the signature of semirings to $\Sigma_s \cup \{\vee\}$. Let E_m be the following set of $(\Sigma_m \cup \{\vee\})$-identities:

- The monoid identities.
- Semi-lattice identities for $\vee$: associativity, commutativity, and idempotence. Recall that these identities define a partial order $\leq$ by $x \leq y$ if $x \vee y = y$.
- The identity: $x \vee (x + y) = x + y$.

We define similarly the set of $(\Sigma_s \cup \{\vee\})$-identities E_s by extending the semiring identities with those for $\vee$ shown above. We prove the following result, which appears to be new:

▶ **Theorem 7.** *The class of ordered monoids is equal to the class of* $(\Sigma_m \cup \{\vee\}), E_m$*-algebras restricted to the monoid operators* Σ_m*.*

The class of ordered semirings is equal to the class of $(\Sigma_s \cup \{\vee\}), E_s$*-algebras restricted to the semiring operators* Σ_s*.*

We prove only the first statement; the second is similar. In one direction, if $(M, +, 0, \vee)$ is a $(\Sigma_m \cup \{\vee\}), E_m$-algebra, then we show that $(M, +, 0)$ is naturally ordered. Let $\leq$ be the partial ordered defined by $\vee$ (thus $a \leq b$ if $a \vee b = b$) and let $\preceq$ be the natural preorder in Eq. (8). We notice that $\preceq$ implies $\leq$, because $a + z = b$ implies $a \vee b = a \vee (a + z) = a + z = b$. Therefore $\preceq$ is antisymmetric (because $\leq$ is a partial order), proving the claim.

For the opposite direction, consider a naturally ordered monoid $(M, +, 0)$, and let $\preceq$ be its natural order. By Szpilrajn's extension theorem [17], there exists a total order $\leq$ that is an extension of $\preceq$, i.e. $a \preceq b$ implies $a \leq b$, and $\leq$ is a total order (a.k.a. linear order). Define $x \vee y \stackrel{\text{def}}{=} \max(x, y)$. We check that $(M, +, 0, \vee)$ is a $(\Sigma_m \cup \{\vee\}), E_m$-algebra. The only non-trivial identity is $x \vee (x + y) = x + y$. This follows from the fact that $x \preceq x + y$ (by the definition of the natural order $\preceq$), which implies $x \leq x + y$, proving that $x \vee (x + y) = x + y$.

5 Difference by Construction

A third approach to defining a difference operator in a semiring $\boldsymbol{S}$ is to construct from $\boldsymbol{S}$ some semiring $\hat{\boldsymbol{S}}$ which has a difference operator. The intuition comes from the standard method of defining $\mathbb{Z}$ from $\mathbb{N}$: first define a semiring on the product $\mathbb{N} \times \mathbb{N}$ by setting $(x, x') + (y, y') \stackrel{\text{def}}{=} (x + x', y + y')$ and $(x, x') \cdot (y, y') \stackrel{\text{def}}{=} (xy + x'y', x'y + xy')$, then consider the equivalence classes $(\mathbb{N} \times \mathbb{N})/\equiv$, where $(x, x') \equiv (y, y')$ if $x + y' = x' + y$. We explore in this section what happens if we apply a similar construction to some arbitrary semiring.

5.1 A Product Construction

▶ **Definition 8.** *Fix a semiring* $\boldsymbol{S}$*. We define the* Σ_{sm}*-algebra* $\hat{\boldsymbol{S}} \stackrel{\text{def}}{=} (S \times S, +, \cdot, (0,0), (1,0), -)$*, where the operations are:*

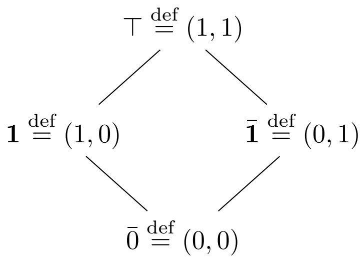
$$(x, x') + (y, y') \stackrel{def}{=} (x + y, x' + y') \qquad (x, x') \cdot (y, y') \stackrel{def}{=} (xy + x'y', xy' + x'y)$$

$$(x, x') - (y, y') \stackrel{def}{=} (x + y', x' + y) \qquad \mathbf{0} \stackrel{def}{=} (0, 0) \quad \mathbf{1} \stackrel{def}{=} (1, 0)$$

(where xy stands for $x \cdot y$, etc.)

We prove:

▶ **Lemma 9.** $\hat{\boldsymbol{S}}$ *is a semiring that also satisfies identity* (7)*, and* $(A2), (A5)$*.*



$\times$	$\mathbf{0}$	$\mathbf{1}$	$\bar{\mathbf{1}}$	$\top$
$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
$\mathbf{1}$	$\mathbf{0}$	$\mathbf{1}$	$\bar{\mathbf{1}}$	$\top$
$\bar{\mathbf{1}}$	$\mathbf{0}$	$\bar{\mathbf{1}}$	$\mathbf{1}$	$\top$
$\top$	$\mathbf{0}$	$\top$	$\top$	$\top$

$-$	$\mathbf{0}$	$\mathbf{1}$	$\bar{\mathbf{1}}$	$\top$
$\mathbf{0}$	$\mathbf{0}$	$\bar{\mathbf{1}}$	$\mathbf{1}$	$\top$
$\mathbf{1}$	$\mathbf{1}$	$\top$	$\mathbf{1}$	$\top$
$\bar{\mathbf{1}}$	$\bar{\mathbf{1}}$	$\bar{\mathbf{1}}$	$\top$	$\top$
$\top$	$\top$	$\top$	$\top$	$\top$

Figure 1 The semiring $\hat{\mathbb{B}}$, as per Definition 8. It is naturally ordered, with a Hasse diagram shown on the left. Addition $+$ is the LUB of the order relation, while multiplication and minus are shown in the table.

Proof. The proof that $\hat{\boldsymbol{S}}$ is a semiring is immediate, and omitted. We check identity (7) and axioms $(A2), (A5)$ directly:

$$\begin{aligned}
\text{Eq (7)}: \qquad & \hat{x} + (\hat{y} - \hat{z}) = (x, x') + ((y, y') - (z, z')) = (x + y + z', x' + y' + z) \\
& (\hat{x} + \hat{y}) - \hat{z} = ((x, x') + (y, y')) - (z, z') = (x + y + z', x' + y + z)
\end{aligned}$$

$$\begin{aligned}
(A2): \qquad (\hat{x} - \hat{y}) - \hat{z} =& (((x, x') - (y, y')) - (z, z')) \\
=& (x + y' + z', x' + y + z) \\
=& ((x, x') - ((y, y') + (z, z'))) = \hat{x} - (\hat{y} + \hat{z})
\end{aligned}$$

$$\begin{aligned}
(A5): \qquad (\hat{y} - \hat{x}) \cdot \hat{z} =& (((y, y') - (x, x')) \cdot (z, z')) \\
=& ((y + x')z + (y' + x)z', (y + x')z' + (y' + x)z) \\
=& ((yz + y'z') + (x'z + xz'), (yz' + y'z) + (x'z' + xz)) \\
=& (yz + y'z', yz' + y'z) - (x'z' + xz, x'z + xz') \\
=& (y, y') \cdot (z, z') - (x, x') \cdot (z, z') = \hat{y} \cdot \hat{z} - \hat{x} \cdot \hat{z}
\end{aligned}$$
◀

In general $\hat{\boldsymbol{S}}$ does not satisfy $(A1)$ and $(A4)$, but this is not a problem, for example, they don't hold in any ring either.

