Semiring Semantics

Erich Grädel

(joint work with Sophie Brinke, Katrin Dannert, Lovro Mrkonjić, Matthias Naaf, and Val Tannen)

Simons Workshop on Logic and Algebra for Query Evaluation, Berkeley 2023

Semiring provenance

originates in the seminal work of Green, Karvounarakis, and Tannen 2007.

Idea: Annotate the facts of a database by values of a commutative semiring $(S, +, \cdot, 0, 1)$.

Semiring provenance

originates in the seminal work of Green, Karvounarakis, and Tannen 2007.

Idea: Annotate the facts of a database by values of a commutative semiring $(S,+,\cdot,0,1)$.

Propagate these annotations through a query, keeping track of whether pieces of information are used jointly or alternatively.

- $+$ interprets alternative use of information ($\vee$, $\exists$, unions)
- $\cdot$ interprets joint use of information ($\wedge$, $\forall$, joins)
- $0 \in S$ interprets false assertions and elements $s \neq 0$ provide annotations for true assertions.
- untracked information is interpreted by $1 \in S$.

This can give detailed insights about which combinations of facts are responsible for the truth of a statement and further information about confidence scores, cost analysis, number of evaluation strategies, access levels, ...

Semirings

$(S,+,\cdot,0,1)$ is a commutative semiring if $(S,+,0)$ and $(S,\cdot,1)$ are commutative monoids, such that $\cdot$ distributes over $+$, $0 \neq 1$, and $0 \cdot a = 0$.

Semirings

$(S,+,\cdot,0,1)$ is a commutative semiring if $(S,+,0)$ and $(S,\cdot,1)$ are commutative monoids, such that $\cdot$ distributes over $+$, $0 \neq 1$, and $0 \cdot a = 0$.

- The Boolean semiring $\mathbb{B} = (\{0,1\},\vee,\wedge,0,1)$ is the standard habitat of logic.

Semirings

$(S,+,\cdot,0,1)$ is a commutative semiring if $(S,+,0)$ and $(S,\cdot,1)$ are commutative monoids, such that $\cdot$ distributes over $+$, $0 \neq 1$, and $0 \cdot a = 0$.

- The Boolean semiring $\mathbb{B} = (\{0,1\},\vee,\wedge,0,1)$ is the standard habitat of logic.

- The natural semiring $\mathbb{N} = (\mathbb{N},+,\cdot,0,1)$ for counting proofs and strategies.

Semirings

$(S,+,\cdot,0,1)$ is a commutative semiring if $(S,+,0)$ and $(S,\cdot,1)$ are commutative monoids, such that $\cdot$ distributes over $+$, $0\neq 1$, and $0\cdot a=0$.

- The Boolean semiring $\mathbb{B}=(\{0,1\},\vee,\wedge,0,1)$ is the standard habitat of logic.

- The natural semiring $\mathbb{N}=(\mathbb{N},+,\cdot,0,1)$ for counting proofs and strategies.

- The tropical semiring $\mathbb{T}=(\mathbb{R}_+^\infty,\min,+,\infty,0)$ for cost interpretations.

Semirings

$(S,+,\cdot,0,1)$ is a commutative semiring if $(S,+,0)$ and $(S,\cdot,1)$ are commutative monoids, such that $\cdot$ distributes over $+$, $0 \neq 1$, and $0 \cdot a = 0$.

- The Boolean semiring $\mathbb{B} = (\{0,1\},\vee,\wedge,0,1)$ is the standard habitat of logic.

- The natural semiring $\mathbb{N} = (\mathbb{N},+,\cdot,0,1)$ for counting proofs and strategies.

- The tropical semiring $\mathbb{T} = (\mathbb{R}_+^\infty,\min,+,\infty,0)$ for cost interpretations.

- The Viterbi semiring $\mathbb{V} = ([0,1],\max,\cdot,0,1)$ for confidence scores.

Semirings

$(S,+,\cdot,0,1)$ is a commutative semiring if $(S,+,0)$ and $(S,\cdot,1)$ are commutative monoids, such that $\cdot$ distributes over $+$, $0 \neq 1$, and $0 \cdot a = 0$.

- The Boolean semiring $\mathbb{B} = (\{0,1\},\vee,\wedge,0,1)$ is the standard habitat of logic.

- The natural semiring $\mathbb{N} = (\mathbb{N},+,\cdot,0,1)$ for counting proofs and strategies.

- The tropical semiring $\mathbb{T} = (\mathbb{R}_+^\infty,\min,+,\infty,0)$ for cost interpretations.

- The Viterbi semiring $\mathbb{V} = ([0,1],\max,\cdot,0,1)$ for confidence scores.

- Min-max semirings $(A,\max,\min,0,1)$, induced by a total order $(A,<)$.
  A particular example is the security semiring induced by $\mathbb{A} = \{0 < \mathsf{T} < \mathsf{S} < \mathsf{C} < \mathsf{P} = 1\}$
  where P is "public", C is "confidential", S is "secret", T is "top secret".

Semirings

- Lattice semirings $(A, \sqcup, \sqcap, 0, 1)$, induced by a partial order.

Semirings

- Lattice semirings $(A, \sqcup, \sqcap, 0, 1)$, induced by a partial order.

- The Łukasiewicz semiring $\mathbb{L} = ([0,1], \max, \otimes, 0, 1)$ with $a \otimes b := \max(a+b-1, 0)$ is popular in the study of many-valued logics, and gives a different notion of confidence or degrees of truth.

Semirings

- Lattice semirings $(A, \sqcup, \sqcap, 0, 1)$, induced by a partial order.

- The Łukasiewicz semiring $\mathbb{L} = ([0,1], \max, \otimes, 0, 1)$ with $a \otimes b := \max(a+b-1, 0)$ is popular in the study of many-valued logics, and gives a different notion of confidence or degrees of truth.

The Łukasiewicz semiring is isomorphic to

- the semiring of doubt $\mathbb{D} = ([0,1], \min, \oplus, 1, 0)$ with $a \oplus b := \min(a+b, 1)$.

Semirings

- Lattice semirings $(A, \sqcup, \sqcap, 0, 1)$, induced by a partial order.

- The Łukasiewicz semiring $\mathbb{L} = ([0,1], \max, \otimes, 0, 1)$ with $a \otimes b := \max(a+b-1, 0)$ is popular in the study of many-valued logics, and gives a different notion of confidence or degrees of truth.

The Łukasiewicz semiring is isomorphic to

- the semiring of doubt $\mathbb{D} = ([0,1], \min, \oplus, 1, 0)$ with $a \oplus b := \min(a+b, 1)$.

We are only interested in commutative semirings that are naturally ordered by addition:
$a \leq b \;:\Longleftrightarrow\; \exists c(a + c = b)$ is antisymmetric, and therefore a partial order.

In particular, this excludes rings.

Provenance semirings: tracking atomic facts

Fundamental question: Which combinations of atomic facts are responsible for the truth of a statement, and how often is a fact used in the evaluation ?

Provenance semirings: tracking atomic facts

Fundamental question: Which combinations of atomic facts are responsible for the truth of a statement, and how often is a fact used in the evaluation ?

Let X be a set of indeterminates, which are used to label the facts that we want to track: $\alpha \mapsto X_\alpha$ (untracked facts are mapped to 0 or 1).

$\mathbb{N}[X]$: semiring of multivariate polynomials in X with coefficients from $\mathbb{N}$.

This is the commutative semiring freely generated by the set X.

