Moscow Journal of Combinatorics and Number Theory

2020
vol. 9 no. 3

Provenance analysis for logic and games

Erich Grädel and Val Tannen

msp

Moscow Journal of Combinatorics and Number Theory
Vol. 9, No. 3, 2020
https://doi.org/10.2140/moscow.2020.9.203

Provenance analysis for logic and games

Erich Grädel and Val Tannen

A model-checking computation checks whether a given logical sentence is true in a given finite structure. Provenance analysis abstracts from such a computation mathematical information on how the result depends on the atomic data that describe the structure. In database theory, provenance analysis by interpretations in commutative semirings has been rather successful for positive query languages (such as unions of conjunctive queries, positive relational algebra, and Datalog). However, it did not really offer an adequate treatment of negation or missing information. Here we propose a new approach for the provenance analysis of logics with negation, such as first-order logic and fixed-point logics. It is closely related to a provenance analysis of the associated model-checking games, and based on new semirings of dual-indeterminate polynomials or dual-indeterminate formal power series. These are obtained by taking quotients of traditional provenance semirings by congruences that are generated by products of positive and negative provenance tokens. Beyond the use for model-checking problems in logics, provenance analysis of games is of independent interest. Provenance values in games provide detailed information about the number and properties of the strategies of the players, far beyond the question whether or not a player has a winning strategy from a given position.

1. Introduction

Provenance analysis aims at understanding how the result of a computational process with a complex input, consisting of multiple items, depends on the various parts of this input. In database theory, provenance analysis based on interpretations in commutative semirings has been developed for positive database query languages, to understand which combinations of the atomic facts in a database can be used for deriving the result of a given query. In this approach, atomic facts are interpreted not just by true or false, but by values in an appropriate semiring, where 0 is the value of false statements, whereas any element $a \neq 0$ of the semiring stands for some shade of truth. These values are then propagated from the atomic facts to arbitrary queries in the language, which permits one to answer questions such as the minimal cost of a query evaluation, the confidence one can have that the result is true, the number of different ways in which the result can be computed, or the clearance level that is required for obtaining the output, under the assumption that some facts are labeled as confidential, secret, top secret, etc. We refer to [Green and Tannen 2017] for a recent account and many references on the semiring framework for database provenance.

Scenarios to which the semiring provenance approach has been successfully applied include unions of conjunctive queries, positive relational algebra, nested relations, Datalog, XQuery, SQL-aggregates and several others, and it has been implemented in software systems such as Orchestra and Propolis.

MSC2010: 03B70, 03C13, 05C57, 68Q19, 91A43.
Keywords: finite model theory, semiring provenance, games.

For details, see, e.g., [Amsterdamer et al. 2011b; Deutch et al. 2014; Foster et al. 2008; Green 2011; Green et al. 2007; Tannen 2013]. A main limitation of this approach is that is has been largely confined to *positive* query languages. Attempts to add operations that capture *difference of relations* have led to interesting and algebraically challenging, but divergent approaches [Amsterdamer et al. 2011a; Geerts and Poggi 2010; Geerts et al. 2016; Green et al. 2009]. In particular there has been no systematic approach in database theory for tracking *negative information*, and no convincing provenance analysis for languages with full negation.

Here, we would like to develop a new approach for a semiring provenance analysis for model-checking problems of logics with negation, in particular first-order logic and fixed-point logic. This approach is based on several ideas:

• Provenance analysis of logics is intimately connected to provenance analysis of games. In the same way as formula evaluation or model checking can be formulated in game-theoretic terms, also the propagation of provenance values from atomic facts to arbitrary formulae can be viewed as a process on the associated games. Also the typical *results* of a provenance analysis of database queries or logical formulae, concerning for instance confidence scores, costs, required clearance level, or number of "proof trees" have natural game-theoretic interpretations. In fact, provenance analysis of games is of independent interest, and provenance values of positions in a game provide detailed information about the number and properties of the strategies of the players, far beyond the question whether or not a player has a winning strategy from a given position.

• We deal with negation by transformation to negation normal form. This is the common approach for the design of model-checking games and game-based evaluation algorithms. But while this is mainly a matter of convenience (to avoid role switches between players during a play), provenance semantics imposes even stronger reasons for transformations to negation normal form. Indeed, beyond Boolean semantics, negation is not a compositional logical operation: the provenance value of $\neg\varphi$ is not necessarily determined by the provenance value of φ.

• On the algebraic side, we introduce new provenance semirings of polynomials and formal power series, which take negation into account. They are obtained by taking quotients of traditional provenance semirings by congruences generated by products of positive and negative provenance tokens; they are called semirings of dual-indeterminate polynomials or dual-indeterminate power series.

Preliminary accounts of our approach, confined to first-order logic and without the connection to games, but discussing potential applications to issues such as model updates, and reverse provenance analysis (e.g., confidence maximization), have been given in [Tannen 2017; Grädel and Tannen 2017]. Here we put also the provenance analysis of games into focus; in fact we develop our approach here from the perspectives of games. We shall first discuss the case of finite acyclic games which are sufficient for the provenance analysis of first-order logic and its fragments. Most of the central issues of our approach, in particular the view of provenance values in terms of valuations of strategies and plays, appear already in this simple scenario. We shall then discuss reachability games on graphs that admit cycles. These are the games that are relevant for the provenance analysis of logics with least (but without greatest) fixed points. For these it will be necessary to restrict from arbitrary commutative semirings to ω-continuous ones. Such an analysis has previously been carried out for Datalog, but to deal with (atomic) negation we have to combine this with the idea of taking quotients by the duality on indeterminates, which will lead us to semirings of dual-indeterminate power series. Finally we shall outline a provenance approach for safety games and greatest fixed points. Our central algebraic tools here are absorptive semirings, in particular the semiring $\mathbb{S}^{\infty}[X]$ of generalized absorptive polynomials, admitting also infinite exponents.

This paper is intended to lay foundations for our general approach to a provenance analysis of logic and games, which should take us far beyond the specific cases studied here. The application of the acyclic case to modal and guarded logics has been analyzed in [Dannert and Grädel 2020]. In [Xu et al. 2018] our approach has been applied to database repairs; it has been shown how our treatment of negation, or absent information, can be used to explain and repair missing query answers and the failure of integrity constraints in databases. Further, the potential of the provenance methods developed here for applications in knowledge representation and description logics has been discussed in [Dannert and Grädel 2019]. Work in progress includes the provenance analysis of temporal and dynamic logics in the setting of absorptive semirings, the study of logics of dependence and independence from the point of view of provenance, and the algorithmic analysis of computing provenance values in various settings.

2. 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 $0 \cdot a = a \cdot 0 = 0$. A semiring is $+$-*positive* if $a + b = 0$ implies $a = 0$ and $b = 0$. This excludes rings. A semiring is *root-integral* if $a \cdot a = 0$ implies $a = 0$. All semirings considered in this paper are commutative, +-positive, and root-integral. Further, a commutative semiring is *positive* if it is +-positive and has no divisors of 0 (i.e., $a \cdot b = 0$ implies $a = 0$ and $b = 0$). The standard semirings considered in provenance analysis are in fact positive, but for an appropriate treatment of negation we shall introduce later in this paper semirings (of dual-indeterminate polynomials or power series) that have divisors of 0.

Notice that a semiring K is positive if, and only if, the unique function $h : K \to \{0, 1\}$ with $h^{-1}(0) = \{0\}$ is a homomorphism from K into the Boolean semiring $\mathbb{B} = (\{0, 1\}, \vee, \wedge, 0, 1)$. A semiring K is (+)-idempotent if $a + a = a$ for all $a \in K$ and $(+, \cdot)$-idempotent if, in addition, $a \cdot a = a$ for all a. Further, K is *absorptive* if $a + ab = a$ for all $a, b \in K$. Obviously, every absorptive semiring is (+)-idempotent.

Elements of a commutative semiring will be used as truth values for logical statements and as values for positions in games. The intuition is that $+$ describes the *alternative use* of information, as in disjunctions or existential quantifications, or for different possible choices of a player in a game, whereas $\cdot$ stands for the *joint use* of information, as in conjunctions or universal quantifications, or for choices in a game that are controlled by the opponent of the given player. Further, 0 is the value of false statements or losing positions, whereas any element $a \neq 0$ of a semiring K stands for a "nuanced" interpretation of true or as a value of a nonlosing position.

Application semirings. We briefly discuss some specific semirings that provide interesting information about a logical statement or a position in a game:

- 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 here for counting winning strategies in games. It also plays an important role for *bag semantics* in databases.

• $\mathbb{T} = (\mathbb{R}_+^\infty, \min, +, \infty, 0)$ is called the *tropical* semiring. It has many applications in several areas of computer science. It is used here for measuring the cost of strategies.

• The *Viterbi* semiring $\mathbb{V} = ([0, 1], \max, \cdot, 0, 1)$ is isomorphic to $\mathbb{T}$ via $x \mapsto e^{-x}$ and $y \mapsto -\ln y$. We will think of the elements of $\mathbb{V}$ as *confidence scores* and use it to describe the confidence that a player can win from a given position or the confidence assigned to a logical statement.

• The *min-max* semiring on a totally ordered set $(A, \leq)$ with least element a and greatest element b is the semiring $(A, \max, \min, a, b)$.

Provenance semirings. Beyond the traditional application semirings, there are some important provenance semirings of polynomials that are used for a general provenance analysis. These semirings have algebraic *universality* properties (they are freely generated) for various classes of semirings. This allows us to compute provenance values once in a general such semiring and then to specialize it via homomorphisms (i.e., evaluation of the polynomials) to specific 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 and with coefficients from $\mathbb{N}$. This is the commutative semiring freely generated by the set X.

• By dropping coefficients from $\mathbb{N}[X]$, we get the semiring $\mathbb{B}[X]$ whose elements are just finite sets of distinct monomials. It is the free (+)-idempotent semiring over X.

• By dropping also exponents, we get the semiring $\mathbb{W}[X]$ of finite sums of monomials that are linear in each argument. It is sometimes called the why-semiring.

• The free absorptive semiring $\mathbb{S}[X]$ over X consists of 0, 1 and all antichains of monomials with respect to the componentwise order on their exponents. It is the quotient of $\mathbb{N}[X]$ by the congruence induced by $p \sim q$ for monomials p, q with $p = qr$.

• Finally $\mathsf{PosBool}(X) = (\mathsf{PosBool}(X), \vee, \wedge, \bot, \top)$ is the semiring whose elements are classes of equivalent positive (monotone) Boolean expressions with variables from X (its elements are in bijection with the positive Boolean expressions in irredundant disjunctive normal form). This is the distributive lattice freely generated by the set X.

3. Games

We consider two-player turn-based games on graphs. Such a game is defined by the game graph on which it is played, and by the objectives of the players.

Definition 2. A *game graph* is a structure $\mathcal{G} = (V, V_0, V_1, T, E)$, where $V = V_0 \cup V_1 \cup T$ is the set of positions, partitioned into the sets V_0, V_1 of the two players and the set T of terminal positions, and where $E \subseteq V \times V$ is the set of moves. We denote the set of immediate successors of a position v by $vE := \{w : (v, w) \in E\}$ and require that $vE = \varnothing$ if, and only if, $v \in T$. A play from an initial position v_0 is a finite or infinite path $v_0 v_1 v_2 \cdots$ through $\mathcal{G}$ where the successor $v_{i+1} \in v_i E$ is chosen by Player 0 if $v_i \in V_0$ and by Player 1 if $v_1 \in V_1$. A play ends when it reaches a terminal node $v_m \in T$.

Definition 3. For every game graph $\mathcal{G} = (V, V_0, V_1, T, E)$, and every initial position $v_0 \in V$, the *tree unraveling* of $\mathcal{G}$ from v_0 is the game tree $\mathcal{T}(\mathcal{G}, v_0)$ consisting of all finite paths from v_0. More precisely,

$\mathcal{T}(\mathcal{G}, v) = (V^\#, V_0^\#, V_1^\#, T^\#, E^\#)$, where $V^\#$ is the set of all finite paths $\pi = v_0 v_1 \cdots v_m$ through $\mathcal{G}$, with $V_\sigma^\# = \{\pi v \in V^\# : v \in V_\sigma\}$, $T^\# = \{\pi t \in V^\# : t \in T\}$, and $E^\# = \{(\pi v, \pi v v') : (v, v') \in E\}$. For most game-theoretic considerations, the games played on $\mathcal{G}$ and its unravelings are equivalent, via the canonical projection $\rho : \mathcal{T}(\mathcal{G}, v_0) \to \mathcal{G}$ that maps every path πv to its end point v.

