Locality Theorems in Semiring Semantics

Clotilde Bizière
ENS Paris, France

Erich Grädel
RWTH Aachen University, Germany

Matthias Naaf
RWTH Aachen University, Germany

Abstract

Semiring semantics of first-order logic generalises classical Boolean semantics by permitting truth values from a commutative semiring, which can model information such as costs or access restrictions. This raises the question to what extent classical model-theoretic properties still apply, and how this depends on the algebraic properties of the semiring.

In this paper, we study this question for the classical locality theorems due to Hanf and Gaifman. We prove that Hanf's locality theorem generalises to all semirings with idempotent operations, but fails for many non-idempotent semirings. We then consider Gaifman normal forms and show that for formulae with free variables, Gaifman's theorem does not generalise beyond the Boolean semiring. Also for sentences, it fails in the natural semiring and the tropical semiring. Our main result, however, is a constructive proof of the existence of Gaifman normal forms for min-max and lattice semirings. The proof implies a stronger version of Gaifman's classical theorem in Boolean semantics: every sentence has a Gaifman normal form which does not add negations.

2012 ACM Subject Classification Theory of computation → Finite Model Theory

Keywords and phrases Semiring semantics, Locality, First-order logic

Digital Object Identifier 10.4230/LIPIcs.MFCS.2023.20

Related Version *Full Version*: `https://arxiv.org/abs/2303.12627` [6]

1 Introduction

Originally motivated by *provenance analysis in databases* (see e.g. [18, 12] for surveys), semiring semantics is based on the idea to evaluate logical statements not just by *true* or *false*, but by values in some commutative semiring $(K, +, \cdot, 0, 1)$. In this context, the standard semantics appears as the special case when the Boolean semiring $\mathbb{B} = (\{\bot, \top\}, \vee, \wedge, \bot, \top)$ is used. Valuations in other semirings provide additional information, beyond truth or falsity: the tropical semiring $\mathbb{T} = (\mathbb{R}_+^\infty, \min, +, \infty, 0)$ is used for *cost analysis*, the natural semiring $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$ for counting evaluation strategies and proofs, and the Viterbi-semiring $\mathbb{V} = ([0, 1]_\mathbb{R}, \max, \cdot, 0, 1)$ models *confidence scores*. Finite or infinite min-max semirings $(K, \max, \min, a, b)$ can model, for instance, different *access levels* to atomic data (see e.g. [10]); valuations of a first-order sentence ψ in such *security semirings* determine the required clearance level that is necessary to access enough information to determine the truth of ψ. Further, semirings of polynomials or formal power series permit us to *track* which atomic facts are used (and how often) to establish the truth of a sentence in a given structure, and this has applications for database repairs [26] and also for the strategy analysis of games [17, 14]. Semiring semantics replaces structures by *K-interpretations*, which are functions $\pi \colon \mathrm{Lit}_A(\tau) \to K$, mapping fully instantiated τ-literals $\varphi(\overline{a})$ over a universe A to values in a commutative semiring K. The value $0 \in K$ is interpreted as *false*, while all other values in K are viewed as nuances of *true* or, perhaps more accurately, as *true, with some* *additional information*. In provenance analysis, this is sometimes referred to as *annotated facts*. The value $1 \in K$ is used to represent *untracked information* and is used in particular to evaluate true equalities and inequalities.

The development of semiring semantics raises the question to what extent classical techniques and results of logic extend to semiring semantics, and how this depends on the algebraic properties of the underlying semirings. Previous investigations in this direction have studied, for instance, the relationship between elementary equivalence and isomorphism for finite semiring interpretations and their definability up to isomorphism [15], Ehrenfeucht-Fraïssé games [7], and 0-1 laws [13].

The purpose of this paper is to study *locality* in semiring semantics. Locality is a fundamental property of first-order logic in classical semantics and an important limitation of its expressive power. It means that the truth of a first-order formula $\psi(\overline{x})$ in a given structure only depends on a neighbourhood of bounded radius around $\overline{x}$, and on the existence of a bounded number of local substructures. Consequently, first-order logic cannot express global properties such as connectivity or acyclicity of graphs. On graphs there are natural and canonical notions of the distance between two points and of a neighbourhood of a given radius around a point. To define these notions for an arbitrary relational structure $\mathfrak{A}$ one associates with it its *Gaifman graph* $G(\mathfrak{A}) = (A, E)$ where two points $a \neq b$ are adjacent if, and only if, they coexist in some atomic fact. There exist several notions of locality; the most common ones are *Hanf locality* and *Gaifman locality*, and the fundamental locality theorems for first-order logic are *Hanf's locality theorem* and *Gaifman's normal form theorem*. In a nutshell, Hanf's theorem gives a criterion for the m-equivalence (i.e. indistinguishability by sentences of quantifier rank up to m) of two structures based on the number of local substructures of any given isomorphism type, while Gaifman's theorem states that every first-order formula is equivalent to a Boolean combination of local formulae and basic local sentences, which has many model-theoretic and algorithmic consequences. We shall present precise statements of these results in Sect. 3 and Sect. 4.

Locality thus provides powerful techniques, also for logics that go beyond first-order logic by counting properties, generalised quantifiers, or aggregate functions, [1, 21, 22, 23]. It has applications in different areas including low-complexity model-checking algorithms [19, 20], approximation schemes for logically defined optimisation problems [8], automata theory [25], computational issues on database transactions [2], and most recently also in learning theory, for the efficient learning of logical concepts [3, 5, 4]. This motivates the question, whether locality is also applicable in semiring semantics. The relevant semiring interpretations in this context are *model-defining*, which means that for any pair of complementary literals $R\overline{a}, \neg R\overline{a}$ precisely one of the values $\pi(R\overline{a})$, $\pi(\neg R\overline{a})$ is 0, and *track only positive information* which means that $\pi(\neg R\overline{a})$ can only take the values 0 or 1. Model defining interpretations π define a unique structure $\mathfrak{A}_\pi$ and we thus obtain a well-defined Gaifman graph $G(\pi) \coloneqq G(\mathfrak{A}_\pi)$, with the associated notions of distance and neighbourhoods. The assumption that only positive information is tracked is necessary to get meaningful locality properties (see Sect. 2).

We clearly cannot generalise all known locality properties of first-order logic to semiring semantics in arbitrary commutative semirings. On semirings whose operations are not idempotent, we cannot expect a Gaifman normal form, since for computing the value of a quantified statement, we have to add or multiply values of subformulae for *all* elements of the structure, which gives an inherent source of non-locality. As a consequence, some of the locality results that we prove hold only under certain algebraic assumptions on the semiring, and further there turns out to be a difference of the locality properties of sentences and those of formulae with free variables. We shall establish the following results.

(1) First-order formulae are Hanf-local for all semirings.
(2) Hanf's locality theorem generalises to all fully idempotent semirings (in which both addition and multiplication are idempotent).
(3) For formulae with free variables, Gaifman's normal form theorem does not generalise beyond the Boolean semiring.
(4) For sentences, Gaifman's normal form theorem also fails in certain important semirings such as the natural semiring and the tropical semiring.
(5) Over min-max semirings (and even lattice semirings), every first-order sentence has a Gaifman normal form.
(6) In classical Boolean semantics, every sentence has a Gaifman normal form which does not introduce new negations.

The results **(1)**, **(2)** on Hanf locality (Sect. 3) are proved by adaptations of the arguments for the Boolean case. The results **(3)** and **(4)** are established in Sect. 5 via specific examples of formulae that defeat locality, using simple algebraic arguments. The most ambitious result and the core of our paper is **(5)**, a version of Gaifman's theorem for min-max semirings (Sect. 6), which we later generalise to lattice semirings (Sect. 7). It requires a careful choice of the right syntactical definitions for local sentences and, since the classical proofs in [11, 9] do not seem to generalise to semiring semantics, a new approach for the proof, based on quantifier elimination. This new approach also leads to a stronger version of Gaifman's theorem in Boolean semantics **(6)**, which might be of independent interest.