Provenance semirings: tracking atomic facts

Fundamental question: Which combinations of atomic facts are responsible for the truth of a statement, and how often is a fact used in the evaluation ?

Let X be a set of indeterminates, which are used to label the facts that we want to track: $\alpha \mapsto X_\alpha$ (untracked facts are mapped to 0 or 1).

$\mathbb{N}[X]$: semiring of multivariate polynomials in X with coefficients from $\mathbb{N}$.

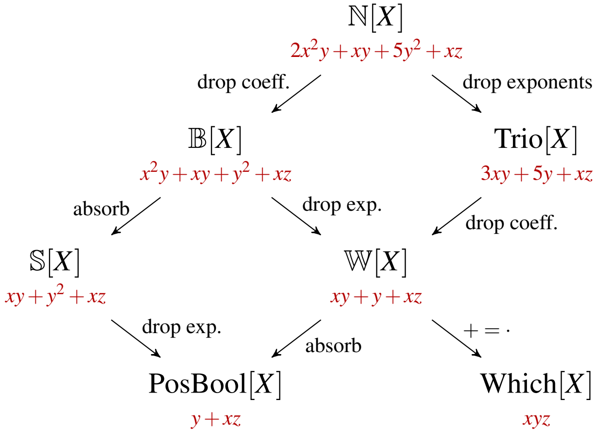
This is the commutative semiring freely generated by the set X.

Universality: Any function $f : X \to S$ into an arbitrary semiring S extends uniquely to a semiring homomorphism $h : \mathbb{N}[X] \to S$.

Think of $h(2x^2 + xy + 3z^2)$ as evaluating $2x^2 + xy + 3z^2$ in S.

Other provenance semirings

Simpler and “less informative” semirings with specific algebraic properties:



A stumbling block for provenance analysis

Semiring provenance has been successful in database theory. But for a long time, it has (essentially) been confined to positive query languages such as positive relational algebra RA^+, (unions of) conjunctive queries, or Datalog.

A stumbling block for provenance analysis

Semiring provenance has been successful in database theory. But for a long time, it has (essentially) been confined to positive query languages such as positive relational algebra RA^+, (unions of) conjunctive queries, or Datalog.

Val Tannen : “Divergent approaches and unsatisfactory state of affairs for queries with negation or difference of relations."

The absence of an adequate treatment of negation has been the main stumbling block against extensions to richer languages and transfers to other fields.

A stumbling block for provenance analysis

Semiring provenance has been successful in database theory. But for a long time, it has (essentially) been confined to positive query languages such as positive relational algebra RA^+, (unions of) conjunctive queries, or Datalog.

Val Tannen : “Divergent approaches and unsatisfactory state of affairs for queries with negation or difference of relations."

The absence of an adequate treatment of negation has been the main stumbling block against extensions to richer languages and transfers to other fields.

Problems: Negation is incompatible with the algebraic semiring operations.

A stumbling block for provenance analysis

Semiring provenance has been successful in database theory. But for a long time, it has (essentially) been confined to positive query languages such as positive relational algebra RA^+, (unions of) conjunctive queries, or Datalog.

Val Tannen : “Divergent approaches and unsatisfactory state of affairs for queries with negation or difference of relations."

The absence of an adequate treatment of negation has been the main stumbling block against extensions to richer languages and transfers to other fields.

Problems: Negation is incompatible with the algebraic semiring operations.

In many scenarios (and especially for provenance), negation is not really a “logical” operation: The meaning of $\neg\varphi$ cannot be derived from the meaning of φ, but depends also on the syntax of φ.

Provenance for logics with negation: a new approach

With Val Tannen, we have proposed a new approach to generalize provenance analysis to full first-order logic and beyond, based on the following ideas:

- Negation is handled via transformation to negation normal form.

Provenance for logics with negation: a new approach

With Val Tannen, we have proposed a new approach to generalize provenance analysis to full first-order logic and beyond, based on the following ideas:

- Negation is handled via transformation to negation normal form.
- New semirings of polynomials with dual indeterminates $\mathbb{N}[X,\overline{X}] := \mathbb{N}[X \cup \overline{X}]/(X\overline{X})$ based on a bijection $X \leftrightarrow \overline{X}$.

Provenance for logics with negation: a new approach

With Val Tannen, we have proposed a new approach to generalize provenance analysis to full first-order logic and beyond, based on the following ideas:

- Negation is handled via transformation to negation normal form.
- New semirings of polynomials with dual indeterminates $\mathbb{N}[X,\overline{X}] := \mathbb{N}[X \cup \overline{X}]/(X\overline{X})$ based on a bijection $X \leftrightarrow \overline{X}$.
- Provenance for logic is intimately connected to provenance analysis for games. Negation is related to the antagonism between the two players in a model checking game.

Provenance for logics with negation: a new approach

With Val Tannen, we have proposed a new approach to generalize provenance analysis to full first-order logic and beyond, based on the following ideas:

- Negation is handled via transformation to negation normal form.
- New semirings of polynomials with dual indeterminates $\mathbb{N}[X,\overline{X}] := \mathbb{N}[X \cup \overline{X}]/(X\overline{X})$ based on a bijection $X \leftrightarrow \overline{X}$.
- Provenance for logic is intimately connected to provenance analysis for games. Negation is related to the antagonism between the two players in a model checking game.
- Provenance for games is of independent interest, and provides interesting insights into the number and properties of strategies for the players, far beyond the question who wins.

Provenance for logics with negation: a new approach

With Val Tannen, we have proposed a new approach to generalize provenance analysis to full first-order logic and beyond, based on the following ideas:

- Negation is handled via transformation to negation normal form.
- New semirings of polynomials with dual indeterminates $\mathbb{N}[X,\overline{X}] := \mathbb{N}[X \cup \overline{X}]/(X\overline{X})$ based on a bijection $X \leftrightarrow \overline{X}$.
- Provenance for logic is intimately connected to provenance analysis for games. Negation is related to the antagonism between the two players in a model checking game.
- Provenance for games is of independent interest, and provides interesting insights into the number and properties of strategies for the players, far beyond the question who wins.
- New kinds of applications: Missing answers, repairs, etc.

Negation normal form

A first-order formula is in negation normal form if negation is applied to atomic formulae only.

For a relational vocabulary $\tau = \{R_1, \dots, R_m\}$, formulae in $\mathrm{FO}(\tau)$ in negation normal form are defined by

$$\varphi ::= \quad x = y \mid x \neq y \mid R_i\bar{x} \mid \neg R_i\bar{x} \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \exists x \varphi \mid \forall x \varphi$$

and arbitrary formulae φ can efficiently be translated into their negation normal form $\mathrm{nnf}(\varphi)$.

Negation normal form

A first-order formula is in negation normal form if negation is applied to atomic formulae only.

For a relational vocabulary $\tau = \{R_1, \dots, R_m\}$, formulae in $\mathrm{FO}(\tau)$ in negation normal form are defined by

$$\varphi ::= \quad x = y \mid x \neq y \mid R_i\bar{x} \mid \neg R_i\bar{x} \mid \varphi \vee \varphi \mid \varphi \wedge \varphi \mid \exists x\varphi \mid \forall x\varphi$$

and arbitrary formulae φ can efficiently be translated into their negation normal form $\mathrm{nnf}(\varphi)$.

The evaluation of a formula in negation normal form on a structure $\mathfrak{A}$ is a positive process on the basis of the positive and negative atomic facts.

Semiring interpretations

Fix a commutative semiring S.

Let A be a finite universe and $\tau = \{R_1, \dots, R_m\}$ be a finite relational vocabulary.