A strategy for a player in a game is a function that selects moves at points that are controlled by that player. A strategy need not be defined at all positions of a player, but it must be closed in the sense that it defines a move from each position that is reachable by a play that is admitted by the strategy. There are several possibilities to define the notion of a strategy formally. For our purposes it is convenient to identify a strategy with the histories of plays that it admits, i.e., to view it as an appropriate subtree of $\mathcal{T}(\mathcal{G}, v_0)$.

Definition 4. A *strategy* of Player σ (for $\sigma \in \{0, 1\}$) from v_0 in a game $\mathcal{G}$ is a subtree of $\mathcal{T}(\mathcal{G}, v_0)$, of the form $\mathcal{S} = (W, F)$ with $W \subseteq V^\#$ and $F \subseteq (W \times W) \cap E^\#$, satisfying the following conditions:

- W is closed under predecessors: if $\pi v \in W$ then also $\pi \in W$.
- If $\pi v \in W \cap V_\sigma^\#$, then $|(\pi v)F| = 1$.
- If $\pi v \in W \cap V_{1-\sigma}^\#$ then $(\pi v)F = (\pi v)E^\#$.

We write $\mathrm{Strat}_\sigma(v_0)$ for the set of all strategies of Player σ from v_0.

In a strategy $\mathcal{S} = (W, F)$, the set W is the part of $\mathcal{T}(\mathcal{G}, v_0)$ on which the strategy is defined, and F is the set of moves that are admitted by the strategy. A strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v_0)$ induces the set $\mathrm{Plays}(\mathcal{S})$ of those plays from v_0 whose moves are consistent with $\mathcal{S}$. We call $\mathcal{S}$ well-founded if it does not admit any infinite plays; this is always the case on finite acyclic game graphs, but need not be the case otherwise. The set of possible *outcomes* of a strategy $\mathcal{S}$ is the set of terminal nodes that are reachable by a play that is consistent with $\mathcal{S}$. A strategy can also be viewed as a function $\mathcal{S} : W \cap V_\sigma^\# \to V$ such that $\mathcal{S}(\pi v) \in vE$ defines the node to which Player σ moves from πv.

The simplest objectives of players are reachability and safety objectives.

Definition 5. A *reachability objective* for Player σ is given by a set $T_\sigma \subseteq T$ of winning terminal positions. With such an objective, Player σ wins every play that reaches a position in T_σ. Dually, a *safety objective* for Player σ is given by a set $L_\sigma \subseteq T$ of "losing" positions that the player has to avoid, or equivalently, by its complement $S_\sigma = V \setminus L_\sigma$, the region of safe positions inside of which the player has to keep the play. With such an objective Player σ wins every play, finite or infinite, that never reaches a position in L_σ.

Notice that the difference between reachability and safety objectives is relevant only in cases where infinite plays are possible. Indeed, in a game that admits only finite plays, Player σ wins a play with the reachability objective T_σ if, and only if, she wins that play with the safety objective given by $L_\sigma = T \setminus T_\sigma$, so we can always reformulate reachability by safety and vice versa. However, in a game that admits infinite plays, Player σ wins with a reachability objective T_σ if, and only if, her opponent, Player $1 - \sigma$, loses with the safety condition $L_{1-\sigma} = T_\sigma$. Hence winning with a reachability objective corresponds to defeating an opponent who plays with a safety objective. If both players play with reachability objectives, then infinite plays are won by neither player.

4. Provenance for well-founded games

We first study the provenance analysis of games for well-founded games, i.e., games that are played on *finite acyclic game graphs* $\mathcal{G} = (V, V_0, V_1, T, E)$, and hence do not admit infinite plays. We introduce *K-valuations* f_0 and f_1 that associate with every position $v \in V$ provenance values $f_0(v)$ and $f_1(v)$, respectively. The idea is that, for $\sigma \in \{0, 1\}$, the function f_σ describes the value of each position from the point of view of Player σ. Such a valuation is induced by its values on the terminal positions, i.e., by a function $f_\sigma : T \to K$, and by a valuation of the moves, i.e., by a function $h_\sigma : E \to K \setminus \{0\}$. Here, the function $f_\sigma : T \to K$ defines the value, for Player σ, of every *terminal position* where, intuitively, $f_\sigma(t) = 0$ means that position t is losing for Player σ. In the simplest case, we can specify reachability objectives T_σ by setting $f_\sigma(t) = 1$ for $t \in T_\sigma$ and $f_\sigma(t) = 0$ otherwise. The functions $h_\sigma : E \to K \setminus \{0\}$ provide a value (or cost) for Player σ of the moves. In many cases valuations of moves are not relevant; we then just put $h_\sigma(vw) = 1$ for all edges $(v, w) \in E$.

The extension of the basic valuations $f_\sigma : T \to K$ and $h_\sigma : E \to K \setminus \{0\}$ to valuations $f_\sigma : V \to K$ for all positions then relies on the idea that a move from v to w contributes to $f_\sigma(v)$ the value $h_\sigma(vw) \cdot f_\sigma(w)$. These contributions are summed up in the case that v is a position for Player σ (i.e., when she chooses herself the successors), and multiplied in the case that v is a position of the opponent (i.e., when she has to cope with any of the possible successors). This is summarized by the following definition.

Definition 6. Let K be a commutative semiring, let $\mathcal{G} = (V, V_0, V_1, T, E)$ be a finite acyclic game graph, and let $\sigma \in \{0, 1\}$ denote one of the two players. A *K-valuation* of $\mathcal{G}$ for Player σ is a function $f_\sigma : V \to K$. It is defined from basic valuations $f_\sigma : T \to K$ and $h_\sigma : E \to K \setminus \{0\}$ via backwards induction, by

$$f_\sigma(v) := \begin{cases} \sum_{w \in vE} h_\sigma(vw) \cdot f_\sigma(w) & \text{if } v \in V_\sigma, \\ \prod_{w \in vE} h_\sigma(vw) \cdot f_\sigma(w) & \text{if } v \in V_{1-\sigma}. \end{cases}$$

An equivalent characterization of the K-valuation f_σ can be obtained by defining provenance values for plays and strategies.

Definition 7. For a play $x = v_0 v_1 \cdots v_m$ from v_0 to a terminal node v_m, we define its valuation for Player σ as $f_\sigma(x) := h_\sigma(v_0 v_1) \cdots h_\sigma(v_{m-1} v_m) \cdot f_\sigma(v_m)$. Let now $\mathcal{S} = (W, F) \subseteq \mathcal{T}(\mathcal{G}, v_0)$ be a strategy for Player σ from v_0 and $\rho_S : (W, F) \to (V, E)$ be the restriction of the canonical homomorphism $\rho : \mathcal{T}(\mathcal{G}, v_0) \to \mathcal{G}$ to $\mathcal{S}$. For any position $v \in V$ and any move $e \in E$, the values

$$\#_{\mathcal{S}}(v) := |\rho_{\mathcal{S}}^{-1}(v)| \quad \text{and} \quad \#_{\mathcal{S}}(e) := |\rho_{\mathcal{S}}^{-1}(e)|$$

indicate how often the position v and the move e appear in the strategy $\mathcal{S}$. We then define the provenance value $\mathcal{S} \in \mathrm{Strat}_\sigma(v_0)$ as

$$F(\mathcal{S}) := \prod_{e \in E} h_\sigma(e)^{\#_{\mathcal{S}}(e)} \cdot \prod_{v \in T} f_\sigma(v)^{\#_{\mathcal{S}}(v)}.$$

In some important special cases, provenance values of strategies coincides with the product of the provenance values over all plays that they admit.

Lemma 8. *If $h_\sigma(e) = 1$ for all moves $e \in E$, or if the underlying semiring is multiplicatively idempotent (i.e., $a^2 = a$ for all a) we have that $F(\mathcal{S}) = \prod_{x \in \mathrm{Plays}(\mathcal{S})} f_\sigma(x)$ for all $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$.*

However, there are simple games where this is not the case. Consider, for instance, the valuation for Player 0 in a game where only the opponent, Player 1, moves: From position v, Player 1 can proceed to w by a move with value $h_0(vw)=a$, and from w he has the choice of moving to either s or t, both options having value 1 for Player 0. There is only one strategy $\mathcal{S}$ for Player 0 (do nothing), with provenance value a. However, the strategy admits two plays, ending in s and t, respectively, both of which have value a. Thus the product over the provenance value of the plays is a^2.

Theorem 9. *For any commutative semiring K and any finite acyclic game $\mathcal{G}$, let $f_\sigma : V \to K$ be the provenance valuation for Player σ, induced by the valuation $f_\sigma : T \to K$ of the terminal nodes and $h_\sigma : E \to K \setminus \{0\}$ of the moves. Then, for every position v*

$$f_\sigma(v) = \sum_{\mathcal{S}\in\mathrm{Strat}_\sigma(v)} F(\mathcal{S}).$$

Proof. For terminal positions v the claim is trivially true. So suppose that $v \in V_\sigma$. Then any strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ can be written in the form $\mathcal{S} = v \cdot \mathcal{S}'$ for some successor $w \in vE$ and some strategy $\mathcal{S}' \in \mathrm{Strat}_\sigma(w)$. Clearly, $\#_{\mathcal{S}}(t) = \#_{\mathcal{S}'}(t)$ for every terminal position $t \in T$. For the moves we have $\#_{\mathcal{S}}(e) = \#_{\mathcal{S}'}(e)$ for all $e \neq (v,w)$ but $\#_{\mathcal{S}}(e) = 1$ and $\#_{\mathcal{S}'}(e) = 0$ for $e = (v,w)$. This implies $F(S) = h(vw)\cdot F(\mathcal{S}')$. By induction hypothesis $f_\sigma(w) = \sum_{\mathcal{S}'\in\mathrm{Strat}_\sigma(w)} F(\mathcal{S}')$. Hence

$$f_\sigma(v) = \sum_{w\in vE} h_\sigma(vw)\cdot f_\sigma(w) = \sum_{w\in vE}\ \sum_{\mathcal{S}'\in\mathrm{Strat}_\sigma(w)} h_\sigma(vw)\cdot F(\mathcal{S}') = \sum_{\mathcal{S}\in\mathrm{Strat}_\sigma(v)} F(\mathcal{S}).$$

Finally, let $v \in V_{1-\sigma}$ with $vE = \{w_1,\dots,w_n\}$. Every strategy $\mathcal{S}\in\mathrm{Strat}_\sigma(v)$ has the form $\mathcal{S} = v(\mathcal{S}_1\cup\cdots\cup\mathcal{S}_n)$, with $\mathcal{S}_i \in \mathrm{Strat}_\sigma(w_i)$. For the terminal nodes $t \in T$ we have $\#_{\mathcal{S}}(t) = \sum_{i\le n}\#_{\mathcal{S}_i}(t)$; similarly, for all moves e from a different position than v, we have $\#_{\mathcal{S}}(e) = \sum_{i\le n}\#_{\mathcal{S}_i}(e)$, but for the moves $e = (v,w_i)$ we have $\#_{\mathcal{S}}(e) = 1$ and $\#_{\mathcal{S}_i}(e) = 0$ for all i. Thus $F(\mathcal{S}) = \prod_{w_i\in vE} h_\sigma(vw_i)\cdot F(\mathcal{S}_i)$. It follows that

$$\begin{aligned} f_\sigma(v) &= \prod_{w_i\in vE} h_\sigma(vw_i)\cdot f_\sigma(w_i) = \prod_{w_i\in vE} h_\sigma(vw_i)\cdot \sum_{\mathcal{S}_i\in\mathrm{Strat}_\sigma(w_i)} F(\mathcal{S}_i) \\ &= \sum_{v\cdot(\mathcal{S}_1\cup\cdots\cup\mathcal{S}_n)\in\mathrm{Strat}_\sigma(v)}\ \prod_{w_i\in vE} h_\sigma(vw_i)\cdot F(\mathcal{S}_i) = \sum_{\mathcal{S}\in\mathrm{Strat}_\sigma(v)} F(\mathcal{S}). \end{aligned}$$ ☐

From this description, we can derive a number of applications of provenance valuations on games. We first consider the information provided by valuations in the general provenance semirings of polynomials. Let $\mathbb{N}[T]$ be the semiring of polynomials with coefficients in $\mathbb{N}$ over indeterminates $t \in T$, where T is the set of terminal positions in an acyclic game graph $\mathcal{G} = (V, V_0, V_1, T, E)$. Let $f_\sigma : V \to \mathbb{N}[T]$ be the valuation induced by setting $f_\sigma(t) = t$ for $t \in T$ and $h_\sigma(vw) = 1$ for all edges (v,w), so that the value of a play is just its outcome, i.e., the terminal position where it ends.

Clearly, we can write $f_\sigma(v)$ as a sum of monomials $m\cdot t_1^{j_1}\cdots t_k^{j_k}$. This provides a detailed description of the number and properties of the strategies that Player σ has from position v.