▶ **Example 10.** If $\mathbb{B}$ is the Boolean semiring, then $\hat{\mathbb{B}}$ consists of four elements, which we denote as $\mathbf{0}, \mathbf{1}, \bar{\mathbf{1}}, \top$, and are shown in Figure 1. The elements can be interpreted as follows: $\mathbf{0}$ = false, $\mathbf{1}$ = positive, $\bar{\mathbf{1}}$ =negative, and $\top$ = over specified (both positive and negative).

More generally, let's apply this construction to a Boolean algebra $\boldsymbol{S} = (2^\Omega, \cup, \cap, \emptyset, \Omega)$. The elements of $\hat{\boldsymbol{S}}$ are pairs of sets (A, B), and can be best viewed as functions $\nu : \Omega \to \hat{\mathbb{B}}$ mapping the elements in the four sets $\Omega \setminus (A \cup B), A \setminus B, B \setminus A, A \cap B$ to $\mathbf{0}, \mathbf{1}, \bar{\mathbf{1}}$, and $\top$ respectively. For example the pair $(\{a, b, d, e, f\}, \{c, d, f\})$ in $\hat{\boldsymbol{S}}$ can be denote more friendly as $\{a, b, \bar{c}, d\bar{d}, e, f\bar{f}\}$, meaning that the elements a, b, e are positive (inserted), c is negative (removed), while d and f were over specified (both inserted and removed).

However, so far we only used the first step from the standard construction of the integers from the natural numbers: taking the cross product. In many cases we need the second step as well: taking the quotient w.r.t. some congruence relation. Generalizing the $\mathbb{N}$-to-$\mathbb{Z}$ construction, our first attempt is to define $\equiv$ as the smallest a congruence relation on $\bar{\boldsymbol{S}}$ satisfying:

$$x + y' = x' + y \implies (x, x') \equiv (y, y') \tag{15}$$

The problem with this definition is that the quotient semiring may become trivial: for example, both $\hat{\mathbb{B}}/\equiv$ and $\widehat{\mathsf{Trop}}/\equiv$ are trivial. To see this, notice that both semirings have an absorptive element $\top$ satisfying $\top + x = \top$, and therefore $(x, x') \equiv (\top, \top)$ for all (x, x'). Thus, we will not consider definition (15), and instead will define a congruence on $\hat{\boldsymbol{S}}$ by using a semiring ideal.

5.2 Ideals in Semirings

Recall that an *ideal* in a ring $\boldsymbol{R}$ is a subset $I \subseteq R$ s.t. $x, y \in I$ implies $x + y \in I$, and $x \in I, u \in R$ implies $u \cdot x \in I$. The equivalence relation $x \equiv_I y$ defined by $x - y \in I$ is a congruence, and the set $R/\equiv_I$ is called the *quotient ring*. We generalize these concepts from rings to semirings.

▶ **Definition 11.** *An* ideal *in a semiring* $\boldsymbol{S}$ *is a set* $I \subseteq S$ *satisfying* $u, v \in I \Rightarrow u + v \in I$*, and* $u \in I, x \in S \Rightarrow u \cdot x \in I$*. Given an ideal* I*, we define two congruence relations:*

$$x \equiv_I y \quad \text{if} \quad \exists u, v \in I, x + u = y + v \tag{16}$$

$$x \cong_I y \quad \text{if} \quad \{(a,b) \mid a, b \in S, ax + b \in I\} = \{(a,b) \mid a, b \in S, ay + b \in I\} \tag{17}$$

It can be checked immediately that both $\equiv_I$ and $\cong_I$ are congruence relations.[11] Therefore both quotients $\boldsymbol{S}/\equiv_I$ and $\boldsymbol{S}/\cong_I$ are semirings, and the canonical mappings $\boldsymbol{S} \to \boldsymbol{S}/\equiv_I$ and $\boldsymbol{S} \to \boldsymbol{S}/\cong_I$ are semiring homomorphisms. Both the definition of an ideal, and of the congruence relation $\equiv_I$ appear in the literature [14–16], and are extensively covered in [10]. The definition of the congruence relation $\cong_I$ appears to be novel.

In a ring, both $\equiv_I, \cong_I$ are equal, and are the same as the standard congruence relation defined by the ideal I; moreover, the congruence class $0/\cong_I$ is precisely the ideal I. For a semiring $\boldsymbol{S}$ and any set A s.t. $0 \in A \subseteq S$, we define the *closure* as $cl(A) \stackrel{\text{def}}{=} \{x \mid \exists a \in A, a + x \in A\}$. We prove:

▶ **Lemma 12.** *In any semiring* $\boldsymbol{S}$*, (1)* $I \subseteq cl(I) = 0/\equiv_I$ *and (2) any congruence class* $x/\cong_I$ *is either a subset of* I*, or disjoint from* I*. In particular,* $0/\cong_I \subseteq I$*.*

Proof. The first statement is immediate, and is well known in the literature. For the second statement, we prove that if $x/\cong_I$ contains some element $u \in I$, then it is a subset of I. Assume $x \cong_I u$ and $u \in I$. Then $x \in I$ follows by setting $a = 1, b = 0$ in (17): then $au + b = u \in I$, and therefore $ax + b = x \in I$. ◀

Henriksen [14] called an ideal I a *k-ideal* if $I = cl(I)$; see also [2, 15, 16]. For example, in the semiring of natural numbers $(\mathbb{N}, +, \cdot, 0, 1)$, the set $I = \{6k + 8\ell \mid k, \ell \in \mathbb{N}\}$ is an ideal $I = \{0, 6, 8, 12, 14, 16, \ldots\}$ which is not a k-ideal: $cl(I)$ is the set of all even numbers, and $cl(I) - I = \{2, 4, 10\}$. We prove:

▶ **Lemma 13.** *The following statements are equivalent: (1)* I *is a k-ideal, (2)* $0/\equiv_I = I$*, (3)* $0/\cong_I = I$*.*

Proof. The equivalence of (1) and (2) is immediate, and was well known in the literature. To prove (1) $\Rightarrow$ (3) we show that for all $u \in I$, $u \cong_I 0$. Let $a, b \in S$ be such that $au + b \in I$; then $au \in I$ and, since $I = cl(I)$, we have $b \in I$, implying $a0 + b \in I$. Conversely, if $a0 + b \in I$, then $b \in I$ and we have $au + b \in I$ because $au \in I$. To prove (3) $\Rightarrow$ (1), assume $u \in I$ and $u + b = 1 \cdot u + b \in I$. Since $u \cong_I 0$ we also have $1 \cdot 0 + b \in I$, proving $b \in I$. ◀

[11] The rationale behind having the parameter a in (17), as opposed to fixing $a = 1$, is to ensure that $\cong_I$ is a congruence w.r.t. multiplication.

