Unified foundations of team semantics via semirings

Timon Barlag, Miika Hannula, Juha Kontinen, Nina Pardal, Jonni Virtema

University of Sheffield, UK

20th International Conference on Principles of
Knowledge Representation and Reasoning, 2–8 September 2023

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

ACADEMY OF FINLAND

Logics with team semantics

Expressive logical formalism for expressing properties of data:

- Notions such as functional dependence, inclusion dependence, and independence between attributes are taken as atomic building blocks of a logic.
- Logics in this setting are high in expressiveness: e.g, equi-expressive with existential second-order logic, i.e. expresses properties of data that are in NP.

Team = set of first-order assignments (i.e., records).

Employee	Research Group	Salary	LineManager
Alice	TCS	50k	Bob
Bob	ML	60k	David
Carol	Security	60k	Carol
David	ML	80k	Carol

- Atom dep(Employee, Salary) (reads: Employee determines Salary)
- Atom LineManager ⊆ Employee (reads: every LineManager is an Employee)

Logics with team semantics

Expressive logical formalism for expressing properties of data:

- Notions such as functional dependence, inclusion dependence, and independence between attributes are taken as atomic building blocks of a logic.
- Logics in this setting are high in expressiveness: e.g, equi-expressive with existential second-order logic, i.e. expresses properties of data that are in NP.

Team = set of first-order assignments (i.e., records).

Employee	Research Group	Salary	LineManager
Alice	TCS	50k	Bob
Bob	ML	60k	David
Carol	Security	60k	Carol
David	ML	80k	Carol

- Atom $\mathrm{dep}(\mathtt{Employee}, \mathtt{Salary})$ (reads: Employee determines Salary)
- Atom $\mathtt{LineManager} \subseteq \mathtt{Employee}$ (reads: every LineManager is an Employee)

Logics with team semantics

More complex properties of data:

- $\exists \texttt{ResearchTheme}\big(\mathrm{dep}(\texttt{ResearchGroup}, \texttt{ResearchTheme}) \wedge \mathrm{dep}(\texttt{ResearchTheme}, \texttt{LineManager})\big)$
  "The data can be extended with values to a new attribute ResearchTheme such the functional dependences mentioned hold".

Employee	ResearchGroup	Salary	LineManager
Alice	TCS	50k	Bob
Bob	ML	60k	David
Carol	Security	60k	Carol
David	ML	80k	Carol

Logics with team semantics

More complex properties of data:

- $\exists \texttt{ResearchTheme}\big(\mathrm{dep}(\texttt{ResearchGroup}, \texttt{ResearchTheme}) \wedge \mathrm{dep}(\texttt{ResearchTheme}, \texttt{LineManager})\big)$
  "The data can be extended with values to a new attribute ResearchTheme such the functional dependences mentioned hold".
- $\mathrm{dep}(\texttt{ResearchGroup}, \texttt{Salary}) \vee \mathrm{dep}(\texttt{ResearchGroup}, \texttt{Salary})$
  "The data can be decomposed into two parts which both satisfy the dependency $\mathrm{dep}(\texttt{ResearchGroup}, \texttt{Salary})$"

Employee	ResearchGroup	Salary	LineManager
Alice	TCS	50k	Bob
Bob	ML	60k	David
Carol	Security	60k	Carol
David	ML	80k	Carol

Multiteam semantics

Multiteam semantics generalises teams to multisets of data:

- Qualitative dependences such as inclusion dependence and independence.
- Quantitative dependences such as marginal multiplicity identity and probabilistic independence between attributes.
- Expressivity relates to quantitative variants of existential second-order logic.

Multiteam = bag of first-order assignments (i.e., records).

ResearchGroup	Salary	LineManager	Multiplicity
TCS	50k	Bob	3
ML	60k	David	2
Security	60k	Carol	2
ML	80k	Carol	1

- Atom ResearchGroup $\perp\!\!\!\perp$ Salary
  (reads: ResearchGroup and Salary are independent of each other)
- Atom ResearchGroup $\approx^*$ LineManager
  (reads: ResearchGroup and LineManager have the same shape of distribution)

Multiteam semantics

Multiteam semantics generalises teams to multisets of data:

- Qualitative dependences such as inclusion dependence and independence.
- Quantitative dependences such as marginal multiplicity identity and probabilistic independence between attributes.
- Expressivity relates to quantitative variants of existential second-order logic.

