Semiring Provenance for Fixed-Point Logic

Katrin M. Dannert
RWTH Aachen University, Germany
dannert@logic.rwth-aachen.de

Erich Grädel
RWTH Aachen University, Germany
graedel@logic.rwth-aachen.de

Matthias Naaf
RWTH Aachen University, Germany
naaf@logic.rwth-aachen.de

Val Tannen
University of Pennsylvania, Philadelphia, PA, USA
val@cis.upenn.edu

Abstract

Semiring provenance is a successful approach, originating in database theory, to providing detailed information on how atomic facts combine to yield the result of a query. In particular, general provenance semirings of polynomials or formal power series provide precise descriptions of the evaluation strategies or "proof trees" for the query. By evaluating these descriptions in specific application semirings, one can extract practical information for instance about the confidence of a query or the cost of its evaluation.

This paper develops semiring provenance for very general logical languages featuring the full interaction between negation and fixed-point inductions or, equivalently, arbitrary interleavings of least and greatest fixed points. This also opens the door to provenance analysis applications for modal μ-calculus and temporal logics, as well as for finite and infinite model-checking games.

Interestingly, the common approach based on Kleene's Fixed-Point Theorem for ω-continuous semirings is not sufficient for these general languages. We show that an adequate framework for the provenance analysis of full fixed-point logics is provided by semirings that are (1) *fully continuous*, and (2) *absorptive*. Full continuity guarantees that provenance values of least and greatest fixed-points are well-defined. Absorptive semirings provide a symmetry between least and greatest fixed-points and make sure that provenance values of greatest fixed points are informative.

We identify *semirings of generalized absorptive polynomials* $\mathbb{S}^{\infty}[X]$ and prove universal properties that make them the most general appropriate semirings for our framework. These semirings have the further property of being (3) *chain-positive*, which is responsible for having *truth-preserving* interpretations that give non-zero values to all true formulae. We relate the provenance analysis of fixed-point formulae with provenance values of plays and strategies in the associated model-checking games. Specifically, we prove that the provenance value of a fixed point formula gives precise information on the evaluation strategies in these games.

2012 ACM Subject Classification Theory of computation → Finite Model Theory

Keywords and phrases Finite Model Theory, Semiring Provenance, Absorptive Semirings, Fixed-Point Logics

Digital Object Identifier 10.4230/LIPIcs.CSL.2021.17

Related Version Due to space reasons, a number of technical details and proofs have been omitted. They are available in the full version of this paper at `https://arxiv.org/abs/1910.07910`.

29th EACSL Annual Conference on Computer Science Logic (CSL 2021).
Editors: Christel Baier and Jean Goubault-Larrecq; Article No. 17; pp. 17:1–17:22

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Provenance analysis for a logical statement ψ, evaluated on a finite structure $\mathfrak{A}$, aims at providing precise information why ψ is true or false in $\mathfrak{A}$. The approach of *semiring provenance*, going back to [16] relies on the idea of annotating the atomic facts by values from a commutative semiring, and to propagate these values through the statement ψ, keeping track whether information is used alternatively (as in disjunctions or existential quantifications) or jointly (as in conjunctions or universal quantifications). Depending on the chosen semiring, the provenance value may then give practical information for instance concerning the *confidence* we may have that $\mathfrak{A} \models \psi$, the *cost* of the evaluation of ψ on $\mathfrak{A}$, the number of successful evaluation strategies for ψ on $\mathfrak{A}$ in a game-theoretic sense, and so on. Beyond such provenance evaluations in specific application semirings, more general and more precise information is obtained by evaluations in so-called *provenance semirings* of polynomials or formal power series. Take, for instance, an abstract set X of *provenance tokens* that are used to label the atomic facts of a structure $\mathfrak{A}$, and consider the semiring $\mathbb{N}[X]$ of polynomials with indeterminates in X and coefficients from $\mathbb{N}$, which is the commutative semiring that is freely generated ("most general") over X. Such a labelling of the atomic facts then extends to a provenance valuation $\pi[\![\psi]\!] \in \mathbb{N}[X]$ for every Boolean query ψ from positive relational algebra RA^+ and, indeed, every negation-free first-order sentence $\psi \in \mathrm{FO}^+$. This provenance valuation gives precise information about the combinations of atomic facts that imply the truth of ψ in $\mathfrak{A}$. Indeed, we can write $\pi[\![\psi]\!]$ as a sum of monomials $m\, x_1^{e_1} \cdots x_k^{e_k}$. Each such monomial indicates that we have precisely m evaluation strategies (or "proof trees") to determine that $\mathfrak{A} \models \psi$ that make use of the atoms labelled by $x_1, \ldots x_k$, and the atom labelled by x_i is used precisely e_i times by the strategy, see [16, 13].

Provenance for least fixed points. A similar analysis has been carried out for Datalog [6, 16]. Due to the need of unbounded least fixed-point iterations in the evaluation of Datalog queries, the underlying semirings have to satisfy the additional property of being ω-continuous. By Kleene's Fixed-Point Theorem, systems of polynomial equations then have least fixed-point solutions that can be computed by induction, reaching the fixed-point after at most ω stages. Most of the common application semirings are ω-continuous, or can easily be extended to one that is so; however, the most general ω-continuous provenance semiring over X is no longer a semiring of polynomials but the semiring of formal power series over X, denoted $\mathbb{N}^\infty[\![X]\!]$, with coefficients in $\mathbb{N}^\infty := \mathbb{N} \cup \{\infty\}$. As above, provenance valuations $\pi[\![\psi]\!] \in \mathbb{N}^\infty[\![X]\!]$ give precise information about the possible evaluation strategies for a Datalog query ψ on $\mathfrak{A}$. Even though $\mathfrak{A}$ is assumed to be finite there may be infinitely many such strategies, but each of them can use each atomic fact only a finite number of times; to put it differently, "proof trees" for $\mathfrak{A} \models \psi$ are still finite. This is closely related to the provenance analysis of reachability games on finite graphs [6, 14].

Negation: a stumbling block for wider applications. Semiring provenance has been applied to a number of other scenarios, such as nested relations, XML, SQL-aggregates, graph databases (see, e.g., the survey [17] as well as [21, 22]), and it is fair to say that in databases, semiring provenance analysis has been rather successful. However, its impact outside of databases has been very limited, despite the fact that the main questions addressed by provenance analysis, namely which parts of a large heterogeneous input structure are responsible for the evaluation of a logical statement, and the applications to cost, confidence, access control and so on are clearly interesting and relevant in many other branches of logics

in computer science. The main obstacle for extending semiring provenance to such fields have been difficulties with handling negation. For a long time, semiring provenance has essentially been restricted to negation-free query languages, and although there have been algebraically interesting attempts to cover difference of relations [1, 10, 11, 15], they have not resulted in a systematic tracking of *negative information*. While there are many applications in databases where one can get quite far with using positive information only, logical applications in most other areas are based on formalisms that use negation in an essential way, often in combination with recursion or fixed-points.

Provenance semirings for logics with negation and recursion. This paper is part of a larger project with the objectives to

- develop semiring provenance systematically for a wide range of logics, including those featuring the notoriously difficult interaction between full negation and recursion,
- to employ algebraic methods for provenance analysis, in particular universal semirings of polynomials to obtain the most general provenance information,
- to exploit the connections between logics and various kinds of games and to use semiring valuations for an analysis of strategies in such games, and
- to explore practical applications of semiring provenance in new areas of logics in computer science, where this has not been used so far, such as knowledge representation, verification, and machine learning.

This project has been initiated in [13], where a provenance analysis of full first-order logic has been proposed. In this approach, negation is dealt with by transformation into negation normal form and, algebraically, by new provenance semirings of dual-indeterminate polynomials, which are obtained by taking quotients of traditional semirings of polynomials, such as $\mathbb{N}[X]$, by congruences generated by products of positive and negative provenance tokens, see Sect. 2 for details. In particular, the semiring $\mathbb{N}[X, \bar{X}]$ of dual-indeterminate polynomials is the most general provenance semiring for full first-order logic FO. These ideas have been used in [3, 4] to provide a provenance analysis of modal and guarded fragments of first-order logic, and to explore applications in description logic. Further, this approach has been applied to database repairs in [23], and it has been shown how this treatment of negation, or absent information, can be used to explain and repair missing query answers and the failure of integrity constraints in databases.

While the connection between provenance analysis of first-order logic and semiring valuations of games had only been hinted at in [13], it has then been developed more systematically in [14], first for games on acyclic graphs, which admit only finite plays, and then also for reachability games on acyclic game graphs. The latter are tightly connected with least fixed-point inductions, used positively. Combining the approach from [13] with the provenance analysis of least fixed-point inductions in ω-continuous semirings of formal power series, one obtains, by an analogous quotient construction, the semiring $\mathbb{N}^{\infty}[\![X, \bar{X}]\!]$ of dual-indeterminate power series [14]. This is the most general provenance semiring for Datalog with negated input predicates and, more generally, also for posLFP, the fragment of the full fixed-point logic LFP that consists of formulae in negation normal form such that all its fixed-point operators are least fixed-points. This is a powerful fixed-point calculus, which suffices to capture all polynomial-time computable properties of ordered finite structures [12]. An important simplification of dealing with posLFP is that the game-based analysis of model checking only requires reachability games rather than the much more complicated parity games that are needed for full LFP. At the end of [14] the problem of generalising semiring valuations and strategy analysis to infinite games with more general

objectives than reachability has been discussed. In particular, a provenance approach for safety games has been proposed, with absorptive semirings as the central algebraic tools, and absorption-dominant strategies as a relevant game-theoretic notion.

Greatest fixed points. What has been missing so far, and what we want to provide in this paper, is an adequate and systematic treatment of greatest fixed points. There is a strong motivation for this: If provenance analysis should ever have an impact in fields such as verification (and we strongly believe it should) then dealing with greatest fixed points, e.g. for safety conditions or bisimulation, and with alternations between least and greatest fixed points is indispensable. The relevant formalisms in verification (such as LTL, CTL, μ-calculus etc.) are negation closed and based on both least and greatest fixed-points, with strict alternation hierarchies (even for finite structures), and without possibilities to eliminate greatest fixed-points. Even in finite model theory, where greatest fixed points can in principle be eliminated from LFP by means of the Stage Comparison Theorem [20, 12], it is usually not desirable to do so. Natural properties involving greatest fixed points (such as bisimilarity) would become very complicated to express, with the need to double the arity of the fixed-point variables. In addition, provenance valuations provide a refined semantics, and statements that are equivalent in the Boolean sense need not have the same provenance value. Therefore, we do not propose an approach that first tries to simplify formulae (e.g. by eliminating fixed-point alternations) and then computes semiring valuations for the translated formulae, but instead lay foundations of a provenance analysis for the general logics with arbitrary interleavings of least and greatest fixed points, such as full LFP or the modal μ-calculus (and for infinite games with more general objectives than reachability).