2 Semiring Semantics

This section gives a brief overview on semiring semantics of first-order logic (see [16] for more details) and the relevant algebraic properties of semirings. A commutative[1] semiring is an algebraic structure $(K, +, \cdot, 0, 1)$ with $0 \neq 1$, such that $(K, +, 0)$ and $(K, \cdot, 1)$ are commutative monoids, $\cdot$ distributes over $+$, and $0 \cdot a = a \cdot 0 = 0$. We focus on semirings that are *naturally ordered*, in the sense that $a \leq b :\Leftrightarrow \exists c(a + c = b)$ is a partial order. For the study of locality properties, an important subclass are the *fully idempotent* semirings, in which both operations are idempotent (i.e., $a + a = a$ and $a \cdot a = a$). Among these, we consider in particular all *min-max* semirings $(K, \max, \min, 0, 1)$ induced by a total order $(K, \leq)$ with minimal element 0 and maximal element 1, and the more general *lattice* semirings $(K, \sqcup, \sqcap, 0, 1)$ induced by a bounded distributive lattice $(K, \leq)$.

For a finite relational vocabulary τ and a finite universe A, we write $\mathrm{Lit}_A(\tau)$ for the set of *instantiated* τ-literals $R\overline{a}$ and $\neg R\overline{a}$ with $\overline{a} \in A^{\mathrm{arity}(R)}$. Given a commutative semiring K, a *K-interpretation* (of vocabulary τ and universe A) is a function $\pi\colon \mathrm{Lit}_A(\tau) \to K$. It is *model-defining* if for any pair of complementary literals L, $\neg L$ precisely one of the values $\pi(L)$, $\pi(\neg L)$ is 0. In this case, π induces a unique (Boolean) τ-structure $\mathfrak{A}_\pi$ with universe A such that, for every literal $L \in \mathrm{Lit}_A(\tau)$, we have that $\mathfrak{A}_\pi \models L$ if, and only if, $\pi(L) \neq 0$.

A K-interpretation $\pi\colon \mathrm{Lit}_A(\tau) \to K$ extends in a straightforward way to a valuation $\pi[\![\varphi(\overline{a})]\!]$ of any instantiation of a formula $\varphi(\overline{x}) \in \mathrm{FO}(\tau)$, assumed to be written in negation normal form, by a tuple $\overline{a} \subseteq A$. The semiring semantics $\pi[\![\varphi(\overline{a})]\!]$ is defined by induction. We first extend π by mapping equalities and inequalities to their truth values, by setting $\pi[\![a = a]\!] := 1$ and $\pi[\![a = b]\!] := 0$ for $a \neq b$ (and analogously for inequalities). Further, disjunctions and existential quantifiers are interpreted as sums, and conjunctions and universal quantifiers as products:

[1] In the following, *semiring* always refers to a commutative semiring.

$$\pi[\![\psi(\overline{a}) \vee \vartheta(\overline{a})]\!] := \pi[\![\psi(\overline{a})]\!] + \pi[\![\vartheta(\overline{a})]\!] \qquad \pi[\![\psi(\overline{a}) \wedge \vartheta(\overline{a})]\!] := \pi[\![\psi(\overline{a})]\!] \cdot \pi[\![\vartheta(\overline{a})]\!]$$
$$\pi[\![\exists x\, \vartheta(\overline{a}, x)]\!] := \sum_{a \in A} \pi[\![\vartheta(\overline{a}, a)]\!] \qquad \pi[\![\forall x\, \vartheta(\overline{a}, x)]\!] := \prod_{a \in A} \pi[\![\vartheta(\overline{a}, a)]\!].$$

Since negation does not correspond to a semiring operation, we insist on writing all formulae in negation normal form. This is a standard approach in semiring semantics (cf. [16]). Equivalence of formulae now takes into account the semiring values and is thus more fine-grained than Boolean equivalence.

▶ **Definition 1** ($\equiv_K$)**.** *Two formulae $\psi(\overline{x})$, $\varphi(\overline{x})$ are K-equivalent (denoted $\psi \equiv_K \varphi$) if $\pi[\![\psi(\overline{a})]\!] = \pi[\![\varphi(\overline{a})]\!]$ for every model-defining K-interpretation π (over finite universe) and every tuple $\overline{a}$. For a class $\mathcal{S}$ of semirings, we write $\psi \equiv_{\mathcal{S}} \varphi$ if $\psi \equiv_K \varphi$ holds for all $K \in \mathcal{S}$.*

Towards locality properties, we define distances between two elements a, b in a K-interpretation π based on the induced structure $\mathfrak{A}_\pi$.

▶ **Definition 2** (Gaifman graph)**.** *The* Gaifman graph *$G(\pi)$ of a model-defining K-interpretation $\pi\colon \mathrm{Lit}_A(\tau) \to K$ is defined as the Gaifman graph $G(\mathfrak{A}_\pi)$ of the induced τ-structure. That is, two elements $a \neq b$ of A are adjacent in $G(\mathfrak{A}_\pi)$ if, and only if, there exists a positive literal $L = Rc_1 \dots c_r \in \mathrm{Lit}_A(\tau)$ such that $\pi(L) \neq 0$ and $a, b \in \{c_1, \dots c_r\}$.*

*We write $d(a, b) \in \mathbb{N}$ for the distance of a and b in $G(\pi)$. We further define the r-*neighbourhood *of an element a in π as $B_r^\pi(a) := \{b \in A : d(a, b) \leq r\}$. For a tuple $\overline{a} \in A^k$ we put $B_r^\pi(\overline{a}) := \bigcup_{i \leq k} B_r^\pi(a_i)$.*

Locality properties are really meaningful only for semiring interpretations $\pi\colon \mathrm{Lit}_A(\tau) \to K$ that *track only positive information*, which means that $\pi(\neg L) \in \{0, 1\}$ for each negative literal $\neg L$. Indeed, if also negative literals carry non-trivial information, then either these must be taken into account in the definition of what "local" means, which will trivialise the Gaifman graph (making it a clique) so locality would become meaningless, or otherwise local information no longer suffices to determine values of even very simple sentences involving negative literals, such as $\exists x \exists y \neg Rxy$. We therefore consider here only K-interpretations over finite universes which are model-defining and track only positive information.

3 Hanf Locality

The first formalisation of locality that we consider is Hanf locality. We present generalisations of both the Hanf locality rank and of Hanf's locality theorem, where the latter is conditional on algebraic properties of the semirings. One point that requires care in the adaptation of the classical proofs (cf. [9, 24]) is the combination of partial isomorphisms on disjoint and non-adjacent neighbourhoods. In the setting of semiring semantics, this depends on the assumption that K-interpretations only track positive information.

▶ **Lemma 3.** *Let π_A and π_B be model-defining K-interpretations that track only positive information. Let $\sigma\colon B_r^{\pi_A}(\overline{a}) \to B_r^{\pi_B}(\overline{b})$ and $\sigma'\colon B_r^{\pi_A}(\overline{a}') \to B_r^{\pi_B}(\overline{b}')$ be two partial isomorphisms between disjoint r-neighbourhoods in π_A and π_B. If $d(\overline{a}, \overline{a}') > 2r+1$ and $d(\overline{b}, \overline{b}') > 2r+1$, then $(\sigma \cup \sigma')\colon B_r^{\pi_A}(\overline{a}, \overline{a}') \to B_r^{\pi_B}(\overline{b}, \overline{b}')$ is also a partial isomorphism.*

In classical Boolean semantics, a formula $\psi(\overline{x})$ is *Hanf-local with Hanf locality rank r*, if for any two tuples $\overline{a}$ in $\mathfrak{A}$ and $\overline{b}$ in $\mathfrak{B}$ we have the equivalence that $\mathfrak{A} \models \psi(\overline{a}) \Leftrightarrow \mathfrak{B} \models \psi(\overline{b})$ whenever there is a bijection $f : A \to B$ such that the r-neighbourhoods $B_r^{\mathfrak{A}}(\overline{a}, c)$ and

$B_r^{\mathfrak{A}}(\bar{b}, f(c))$ are isomorphic for all $c \in A$. It is known that every first-order formula is Hanf-local, with locality rank depending only on the quantifier rank. The proof of this fact [24] relies on an inductive argument which, from given bijections between $(3r+1)$-neighbourhoods of k-tuples, builds bijections between r-neighbourhoods of $(k+1)$-tuples. Based on Lemma 3, the inductive argument can be adapted to semiring semantics to get the following result, which does not assume any specific properties of the underlying semiring (for details, see [6]).