Multiteam = bag of first-order assignments (i.e., records).

ResearchGroup	Salary	LineManager	Multiplicity
TCS	50k	Bob	3
ML	60k	David	2
Security	60k	Carol	2
ML	80k	Carol	1

- Atom $\mathtt{ResearchGroup} \perp\!\!\!\perp \mathtt{Salary}$
  (reads: ResearchGroup and Salary are independent of each other)
- Atom $\mathtt{ResearchGroup} \approx^{*} \mathtt{LineManager}$
  (reads: ResearchGroup and LineManager have the same shape of distribution)

Probabilistic team semantics

Generalises teams to discrete distributions of data:

- Qualitative dependences such as inclusion dependence and independence.
- Quantitative dependences such as probabilistic independence.
- Expressivity relates to quantitative variants of existential second-order logic.

Probabilistic team = discrete real valued (distributions) of first-order assignments.

Shape	Mass	Volume	weight
Sphere	50kg	$10cm^3$	2/10
Sphere	60kg	$20cm^3$	3/7
Cube	10kg	$30cm^3$	4/7
Torus	50kg	$10cm^3$	1/7

- Atom Shape $\perp\!\!\!\perp$ Mass
  (reads: in the experiment Shape and Mass are picked independently)
- Formula $\mathrm{dep}(\mathtt{Volume}, \mathtt{Mass}) \vee \mathrm{dep}(\mathtt{Volume}, \mathtt{Mass})$
  (reads: There are at most two ways in the data how Volume fuctionally determines Mass.

Probabilistic team semantics

Generalises teams to discrete distributions of data:

- Qualitative dependences such as inclusion dependence and independence.
- Quantitative dependences such as probabilistic independence.
- Expressivity relates to quantitative variants of existential second-order logic.

Probabilistic team = discrete real valued (distributions) of first-order assignments.

Shape	Mass	Volume	weight
Sphere	50kg	$10cm^3$	2/10
Sphere	60kg	$20cm^3$	3/7
Cube	10kg	$30cm^3$	4/7
Torus	50kg	$10cm^3$	1/7

- Atom $\texttt{Shape} \perp\!\!\!\perp \texttt{Mass}$
  (reads: in the experiment Shape and Mass are picked independently)
- Formula $\mathrm{dep}(\texttt{Volume}, \texttt{Mass}) \vee \mathrm{dep}(\texttt{Volume}, \texttt{Mass})$
  (reads: There are at most two ways in the data how Volume fuctionally determines Mass.

Unifying approach to team semantics

- Goal: Give a general recipe for different flavours of team semantics.
- What do we need?
  - Abstraction of a team.
  - A uniform way to define semantics of connectives.
  - A uniform way to define semantics of atoms.
  - A way of obtaining team, multi team, and probabilistic team semantics as instances!
- Solution: Define notions with logical formulae that are interpreted as algebraic expressions over some semiring!

Unifying approach to team semantics

- Goal: Give a general recipe for different flavours of team semantics.
- What do we need?
  - Abstraction of a team.
  - A uniform way to define semantics of connectives.
  - A uniform way to define semantics of atoms.
  - A way of obtaining team, multi team, and probabilistic team semantics as instances!
- Solution: Define notions with logical formulae that are interpreted as algebraic expressions over some semiring!