We return now to our product semiring $\hat{\boldsymbol{S}}$ in Definition 8, and define the following ideal in $\hat{\boldsymbol{S}}$.

$$\Delta \stackrel{\text{def}}{=} \{(x, x) \mid x \in S\} \tag{18}$$

It can be checked immediately that Δ is an ideal in $\hat{\boldsymbol{S}}$, which we call the *diagonal ideal.* In general, Δ is not a k-ideal, and, therefore, the congruences $\equiv_\Delta$ and $\cong_\Delta$ are distinct. One can check that each of them is also a congruence w.r.t. to the difference operator. We examine now the quotients $\boldsymbol{S}/\equiv_\Delta$ and $\boldsymbol{S}/\cong_\Delta$, starting with $\equiv_\Delta$.

▶ **Lemma 14.** *For any semiring $\boldsymbol{S}$, the Σ_{sm}-algebra $\hat{\boldsymbol{S}}/\equiv_\Delta$ is the ring freely generated by $\boldsymbol{S}$.*

Proof. By Lemma 9 $\hat{\boldsymbol{S}}$ satisfies the semiring identities and identity (7), and therefore so does $\hat{\boldsymbol{S}}/\cong_\Delta$. It remains to prove it satisfies identity (6): $\hat{x} - \hat{x} = (x, x') - (x, x') = (x + x', x + x') \in \Delta$, which implies $(\hat{x} - \hat{x})/\equiv_\Delta = \hat{0}/\equiv_\Delta$, as required. To prove that it is the freely generated ring, we check the diagram (5) from Theorem 1: given a ring $\boldsymbol{R}$ and a semiring homomorphism $h : \boldsymbol{S} \to \boldsymbol{R}$, first extend it to a Σ_{sm}-homomorphism $\hat{h} : \hat{S} \to \boldsymbol{R}$ by $\hat{h}(x, x') \stackrel{\text{def}}{=} h(x) - h(x')$, then observe that $(x, x') \equiv_\Delta (y, y')$ implies that there exists $(u, u) \in \Delta$ such that $(x + u, x' + u) = (y + u, y' + u)$, which implies $\hat{h}(x + u, x' + u) = h(x + u) - h(x' + u) = h(x) - h(x') = \hat{h}(y + u, y' + u) = h(y) - h(y')$, in other words $\hat{h}(x, x') = \hat{h}(y, y')$. Therefore, we can uniquely extend $\hat{h} : \hat{\boldsymbol{S}} \to \boldsymbol{R}$ to $\hat{\boldsymbol{S}}/\equiv_\Delta \to \boldsymbol{R}$. This completes the proof of the lemma. ◀

The lemma gives us a constructive way to obtain the ring freely generated by the semiring $\boldsymbol{S}$, but, as we saw in Lemma 2, the freely generated ring can sometimes be trivial. This justifies exploring the second alternative for our construction: $\hat{\boldsymbol{S}}/\cong_\Delta$. We prove:

▶ **Lemma 15.** *For any semiring $\boldsymbol{S}$, $\hat{\boldsymbol{S}}/\cong_\Delta$ is a semiring that satisfies the identities* (7)*, and* $(A2), (A5)$. *Moreover: (1) the mapping $\eta : \boldsymbol{S} \to \hat{\boldsymbol{S}}/\cong_\Delta$, $\eta(x) \stackrel{def}{=} (x, 0)/\cong_\Delta$ is an injective homomorphism, and (2) the mapping $x \mapsto (0, x)/\cong_\Delta$ is injective (but not a homomorphism).*

Proof. By Lemma 9 $\hat{\boldsymbol{S}}$ satisfies the identities (7), and $(A2), (A5)$, therefore so does $\hat{\boldsymbol{S}}/\cong_\Delta$. We prove (1). It is straightforward to check that η is a homomorphism, we prove that it is injective. Assume $(x, 0) \cong_\Delta (y, 0)$ and set $\hat{a} = (1, 0)$, $\hat{b} = (0, x)$ in (17). Then $\hat{a} \cdot (x, 0) + \hat{b} = (x, x) \in \Delta$, and therefore we must have $\hat{a} \cdot (y, 0) + \hat{b} = (y, x) \in \Delta$, which implies $x = y$ as required. The proof of (2) is similar and omitted. ◀

The lemma proves that $\hat{\boldsymbol{S}}/\cong_\Delta$ contains two copies of $\boldsymbol{S}$:

- a copy $\{(x, 0)/\cong_I \mid x \in S\}$ of elements that we call *positive elements*, and
- a copy $\{(0, x)/\cong_I \mid x \in S\}$ of elements that we call *negative elements.*

Our last result proves that, in the case of the semirings $\mathbb{B}$ and Trop, then quotients $\hat{\mathbb{B}}/\cong_\Delta$ and $\widehat{\mathsf{Trop}}/\cong_\Delta$ consists precisely of the positive elements, negative elements, and *over determined* elements $\{(x, x)/\cong_I \mid x \in S\}$. We prove this separately for Boolean algebras and for Trop.

▶ **Lemma 16.** *$\hat{\mathbb{B}}/\cong_I$ is isomorphic to $\hat{\mathbb{B}}$ (shown in Fig. 1). As a consequence, if $\boldsymbol{S}$ is a Boolean algebra, as in Example 10, then $\hat{\boldsymbol{S}}/\cong_\Delta$ is isomorphic to $\hat{\boldsymbol{S}}$.*

Proof. It suffices to check that no two elements in $\hat{\mathbb{B}}$ are congruent, by checking six inequalities of the form $\hat{x} \cong_\Delta \hat{y}$. In each case we will show that there exists $\hat{b}$ such that $\hat{x} + \hat{b} \in \Delta$ and $\hat{y} + \hat{b} \notin \Delta$ (in other words, $\hat{a} = \mathbf{1}$ in all cases):

- $\mathbf{0} \ncong_\Delta \mathbf{1}$ and $\top \ncong_\Delta \mathbf{1}$: choose $\hat{b} = \bar{\mathbf{1}}$.
- $\mathbf{0} \ncong_\Delta \bar{\mathbf{1}}$ and $\top \ncong_\Delta \bar{\mathbf{1}}$: choose $\hat{b} = \mathbf{1}$.
- $\mathbf{1} \ncong_\Delta \bar{\mathbf{1}}$: choose $\hat{b} = \bar{\mathbf{1}}$.
- $\top \ncong_\Delta \mathbf{0}$: choose $\hat{b} = \mathbf{1}$. ◀

We prove next the same result for Trop, which we state in a slightly more general form. Recall that a *diod* is a semiring $\boldsymbol{S}$ where addition is idempotent. It can be shown that a diod is naturally ordered, and addition is the LUB, in other words $\boldsymbol{S} = (S, \vee, \cdot, 0, 1)$. Call a diod *strict* if its natural order $\preceq$ is total, and multiplication is cancelative, meaning that $a \cdot x = a \cdot y$ implies $x = y$ when $a \neq 0$; the semiring Trop is strict. We prove:

▶ **Lemma 17.** *Let $\boldsymbol{S}$ be a strict diod. Then $\hat{\boldsymbol{S}}/\cong_\Delta$ consists of the following congruence classes:*
- *Zero,* $(0,0)/\cong_\Delta = \{(0,0)\}$.
- *The* positive *elements* $(x,0)/\cong_\Delta = \{(x,z) \mid x \succ z\}$, *for* $x \in S$, $x \neq 0$.
- *The* negative *elements* $(0,x)/\cong_\Delta = \{(z,x) \mid z \prec x\}$, *for* $x \in S$, $x \neq 0$
- *The* over determined *elements* $(x,x)/\cong_\Delta = \{(x,x)\}$, *for* $x \in S$, $x \neq 0$.