Theorem 10. *The valuation $f_\sigma(v)\in\mathbb{N}[T]$ is the sum of those monomials $m\cdot t_1^{j_1}\cdots t_k^{j_k}$ (with $m, j_1,\dots,j_k>0$) such that Player σ has precisely m strategies $\mathcal{S}\in\mathrm{Strat}_\sigma(v)$ with the property that the set of possible outcomes for $\mathcal{S}$ is precisely $\{t_1,\dots,t_k\}$, and precisely j_i plays that are consistent with $\mathcal{S}$ have the outcome t_i.*

This is an immediate consequence of Theorem 9 and Lemma 8. In many cases, somewhat less-detailed information is sufficient, which can be obtained by valuations in less-informative provenance semirings than $\mathbb{N}[T]$:

• Evaluating $f_\sigma(v)$ in the idempotent semiring $\mathbb{B}[T]$ gives us the sum of monomials $t_1^{j_1} \cdots t_k^{j_k}$ for which Player σ has at least one strategy whose multiset of admitted outcomes consists of $t_1, \dots, t_k$ with multiplicities $j_1, \dots, j_k$, respectively.

• If we evaluate $f_\sigma(v)$ in $\mathbb{W}[T]$ we get the sum of monomials $t_1 \cdots t_m$ such that Player σ has a strategy whose set of outcomes is $\{t_1, \dots, t_m\}$. The information on multiplicities of strategies and outcomes is dropped.

• An interesting case is the evaluation in the absorptive semiring $\mathbb{S}[X]$. For two strategies $\mathcal{S}, \mathcal{S}' \in \mathrm{Strat}_\sigma(v)$, we say that $\mathcal{S}$ *absorbs* $\mathcal{S}'$ if for every terminal position $t \in T$, the strategy $\mathcal{S}$ admits fewer plays with outcome t than $\mathcal{S}'$. We call $\mathcal{S}$ *absorption-dominant* if it is not absorbed by any other strategy. Now, $f_\sigma(v) \in \mathbb{S}[X]$ is the sum of monomials $t_1^{j_1} \cdots t_k^{j_k}$ that describe precisely the (multiset of outcomes of the) absorption-dominant strategies of Player σ from v. See Section 11 below for a more detailed analysis of absorption among strategies.

• Finally, the evaluation of $f_\sigma(v) \in \mathsf{PosBool}[T]$ consists of those monomials $t_1 \cdots t_k$ such that $\{t_1, \dots, t_k\}$ is a minimal set among the sets of outcomes of strategies $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$.

Fix any reachability objective $W \subseteq T$. In any of these provenance semirings, we can write the polynomial $f_\sigma(v)$ as a sum $f_\sigma(v) = f_\sigma^W(v) + g_\sigma^W(v)$, where $f_\sigma^W(v)$ is the sum of those monomials that only contain indeterminates in W and $g_\sigma^W(v)$ contains the rest.

Theorem 11. *For every subset $W \subseteq T$ and every $v \in V$, Player σ has a strategy to reach W from v if, and only if, $f_\sigma^W(v) \neq 0$ (in any of the provenance semirings given above). Moreover, if we set $f(t) = 1$ for $t \in W$ and $f(t) = 0$ for $t \in T \setminus W$, and evaluate f_σ in the semiring $\mathbb{N}$ of natural numbers, then $f_\sigma(v)$ is the number of distinct winning strategies for Player σ to reach W from v.*

Evaluation in other application semirings gives further interesting information about strategies:

Cost of strategies. Given a game $\mathcal{G}$, we associate with Player 0 *cost functions* $f_0 : T \to \mathbb{R}_+$ and $h : E \to \mathbb{R}_+$ for the terminal positions and the moves. We define the cost of a strategy $\mathcal{S} \in \mathrm{Strat}_0(v)$ as the sum of the costs of all moves and outcomes that it admits, weighted by the number of their occurrences.

Proposition 12. *The* cost of an optimal strategy *from v in $\mathcal{G}$ is given by the valuation $f_0(v)$ in the tropical semiring* $\mathbb{T} = (\mathbb{R}_+^\infty, \min, +, \infty, 0)$.

Proof. Since the product in $\mathbb{T}$ is addition in $\mathbb{R}_+^\infty$, the cost of a strategy $\mathcal{S}$ for Player 0, as defined above, coincides with the valuation $f_0(\mathcal{S})$ in $\mathbb{T}$. The summation in $\mathbb{T}$ is minimization in $\mathbb{R}_+^\infty$, so from Theorem 9 we get that

$$f_0(v) = \min_{\mathcal{S} \in \mathrm{Strat}_0(v)} F(\mathcal{S})$$

describes indeed the minimal cost of a strategy for Player 0 from position v. □

Clearance levels. The access control semiring is $\mathbb{A} = (\{\mathsf{P} < \mathsf{C} < \mathsf{S} < \mathsf{T} < 0\}, \min, \max, 0, \mathsf{P})$, where P is "public", C is "confidential", S is "secret", T is "top secret", and 0 is "so secret that nobody can access it!". Let $f_\sigma : T \to \mathbb{A}$ and $h_\sigma : E \to \mathbb{A} \setminus \{0\}$ define access levels for the terminal positions and the moves for Player σ, in the sense that Player σ can make a move e if, and only if, his personal clearance level is at least $h(e)$ and similarly, he can access a terminal position t if, and only if, his clearance level is at least $f_\sigma(t)$.

Proposition 13. *The valuation $f_\sigma(v) \in \mathbb{A}$ describes the **minimal clearance level** that Player 0 needs to win from position v, i.e., to have a strategy that guarantees reaching a terminal position that is accessible for him.*

The proof is a straightforward induction.

Confidence in games. Suppose that $f_\sigma : T \to [0, 1]$ describes the confidence that Player σ puts into t being a winning position for her. We want to compute *confidence scores* $f_\sigma(v)$ to describe the confidence of Player σ that she can win from v. It is natural to define the confidence score $f_\sigma(v)$ as the *maximum* of the confidence scores of the successors $w \in vE$ in the case that $v \in V_\sigma$. For confidence scores of combinations of events whose choice is taken by an opponent, such as for the possible moves from a position $v \in V_{1-\sigma}$, there are different approaches in the literature. A popular one, with which we work here, takes the *product* of the confidence scores of the events from which the opponent chooses. Adopting this definition, the following proposition is immediate.

Proposition 14. *Confidence scores are computed as semiring valuations $f_\sigma : V \to \mathbb{V}$ in the Viterbi semiring* $\mathbb{V} = ([0, 1], \max, \cdot, 0, 1)$.

Min-max games. Finally note that valuations in a min-max semiring $(A, \max, \min, a, b)$ describe the value of positions in games where Player 0 tries to maximize and Player 1 tries to minimize the outcome of the play.

Separating valuations. The K-valuations f_0, f_1 for the two players in a game $\mathcal{G}$, as defined by Definition 6, are a priori completely independent of each other. This admits the treatment of a wide variety of games, without any restrictions on how the objectives of the two players relate to each other. For instance, in a completely cooperative game, the basic valuations of the terminal positions would be the same for Player 0 and Player 1. However, in many games, the objectives of the two players are antagonistic, and valuations f_0 and f_1 should reflect this. This motivates the following definition.

Definition 15. Let $\mathcal{G}$ be a game graph, with valuations f_0, f_1 for the two players in a semiring K, and let $U \subseteq V$ be a set of positions. We say that:

(1) f_0, f_1 for the two players are *separating* on U if for all $u \in U$, either $f_0(u) = 0$ or $f_1(u) = 0$.

(2) f_0, f_1 are *weakly separating* on U if $f_0(u) f_1(u) = 0$ for all $u \in U$. Notice that in the case where K has no divisors of 0, weakly separating valuations are in fact separating.

(3) f_0 and f_1 are *strongly separating* on U if they are separating, and in addition, $f_0(u) + f_1(u) \neq 0$ for all $u \in U$.

Proposition 16. *If two valuations f_0 and f_1 are* (*weakly*) *separating on the terminal positions of $\mathcal{G}$, then they are* (*weakly*) *separating on all positions of $\mathcal{G}$.*

Proof. Recall that all our semirings are assumed to be +-positive. For $v \in V_\sigma$, we have

$$f_\sigma(v) = \sum_{w \in vE} h(vw) f_\sigma(w) \quad \text{and} \quad f_{1-\sigma}(v) = \prod_{w \in vE} h(vw) f_{1-\sigma}(w).$$

It follows that f_0 and f_1 are separating on v if they are separating on all $w \in vE$. Further,

$$\begin{aligned} f_\sigma(v) f_{1-\sigma}(v) = &\Big(\sum_{w \in vE} h_\sigma(vw) f_\sigma(w)\Big)\Big(\prod_{w \in vE} h_{1-\sigma}(vw) f_{1-\sigma}(w)\Big) = \\ &\sum_{w \in vE} \Big(h_\sigma(vw) f_\sigma(w) \prod_{w' \in vE} h_{1-\sigma}(vw') f_{1-\sigma}(w')\Big) = \\ &\sum_{w \in vE} \Big(h_\sigma(vw) h_{1-\sigma}(vw) f_\sigma(w) f_{1-\sigma}(w) \prod_{w' \in vE \setminus \{w\}} h_{1-\sigma}(vw') f_{1-\sigma}(w')\Big). \end{aligned}$$

This proves that f_0 and f_1 are weakly separating on v if they are so on all $w \in vE$. □

The corresponding implication for strongly separating valuations does not hold for all +-positive semirings, but it holds for positive ones.

Proposition 17. *If two valuations f_0 and f_1 into a positive semiring are strongly separating on the terminal positions of $\mathcal{G}$, then they are so on all positions of $\mathcal{G}$.*

Proof. We prove this by induction. Assume that f_0 and f_1 are strongly separating on all $w \in vE$. Then $f_\sigma(v) + f_{1-\sigma}(v) = 0$ only if $f_\sigma(w) = 0$ for all $w \in vE$ and $f_{1-\sigma}(w) = 0$ for at least one $w \in vE$. But this implies $f_0(w) + f_1(w) = 0$ for some $w \in vE$, which contradicts our assumption. □

Note that for the Boolean semiring $K = \mathbb{B}$, this is just Zermelo's theorem on the determinacy of reachability games on well-founded game graphs: from every position, one of the two players has a winning strategy.

Counting positional winning strategies? A strategy is *positional* if it only depends on the current position, and not on the history of the play, i.e., if $\mathcal{S}(\pi v) = \mathcal{S}(\pi' v)$ for all v and all paths πv, $\pi' v$ that lead to v. A positional strategy can be described by a function $s : V_\sigma \to V$ or by a subgraph $\mathcal{S}$ of $\mathcal{G}$ (rather than of $\mathcal{T}(\mathcal{G}, v_0)$).

Given that in the study of games there is (for instance for algorithmic reasons) a strong interest in positional strategies, it is reasonable to ask whether there exist valuations in different semirings that count just the positional strategies. However, invariance under counting bisimulation shows that this is not possible.

Definition 18. Let $\mathcal{G} = (V, V_0, V_1, T, E)$ and $\mathcal{G}' = (V', V'_0, V'_1, T', E')$ be two game graphs. A *counting bisimulation* between $\mathcal{G}$ and $\mathcal{G}'$ is a relation $Z \subseteq V \times V$ such that for every pair $(v, v') \in Z$ we have

(1) $v \in V_\sigma$ if, and only if, $v' \in V'_\sigma$ and $v \in T$ if, and only if, $v' \in T'$, and

(2) there is a local bijection $z_{vv'} : vE \to v'E'$ between the immediate successors of v and v' such that $(w, z_{vv'}(w)) \in Z$ for every $w \in vE$.

We write $\mathcal{G}, v \sim \mathcal{G}', v'$ if there is a counting bisimulation Z between $\mathcal{G}$ and $\mathcal{G}'$ such that $(v, v') \in Z$. Notice that for any game graph $\mathcal{G}$, the relation $Z = \{(v, \pi v) : v \in V,\ \pi v \in V^\#\}$ is a counting bisimulation between $\mathcal{G}$ and its unraveling $\mathcal{T}(\mathcal{G}, v_0)$.

K-valuations of games are invariant under counting bisimilarity in the following sense. Let $\mathcal{G}$ and $\mathcal{G}'$ be two acyclic game graphs with K-valuations $f_\sigma : T \to K$ and $f'_\sigma : T' \to K$ of the terminal positions and $h : E \to K$ and $h' : E' \to K$ of the moves. We say that a counting bisimulation $Z \subseteq V \times V'$ respects these valuations if $f_\sigma(t) = f'_\sigma(t')$ for all $(t, t') \in Z \cap T \times T'$, and $h_\sigma(vw) = h'_\sigma(v'w')$ whenever $(v, v') \in Z$ and $(w, w') \in Z$.