▶ **Proposition 4** (Hanf locality in semiring semantics). *Let K be an arbitrary semiring. For every first-order formula $\varphi(\bar{x})$, there exists $r \in \mathbb{N}$, depending only on the quantifier rank of φ, such that for all model-defining K-interpretations π_A, π_B that track only positive information, and all tuples $\bar{a}$, $\bar{b}$ we have that $\pi_A[\![\varphi(\bar{a})]\!] = \pi_B[\![\varphi(\bar{b})]\!]$ whenever there is a bijection $f\colon A \to B$ such that $B_r^{\pi_A}(\bar{a}, c) \cong B_r^{\pi_B}(\bar{b}, f(c))$ for all $c \in A$.*

A much more fundamental result is *Hanf's locality theorem* which provides a sufficient combinatorial criterion for the m-equivalence of two structures, i.e. for their indistinguishability by sentences of quantifier rank up to m. We follow the classical proof in [9], which proceeds by showing that Hanf's criterion admits the construction of a back-and-forth system $(I_j)_{j\le m}$ which, by the Ehrenfeucht-Fraïssé theorem, implies the m-equivalence of the two structures. It turns out that this method carries over to K-interpretations precisely in the case that the semiring K is fully idempotent. We further show that for semirings that are not fully idempotent, there actually are counterexamples to Hanf's locality theorem.

To define back-and-forth systems between K-interpretations, first notice that the notion of partial isomorphisms generalises in an obvious way to K-interpretations (cf. [6]).

▶ **Definition 5** (Back-and-forth system). *Let π_A and π_B be two K-interpretations and let $k \ge 0$. A m-*back-and-forth system* for π_A and π_B is a sequence $(I_j)_{j\le m}$ of finite sets of partial isomorphisms between π_A and π_B such that*

- *$\varnothing \in I_m$, and*
- *for all $j < m$, the set I_{j+1} has back-and-forth extensions in I_j, i.e., whenever $\bar{a} \mapsto \bar{b} \in I_{j+1}$ then for every $c \in A$ there exists $d \in B$, and vice versa, such that $(\bar{a}c) \mapsto (\bar{b}d)$ is in I_j.*

We write $(I_j)_{j\le m}\colon \pi_A \cong_m \pi_B$ if $(I_j)_{j\le m}$ is a m-back-and-forth system for π_A and π_B.

Back-and-forth systems can be seen as algebraic descriptions of winning strategies in Ehrenfeucht-Fraïssé games, and in classical semantics, an m-back-and-forth system between two structures exists if, and only if, the structures are m-equivalent. However, in semiring semantics this equivalence may, in general, fail in both directions [7]. A detailed investigation of the relationship between elementary equivalence, Ehrenfeucht-Fraïssé games, and back-and-forth-systems in semiring semantics is outside the scope of this paper, and will be presented in forthcoming work. For the purpose of studying Hanf locality, we shall need just the fact that in the specific case of fully idempotent semirings, m-back-and-forth systems do indeed provide a sufficient criterion for m-equivalence.

▶ **Proposition 6.** *Let π_A and π_B be K-interpretations into a fully idempotent semiring K. If there is an m-back-and-forth system $(I_j)_{j\le m}$ for π_A and π_B, then $\pi_A \equiv_m \pi_B$.*

Proof. We show by induction that for every first-order formula $\psi(\bar{x})$ of quantifier rank $j \le m$ and every partial isomorphism $\bar{a} \mapsto \bar{b} \in I_j$ we have that $\pi_A[\![\psi(\bar{a})]\!] = \pi_B[\![\psi(\bar{b})]\!]$. For $j = 0$ this is trivial. For the inductive case it suffices to consider formulae $\psi(\bar{x}) = \exists y\, \varphi(\bar{x}, y)$ and $\psi(\bar{x}) = \forall y\, \varphi(\bar{x}, y)$, and a map $\bar{a} \mapsto \bar{b} \in I_{j+1}$. We have that

$$\pi_A[\![\exists y\, \varphi(\overline{a}, y)]\!] = \sum_{c \in A} \pi_A[\![\varphi(\overline{a}, c)]\!] \qquad \text{and} \qquad \pi_B[\![\exists y\, \psi(\overline{b}, y)]\!] = \sum_{d \in B} \pi_B[\![\varphi(\overline{b}, d)]\!],$$
$$\pi_A[\![\forall y\, \varphi(\overline{a}, y)]\!] = \prod_{c \in A} \pi_A[\![\varphi(\overline{a}, c)]\!] \qquad \text{and} \qquad \pi_B[\![\forall y\, \psi(\overline{b}, y)]\!] = \prod_{d \in B} \pi_B[\![\varphi(\overline{b}, d)]\!].$$

Since the semiring is fully idempotent, the valuations $\pi_A[\![\exists y\, \varphi(\overline{a}, y)]\!]$ and $\pi_A[\![\forall y\, \varphi(\overline{a}, y)]\!]$ only depend on the *set* of all values $\pi_A[\![\varphi(\overline{a}, c)]\!]$ for $c \in A$, and not on their multiplicities. It thus suffices to prove that the sets of values are identical for $(\pi_A, \overline{a})$ and $(\pi_B, \overline{b})$, i.e.

$$\{\pi_A[\![\varphi(\overline{a}, c)]\!] : c \in A\} = \{\pi_B[\![\varphi(\overline{b}, d)]\!] : d \in B\}.$$

But this follows immediately from the fact that $\overline{a} \mapsto \overline{b}$ has back and forth extensions in I_j, and from the induction hypothesis: for each $c \in A$ there exists some $d \in B$, and vice versa, such that the map $(\overline{a}, c) \mapsto (\overline{b}, d)$ is in I_j, and therefore $\pi_A[\![\varphi(\overline{a}, c)]\!] = \pi_B[\![\varphi(\overline{b}, d)]\!]$. ◀

To formulate Hanf's criterion for K-interpretations π_A, π_B, we write $\pi_A \rightleftharpoons_{r,t} \pi_B$, for $r, t \in \mathbb{N}$, if for every isomorphism type ι of r-neighbourhoods, either π_A and π_B have the same number of realisations of ι, or both have at least t realisations.

▶ **Theorem 7** (Hanf's theorem for fully idempotent semirings). *Let K be a fully idempotent semiring. For all $m, \ell \in \mathbb{N}$ there exist $r = r(m) \in \mathbb{N}$ and $t = t(m, \ell) \in \mathbb{N}$ such that for all model-defining K-interpretations π_A and π_B that track only positive information and whose Gaifman graphs have maximal degree $\leq \ell$, we have that $\pi_A \equiv_m \pi_B$ whenever $\pi_A \rightleftharpoons_{r,t} \pi_B$.*

Proof. Given $m, \ell \in \mathbb{N}$, let $r_0 = 0$, inductively define $r_{i+1} = 3r_i + 1$, and set $r = r_{m-1}$. Further, let $t = m \cdot e + 1$, where $e \coloneqq 1 + \ell + \ell^2 + \cdots + \ell^r$ is the maximal number of elements in an r-neighbourhood of a point, in K-interpretations with Gaifman graphs with maximal degree ℓ. Assume that π_A and π_B are K-interpretations with that property, such that $\pi_A \rightleftharpoons_{r,t} \pi_B$.