$\mathrm{Lit}_A(\tau)$: all fully instantiated literals $R\overline{a}$ and $\neg R\overline{a}$ with $R \in \tau$ and $\overline{a} \in A^k$.

Semiring interpretations

Fix a commutative semiring S.

Let A be a finite universe and $\tau = \{R_1, \dots, R_m\}$ be a finite relational vocabulary.

$\mathrm{Lit}_A(\tau)$: all fully instantiated literals $R\overline{a}$ and $\neg R\overline{a}$ with $R \in \tau$ and $\overline{a} \in A^k$.

A S-interpretation for A and τ is a function $\pi : \mathrm{Lit}_A(\tau) \to S$.
Further, let π map equalities $a = b$ and $a \neq b$ to their truth values 0 or 1.

Semiring interpretations

Fix a commutative semiring S.

Let A be a finite universe and $\tau = \{R_1, \dots, R_m\}$ be a finite relational vocabulary.

$\text{Lit}_A(\tau)$: all fully instantiated literals $R\overline{a}$ and $\neg R\overline{a}$ with $R \in \tau$ and $\overline{a} \in A^k$.

A S-interpretation for A and τ is a function $\pi : \text{Lit}_A(\tau) \to S$.
Further, let π map equalities $a = b$ and $a \neq b$ to their truth values 0 or 1.

We call $\pi : \text{Lit}_A(\tau) \to S$ model-defining if, for all atoms $R\overline{a}$, precisely one of the values $\pi(R\overline{a})$ and $\pi(\neg R\overline{a})$ is zero. Then π specifies a unique structure $\mathfrak{A}_\pi$.

Semiring semantics for first-order logic

We can extend any S-interpretation $\pi : \mathrm{Lit}_A(\tau) \to S$ to a S-valuation $\pi : \mathrm{FO}(\tau) \to S$ giving values $\pi[\![\varphi]\!] \in S$ to all $\varphi \in \mathrm{FO}(\tau)$.

$$\pi[\![\varphi \vee \psi]\!] := \pi[\![\varphi]\!] + \pi[\![\psi]\!] \qquad \pi[\![\varphi \wedge \psi]\!] := \pi[\![\varphi]\!] \cdot \pi[\![\psi]\!]$$

$$\pi[\![\exists x \varphi(x)]\!] := \sum_{a\in A} \pi[\![\varphi(a)]\!] \qquad \pi[\![\forall x \varphi(x)]\!] := \prod_{a\in A} \pi[\![\varphi(a)]\!]$$

$$\pi[\![\neg\varphi]\!] := \pi[\![\mathrm{nnf}(\neg\varphi)]\!].$$

Semiring semantics for first-order logic

We can extend any S-interpretation $\pi : \mathrm{Lit}_A(\tau) \to S$ to a S-valuation $\pi : \mathrm{FO}(\tau) \to S$ giving values $\pi[\![\varphi]\!] \in S$ to all $\varphi \in \mathrm{FO}(\tau)$.

$$\pi[\![\varphi \vee \psi]\!] := \pi[\![\varphi]\!] + \pi[\![\psi]\!] \qquad \pi[\![\varphi \wedge \psi]\!] := \pi[\![\varphi]\!] \cdot \pi[\![\psi]\!]$$

$$\pi[\![\exists x \varphi(x)]\!] := \textstyle\sum_{a \in A} \pi[\![\varphi(a)]\!] \qquad \pi[\![\forall x \varphi(x)]\!] := \textstyle\prod_{a \in A} \pi[\![\varphi(a)]\!]$$

$$\pi[\![\neg\varphi]\!] := \pi[\![\mathrm{nnf}(\neg\varphi)]\!].$$

Remark. For semiring semantics of FO on infinite universes, we need to extend the semiring operations + and · to infinitary sum and product operators.

Semirings of dual-indeterminate polynomials

Annotate atoms by indeterminates in X, and negated atoms by indeteminates in $\overline{X}$, with a bijection $X \leftrightarrow \overline{X}$ mapping $x \in X$ to its complementary token $\overline{x} \in \overline{X}$.

Semirings of dual-indeterminate polynomials

Annotate atoms by indeterminates in X, and negated atoms by indeteminates in $\overline{X}$, with a bijection $X \leftrightarrow \overline{X}$ mapping $x \in X$ to its complementary token $\overline{x} \in \overline{X}$.

$\mathbb{N}[X, \overline{X}] := \mathbb{N}[X \cup \overline{X}]/(X\overline{X})$ is the quotient semiring of $\mathbb{N}[X \cup \overline{X}]$ by the congruence generated by the equations $x \cdot \overline{x} = 0$. Corresponds to polynomials in $\mathbb{N}[X \cup \overline{X}]$ such that no monomial contains complementary tokens.

Semirings of dual-indeterminate polynomials

Annotate atoms by indeterminates in X, and negated atoms by indeteminates in $\overline{X}$, with a bijection $X \leftrightarrow \overline{X}$ mapping $x \in X$ to its complementary token $\overline{x} \in \overline{X}$.

$\mathbb{N}[X,\overline{X}] := \mathbb{N}[X \cup \overline{X}]/(X\overline{X})$ is the quotient semiring of $\mathbb{N}[X \cup \overline{X}]$ by the congruence generated by the equations $x \cdot \overline{x} = 0$. Corresponds to polynomials in $\mathbb{N}[X \cup \overline{X}]$ such that no monomial contains complementary tokens.

Universality. Any map $h : (X \cup \overline{X}) \to S$ into a semiring S, with $h(x) \cdot h(\overline{x}) = 0$ for $x \in X$, extends uniquely to a semiring homomorphism $h : \mathbb{N}[X,\overline{X}] \to S$.

Proof trees and evaluation strategies

An evaluation tree for a sentence $\psi \in \text{FO}$ and a semiring interpretation $\pi : \text{Lit}_A(\tau) \to S$ is the same thing as a strategy in the associated evaluation game.

Let $\#_\alpha(T)$ denote the number of leaves of the tree T labelled by the literal α.

Valuation of T:

$$\pi[\![T]\!] := \prod_{\alpha \in \text{Lit}_A(\tau)} \pi(\alpha)^{\#_\alpha(T)}.$$

A proof tree for $\psi \in \text{FO}$ and $\pi : \text{Lit}_A(\tau) \to S$ is an evaluation tree with $\pi(T) \neq 0$

Theorem. For every semiring interpretation $\pi : \text{Lit}_A(\tau) \to S$ and every $\psi \in \text{FO}$

$$\pi[\![\psi]\!] = \sum \left\{ \pi[\![T]\!] \ : \ T \text{ is a proof tree for } \psi \text{ and } \pi \right\}$$

Proof trees and dual-indeterminate polynomials

Consider a model-defining semiring interpretation $\pi : \mathrm{Lit}_A(\tau) \to \mathbb{N}[X, \overline{X}]$ that maps each literal to either an indeterminate in $X \cup \overline{X}$ or to a truth value 0 or 1 ,

Proof trees and dual-indeterminate polynomials

Consider a model-defining semiring interpretation $\pi : \mathrm{Lit}_A(\tau) \to \mathbb{N}[X, \overline{X}]$ that maps each literal to either an indeterminate in $X \cup \overline{X}$ or to a truth value 0 or 1 ,

What does the provenance polynomial $\pi[\![\psi]\!]$ tell us about the model-checking problem $\mathfrak{A}_\pi \models \psi$?