Provenance semirings for arbitrary fixed points. We first address the question, what kind of semirings are adequate for a meaningful and informative provenance analysis of unrestricted fixed point logics (Sect. 4). The common approach for dealing with least fixed point inductions, based on ω-continuous semirings and Kleene's Fixed-Point Theorem, is not sufficient to guarantee that both least and greatest fixed point are well-defined. Instead, we require that the semirings are *fully continuous* which means that every chain C has not only a supremum $\bigsqcup C$, but also an infimum $\bigsqcap C$, and that both semiring operations are compatible with these suprema and infima. For an informative provenance semantics, there is a second important condition that is connected with the symmetry between least and greatest fixed point computations. In the Boolean setting, fixed-point logic is based on complete lattices which are inherently symmetric. Moreover, conjunction and disjunction are dual in the sense that one leads to larger lattice elements while the other is decreasing. In the semiring setting, we compute fixed points with respect to the natural order induced by addition. The only constraint that relates this order with multiplication is distributivity, but this alone does not suffice to ensure a similar duality. We achieve this by requiring that the semiring is *absorptive*. This means that $a + ab = a$ for all a, b, and we shall see that this is equivalent with 1 being the greatest element or with multiplication being decreasing, giving us the desired duality with 0 and addition. As a result, absorptive and fully continuous semirings guarantee a well-defined and informative provenance semantics for arbitrary fixed-point formulae.

Generalized absorptive polynomials. For a most general provenance analysis, we further want the semiring semantics to be *truth-preserving*, which means that it gives non-zero values to true formulae. In positive semirings, this is guaranteed if infima of non-zero values are also non-zero, which we call *chain-positivity*. Our fundamental examples of absorptive, fully

continuous, and chain-positive semirings are the semirings $\mathbb{S}^\infty[X]$ of *generalized absorptive polynomials* and its dual-indeterminate version $\mathbb{S}^\infty[X, \bar{X}]$, as introduced in [14]. Informally such a polynomial is a sum of monomials, with possibly infinite exponents, that are maximal with respect to absorption. For instance a monomial $x^2y^\infty z$ occurring in a provenance value $\pi[\![\psi]\!]$ indicates an absorption-dominant evaluation strategy that uses the atom labelled by x twice, the atom labelled by y an infinite number of times, and the atom labelled by z once. This monomial absorbs all those that have larger exponents for all variables, such as for instance $x^3y^\infty z^\infty u$, but not, say, $x^\infty y^3$. Absorptive polynomials thus describe model-checking proofs or evaluation strategies with a minimal use of atomic facts. A precise definition and analysis of these semirings will be given in Sect. 5. We prove that they do indeed have universal properties (see Theorem 17) that make $\mathbb{S}^\infty[X, \bar{X}]$ the most general absorptive semiring for LFP and thus also an indispensable tool to prove general results about provenance semantics in absorptive, fully continuous semirings.

Game-theoretic analysis. In the final Sect. 6 we illustrate the power of provenance interpretations for LFP in absorptive, fully-continuous semirings, and particularly in $\mathbb{S}^\infty[X, \bar{X}]$ by relating them to provenance values of plays and strategies in the associated model-checking games which in this case are parity games. Specifically we prove that, as in the case of FO and posLFP, the provenance value of an LFP-formula φ gives precise information on the evaluation strategies in these games.

2 Preliminaries: Commutative Semirings

▶ **Definition 1.** A *commutative 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 $0 \cdot a = a \cdot 0 = 0$. It is *naturally ordered* if the relation $a \leq b :\Longleftrightarrow a + c = b$ for some $c \in K$ is a partial order. Further, a commutative semiring is *positive* if $a + b = 0$ implies $a = 0$ and $b = 0$ and if it has no divisors of 0 (i.e., $a \cdot b = 0$ implies that $a = 0$ and $b = 0$).

All semirings considered in this paper are commutative and naturally ordered (which excludes rings). In the following we just write "semiring" to denote a commutative, naturally ordered semiring. Standard semirings considered in provenance analysis are in fact also positive, but for an appropriate treatment of negation we need semirings (of dual-indeterminate polynomials or power series) that have divisors of 0.

Elements of semirings will be used as truth values for logical statements. The intuition is that $+$ describes the *alternative use* of information, as in disjunctions or existential quantifications, whereas $\cdot$ stands for the *joint use* of information, as in conjunctions or universal quantifications. Further, 0 is the value of false statements, whereas any element $a \neq 0$ of a semiring K stands for a "nuanced" interpretation of true. We briefly discuss some specific semirings that provide interesting information about a logical statement.

- The *Boolean semiring* $\mathbb{B} = (\{0, 1\}, \vee, \wedge, 0, 1)$ is the standard habitat of logical truth.
- $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$ is used for counting evaluation strategies for a logical statement.
- $\mathbb{T} = (\mathbb{R}_+^\infty, \min, +, \infty, 0)$ is called the *tropical* semiring. It can be used for measuring the cost of evaluation strategies.
- The *Viterbi* semiring $\mathbb{V} = ([0, 1], \max, \cdot, 0, 1)$ is used to compute *confidence scores* for logical statements. It is in fact isomorphic to $\mathbb{T}$.
- The *min-max* semiring on a totally ordered set $(A, \leq)$ with least element a and greatest element b is the semiring $(A, \max, \min, a, b)$.

Beyond these *application semirings*, (most general) abstract provenance can be calculated in freely generated (universal) *provenance semirings* of polynomials or formal power series. The abstract provenance can then be specialised via homomorphisms to provenance values in different application semirings as needed.

- For any set X, the semiring $\mathbb{N}[X] = (\mathbb{N}[X], +, \cdot, 0, 1)$ consists of the multivariate polynomials in indeterminates from X with coefficients from $\mathbb{N}$. This is the commutative semiring freely generated by X. Admitting also infinite sums of monomials we obtain the semiring $\mathbb{N}^\infty[\![X]\!]$ of formal power series over X, with coefficients in $\mathbb{N}^\infty := \mathbb{N} \cup \{\infty\}$.
- Given two disjoint sets $X, \bar{X}$ of "positive" and "negative" provenance tokens, together with a one-to-one correspondence $X \leftrightarrow \bar{X}$, mapping each positive token x to its corresponding negative token $\bar{x}$, the semiring $\mathbb{N}[X, \bar{X}]$ is the quotient of the semiring of polynomials $\mathbb{N}[X \cup \bar{X}]$ by the congruence generated by the equalities $x \cdot \bar{x} = 0$ for all $x \in X$. This is the same as quotienting by the ideal generated by the polynomials $x\bar{x}$ for all $x \in X$. The congruence classes in $\mathbb{N}[X, \bar{X}]$ are in one-to-one correspondence with the polynomials in $\mathbb{N}[X \cup \bar{X}]$ such that none of their monomials contain complementary tokens. We call these *dual-indeterminate polynomials*. $\mathbb{N}[X, \bar{X}]$ is freely generated by $X \cup \bar{X}$ for homomorphisms such that $h(x) \cdot h(\bar{x}) = 0$. By a completely analogous quotient construction, we obtain the semiring $\mathbb{N}^\infty[\![X, \bar{X}]\!]$ of dual-indeterminate power series.

3 Provenance Semantics for Fixed-Point Logic

Semiring provenance is well understood for first-order logic and for logics with only least fixed points, used positively. To extend it to logics with arbitrary interleavings of least and greatest fixed points, we discuss the general fixed-point logic LFP that extends first-order logic by least and greatest fixed-point operators, but our insights also apply to weaker logics such as the modal μ-calculus, dynamic logics, or temporal logics such as CTL.

Least Fixed-Point Logic. Least fixed-point logic, denoted LFP, extends first order logic by least and greatest fixed points of definable monotone operators on relations: If $\psi(R, \mathbf{x})$ is a formula of vocabulary $\tau \cup \{R\}$, in which the relational variable R occurs only positively and the length of $\mathbf{x}$ matches the arity of R, then $[\mathbf{lfp}\, R\mathbf{x} \,.\, \psi](\mathbf{x})$ and $[\mathbf{gfp}\, R\mathbf{x} \,.\, \psi](\mathbf{x})$ are also formulae (of vocabulary τ). The semantics of these formulae is that $\mathbf{x}$ is contained in the least (respectively the greatest) fixed point of the update operator $F_\psi : R \mapsto \{\mathbf{a} : \psi(R, \mathbf{a})\}$. Due to the positivity of R in ψ, any such operator F_ψ is monotone and has, by the Knaster-Tarski-Theorem, a least fixed point $\mathbf{lfp}(F_\psi)$ and a greatest fixed point $\mathbf{gfp}(F_\psi)$. See e.g. [12] for background on LFP. The duality between least and greatest fixed points implies that $[\mathbf{gfp}\, R\mathbf{x} \,.\, \psi](\mathbf{x}) \equiv \neg[\mathbf{lfp}\, R\mathbf{x} \,.\, \neg\psi[R/\neg R]](\mathbf{x})$. By this duality together with de Morgan's laws, every LFP-formula can be brought into *negation normal form*, where negation applies to atoms only. The fragment posLFP of LFP consists of the formulae in negation normal form in which all fixed-point operators are least fixed points. It is well-known that LFP, and even posLFP, captures all polynomial-time computable properties of ordered finite structures [12].

Provenance Semantics. Instead of truth-values, we now assign semiring values to literals. For a finite universe A and a finite relational vocabulary τ we denote the set of atoms as $\mathrm{Atoms}_A(\tau) = \{R\mathbf{a} \colon R \in \tau,\ \mathbf{a} \in A^{\mathrm{arity}(R)}\}$. The set $\mathrm{NegAtoms}_A(\tau)$ contains all negations $\neg R\mathbf{a}$ of atoms in $\mathrm{Atoms}_A(\tau)$ and we define the set of τ-literals on A as

$$\mathrm{Lit}_A(\tau) := \mathrm{Atoms}_A(\tau) \cup \mathrm{NegAtoms}_A(\tau) \cup \{a = b \colon a, b \in A\} \cup \{a \neq b \colon a, b \in A\}.$$

▶ **Definition 2.** For any semiring K, a *K-interpretation* (for A and τ) is a function $\pi\colon \mathrm{Lit}_A(\tau) \to K$ mapping true equalities and inequalities to 1 and false ones to 0.