Unifying approach to team semantics: math definitions

- Examples of semirings:
  - The *Boolean semiring* $\mathbb{B} = (\mathbb{B}, \vee, \wedge, 0, 1)$.
  - The *semiring of natural numbers* $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$.
  - The *probability semiring* $\mathbb{R}_{\geq 0} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$.
  - The semiring of multivariate polynomials $\mathbb{N}[X] = (\mathbb{N}[X], +, \cdot, 0, 1)$.
- $(K, +, 0)$ is a monoid, if $+$ is associative, and 0 is an identity element.
- Semiring is a structure $(K, +, \cdot, 0, 1)$, where
  - $+$ and $\cdot$ are binary operations on K,
  - $(K, +, 0)$ is a commutative monoid with identity element 0,
  - $(K, \cdot, 1)$ is a monoid with identity element 1,
  - $\cdot$ left and right distributes over $+$,
  - $x \cdot 0 = 0 = 0 \cdot x$, for all $x \in K$.
- Sometimes we need to assume that K is a positive (no zero divisors and $a + b = 0$ implies $a = b = 0$)
- Often we need to assume that K is commutative.

Unifying approach to team semantics: math definitions

- Examples of semirings:
  - The *Boolean semiring* $\mathbb{B} = (\mathbb{B}, \vee, \wedge, 0, 1)$.
  - The *semiring of natural numbers* $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$.
  - The *probability semiring* $\mathbb{R}_{\geq 0} = (\mathbb{R}_{\geq 0}, +, \cdot, 0, 1)$.
  - The semiring of multivariate polynomials $\mathbb{N}[X] = (\mathbb{N}[X], +, \cdot, 0, 1)$.
- $(K, +, 0)$ is a monoid, if $+$ is associative, and 0 is an identity element.
- Semiring is a structure $(K, +, \cdot, 0, 1)$, where
  - $+$ and $\cdot$ are binary operations on K,
  - $(K, +, 0)$ is a commutative monoid with identity element 0,
  - $(K, \cdot, 1)$ is a monoid with identity element 1,
  - $\cdot$ left and right distributes over $+$,
  - $x \cdot 0 = 0 = 0 \cdot x$, for all $x \in K$.
- Sometimes we need to assume that K is a positive (no zero divisors and $a + b = 0$ implies $a = b = 0$)
- Often we need to assume that K is commutative.

K-teams

Given a semiring $(K, +, \cdot, 0, 1)$, a finite set of first-order variables VAR, and a first-order structure $\mathfrak{A}$ with domain A

- A K-team maps every assignment $s : VAR \to A$ to a value in the semiring. (It is a function $X : A^{VAR} \to K$)
- The sum $X + Y$ of two K-teams is defined such that $s \mapsto X(s) + Y(s)$.
- For the Boolean semiring $(\mathbb{B}, \vee, \wedge, 0, 1)$, we obtain set-based teams. Addition corresponds to set union (via characteristic functions of sets).
- A marginalisation $X \upharpoonright VAR'$ is defined such that

$$s' \mapsto \sum_{(s \upharpoonright VAR') = s'} X(s)$$

  for $s : VAR' \to A$.
- For the semiring of natural numbers, we obtain multiteams. Addition corresponds to disjoint union of multisets and marginalisation is standard.

K-teams

Given a semiring $(K, +, \cdot, 0, 1)$, a finite set of first-order variables VAR, and a first-order structure $\mathfrak{A}$ with domain A

- A K-team maps every assignment $s : VAR \to A$ to a value in the semiring. (It is a function $X : A^{VAR} \to K$)
- The sum $X + Y$ of two K-teams is defined such that $s \mapsto X(s) + Y(s)$.
- For the Boolean semiring $(\mathbb{B}, \vee, \wedge, 0, 1)$, we obtain set-based teams. Addition corresponds to set union (via characteristic functions of sets).
- A marginalisation $X \upharpoonright VAR'$ is defined such that

  $$s' \mapsto \sum_{(s \upharpoonright VAR') = s'} X(s)$$

  for $s : VAR' \to A$.
- For the semiring of natural numbers, we obtain multiteams. Addition corresponds to disjoint union of multisets and marginalisation is standard.