Proof trees and dual-indeterminate polynomials

Consider a model-defining semiring interpretation $\pi : \mathrm{Lit}_A(\tau) \to \mathbb{N}[X,\overline{X}]$ that maps each literal to either an indeterminate in $X \cup \overline{X}$ or to a truth value 0 or 1 ,

What does the provenance polynomial $\pi[\![\psi]\!]$ tell us about the model-checking problem $\mathfrak{A}_\pi \models \psi$?

$$\pi[\![\psi]\!] = \sum \Big\{ \pi[\![T]\!] \ : \ T \text{ is a proof tree for } \psi \text{ and } \pi \Big\}$$

is a sum of monomials $mx_1^{e_1} \cdots x_k^{e_k}$. Each such monomial tells us that there are precisely m proof trees establishing that $\mathfrak{A}_\pi \models \psi$ which

- use among the tracked literals only those labelled by $x_1, \ldots, x_k$,
- use literals labelled by x_i precisely e_i times,
- and may use true untracked true literals (that have value 1) arbitrarily.

Proof trees and dual-indeterminate polynomials

Consider a model-defining semiring interpretation $\pi : \mathrm{Lit}_A(\tau) \to \mathbb{N}[X, \overline{X}]$ that maps each literal to either an indeterminate in $X \cup \overline{X}$ or to a truth value 0 or 1 ,

What does the provenance polynomial $\pi[\![\psi]\!]$ tell us about the model-checking problem $\mathfrak{A}_\pi \models \psi$?

$$\pi[\![\psi]\!] = \sum \left\{ \pi[\![T]\!] \ : \ T \text{ is a proof tree for } \psi \text{ and } \pi \right\}$$

is a sum of monomials $mx_1^{e_1} \cdots x_k^{e_k}$. Each such monomial tells us that there are precisely m proof trees establishing that $\mathfrak{A}_\pi \models \psi$ which
- use among the tracked literals only those labelled by $x_1, \dots, x_k$,
- use literals labelled by x_i precisely e_i times,
- and may use true untracked true literals (that have value 1) arbitrarily.

In particular $\mathfrak{A}_\pi \models \psi$ if, and only if, $\pi[\![\psi]\!] \neq 0$.

Provenance information for classes of structures

Model-compatible interpretations $\pi : \mathrm{Lit}_A(\tau) \to \mathbb{N}[X, \overline{X}]$. For every atom $R\overline{a}$, either

(1) $\pi(R\overline{a}) = x$ and $\pi(\neg R\overline{a}) = \overline{x}$, for some $x \in X$, or

(2) $\pi(R\overline{a}) = 1$ and $\pi(\neg R\overline{a}) = 0$, or vice versa.

Provenance information for classes of structures

Model-compatible interpretations $\pi : \mathrm{Lit}_A(\tau) \to \mathbb{N}[X, \overline{X}]$. For every atom $R\overline{a}$, either

(1) $\pi(R\overline{a}) = x$ and $\pi(\neg R\overline{a}) = \overline{x}$, for some $x \in X$, or

(2) $\pi(R\overline{a}) = 1$ and $\pi(\neg R\overline{a}) = 0$, or vice versa.

A model-compatible interpretation is consistent with at least one τ-structure on A, but in general with a larger set of such structures.

$\mathrm{Must}_\pi := \{\varphi \in \mathrm{Lit}_A(\varphi) : \pi(\varphi) = 1\}$ (true in all models of π)

$\mathrm{May}_\pi := \{\varphi \in \mathrm{Lit}_A(\varphi) : \pi(\varphi) \in X \cup \overline{X}\}$ (true in some models of π)

Provenance information for classes of structures

Model-compatible interpretations $\pi : \mathrm{Lit}_A(\tau) \to \mathbb{N}[X,\overline{X}]$. For every atom $R\overline{a}$, either
(1) $\pi(R\overline{a}) = x$ and $\pi(\neg R\overline{a}) = \overline{x}$, for some $x \in X$, or
(2) $\pi(R\overline{a}) = 1$ and $\pi(\neg R\overline{a}) = 0$, or vice versa.

A model-compatible interpretation is consistent with at least one τ-structure on A, but in general with a larger set of such structures.

$\mathrm{Must}_\pi := \{\varphi \in \mathrm{Lit}_A(\varphi) : \pi(\varphi) = 1\}$ (true in all models of π)
$\mathrm{May}_\pi := \{\varphi \in \mathrm{Lit}_A(\varphi) : \pi(\varphi) \in X \cup \overline{X}\}$ (true in some models of π)

Conclusion. For the resulting valuation $\pi : \mathrm{FO}(\tau) \to \mathbb{N}[X,\overline{X}]$, the provenance polynomial $\pi[\![\psi]\!]$ describes all proof trees for ψ whose leaves are in $\mathrm{Must}_\pi \cup \mathrm{May}_\pi$. Every monomial corresponds to one proof tree, and gives precise information about the literals on which the proof tree depends, giving a complete description of all models of ψ that are compatible with π.

Research programmes for semiring semantics

Beyond first-order logic

Semiring semantics for other logical formalisms, in particular for fixed-point logics such as LFP and the modal μ-calculus

Strategy analysis of games

Basic “first-order” semiring valuations provide a strategy analysis for acyclic reachability games. With ω-continuous semirings, this extends to reachability games with cycles. What about more complicated games, such as Büchi games, parity games, games with incomplete information ... ?

The model theory of semiring semantics

To what extent do classical results of logic generalise to semiring semantics, and how does this depend on the algebraic properties of the underlying semirings ?

Fixed-point logics

Semiring semantics has meanwhile be extended beyond FO to many other logics. The most interesting challenges are provided by fixed-point logics such as LFP and the modal μ-calculus L_μ.

Fixed-point logics

Semiring semantics has meanwhile be extended beyond FO to many other logics. The most interesting challenges are provided by fixed-point logics such as LFP and the modal μ-calculus L_μ.

Semiring provenance for Datalog had already been done in papers by Green, Karvounarakis, Tannen 2007 and Deutch, Milo, Roy, Tannen 2014, based on ω-continuous semirings. The universal one are semirings $\mathbb{N}^\infty[[X]]$ of formal power series.

Fixed-point logics

Semiring semantics has meanwhile be extended beyond FO to many other logics. The most interesting challenges are provided by fixed-point logics such as LFP and the modal μ-calculus L_μ.

Semiring provenance for Datalog had already been done in papers by Green, Karvounarakis, Tannen 2007 and Deutch, Milo, Roy, Tannen 2014, based on ω-continuous semirings. The universal one are semirings $\mathbb{N}^\infty[\![X]\!]$ of formal power series.

With dual indeterminates, this leads to semirings $\mathbb{N}^\infty[\![X,\overline{X}]\!]$ which provide semiring semantics for semipositive Datalog and the positive fragment posLFP of fixed-point logic.

However the general fixed-point logics LFP and L_μ may have arbitrary interleavings of least and greatest fixed points, and ω-continuous semirings are not adequate for these.

Semiring semantics for fixed-point logic

What are the algebraic conditions required for semirings for fixed-point logics?

Full continuity: each chain $C \subseteq S$ has a supremum $\bigsqcup C$ and an infimum $\bigsqcap C$ in S, with $a + \bigsqcup C = \bigsqcup(a + C)$, $a \cdot \bigsqcup C = \bigsqcup(a \cdot C)$ and analogously for $\bigsqcap C$.

Fully continuous semirings suffice to get a well-defined semantics for LFP, but for a meaningful semantics that provides insights why a formula holds, an additional condition is necessary.