For a finite universe A, we can extend K-interpretations π to provide provenance values $\pi[\![\varphi]\!]$ for any first-order formula φ in a natural way [13], by interpreting disjunctions and existential quantification via addition, and conjunctions and universal quantification via multiplication. Negation is not interpreted directly by an algebraic operation. We deal with it syntactically, by evaluating the negation normal form $\mathrm{nnf}(\psi)$ instead. To interpret fixed-point formulae $[\mathbf{lfp}\, R\mathbf{x}\,.\,\psi](\mathbf{a})$ and $[\mathbf{gfp}\, R\mathbf{x}\,.\,\psi](\mathbf{a})$, we generalize the update operators F_ψ to semiring semantics. If R has arity m, then its K-interpretations on A are functions $g : A^m \to K$. These functions are ordered, by $g \leq g'$ if, and only if, $g(\mathbf{a}) \leq g'(\mathbf{a})$ for all $\mathbf{a} \in A^m$ (recall that our semirings are naturally ordered). Given a K-interpretation $\pi\colon \mathrm{Lit}_A(\tau) \to K$, we denote by $\pi[R \mapsto g]$ the K-interpretation of $\mathrm{Lit}_A(\tau) \cup \mathrm{Atoms}_A(\{R\})$ obtained from π by adding values $g(\mathbf{c})$ for the atoms $R\mathbf{c}$. (Notice that R appears only positively in φ, so negated R-atoms are not needed). The formula $\varphi(R, \mathbf{x})$ now defines, together with π, a monotone update operator F^φ_π on functions $g : A^m \to K$. More precisely, it maps g to the function

$$F^\varphi_\pi(g)\colon\ \mathbf{a} \mapsto \pi[R \mapsto g][\![\varphi(R, \mathbf{a})]\!].$$

We obtain a well-defined provenance semantics for LFP if we can make sure that the update operators F^φ_π have least and greatest fixed points $\mathbf{lfp}(F^\varphi_\pi)$, $\mathbf{gfp}(F^\varphi_\pi)\colon A^m \to K$. However, this is not guaranteed in all semirings, and also the common approach to least fixed-point inductions based on ω-continuous semirings is not sufficient here, as these, in general, do not guarantee the existence of *greatest* fixed points. This raises the fundamental question: which semirings are really appropriate for LFP? We shall discuss this in detail in the next section. Once we have fixed a notion of appropriate semirings for LFP, we obtain a provenance semantics for LFP as follows.

▶ **Definition 3.** A K-interpretation $\pi\colon \mathrm{Lit}_A(\tau) \to K$ (for finite A and τ) in an *appropriate* semiring K extends to a *K-valuation* $\pi\colon \mathrm{LFP}(\tau) \to K$ by mapping an LFP-sentence $\psi(\mathbf{a})$ to a value $\pi[\![\psi]\!]$ using the following rules

$$\pi[\![\psi \vee \varphi]\!] := \pi[\![\psi]\!] + \pi[\![\varphi]\!] \qquad \pi[\![\psi \wedge \varphi]\!] := \pi[\![\psi]\!] \cdot \pi[\![\varphi]\!] \qquad \pi[\![\exists x \psi(x)]\!] := \sum_{a \in A} \pi[\![\varphi(a)]\!]$$

$$\pi[\![\forall x \psi(x)]\!] := \prod_{a \in A} \pi[\![\varphi(a)]\!] \qquad \pi[\![[\mathbf{lfp}\, R\mathbf{x}.\varphi(R, \mathbf{x})](\mathbf{a})]\!] := \mathbf{lfp}(F^\varphi_\pi)(\mathbf{a})$$

$$\pi[\![\neg\psi]\!] := \pi[\![\mathrm{nnf}(\neg\psi)]\!] \qquad \pi[\![[\mathbf{gfp}\, R\mathbf{x}.\varphi(R, \mathbf{x})](\mathbf{a})]\!] := \mathbf{gfp}(F^\varphi_\pi)(\mathbf{a}).$$

We remark that there is an important difference between the classical Boolean semantics and provenance semantics concerning the relationship of fixed-point logics with first-order logic. The (Boolean) evaluation of a fixed-point formula on a finite structure is computed by fixed-point inductions that terminate after a polynomial number of stages (with respect to the size of the structure). Hence, on any fixed finite universe, a fixed-point formula can be unraveled to an equivalent first-order formula. This is not the case for the provenance valuations in infinite semirings. Even for very simple Datalog queries, a fixed-point induction need not terminate after a finite number of steps. Provenance valuations provide more information than just the truth or falsity of a statement, and in a general setting, this provenance information, for instance about the number and properties of successful evaluation strategies, may also be infinite.

4 Semirings for Fixed-Point Logic

Given a naturally ordered semiring K, a *chain* is a totally ordered subset $C \subseteq K$. For $\circ \in \{+, \cdot\}$ we write $a \circ C$ for $\{a \circ c \mid c \in C\}$. Provided they exist, we write $\bigsqcup C$ and $\bigsqcap C$ for the *supremum* (least upper bound) and *infimum* (greatest lower bound) of $C \subseteq K$, and further $\bot$ and $\top$ for the least and greatest elements of K. We say that a function $f : K_1 \to K_2$ is fully chain-continuous or, for short, *fully continuous* if it preserves suprema and infima of nonempty chains, i.e., $f(\bigsqcup C) = \bigsqcup f(C)$ and $f(\bigsqcap C) = \bigsqcap f(C)$ for all chains $\varnothing \neq C \subseteq K_1$.

▶ **Definition 4.** A naturally ordered semiring K is *fully chain-complete* if every chain $C \subseteq K$ has a supremum $\bigsqcup C$ and an infimum $\bigsqcap C$ in K. It is additionally *fully continuous* if its operations are fully continuous in both arguments, i.e., $a \circ \bigsqcup C = \bigsqcup(a \circ C)$ and $a \circ \bigsqcap C = \bigsqcap(a \circ C)$ for all $a \in K$, chains $\varnothing \neq C \subseteq K$ and $\circ \in \{+, \cdot\}$.

Examples of fully continuous semirings include the Viterbi semiring, the semiring $\mathbb{N}^\infty$ of natural numbers extended by infinity, and semirings of formal power series $\mathbb{N}^\infty[\![X]\!]$ and $\mathbb{N}^\infty[\![X, \bar{X}]\!]$. For positive least fixed-point inductions, as in Datalog [16] or posLFP [14], the common approach is to use ω-*continuous* semirings. There, only suprema of ω-chains are required and both operations must preserve suprema. It would be tempting to work with a minimal generalization that imposes similar properties for descending ω-chains, using a dual version of Kleene's Fixed-Point Theorem. However, the following example shows that this approach will not work in general with alternating fixed points.

▶ **Example 5.** Let K be a naturally ordered semiring that has both suprema of ascending ω-chains and infima of descending ω-chains and let $f : K \times K \to K$ be a function that preserves these suprema and infima in each argument. For each $x \in K$, we can consider the function $g_x : K \to K$, $g_x(y) = f(x, y)$ and, further, the function $G : K \to K$, $G(x) = \mathbf{gfp}(g_x)$. Note that G is well-defined due to the preservation property of f and a dual version of Kleene's Fixed-Point Theorem. Now consider $\mathbf{lfp}(G)$. To guarantee the existence of this fixed point via Kleene's theorem, G has to preserve suprema of ω-chains. This is, however, not the case, in general. One counterexample is the the function $f(x, y) = x \diamond y$ in the (fully continuous) Łukasiewicz semiring $\mathbb{L} = ([0, 1], \max, \diamond, 0, 1)$ with $a \diamond b = \max(0, a + b - 1)$ on the ω-chain $(x_n)_{n<\omega}$ defined by $x_n = 1 - \frac{1}{1+n}$. Then $G(\bigsqcup_{n<\omega} x_n) = G(1) = \mathbf{gfp}(g_1) = 1$, whereas $\bigsqcup_{n<\omega} G(x_n) = \bigsqcup_{n<\omega} \mathbf{gfp}\left(g_{x_n}\right) = \bigsqcup_{n<\omega} 0 = 0$. Hence Kleene's theorem is not applicable (although the least fixed point exists with $\mathbf{lfp}(G) = 0$). ⌟

Instead, we rely on K being fully chain-complete to guarantee the existence of fixed points of monotone functions. We can then extend [20] the Kleene iteration $\bot$, $f(\bot)$, $f^2(\bot)$, $f^3(\bot)$, ... for $\mathbf{lfp}(f)$ to a transfinite fixed-point iteration $(x_\beta)_{\beta \in \mathrm{On}}$ by setting $x_0 = \bot$, $x_{\beta+1} = f(x_\beta)$ for ordinals β and $x_\lambda = \bigsqcup\{x_\beta \mid \beta < \lambda\}$ for limit ordinals λ. If f is monotone, this iteration forms a chain and is well-defined due to the chain-completeness of K. The iteration for $\mathbf{gfp}(f)$ can be defined analogously by $x_\lambda = \bigsqcap\{x_\beta \mid \beta < \lambda\}$ for limit ordinals and it follows that both $\mathbf{lfp}(f)$ and $\mathbf{gfp}(f)$ exist for any monotone function $f : K \to K$ on a fully chain-complete semiring K. Coming back to the question of appropriate semirings for LFP, we observe that the update operators F^φ_π are always monotone (this can be seen by structural induction on φ, using the monotonicity of $\cdot$ and $+$).

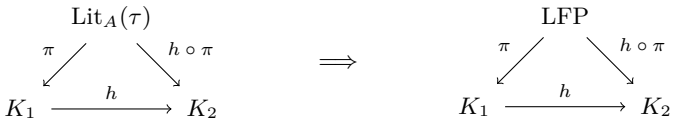
▶ **Theorem 6.** *Semiring semantics for* LFP *is well-defined for fully chain-complete semirings.*

We further remark that full chain-completeness is more general than the common notion of complete lattices, used in the Knaster-Tarski fixed-point theory, as we only require suprema (and infima) of chains instead of arbitrary sets. However, based on results in [19] it follows

that every idempotent, fully chain-complete semiring is in fact a complete lattice (under its natural order); this applies in particular to the absorptive, fully continuous semirings discussed below.

The following fundamental property for provenance analysis (cf. [13]) establishes a closer connection between logic (the semantics of φ) and algebra (the semiring homomorphism h) and enables us to compute provenance information in a general semiring and then specialize the result to application semirings by applying homomorphisms.

▶ **Proposition 7** (Fundamental Property). *Let K_1, K_2 be fully chain-complete semirings and let $h : K_1 \to K_2$ be a fully continuous semiring homomorphism with $h(\top) = \top$. Then for every K_1-interpretation π, the mapping $h \circ \pi$ is a K_2-interpretation and for every $\varphi \in \mathrm{LFP}$, we have $h(\pi[\![\varphi]\!]) = (h \circ \pi)[\![\varphi]\!]$. In diagrammatic form:*



Fully continuous semirings. While fully chain-complete semirings suffice to guarantee well-defined semantics, our main results (the universal property in Theorem 17 and the connection to games in Sect. 6) require the technically slightly stronger notion of fully continuous semirings, in which addition and multiplication preserve suprema and infima of chains. This is an adaption of the standard notion of ω-continuity to our setting and all natural examples of fully chain-complete semirings we are aware of are in fact fully continuous. On a different note, the notion of chain-completeness is based on chains of arbitrary length. We do not know whether working with ascending and descending ω-chains would suffice in all cases, but we show in Sect. 5 that it suffices in absorptive, fully continuous semirings.

Absorptive and chain-positive semirings. Although the *existence* of fixed points is guaranteed in fully continuous semirings, we observe (in Example 10 below) that one may have valuations of greatest fixed-point formulae in such semirings that are not really informative and do not provide useful insights why a formula holds. This can be tied to two separate problems: the *lack of symmetry* between least and greatest fixed-point inductions in some such semirings, and the fact that such semirings are not necessarily *truth-preserving*, i.e. they may evaluate true statements to 0. To deal with these problems we propose to work with fully continuous semirings that are *absorptive*, to provide useful provenance information for greatest fixed points, and *chain-positive*, to guarantee truth-preservation.