Proposition 19. *Let Z be a counting bisimulation between $\mathcal{G}$ and $\mathcal{G}'$ that respects the basic valuations of the terminal positions and the moves. Then Z respects the valuations of all positions*; *i.e.*, $f_\sigma(v) = f'_\sigma(v')$ *for all* $(v, v') \in Z$.

Proof. Let $(v, v') \in Z$. If v and v' are terminal positions, then $f_\sigma(v) = f'_\sigma(v')$ by assumption. Otherwise, v and v' are both positions of the same player. If they belong to Player σ, then $f_\sigma(v) = \sum_{w \in vE} f_\sigma(w)$. The local bijection $z_{vv'}$ maps every $w \in vE$ to some $w' \in v'E'$ such that, by induction hypothesis, $f_\sigma(w) = f'_\sigma(w')$. Hence $f'_\sigma(v') = \sum_{w' \in v'E'} f'_\sigma(w') = \sum_{w \in vE} f_\sigma(w) = f_\sigma(v)$. If v and v' belong to Player $(1 - \sigma)$ the reasoning is completely analogous, taking a product rather than a sum. □

In particular K-valuations of acyclic games do not change if we replace a game graph $\mathcal{G}$ by one of its unravelings $\mathcal{T}(\mathcal{G}, v)$. Indeed, every valuation $f_\sigma : T \to K$ on the terminal positions of a game graph $\mathcal{G}$ extends to the same valuation for v on $\mathcal{G}$ as on the tree unraveling $\mathcal{T}(\mathcal{G}, v)$. On the other side, every strategy on a tree-shaped game graph is positional. Thus the number of positional winning strategies is certainly not invariant under unraveling and hence not definable by valuations in a semiring.

5. Provenance for first-order logic via model-checking games and dual-indeterminate polynomials

Given a finite relational vocabulary τ and a finite nonempty universe A, we denote by $\mathrm{Atoms}_A(\tau)$ the set of all atoms $R\bar{a}$ with $R \in \tau$ and $\bar{a} \in A^k$. Further, let $\mathrm{NegAtoms}_A(\tau)$ be the set of all negated atoms $\neg R\bar{a}$, where $R\bar{a} \in \mathrm{Atoms}_A(\tau)$, and consider the set of all τ-literals on A,

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

where op stands for $=$ or $\neq$.

Definition 20. Given any commutative semiring K, a *K-interpretation* (for τ and A) is a function $\pi : \mathrm{Lit}_A(\tau) \to K$ that maps equalities and inequalities to their truth values 0 or 1.

We have defined in [Grädel and Tannen 2017] how a semiring interpretation extends to a full valuation $\pi : \mathrm{FO}(\tau) \to K$ mapping any fully instantiated formula $\psi(\bar{a})$ (or equivalently, any first-order sentence of vocabulary $\tau \cup A$) to a value $\pi[\![\psi]\!]$ by setting

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

Negation is handled via negation normal forms: we set $\pi[\![\neg\varphi]\!] := \pi[\![\mathrm{nnf}(\neg\varphi)]\!]$ where $\mathrm{nnf}(\varphi)$ is the negation normal form of φ.

This is equivalent to the game provenance, as defined above, for the model-checking game associated with the formula ψ and the K-interpretation $\pi : \mathrm{Lit}_A(\tau) \to K$. Notice that classically, model-checking games are defined for a formula (assumed to be given in negation normal form) and a fixed structure $\mathfrak{A}$;

see, e.g., [Apt and Grädel 2011, Chapter 4]. However, the game graph of such a model-checking game depends only on the formula ψ and the *universe* A of the given structure $\mathfrak{A}$. It is only the labeling of the terminal positions of the game, 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 of a model-checking game readily generalizes to our more abstract provenance scenario.

Definition 21. Let $\psi(\bar{x}) \in \mathrm{FO}(\tau)$ be a first-order formula in negation normal form with a relational vocabulary τ, and let A be a (finite) universe. The model-checking game $\mathcal{G}(A, \psi)$ has positions $\varphi(\bar{a})$, obtained from a subformula $\varphi(\bar{x})$ of ψ, by instantiating the free variables $\bar{x}$ by a tuple $\bar{a}$ of elements of A. At a disjunction $(\psi \vee \varphi)$, Player 0 (Verifier) moves to either ψ or φ, and at a conjunction, Player 1 (Falsifier) makes an analogous move. At a position $\exists x \varphi(\bar{a}, x)$, The Verifier selects an element b and moves to $\varphi(\bar{a}, b)$, whereas at positions $\forall x \varphi(\bar{a}, x)$ the move to the next position $\varphi(\bar{a}, b)$ is done by the Falsifier. The terminal positions of $\mathcal{G}(A, \psi)$ are the literals in $\mathrm{Lit}_A(\tau)$.

A K-interpretation $\pi : \mathrm{Lit}_A(\tau) \to K$ thus provides a valuation of the set $T \subseteq \mathrm{Lit}_A(\tau)$ of terminal positions of the model-checking game $\mathcal{G}(A, \psi)$ for any sentence $\psi \in \mathrm{FO}(\tau \cup A)$. We view it as a valuation f_0 for Player 0. The associated valuation f_1 for Player 1 is obtained by setting $f_1(\varphi) = \pi[\![\neg\varphi]\!]$ for any literal $\varphi \in \mathrm{Lit}_A(\tau)$. Both valuations then extend to full valuations f_0 and f_1 of all positions of $\mathcal{G}(A, \psi)$, including the position ψ itself. The following result is proved by a straightforward induction on formulae.

Theorem 22. *For all positions φ of $\mathcal{G}(A, \psi)$ we have $f_0(\varphi) = \pi[\![\varphi]\!]$ and $f_1(\varphi) = \pi[\![\neg\varphi]\!]$.*

Although this theorem holds without any restrictions on the semiring K and the K-interpretation π, not all such K-interpretations are really meaningful for logic. Indeed the provenance value of complementary literals $R\bar{a}$ and $\neg R\bar{a}$ have to be related in a reasonable way, and as a consequence also the general provenance semirings of polynomials need to be modified. In the simplest case a K-interpretation defines a unique τ-structure.

Definition 23. A semiring interpretation $\pi : \mathrm{Lit}_A(\tau) \to K$ is *model-defining* if for every atom $\varphi \in \mathrm{Atoms}_A(\tau)$ one of $\pi(\varphi)$ and $\pi(\neg\varphi)$ is 0, and the other is $\neq 0$. It uniquely defines the τ-structure $\mathfrak{A}_\pi$ that has universe A, and in which precisely those literals φ are true for which $\pi(\varphi) \neq 0$.

Notice that if K is not the Boolean semiring, then several different K-interpretations may define the same structure. Further, K-interpretations are interesting, and have a number of applications, also in cases where they do not specify a single model; see [Grädel and Tannen 2017].

Dual-indeterminate polynomials. Let $X, \overline{X}$ be two disjoint sets together with a one-to-one correspondence $X \leftrightarrow \overline{X}$. We denote by $p \in X$ and $\bar{p} \in \overline{X}$ two elements that are in this correspondence. We refer to the elements of $X \cup \overline{X}$ as *provenance tokens* and we shall use "positive" and "negative" tokens p and $\bar{p}$ to annotate atoms $R\bar{a} \in \mathrm{Atoms}_A(\tau)$ and negated atoms $\neg R\bar{a} \in \mathrm{NegAtoms}_A(\tau)$, respectively. By convention, if we annotate $R(\bar{a})$ with p then the "negative" token $\bar{p}$ can only be used to annotate $\neg R(\bar{a})$, and vice versa. We refer to p and $\bar{p}$ as *complementary* tokens.

Definition 24. The semiring $\mathbb{N}[X, \overline{X}]$ is the quotient of the semiring of polynomials $\mathbb{N}[X \cup \overline{X}]$ by the congruence generated by the equalities $p \cdot \bar{p} = 0$ for all $p \in X$. This is the same as quotienting by the ideal generated by the polynomials $p\bar{p}$ for all $p \in X$. Observe that two polynomials $g, g' \in \mathbb{N}[X \cup \overline{X}]$ are congruent if, and only if, they become identical after deleting from each of them the monomials that contain complementary tokens. Hence, the congruence classes in $\mathbb{N}[X, \overline{X}]$ are in one-to-one correspondence with the polynomials in $\mathbb{N}[X \cup \overline{X}]$ such that none of their monomials contain complementary tokens. We shall call these *dual-indeterminate polynomials*.

Note that $\mathbb{N}[X, \overline{X}]$ is +-positive and root-integral, but not positive, since it has divisors of 0. Further, we have the following *universality property*:

Proposition 25. *Every function $f : X \cup \overline{X} \to K$ into any commutative semiring K with the property that $f(p) \cdot f(\bar{p}) = 0$ for all $p \in X$ extends uniquely to a semiring homomorphism $h : \mathbb{N}[X, \overline{X}] \to K$ that coincides with f on $X \cup \overline{X}$.*

Definition 26. A *provenance-tracking* interpretation is a mapping $\pi : \mathrm{Lit}_A(\tau) \to X \cup \overline{X} \cup \{0, 1\}$ such that $\pi(\mathrm{Atoms}_A(\tau)) \subseteq X \cup \{0, 1\}$ and $\pi(\mathrm{NegAtoms}_A(\tau)) \subseteq \overline{X} \cup \{0, 1\}$. Further, π maps equalities and inequalities to their truth values 0 or 1.

The idea is that if π annotates a positive or negative atom with a token, then we wish to track that literal through the model-checking computation. On the other hand annotating with 0 or 1 is done when we do not track the literal, yet we need to recall whether it holds or not in the model. See [Grädel and Tannen 2017] for more details and potential applications of provenance-tracking interpretations.

6. Semirings of dual-indeterminate power series and least fixed-point solutions

It is known that the general properties of commutative semirings are not sufficient to deal with unbounded iterations as they occur in fixed-point logic. Even for Datalog, one of the simplest fixed-point formalisms that omits the complications arising with universal quantification and negation, appropriate semirings have the additional property of being *ω-continuous*. The general ω-continuous provenance semirings are no longer semirings of polynomials, but semirings of formal power series, such as $\mathbb{N}^\infty[\![X]\!]$. We combine this here with our approach for dealing with negation by taking quotients with respect to the congruence generated by products $p\bar{p}$ of positive and negative provenance tokens. What we obtain are ω-continuous provenance semirings of dual-indeterminate power series, such as $\mathbb{N}^\infty[\![X, \overline{X}]\!]$, as well as idempotent, absorptive, and other variants thereof.

A semiring K is *naturally ordered* if the relation $a \leq b :\Leftrightarrow \exists x(a + x = b)$ is a partial order. Note that this relation is reflexive and transitive in every semiring, but it is not always antisymmetric. An ω-chain is a sequence $(a_i)_{i\in\omega}$ with $a_i \leq a_{i+1}$ for all $i \in \omega$.

Definition 27. A commutative semiring K is *ω-continuous* if it is naturally ordered and satisfies the following additional conditions:

- Every ω-chain $(a_i)_i \in \omega$ has a supremum $\sup_{i\in\omega} a_i$ in K. As a consequence, we have a well-defined infinite summation operator $\sum$ such that for every sequence $(b_i)_{i\in\omega}$
$$\sum_{i\in\omega} b_i := \sup\{a_0 + \cdots + a_n : n \in \omega\}.$$
- For every sequence $(a_i)_{i\in\omega}$ in K, every $c \in K$, and every partition $(I_j)_{j\in J}$ of ω, we have $c \cdot \sum_{i\in\omega} a_i = \sum_{i\in\omega} c \cdot a_i$ and $\sum_{j\in J} \sum_{i\in I_j} a_i = \sum_{i\in\omega} a_i$.

In an ω-continuous semiring we further have the Kleene star operation,

$$a^* := \sum_{i\in\omega} a^i = \sup_{i\in\omega}(1 + a + a^2 + \cdots + a^i).$$

A function $f : K \to K$ is ω-continuous if $\sup_{i\in\omega} f(a_i) = f(\sup_{i\in\omega} a_i)$ for every ω-chain $(a_i)_{i\in\omega}$. A consequence of the definition is that any function defined by a polynomial or a power series is ω-continuous in each argument.

Definition 28. Given a semiring K and a finite set X of indeterminates, we denote by $K[\![X]\!]$ the semiring of formal power series (i.e., possibly infinite sums of monomials) with coefficients in K and indeterminates in X, with addition and multiplication defined in the obvious way. If K is ω-continuous and $|X| = n$, then every formal power series $f \in K[\![X]\!]$ induces a well-defined function $f : K^n \to K$ which is ω-continuous in each argument. Further, if K is ω-continuous, then so is $K[\![X]\!]$ [Kuich 1997].