Absorption: $a + ab = a$ for all $a, b \in S$. This makes multiplication decreasing: $a \cdot b \leq a$ and $a \leq 1$.

Theorem. (Dannert-G.-Naaf-Tannen 2021)
In absorptive, fully chain-continuous semirings S, each monotone function $f : S \to S$ has a least fixed point $\mathbf{lfp}(f)$ and a greatest fixed point $\mathbf{gfp}(f)$. Together with the semiring semantics for FO, this provides meaningful semiring semantics for LFP.

Semirings for LFP

Many common application semirings are fully continuous and absorptive such as the tropical semiring, min-max semirings, the Lukasiewicz semiring. However, the general provenance semirings $\mathbb{N}[X]$ and $\mathbb{N}^\infty[[X]]$ are neither fully continuous nor absorptive.

Semirings for LFP

Many common application semirings are fully continuous and absorptive such as the tropical semiring, min-max semirings, the Lukasiewicz semiring. However, the general provenance semirings $\mathbb{N}[X]$ and $\mathbb{N}^\infty[\![X]\!]$ are neither fully continuous nor absorptive.

Instead, the general semirings for LFP are the semirings $\mathbb{S}^\infty[X]$ of generalized absorptive polynomials

$$f = x^2y^3z + x^\infty y + z^\infty$$

- no coefficients
- exponents in $\mathbb{N} \cup \{\infty\}$.
- absorption among monomials (those with larger exponents are absorbed).

Semirings $\mathbb{S}^\infty[X]$ and $\mathbb{S}^\infty[X, \overline{X}]$ have universality properties that make them the “right” general semirings for fixed-point logics. (Dannert, G., Naaf, Tannen, CSL 21)

Strategy analysis

With fully continuous and absorptive semirings, we can also provide a strategy analysis of more complicated games than reachability games, in particular Büchi games and parity games.

The core of such an analysis are Sum-of-Strategies Theorems. For a game $\mathcal{G}$ with positions in V and moves in E, we need

- An appropriate semiring S, for instance $S = \mathbb{S}^\infty[X]$ based of a labelling $\pi : e \mapsto x_e$ of moves in the game by indeterminates in X.
- A valuation $F : V \to S$ of the game positions. For instance $F(v) := \pi[\![\text{win}(v)]\!]$ where $\text{win}(x) \in \text{LFP}$ expresses that there is a winning strategy from x.
- A class of strategies, with a valuation $F(\mathscr{S}) \in S$ for each strategy $\mathscr{S}$. For instance

$$F(\mathscr{S}) := \sum_{e \in E} x_e^{\#_e(\mathscr{S})}.$$

Case study: Büchi games

Büchi games: Player 0 wins if she manages to hit some good position infinitely often. Winning positions in Büchi games are definable by a formula $\text{win}(x) \in \text{LFP}$ that requires an alternation between a greatest and a least fixed point.

Case study: Büchi games

Büchi games: Player 0 wins if she manages to hit some good position infinitely often. Winning positions in Büchi games are definable by a formula $\mathrm{win}(x) \in \mathrm{LFP}$ that requires an alternation between a greatest and a least fixed point.

The appropriate winning strategies are the absorption-dominant ones: A somewhat larger class than the positional strategies, minimizing the multiset of moves that are played.

The Sum-of-Strategies Theorem. Let $\mathcal{G}$ be a Büchi game and $\pi : \mathcal{G} \to S$ be an edge-tracking interpretation into an absorptive, fully continuous semiring. Then,

$$\pi[\![\mathrm{win}(v)]\!] = \sum \left\{ \pi[\![\mathscr{S}]\!] \quad : \quad \begin{array}{l} \mathscr{S} \text{ is an absorption-dominant} \\ \quad \text{winning strategy from } v \end{array} \right\}$$

Strategy analysis

From the polynomial $\pi[\![\mathrm{win}(v)]\!] \in \mathbb{S}^\infty[X]$ we can derive:

(1) whether Player 0 wins from v: this holds if $\pi[\![\mathrm{win}(v)]\!] \neq 0$,

(2) edge profiles of all absorption-dominant winning strategies from v,

(3) the number and shapes of all positional winning strategies from v,

(4) whether Player 0 can still win if a subset $X \subseteq E$ is forbidden.

Game repairs

Assume that Player 0 loses $\mathcal{G}$ from v.

What are minimal modifications to $\mathcal{G}$ that make Player 0 win?

This can be determined by a different semiring valuation $\pi : \mathcal{G} \to \mathbb{S}^\infty[X, \overline{X}]$ taking into account sets of edges we are allowed to delete or add.

The approach is not limited to tracking edges.

For instance, we can also track winning conditions:
how to choose or modify the target set so that Player 0 wins?

The model theory of semiring semantics

To what extent do classical results of logic generalise to semiring semantics?

The model theory of semiring semantics

To what extent do classical results of logic generalise to semiring semantics?

- Elementary equivalence versus isomorphism. For finite structures, $\mathfrak{A} \equiv \mathfrak{B} \Longleftrightarrow \mathfrak{A} \cong \mathfrak{B}$.
  Every finite structure can be axiomatised, up to isomorphism, by a first-order sentence.

The model theory of semiring semantics

To what extent do classical results of logic generalise to semiring semantics?

- Elementary equivalence versus isomorphism. For finite structures, $\mathfrak{A} \equiv \mathfrak{B} \iff \mathfrak{A} \cong \mathfrak{B}$. Every finite structure can be axiomatised, up to isomorphism, by a first-order sentence.
- 0-1 laws. Every first-order sentence is either almost surely true or almost surely false on random finite structures.

The model theory of semiring semantics

To what extent do classical results of logic generalise to semiring semantics?

- Elementary equivalence versus isomorphism. For finite structures, $\mathfrak{A} \equiv \mathfrak{B} \Longleftrightarrow \mathfrak{A} \cong \mathfrak{B}$. Every finite structure can be axiomatised, up to isomorphism, by a first-order sentence.
- 0-1 laws. Every first-order sentence is either almost surely true or almost surely false on random finite structures.
- Locality. By Theorems of Hanf and Gaifman, first-order formulae can only express local properties. In fact, every first-order formula is equivalent to one in Gaifman normal form.

The model theory of semiring semantics

To what extent do classical results of logic generalise to semiring semantics?

- Elementary equivalence versus isomorphism. For finite structures, $\mathfrak{A} \equiv \mathfrak{B} \Longleftrightarrow \mathfrak{A} \cong \mathfrak{B}$. Every finite structure can be axiomatised, up to isomorphism, by a first-order sentence.
- 0-1 laws. Every first-order sentence is either almost surely true or almost surely false on random finite structures.
- Locality. By Theorems of Hanf and Gaifman, first-order formulae can only express local properties. In fact, every first-order formula is equivalent to one in Gaifman normal form.
- Ehrenfeucht-Fraïssé games provide a sound and complete method for establishing logical equivalences.

All the definitions involved in these results generalise in a straightforward way to semiring semantics. But what about the results themselves, and the associated methods?

Elementary equivalence versus isomorphism

Both notions naturally generalize to semiring interpretations $\pi : \mathrm{Lit}_A(\tau) \to S$

$\pi_A \equiv \pi_B$ if $\pi_A[\![\varphi]\!] = \pi_B[\![\varphi]\!]$ for all $\varphi \in \mathrm{FO}$

$\pi_A \cong \pi_B$ if ....

In Boolean semantics, for finite structures, we have that $\mathfrak{A} \equiv \mathfrak{B} \iff \mathfrak{A} \cong \mathfrak{B}$.

This fails in semiring semantics, for some semrings.