We first address the issue of symmetry between least and greatest fixed points. In the Boolean setting, these are computed in the complete lattice of subsets which is inherently symmetric. For instance, a greatest fixed point of a monotone operator is the complement of the least fixed point of the dual operator (which is essential for a negation normal form). Moreover, conjunction and disjunction are symmetric in the sense that one increases values, acting as set union in the lattice of subsets, while the other is decreasing. In the semiring setting, we compute fixed points with respect to the natural order induced by addition. This order is always a complete lattice in absorptive semirings (in fact, idempotent semirings suffice) and it is clear that addition is increasing in the sense that $a + b \geq a$ for all a, b. The issue is with multiplication: The only constraint relating addition and multiplication is distributivity, but this alone does not ensure a symmetry similar to the Boolean setting. We achieve a sufficient degree of symmetry by requiring that the semiring is absorptive.

▶ **Definition 8.** A semiring K is *absorptive* if $a + ab = a$ for all $a, b \in K$, which is equivalent to saying that $1 + b = 1$, for all $b \in K$.

Clearly, every absorptive semiring is *idempotent*: $a + a = a$ for all a. For naturally ordered semirings, absorption indeed provides symmetry: multiplication becomes decreasing and 1 becomes the greatest element, symmetric to addition and the least element 0. We remark that while this adds a certain, fruitful degree of symmetry, it does not enforce complete symmetry of addition and multiplication (as in the Boolean setting). For instance, multiplication need not be idempotent. In particular, absorptive semirings need not be lattices (with $+$ and $\cdot$ as lattice operations), even if the natural order is always a complete lattice.

▶ **Proposition 9.** *In a naturally ordered semiring K, the following are equivalent:*

1. *K is absorptive,*
2. *K has the greatest element $\top = 1$, i.e., $a \leq 1$ for all $a \in K$,*
3. *multiplication in K is decreasing, i.e., $a \cdot b \leq b$ for all $a, b \in K$.*

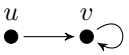
This symmetry helps, for instance, to avoid problems of increasing multiplication as in $\mathbb{N}^\infty$. Fixed-point theory often relies on symmetry and it is thus no surprise that more symmetry leads to more useful provenance information. This can be seen in Example 10 below when comparing the computations of greatest fixed points in the non-absorptive semiring $\mathbb{N}^\infty$ and the more informative Viterbi semiring.

A further motivation for absorptive semirings is that they give information about reduced proofs of a formula. The property $a + ab = a$ implies, for example, that a proof containing two literals mapped to a and b, thus having the value ab, is absorbed by a proof only using one literal, with provenance value a. This has the further benefit that, unlike formal power series $\mathbb{N}^\infty[\![X]\!]$, provenance information is always finitely representable (see Sect. 5).

We next address the issue of truth-preservation, which is defined as follows. As in [13], we say that a K-interpretation $\pi\colon \mathrm{Lit}_A(\tau) \to K$ is *model-defining* if for all atoms $R\mathbf{a}$ exactly one of the two values $\pi[\![R\mathbf{a}]\!]$ and $\pi[\![\neg R\mathbf{a}]\!]$ is zero. A model-defining K-interpretation induces a unique structure $\mathfrak{A}_\pi$ with universe A and $\mathbf{a} \in R^{\mathfrak{A}}$ if, and only if, $\pi(R\mathbf{a}) \neq 0$. For a truthful provenance analysis for a logic L, this should lift to arbitrary sentences $\varphi \in L$, so that $\mathfrak{A}_\pi \models \varphi$ if, and only if, $\pi[\![\varphi]\!] \neq 0$. If this is indeed the case for all model-defining K-interpretations π, we say that K is *truth-preserving* for L. This is illustrated in the following example.

▶ **Example 10.** The existence of an infinite path from u in a graph G is expressed by the LFP-formula

$$\varphi(u) = [\mathbf{gfp}\, Rx\,.\, \exists y(Exy \wedge Ry)](u)$$



For the Boolean semiring $\mathbb{B} = \{0, 1\}$ there is a unique $\mathbb{B}$-interpretation π that defines the displayed graph G. Provenance semantics in $\mathbb{B}$ coincides with standard semantics and we indeed obtain $\pi[\![\varphi(u)]\!] = 1$. The Viterbi semiring $\mathbb{V}$ instead allows us to assign confidence scores to the edges. If we set $\pi(Euv) = \pi(Evv) = 1$ as in the Boolean interpretation, we again obtain an overall confidence of $\pi[\![\varphi(u)]\!] = 1$. However, if we instead lower the score of the self-loop to $\pi(Evv) = 1 - \varepsilon$, we obtain an overall confidence of $\pi[\![\varphi(u)]\!] = 0$ due to the fixed-point iteration $1, 1 - \varepsilon, (1 - \varepsilon)^2, \ldots$. So while π still defines the model shown above, the formula evaluates to 0 which we usually interpret as *false*, illustrating that the Viterbi semiring is *not truth-preserving*. Since the loop occurs infinitely often in the unique infinite path from u, the value 0 makes sense as a confidence score. Thus, although it is not truth-preserving, the Viterbi semiring does provide useful information.

Consider next the semiring of formal power series $\mathbb{N}^\infty[\![X]\!]$. If we choose $\pi(Euv) = x$ and $\pi(Evv) = y$ (and keep the values 0 or 1 for the remaining literals), then $\pi[\![\varphi(u)]\!] = 0$, as result of the infinite iteration $\top$, $y \cdot \top$, $y^2 \cdot \top$, $y^3 \cdot \top$, ... with infimum 0 at node v (here, $\top$ is the power series in which all monomials have coefficient ∞). Thus, the semiring $\mathbb{N}^\infty[\![X]\!]$ is *not truth-preserving* either.

In the semiring $\mathbb{N}^\infty$, used to count proofs of formulae in FO and posLFP, the consideration of greatest fixed points imposes problems: Intuitively, the graph only has one infinite path that we would view as a proof of $\varphi(u)$. But setting $\pi(Euv) = \pi(Evv) = 1$ results in $\pi[\![\varphi(u)]\!] = \infty$, since the iteration for the evaluation of φ at v is ∞, $1 \cdot \infty$, $1 \cdot \infty$, ... which stagnates immediately. Although $\mathbb{N}^\infty$ is truth-preserving, the example hints at another general issue: Multiplication with non-zero values in $\mathbb{N}^\infty$ always increases values. The same is true for addition, so fixed-point iterations of **gfp**-formula are likely to result in ∞ and *do not give meaningful provenance information*, e.g. about the number of proofs. Since the computation in $\mathbb{N}^\infty[\![X]\!]$ yields 0, we further see that we cannot obtain the result in $\mathbb{N}^\infty$ from the computation in $\mathbb{N}^\infty[\![X]\!]$ by polynomial evaluation. Hence evaluation of formal power series *does not preserve provenance semantics* in general. This is a further reason why formal power series are not the right provenance semirings for LFP. ⌟

We shall define and investigate in the next section the semiring of generalized absorptive polynomials $\mathbb{S}^\infty[X]$ which, contrary to other fully continuous and absorptive semirings, is truth-preserving due to the following algebraic property.

▶ **Definition 11.** A fully chain-complete semiring K is *chain-positive* if for each non-empty chain $C \subseteq K$ of non-zero elements, the infimum $\bigsqcap C$ is non-zero as well.

▶ **Proposition 12.** *Every chain-positive, positive semiring is truth-preserving for* LFP.

Chain-positivity is not an indispensable requirement for provenance analysis, as shown by the Viterbi semiring (which is absorptive and fully continuous). However, we need this property for provenance semirings which should give insights into proofs or evaluation strategies and thus have to preserve truth.

5 Generalized Absorptive Polynomials

We now discuss the semirings $\mathbb{S}^\infty[X]$ and $\mathbb{S}^\infty[X, \bar{X}]$ of generalized absorptive polynomials. They were introduced in [14] and generalize the semiring of absorptive polynomials $\mathsf{Sorp}(X)$ from [6] by admitting exponents in $\mathbb{N}^\infty$ to guarantee chain-positivity. We show that these semirings are, in a well-defined sense, the most general absorptive, fully continuous semirings and we argue that $\mathbb{S}^\infty[X, \bar{X}]$ is the right provenance semiring for LFP.

▶ **Definition 13.** Let X be a *finite* set of provenance tokens. We generalize the notion of a monomial over X to admit exponents from $\mathbb{N}^\infty$. Monomials are here functions $m : X \to \mathbb{N}^\infty$, written $x_1^{m(x_1)} \cdots x_n^{m(x_n)}$. Multiplication adds the exponents, and $x^\infty \cdot x^n = x^\infty$. We say that m_2 *absorbs* m_1, denoted $m_2 \succeq m_1$, if m_2 has smaller exponents than m_1, i.e., $m_2(x) \leq m_1(x)$ for all $x \in X$. This is the pointwise partial order given by the reverse order on $\mathbb{N}^\infty$.

The set of monomials inherits a lattice structure from $\mathbb{N}^\infty$ and is, of course, infinite. However, it has some crucial finiteness properties.

▶ **Proposition 14.** *Every antichain of monomials is finite. Further, while there are infinitely descending chains of monomials, such as* $1 = x^0 \succ x^1 \succ x^2 \succ \ldots$ *there are no infinitely ascending such chains.*

Indeed, $(\mathbb{N}^\infty, \leq)$ is a well-order. The set of monomials $m : X \to \mathbb{N}^\infty$ with the *reverse order* of the absorption order is isomorphic to $(\mathbb{N}^\infty)^k$ with $k = |X|$ and with the componentwise order inherited from $(\mathbb{N}^\infty, \leq)$. This is a well-quasi-order and therefore has no infinite descending chains and no infinite antichains. This implies that in the set of monomials over X with the absorption order, all ascending chains and all antichains are finite.

▶ **Definition 15.** We define $\mathbb{S}^\infty[X]$ as the set of antichains of monomials with indeterminates from X and exponents in $\mathbb{N}^\infty$. We write such antichains as formal sums of their monomials and call them *(generalized absorptive) polynomials*. Addition and multiplication of polynomials proceed as usual, but keeping only the maximal monomials (w.r.t. $\succeq$) in the result (and disregarding coefficients).

Since antichains of monomials are finite, there is no difference between polynomials and power series here. The natural order on $\mathbb{S}^\infty[X]$ can be characterized by monomial absorption: $P \leq Q$ if, and only if, for each $m \in P$ there is $m' \in Q$ with $m' \succeq m$. With Proposition 14, it follows that there are no infinitely ascending chains of polynomials, and further that the supremum of $S \subseteq \mathbb{S}^\infty[X]$ is $\bigsqcup S = \text{maximals}\,(\bigcup S)$ which is the set of $\succeq$-maximal monomials in $\bigcup S$. Due to the exponent ∞ and the finiteness of X, there is a smallest monomial $m_\infty \neq 0$ with $m_\infty(x) = \infty$ for all $x \in X$. This ensures chain-positivity of $\mathbb{S}^\infty[X]$.