A *system of power series* with indeterminates $X_1, \dots, X_n$ is a sequence $G = (g_1 \cdots g_n)$ with $g_i \in K[\![X]\!]$ for each i. It induces a function $G : K^n \to K^n$ that is monotone in each argument. By Kleene's fixed-point theorem G has a least fixed point $\operatorname{lfp} G$ which coincides with the supremum of the Kleene approximants G^k, defined by $G^0 = 0,\ G^{k+1} = G(G^k)$; i.e., $\operatorname{lfp} G = \sup_{k\in\omega} G^k$. We also refer to $\operatorname{lfp} G$ as the *least fixed-point solution* of the equation system

$$X_1 = g_1(X_1, \dots, x_n), \quad \dots, \quad X_n = g_n(X_1, \dots, X_n);$$

in short, $X = G(X)$.

Dual-indeterminate power series. Semirings $K[\![X]\!]$ of power series turn out to be appropriate as general provenance semirings for (not necessarily acyclic) reachability games, without any further structure on the terminal nodes, as well as for purely positive fixed-point formalisms, without negation even on the atomic level. However, as soon as we want to deal with fixed-point logics with (atomic) negation we again need to take quotients with respect to the congruence generated by an appropriate correspondence $X \leftrightarrow \overline{X}$ between positive and negative tokens (with the same conventions as in Definition 24).

Definition 29. The semiring $K[\![X, \overline{X}]\!]$ is the quotient of the semiring of power series $K[\![X \cup \overline{X}]\!]$ by the congruence generated by the equalities $p \cdot \bar{p} = 0$ for all $p \in X$. The congruence classes in $K[\![X, \overline{X}]\!]$ are in one-to-one correspondence with the power series in $K[\![X \cup \overline{X}]\!]$ such that none of their monomials contain complementary tokens. We call these *dual-indeterminate power series*.

Again we have a universality property.

Proposition 30. *Every function $f : X \cup \overline{X} \to K$ into an ω-continuous semiring K with the property that $f(p) \cdot f(\bar{p}) = 0$ for all $p \in X$ extends uniquely to an ω-continuous semiring homomorphism $h : \mathbb{N}[\![X, \overline{X}]\!] \to K$ that coincides with f on $X \cup \overline{X}$.*

7. Provenance for reachability games with cycles

We now extend our provenance approach to games that admit infinite plays. We assume that the game graphs are finite, but no longer acyclic. Given a valuation $f_\sigma : T \to K$ in a semiring K for the terminal nodes of a game graph $\mathcal{G}$, the rules defining valuations for the other nodes have now to be read as an equation system (G_σ) in indeterminates X_v (for $v \in V$):

$$\begin{aligned} X_v &= f_\sigma(v) && \text{for } v \in T, \\ X_v &= \sum_{w \in vE} h_\sigma(vw) \cdot X_w && \text{if } v \in V_\sigma, \\ X_v &= \prod_{w \in vE} h_\sigma(vw) \cdot X_w && \text{if } v \in V_{1-\sigma}. \end{aligned} \tag{G_σ}$$

If we assume that the underlying semiring K is ω-continuous, then such a system (G_σ) always has a least fixed-point solution $\operatorname{lfp} G_\sigma$, which can be computed as the limit of its Kleene approximants $G^n : V \to K$ for $n \in \omega$. These Kleene approximants can be seen as valuations in the unravelings $\mathcal{G}^n$ of the game $\mathcal{G}$ up to n moves, defined as follows.

Recall that, for every game graph $\mathcal{G} = (V, V_0, V_1, T, E)$, and every initial position $v_0 \in V$, we have the *tree unraveling* $\mathcal{T}(\mathcal{G}, v_0) = (V^\#, V_0^\#, V_1^\#, T^\#, E^\#)$ consisting of all finite paths from v_0, with the canonical projection $\rho : \mathcal{T}(\mathcal{G}, v_0) \to \mathcal{G}$ that maps every path πv to its end point v.

Definition 31. Given $\mathcal{G}$ with basic valuations $f_\sigma : T \to K$ and $h_\sigma : E \to K \setminus \{0\}$ of the terminal positions and moves, the *truncation* $\mathcal{G}^n = (V^{(n)}, V_0^{(n)}, V_1^{(n)}, T^{(n)}, E^{(n)})$, for $n > 0$, is the restriction of the union of the trees $\mathcal{T}(\mathcal{G}, v)$ (with $v \in V$) to paths of less than n moves, and $\rho^n : \mathcal{G}^n \to \mathcal{G}$ is the restriction of the canonical homomorphism ρ to $\mathcal{G}^n$. Notice that the truncation induces new terminal nodes:

$$T^{(n)} := \{\pi v \in V^{(n)} : v \in T\} \cup \{\pi v \in V^{(n)} : |\pi| = n - 1,\ v \in V \setminus T\}.$$

In $\mathcal{G}^n$, we define the basic valuation of the moves, $h_\sigma^n : E^{(n)} \to K \setminus \{0\}$, in the obvious way, by $h_\sigma^n(e) := h_\sigma(\rho^n(e))$. For the valuation of the terminal nodes $\pi v \in T^{(n)}$, we put $f_\sigma^n(\pi v) = f_\sigma(v)$ if $v \in T$, and $f_\sigma^n(\pi v) = 0$ otherwise, i.e., if πv is an initial segment of a play in $\mathcal{G}$, with $n - 1$ moves, that has not reached a terminal position in T.

The games $\mathcal{G}^n$ are finite acyclic games, and the basic valuations extend to valuations $f_\sigma^n : V^{(n)} \to K$ for all nodes of $\mathcal{G}^n$. By induction, it readily follows that, for all nodes v of $\mathcal{G}$, the Kleene approximants G^n of (G_σ) coincide with these valuations.

Lemma 32. *For all n and all positions v of G, we have $G^n(v) = f_\sigma^n(v)$.*

We denote the strategy space of Player σ from v in $\mathcal{G}^n$ by $\mathrm{Strat}_\sigma^{(n)}(v)$. Since the games $\mathcal{G}^n$ are acyclic, Theorem 9 applies.

Lemma 33. *For every n, and every position v, we have $f_\sigma^n(v) = \sum_{\mathcal{T} \in \mathrm{Strat}_\sigma^{(n)}(v)} F(\mathcal{T})$.*

Valuations of plays and strategies in games with cycles. To generalize Theorem 9 to reachability games with cycles, we first need to extend the valuations of plays and strategies to such games. As in Section 4 a *finite* play $x = v_0 v_1 \cdots v_m$ in $\mathcal{G}$ from v_0 to a terminal node $v_m \in T$ gets the valuation $f_\sigma(x) = h_\sigma(v_0 v_1) \cdots h_\sigma(v_{m-1} v_m) \cdot f_\sigma(v_m)$. The provenance value of an infinite play is defined to be 0. For a strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$, we put $F(\mathcal{S}) := 0$ if $\mathcal{S}$ admits any infinite play. Hence a strategy $\mathcal{S}$ can have a nonzero provenance value only when it admits just finite plays. By König's lemma, it then admits only a finite number of plays, and putting, as in Section 4,

$$F(\mathcal{S}) := \prod_{e\in E} h_\sigma(e)^{\#_{\mathcal{S}}(e)} \cdot \prod_{v\in T} f_\sigma(v)^{\#_{\mathcal{S}}(v)},$$

we have $F(\mathcal{S})$ is well-defined for such strategies, as the values $\#_{\mathcal{S}}(e)$ and $\#_{\mathcal{S}}(v)$ are finite, for all $e \in E$ and $v \in T$. Although the number of different strategies $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ may well be infinite, Theorem 9 generalizes to reachability games with cycles, with a proof based on Kleene's fixed-point theorem, and the unravelings of $\mathcal{G}$ to finite acyclic games $\mathcal{G}^n$.

Notice that every strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ for the original game $\mathcal{G}$ induces, in every game $\mathcal{G}^n$, a strategy $\mathcal{S}^{(n)} \in \mathrm{Strat}_\sigma^{(n)}(v)$ for the game $\mathcal{G}_n$.

Lemma 34. *For every strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ in $\mathcal{G}$ with $F(\mathcal{S}) \neq 0$ there exists some $n_{\mathcal{S}} < \omega$ such that*

- *$\mathcal{S} = \mathcal{S}^{(n)}$ for all $n \geq n_{\mathcal{S}}$,*
- *$F(\mathcal{S}^{(m)}) = 0$ for all $m < n_{\mathcal{S}}$.*

Proof. This readily follows from the fact that a strategy $\mathcal{S}$ with $F(\mathcal{S}) \neq 0$ admits only a finite number of plays, all of which are finite. Let $n_{\mathcal{S}}$ be the maximal length of these plays. Then, for $n \geq n_{\mathcal{S}}$, all plays in $\mathcal{S}$ are already contained in $\mathcal{S}^{(n)}$. For any $m < n_s$, the induced strategy admits an unfinished play; hence $F(\mathcal{S}^m) = 0$. □

Every strategy $\mathcal{T} \in \mathrm{Strat}_\sigma^{(n)}(v)$ can be obtained as the induced strategy of some $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ such that $\mathcal{T} = \mathcal{S}^{(n)}$. In general $\mathcal{S}$ is not uniquely determined by $\mathcal{T}$ and n. Nevertheless, we have the following.

Lemma 35. *For every position v of $\mathcal{G}$ and every $n < \omega$, we have that in $\mathcal{G}^n$*

$$\sum_{\mathcal{S}\in\mathrm{Strat}_\sigma(v)} F(\mathcal{S}^{(n)}) = \sum_{\mathcal{T}\in\mathrm{Strat}_\sigma^{(n)}(v)} F(\mathcal{T}).$$

Proof. If we have two strategies $\mathcal{S}_1 \neq \mathcal{S}_2$ in $\mathrm{Strat}_\sigma(v)$ with $\mathcal{T} = \mathcal{S}_1^{(n)} = \mathcal{S}_2^{(n)}$, then $\mathcal{T}$ must contain an unfinished play (otherwise $\mathcal{T} = \mathcal{S}_1 = \mathcal{S}_2$), which implies that $F(\mathcal{T}) = 0$. Thus, although the strategy spaces $\mathrm{Strat}_\sigma(v)$ are in general infinite, whereas $\mathrm{Strat}_\sigma^{(n)}(v)$ is finite for each fixed n, those strategies that provide nonzero values to the sums are in one-to-one correspondence, and the two sums have the same value. □

Putting these observations together, we obtain the desired generalization of Theorem 9.

Theorem 36. *For every game graph $\mathcal{G}$ with basic valuations f_σ and h_σ of the terminal positions and moves in an ω-continuous semiring K, we have that, for every position v,*

$$f_\sigma(v) := (\mathrm{lfp}\, G_\sigma)(v) = \sum_{\mathcal{S}\in\mathrm{Strat}_\sigma(v)} F(\mathcal{S}).$$

In the cases where $h(e) = 1$ for all e, or where K is multiplicatively idempotent, we further have that

$$f_\sigma(v) = \sum_{\mathcal{S}\in\mathrm{Strat}_\sigma(v)} \prod_{x\in\mathrm{Plays}(\mathcal{S})} f_\sigma(x).$$

Proof. By the lemmata above, we have that, for every $n < \omega$,

$$G^n(v) = f_\sigma^n(v) = \sum_{\mathcal{T} \in \mathrm{Strat}_\sigma^{(n)}(v)} F(\mathcal{T}) = \sum_{\mathcal{S} \in \mathrm{Strat}_\sigma(v)} F(\mathcal{S}^{(n)}).$$

Since, for every strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ we have $F(\mathcal{S}) = F(\mathcal{S}^n)$ for sufficiently large n, the result follows by taking suprema. $\square$

For the case of game valuations $f_\sigma : V \to \mathbb{N}[\![T]\!]$, given by the basic valuations $f_\sigma(t) = t$ for terminal positions $t \in T$ and $h_\sigma(vw) = 1$ for all moves $(v, w) \in E$, we again get precise information about the number of strategies that a player has for a specific outcome. Indeed, $f_\sigma(v)$ is a (possibly) infinite sum of monomials $m \cdot t_1^{j_1} \cdots t_k^{j_k}$.

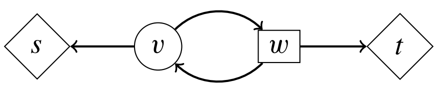
Corollary 37. *Let $f_\sigma : V \to \mathbb{N}[\![T]\!]$ be the valuation of Player σ for the game $\mathcal{G}$ in $\mathbb{N}[\![T]\!]$. For every monomial $m \cdot t_1^{j_1} \cdots t_k^{j_k}$ in $f_\sigma(v)$ (with $m \in \mathbb{N}$ and $j_i > 0$) Player σ has precisely m strategies $\mathcal{S}$ from v with the property that the set of possible outcomes for $\mathcal{S}$ is precisely $\{t_1, \ldots, t_k\}$, and precisely j_i plays that are consistent with $\mathcal{S}$ have the outcome t_i.*

Let $\mathcal{G} = (V, V_0, V_1, T, E)$ be a game with reachability objectives T_0, T_1 for the two players such that $T_0 \cap T_1 = \varnothing$. Let $W_0, W_1 \subseteq V$ be the winning regions for the two players; i.e., W_σ is the set of those positions $v \in V$ such that Player σ has a strategy from v to force the play to T_σ. Note that V is the disjoint union of the W_0, W_1 and U, the set of those positions from which none of the two players has a winning strategy. By Zermelo's theorem both players have strategies to guarantee that each play from U will be at least a draw.