K-teams

Given a semiring $(K, +, \cdot, 0, 1)$, a finite set of first-order variables VAR, and a first-order structure $\mathfrak{A}$ with domain A

- A K-team maps every assignment $s : VAR \to A$ to a value in the semiring. (It is a function $X : A^{VAR} \to K$)
- The sum $X + Y$ of two K-teams is defined such that $s \mapsto X(s) + Y(s)$.
- For the Boolean semiring $(\mathbb{B}, \vee, \wedge, 0, 1)$, we obtain set-based teams. Addition corresponds to set union (via characteristic functions of sets).
- A marginalisation $X \upharpoonright VAR'$ is defined such that

  $$s' \mapsto \sum_{(s \upharpoonright VAR') = s'} X(s)$$

  for $s : VAR' \to A$.
- For the semiring of natural numbers, we obtain multiteams. Addition corresponds to disjoint union of multisets and marginalisation is standard.

Basic results on team, multiteams, and probabilistic teams

- Team semantics: empty team property

$$\mathfrak{A} \models_{\emptyset} \varphi, \text{ for any formula } \varphi.$$

- Multiteam semantics: closure under scalar multiplication of teams

$$\mathfrak{A} \models_X \varphi \text{ implies } \mathfrak{A} \models_{c \cdot X} \varphi, \text{ for any } c \in \mathbb{N}.$$

- Probabilistic (real-weighted) team semantics: distribution invariance

$$\mathfrak{A} \models_X \varphi \text{ if and only if } \mathfrak{A} \models_Y \varphi, \text{ provided that dist(X)=dist(Y).}$$

- General notion: If X is a K-team, then for each $c \in K$

$$\mathfrak{A} \models_X \varphi \text{ implies } \mathfrak{A} \models_{c \cdot X} \varphi.$$

- The stronger results for the probabilistic semiring follows via the existence of multiplicative inverses.

Basic results on team, multiteams, and probabilistic teams

- Team semantics: empty team property

$$\mathfrak{A} \models_{\emptyset} \varphi, \text{ for any formula } \varphi.$$

- Multiteam semantics: closure under scalar multiplication of teams

$$\mathfrak{A} \models_X \varphi \text{ implies } \mathfrak{A} \models_{c \cdot X} \varphi, \text{ for any } c \in \mathbb{N}.$$

- Probabilistic (real-weighted) team semantics: distribution invariance

$$\mathfrak{A} \models_X \varphi \text{ if and only if } \mathfrak{A} \models_Y \varphi, \text{ provided that dist(X)=dist(Y).}$$

- General notion: If X is a K-team, then for each $c \in K$

$$\mathfrak{A} \models_X \varphi \text{ implies } \mathfrak{A} \models_{c \cdot X} \varphi.$$

- The stronger results for the probabilistic semiring follows via the existence of multiplicative inverses.

Basic results on team, multiteams, and probabilistic teams

- (Qualitative) inclusion logic $FO(\subseteq)$ is closed under unions

$$\mathfrak{A} \models_X \varphi \quad \& \quad \mathfrak{A} \models_Y \varphi \text{ implies } \mathfrak{A} \models_{X \cup Y} \varphi$$

- (Quantitative) real-weighted/multiteam inclusion logic $FO(\leq)$ is closed under disjoint unions

$$\mathfrak{A} \models_X \varphi \quad \& \quad \mathfrak{A} \models_Y \varphi \text{ implies } \mathfrak{A} \models_{X \uplus Y} \varphi.$$

- (Quantitative) Probabilistic inclusion logic $FO(\leq)$ is closed under scaled unions

$$\mathfrak{A} \models_X \varphi \quad \& \quad \mathfrak{A} \models_Y \varphi \text{ implies } \mathfrak{A} \models_{\alpha \cdot X \uplus (1-\alpha) Y} \varphi, \text{ for all } \alpha \in [0,1].$$

- General notion: Closure under addition. If X is a K-team

$$\mathfrak{A} \models_X \varphi \quad \& \quad \mathfrak{A} \models_Y \varphi \text{ implies } \mathfrak{A} \models_{X+Y} \varphi,$$

where $+$ inherits its semantics from K.