▶ **Proposition 16.** $(\mathbb{S}^\infty[X], +, \cdot, 0, 1)$ *is absorptive, fully continuous, and chain-positive.*

The central property of $\mathbb{S}^\infty[X]$ is the following universal property which says that it is the absorptive, fully continuous semiring freely generated by X w.r.t. fully continuous homomorphisms. These homomorphisms enable us to apply the fundamental property.

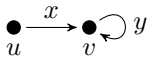
▶ **Theorem 17** (Universality). *Every mapping $h : X \to K$ into an absorptive, fully continuous semiring K uniquely extends to a fully continuous semiring homomorphism $h : \mathbb{S}^\infty[X] \to K$.*

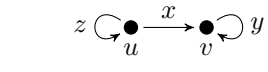
In absorptive semirings, the powers of an element a form a descending ω-chain $1 \geq a \geq a^2 \geq \cdots$ whose infimum we denote by a^∞. Since we want h to be fully continuous, the mapping $h(x) = a$ implies $h(x^\infty) = h(\bigsqcap_n x^n) = \bigsqcap_n h(x)^n = a^\infty$. By similar arguments, it is straightforward to see that h is uniquely defined. The nontrivial part of the proof is that the induced homomorphism h is fully continuous. Ascending chains are finite and thus impose no difficulties, so what remains is to prove that $h(\bigsqcap C) = \bigsqcap h(C)$ for chains $C \neq \varnothing$. Our proof constructs from $\bigsqcap C$ a canonical chain and makes use of Kőnig's lemma (recall that polynomials are finite) to relate the original chain C to the canonical chain, in a way that is preserved by h (see Appendix A for the full proof).

The fact that the universal property guarantees fully continuous homomorphisms should not be taken lightly: We have seen in Example 10 that this is not the case for formal power series. There, polynomial evaluation induces homomorphisms that are, in general, not fully continuous and hence do not preserve greatest fixed points. The following example shows how we can specialize provenance values in $\mathbb{S}^\infty[X]$ to application semirings.

▶ **Example 18.** We recall the setting from Example 10 and first consider the model-defining $\mathbb{S}^\infty[X]$-interpretation tracking the two edges labelled x and y, as indicated in the left graph.

$$\varphi(u) = [\mathbf{gfp}\, Rx\,.\, \exists y(Exy \wedge Ry)](u)$$





We obtain $\pi[\![\varphi(u)]\!] = xy^\infty$ corresponding to the infinite path $uvvv\ldots$. The confidence values from Example 10 can be obtained by polynomial evaluation: For $h(x) = h(y) = 1$, we get $(h \circ \pi)[\![\varphi(u)]\!] = 1 \cdot 1^\infty = 1$ and for $h'(x) = 1$, $h'(y) = 1 - \varepsilon$ we get $(h' \circ \pi)[\![\varphi(u)]\!] = 1 \cdot (1-\varepsilon)^\infty = 0$.

We next consider the graph on the right by setting $\pi(Euu) = z$. There are now infinitely many infinite paths from u to v. However, we obtain only finitely many monomials due to absorption: $\pi[\![\varphi(u)]\!] = xy^\infty + z^\infty$. These correspond to the *simplest* infinite paths since monomials such as z^2xy^∞ (corresponding to the path $uuuvvv\ldots$) are absorbed by xy^∞. ⌟

One consequence of the universal property is the existence of a *most general* $\mathbb{S}^\infty[X]$-*interpretation* π_0 by introducing variables $X = \{x_L \mid L \in \mathrm{Atoms}_A(\tau) \cup \mathrm{NegAtoms}_A(\tau)\}$ for all literals and setting $\pi_0(L) = x_L$. Any other K-interpretation π (where K is fully continuous and absorptive) results from π_0 by the evaluation $x_L \mapsto \pi(L)$ which lifts to a fully continuous homomorphism h. After computing $\pi_0[\![\varphi]\!]$ once, the computation for any K-interpretation π is then simply a matter of applying polynomial evaluation, since $\pi[\![\varphi]\!] = h(\pi_0[\![\varphi]\!])$.

The most general $\mathbb{S}^\infty[X]$-interpretation can also be used to prove that the update operators F^φ_π induced by LFP-formulae in $\mathbb{S}^\infty[X]$ are fully continuous. Hence Kleene's Fixed-Point Theorem applies and guarantees that the fixed-point iterations for $\mathbf{lfp}(F^\varphi_\pi)$ and $\mathbf{gfp}(F^\varphi_\pi)$ have closure ordinal at most ω. Using the universal property, the statement on the closure ordinal generalizes to all absorptive, fully continuous semirings – even to semirings in which update operators are not continuous in general (such as the semiring $\mathbb{L}$ in Example 5).

▶ **Proposition 19.** *Given a K-interpretation π into an absorptive, fully continuous semiring, all fixed-point iterations for $\mathbf{lfp}(F^\varphi_\pi)$ and $\mathbf{gfp}(F^\varphi_\pi)$ have closure ordinal at most ω.*

What we still have to provide for an adequate provenance analysis is a proper treatment of negation: If we track a literal and its negation by different variables x and y, respectively, we may obtain inconsistent monomials such as xy. As in other semirings of polynomials and power series we can also here take pairs of positive and negative indeterminates, with a correspondence $X \leftrightarrow \bar{X}$, and build the quotient with respect to the congruence generated by the equation $x \cdot \bar{x} = 0$. We thus obtain a new semiring $\mathbb{S}^\infty[X, \bar{X}]$ which, as a quotient, retains the properties of being absorptive, fully continuous and chain-positive. Of course, $\mathbb{S}^\infty[X, \bar{X}]$ is no longer positive, as x and $\bar{x}$ are divisors of 0. Most importantly, $\mathbb{S}^\infty[X, \bar{X}]$ inherits the universal property: If $h : X \cup \bar{X} \to K$ respects dual-indeterminates, so $h(x) \cdot h(\bar{x}) = 0$ for all $x \in X$, then it extends uniquely to a fully continuous homomorphism $h : \mathbb{S}^\infty[X, \bar{X}] \to K$. Together with the fundamental property, $\mathbb{S}^\infty[X, \bar{X}]$ is thus the most general appropriate provenance semiring for LFP that can represent negation, hence providing a natural framework for a provenance analysis for LFP and other fixed-point calculi.

Instead of model-defining interpretations, we consider *model-compatible* interpretations π. That is, for each atom $R\mathbf{a}$ we either have $\pi(R\mathbf{a}) = x$ and $\pi(\neg R\mathbf{a}) = \bar{x}$, or $\{\pi(R\mathbf{a}), \pi(\neg R\mathbf{a})\} = \{0, 1\}$. Additionally, π must not use the same indeterminate for two different atoms. We say that a model $\mathfrak{A}$ is *compatible* with π if $\mathfrak{A} \models L$ for all literals L with $\pi(L) = 1$ and denote the set of compatible models by Mod_π. Model-compatible interpretations can be used to reason about several models at once. Mapping certain literals to indeterminate pairs x and $\bar{x}$ leaves open the truth of these literals, but still encodes the semantics of opposing literals:

▶ **Proposition 20.** *Let π be a model-compatible $\mathbb{S}^\infty[X, \bar{X}]$-interpretation. An LFP-formula φ is Mod_π-satisfiable (Mod_π-valid) if, and only if, $\pi[\![\varphi]\!] \neq 0$ ($\pi[\![\neg\varphi]\!] = 0$).*

6 Game-theoretic analysis

It has been shown in [14] that the provenance analysis for FO and posLFP is intimately connected with the provenance analysis of reachability games. Evaluation strategies to establish the truth of first-order formulae are really winning strategies for reachability games

on acyclic game graphs. For posLFP the situation is similar, but the associated model-checking games may have cycles and thus admit infinite plays, but the winning plays for the verifying player have to reach a winning position (a true literal) in a finite number of steps. By annotating such terminal positions with semiring values and propagating these values along the edges to the remaining positions, one obtains provenance values that coincide with the syntactically defined semantics $\pi[\![\psi]\!]$.

For full LFP or the modal μ-calculus, the model-checking games are parity games which are considerably more complex and do not allow for a simple propagation of values from terminal positions. We do not present here a general provenance analysis of parity games, but we show how provenance values $\pi[\![\varphi]\!]$ for fixed-point formulae can be understood from a game-theoretic point of view. For first-order logic or posLFP, provenance values $\pi[\![\varphi]\!]$ in $\mathbb{N}[X, \bar{X}]$ or $\mathbb{N}^\infty[\![X, \bar{X}]\!]$ are sums of monomials that correspond to the evaluation strategies for φ and provide information about the literals used by these strategies. We present an analogue of this statement for full fixed-point logic and the semiring $\mathbb{S}^\infty[X, \bar{X}]$.

Model-checking games for LFP. Model-checking games are classically defined for a formula and a fixed structure $\mathfrak{A}$ (see e.g. [2, Chap. 4]). However, the *game graph* of such a game depends only on the formula ψ and the *universe* of the given structure, and it is only the labelling of the terminal positions as winning for either the Verifier (Player 0) or the Falsifier (Player 1), that depends on which of the literals in $\mathrm{Lit}_A(\tau)$ are true in $\mathfrak{A}$. Hence the definition readily generalizes to a more abstract provenance scenario where we instead label terminal positions by semiring values. As the definition of the model-checking game $\mathcal{G}(A, \psi)$ itself is standard, we refer to the full version [5] or [2] for details. Most importantly, positions in the game $\mathcal{G}(A, \psi)$ correspond to subformulae of ψ and terminal positions are literals in $\mathrm{Lit}_A(\tau)$. The game may have cycles that admit infinite plays which are won according to the *parity condition*: We assign to each fixed-point variable a priority, and an infinite play is then won by Verifier precisely if the least priority occurring infinitely often is even.

Provenance values for plays and strategies. Given a parity game $\mathcal{G}(A, \psi)$, every K-interpretation $\pi : \mathrm{Lit}_A(\tau) \to K$ provides a valuation of the terminal positions. Based on this, we define provenance values for plays and strategies.

▶ **Definition 21.** A *finite* play $\rho = (\varphi_0, \ldots, \varphi_t)$ ends in a terminal position $\varphi_t \in \mathrm{Lit}_A(\tau)$ which we call the *outcome* of ρ. We simply identify the provenance value of ρ with the value of its outcome, i.e. we put $\pi[\![\rho]\!] \coloneqq \pi[\![\varphi_t]\!]$. For an *infinite* play ρ we put $\pi[\![\rho]\!] \coloneqq 1$ if ρ is a wining play for the Verifier, and $\pi[\![\rho]\!] \coloneqq 0$ otherwise.