We construct an m-back-and-forth system $(I_j)_{j \leq m}$ for (π_A, π_B) by setting

$$I_j \coloneqq \{\overline{a} \mapsto \overline{b} : |\overline{a}| = |\overline{b}| = m - j \text{ and } B^{\pi_A}_{r_j}(\overline{a}) \cong B^{\pi_B}_{r_j}(\overline{b})\}.$$

We have $I_m = \{\varnothing\}$, and since $\pi_A \rightleftharpoons_{r,t} \pi_B$, we have for every $a \in A$ some $b \in B$, and vice versa, such that $B^{\pi_A}_r(a) \cong B^{\pi_B}_r(b)$, so I_m has back-and-forth extensions in I_{m-1}. Consider now a partial isomorphism $\overline{a} \mapsto \overline{b}$ in I_{j+1}. There is an isomorphism $\rho\colon B^{\pi_A}_{3r_j+1}(\overline{a}) \cong B^{\pi_B}_{3r_j+1}(\overline{b})$. By symmetry, it suffices to prove the forth-property: for every $a \in A$ we must find some $b \in B$ such that $\overline{a}a \mapsto \overline{b}b \in I_j$ which means that $B^{\pi_A}_{r_j}(\overline{a}a) \cong B^{\pi_B}_{r_j}(\overline{b}b)$.

Case 1 (a close to $\overline{a}$). If $a \in B^{\pi_a}_{2r_j+1}(\overline{a})$, then we choose $b = \rho(a) \in B^{\pi_B}_{2r_j+1}(\overline{b})$. This is a valid choice since $B^{\pi_A}_{r_j}(\overline{a}a) \subseteq B^{\pi_A}_{3r_j+1}(\overline{a})$ so ρ also provides an isomorphism between $B^{\pi_A}_{r_j}(\overline{a}a)$ and $B^{\pi_B}_{r_j}(\overline{b}b)$.

Case 2 (a far from $\overline{a}$). If $a \notin B^{\pi_a}_{2r_j+1}(\overline{a})$, then $B^{\pi_A}_{r_j}(a) \cap B^{\pi_A}_{r_j}(\overline{a}) = \varnothing$. Hence, it suffices to find $b \in B$ such that $B^{\pi_B}_{r_j}(b)$ has the same isomorphism type as $B^{\pi_A}_{r_j}(a)$ (call this ι) with the property that b has distance at least $2r_j + 2$ to $\overline{b}$. Since π_A and π_B only track positive information the isomorphisms can be combined by Lemma 3 to show that $B^{\pi_A}_{r_j}(\overline{a}a) \cong B^{\pi_B}_{r_j}(\overline{b}b)$.

Assume that no such b exists. Let s be the number of elements realising ι in π_B. Since all of them are have distance at most $2r_j + 1$ from $\overline{b}$ and there are at most t elements in r-neighbourhoods around $\overline{b}$, we have that $s \leq t$. On the other side there are at least $s + 1$ elements realising ι in π_A, namely s elements in $B^{\pi_A}_{2r_j+1}(\overline{a})$ (due to ρ) and a. But this contradicts the fact that ι either has the same number of realisations in π_A and π_B, or more than t realisations in both interpretations. Hence such an element b exists, and we have proved that $(I_j)_{j\leq m}$ is indeed a m-back-and-forth system for (π_A, π_B).

By Proposition 6 this implies that $\pi_A \equiv_m \pi_B$. ◀

On the other side, we observe that Hanf's locality theorem in general *fails* for semirings with non-idempotent operations.

▶ **Example 8** (Counterexample Hanf). Consider the natural semiring $(\mathbb{N}, +, \cdot, 0, 1)$ and $\psi = \exists x\, Ux$ over signature $\tau = \{U\}$. For each n, we define a model-defining K-interpretation π_n with universe $\{a_1, \dots, a_n\}$ by setting $\pi(Ua_i) = 1$ for all i. Then $\pi_n[\![\psi]\!] = \sum_i \pi(Ua_i) = n$.

As we only have unary predicates, all neighbourhoods are trivial. That is, they consist of just one element and all of them have the same isomorphism type. Thus, π_n realises this single isomorphism type precisely n times, which means that $\pi_n \rightleftharpoons_{r,t} \pi_t$ for all r, t with $n \geq t$. But $\pi_n[\![\psi]\!] \neq \pi_t[\![\psi]\!]$ for $n \neq t$, so Hanf's theorem fails for the natural semiring.

This example readily generalises to all semirings containing an element $s \in K$ for which there are arbitrarily large numbers $n, m \in \mathbb{N}$ with $m \cdot s \neq n \cdot s$ or $s^m \neq s^n$ ($m \cdot s$ and s^m refer to the m-fold addition and multiplication of s, respectively). Indeed, we can map all atoms Ua_i to s and observe that Hanf's theorem fails for either $\psi = \exists x\, Ux$ or $\psi = \forall x\, Ux$. ⌟

4 Gaifman Normal Forms in Semirings Semantics

We briefly recall the classical notion of Gaifman normal forms (cf. [11, 9]), which capture locality in a syntactic way. Gaifman normal forms are Boolean combinations of *local formulae* $\varphi^{(r)}(x)$ and *basic local sentences*. A local formula $\varphi^{(r)}(x)$ is a formula in which all quantifiers are *relativised* to the r-neighbourhood of x, for instance $\exists y\, \vartheta(x, y)$ is relativised to $\exists y(d(x, y) \leq r \wedge \vartheta(x, y))$. Here, $d(x, y) \leq r$ asserts that x and y have distance $\leq r$ in the Gaifman graph, which can easily be expressed in first-order logic (in Boolean semantics). A basic local sentence asserts that there exist *scattered* elements, i.e., elements with distinct r-neighbourhoods, which all satisfy the same r-local formula: $\exists x_1 \dots \exists x_m(\bigwedge_{i\neq j} d(x_i, x_j) > 2r \wedge \bigwedge_i \varphi^{(r)}(x_i))$. By Gaifman's theorem, every formula has an equivalent Gaifman normal form, which intuitively means that it only makes statements about distinct local neighbourhoods.

Moving to semiring semantics, we keep the notion of Gaifman normal forms close to the original one, with two exceptions. First, we only consider formulae in negation normal form. This means that we restrict to *positive* Boolean combinations and, in turn, permit the duals of basic local sentences (i.e., the negations of basic local sentences, in negation normal form). Second and most importantly, we lose the ability to express relativised quantifiers[2] in our logic. Instead, we extend first-order logic by adding relativised quantifiers (*ball quantifiers*) of the form $Qy{\in}B_r^\tau(x)$ for $Q \in \{\exists, \forall\}$ with the following semantics: given a formula $\varphi(x, y)$, a K-interpretation $\pi\colon \mathrm{Lit}_A(\tau) \to K$, and an element a, we define

$$\pi[\![\exists y{\in}B_r^\tau(a)\ \varphi(a, y)]\!] \coloneqq \sum_{b\in B_r^\pi(a)} \pi[\![\varphi(a, b)]\!], \qquad \pi[\![\forall y{\in}B_r^\tau(a)\ \varphi(a, y)]\!] \coloneqq \prod_{b\in B_r^\pi(a)} \pi[\![\varphi(a, b)]\!].$$

We drop τ and write $\exists y{\in}B_r(a)$ or $\forall y{\in}B_r(a)$ if the signature is clear from the context.

[2] We could use the same formula for $d(x, y) \leq r$ as in the Boolean case. However, this formula would not just evaluate to 0 or 1, but would include the values of all edges around x, so each relativised quantifier would have the unintended side-effect of multiplying with the edge values in the neighbourhood. One can show that this side-effect would make Gaifman normal forms impossible (see [6] for details).