Basic results on team, multiteams, and probabilistic teams

- (Qualitative) inclusion logic $FO(\subseteq)$ is closed under unions

$$\mathfrak{A} \models_X \varphi \quad \& \quad \mathfrak{A} \models_Y \varphi \text{ implies } \mathfrak{A} \models_{X \cup Y} \varphi$$

- (Quantitative) real-weighted/multiteam inclusion logic $FO(\leq)$ is closed under disjoint unions

$$\mathfrak{A} \models_X \varphi \quad \& \quad \mathfrak{A} \models_Y \varphi \text{ implies } \mathfrak{A} \models_{X \uplus Y} \varphi.$$

- (Quantitative) Probabilistic inclusion logic $FO(\leq)$ is closed under scaled unions

$$\mathfrak{A} \models_X \varphi \quad \& \quad \mathfrak{A} \models_Y \varphi \text{ implies } \mathfrak{A} \models_{\alpha \cdot X \uplus (1-\alpha) Y} \varphi, \text{ for all } \alpha \in [0,1].$$

- General notion: Closure under addition. If X is a K-team

$$\mathfrak{A} \models_X \varphi \quad \& \quad \mathfrak{A} \models_Y \varphi \text{ implies } \mathfrak{A} \models_{X+Y} \varphi,$$

where $+$ inherits its semantics from K.

What we have seen so far?

- The concept of a K-team and some hints to K-team semantics.
- How teams, multiteams, and probablistic teams arise from K-teams.
  - Semantics for the disjuctions arises from addition:
    $\mathfrak{A} \models_X \varphi \vee \psi \quad$ iff $\quad \mathfrak{A} \models_Y \varphi$ and $\mathfrak{A} \models_Z \varphi$, for some Y and Z s.t $X = Y + Z$.
  - Semantics for existential quantifiers arises from marginalisations.
- How some basic results for different team semantics variants arise from K-teams.

What we have seen so far?

- The concept of a K-team and some hints to K-team semantics.
- How teams, multiteams, and probablistic teams arise from K-teams.
  - Semantics for the disjuctions arises from addition:
    $\mathfrak{A} \models_X \varphi \vee \psi \quad$ iff $\quad \mathfrak{A} \models_Y \varphi$ and $\mathfrak{A} \models_Z \varphi$, for some Y and Z s.t $X = Y + Z$.
  - Semantics for existential quantifiers arises from marginalisations.
- How some basic results for different team semantics variants arise from K-teams.

What we have seen so far?

- The concept of a K-team and some hints to K-team semantics.
- How teams, multiteams, and probablistic teams arise from K-teams.
  - Semantics for the disjuctions arises from addition:
    $\mathfrak{A} \models_X \varphi \vee \psi \quad \text{iff} \quad \mathfrak{A} \models_Y \varphi$ and $\mathfrak{A} \models_Z \varphi$, for some Y and Z s.t $X = Y + Z$.
  - Semantics for existential quantifiers arises from marginalisations.
- How some basic results for different team semantics variants arise from K-teams.

Logic for defining atoms

We extend semiring semantics for first-order logic (Grädel and Tannen 2017) with formula (in)equalities.

- The value of a formula is an element of a semiring.
- The value can denote a truth value, a number distinct of proofs, or something else.
- The value can be a multivariate polynomial carrying some provenance information.
- How the value of a formulae is computed?
  - For literals the value is given by a K-interpretation function.
  - For disjunction, the value is the sum of the values of the disjuncts.
  - For conjunction, the value is the product of the values of the conjuncts.
  - For the quantifiers, the value is a sum or product of all the possible interpretations of the quantified variable
  - For formula (in)equalities

$$[\![\phi * \psi]\!] = \begin{cases} 1 & \text{if } [\![\phi]\!] * [\![\psi]\!] \\ 0 & \text{otherwise,} \end{cases}$$

where $* \in \{=, \neq, \leq, \not\leq\}$.

Logic for defining atoms

We extend semiring semantics for first-order logic (Grädel and Tannen 2017) with formula (in)equalities.

- The value of a formula is an element of a semiring.
- The value can denote a truth value, a number distinct of proofs, or something else.
- The value can be a multivariate polynomial carrying some provenance information.
- How the value of a formulae is computed?
  - For literals the value is given by a K-interpretation function.
  - For disjunction, the value is the sum of the values of the disjuncts.
  - For conjunction, the value is the product of the values of the conjuncts.
  - For the quantifiers, the value is a sum or product of all the possible interpretations of the quantified variable
  - For formula (in)equalities

$$[\![\phi * \psi]\!] = \begin{cases} 1 & \text{if } [\![\phi]\!] * [\![\psi]\!] \\ 0 & \text{otherwise,} \end{cases}$$

where $* \in \{=, \neq, \leq, \not\leq\}$.