Proof. Let $\sim$ be the following equivalence relation on $\hat{\boldsymbol{S}}$:

$$(x,y) \sim (u,v) \quad \text{if } ((y \prec x = u \succ v) \text{ or } (x \prec y = v \succ u) \text{ or } (x = y = u = v))$$

To prove the lemma we have to show that $\cong_\Delta = \sim$.

We start by showing $\cong_\Delta \subseteq \sim$, and for that we show that $(x,y) \not\sim (u,v)$ implies $(x,y) \not\cong_\Delta (u,v)$. Assume $(x,y) \not\sim (u,v)$. There are three cases. **Case 1**: $x \succ y$ and $u = v$. Setting $\hat{a} \stackrel{\text{def}}{=} (1,0)$ and $\hat{b} \stackrel{\text{def}}{=} (0,0)$ we have

$$\hat{a} \cdot (x,y) \bigvee \hat{b} = (x,y) \notin \Delta$$
$$\hat{a} \cdot (u,u) \bigvee \hat{b} = (u,u) \in \Delta$$

proving $(x,y) \not\cong_\Delta (u,u)$. **Case 2**: $x \succ y$ and $u \prec v$. Then we set $\hat{a} \stackrel{\text{def}}{=} (1,0)$, $\hat{b} \stackrel{\text{def}}{=} (v,0)$ and we have:

$$\hat{a} \cdot (x,y) \bigvee \hat{b} = (x,y) \bigvee (v,0) = (x \vee v, y) \notin \Delta \text{ because } x \vee v \succeq x \succ y$$
$$\hat{a} \cdot (u,v) \bigvee \hat{b} = (u,v) \bigvee (v,0) = (u \vee v, v) = (v,v) \in \Delta$$

Case 3: $x \prec y$ and $u = v$ is similar to case 1 and omitted.

Next, we prove that $(x,y) \sim (u,v)$ implies $(x,y) \cong_\Delta (u,v)$. For that it suffices to assume that $y \prec x = u \succ v$: the second case $x \prec y = v \succ u$ is symmetric, and the third case $x = y = u = v$ is trivial. Thus, it suffices to prove:

$$\text{If } y \prec x \succ v \text{ then } (x,y) \cong_\Delta (x,v) \tag{19}$$

We apply the definition of $\cong_\Delta$ in Eq. (17) to (x,y) and (x,v). It suffices to prove that $\hat{a} \cdot (x,y) \bigvee \hat{b} \in \Delta$, implies $\hat{a} \cdot (x,v) \bigvee \hat{b} \in \Delta$, for any two elements $\hat{a} = (a,a')$ and $\hat{b} = (b,b')$ in $\hat{S}$; the other direction of the implication is proven similarly and we will omit it. The semiring operations in $\hat{\boldsymbol{S}}$ were defined in Def. 8, and the condition $\hat{a} \cdot (x,y) \bigvee \hat{b} \in \Delta$ is equivalent to:

$$ax \vee a'y \vee b = ay \vee a'x \vee b' \tag{20}$$

which can be further rewritten two:

$$\begin{aligned} (ax \vee b) \vee a'y &= (a'x \vee b') \vee ay \\ A \vee a'y &= A' \vee ay \end{aligned} \tag{21}$$

where $A \stackrel{\text{def}}{=} ax \vee b \succeq ay$ and $A' \stackrel{\text{def}}{=} a'x \vee b' \succeq a'y$.

We claim that condition (21) and the fact that the semiring $\boldsymbol{S}$ is strict implies:

$$A = A' \tag{22}$$

The claim completes the proof because, equality (21) continues to hold if we replace y with v, because $A \succeq av$, $A' = A \succeq a'v$ and therefore $A \vee a'v = A$ and $A' \vee ay = A \vee ay = A$. Thus, it remains to prove that (21) and the fact that $\boldsymbol{S}$ is strict implies (22).

From (21) we derive:

$$A \vee a'y = A' \vee ay = (A \vee a'y) \vee (A' \vee ay) = A \vee A' \tag{23}$$

If $a = a' = 0$ then we immediately obtain $A = A'$.

Assume w.l.o.g. that $a \neq 0$. Since $\boldsymbol{S}$ is strict and $x \succ y$, we derive $ax \succ ay$ and therefore $A \succ ay$. It means that the four equal quantities in (23) are $\succ ay$. Since $\preceq$ is a total order, the least upper bound $A' \vee ay$ is either A' or $A' \vee ay = ay$: the latter impossible (because we proved that $A' \vee ay \succ ay$), therefore all four quantities in (23) are equal to A'.

In particular it holds that $A \vee a'y = A'$. We now consider two cases. When $a' = 0$, then we immediately derive $A = A'$. When $a' \neq 0$ then $A' \succ a'y$: since the least upper bound $A \vee a'y$ is either A or $a'y$, and it cannot be $a'y$, it that it is equal to A. Since all terms in (23) are equal to A', we conclude $A = A'$, as required. ◀

6 Discussion

We have examined three alternative ways to add difference to a semiring: by specifying the desired identities, by using the natural order, or by construction. The construction-based approach appears to be novel: we have only investigated a couple of options for the construction and proved only a few properties, leaving many open questions. For example, one open question is whether the class of Σ_{sm}-algebras of the form $\hat{\boldsymbol{S}}/\cong_\Delta$ is an equational class.

However, our investigation is far from complete. We mention here only one example that deserves further exploration. Grädel and Tannen [11] and later Dannert, Grädel, Naaf and Tannen [8] gave an interpretation to negative information by considering dual-indeterminate polynomials, $\mathbb{N}[X, \bar{X}]$. Such a polynomial has two kinds of variables (also called *provenance tokens*): positive variables x and negative variables $\bar{x}$. For example $3x + 4y\bar{z}^2 + \bar{x}\bar{z}$. They further assumed $x\bar{x} = 0$ for every variable x, which is equivalent to taking the quotient w.r.t. the ideal I generated by the monomials of the form $x\bar{x}$. One can naturally define a difference operator in this semiring as $f - g \stackrel{\text{def}}{=} f + \bar{g}$, where $\bar{g}$ is obtained by converting each variable from positive to negative and vice versa. Thus, $\mathbb{N}[X, \bar{X}]$ (or, more precisely, $\mathbb{N}[X, \bar{X}]/\equiv_I$) becomes an Σ_{sm}-algebra. What are the identities (in addition to the semiring identities) satisfied by this algebra? Is there any connection to the $\hat{\boldsymbol{S}}/\cong_\Delta$ construction that we explored in Sec. 5? We leave these questions for future work.