$\varphi^{(r)}(x,y) =$
$\forall z_1{\in}B_{r_1}(x)\,\big(\exists z_2{\in}B_{r_2}(z_1)\,\forall z_3{\in}B_{r_3}(z_1)\,\neg Ez_2z_3\big)$
$\quad\vee\ \forall z_4{\in}B_{r_4}(y)\,Exz_4$

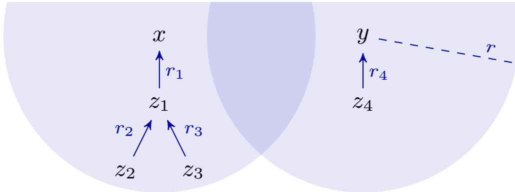

Figure 1 Example of a local formula and the corresponding quantification dag $D(\varphi)$, with circles indicating $B_r(xy)$. In this example, $\varphi^{(r)}(x,y)$ is r-local for all $r \geq \max(r_1+r_2, r_1+r_3, r_4)$.

This alone is not as expressive as the Boolean notion. For instance, consider $\varphi^{(r)}(x) = \exists y(d(x,y) \leq \frac{r}{2} \wedge \exists z(d(x,z) \leq r \wedge d(y,z) \leq \frac{r}{2} \wedge \dots))$ which quantifies z local around y. Using ball quantifiers, we want to write this as $\varphi^{(r)}(x) = \exists y{\in}B_{\frac{r}{2}}(x)(\exists z{\in}B_{\frac{r}{2}}(y)\dots)$, so we also permit ball quantifiers around previously quantified variables (here y), as long as they stay within the r-neighbourhood of x (here: $\frac{r}{2}+\frac{r}{2} \leq r$).

To formalise this condition, we consider the *quantification dag* $D(\varphi)$ of a formula $\varphi(\overline{x})$ which contains nodes for all variables in φ and where for every quantifier $Qz{\in}B_{r'}(y)$ in φ, we add an edge $z \to y$ with distance label r' (see Figure 1). If the summed distance of any path ending in a free variable $x \in \overline{x}$ is at most r, then φ is r-local.

▶ **Definition 9** (Local formula). *An* r-local τ-formula around $\overline{x}$*, denoted* $\varphi^{(r)}(\overline{x})$*, is built from* τ*-literals by means of* $\wedge$*,* $\vee$ *and ball quantifiers* $Qz{\in}B^{\tau}_{r'}(y)$ *such that in the associated quantification dag* $D(\varphi)$*, all paths ending in a free variable* $x \in \overline{x}$ *have total length at most* r*.*

We emphasise that in the Boolean case, Definition 9 is equivalent to the standard notion, so we do not add expressive power. For convenience, we allow quantification $Qz{\in}B_{r'}(\overline{y})\,\varphi(\overline{y},z)$ around a tuple $\overline{y}$, which can easily be simulated by regular ball quantifiers.

For basic local sentences, we further need to quantify over scattered tuples. To this end, we also add *scattered quantifiers* $\exists^{r\text{-sc}}(\overline{y})$ and $\forall^{r\text{-sc}}(\overline{y})$ with the following semantics:

$$\pi[\![\exists^{r\text{-sc}}(\overline{y})\,\varphi(\overline{y})]\!] = \sum_{\substack{\overline{a}\subseteq A\\ d(a_i,a_j)>2r \text{ for } i\neq j}} \pi[\![\varphi(\overline{a})]\!], \qquad \pi[\![\forall^{r\text{-sc}}(\overline{y})\,\varphi(\overline{y})]\!] = \prod_{\substack{\overline{a}\subseteq A\\ d(a_i,a_j)>2r \text{ for } i\neq j}} \pi[\![\varphi(\overline{a})]\!].$$

We remark that in idempotent semirings, which will be the main focus of our positive results, the addition of ball quantifiers makes it possible to express $d(x,y) \leq r$ by a formula that only assumes values 0 or 1, such as $\exists x'{\in}B_{\frac{r}{2}}(x)\,\exists y'{\in}B_{\frac{r}{2}}(y)\,(x'=y')$, which is $\frac{r}{2}$-local around xy, or alternatively $\exists x'{\in}B_r(x)\,(x'=y)$, which is r-local only around x. Analogously for $d(x,y) > r$, so we permit the use of distance formulae to simplify notation whenever we work in idempotent semirings. Scattered quantifiers can then easily be expressed as $\exists^{r\text{-sc}}(y_1,\dots y_m)\,\vartheta(\overline{y}) := \exists y_1 \dots \exists y_m\big(\bigwedge_{i<j} d(y_i,y_j) > 2r \wedge \vartheta(\overline{y})\big)$ and $\forall^{r\text{-sc}}(y_1,\dots y_m)\,\vartheta(\overline{y}) := \forall y_1 \dots \forall y_m\big(\bigvee_{i<j} d(y_i,y_j) \leq 2r \vee \vartheta(\overline{y})\big)$.

▶ **Definition 10** (Local sentence). *A* basic local sentence *is a sentence of the form*

$$\exists^{r\text{-}sc}(y_1,\dots y_m) \bigwedge_{i\leq m} \varphi^{(r)}(y_i) \quad \textit{or} \quad \forall^{r\text{-}sc}(y_1,\dots y_m) \bigvee_{i\leq m} \varphi^{(r)}(y_i).$$

A local sentence *is a positive Boolean combination of basic local sentences.*

Based on these notions we can now formulate precisely the questions about Gaifman normal forms in semiring semantics:

(1) For which semirings K does every first-order *sentence* have a K-equivalent local sentence?
(2) For which semirings K is it the case that every first-order *formula* is K-equivalent to a positive Boolean combination of local formulae and basic local sentences?

5 Counterexamples Against Gaifman Normal Forms

This section presents two examples for which a Gaifman normal form does not exist. Both use the vocabulary $\tau = \{U\}$ with only unary predicates, so that the Gaifman graph $G(\pi)$ of any K-interpretation $\pi\colon \mathrm{Lit}_A(\tau) \to K$ is trivial and the r-neighbourhood of a point, for any r, consists only of the point itself. Thus, local formulae $\varphi^{(r)}(x)$ around x can always be written as positive Boolean combinations of literals Ux, $\neg Ux$ and equalities $x = x$, $x \neq x$. Scattered tuples are simply distinct tuples, so we write $\exists^{\text{distinct}}(\overline{x})$ instead of $\exists^{r\text{-sc}}(\overline{x})$.

5.1 A Formula Without a Gaifman Normal Form

Consider the formula $\psi(x) \coloneqq \exists y(Uy \wedge y \neq x)$ which, in classical Boolean semantics, has the Gaifman normal form $\varphi(x) \coloneqq \exists^{\text{distinct}}(y, z)(Uy \wedge Uz) \vee (\neg Ux \wedge \exists y Uy)$. However, in semiring semantics it is in general not the case that $\psi(x) \equiv_K \varphi(x)$. Here we consider the specific case of a universe with two elements $A = \{a, b\}$ and K-interpretations π_{st} with $\pi_{st}(Ua) = s$ and $\pi_{st}(Ub) = t$, where $s, t \in K \setminus \{0\}$ and $s \neq t$. Then $\pi_{st}[\![\psi(a)]\!] = t$ but $\pi_{st}[\![\varphi(a)]\!] = st + ts$. So, unless K is the Boolean semiring, we find elements s, t where $\pi_{st}[\![\psi(a)]\!] \neq \pi_{st}[\![\varphi(a)]\!]$.

Of course, it might still be the case that there is a different Gaifman normal form of $\psi(x)$ for semiring interpretations in a specific semiring K. We prove that this is not the case.