Logic for defining atoms

We extend semiring semantics for first-order logic (Grädel and Tannen 2017) with formula (in)equalities.

- The value of a formula is an element of a semiring.
- The value can denote a truth value, a number distinct of proofs, or something else.
- The value can be a multivariate polynomial carrying some provenance information.
- How the value of a formulae is computed?
  - For literals the value is given by a K-interpretation function.
  - For disjunction, the value is the sum of the values of the disjuncts.
  - For conjunction, the value is the product of the values of the conjuncts.
  - For the quantifiers, the value is a sum or product of all the possible interpretations of the quantified variable
  - For formula (in)equalities

$$[\![\phi * \psi]\!] = \begin{cases} 1 & \text{if } [\![\phi]\!] * [\![\psi]\!] \\ 0 & \text{otherwise,} \end{cases}$$

where $* \in \{=, \neq, \leq, \not\leq\}$.

Logical definitions of atoms

- Consider an inclusion atom $\vec{x}_i \leq \vec{x}_j$.
  - The defining formula is of the form $\forall \vec{u} \Big(\theta_{\vec{i}}(\vec{u}) \leq \theta_{\vec{j}}(\vec{u}) \Big)$.
  - The formula $\theta_{\vec{i}}$ extracts the marginalisation for $\vec{x}_i$ from the K-team.
  - On the Boolean semiring the above yields the (qualitative) inclusion atoms.
  - On the probabilistic semiring we obtain the marginal distribution identity atoms.
- Consider an independence atom $\vec{x}_j \perp\!\!\!\perp_{\vec{x}_i} \vec{x}_k$.
  - Defining formula: $\forall \vec{u}\vec{v}\vec{w} \Big(\big(\theta_{\vec{i},\vec{j}}(\vec{u},\vec{v}) \wedge \theta_{\vec{i},\vec{k}}(\vec{u},\vec{w}) \big) = \big(\theta_{\vec{i}}(\vec{u}) \wedge \theta_{\vec{i},\vec{j},\vec{k}}(\vec{u},\vec{v},\vec{w}) \big) \Big)$
  - The formulae θ extract the relevant marginalisation from the K-team.
  - The above is similar to how the probabilistic conditional independence $y \perp\!\!\!\perp_x z$ could be written in probability theory:
    $P(xy = ab) \cdot P(xz = ac) = P(xyz = abc) \cdot P(x = a)$, for all values a,b,c
  - On the Boolean semiring the above yields the (qualitative) independence atoms.
  - On the probabilistic semiring we obtain the probabilistic independence atoms.

Logical definitions of atoms

- Consider an inclusion atom $\vec{x}_i \leq \vec{x}_j$.
  - The defining formula is of the form $\forall \vec{u}\Big(\theta_{\vec{i}}(\vec{u}) \leq \theta_{\vec{j}}(\vec{u})\Big)$.
  - The formula $\theta_{\vec{i}}$ extracts the marginalisation for $\vec{x}_i$ from the K-team.
  - On the Boolean semiring the above yields the (qualitative) inclusion atoms.
  - On the probabilistic semiring we obtain the marginal distribution identity atoms.
- Consider an independence atom $\vec{x}_j \perp\!\!\!\perp_{\vec{x}_i} \vec{x}_k$.
  - Defining formula: $\forall \vec{u}\vec{v}\vec{w}\Big(\big(\theta_{\vec{i},\vec{j}}(\vec{u},\vec{v}) \wedge \theta_{\vec{i},\vec{k}}(\vec{u},\vec{w})\big) = \big(\theta_{\vec{i}}(\vec{u}) \wedge \theta_{\vec{i},\vec{j},\vec{k}}(\vec{u},\vec{v},\vec{w})\big)\Big)$
  - The formulae θ extract the relevant marginalisation from the K-team.
  - The above is similar to how the probabilistic conditional independence $y \perp\!\!\!\perp_x z$ could be written in probability theory:
    $P(xy = ab) \cdot P(xz = ac) = P(xyz = abc) \cdot P(x = a)$, for all values a,b,c
  - On the Boolean semiring the above yields the (qualitative) independence atoms.
  - On the probabilistic semiring we obtain the probabilistic independence atoms.