References

1 Mahmoud Abo Khamis, Hung Q. Ngo, Reinhard Pichler, Dan Suciu, and Yisu Remy Wang. Convergence of datalog over (pre-) semirings. In Leonid Libkin and Pablo Barceló, editors, *PODS '22: International Conference on Management of Data, Philadelphia, PA, USA, June 12 - 17, 2022*, pages 105–117. ACM, 2022. `doi:10.1145/3517804.3524140`.

2 P. J. Allen. A fundamental theorem of homomorphisms for semirings, 1969. URL: `https://api.semanticscholar.org/CorpusID:8723880`.

3 K Amer. Equationally complete classes of commutative monoids with monus. *Algebra Universalis*, 18:129–131, 1984. doi:10.1007/BF01182254.

4 Yael Amsterdamer, Daniel Deutch, and Val Tannen. On the limitations of provenance for queries with difference. In Peter Buneman and Juliana Freire, editors, *3rd Workshop on the Theory and Practice of Provenance, TaPP'11, Heraklion, Crete, Greece, June 20-21, 2011*. USENIX Association, 2011. URL: https://www.usenix.org/conference/tapp11/limitations-provenance-queries-difference.

5 Bruno Bosbach. Komplementäre Halbgruppen. Ein Beitrag zur instruktiven Idealtheorie kommutativer Halbgruppen. *Math. Annalen*, 161:279–295, 1965.

6 Stanley Burris and Hanamantagouda P. Sankappanavar. *A course in universal algebra*, volume 78 of *Graduate texts in mathematics*. Springer, 1981.

7 Shumo Chu, Brendan Murphy, Jared Roesch, Alvin Cheung, and Dan Suciu. Axiomatic foundations and algorithms for deciding semantic equivalences of SQL queries. *Proc. VLDB Endow.*, 11(11):1482–1495, 2018. doi:10.14778/3236187.3236200.

8 Katrin M. Dannert, Erich Grädel, Matthias Naaf, and Val Tannen. Semiring provenance for fixed-point logic. In Christel Baier and Jean Goubault-Larrecq, editors, *29th EACSL Annual Conference on Computer Science Logic, CSL 2021, January 25-28, 2021, Ljubljana, Slovenia (Virtual Conference)*, volume 183 of *LIPIcs*, pages 17:1–17:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.CSL.2021.17.

9 Floris Geerts and Antonella Poggi. On database query languages for k-relations. *J. Appl. Log.*, 8(2):173–185, 2010. doi:10.1016/j.jal.2009.09.001.

10 Jonathan S Golan. *Semirings and their Applications*. Springer Science & Business Media, 2013.

11 Erich Grädel and Val Tannen. Semiring provenance for first-order model checking. *CoRR*, abs/1712.01980, 2017. arXiv:1712.01980.

12 Todd J. Green, Zachary G. Ives, and Val Tannen. Reconcilable differences. In Ronald Fagin, editor, *Database Theory - ICDT 2009, 12th International Conference, St. Petersburg, Russia, March 23-25, 2009, Proceedings*, volume 361 of *ACM International Conference Proceeding Series*, pages 212–224. ACM, 2009. doi:10.1145/1514894.1514920.

13 Todd J. Green, Gregory Karvounarakis, and Val Tannen. Provenance semirings. In Leonid Libkin, editor, *Proceedings of the Twenty-Sixth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, June 11-13, 2007, Beijing, China*, pages 31–40. ACM, 2007. doi:10.1145/1265530.1265535.

14 M. Henriksen. Ideals in semirings with commutative addition. *Notices of the American Mathematical Society*, 6:321, 1958.

15 D. R. Latorre. A note on quotient semirings. *Proceedings of the American Mathematical Society*, 24(3):463–465, 1970. URL: http://www.jstor.org/stable/2037388.

16 M. Sen and Mahim Adhikari. On k-ideals of semirings. *International Journal of Mathematics and Mathematical Sciences*, 15, January 1992. doi:10.1155/S0161171292000437.

17 Edward Szpilrajn. Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae*, 16(1):386–389, 1930. URL: http://eudml.org/doc/212499.

A Monus and Query Optimization

A natural question emerging from [4] is to find a set of m-semiring identities E that is sound for query optimization. We show here that such a set is undecidable, and, in fact, it is co-r.e.-complete. More precisely, we seek a set of identities E such that, if two relational algebra (RA) queries are equivalent under standard semantics, then they remain equivalent over K-relations, when we use a semiring that satisfies E. In that case we say that the set E is *sound*. A trivial sound set is $E = \{0 = 1\}$, because it only holds in the trivial semiring, where all queries becomes equivalent. To avoid such degenerate solutions, we ask for a *minimal* sound set of identities E. We describe here this minimal set, by considering two flavors of soundness, depending on what semantics we adopt for RA expressions: set, or bag semantics. For example, if we use set semantics, then an optimizer could replace $R \cup R$ with R, and we need to add idempotence ($x + x = x$) to E to ensure soundness, but for bag semantics we don't need idempotence. In this section we prove the following. The minimal sound set of identities for bag semantics is $E_{\Sigma_{sm}}(\mathbb{N})$, i.e. the set of all identities satisfied by $(\mathbb{N}, +, \cdot, 0, 1, \dot{-})$: this set is co-r.e.-complete, and, thus, undecidable, and not finitely axiomatizable. Thus, adding just ($A5$), or any finite set of identities to the m-semirings identities is insufficient to ensure soundness. We also prove that the minimal sound set of identities for set semantics is $E_{\Sigma_{sm}}(\mathbb{B})$, the set of identities satisfied by $(\mathbb{B}, \vee, \wedge, 0, 1, \dot{-})$, which are the identities of Boolean algebras: there are well known finite axiomatizations for this set.

To state and prove this result formally we need a brief review of relational algebra and its interpretation over K-relations, based on [13].

Relational algebra[12], RA, consists of the six operators $\bowtie, \sigma, \Pi, \cup, -, \rho$. ($\rho$ is "renaming".) If we drop $-$, then it is called the *positive relational algebra*, and denoted RA^+. Recall the signatures Σ_s and Σ_{sm} from Eq. (2). Let $\boldsymbol{S}$ be any Σ_s-algebra (not necessarily a semiring), and define an $\boldsymbol{S}$-relation of arity k to be a function $R : \mathsf{Dom}^k \to \boldsymbol{S}$ of finite support (i.e. $\{t \mid R(t) \neq 0\}$ is finite). When $R(t) = u \in \boldsymbol{S}$, then we say that a tuple t is *annotated* with the element u. Tannen [13] associated to each operator in RA^+ an operation on $\boldsymbol{S}$-relations, in a natural way. For example union $R_1 \cup R_2$ returns the $\boldsymbol{S}$-relation $(R_1 \cup R_2)(t) \stackrel{\text{def}}{=} R_1(t) + R_2(t)$, natural join returns $(R_1 \bowtie R_2)(t) \stackrel{\text{def}}{=} R_1(\Pi_{\text{attrs}(R_1)}(t_1)) \cdot R_2(\Pi_{\text{attrs}(R_2)}(t_2))$, etc;[13] we refer the reader to Definition 3.2. in [13]. If $\boldsymbol{S}$ is an Σ_{sm} algebra, then we extend this definition from RA^+ to RA, by defining $(R_1 - R_2)(t) \stackrel{\text{def}}{=} R_1(t) - R_2(t)$.