Corollary 38. *Let $f_\sigma : T \to K$ be a valuation of the terminal positions of $\mathcal{G}$ in an ω-continuous semiring, with $f_\sigma(t) \neq 0$ if, and only if, $t \in T_\sigma$. The least fixed-point solution of the equation system F_σ extends this to a valuation $f_\sigma : V \to K$, with $f_\sigma(v) \neq 0$ if, and only if, $v \in W_\sigma$.*

Notice that weakly contradictory valuations f_0 and f_1 on the terminal positions extend to weakly contradictory valuations on all positions. However, even valuations into ω-continuous semirings that are strongly contradictory on the terminal positions are in general only weakly contradictory on the set of all positions, unless $W_0 \cup W_1 = V$, since $f_0(U) = f_1(U) = 0$.

Example 39. We illustrate our findings by the following very simple example of a game where Player 0 moves from v, Player 1 moves from w, and s and t are terminal nodes:



The corresponding equation system for Player 0 has the equations $X_v = s + X_w$ and $X_w = t \cdot X_v$. In $\mathbb{N}^\infty[\![s, t]\!]$ the least fixed-point solution is $f(v) = s \cdot (1 + t + t^2 \cdots)$ and $f(w) = s \cdot (t + t^2 + \cdots)$. If we evaluate it for the reachability objectives $\{s\}$ and $\{t\}$, respectively, we obtain $f(v)[0, t] = f(w)[0, t] = 0$, which illustrates that neither from v nor from w does Player 0 have a strategy to reach t. On the other side, $f(v)[s, 0] = s$ and $f(w)[s, 0] = 0$, which is consistent with the fact that Player 0 has a strategy to reach s from v but not from w.

But the formal power series $f(v)$ and $f(w)$ reveal more information than that. For instance, the fact that $f(v)$ contains, for every n, the monomial $s \cdot t^n$ implies that Player 0 has precisely one strategy S

from v that admits precisely $n+1$ consistent plays, one of which has outcome s and the other n have outcome t; this is the strategy where Player 0 moves from v to w the first n times, and then to s. Notice that Player 0 also has one further strategy, namely the (positional) strategy to move always to w. However, this strategy does not guarantee that the play terminates and therefore has value 0, so it is not visible in the provenance values $f(v)$ and $f(w)$.

8. Provenance analysis for positive LFP

Least fixed-point logic, denoted LFP, extends first-order logic by least and greatest fixed points of definable monotone operators on relations: if $\psi(R, \bar{x})$ is a formula of vocabulary $\tau \cup \{R\}$, in which the relational variable R occurs only positively, and if $\bar{x}$ is a tuple of variables such that the length of $\bar{x}$ matches the arity of R, then $[\mathrm{lfp}\, R\bar{x}\,.\,\psi](\bar{x})$ and $[\mathrm{gfp}\, R\bar{x}\,.\,\psi](\bar{x})$ are also formulae (of vocabulary τ). The semantics of these formulae is that $\bar{x}$ is contained in the least (respectively the greatest) fixed point of the update operator $F_\psi : R \mapsto \{\bar{a} : \psi(R, \bar{a})\}$. Due to the positivity of R in ψ, any such operator F_ψ is monotone and therefore has, by the Knaster–Tarski theorem, a least fixed point $\mathrm{lfp}(F_\psi)$ and a greatest fixed point $\mathrm{gfp}(F_\psi)$. See, e.g., [Grädel et al. 2007] for background on LFP.

Note that in formulae $[\mathrm{lfp}\, R\bar{x}\,.\,\psi](\bar{x})$ one may allow ψ to have other free variables besides $\bar{x}$; these are called parameters of the fixed-point formula. However, at the expense of increasing the arity of the fixed-point predicates and the number of variables, one can always eliminate parameters. For the construction of model-checking games and also for provenance analysis it is convenient to assume that formulae are parameter-free. The duality between least and greatest fixed points implies that for any ψ

$$[\mathrm{gfp}\, R\bar{x}\,.\,\psi](\bar{x}) \equiv \neg[\mathrm{lfp}\, R\bar{x}\,.\,\neg\psi[R/\neg R]](\bar{x}).$$

Using 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 of positive least fixed points. We denote by posLFP the fragment of LFP consisting of formulae in negation normal form such that all their fixed-point operators are least fixed points. It is known that, on finite structures (but not in general), posLFP has the same expressive power as full LFP, and thus captures all polynomial-time computable properties of ordered finite structures [Grädel et al. 2007] .

An advantage of dealing with posLFP, rather than full LFP, is that it admits much simpler model-checking games. Indeed the appropriate games for LFP are *parity games*, whereas for posLFP, reachability games are sufficient. This can be exploited to define provenance interpretations for fixed-point formulae, along the lines described in the previous section.

Definition 21 of model-checking games $\mathcal{G}(A, \psi)$ for $\psi \in \mathrm{FO}(\tau)$ extends to formulae $\psi(\bar{x}) \in \mathrm{posLFP}(\tau)$ as follows: for every subformula of ψ of the form $\vartheta := [\mathrm{lfp}\, R\bar{x}\,.\,\varphi(R, \bar{x})](\bar{x})$ we add moves from positions $\vartheta(\bar{a})$ to $\varphi(\bar{a})$, and from positions $R\bar{a}$ to $\varphi(\bar{a})$ for every tuple $\bar{a}$. Since these moves are unique it makes no difference to which of the two players we assign the positions $\vartheta(\bar{a})$ and $R\bar{a}$. The resulting game graphs $\mathcal{G}(A, \psi)$ may contain cycles, but the set T of terminal nodes is again a subset of $\mathrm{Lit}_A(\tau)$.

A K-interpretation $\pi : \mathrm{Lit}_A(\tau) \to K$ into an ω-continuous semiring thus provides a valuation of the terminal positions of the game graph $\mathcal{G}(A, \psi)$ for any $\psi \in \mathrm{posLFP}(\tau)$. By Corollary 38 this extends to a valuation $f_0 : V \to K$ on the set V of all positions $\varphi(\bar{a})$ of $\mathcal{G}(A, \psi)$, including position ψ itself.

Definition 40. For any instantiated subformula φ of a sentence $\psi \in \text{posLFP}$, we define the provenance value $\pi[\![\varphi]\!]$ by its game valuation: $\pi[\![\varphi]\!] := f_0(\varphi)$.

In particular, if π is model-defining, then f_0 provides truth values for all fully instantiated subformula φ of ψ on the structure $\mathfrak{A}_\pi$ that π describes. Indeed $\mathfrak{A}_\pi \models \varphi$ if, and only if, $\pi[\![\varphi]\!] \neq 0$, and in that case the value $\pi[\![\varphi]\!]$ gives us additional information, how and why φ holds in $\mathfrak{A}$, for instance by information on the winning strategies that the Verifier has available for establishing the truth of φ in $\mathfrak{A}_\pi$. However, contrary to the case of first-order logic, in the case where $\mathfrak{A}_\pi \not\models \varphi$, and hence $\pi[\![\varphi]\!] = 0$, we do not get additional information on the reasons why φ is false. The possibility to move to $\neg\varphi$ (or more precisely, its negation normal form) and to do the provenance analysis for that formula does not exist here since $\neg\varphi$ is not a formula of posLFP. In fact, the model-checking game for $\neg\varphi$ is not a reachability game, but a safety game. To deal with safety games and greatest fixed points we shall have to impose additional restrictions on the underlying semirings. We shall discuss this below.

One can define provenance values for posLFP-sentences also directly by a fixed-point interpretation in ω-commutative semirings. The goal is to extend, by induction over the syntax, a K-interpretation $\pi : \text{Lit}_A(\tau) \to K$ to valuations $\pi[\![\psi]\!] \in K$ for all sentences $\psi \in \text{posLFP}(\tau \cup A)$. The rules for first-order operations are defined already, so we just have to consider sentences of form $\psi(\bar{a}) = [\text{lfp}\, R\bar{x}.\varphi(R, \bar{x})](\bar{a})$, with $\varphi \in \text{posLFP}(\tau \cup \{R\})$. If R has arity m, then its K-interpretations of A are functions $g : A^m \to K$. These functions are ordered by $g \leq g'$ if, and only if, $g(\bar{a}) \leq g'(\bar{a})$ for all $\bar{a} \in A^m$. Given a K-interpretation $\pi : \text{Lit}_A(\tau) \to K$, we denote by $\pi[R \mapsto g]$ the K-interpretation of $\text{Lit}_A(\tau) \cup \text{Atoms}_A(\{R\})$ obtained from π by adding values $g(\bar{c})$ for the atoms $R\bar{c}$. (Notice that R appears only positively in φ, so negated atoms are not needed.)

The formula $\varphi(R, \bar{x})$ 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

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

By Kleene's fixed-point theorem, the operator F^φ_π has a least fixed point $\text{lfp}(F^\varphi_\pi)$ which coincides with the limit of the sequence $(g^n)_{n<\omega}$ with $g^0 := 0$ and $g^{n+1} := F^\varphi_\pi(g^n)$, and which we may define as the provenance value of $[\text{lfp}\, R\bar{x}.\varphi(R, \bar{x})](\bar{a})$. The two definitions coincide.

Proposition 41. *For every formula* $[\text{lfp}\, R\bar{x}.\varphi(R, \bar{x})] \in \text{posLFP}$ *and every K-interpretation* $\pi : \text{Lit}_A(\tau) \to K$ *into an ω-continuous semiring*, $\pi[\![\, [\text{lfp}\, R\bar{x}.\varphi(R, \bar{x})](\bar{a})]\!] = \text{lfp}(F^\varphi_\pi)(\bar{a})$.

The proof is a rather straightforward adaptation of the correctness proof for model-checking games for LFP; see, e.g., [Grädel et al. 2007, Chapter 3.3].

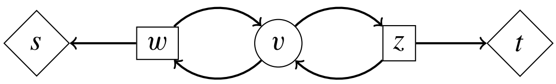
9. Beyond reachability: safety games and greatest fixed points

While the restriction of LFP to its positive fragment comes with no loss of expressive power (on finite structures) and while posLFP is sufficiently powerful to capture a number of interesting and relevant other fixed-point formalisms in computer science, it is nevertheless not really satisfactory. One reason is that the transformation from a fixed-point formula with nonatomic negation into one in posLFP is (contrary to transformations into negation normal form) not a simple syntactic translation. It goes through the stage comparison theorem and can make a formula much longer and more complicated. Further, such transformations are not available for important fixed-point formalism such as the modal μ-calculus, stratified Datalog, transitive closure logics, and even simple temporal languages such as CTL. On the game-theoretic side, reachability games are just the simplest kind of games on graphs, and in many applications players have different and more ambitious goals such as safety, Büchi, parity or Muller objectives. It is thus an important and interesting challenge to lay the foundations of a provenance analysis for full LFP and infinite games with more general objectives, and to apply this approach to the numerous other fixed-point formalisms, in particular in databases and verification.

We defer a detailed treatment of this to forthcoming work. Here we discuss some of the mathematical concepts and challenges that arise in this project, and apply them to the provenance of *safety games*. Recall that the computation of winning positions for safety objectives is a simple, but also in some sense universal, application of greatest fixed points.

The first observation is that we need to impose additional requirements on the semirings that we consider. While ω-continuous semirings are appropriate for a provenance analysis of least fixed points and reachability objectives, they are not always adequate for greatest fixed points. The property of ω-continuity is not sufficient to guarantee the existence of greatest fixed points, and in cases where they exist they do not necessarily provide the information that we are interested in.

Example 42. We consider the game graph



with associated equation system for Player 0 consisting of $X_v = X_w + X_z$, $X_w = f(s) \cdot X_v$, and $X_z = f(t) \cdot X_v$. The least fixed-point solution (in whatever semiring) has values $f(v) = f(w) = f(z) = 0$, which reflects the fact that Player 0 has no strategy to guarantee a finite play. It is not difficult to see that in $\mathbb{N}^\infty[\![s, t]\!]$ this is in fact the unique fixed point, hence in particular the greatest one, which however gives us no information about safety strategies. In $\mathbb{N}^\infty$ instead, under a valuation of the terminal nodes with $f(s) = a \neq 0$ and $f(t) = 0$, we get the greatest fixed point $f(v) = f(w) = \infty$ and $f(z) = 0$. In particular, greatest fixed points do not specialize correctly from $\mathbb{N}^\infty[\![s, t]\!]$ to $\mathbb{N}^\infty$.