Theorem (G., Mrkonjic, 2021) There exist finite S-interpretations $\pi_A \not\cong \pi_B$ (for instance in min-max semirings with ≥ 3 elements) such that $\pi_A \equiv \pi_B$.

Elementary equivalence versus isomorphism

Both notions naturally generalize to semiring interpretations $\pi : \mathrm{Lit}_A(\tau) \to S$

$\pi_A \equiv \pi_B$ if $\pi_A[\![\varphi]\!] = \pi_B[\![\varphi]\!]$ for all $\varphi \in \mathrm{FO}$

$\pi_A \cong \pi_B$ if ....

In Boolean semantics, for finite structures, we have that $\mathfrak{A} \equiv \mathfrak{B} \iff \mathfrak{A} \cong \mathfrak{B}$.

This fails in semiring semantics, for some semrings.

Theorem (G., Mrkonjic, 2021) There exist finite S-interpretations $\pi_A \ncong \pi_B$ (for instance in min-max semirings with ≥ 3 elements) such that $\pi_A \equiv \pi_B$.

Indeed, finite semiring interpretations are not always first-order definable up to isomorphism.
And even if they are, they may need an infinite axiom system.
And even if a finite axiom system suffices, a single axiom might not.

0-1 laws

Random structures naturally generalise to random S-interpretations:
- Fix a probability distribution p on $S \setminus \{0\}$.
- For each atom $R\bar{a} \in \mathrm{Lit}_{[n]}(\tau)$, make a random choice whether $R\bar{a}$ or $\neg R\bar{a}$ is true, and randomly assign to the true literal a value according to p.

For many semirings S we can prove the following 0-1 law for FO (G. , Helal, Naaf, Wilke, 2022):

With probabilities converging to 1 exponentially fast, valuations $\pi[\![\psi]\!]$ almost surely concentrate on one specific value $j \in S$.

The induced partition $(\Phi_j)_{j\in S}$ of FO into classes of sentences that almost surely evaluate to j, collapses to just three classes Φ_0, Φ_1, and Φ_ε: Every sentence almost surely evaluate to 0, 1, or to $\varepsilon = \prod\{j \in S : j \neq 0\}$.

Locality

Hanf's Theorem: A locality criterion for m-equivalence of two structures based on the number of local substructures of any given isomorphism type.

Gaifman normal form: Every $\psi \in \mathrm{FO}$ is equivalent to a Boolean combination of local formulae and sentences "there exist m disjoint neighbourhoods of radius r satisfying a local property $\varphi^{(r)}$".

In semiring semantics, we have the following results (Bizière, G, Naaf 2023):

- Hanf's Theorem generalises to all fully idempotent semirings, but fails for others.
- Formulae in general do not have Gaifman normal forms over semirings $S \neq \mathbb{B}$. Also for sentences, Gaifman's Theorem fails in certain semirings such as $\mathbb{N}$ and the tropical semiring.
- Positive result: Gaifman normal forms for sentences exist over min-max semirings, and even lattice semirings.

Ehrenfeucht-Fraïssé Games

$G_m(\mathfrak{A}, \mathfrak{B})$: m-move EF-game on τ-structures $\mathfrak{A}$ and $\mathfrak{B}$

i-th move: Spoiler (I) selects $a_i \in \mathfrak{A}$ or $b_i \in \mathfrak{B}$, Duplicator (II) answers with $b_i \in \mathfrak{B}$ or $a_i \in \mathfrak{A}$.

after m moves, II has won if $\{(a_1, b_1), \dots, (a_m, b_m)\}$ is a local isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

Theorem. For any two structures $\mathfrak{A}$ and $\mathfrak{B}$, the following are equivalent

(1) $\mathfrak{A} \equiv_m \mathfrak{B}$

(2) Duplicator wins $G_m(\mathfrak{A}, \mathfrak{B})$

Ehrenfeucht-Fraïssé Games

$G_m(\mathfrak{A},\mathfrak{B})$: m-move EF-game on τ-structures $\mathfrak{A}$ and $\mathfrak{B}$

i-th move: Spoiler (I) selects $a_i \in \mathfrak{A}$ or $b_i \in \mathfrak{B}$, Duplicator (II) answers with $b_i \in \mathfrak{B}$ or $a_i \in \mathfrak{A}$.

after m moves, II has won if $\{(a_1,b_1),\dots,(a_m,b_m)\}$ is a local isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

Theorem. For any two structures $\mathfrak{A}$ and $\mathfrak{B}$, the following are equivalent

(1) $\mathfrak{A} \equiv_m \mathfrak{B}$

(2) Duplicator wins $G_m(\mathfrak{A},\mathfrak{B})$

The game $G(\mathfrak{A},\mathfrak{B})$: I selects $m \in \mathbb{N}$. Then $G_m(\mathfrak{A},\mathfrak{B})$ is played.

$$\text{II wins } G(\mathfrak{A},\mathfrak{B}) \iff \text{II wins } G_m(\mathfrak{A},\mathfrak{B}) \text{ for all } m \iff \mathfrak{A} \equiv_m \mathfrak{B} \text{ for all } m \iff \mathfrak{A} \equiv \mathfrak{B}.$$

Ehrenfeucht-Fraïssé Games

$G_m(\mathfrak{A},\mathfrak{B})$: m-move EF-game on τ-structures $\mathfrak{A}$ and $\mathfrak{B}$

i-th move: Spoiler (I) selects $a_i \in \mathfrak{A}$ or $b_i \in \mathfrak{B}$, Duplicator (II) answers with $b_i \in \mathfrak{B}$ or $a_i \in \mathfrak{A}$.

after m moves, II has won if $\{(a_1,b_1),\dots,(a_m,b_m)\}$ is a local isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

Theorem. For any two structures $\mathfrak{A}$ and $\mathfrak{B}$, the following are equivalent

(1) $\mathfrak{A} \equiv_m \mathfrak{B}$

(2) Duplicator wins $G_m(\mathfrak{A},\mathfrak{B})$

The game $G(\mathfrak{A},\mathfrak{B})$: I selects $m \in \mathbb{N}$. Then $G_m(\mathfrak{A},\mathfrak{B})$ is played.

$$\text{II wins } G(\mathfrak{A},\mathfrak{B}) \iff \text{II wins } G_m(\mathfrak{A},\mathfrak{B}) \text{ for all } m \iff \mathfrak{A} \equiv_m \mathfrak{B} \text{ for all } m \iff \mathfrak{A} \equiv \mathfrak{B}.$$

Question: What about $G_m(\pi_A,\pi_B)$ versus $\pi_A \equiv_m \pi_B$ for semiring interpretations π_A and π_B?

Soundness and Completeness

The game G_m is sound for $\equiv_m$ on a semiring S if for all S-interpretations π_A and π_A:

$$\text{II wins } G_m(\pi_A, \pi_B) \quad \Longrightarrow \quad \pi_A \equiv_m \pi_B$$

G_m is complete for $\equiv_m$ on a semiring S if for all S-interpretations π_A and π_A:

$$\pi_A \equiv_m \pi_B \quad \Longrightarrow \quad \text{II wins } G_m(\pi_a, \pi_B)$$

By the Ehrenfeucht-Fraïssé-Theorem G_m is sound and complete for $\equiv_m$ on the Boolean semiring $\mathbb{B}$.

It follows that the unrestricted game G is sound and complete for $\equiv$ on $\mathbb{B}$.