We denote by $\mathrm{Strat}(\varphi)$ the set of evaluation strategies for the subformula φ of ψ, i.e. the set of all (not necessarily positional) strategies that the Verifier has from position φ in the parity game $\mathcal{G}(A, \psi)$. Every strategy $\mathcal{S} \in \mathrm{Strat}(\varphi)$ induces the set $\mathrm{Plays}(\mathcal{S})$ of plays that are consistent with $\mathcal{S}$. Intuitively, the provenance value of a strategy is simply the product over the provenance values of all plays that it admits. However, a strategy may well admit an infinite set of plays and while it is possible to define infinite products in our setting, we instead observe that the set of possible outcomes is of course finite, since there exist only finitely many literals. As a consequence, we define the provenance value for a strategy by grouping those plays with identical outcome.

▶ **Definition 22.** The *provenance value* of a strategy $\mathcal{S}$ is $\pi[\![\mathcal{S}]\!] \coloneqq \prod_{L \in \mathrm{Lit}_A(\tau)} \pi(L)^{\#_{\mathcal{S}}(L)}$ if all infinite $\rho \in \mathrm{Plays}(\mathcal{S})$ are winning for Verifier, and $\pi[\![\mathcal{S}]\!] \coloneqq 0$ otherwise. Here, $\#_{\mathcal{S}}(L) \in \mathbb{N} \cup \{\infty\}$ denotes the number of plays $\rho \in \mathrm{Plays}(\mathcal{S})$ with outcome L.

The case for $\#_{\mathcal{S}}(L) = \infty$ is well-defined, as the infinitary power $a^\infty = \bigsqcap_n a^n$ can be defined in all absorptive, fully continuous semirings. For model-compatible interpretations in $\mathbb{S}^\infty[X, \bar{X}]$, the value $\pi[\![\mathcal{S}]\!]$ is a single monomial. The following central result justifies our game-theoretic analysis and precisely characterizes provenance semantics $\pi[\![\psi]\!]$ in terms of strategies in the associated model-checking game.

▶ **Theorem 23.** *Let $\psi \in \mathrm{LFP}$, and and let $\pi\colon \mathrm{Lit}_A(\tau) \to K$ be a K-interpretation into an absorptive, fully continuous semiring K. Then $\pi[\![\psi]\!] = \bigsqcup\{\pi[\![\mathcal{S}]\!] \mid \mathcal{S} \in \mathrm{Strat}(\psi)\}$.*

Model-checking games become large even for simple formulae over a small universe A; we thus refer to Appendix B and [5, Sect. 6.1] for examples. The proof of this result [5, Sect. 6.2] is not short either. The key idea is to view strategies $\mathcal{S}$ in the game of, say, $[\mathbf{gfp}\, R\mathbf{x}\,.\,\varphi](\mathbf{a})$ as trees and then define prefixes $\mathcal{S}|_n$ of these trees based on the number of fixed-point literals $R\mathbf{b}$ along a path. We prove by induction that these prefixes of increasing size correspond exactly to the steps of the fixed-point iteration via F^φ_π. For greatest fixed points, strategies can be infinite which leads to subtle obstacles. Perhaps the most challenging step is the so-called *puzzle lemma* which shows that, roughly speaking, computing infima of strategy prefixes leads to meaningful values corresponding to actual (infinite) strategies.

Consider now specifically the semiring $\mathbb{S}^\infty[X, \bar{X}]$ and model-compatible interpretations. By the above theorem, the provenance value of a sentence ψ is then a sum of monomials $x_1^{e_1} \cdots x_k^{e_k}$, each of which corresponds to a strategy $\mathcal{S}$ for Verifier that uses precisely the literals labelled by $x_1, \ldots, x_k$, and each literal x_i is used precisely e_i many times, that is, there are e_i plays consistent with $\mathcal{S}$ that have outcome x_i. By using dual indeterminates, we make sure that these literals are consistent and hence represent actual evaluation strategies.

In this sense, provenance semantics in absorptive, fully continuous semirings, and most prominently in $\mathbb{S}^\infty[X, \bar{X}]$, provide detailed information about evaluation strategies. Because of absorption, we do not obtain information about all evaluation strategies, as in first-order logic and $\mathbb{N}[X, \bar{X}]$, but instead only about the absorption-dominant strategies, corresponding to absorption-maximal monomials. These are strategies that allow the fewest different possible outcomes and are thus the simplest or canonical evaluation strategies.

7 Related Work

While our approach and our general project, as outlined in the introduction, is rooted in the work on semiring provenance in databases, there have also been a number of other areas of logic in computer science where semiring semantics have been used.

A prominent instance is the work on weighted automata (see, e.g., the Handbook [8]). In particular, weighted automata over finite and infinite words are discussed in [7, 9], and their expressive power is related to weighted monadic second-order logic (MSO) on words. In this setting, the weight of a word is defined as the sum over the weights of accepting paths, and the overall behaviour of an automaton is described by a formal power series over a semiring. To deal with infinite words, infinite sum and product operations of the semiring are assumed [9], roughly comparable to our assumptions on suprema and infima. Whereas the power series assign semiring values to *words*, we instead use indeterminates to track (combinations of) *literals* which then provide us with provenance information. As we have seen, formal power series are not the right tool for this purpose when confronted with greatest fixed points, so we consider absorptive polynomials $\mathbb{S}^\infty[X]$ instead. Moreover, in our setting the sum-of-strategies characterization is not a definition, but a non-trivial result. The definition of weighted MSO is similar to our semiring semantics for LFP and is also based

on negation normal form. Main differences are that only logics over words are considered, and that semiring values are part of the formulae, whereas we assign values to literals. This reflects the different point of view: Weighted MSO is used to define series recognizable by weighted automata, whereas our goal is the provenance analysis of the logic itself.

Lluch-Lafuente and Montanari [18] have studied a semantics of CTL and μ-calculus in so-called constraint semirings, to reason about issues of quality of service such as delay or bandwidth. The choice of constraint semirings is motivated by applications for a particular class of constraint satisfaction problems, called soft CSP, and by useful closure properties, such as closure under Cartesian products, exponentials, and power constructions. Although this is not mentioned explicitly, constraint semirings are in fact also absorptive and satisfy a continuity requirement for suprema. However, the approach to negation is different from ours, requiring the extension of the semiring by new functions, and they do not have an abstract approach on the basis of polynomials with universal properties and reasoning over multiple constraint semirings. A main result of [18] is that the usual embedding of CTL into the μ-calculus fails for this semantics, which is another instance showing that a refined semiring semantics may distinguish between formulae that are equivalent under Boolean semantics.

8 Conclusion and Outlook

Let us summarize the contributions of this paper: We have laid foundations for the semiring provenance analysis of full fixed-point logics, with arbitrary interleavings of least and greatest fixed points, as part of the general project of developing provenance semantics of logical languages used in various branches of computer science. We have seen that absorptive and fully continuous semirings provide an adequate framework for this. We have identified the semiring of dual-indeterminate generalized absorptive polynomials $\mathbb{S}^\infty[X, \bar{X}]$ as the "right" provenance semiring for LFP. It satisfies the further algebraic property of chain-positivity which guarantees that provenance interpretations are truth-preserving, and we have established an important universal property of this semiring. Finally, we have shown how provenance for LFP is related to strategies in model-checking games.

Next steps will include the specific analysis of important logics such as temporal logics, dynamic logics, the modal μ-calculus, description logics (see initial work in [3]) etc. Applications require in particular the study of *algorithms* for computing provenance values – a non-trivial task, considering that greatest fixed-point iterations in semirings such as $\mathbb{S}^\infty[X, \bar{X}]$ can be infinite. Nevertheless, absorption and the infinitary power a^∞ can be used to short-circuit these iterations; forthcoming work will include results that show how an effective, and in important cases also efficient, computation of provenance values is possible in absorptive, fully continuous semirings.

References

1 Y. Amsterdamer, D. Deutch, and V. Tannen. On the limitations of provenance for queries with difference. In *3rd Workshop on the Theory and Practice of Provenance, TaPP'11*, 2011. See also CoRR abs/1105.2255.

2 K. Apt and E. Grädel, editors. *Lectures in Game Theory for Computer Scientists*. Cambridge University Press, 2011. doi:10.1017/CBO9780511973468.

3 K. Dannert and E. Grädel. Provenance analysis: A perspective for description logics? In C. Lutz et al., editor, *Description Logic, Theory Combination, and All That*, Lecture Notes in Computer Science Nr. 11560. Springer, 2019. doi:10.1007/978-3-030-22102-7_12.

4 K. Dannert and E. Grädel. Semiring provenance for guarded logics. In *Hajnal Andréka and István Németi on Unity of Science: From Computing to Relativity Theory through Algebraic Logic*, Outstanding Contributions to Logic. Springer, 2021.

5 K. Dannert, E. Grädel, M. Naaf, and V. Tannen. Generalized absorptive polynomials and provenance semantics for fixed-point logic. arXiv: 1910.07910 [cs.LO], 2019. URL: `https://arxiv.org/abs/1910.07910`.

6 D. Deutch, T. Milo, S. Roy, and V. Tannen. Circuits for datalog provenance. In *Proc. 17th International Conference on Database Theory ICDT*, pages 201–212. OpenProceedings.org, 2014. `doi:10.5441/002/icdt.2014.22`.

7 M. Droste and P. Gastin. Weighted automata and weighted logics. In *Handbook of weighted automata*, pages 175–211. Springer, 2009. `doi:10.1007/978-3-642-01492-5_5`.

8 M. Droste, W. Kuich, and H. Vogler, editors. *Handbook of Weighted Automata*. Springer, 2009. `doi:10.1007/978-3-642-01492-5`.

9 M. Droste and G. Rahonis. Weighted automata and weighted logics on infinite words. In *International Conference on Developments in Language Theory*, pages 49–58. Springer, 2006. `doi:10.1007/11779148_6`.

10 F. Geerts and A. Poggi. On database query languages for K-relations. *J. Applied Logic*, 8(2):173–185, 2010. `doi:10.1016/j.jal.2009.09.001`.

11 F. Geerts, T. Unger, G. Karvounarakis, I. Fundulaki, and V. Christophides. Algebraic structures for capturing the provenance of SPARQL queries. *J. ACM*, 63(1):7:1–7:63, 2016. `doi:10.1145/2810037`.

12 E. Grädel, P. G. Kolaitis, L. Libkin, M. Marx, J. Spencer, M. Y. Vardi, Y. Venema, and S. Weinstein. *Finite Model Theory and Its Applications*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2007. `doi:10.1007/3-540-68804-8`.

13 E. Grädel and V. Tannen. Semiring provenance for first-order model checking. arXiv:1712.01980 [cs.LO], 2017. URL: `https://arxiv.org/abs/1712.01980`.

14 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`.

15 T. Green, Z. Ives, and V. Tannen. Reconcilable differences. In *Database Theory - ICDT 2009*, pages 212–224, 2009. `doi:10.1145/1514894.1514920`.

16 T. Green, G. Karvounarakis, and V. Tannen. Provenance semirings. In *Principles of Database Systems PODS*, pages 31–40, 2007. `doi:10.1145/1265530.1265535`.

17 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`.

18 A. Lluch-Lafuente and U. Montanari. Quantitative mu-calculus and CTL defined over constraint semirings. *Theoretical Compututer Science*, 346(1):135–160, 2005. `doi:10.1016/j.tcs.2005.08.006`.