▶ **Proposition 11.** *In any naturally ordered semiring with at least three elements, the formula $\psi(x) = \exists y(Uy \wedge y \neq x)$ does not have a Gaifman normal form.*

For the proof, we describe the values that the building blocks of Gaifman normal forms may assume in π_{st}. Recall that a local formula $\alpha(x)$ is equivalent to a positive Boolean combination of literals Ux, $\neg Ux$, and equalities. Since $\pi_{st}(\neg Ux) = 0$ for all $x \in A$, we can view the evaluation $\pi_{st}[\![\alpha(a)]\!]$ as an expression built from the semiring operations, the value $\pi_{st}(Ua) = s$ and constants $0, 1$. Analogously for $\pi_{st}[\![\alpha(b)]\!]$, but using $\pi_{st}(Ub) = t$ instead of s. Hence there is a polynomial $p_\alpha(X) \in K[X]$ such that $\pi_{st}[\![\alpha(a)]\!] = p_\alpha(s)$ and $\pi_{st}[\![\alpha(b)]\!] = p_\alpha(t)$, for all interpretations π_{st}. For the evaluation of a basic local sentence $\beta = \exists^{\text{distinct}}(y, z)(\alpha(y) \wedge \alpha(z))$, we then obtain $\pi_{st}[\![\beta]\!] = p_\alpha(s)p_\alpha(t) + p_\alpha(t)p_\alpha(s)$. That is, β can be described by a polynomial $p_\beta(X, Y) \in K[X, Y]$ such that $\pi_{st}[\![\beta]\!] = p_\beta(s, t)$ and p_β is *symmetric* (that is, $p_\beta(X, Y) = p_\beta(Y, X)$). The same holds for universal basic local sentences $\beta = \forall^{\text{distinct}}(y, z)(\alpha(y) \vee \alpha(z))$.

Every Gaifman normal form $\varphi(x)$ can thus be represented by a polynomial $f_\varphi(X, Y) = \sum_i h_i(X)g_i(X, Y)$, with symmetric g_i, such that $\pi_{st}[\![\varphi(a)]\!] = f_\varphi(s, t)$ for all s, t. Proposition 11 then follows from the following algebraic observation (see [6] for a proof).

▶ **Lemma 12.** *Let K be a naturally ordered semiring with at least three elements. For any polynomial $f(X, Y) = \sum_i h_i(X)g_i(X, Y)$ where the g_i are symmetric polynomials, there exist values $s, t \in K \setminus \{0\}$ such that $f(s, t) \neq t$.*

5.2 A Sentence Without a Gaifman Normal Form

While Gaifman normal forms need not exist for formulae, in all relevant semirings beyond the Boolean one, they might still exist for sentences. Indeed, we shall prove a positive result for min-max semirings. However, such a result seems only possible for semirings where both operations are idempotent, similar to Hanf's theorem. For other semirings one can find rather simple counterexamples, as we illustrate for the tropical semiring $\mathbb{T} = (\mathbb{R}_+^\infty, \min, +, \infty, 0)$.

▶ **Proposition 13.** *The sentence $\psi \coloneqq \exists z \forall x \exists y(Uy \vee x = z)$ has no Gaifman normal form in the tropical semiring.*

The proof again works by describing the values of basic local sentences, this time in $\mathbb{T}$-interpretations of increasingly large size. One can then show that these values are either constant or grow too fast, compared to the value of ψ (see [6] for details). A similar construction works for the natural semiring $(\mathbb{N}, +, \cdot, 0, 1)$ and we conjecture that it can be adapted to any infinite semiring with operations that are not idempotent.

6 Gaifman's Theorem for Min-Max Semirings

In this section, we prove our main result: a version of Gaifman's theorem for sentences evaluated in min-max semirings (which can be lifted to lattice semirings, see Sect. 7). We write $\mathcal{M}$ for the class of min-max semirings and refer to $\equiv_{\mathcal{M}}$ as *minmax-equivalence*.

▶ **Theorem 14** (Gaifman normal form). *Let τ be a finite relational signature. Every* $\mathrm{FO}(\tau)$*-sentence ψ is minmax-equivalent ($\equiv_{\mathcal{M}}$) to a local sentence.*

Contrary to Hanf's locality theorem, we cannot follow the classical proofs of Gaifman's theorem. For instance, the proof in [9] is based on the Ehrenfeucht-Fraïssé method and makes use of characteristic sentences, which in general do not exist in semiring semantics over min-max semirings (cf. [15]). Gaifman's original proof [11] is a constructive quantifier elimination argument (which is similar to our approach), but makes use of negation to encode case distinctions in the formula, which is not possible in semiring semantics. Another argument why Gaifman's proof does not go through is that it applies to formulae, whereas formulae need not have Gaifman normal forms in our setting (cf. Sect. 5.1).

Instead, we present a novel proof of Gaifman's theorem that applies to the Boolean case as well as to min-max semirings. While our strategy is similar to Gaifman's – a constructive elimination of quantifier alternations – we have to phrase all results in terms of sentences and need to be more careful to derive equivalences that hold in all min-max semirings. These restrictions lead to a slight strengthening of Gaifman's classical result (see Sect. 7).

6.1 Toolbox

The proof is rather technical, but is based on a few simple observations. First notice that min-max semirings share many algebraic properties with the Boolean semiring. As a consequence, many classical logical equivalences are also minmax-equivalences, such as distributivity or idempotence. In particular, we can make use of disjunctive normal forms, conjunctive normal forms and prenex normal forms. Moreover, we can exploit the inherent symmetry of min-max semirings to simplify our proofs: arguments for existential sentences can be dualised for universal sentences (see [6] for details). However, we still have to consider quantifier alternations, which pose the main challenge.

Concerning locality, we make two simple but crucial observations. For the first one, consider a local formula $\varphi^{(r)}(x, y)$ around two variables x and y. Such a formula may assert that x and y are close to each other, for instance $\varphi^{(r)}(x, y) = Exy$. But if x and y do not occur together within one literal, then $\varphi^{(r)}$ intuitively makes independent statements about the neighbourhood of x, and the neighbourhood of y, so we can split $\varphi^{(r)}$ into two separate local formulae. For the general case $\varphi^{(r)}(\overline{x})$ in several variables, we group $\overline{x}$ into tuples $\overline{x}^1, \dots, \overline{x}^n$ with the idea that $\varphi^{(r)}$ makes independent statements about each group $\overline{x}^i$.

▶ **Lemma 15** (Separation). *Let $\varphi^{(r)}(\overline{x}^1, \dots, \overline{x}^n)$ be a local formula around $\overline{x}^1 \dots \overline{x}^n$ and define X_i as the set of variables connected to some $x \in \overline{x}^i$ in $D(\varphi)$. If each literal of $\varphi^{(r)}(\overline{x}^1, \dots, \overline{x}^n)$ uses only variables in $\overline{x}^i \cup X_i$ for a single i, then $\varphi^{(r)}(\overline{x}^1, \dots, \overline{x}^n)$ is minmax-equivalent to a positive Boolean combination of r-local formulae around a single group $\overline{x}^i$.*

The second observation is that we can perform a clustering of any tuple $(a_1, \dots, a_n) \in A^n$ into classes $I_1, \dots, I_k$ so that elements within one class have "small" distance to each other, whereas different classes are "far apart". This simple combinatorial observation is a fruitful tool to construct Gaifman normal forms: it becomes easy to quantify elements with a known clustering, and by the following lemma we can then do a disjunction over all clusterings.

▶ **Definition 16** (Configuration). *Let π be a K-interpretation with universe A. Let $P = \{I_1, \dots, I_k\}$ be a partition of $\{1, \dots, n\}$ and define representatives $i_l = \min I_l$ of each class. We say that a tuple $(a_1, \dots, a_n) \in A^n$ is in* configuration *(P, r), if*

(a) *$d(a_{i_l}, a_i) \leq 5^{n-k}r - r$, for all $i \in I_l$, $l \in \{1, \dots, k\}$, and*

(b) *$d(a_{i_l}, a_{i_{l'}}) > 4 \cdot 5^{n-k}r$, for all $l \neq l'$ (representatives are $(2 \cdot 5^{n-k}r)$-scattered).*

Such a partition always exists: condition **(a)** remains true if we merge two classes violating **(b)**, so starting from $P = \{\{1\}, \dots, \{n\}\}$ we can merge classes until **(b)** holds.