Let Q_1, Q_2 be two RA-expressions. We write $Q_1 \equiv_{\boldsymbol{S}} Q_2$ if these two expressions return the same output for any input $\boldsymbol{S}$-relations. For example, if $+$ is commutative in $\boldsymbol{S}$ and $Q_1 = R \cup R'$, $Q_2 = R' \cup R$, then $Q_1 \equiv_{\boldsymbol{S}} Q_2$. Since $\mathbb{N}$-relations are bags, the equivalence $Q_1 \equiv_{\mathbb{N}} Q_2$ holds iff Q_1, Q_2 are equivalent RA-expressions under bag semantics. Similarly, $Q_1 \equiv_{\mathbb{B}} Q_2$ iff Q_1, Q_2 are equivalent under set semantics, for example $R \cup R \equiv_{\mathbb{B}} R$.

[12] The term *algebra* in RA is used with some abuse, since it is not a Σ-algebra, in the sense of Sec. 3. This is because the operators can only be applied to arguments with the right schemas, for example $R \cup S$ is defined only if the relations R, S have the same arity.

[13] Notice that the result of an operation may be an $\boldsymbol{S}$-relation with infinite support; this was apparently overlooked in [13]. However, this does not affect either the results in [13], nor those in this section, because, when $\boldsymbol{S}$ is a semiring, then all operations return $\boldsymbol{S}$-relations with finite support, assuming the inputs also have finite support.

▶ **Definition 18.** *A set of Σ_{sm}-identities E is* sound *for RA under bag semantics if, for any Σ_{sm}, E-algebra $\boldsymbol{S}$ and any two RA queries Q_1, Q_2, if $Q_1 \equiv_{\mathbb{N}} Q_2$ then $Q_1 \equiv_{\boldsymbol{S}} Q_2$. Similarly, E is sound for RA under set semantics if, for any Σ_{sm}, E-algebra $\boldsymbol{S}$ and any two RA queries Q_1, Q_2, if $Q_1 \equiv_{\mathbb{B}} Q_2$ then $Q_1 \equiv_{\boldsymbol{S}} Q_2$.*

Similarly, a set of Σ_s-identities is sound for RA$^+$ under bag (set) semantics if the condition above holds when Q_1, Q_2 are restricted to RA$^+$.

Our goal is to find a minimal set E that is sound for bag (set) semantics.

If $\boldsymbol{S}$ is any Σ_{sm} algebra then we denote by $E_{\mathrm{RA}}(\boldsymbol{S})$ the set of identities $Q_1 \equiv_{\boldsymbol{S}} Q_2$, where Q_1, Q_2 are RA queries, and denote by $E_{\mathrm{RA}}(\mathcal{C}) \stackrel{\text{def}}{=} \bigcap_{\boldsymbol{S}\in\mathcal{C}} E_{\mathrm{RA}}(\boldsymbol{S})$ where $\mathcal{C}$ is a class of Σ_s-algebras. We define similarly $E_{\mathrm{RA}^+}(\boldsymbol{S}), E_{\mathrm{RA}^+}(\mathcal{C})$ by restricting to Σ_s algebras and RA$^+$ queries. Tannen answered the soundness question for RA$^+$ under bag semantics in Proposition 3.4 of [13]:

▶ **Theorem 19** (Implicit in [13])**.** *(a) $E_{RA^+}(\mathbb{N}) \subseteq E_{RA^+}(\boldsymbol{S})$ iff $\boldsymbol{S}$ is a semiring. (b) If $\mathcal{C}$ is an equational class of Σ_s-algebras, then $E_{RA^+}(\mathcal{C}) = E_{RA^+}(\mathbb{N})$ iff $\mathcal{C}$ is the class of semirings.*

Part (a) proves that if E are the identities of semirings, then E is sound for RA$^+$ and bag semantics. Part (b) proves that E is the smallest sound set of identities. We will prove below a more general result that extends Theorem 19 from RA$^+$ to RA; the same proof can be used to prove Theorem 19.

Before we can extend the theorem, we need a brief review of Galois connections. Given two sets U, V, a Galois connection is a pair of functions $F : 2^U \to 2^V$, $G : 2^V \to 2^U$ such that (a) F, G are anti-monotone, and (b) the following condition holds:

$$Y \subseteq F(X) \text{ iff } X \subseteq G(Y) \tag{24}$$

In any Galois connection the following hold: $F(G(F(X))) = F(X)$ and $G(F(G(Y))) = G(Y)$.

Consider now a signature Σ, and an infinite set of variables X, and recall that a Σ-identity is a pair $e = (e_1, e_2)$ where $e_1, e_2 \in T_\Sigma(X)$. For any Σ-algebra $\boldsymbol{A}$, denote by $E_\Sigma(\boldsymbol{A})$ the set of identities that hold on $\boldsymbol{A}$. For example, $E_{\Sigma_{sm}}(\mathbb{N})$ contains all identities satisfied by $\mathbb{N}$, which includes the semiring identities, the CMM identities $(A1) - (A4)$, the identity $(A5)$, and many more. Furthermore, for an identity e denote by $\mathcal{C}_\Sigma(e)$ the class of Σ-algebras that satisfy e. Then the following two mappings form a Galois connection:

$$E_\Sigma(\mathcal{C}) \stackrel{\text{def}}{=} \bigcap_{\boldsymbol{A}\in\mathcal{C}} E_\Sigma(\boldsymbol{A}) \qquad\qquad \mathcal{C}_\Sigma(E) \stackrel{\text{def}}{=} \bigcap_{e\in E} \mathcal{C}_\Sigma(e) \tag{25}$$

The generalization of Theorem 19 to RA is the following:

▶ **Theorem 20.** *(a) $E_{RA}(\mathbb{N}) \subseteq E_{RA}(\boldsymbol{S})$ iff $\boldsymbol{S}$ is a $E_{\Sigma_{sm}}(\mathbb{N})$-algebra. (b) If $\mathcal{C}$ is any equational class of Σ_{sm}-algebras, then $E_{RA}(\mathcal{C}) = E_{RA}(\mathbb{N})$ iff $\mathcal{C}$ is the class defined by the identities $E_{\Sigma_{sm}}(\mathbb{N})$.*

Thus, in order to ensure soundness, the semiring $\boldsymbol{S}$ must satisfy all identities satisfied by the m-semiring $(\mathbb{N}, +, \cdot, 0, 1, \dot{-})$, which we denoted $E_{\Sigma_{sm}}(\mathbb{N})$. Before we prove the theorem, we show that this set is co-r.e. complete.