19 G. Markowsky. Chain-complete posets and directed sets with applications. *Algebra universalis*, 6(1):53–68, 1976.

20 Y. Moschovakis. *Elementary induction on abstract structures*. North Holland, 1974.

21 Y. Ramusat, S. Maniu, and P. Senellart. Semiring provenance over graph databases. In *10th USENIX Workshop on the Theory and Practice of Provenance (TaPP 2018)*, London, 2018.

22 P. Senellart. Provenance and probabilities in relational databases: From theory to practice. *SIGMOD Record*, 46(4):5–15, 2017. `doi:10.1145/3186549.3186551`.

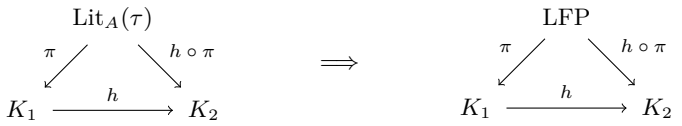
23 J. Xu, W. Zhang, A. Alawini, and V. Tannen. Provenance analysis for missing answers and integrity repairs. *IEEE Data Eng. Bull.*, 41(1):39–50, 2018.

A Proofs

This appendix contains proofs of two key results, the Fundamental Property (Proposition 7) and the Universality (Theorem 17) of the semiring $\mathbb{S}^{\infty}[X]$. Proofs of all remaining results, in particular for Sect. 6, are available in the full version of this paper [5].

A.1 Fundamental Property

▶ **Proposition 7** (Fundamental Property). *Let K_1, K_2 be fully chain-complete semirings and let $h : K_1 \to K_2$ be a fully continuous semiring homomorphism with $h(\top) = \top$. Then for every K_1-interpretation π, the mapping $h \circ \pi$ is a K_2-interpretation and for every $\varphi \in \mathrm{LFP}$, we have $h(\pi[\![\varphi]\!]) = (h \circ \pi)[\![\varphi]\!]$. In diagrammatic form:*



Proof. The proof is a mostly straightforward induction on the structure of φ. For fixed-point formulae, we proceed by transfinite induction on the fixed-point iterations and rely on full continuity of the homomorphism h for limit ordinals. Formally, we prove that for all LFP-formulae $\varphi(\mathbf{x})$ in negation normal form, $h(\pi[\![\varphi(\mathbf{a})]\!]) = (h \circ \varphi)[\![\varphi(\mathbf{a})]\!]$ holds for all K-interpretations π and all tuples $\mathbf{a}$ from the universe A.

- For literals, we trivially have $h(\pi[\![R\mathbf{a}]\!]) = h(\pi(R\mathbf{a})) = (h \circ \pi)(R\mathbf{a}) = (h \circ \pi)[\![R\mathbf{a}]\!]$.
- For $\varphi = \varphi_1 \wedge \varphi_2$ we use that h is a semiring homomorphism:

$$h(\pi[\![\varphi]\!]) = h(\pi[\![\varphi_1]\!] \cdot \pi[\![\varphi_2]\!]) = h(\pi[\![\varphi_1]\!]) \cdot h(\pi[\![\varphi_2]\!]) = (h \circ \pi)[\![\varphi_1]\!] \cdot (h \circ \pi)[\![\varphi_2]\!] = (h \circ \pi)[\![\varphi]\!].$$

  The proof for $\vee$, $\exists$ and $\forall$ is analogous (recall that we assume a finite universe, so quantifiers translate to finite sums or products).
- For $\varphi = [\mathbf{gfp}\, R\mathbf{x} \,.\, \vartheta](\mathbf{y})$ with R of arity k, we consider the fixed-point iteration $(g_\beta)_{\beta \in \mathrm{On}}$ for π in K_1 and the iteration $(f_\beta)_{\beta \in \mathrm{On}}$ for $h \circ \pi$ in K_2. We show by induction that $h \circ g_\beta = f_\beta$ for all ordinals $\beta \in \mathrm{On}$, so h preserves all steps of the fixed-point iteration.

  - Initially, $g_0 \colon A^k \to K_1$, $\mathbf{a} \mapsto \top$ and $f_0 \colon A^k \to K_2$, $\mathbf{a} \mapsto \top$. So $h \circ g_0 = f_0$ by $h(\top) = \top$.
  - For successor ordinals, we can apply the induction hypothesis. By definition,

$$g_{\beta+1}(\mathbf{a}) = F^{\vartheta}_{\pi}(g_\beta)(\mathbf{a}) = \pi[R \mapsto g_\beta][\![\vartheta(\mathbf{a})]\!],$$
$$f_{\beta+1}(\mathbf{a}) = F^{\vartheta}_{h\circ\pi}(f_\beta)(\mathbf{a}) = (h \circ \pi)[R \mapsto f_\beta][\![\vartheta(\mathbf{a})]\!] \overset{(*)}{=} (h \circ \pi[R \mapsto g_\beta])[\![\vartheta(\mathbf{a})]\!].$$

    In $(*)$, we use the induction hypothesis $h \circ g_\beta = f_\beta$. Using the (outer) induction hypothesis on ϑ in $(\dagger)$, we obtain

$$(h \circ g_{\beta+1})(\mathbf{a}) = h(\pi[R \mapsto g_\beta][\![\vartheta(\mathbf{a})]\!]) \overset{(\dagger)}{=} (h \circ \pi[R \mapsto g_\beta])[\![\vartheta(\mathbf{a})]\!] = f_{\beta+1}(\mathbf{a}).$$

  - For limit ordinals, we exploit that h is fully continuous:

$$h(g_\lambda(\mathbf{a})) = h(\bigsqcap\{g_\beta(\mathbf{a}) \mid \beta < \lambda\})$$
$$= \bigsqcap\{h(g_\beta(\mathbf{a})) \mid \beta < \lambda\} = \bigsqcap\{f_\beta(\mathbf{a}) \mid \beta < \lambda\} = f_\lambda(\mathbf{a}).$$

  This closes the proof for **gfp**-formulae, as for sufficiently large β, we have

$$h(\pi[\![\varphi(\mathbf{a})]\!]) = h(g_\beta(\mathbf{a})) = f_\beta(\mathbf{a}) = (h \circ \pi)[\![\varphi(\mathbf{a})]\!].$$

  The proof for **lfp**-formulae is analogous. ◀

A.2 Universal Property of Absorptive Polynomials

▶ **Theorem 17** (Universality). *Every mapping $h : X \to K$ into an absorptive, fully continuous semiring K uniquely extends to a fully continuous semiring homomorphism $h : \mathbb{S}^\infty[X] \to K$.*

Towards the proof, we need an auxiliary lemma on descending ω-chains, that is, sequences of the form $(a_i)_{i<\omega}$ with $a_0 \geq a_1 \geq \dots$ of elements from an absorptive, fully continuous semiring. For instance, the powers of an element a form a descending ω-chain, since $a \geq a^2 \geq a^3 \geq \dots$. We refer to the infimum of this chain as infinitary power $a^\infty = \prod_n a^n$.

▶ **Lemma A1** (Splitting Lemma). *Let K be a fully continuous semiring and let $(a_i)_{i<\omega}$ and $(b_i)_{i<\omega}$ be two descending ω-chains. Then, $\prod_{i<\omega}(a_i \circ b_i) = \left(\prod_{i<\omega} a_i\right) \circ \left(\prod_{j<\omega} b_j\right)$, with $\circ \in \{+, \cdot\}$. Analogous statements hold for suprema.*

Proof. We only show the statement for infima, the proof for suprema is analogous. We have the following equality, where $(*)$ holds since K is fully continuous:

$$\prod_{i<\omega} a_i \circ b_i \overset{(1)}{=} \prod_{i<\omega}\prod_{j<\omega} a_i \circ b_j \overset{(*)}{=} \prod_{i<\omega}(a_i \circ \prod_{j<\omega} b_j) \overset{(*)}{=} \prod_{i<\omega} a_i \ \circ \prod_{j<\omega} b_j$$

We prove both directions of (1). Fix i, j and let $k = \max(i, j)$. Then $a_i \circ b_j \geq a_k \circ b_k \geq \prod_k a_k \circ b_k$ by monotonicity of $\circ$. As i, j are arbitrary, this proves $\prod_i \prod_j a_i \circ b_j \geq \prod_k a_k \circ b_k$.

For the other direction, we have $a_i \circ b_i \geq a_i \circ \prod_j b_j$ for every i by monotonicity of $\circ$. By continuity, $a_i \circ b_i \geq \prod_j a_i \circ b_j$ for every i, and thus $\prod_i a_i \circ b_i \geq \prod_i \prod_j a_i \circ b_j$. ◀

Recall that full continuity of the induced homomorphism h means that it preserves suprema and infima of all chains, i.e., of all totally ordered sets. By observing that $\mathbb{S}^\infty[X]$ is countable (given that X is finite), it in fact suffices to show that suprema and infima of ω-chains are preserved:

▶ **Lemma A2** (Countable Chains). *Let K, K' be fully chain-complete semirings and $C \subseteq K$ a countable chain. Then there is a descending ω-chain $(x_i)_{i<\omega}$ such that $\prod C = \prod_i x_i$. Moreover, if $f : K \to K'$ is a monotone function, then additionally $\prod f(C) = \prod_i f(x_i)$. Analogous statements hold for suprema.*

Proof. We only show the statement involving f, as it implies the first, and only consider infinite C (otherwise the statement is trivial). Fix a bijection $g : \omega \to C$ and recursively define $x_0 = g(0)$ and $x_{i+1} = \min(g(i+1), x_i)$. This defines an ω-chain with $x_i \in C$ and thus $\prod_i f(x_i) \geq \prod f(C)$. Conversely, for every $c \in C$ there is an i with $g(i) = c$ and thus $c \geq x_i$. By monotonicity, $f(c) \geq f(x_i)$ and thus $\prod f(C) \geq \prod_i f(x_i)$. ◀

We are now ready to prove the universal property. The main difficulty is to show that h preserves infima of chains; we achieve this by simplifying the chain to a well-behaved *canonical chain* with similar convergence properties, as illustrated in Figure 1.

Proof of Theorem 17. Due to the additivity and multiplicity requirements for homomorphisms, h uniquely extends to monomials. For the exponent ∞, notice that continuity requires $h(x^\infty) = \prod_{n<\omega} h(x)^n$ for $x \in X$. It further follows that $h(m_1 + m_2) = h(m_1) + h(m_2)$, hence h is uniquely defined on $\mathbb{S}^\infty[X]$. Care has to be taken regarding absorption: If $m_1 \preceq m_2$, then $m_1 + m_2 = m_2$. Since h is order-preserving and K is absorptive, we also have $h(m_1 + m_2) = h(m_1) + h(m_2) = h(m_2)$ and it follows that h is well-defined.