▶ **Lemma 17** (Clustering). *Let π be a K-interpretation on A. For all tuples $(a_1, \dots, a_n) \in A^n$ and all $r \geq 1$, there is a partition P such that $(a_1, \dots, a_n)$ is in configuration (P, r).*

6.2 Proof Outline for Gaifman's Theorem

The heart of our proof is the elimination of quantifier alternations. Due to space reasons, we refer to the full version [6] for detailed proofs. Here we present an overview of the main steps. Each steps proves, building on the previous ones, that sentences of a certain fragment can be translated to minmax-equivalent local sentences. These fragments consist of

(1) sentences of the form $\exists^{r\text{-sc}}(x_1, \dots x_m) \bigwedge_{i \leq m} \varphi_i^{(r)}(x_i)$;

(2) existential sentences $\exists \overline{x}\, \varphi^{(r)}(\overline{x})$;

(3) existential-universal sentences $\exists \overline{y} \forall \overline{x}\, \varphi^{(r)}(\overline{y}, \overline{x})$;

(4) all first-order sentences (Theorem 14).

We first note that Theorem 14 (step **(4)**) is a rather simple consequence of **(2)** and **(3)**. By applying **(3)** and putting the resulting local sentence in prenex normal form, we can bring $\exists^*\forall^*$-sentences into $\forall^*\exists^*$-form. We can thus inductively[3] eliminate quantifier alternations by swapping quantifiers, until at most one alternation remains and **(2)** or **(3)** apply directly.

For step **(1)**, note that the difference to a basic local sentence is that we permit different local formulae $\varphi_i^{(r)}$ for each x_i (such sentences have been called *asymmetric* in [20, 8]). Our proof is an inductive construction of the equivalent local sentence. This step is quite technical, but greatly simplifies the following constructions.

To prove **(2)**, we have to rewrite the $\exists^*$-prefix as a scattered quantifier $\exists^{r\text{-sc}}(\overline{x})$. This essentially follows from the Clustering and Separation Lemmas: for a given partition $P = \{I_1, \dots, I_k\}$ we can do a scattered quantification of the representatives $x_{i_1}, \dots, x_{i_k}$, and then quantify the elements of each class I_l locally around its representative x_{i_l}.

Step **(3)** is the core of the elimination argument and the most difficult step of the proof. We roughly follow the structure of Gaifman's proof [11] and, for a sentence $\exists \overline{y} \forall \overline{x}\, \varphi^{(r)}(\overline{y}, \overline{x})$, first split $\forall \overline{x}$ into those elements close to $\overline{y}$ (which we can quantify locally within φ) and those elements far from $\overline{x}$, using the Separation Lemma. Eventually, we arrive at a positive Boolean combination of sentences $\exists \overline{y} (\varphi_{\text{close}}^{(r)}(\overline{y}) \wedge \forall x {\notin} B_s(\overline{y})\, \varphi_{\text{far}}^{(s)}(x))$. Here, the far elements

[3] An attentive reader may notice that we have to deal with formulae with free variables in the induction, but **(2)**, **(3)** only apply to sentences. We resolve this issue by temporarily substituting atoms with free variables by fresh relation symbols, without affecting the Gaifman graph (see [6, Abstraction Lemma]).

are covered by the *outside quantifier* $\forall x \notin B_s(\overline{y})$ with the obvious semantics. As in Gaifman's proof, the main challenge is the elimination of this outside quantifier. Gaifman approaches this by using negation to encode case distinctions in the Gaifman normal form. Our proof instead consists of a series of surprisingly difficult syntactical transformations that avoid negation, eventually leading to a minmax-equivalent local sentence without outside quantifiers.

7 Strengthening Gaifman's Theorem

In this section, we rephrase our main result in terms of Boolean semantics, which leads to a novel strengthening of Gaifman's classical theorem. Interestingly, Theorem 14 can be regained from the Boolean result by algebraic techniques, and even lifted to lattice semirings. These insights suggest that the merit of our proof, and the reason why it is more complicated than Gaifman's original proof, is the construction of a Gaifman normal form without the use of negation. Since our proof applies in particular to the Boolean semiring and hence to standard Boolean semantics (the only difference is that we use ball quantifiers instead of distance formulae, but these are interchangeable), we obtain the following corollary.

▶ **Corollary 18** (Gaifman normal form without negation)**.** *Let τ be a finite relational signature. In Boolean semantics, every* $\mathrm{FO}(\tau)$*-sentence ψ has an equivalent local sentence ψ' such that every relation symbol occurring only positively (only negatively) in ψ also occurs only positively (only negatively) in ψ', not counting occurrences within distance formulae.*

We believe that this result may be of independent interest. A similar adaptation of Gaifman's theorem has been considered in [20], namely that *existential* sentences are equivalent to *positive* Boolean combinations of *existential* basic local sentences. Our proof of step **(2)** implies a similar result (cf. [6]), as we also construct a positive Boolean combination of existential basic local sentences. However, we permit distance formulae $d(x, y) > 2r$ within local formulae (which are abbreviations for universal quantifiers), while [20] does not. Moreover, the approximation schemes of [8] are based on a version of Gaifman's theorem for sentences positive in a single unary relation (i.e., no negations are added in front of this relation). Their proof uses a version of Ehrenfeucht-Fraïssé games, which is quite different from our syntactical approach. Since unary relations do not occur in distance formulae, Corollary 18 subsumes their result. Interestingly, [20, 8] both share our observation that the proof of the respective version of Gaifman's theorem is surprisingly difficult.

To prove Theorem 14 from Corollary 18, one can show (cf. [6]) that with some preparation, Boolean equivalences $\equiv_{\mathbb{B}}$ can be lifted to lattice-equivalences $\equiv_{\mathcal{L}}$ (which subsume $\equiv_{\mathcal{M}}$). This is done by applying *separating homomorphisms* of [15] to turn a falsifying K-interpretation π, witnessing $\not\equiv_{\mathcal{L}}$, into a falsifying Boolean structure $h \circ \pi$, witnessing $\not\equiv_{\mathbb{B}}$. Such homomorphisms h exist for all min-max semirings [15] and also for the more general lattice semirings [7, 6]. We obtain the following generalisation of Theorem 14 by lifting the Boolean result.

▶ **Corollary 19.** *Let τ be a finite relational signature. Every* $\mathrm{FO}(\tau)$*-sentence ψ is lattice-equivalent ($\equiv_{\mathcal{L}}$) to a local sentence.*

We remark that the lifting argument implies that for many sentences (to be precise, those where no relation occurs both positively and negatively, cf. [6]), the Gaifman normal form in min-max and lattice semirings coincides with the one for Boolean semantics in Corollary 18 (but not necessarily with Gaifman's original construction). A further consequence is that the counterexample for formulae in Sect. 5.1 also applies to Corollary 18.

8 Conclusion

Semiring semantics is a refinement of classical Boolean semantics, which provides more detailed information about a logical statement than just its truth or falsity. This leads to a finer distinction between formulae: statements that are equivalent in the Boolean sense may have different valuations in semiring interpretations, depending on the underlying semiring. It is an interesting and non-trivial question, which logical equivalences and, more generally, which model-theoretic methods, can be carried over from classical semantics to semiring semantics, and how this depends on the algebraic properties of the underlying semiring.

Here we have studied this question for locality properties of first-order logic, in particular for Hanf's locality theorem and for Gaifman normal forms. Our setting assumes semiring interpretations which are model-defining and track only positive information, since these are the conditions that provide well-defined and meaningful locality notions. However, from the outset, it has been clear that one cannot expect to transfer all locality properties of first-order logic to semiring semantics in arbitrary commutative semirings. Indeed, semiring semantics evaluates existential and universal quantifiers by sums and products over all elements of the universe, which gives an inherent source of non-locality if these operations are not idempotent.