We shall see below that we get interesting information on safety strategies by provenance values in the absorptive semiring $\mathbb{S}^\infty[s, t]$.

To make sure that also greatest fixed points of polynomial equation systems exist, we shall require that our semirings are not just ω-continuous, but also *ω-cocontinuous*, i.e., that every descending ω-chain $(a_i)_{i\in\omega}$, with $a_{i+1} \leq a_i$ for all $i \in \omega$, has an infimum $\inf_{i\in\omega} a_i$ in K, which is compatible with the semiring operations in the sense that, for every $c \in K$,

$$c + \inf_{i\in\omega} a_i = \inf_{i\in\omega}(c + a_i) \quad \text{and} \quad c \cdot \inf_{i\in\omega} a_i = \inf_{i\in\omega}(c \cdot a_i).$$

We call such semirings *fully ω-continuous*. Our most important example of such a semiring is $\mathbb{S}^\infty[X]$, the semiring of generalized absorptive polynomials, which we are going to discuss next.

10. Absorptive semirings and generalized absorptive polynomials

Recall that a semiring K is *absorptive* if $a + ab = a$ for all $a, b \in K$, which is equivalent to $1 + a = 1$ for all $a \in K$. Examples include the Viterbi semiring, the tropical semiring, min-max semirings, and further the semiring $\mathbb{S}[X]$ of absorptive polynomials over X. Absorptive semirings are (+)-idempotent and naturally ordered, 1 is the top element, and multiplication decreases elements: $ab \leq b$. In particular, the powers of an element form a descending ω-chain $1 \geq a \geq a^2 \geq \cdots$. If this chain has an infimum then we denote it by a^∞.

In the semiring $\mathbb{S}[X]$, the infima of descending ω-chains $(x^n)_{n<\omega}$ are always 0 and thus not very informative. We therefore complete $\mathbb{S}[X]$ to the semiring $\mathbb{S}^\infty[X]$ by admitting exponents in $\mathbb{N}^\infty$.

Definition 43. Let X be a *finite* set of provenance tokens. A *monomial* over X with exponents from $\mathbb{N}^\infty$ is a function $m : X \to \mathbb{N}^\infty$. Informally, we write m as $x_1^{m(x_1)} \cdots x_n^{m(x_n)}$. Monomial multiplication adds the exponents. Observe also that $x^\infty \cdot x^n = x^\infty$. For any two monomials, m_1, m_2 we say that m_2 *absorbs* m_1 if m_2 has smaller exponents than m_1. Formally, $m_1 \preceq m_2$ if, and only if, $m_1(x) \geq m_2(x)$ for all $x \in X$. Since monomials are functions, this is the pointwise partial order given by the order on $\mathbb{N}^\infty$.

Because $\mathbb{N}^\infty$ is a lattice (with top and bottom) the monomials also inherit a lattice structure. The set of all monomials is, of course, infinite. However, it has some crucial finiteness properties.

Proposition 44. *Every ascending chain and every antichain of monomials is finite.*

Proof. Clearly $(\mathbb{N}^\infty, \leq)$ is a well-order. For any finite set X, 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-quasiorder 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 45. We define $\mathbb{S}^\infty[X]$ as the set of antichains of monomials with indeterminates from X and exponents in $\mathbb{N}^\infty$. Writing an antichain as a (formal) sum of its monomials we identify it with a polynomial with coefficients 0 or 1, and call these *generalized absorptive polynomials*. We define polynomial addition and multiplication as usual, except that for coefficients $1 + 1 = 1$, and that we keep only the maximal monomials in the result. The empty antichain corresponds to the 0 polynomial. The 1 polynomial consists of just the monomial in which every indeterminate has exponent 0.

Proposition 46. $(\mathbb{S}^\infty[X], +, \cdot, 0, 1)$ *is an absorptive commutative semiring. Further it is a complete lattice with respect to the natural order, which is fully ω-continuous and moreover completely distributive.*

As a consequence, we can compute not only least fixed point solutions for systems of polynomial equations but also greatest fixed points. In contrast to other semirings with such properties, such as for instance the Viterbi semiring, $\mathbb{S}^\infty[X]$ has one further crucial property. It is *chain-positive* which means that the infimum of every chain of nonzero elements is also nonzero.

As in other semirings of polynomials and power series we can also here take pairs of positive and negative indeterminates, with a correspondence $X \leftrightarrow \overline{X}$ and build the quotient with respect to the congruence generated by the equation $x \cdot \bar{x} = 0$. We thus obtain a new semiring $\mathbb{S}^\infty[X, \overline{X}]$ which provides a natural framework for a provenance analysis for full LFP and other fixed-point calculi. We shall develop this in forthcoming work.

Here we use the semiring $\mathbb{S}^\infty[T]$ to describe a provenance analysis for safety games where T is the set of terminal positions of the given game graph.

11. Absorption among strategies

Definition 47. Let $\mathcal{G} = (V, V_0, V_1, T, E)$ be a finite game graph, and $v \in V$. For two strategies $\mathcal{S}, \mathcal{S}' \in \mathrm{Strat}_\sigma(v)$, we say that $\mathcal{S}$ *absorbs* $\mathcal{S}'$ (in symbols $\mathcal{S} \succeq_a \mathcal{S}'$) if

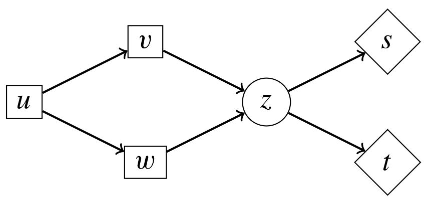
- for all $t \in T$, $\mathcal{S}$ admits at most as many plays with outcome t as $\mathcal{S}'$ does, and
- if $\mathcal{S}$ admits an infinite play, then so does $\mathcal{S}'$.

We call $\mathcal{S}$ *absorption-dominant* if it is maximal with respect to $\succeq_a$.

Absorption-dominant strategies are interesting both for games in general and for logic because they can win "with minimal effort". As a simple example, consider a model-checking game for a formula $\varphi \vee (\varphi \wedge \psi)$. The Verifier can either establish φ or $\varphi \wedge \psi$, but any strategy that establishes the truth of $\varphi \wedge \psi$ will have more plays and more outcomes than one that proves just φ, and will thus be absorbed by it. The absorption-dominant strategies for $\varphi \vee (\varphi \wedge \psi)$ are thus precisely the absorption-dominant strategies for φ.

Notice however that, despite this minimality, absorption-dominant strategies need not be positional, not even in acyclic games.

Example 48. Consider the game



There are four strategies in $\mathrm{Strat}_0(u)$ with provenance values s^2, st, st, and t^2. The positional ones are those with values s^2 and t^2, but all four strategies are absorption-dominant.

However, absorption-dominant strategies are *weakly positional* in the sense that if a node is reached several times during the same play, then, without loss of strategic power, the player can always make the same choice at that node. Absorption among strategies makes sense for both acyclic and cyclic games. In acyclic games, absorption-dominant strategies are described by provenance polynomials in $\mathbb{S}[T]$ (with only finite exponents). But they are even more interesting for the analysis of reachability *and safety* games that admit infinite plays. The fundamental difference between valuations for reachability and safety strategies concerns the valuations of infinite plays. If, as we assume here, reachability and safety goals are defined for terminal nodes, then an infinite play is losing for every reachability objective but winning for every safety objective. As a consequence, the strategies $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ that enforce all plays to be nonterminating absorb all other strategies in $\mathrm{Strat}_\sigma(v)$ that admit at least one infinite play.

We thus extend the valuations of plays in a game $\mathcal{G}$ (with finite game graph that may contain cycles) to two different valuation functions f^μ_σ and f^ν_σ. For simplicity, we assume trivial valuations on the edges, so for a finite play x ending in t, we just put $f^\mu_\sigma(x) = f^\nu_\sigma(x) = f_\sigma(t)$ but if x is an infinite play, we put $f^\mu_\sigma(x) = 0$ and $f^\nu_\sigma(x) = 1$.

A strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ may well admit an infinite set of plays. Taking the semiring $\mathbb{S}^\infty[T]$ with the basic valuation $f_\sigma(t) := t$ for the terminal nodes, strategies are described by monomials (or 0), and we put

$$F^{\mu}(\mathcal{S}) := \prod_{x \in \mathrm{Plays}(\mathcal{S})} f^{\mu}_{\sigma}(x) \quad \text{and} \quad F^{\nu}(\mathcal{S}) := \prod_{x \in \mathrm{Plays}(\mathcal{S})} f^{\nu}_{\sigma}(x) = \prod_{t \in T} t^{\#_{\mathcal{S}}(t)}.$$

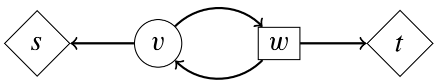
We extend the absorption order $\succeq$ on monomials by $m \succeq 0$ for all m.

Lemma 49. *Let $\mathcal{G}$ be any finite game graph, with valuations of strategies in $\mathbb{S}^{\infty}[T]$ induced by $f_{\sigma}(t) = t$ for all $t \in T$. For all strategies $\mathcal{S}, \mathcal{S}' \in \mathrm{Strat}_{\sigma}(v)$ we have:*

- $0 \neq F^{\nu}(\mathcal{S}) \succeq F^{\mu}(\mathcal{S}) \neq 1$.
- *$F^{\mu}(\mathcal{S}) = 0$ if, and only if, $\mathcal{S}$ admits an infinite play. Otherwise $F^{\mu}(\mathcal{S}) = F^{\nu}(\mathcal{S})$.*
- *$F^{\nu}(\mathcal{S}) = 1$ if, and only if, $\mathcal{S}$ admits only infinite plays.*
- *$\mathcal{S}$ absorbs $\mathcal{S}'$ if, and only if, both $F^{\nu}(\mathcal{S}) \succeq F^{\nu}(\mathcal{S}')$ and $F^{\mu}(\mathcal{S}) \succeq F^{\mu}(\mathcal{S}')$.*

Proof. Only the last item requires proof. Suppose that $\mathcal{S}$ absorbs $\mathcal{S}'$. If $\mathcal{S}$ admits only finite plays, then $F^{\mu}(\mathcal{S}) = F^{\nu}(\mathcal{S}) \succeq F^{\nu}(\mathcal{S}') \succeq F^{\mu}(\mathcal{S}')$. If $\mathcal{S}$ admits an infinite play, then so does $\mathcal{S}'$ and $F^{\nu}(\mathcal{S}) \succeq F^{\nu}(\mathcal{S}') \succeq F^{\mu}(\mathcal{S}') = F^{\mu}(\mathcal{S}) = 0$. In both cases, $F^{\nu}(\mathcal{S}) \succeq F^{\nu}(\mathcal{S}')$ and $F^{\mu}(\mathcal{S}) \succeq F^{\mu}(\mathcal{S}')$. Conversely, assume that $\mathcal{S}$ does not absorb $\mathcal{S}'$. Then either there is a terminal t such that $\mathcal{S}$ admits more plays with outcome t than $\mathcal{S}'$ does, or $\mathcal{S}$ admits an infinite play, but $\mathcal{S}'$ does not. In the first case, $F^{\nu}(\mathcal{S}) \not\succeq F^{\nu}(\mathcal{S}')$ and in the second case $0 = F^{\mu}(\mathcal{S}) \not\succeq F^{\mu}(\mathcal{S}') \neq 0$. $\square$

Example 50. We return to the game described in Example 39



with equation system G_0 consisting of $X_v = s + X_w$ and $X_w = t \cdot X_v$. In $\mathbb{N}^{\infty}[\![s, t]\!]$ the least fixed-point solution is $f(v) = s \cdot (1 + t + t^2 \cdots)$ and $f(w) = s \cdot (t + t^2 + \cdots)$. In $\mathbb{S}^{\infty}[s, t]$ the least fixed-point solution $f^{\mu} = \mathrm{lfp}\, G_0$ has values $f^{\mu}(v) = s$ and $f^{\mu}(w) = st$, which describes the possible outcomes of the unique absorption-dominant strategy that enforces finite plays. The only other absorption-dominant strategy (moving from v to w) has value 0 because it admits an infinite play.

However, the greatest fixed-point solution $f^{\nu} = \mathrm{gfp}\, G_0$ of this equation system in $\mathbb{S}^{\infty}[s, t]$ has values $f^{\nu}(v) = s + t^{\infty}$ and $f^{\nu}(w) = st + t^{\infty}$. Here this second strategy has value t^{∞} since it admits infinitely many plays ending in t (and one infinite play with value 1).