▶ **Theorem 21.** *$E_{\Sigma_{sm}}(\mathbb{N})$ is co-r.e. complete.*

Proof. Membership in co-r.e. is immediate: to check that an identity $e_1 = e_2$ is false in $\mathbb{N}$, it suffices to iterate over all possible assignments to the variables of e_1, e_2 and stop when one such assignment makes $e_1 \neq e_2$. To prove completeness, it suffices to prove that membership is undecidable, and for that we will use Matiyasevich theorem on the undecidability of Hilbert's tenth problem. It implies that the following problem is undecidable: *given a multivariate polynomial* $F \in \mathbb{Z}[X]$ *with variables* $x_1, x_2, \ldots$ *decide if there exists values* $x_1, x_2, \ldots \in \mathbb{N}$ *s.t.* $F(x_1, x_2, \ldots) = 0$. It follows immediately that the following problem is undecidable: given two polynomials $F, G \in \mathbb{N}[X]$, decide if the following holds:

$$\exists x_1, x_2, \ldots \in \mathbb{N} : \quad F(x_1, x_2, \ldots) = G(x_1, x_2, \ldots)$$

Its negation (which is also undecidable) is the statement:

$$\forall x_1, x_2, \ldots \in \mathbb{N} : \quad F(x_1, x_2, \ldots) \neq G(x_1, x_2, \ldots) \tag{26}$$

which we abbreviate by $F \neq G$. We use the following equivalences in $(\mathbb{N}, +, \cdot, 0, 1, \dot{-})$:

$$F \neq G \text{ iff } (F \dot{-} G) + (G \dot{-} F) > 0 \text{ iff } 1 \dot{-} ((F \dot{-} G) + (G \dot{-} F)) = 0$$

Since (26) is undecidable, checking identities of the form $1 \dot{-} ((F \dot{-} G) + (G \dot{-} F)) = 0$ in the m-semiring $(\mathbb{N}, +, \cdot, 0, 1, \dot{-})$ is also undecidable. Thus, $E_{\Sigma_{sm}}(\mathbb{N})$ is undecidable, therefore co-r.e. complete. ◀

To summarize, the minimal set of identities that ensures that an m-semiring is sound for all RA-identities is the set $E_{\Sigma_{sm}}(\mathbb{N})$, which is infinite, co-r.e.-complete, and, thus, it is not finitely generated. Obviously, the set $E_{\Sigma_{sm}}(\mathbb{N})$ is not practical. A possible workaround could be to find a non-minimal sound set, which is still useful for practical purposes: we leave this for future work.

We briefly discuss what happens when we interpret RA using set semantics instead of bag semantics, or consider both semantics. Theorem 20 continues to hold if we replace $E_{\mathrm{RA}}(\mathbb{N})$ and $E_{\Sigma_{sm}}(\mathbb{N})$ with $E_{\mathrm{RA}}(\mathbb{B})$ and $E_{\Sigma_{sm}}(\mathbb{B})$, thus the necessary and sufficient identities in this case are $E_{\Sigma_{sm}}(\mathbb{B})$. These are precisely the identities of Boolean algebras, which are generated by a finite set, and membership is decidable. This does *not* imply that $E_{\mathrm{RA}}(\mathbb{B})$ is decidable: in fact $E_{\mathrm{RA}}(\mathbb{B})$ consists of all pairs of RA-expressions that are equivalent under set semantics, and is undecidable by Trakhtenbrot's theorem.

Theorems 20 also specializes to RA^+, and we derive the following version of Theorem 19: the minimal set of Σ_s-identities E that is sound for RA^+ under bag semantics is $E_{\Sigma_s}(\mathbb{N})$, which is equivalent to the semiring axioms,[14], hence we recover Theorem 19. Similarly, we obtain that the minimal set of Σ_s-identities that are sound for RA^+ under set semantics is $E_{\Sigma_s}(\mathbb{B})$, which is equivalent to the set of identities of bounded, distributive lattices.[15] Finally, we briefly discuss what happens if we extend RA to support mixed set/bag semantics, by adding an operator δ to RA which eliminates duplicates. This requires us to add a new operation to the semiring $\boldsymbol{S}$, lets call it also δ, which satisfies $\delta(0) = 0$ and $\delta(x) = 1$ for all $x \neq 0$. Unfortunately, the δ-semirings do not form an equational class, because the product of two δ-semirings, $\boldsymbol{S}_1 \times \boldsymbol{S}_2$, is not a δ-semiring: for $(x, 0) \in \boldsymbol{S}_1 \times \boldsymbol{S}_2$, we have $\delta(x, 0) = (\delta_1(x), \delta_2(0)) = (1, 0)$, which is neither $(0, 0)$, nor $(1, 1)$. An operation of this kind was considered in [7] under the name *squash*, and defined using a conditional axiom.

[14] To see this, consider any two expressions $e_1, e_2 \in T_{\Sigma_s}(X)$ that are equivalent in $(\mathbb{N}, +, \cdot, 0, 1)$. Using the semiring identities only we can write e_1, e_2 in a canonical form, as a sum of monomials, i.e. $e_1, e_2 \in \mathbb{N}[X]$. Since they are equivalent in $(\mathbb{N}, +, \cdot, 0, 1)$, they must be identical polynomials. Thus, the equivalence $e_1 = e_2$ follows using only semiring axioms.

[15] The proof is similar to the above. Any two expressions e_1, e_2 equivalent in $\mathbb{B}$ can be transformed into DNF expressions using only the identities of distributive lattices, and their DNF expressions must be isomorphic.

In the rest of this section we prove Theorem 20. The proof follows from three lemmas. If $f : S_1 \to S_2$ is any function, then, for any k-ary $\boldsymbol{S}_1$-relation R we will denote by $f \circ R$ the $\boldsymbol{S}_2$-relation defined by $(f \circ R)(t) \stackrel{\text{def}}{=} f(R(t))$, for all $t \in \mathsf{Dom}^k$. If $\bar{R} = (R_1, R_2, \ldots)$ is a tuple of relations, then we write $f \circ \bar{R}$ for $(f \circ R_1, f \circ R_2, \ldots)$.

▶ **Lemma 22.** *Let $f : \boldsymbol{S}_1 \to \boldsymbol{S}_2$ be a homomorphism between Σ_{sm}-algebra, and Q be an RA query. Let $\bar{R}$ be a tuple of $\boldsymbol{S}_1$-relations. Then $Q(f \circ \bar{R}) = f \circ Q(\bar{R})$.*

Proof. The proof follows immediately by induction on the structure of Q. We illustrate here only for the case when Q is the union of two sub-queries; all other cases are similar. Assume $Q = Q_1 \cup Q_2$ and let t be a tuple in the output. Then:

$$\begin{aligned}
(Q(f \circ \bar{R}))(t) = & \left(Q_1(f \circ \bar{R}) \cup Q_2(f \circ \bar{R})\right)(t) && = \left(Q_1(f \circ \bar{R})\right)(t) + \left(Q_2(f \circ \bar{R})\right)(t) \\
= & \left(f \circ Q_1(\bar{R})\right)(t) + \left(f \circ Q_2(\bar{R})\right)(t) && = f\left(Q_1(\bar{R})(t)\right) + f\left(Q_2(\bar{R})(t)\right) \\
= & f\left(Q_1(\bar{R})(t) + Q_2(\bar{R})(t)\right) && = f\left(\left(Q_1(\bar{R}) \cup Q_2(\bar{R})\right)(t)\right) = \left(f \circ Q(\bar{R})\right)(t)
\end{aligned}$$

◀

▶ **Lemma 23.** *Let $\boldsymbol{S}_1, \boldsymbol{S}_2$ be two Σ_{sm}-algebras. Then $E_{RA}(\boldsymbol{S}_1) \subseteq E_{RA}(\boldsymbol{S}_2)$ iff $E_{\Sigma_{sm}}(\boldsymbol{S}_1) \subseteq E_{\Sigma_{sm}}(\boldsymbol{S}_2)$. The same statement holds for Σ_s-algebras and RA$^+$.*

In other words, in order to compare the RA-identities (RA$^+$-identities) satisfied by $\boldsymbol{S}_1$-relations with those satisfied by $\boldsymbol{S}_2$-relations, it suffices to compare the algebraic identities satisfied by $\boldsymbol{S}_1$ with those satisfied by $\boldsymbol{S}_2$.