Soundness and Completeness

The game G_m is sound for $\equiv_m$ on a semiring S if for all S-interpretations π_A and π_A:

$$\text{II wins } G_m(\pi_A, \pi_B) \quad \Longrightarrow \quad \pi_A \equiv_m \pi_B$$

G_m is complete for $\equiv_m$ on a semiring S if for all S-interpretations π_A and π_A:

$$\pi_A \equiv_m \pi_B \quad \Longrightarrow \quad \text{II wins } G_m(\pi_a, \pi_B)$$

By the Ehrenfeucht-Fraïssé-Theorem G_m is sound and complete for $\equiv_m$ on the Boolean semiring $\mathbb{B}$.

It follows that the unrestricted game G is sound and complete for $\equiv$ on $\mathbb{B}$.

However, for other semirings the games need be neither sound nor complete.

To what extent do the games work for semirings?

Question: For which semirings are the EF-games G_m and G sound, for which are they complete?

To what extent do the games work for semirings?

Question: For which semirings are the EF-games G_m and G sound, for which are they complete?

There are also other variants of model comparison games. for which we pose the same question;

The m-move bijection game $BG_m(\pi_A, \pi_B)$: (Hella, for logics with counting quantifiers)

i-th move: Duplicator selects a bijection $h : A \to B$ with $h(a_j) = b_j$ for $j < i$

Spoiler selects a new pair (a_i, b_i) where $b_i = h(a_i)$.

To what extent do the games work for semirings?

Question: For which semirings are the EF-games G_m and G sound, for which are they complete?

There are also other variants of model comparison games. for which we pose the same question;

The m-move bijection game $BG_m(\pi_A, \pi_B)$: (Hella, for logics with counting quantifiers)

i-th move: Duplicator selects a bijection $h : A \to B$ with $h(a_j) = b_j$ for $j < i$

Spoiler selects a new pair (a_i, b_i) where $b_i = h(a_i)$.

The parametrised m-counting game $CG^n_m \pi_A, \pi_B)$:

i-th move: Spoiler selects a set $X \subseteq A$ or $X \subseteq B$ with $|X| \leq n$.

Duplicator answers with $Y \subseteq B$ or $Y \subseteq A$ such that $|Y| = |X|$.

Spoiler selects an element of Y, Duplicator answers with an element of X.

This gives the new pair (a_i, b_i).

Note that $CG^1_m = G_m$

Classification of semirings for model comparison games

- The games G_m are sound on S, for all m, if and only if, S is fully idempotent.

Classification of semirings for model comparison games

- The games G_m are sound on S, for all m, if and only if, S is fully idempotent.
- But if all games G_m are sound and complete on S, then $S = \mathbb{B}$.

Classification of semirings for model comparison games

- The games G_m are sound on S, for all m, if and only if, S is fully idempotent.
- But if all games G_m are sound and complete on S, then $S = \mathbb{B}$.
- Nevertheless, the game G is sound on more semirings, such as $\mathbb{W}[X], \mathbb{N}^\infty, \mathbb{S}^\infty[X], \mathbb{N}, \mathbb{S}[X], \mathbb{B}[X], \mathbb{N}[X]$

Classification of semirings for model comparison games

- The games G_m are sound on S, for all m, if and only if, S is fully idempotent.
- But if all games G_m are sound and complete on S, then $S = \mathbb{B}$.
- Nevertheless, the game G is sound on more semirings, such as $\mathbb{W}[X], \mathbb{N}^\infty, \mathbb{S}^\infty[X], \mathbb{N}, \mathbb{S}[X], \mathbb{B}[X], \mathbb{N}[X]$
- But G is unsound on $\mathbb{V}, \mathbb{T}, \mathbb{L}, \mathbb{D}$

Classification of semirings for model comparison games

- The games G_m are sound on S, for all m, if and only if, S is fully idempotent.
- But if all games G_m are sound and complete on S, then $S = \mathbb{B}$.
- Nevertheless, the game G is sound on more semirings, such as $\mathbb{W}[X], \mathbb{N}^\infty, \mathbb{S}^\infty[X], \mathbb{N}, \mathbb{S}[X], \mathbb{B}[X], \mathbb{N}[X]$
- But G is unsound on $\mathbb{V}, \mathbb{T}, \mathbb{L}, \mathbb{D}$
- The bijection games BG_m are sound on every semiring.

Classification of semirings for model comparison games

- The games G_m are sound on S, for all m, if and only if, S is fully idempotent.
- But if all games G_m are sound and complete on S, then $S = \mathbb{B}$.
- Nevertheless, the game G is sound on more semirings, such as $\mathbb{W}[X], \mathbb{N}^\infty, \mathbb{S}^\infty[X], \mathbb{N}, \mathbb{S}[X], \mathbb{B}[X], \mathbb{N}[X]$
- But G is unsound on $\mathbb{V}, \mathbb{T}, \mathbb{L}, \mathbb{D}$
- The bijection games BG_m are sound on every semiring.
- The m-counting game CG_m^n is sound on n-idempotent semirings.

Classification of semirings for model comparison games

- The games G_m are sound on S, for all m, if and only if, S is fully idempotent.
- But if all games G_m are sound and complete on S, then $S = \mathbb{B}$.
- Nevertheless, the game G is sound on more semirings, such as $\mathbb{W}[X], \mathbb{N}^\infty, \mathbb{S}^\infty[X], \mathbb{N}, \mathbb{S}[X], \mathbb{B}[X], \mathbb{N}[X]$
- But G is unsound on $\mathbb{V}, \mathbb{T}, \mathbb{L}, \mathbb{D}$
- The bijection games BG_m are sound on every semiring.
- The m-counting game CG^n_m is sound on n-idempotent semirings.
- On $\mathbb{N}$ and $\mathbb{N}[X]$, all these games are complete.

Classification of semirings for model comparison games

- The games G_m are sound on S, for all m, if and only if, S is fully idempotent.
- But if all games G_m are sound and complete on S, then $S = \mathbb{B}$.
- Nevertheless, the game G is sound on more semirings, such as $\mathbb{W}[X], \mathbb{N}^\infty, \mathbb{S}^\infty[X], \mathbb{N}, \mathbb{S}[X], \mathbb{B}[X], \mathbb{N}[X]$
- But G is unsound on $\mathbb{V}, \mathbb{T}, \mathbb{L}, \mathbb{D}$
- The bijection games BG_m are sound on every semiring.
- The m-counting game CG^n_m is sound on n-idempotent semirings.
- On $\mathbb{N}$ and $\mathbb{N}[X]$, all these games are complete.
- All these games are incomplete on $\mathbb{V}, \mathbb{T}, \mathbb{L}, \mathbb{D}, \mathbb{N}^\infty, \mathbb{W}[X], \mathbb{S}[X], \mathbb{B}[X]$, and $\mathbb{S}^\infty[X]$.