Theorem 51. *Let $\mathcal{G} = (V, V_0, V_1, T, E)$ be a game graph and let G_{σ} be the associated equation system for Player σ. In the semiring $\mathbb{S}^{\infty}[T]$ this system has least and greatest fixed point solutions $\mathrm{lfp}\, G_{\sigma}$ and $\mathrm{gfp}\, G_{\sigma}$ with*

$$(\mathrm{lfp}\, G_{\sigma})(v) := \sum_{\mathcal{S} \in \mathrm{Strat}_{\sigma}(v)} F^{\mu}(\mathcal{S}) \quad \textit{and} \quad (\mathrm{gfp}\, G_{\sigma})(v) := \sum_{\mathcal{S} \in \mathrm{Strat}_{\sigma}(v)} F^{\nu}(\mathcal{S}).$$

The values of these sums do not change if we restrict them to the absorption-dominant strategies.

Proof. Since $\mathbb{S}^{\infty}[T]$ is ω-continuous, the claim for $(\mathrm{lfp}\, G_{\sigma})$ follows from Theorem 36. For the greatest fixed-point solution we use that $\mathbb{S}^{\infty}[T]$ is also ω-cocontinuous and has the structure of a complete lattice. Thus, $(\mathrm{gfp}\, G_{\sigma})$ is the limit of the descending chain $(G^n)_{n<\omega}$ of approximants starting with $G^0 = 1$, and G^{n+1} is defined by applying the equation system to $G^n : V \to \mathbb{S}^{\infty}[T]$.

As in the proof of Theorem 36 we argue with the unfoldings $\mathcal{G}^n$ of $\mathcal{G}$ up to $n-1$ moves, but we now put $f^n_\sigma(\pi v) = 1$ for the final node of an "unfinished" play, i.e., with $|\pi| = n-1$ and $v \in V \setminus T$. The valuations f^n_σ extend to all nodes of the (acyclic) game $\mathcal{G}^n$, and again, coincide with the Kleene approximants G^n: for every n and every v we have $G^n(v) = f^n_\sigma(v)$. The different valuation of the terminal nodes in $\mathcal{G}^n$ also has the effect that for any $\mathcal{T} \in \mathrm{Strat}^{(n)}_\sigma(v)$ we have $F(\mathcal{T}) = \prod_{t \in T} t^{\#_\mathcal{T}(t)}$, which is a monomial with only finite exponents. Since $\mathcal{G}^n$ is acyclic

$$f^n_\sigma(v) = \sum_{\mathcal{T} \in \mathrm{Strat}^{(n)}_\sigma(v)} F(\mathcal{T}).$$

Every strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$ for the original game $\mathcal{G}$ induces for each game $\mathcal{G}^n$ a strategy $\mathcal{S}^{(n)}$. Conversely, every strategy $\mathcal{T} \in \mathrm{Strat}^{(n)}_\sigma(v)$ is induced by at least one strategy $\mathcal{S} \in \mathrm{Strat}_\sigma(v)$. Since the semiring $\mathbb{S}^\infty[T]$ is idempotent, we have that, for every $n < \omega$,

$$\sum_{\mathcal{S} \in \mathrm{Strat}_\sigma(v)} F(\mathcal{S}^{(n)}) = \sum_{\mathcal{T} \in \mathrm{Strat}^{(n)}_\sigma(v)} F(\mathcal{T}).$$

As graphs, the sequence $(\mathcal{S}^{(n)})_{n \in \omega}$ of induced strategies is increasing, i.e., $\mathcal{S}^{(1)} \subseteq S^{(2)} \subseteq \cdots$, but their values in $\mathbb{S}^\infty[T]$ are decreasing, i.e., $F(\mathcal{S}^{(1)}) \succeq F(S^{(2)}) \subseteq \cdots$. Further, $F^\nu(\mathcal{S}) = \prod_{t \in T} t^{\#_S(t)}$ with exponents $\#_\mathcal{S}(t) \in \mathbb{N}^\infty$ and the corresponding exponents in the monomial $F(\mathcal{S}^{(n)})$ tell us how often a terminal position $t \in T$ has been reached by $\mathcal{S}$ after $n-1$ moves. In particular,

$$F^\nu(\mathcal{S}) = \lim_{n \to \infty} F(\mathcal{S}^{(n)}).$$

It is clear that these limits commute with summation over strategies, so we have

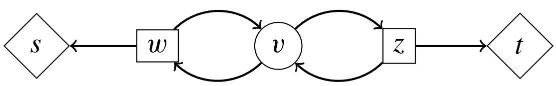
$$\sum_{\mathcal{S} \in \mathrm{Strat}_\sigma(v)} F^\nu(\mathcal{S}) = \lim_{n \to \infty} \sum_{\mathcal{S} \in \mathrm{Strat}_\sigma(v)} F(\mathcal{S}^{(n)}) = \lim_{n \to \infty} \sum_{\mathcal{T} \in \mathrm{Strat}^{(n)}_\sigma(v)} F(\mathcal{T}).$$

Putting everything together we get

$$(\mathrm{gfp}\, G_\sigma)(v) = \lim_{n \to \infty} G^n(v) = \lim_{n \to \infty} f^n_\sigma(v) = \lim_{n \to \infty} \sum_{\mathcal{T} \in \mathrm{Strat}^{(n)}} F(\mathcal{T}) = \sum_{\mathcal{S} \in \mathrm{Strat}_\sigma(v)} F^\nu(\mathcal{S}). \qquad \square$$

These least and greatest fixed points give precise descriptions of the absorption-dominant reachability and safety strategies of the players for each position of the game.

Example 52. We return to the Example 42:



Recall that the associated equation system for Player 0 has the equations $X_v = X_w + X_z$, $X_w = f(s) \cdot X_v$, and $X_z = f(t) \cdot X_v$.

The greatest fixed-point solution in $\mathbb{S}^\infty[s,t]$, computed by iterating from the top element $f = 1$ results in $f^\nu(v) = s^\infty + t^\infty$, $f^\nu(w) = s^\infty + st^\infty$", and $f^\nu(z) = s^\infty t + t^\infty$. Notice that indeed, $f^\nu(v) = f^\nu(w) + f^\nu(z)$ because st^∞ is absorbed by t^∞, and $s^\infty t$ by s^∞. The greatest fixed-point solution indicates that Player 0 has two absorptive strategies (move always to w or move always to z), and gives, for each of the terminal nodes s and t the number of plays ending in that node that the strategy admits. For instance, if the safety objective requires avoiding t, then v and w, the strategy moving to w has infinitely many winning plays ending in s (and one nonterminating play with value 1), but since $f(z)[s, 0] = 0$, Player 0 has no safety strategy from z that avoids t.

12. Outlook

In this paper we have extended the semiring framework for provenance analysis by new elements, so that it can be applied to logics with negation, in particular first-order logic and fixed-point logics, and to an analysis of games that provides detailed information about the number and properties of the strategies of the players.

Our treatment of negation is based on transformations to negation normal form and the use of newly introduced semirings of dual-indeterminate polynomials and dual-indeterminate power series. In particular, ω-continuous semirings $\mathbb{N}^\infty[\![X, \overline{X}]\!]$ of dual-indeterminate power series provide an adequate general framework for logics with least fixed points, such as posLFP (and Datalog) and the semiring of absorptive generalized dual-indeterminate polynomials $\mathbb{S}^\infty[\![X, \overline{X}]\!]$ permits an adequate treatment of greatest fixed points. We have thus laid foundations for a provenance analysis of general fixed-point logics, and we are currently applying this also to modal, temporal, and dynamic logics.

On the level of games, we have seen that provenance valuations in ω-continuous and absorptive semirings give us very detailed information about strategies for possibly infinite games with reachability and safety objectives. We are currently expanding this to games with more complicated objectives, such as Büchi, co-Büchi or parity games. Since these objectives no longer depend on terminal nodes but on the data occurring in infinite plays, a somewhat different framework has to be used, depending for instance on basic valuations of the edges of the game graph.

References

[Amsterdamer et al. 2011a] Y. Amsterdamer, D. Deutch, and V. Tannen, "On the limitations of provenance for queries with difference", presented at the 3rd Workshop on the Theory and Practice of Provenance, TaPP'11, 2011, available at https://www.usenix.org/legacy/events/tapp11/tech/final_files/Amsterdamer.pdf.

[Amsterdamer et al. 2011b] Y. Amsterdamer, D. Deutch, and V. Tannen, "Provenance for aggregate queries", pp. 153–164 in *PODS '11: proceedings of the 30th ACM SIGMOD-SIGACT-SIGART symposium on principles of database systems* (Athens, Greece), Association for Computing Machinery, New York, 2011.

[Apt and Grädel 2011] K. Apt and E. Grädel (editors), *Lectures in game theory for computer scientists*, Cambridge University Press, 2011. Zbl

[Dannert and Grädel 2019] K. M. Dannert and E. Grädel, "Provenance analysis: a perspective for description logics?", pp. 266–285 in *Description logic, theory combination, and all that*, edited by C. Lutz et al., Lecture Notes in Comput. Sci. **11560**, Springer, 2019. MR

[Dannert and Grädel 2020] K. Dannert and E. Grädel, "Semiring provenance for guarded logics", preprint, 2020, available at https://www.logic.rwth-aachen.de/pub/DannertGraedel.pdf. To appear in *Hajnal Andréka and István Németi on unity of science: from computing to relativity theory through algebraic logic*.

[Deutch et al. 2014] D. Deutch, T. Milo, S. Roy, and V. Tannen, "Circuits for datalog provenance", pp. 201–212 in *Database theory — ICDT 2014*, edited by N. Schweikardt et al., OpenProceedings, Konstanz, 2014. MR

[Foster et al. 2008] J. Foster, T. Green, and V. Tannen, "Annotated XML: queries and provenance", pp. 271–280 in *PODS '08: proceedings of the 27th ACM SIGMOD-SIGACT-SIGART symposium on principles of database systems* (Vancouver, BC, 2008), Association for Computing Machinery, New York, 2008.

[Geerts and Poggi 2010] F. Geerts and A. Poggi, "On database query languages for K-relations", *J. Appl. Log.* **8**:2 (2010), 173–185. MR Zbl

[Geerts et al. 2016] F. Geerts, T. Unger, G. Karvounarakis, I. Fundulaki, and V. Christophides, "Algebraic structures for capturing the provenance of SPARQL queries", *J. ACM* **63**:1 (2016), art. id. 7. MR Zbl

[Grädel and Tannen 2017] E. Grädel and V. Tannen, "Semiring provenance for first-order model checking", preprint, 2017. arXiv

[Grädel et al. 2007] 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*, Springer, 2007. MR

[Green 2011] T. J. Green, "Containment of conjunctive queries on annotated relations", *Theory Comput. Syst.* **49**:2 (2011), 429–459. MR Zbl

[Green and Tannen 2017] T. Green and V. Tannen, "The semiring framework for database provenance", pp. 93–99 in *PODS '17: Proceedings of the 36th ACM SIGMOD-SIGACT-SIGART symposium on principles of database systems* (Chicago, IL), Association for Computing Machinery, New York, 2017.

[Green et al. 2007] T. Green, G. Karvounarakis, and V. Tannen, "Provenance semirings", pp. 31–40 in *PODS '07: Proceedings of the 26th ACM SIGMOD-SIGACT-SIGART symposium on principles of database systems* (Beijing, China), Association for Computing Machinery, New York, 2007.

[Green et al. 2009] T. Green, G. Karvounarakis, and V. Tannen, "Reconcilable differences", pp. 212–224 in *ICDT '09: Proceedings of the 12th International Conference on Database Theory*, Association for Computing Machinery, New York, 2009.

[Kuich 1997] W. Kuich, "Semirings and formal power series: their relevance to formal languages and automata", pp. 609–677 in *Handbook of formal languages, Vol. 1*, edited by G. Rozenberg and A. Salomaa, Springer, 1997. MR

[Tannen 2013] V. Tannen, "Provenance propagation in complex queries", pp. 483–493 in *In search of elegance in the theory and practice of computation*, edited by V. Tannen et al., Lecture Notes in Comput. Sci. **8000**, Springer, 2013. MR Zbl

[Tannen 2017] V. Tannen, "Provenance analysis for FOL model checking", *ACM SIGLOG News* **4**:1 (2017), 24–36.

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

Received 2 Aug 2019.

ERICH GRÄDEL:

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

VAL TANNEN:

val@cis.upenn.edu
University of Pennsylvania, Philadelphia, PA, United States

Moscow Journal of Combinatorics and Number Theory

msp.org/moscow

See inside back cover or msp.org/moscow for submission instructions.

The subscription price for 2020 is US $310/year for the electronic version, and $365/year (+$20, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Moscow Journal of Combinatorics and Number Theory (ISSN 2640-7361 electronic, 2220-5438 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

MJCNT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Moscow Journal of Combinatorics and Number Theory

vol. 9 no. 3 2020

Thematic Issue on

Complexity