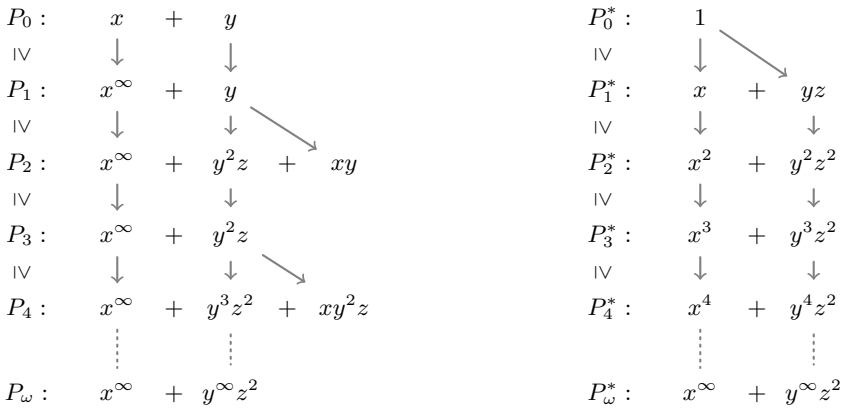


Figure 1 An example of an ω-chain of polynomials (left) and the corresponding canonical chain (right) for the proof of Theorem 17. The arrows indicate absorption between monomials of consecutive polynomials and induce a directed graph on which we then apply Kőnig's lemma.

It remains to show that h is fully continuous. Ascending chains are always finite (because of $\bigsqcup S = \text{maximals}\,(\bigcup S)$), so we only have to consider descending chains. By Lemma A2, it further suffices to consider ω-chains. Hence it suffices to prove that

$$\bigsqcap_{i<\omega} h(P_i) = h\big(\bigsqcap_{i<\omega} P_i\big)$$

for any descending ω-chain $(P_i)_{i<\omega}$ of polynomials in $\mathbb{S}^\infty[X]$. The homomorphism h preserves addition and is thus monotone, which entails the direction "$\geq$".

For the other direction, we first consider the case of single monomials. Let $(m_i)_{i<\omega}$ be a descending ω-chain of monomials. Recall that X is finite, so we can write $m_i = \prod_{x\in X} x^{m_i(x)}$. As the m_i form a descending chain, the exponents $(m_i(x))_{i<\omega}$ form an ascending chain for each $x \in X$. By Lemma A1 and the definition of h,

$$\bigsqcap_{i<\omega} h(m_i) = \prod_{x\in X} \bigsqcap_{i<\omega} h(x)^{m_i(x)} \overset{(*)}{=} \prod_{x\in X} h(x)^{(\bigsqcup_{i<\omega} m_i(x))} = h(\bigsqcap_{i<\omega} m_i).$$

where $(*)$ can easily be seen by case distinction whether $\bigsqcup_{i<\omega} m_i(x)$ is finite or ∞.

For the general case of polynomials, let $P_\omega = \bigsqcap_{i<\omega} P_i$ be the infimum, which is of the form $P_\omega = m_1 + \cdots + m_n$. We define a second, canonical ω-chain $(P_i^*)_{i<\omega}$ with the same infimum. To this end, we define the canonical monomial chain $(m_j^*)_{j<\omega}$ of a given monomial m as follows (see Figure 1 for an example),

$$m_j^*(x) = \min(j, m(x)), \qquad \text{for all } x \in X,$$

which satisfies the following properties needed for the proof:

1. If m, v are two monomials with $m \preceq v$, then $m_j^* \preceq v_j^*$ for all $j < \omega$.
2. If $m = \bigsqcap_{i<\omega} m_i$ for an ω-chain $(m_i)_{i<\omega}$ of monomials, then $\forall j\, \exists i : m_j^* \succeq m_i$.
3. In particular, $\bigsqcap_{j<\omega} m_j^* = m$.

The canonical polynomial chain $(P_j^*)_{j<\omega}$ is then defined by $P_j^* = (m_1)_j^* + \cdots + (m_n)_j^*$ for each $j < \omega$. We make the following observation:

▷ Claim.

$$\forall j\, \exists i\colon P_j^* \geq P_i.$$

We first show that the claim implies the theorem:

$$\begin{aligned}\prod_{i<\omega} h(P_i) &\overset{(1)}{\leq} \prod_{j<\omega} h(P_j^*) = \prod_{j<\omega} \Big(h((m_1)_j^*) + \cdots + h((m_n)_j^*)\Big)\\ &\overset{(2)}{=} \prod_{j<\omega} h((m_1)_j^*) + \cdots + \prod_{j<\omega} h((m_n)_j^*)\\ &\overset{(3)}{=} h\Big(\prod_{j<\omega} (m_1)_j^*\Big) + \cdots + h\Big(\prod_{j<\omega} (m_n)_j^*\Big)\\ &\overset{(4)}{=} h(m_1) + \cdots + h(m_n) = h(P_\omega),\end{aligned}$$

where (1) follows from the claim, (2) holds by Lemma A1, (3) was shown above and (4) holds due to property 3 above. Hence the claim suffices to prove the theorem.

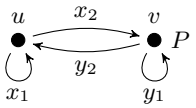
To prove the claim, assume towards a contradiction that there is a j such that $P_j^* \not\geq P_i$ for all $i < \omega$. Let us fix an $i < \omega$ for the moment. Because of $P_j^* \not\geq P_i$, there is a monomial $m_i \in P_i$ with $P_j^* \not\geq m_i$. Because of $P_{i-1} \geq P_i$, there is further $m_{i-1} \in P_{i-1}$ with $m_{i-1} \succeq m_i$. But then also $P_j^* \not\geq m_{i-1}$ (as otherwise $P_j^* \geq m_{i-1} \geq m_i$). By repeating this argument, we obtain a finite chain $m_0 \succeq m_1 \succeq \cdots \succeq m_i$ of monomials with the property that $m_k \in P_k$ and $P_j^* \not\geq m_k$ for all $0 \leq k \leq i$.

This argument applies to all $i < \omega$, so we obtain arbitrarily long finite chains with this property. By Kőnig's lemma (recall that all polynomials P_i are finite), there must be an infinite monomial chain $(m_i)_{i<\omega}$ with $m_i \in P_i$ and $P_j^* \not\geq m_i$ for all $i < \omega$. Let $m_\omega = \prod_{i<\omega} m_i$. Because of $m_i \leq P_i$ for all i, we have $m_\omega \leq P_\omega$, so there is a monomial $v \in P_\omega$ with $m_\omega \preceq v$. By considering the corresponding canonical monomial chains $(v_k^*)_{k<\omega}$ and $((m_\omega)_k^*)_{k<\omega}$ at $k = j$, we obtain a contradiction: We know from the above properties that there is an i with $(m_\omega)_j^* \succeq m_i$ and further $v_j^* \geq (m_\omega)_j^*$. Because of $v_j^* \in P_j^*$, we obtain $P_j^* \geq v_j^* \geq (m_\omega)_j^* \geq m_i$, contradicting our assumption. The claim follows, closing the overall proof. ◀

B Example of a Model-Checking Game

The proof of the main result in Sect. 6 requires some preparations and is deferred to the full version [5]. Here, we attempt to convince the reader by an example instead of rigorous arguments. Due to space reasons, we consider a small graph comprised of only two nodes. The formula $\varphi(u)$, on the other hand, features alternating least and greatest fixed points and is thus non-trivial to analyse. It expresses that there is a path from u on which P holds infinitely often. We evaluate $\varphi(u)$ using the model-compatible $\mathbb{S}^\infty[X, \bar{X}]$-interpretation π over $A = \{u, v\}$ indicated on the right, with $\pi(Pu) = 0$ and $\pi(Pv) = 1$.

$$\varphi(u) = [\mathbf{gfp}\, Xx\,.\,[\mathbf{lfp}\, Yx\,.\,\exists y\big(Exy \wedge ((Xy \wedge Py) \vee Yy)\big)](x)\,](u)$$



Intuitively, witnesses for $\varphi(u)$ are infinite paths that infinitely often visit v. There are infinitely many such paths, but the simplest ones (in terms of the different edges they use) are the paths $uvvvv\ldots$ and $uvuvuv\ldots$ which correspond to the monomials $x_2 y_1^\infty$ and $x_2^\infty y_2^\infty$. And indeed, $\pi[\![\varphi(u)]\!] = x_2 y_1^\infty + x_2^\infty y_2^\infty$. Notice that the edge x_1 does not appear in the result and we can conclude that its existence does not affect the truth of $\varphi(u)$.

Let us now consider the evaluation strategies for $\varphi(u)$ from the game-theoretic perspective. The complete model-checking game is shown in Figure 2, where rounded nodes belong to Verifier, rectangular nodes to Falsifier, and the small numbers indicate the priorities

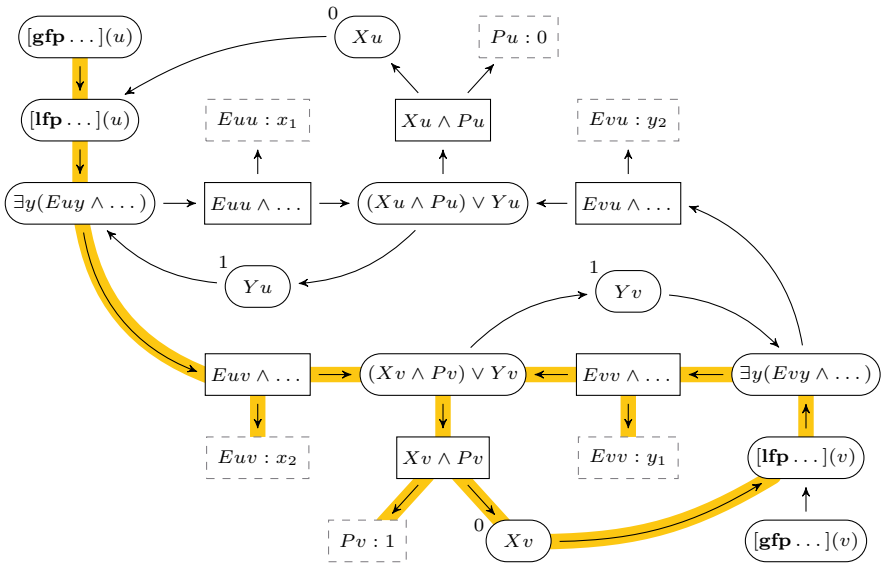


Figure 2 Model-checking game for $\varphi(u)$, with highlighted winning strategy.

assigned to fixed-point relations. Terminal positions have dashed borders and include the value assigned by π. Verifier can make decisions at four positions (two nodes labelled $\exists y(\dots)$ and two disjunctions in the center), hence there are 16 positional strategies in total. One of these strategies is highlighted (yellow color) and has the provenance value $x_2 y_1^\infty$, having one play ending in Euv and arbitrarily long plays ending either in Evv or in Pv (depending on the choices of Falsifier). Most of the other strategies allow infinite plays with least priority 1 which lead to provenance value 0 (for instance by choosing the cycle $\exists y(\dots) \to Euu \wedge \dots \to (Xu \wedge Pu) \vee Yu \to Yu$). The only remaining strategy has the provenance value $x_2^\infty y_2^\infty$. One can further observe that non-positional strategies only lead to monomials with additional variables which are then absorbed, so summing over all strategies gives $x_2 y_1^\infty + x_2^\infty y_2^\infty$ as expected.