Proof. We start with the $\Leftarrow$ direction, and assume $E_{\Sigma_{sm}}(\boldsymbol{S}_1) \subseteq E_{\Sigma_{sm}}(\boldsymbol{S}_2)$. Let $(Q_1, Q_2) \in E_{\text{RA}}(\boldsymbol{S}_1)$, in other words $Q_1(\bar{R}) = Q_2(\bar{R})$ for any input $\boldsymbol{S}_1$-relation instance $\bar{R}$. Consider some input $\boldsymbol{S}_2$-instance $\bar{R}'$: we need to prove that $Q_1(\bar{R}') = Q_2(\bar{R}')$. Let X be a set of variables, s.t. $|X| = |S_2|$, let $\eta : X \to T_{\Sigma_{sm}}(X)$ be the canonical injection, $h : X \to S_2$ be a bijection, $\bar{h} : T_{\Sigma_{sm}}(X) \to \boldsymbol{S}_2$ its extension to a homomorphism, and $\bar{R}'' \stackrel{\text{def}}{=} h^{-1} \circ \bar{R}'$. Thus, if a tuple in $\bar{R}''$ is annotated with variable $x \in X$, then the same tuple is annotated in $\bar{R}'$ with $h(x) \in S_2$: formally, $h \circ \bar{R}'' = \bar{h} \circ \eta \circ \bar{R}'' = \bar{R}'$. Then $Q_i(\bar{R}') = Q_i(\bar{h} \circ \eta \circ \bar{R}'') = \bar{h} \circ Q_i(\eta \circ \bar{R}'')$ (by Lemma 22) for $i = 1, 2$. (We cannot apply Q_i to $\bar{R}''$ because its annotations are variables in X; instead we apply Q_i to $\eta \circ \bar{R}''$, whose annotations are the same variables, but viewed in $T_{\Sigma_{sm}}(X)$, where the Σ_{sm}-operations are defined.) Let t be some output tuple, and define:

$$e_i \stackrel{\text{def}}{=} (Q_i(\eta \circ \bar{R}''))(t), \quad i = 1, 2$$

Thus, e_1, e_2 are expressions in $T_{\Sigma_{sm}}(X)$ that annotate the tuple t in the outputs of Q_1, Q_2 on $\bar{R}''$ respectively. We claim that $(e_1, e_2) \in E_{\Sigma_{sm}}(\boldsymbol{S}_1)$ (i.e. they form an identity that holds in $\boldsymbol{S}_1$). For that, we need to show that for any function $g : X \to S_1$, the equality $\bar{g}(e_1) = \bar{g}(e_2)$ holds. To prove that we use the fact that Q_1, Q_2 return the same output on the $\boldsymbol{S}_1$ relation $\bar{g} \circ \eta \circ \bar{R}''$:

$$\begin{aligned}
\left(Q_1(\bar{g} \circ \eta \circ \bar{R}'')\right)(t) &= \left(Q_2(\bar{g} \circ \eta \circ \bar{R}'')\right)(t) \\
\bar{g}\left(Q_1(\eta \circ \bar{R}'')(t)\right) &= \bar{g}\left(Q_2(\eta \circ \bar{R}'')(t)\right) \\
\bar{g}(e_1) &= \bar{g}(e_2)
\end{aligned}$$

Since g was arbitrary, we conclude that $(e_1, e_2) \in E_{\Sigma_{sm}}(\boldsymbol{S}_1)$, and therefore $(e_1, e_2) \in E_{\Sigma_{sm}}(\boldsymbol{S}_2)$, which implies $\bar{h}(e_1) = \bar{h}(e_2)$, proving that $Q_1(\bar{R}')(t) = Q_2(\bar{R}'')(t)$.

We briefly sketch the $\Rightarrow$ direction of the proof, for Σ_{sm} and RA. Given $(e_1, e_2) \in E_{\Sigma_{sm}}(\boldsymbol{S}_1)$, we convert both e_1 and e_2 into RA-expressions over unary relations. For example, if e_1 is $(x_1^2 \cdot x_2 \dot{-} x_3) + x_1$ then Q_1 is $(R_1 \bowtie R_1 \bowtie R_2 - R_3) \cup R_1$, where R_1, R_2, R_3 are unary relations with the same attribute name. It follows immediately that $(Q_1, Q_2) \in E_{\mathrm{RA}}(\boldsymbol{S}_1)$, hence it is also in $E_{\mathrm{RA}}(\boldsymbol{S}_2)$, and this implies that the identity $e_1 = e_2$ also holds in $\boldsymbol{S}_2$. ◀

The last lemma is:

▶ **Lemma 24.** *Fix a signature Σ. (a) For any two Σ-algebras $\boldsymbol{A}, \boldsymbol{B}$, the following holds: $E_\Sigma(\boldsymbol{A}) \subseteq E_\Sigma(\boldsymbol{B})$ iff $\boldsymbol{B} \in \mathcal{C}_\Sigma(E_\Sigma(\boldsymbol{A}))$. (b) If $\mathcal{C}$ is an equational class of Σ-algebras, then $E_\Sigma(\mathcal{C}) = E_\Sigma(\boldsymbol{A})$ iff $\mathcal{C} = \mathcal{C}_\Sigma(E_\Sigma(\boldsymbol{A}))$.*

Proof. Part (a) is by the definition of the Galois connection (24). We prove now part (b). For one direction of (b), assume $\mathcal{C} = \mathcal{C}_\Sigma(E_\Sigma(\boldsymbol{A}))$. Then $E_\Sigma(\mathcal{C}) = E_\Sigma(\mathcal{C}_\Sigma(E_\Sigma(\boldsymbol{A}))) = E_\Sigma(\boldsymbol{A})$. For the other direction, assume $\mathcal{C}$ is an equational class and $E_\Sigma(\mathcal{C}) = E_\Sigma(\boldsymbol{A})$. Then $\mathcal{C}_\Sigma(E_\Sigma(\mathcal{C})) = \mathcal{C}_\Sigma(E_\Sigma(\boldsymbol{A}))$, and the claim follows from the fact that $\mathcal{C}_\Sigma(E_\Sigma(\mathcal{C})) = \mathcal{C}$ because $\mathcal{C}$ is an equational class, i.e. $\mathcal{C} = \mathcal{C}_\Sigma(E)$ for some E. ◀

Lemmas 23 and 24 immediately imply Theorem 20.