Most positive locality results thus require that the underlying semirings are fully idempotent. Under this assumption, one can adapt the classical proof of Hanf's locality theorem to the semiring setting, relying on a back-and-forth argument that itself requires fully idempotent semirings. The question whether there exist Gaifman normal forms in semiring semantics turned out to be more subtle. Indeed, for formulae with free variables Gaifman normal forms need not exist once one goes beyond the Boolean semiring. Also for sentences, one can find examples that do not admit Gaifman normal forms in semirings that are not fully idempotent. We have presented such an example for the tropical semiring.

Our main result, however, is a positive one and establishes the existence of Gaifman normal forms over the class of all min-max and lattice semirings. Intuitively, it relies on the property that in min-max semirings, the value of a quantified statement $\exists x\, \varphi(x)$ or $\forall x\, \varphi(x)$ coincides with a value of $\varphi(a)$, for some witness a. This needs, for instance, not be the case in lattice semirings, and hence the generalisation to lattice semirings uses a different approach based on separating homomorphisms. It is still an open question whether, in analogy to Hanf's theorem, Gaifman normal forms exist over all fully idempotent semirings. The proof of our main result, which is based on quantifier elimination arguments, turned out to be surprisingly difficult; we identified the lack of a classical negation operator as the main reason for its complexity. An interesting consequence of this restriction is a stronger version of Gaifman's classical theorem in Boolean semantics: every sentence has a Gaifman normal form which, informally speaking, does not add negations.

For applications such as provenance analysis, min-max semirings are relevant, for instance, for studying access levels and security issues. A much larger interesting class of semirings with wider applications are the absorptive ones, including the tropical semiring, in which addition is idempotent, but multiplication in general is not. We have seen that Gaifman normal forms for such semirings need not exist for all sentences. The question arises whether one can establish weaker locality properties for absorptive semirings, applicable perhaps to just a relevant fragment of first-order logic.

References

1. M. Arenas, P. Barceló, and L. Libkin. Game-based notions of locality over finite models. *Ann. Pure Appl. Log.*, 152(1-3):3–30, 2008. doi:10.1016/j.apal.2007.11.012.
2. M. Benedikt, T. Griffin, and L. Libkin. Verifiable properties of database transactions. *Inf. Comput.*, 147(1):57–88, 1998. doi:10.1006/inco.1998.2731.

3 S. van Bergerem. Learning concepts definable in first-order logic with counting. In *34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019*, pages 1–13, 2019. doi:10.1109/LICS.2019.8785811.

4 S. van Bergerem, M. Grohe, and M. Ritzert. On the parameterized complexity of learning first-order logic. In *PODS '22: International Conference on Management of Data*, pages 337–346. ACM, 2022. doi:10.1145/3517804.3524151.

5 S. van Bergerem and N. Schweikardt. Learning concepts described by weight aggregation logic. In *29th EACSL Annual Conference on Computer Science Logic, CSL 2021*, pages 10:1–10:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. doi:10.4230/LIPIcs.CSL.2021.10.

6 C. Bizière, E. Grädel, and M. Naaf. Locality theorems in semiring semantics, 2023. Full version of this paper. arXiv:2303.12627.

7 S. Brinke. *Ehrenfeucht-Fraïssé Games for Semiring Semantics*. M. Sc. thesis, RWTH Aachen University, 2023. URL: https://logic.rwth-aachen.de/pub/brinke/masterarbeit.pdf.

8 A. Dawar, M. Grohe, S. Kreutzer, and N. Schweikardt. Approximation schemes for first-order definable optimisation problems. In *21th IEEE Symposium on Logic in Computer Science (LICS 2006)*, pages 411–420. IEEE, 2006. doi:10.1109/LICS.2006.13.

9 H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer, 2nd edition, 1995. doi:10.1007/3-540-28788-4.

10 J. Foster, T. Green, and V. Tannen. Annotated XML: queries and provenance. In *Proceedings of PODS*, pages 271–280, 2008. doi:10.1145/1376916.1376954.

11 H. Gaifman. On local and non-local properties. In J. Stern, editor, *Proceedings of the Herbrand Symposium*, volume 107 of *Studies in Logic and the Foundations of Mathematics*, pages 105–135. Elsevier, 1982. doi:10.1016/S0049-237X(08)71879-2.

12 B. Glavic. Data provenance. *Foundations and Trends in Databases*, 9(3-4):209–441, 2021. doi:10.1561/1900000068.

13 E. Grädel, H. Helal, M. Naaf, and R. Wilke. Zero-one laws and almost sure valuations of first-order logic in semiring semantics. In Christel Baier and Dana Fisman, editors, *LICS '22: 37th Annual ACM/IEEE Symposium on Logic in Computer Science, Haifa, Israel, August 2 - 5, 2022*, pages 41:1–41:12. ACM, 2022. doi:10.1145/3531130.3533358.

14 E. Grädel, N. Lücking, and M. Naaf. Semiring provenance for Büchi games: Strategy analysis with absorptive polynomials. In *Proceedings 12th International Symposium on Games, Automata, Logics, and Formal Verification (GandALF 2021)*, volume 346 of *EPTCS*, pages 67–82, 2021. doi:10.4204/EPTCS.346.5.

15 E. Grädel and L. Mrkonjić. Elementary equivalence versus isomorphism in semiring semantics. In *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*, volume 198, pages 133:1–133:20, 2021. doi:10.4230/LIPIcs.ICALP.2021.133.

16 E. Grädel and V. Tannen. Semiring provenance for first-order model checking, 2017. arXiv:1712.01980.

17 E. Grädel and V. Tannen. Provenance analysis for logic and games. *Moscow Journal of Combinatorics and Number Theory*, 9(3):203–228, 2020. doi:10.2140/moscow.2020.9.203.

18 T. Green and V. Tannen. The semiring framework for database provenance. In *Proceedings of PODS*, pages 93–99. ACM, 2017. doi:10.1145/3034786.3056125.

19 M. Grohe, S. Kreutzer, and S. Siebertz. Deciding first-order properties of nowhere dense graphs. *J. ACM*, 64(3):17:1–17:32, 2017. doi:10.1145/3051095.

20 M. Grohe and S. Wöhrle. An existential locality theorem. *Ann. Pure Appl. Log.*, 129(1-3):131–148, 2004. doi:10.1016/j.apal.2004.01.005.

21 L. Hella, L. Libkin, and J. Nurmonen. Notions of locality and their logical characterizations over finite models. *J. Symb. Log.*, 64(4):1751–1773, 1999. doi:10.2307/2586810.

22 L. Hella, L. Libkin, J. Nurmonen, and L. Wong. Logics with aggregate operators. *J. ACM*, 48(4):880–907, 2001. doi:10.1145/502090.502100.

23 D. Kuske and N. Schweikardt. Gaifman normal forms for counting extensions of first-order logic. In *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 133:1–133:14, 2018. doi:10.4230/LIPIcs.ICALP.2018.133.

24 L. Libkin. *Elements of Finite Model Theory*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004. doi:10.1007/978-3-662-07003-1.

25 T. Schwentick and K. Barthelmann. Local normal forms for first-order logic with applications to games and automata. In *STACS 98, 15th Annual Symposium on Theoretical Aspects of Computer Science, Paris, France, February 25-27, 1998, Proceedings*, volume 1373 of *Lecture Notes in Computer Science*, pages 444–454. Springer, 1998. doi:10.1007/BFb0028580.

26 J. Xu, W. Zhang, A. Alawini, and V. Tannen. Provenance analysis for missing answers and integrity repairs. *IEEE Data Eng. Bull.*, 41(1):39–50, 2018. URL: http://sites.computer.org/debull/A18mar/p39.pdf.