Ehrenfeucht-Fraïssé Games for Application Semirings

		$S \not\cong \mathbb{B}$ fully idempotent	$\mathbb{V} \cong \mathbb{T}$	$\mathbb{L} \cong \mathbb{D}$	$\mathbb{N}$	$\mathbb{N}^\infty$
Soundness of	G_m for $\equiv_m$	✓	✗	✗	✗	✗
	CG^n_m for $\equiv_m$	✓	✗	✗	✗	✗
	BG_m for $\equiv_m$	✓	✓	✓	✓	✓
	G for $\equiv$	✓	✗	✗	✓	✓
Completeness of	G_m for $\equiv_m$	✗	✗	✗	✓	✗
	CG^n_m for $\equiv_m$	✗	✗	✗	✓	✗
	BG_m for $\equiv_m$	✗	✗	✗	✓	✗
	G for $\equiv$	✗	✗	✗	✓	✗

Ehrenfeucht-Fraïssé Games for Provenance Semirings

		PosBool$[X]$	$\mathbb{W}[X]$	$\mathbb{S}[X]$, $\mathbb{B}[X]$	$\mathbb{N}[X]$	$\mathbb{S}^\infty[X]$
Soundness of	G_m for $\equiv_m$	✓	✗	✗	✗	✗
	CG^n_m for $\equiv_m$	✓	✓	✗	✗	✗
	BG_m for $\equiv_m$	✓	✓	✓	✓	✓
	G for $\equiv$	✓	✓	✓	✓	✓
Completeness of	G_m for $\equiv_m$	✗	✗	✗	✓	✗
	CG^n_m for $\equiv_m$	✗	✗	✗	✓	✗
	BG_m for $\equiv_m$	✗	✗	✗	✓	✗
	G for $\equiv$	✗	✗	✗	✓	✗

How to prove elementary equivalence

Let π_A, π_B be two S-interpretations. We want to prove that $\pi_A \equiv \pi_B$ although π_A and π_B are quite different.

Find a separating set of homomorphisms $h : S \to \mathbb{B}$ such that for all $s, t \in S$ we have that $h(s) \neq h(t)$ for some $h \in H$. Prove that $h \circ \pi_A \equiv h \circ \pi_B$ for all $h \in H$. Since these are $\mathbb{B}$-interpretations, i.e. classical structures, we can do this by Ehrenfeucht-Fraïssé games.

Claim. This implies $\pi_A \equiv \pi_B$

Otherwise there exists ψ such that $\pi_A[\![\psi]\!] = s \neq t = \pi_B[\![\psi]\!]$. But then

$$(h \circ \pi_A)[\![\psi]\!] = h(\pi_A[\![\psi]\!]) = h(s) \neq h(t) = h(\pi_B[\![\psi]\!]) = (h \circ \pi_B)[\![\psi]\!]$$

which is impossible since $h \circ \pi_A \equiv h \circ \pi_B$.

Example

Let $S = \text{PosBool}[X]$. Every $Y \subseteq X$ induces a unique homomorphism $h_Y : \text{PosBool}[X] \to \mathbb{B}$ with $h_Y(x) = \top$ for $x \in Y$ and $h_Y(x) = \bot$ for $x \in X \setminus Y$. For $p \in \text{PosBool}[X]$, we have that $h_Y(p) = \top$ if, and only if, p contains a monomial with only variables from Y.

$\{h_Y : Y \subseteq X\}$ is a separating set of homorphisms.

Claim. The following two $\text{PosBool}[x,y]$-interpretations π_{xy}, π_{yx} are elementarily equivalent.

π_{xy} :

A	P	Q	$\neg P$	$\neg Q$
a	0	y	x	0
b	x	0	0	y
c	y	x	0	0
d	0	0	y	x

π_{yx} :

A	P	Q	$\neg P$	$\neg Q$
a	y	0	0	x
b	0	x	y	0
c	x	y	0	0
d	0	0	x	y

Proof

The separating set of homomorphisms $h : \text{PosBool}[x,y] \to \mathbb{B}$ consists of $h_\varnothing, h_{\{x\}}, h_{\{y\}}$ and $h_{\{x,y\}}$.

For each of these, we have to show that $h \circ \pi_{xy} \equiv h \circ \pi_{yx}$

For $h_\varnothing$ this is trivial.

$h_\varnothing \circ \pi_{xy}$:

A	P	Q	$\neg P$	$\neg Q$
a	$\bot$	$\bot$	$\bot$	$\bot$
b	$\bot$	$\bot$	$\bot$	$\bot$
c	$\bot$	$\bot$	$\bot$	$\bot$
d	$\bot$	$\bot$	$\bot$	$\bot$

$h_\varnothing \circ \pi_{yx}$:

A	P	Q	$\neg P$	$\neg Q$
a	$\bot$	$\bot$	$\bot$	$\bot$
b	$\bot$	$\bot$	$\bot$	$\bot$
c	$\bot$	$\bot$	$\bot$	$\bot$
d	$\bot$	$\bot$	$\bot$	$\bot$

Proof: $h = h_{\{x\}}$

π_{xy} :

A	P	Q	$\neg P$	$\neg Q$
a	0	y	x	0
b	x	0	0	y
c	y	x	0	0
d	0	0	y	x

π_{yx} :

A	P	Q	$\neg P$	$\neg Q$
a	y	0	0	x
b	0	x	y	0
c	x	y	0	0
d	0	0	x	y

$h_{\{x\}} \circ \pi_{xy}$:

A	P	Q	$\neg P$	$\neg Q$
a	$\bot$	$\bot$	$\top$	$\bot$
b	$\top$	$\bot$	$\bot$	$\bot$
c	$\bot$	$\top$	$\bot$	$\bot$
d	$\bot$	$\bot$	$\bot$	$\top$

$h_{\{x\}} \circ \pi_{yx}$:

A	P	Q	$\neg P$	$\neg Q$
a	$\bot$	$\bot$	$\bot$	$\top$
b	$\bot$	$\top$	$\bot$	$\bot$
c	$\top$	$\bot$	$\bot$	$\bot$
d	$\bot$	$\bot$	$\top$	$\bot$

Proof: $h = h_{\{y\}}$

π_{xy} :

A	P	Q	$\neg P$	$\neg Q$
a	0	y	x	0
b	x	0	0	y
c	y	x	0	0
d	0	0	y	x

π_{yx} :

A	P	Q	$\neg P$	$\neg Q$
a	y	0	0	x
b	0	x	y	0
c	x	y	0	0
d	0	0	x	y

$h_{\{y\}} \circ \pi_{xy}$:

A	P	Q	$\neg P$	$\neg Q$
a	$\bot$	$\top$	$\bot$	$\bot$
b	$\bot$	$\bot$	$\bot$	$\top$
c	$\top$	$\bot$	$\bot$	$\bot$
d	$\bot$	$\bot$	$\top$	$\bot$

$h_{\{y\}} \circ \pi_{yx}$:

A	P	Q	$\neg P$	$\neg Q$
a	$\top$	$\bot$	$\bot$	$\bot$
b	$\bot$	$\bot$	$\top$	$\bot$
c	$\bot$	$\top$	$\bot$	$\bot$
d	$\bot$	$\bot$	$\bot$	$\top$

Proof: $h = h_{\{x,y\}}$

π_{xy} :

A	P	Q	$\neg P$	$\neg Q$
a	0	y	x	0
b	x	0	0	y
c	y	x	0	0
d	0	0	y	x

π_{yx} :

A	P	Q	$\neg P$	$\neg Q$
a	y	0	0	x
b	0	x	y	0
c	x	y	0	0
d	0	0	x	y

$h_X \circ \pi_{xy}$:

A	P	Q	$\neg P$	$\neg Q$
a	$\bot$	$\top$	$\top$	$\bot$
b	$\top$	$\bot$	$\bot$	$\top$
c	$\top$	$\top$	$\bot$	$\bot$
d	$\bot$	$\bot$	$\top$	$\top$

$h_X \circ \pi_{yx}$:

A	P	Q	$\neg P$	$\neg Q$
a	$\top$	$\bot$	$\bot$	$\top$
b	$\bot$	$\top$	$\top$	$\bot$
c	$\top$	$\top$	$\bot$	$\bot$
d	$\bot$	$\bot$	$\top$	$\top$