Logical definitions of atoms

- Consider an inclusion atom $\vec{x}_i \leq \vec{x}_j$.
  - The defining formula is of the form $\forall \vec{u}\Big(\theta_{\vec{i}}(\vec{u}) \leq \theta_{\vec{j}}(\vec{u})\Big)$.
  - The formula $\theta_{\vec{i}}$ extracts the marginalisation for $\vec{x}_i$ from the K-team.
  - On the Boolean semiring the above yields the (qualitative) inclusion atoms.
  - On the probabilistic semiring we obtain the marginal distribution identity atoms.
- Consider an independence atom $\vec{x}_j \perp\!\!\!\perp_{\vec{x}_i} \vec{x}_k$.
  - Defining formula: $\forall \vec{u}\vec{v}\vec{w}\Big(\big(\theta_{\vec{i},\vec{j}}(\vec{u},\vec{v}) \wedge \theta_{\vec{i},\vec{k}}(\vec{u},\vec{w})\big) = \big(\theta_{\vec{i}}(\vec{u}) \wedge \theta_{\vec{i},\vec{j},\vec{k}}(\vec{u},\vec{v},\vec{w})\big)\Big)$
  - The formulae θ extract the relevant marginalisation from the K-team.
  - The above is similar to how the probabilistic conditional independence $y \perp\!\!\!\perp_x z$ could be written in probability theory:
    $P(xy = ab) \cdot P(xz = ac) = P(xyz = abc) \cdot P(x = a)$, for all values a,b,c
  - On the Boolean semiring the above yields the (qualitative) independence atoms.
  - On the probabilistic semiring we obtain the probabilistic independence atoms.

Logical definitions of atoms

- Consider an inclusion atom $\vec{x}_i \leq \vec{x}_j$.
  - The defining formula is of the form $\forall \vec{u}\Big(\theta_{\vec{i}}(\vec{u}) \leq \theta_{\vec{j}}(\vec{u})\Big)$.
  - The formula $\theta_{\vec{i}}$ extracts the marginalisation for $\vec{x}_i$ from the K-team.
  - On the Boolean semiring the above yields the (qualitative) inclusion atoms.
  - On the probabilistic semiring we obtain the marginal distribution identity atoms.
- Consider an independence atom $\vec{x}_j \perp\!\!\!\perp_{\vec{x}_i} \vec{x}_k$.
  - Defining formula: $\forall \vec{u}\vec{v}\vec{w}\Big(\big(\theta_{\vec{i},\vec{j}}(\vec{u},\vec{v}) \wedge \theta_{\vec{i},\vec{k}}(\vec{u},\vec{w})\big) = \big(\theta_{\vec{i}}(\vec{u}) \wedge \theta_{\vec{i},\vec{j},\vec{k}}(\vec{u},\vec{v},\vec{w})\big)\Big)$
  - The formulae θ extract the relevant marginalisation from the K-team.
  - The above is similar to how the probabilistic conditional independence $y \perp\!\!\!\perp_x z$ could be written in probability theory:
    $P(xy = ab) \cdot P(xz = ac) = P(xyz = abc) \cdot P(x = a)$, for all values a,b,c
  - On the Boolean semiring the above yields the (qualitative) independence atoms.
  - On the probabilistic semiring we obtain the probabilistic independence atoms.

Logical definitions of atoms

- Consider an inclusion atom $\vec{x}_i \leq \vec{x}_j$.
  - The defining formula is of the form $\forall \vec{u}\Big(\theta_{\vec{i}}(\vec{u}) \leq \theta_{\vec{j}}(\vec{u})\Big)$.
  - The formula $\theta_{\vec{i}}$ extracts the marginalisation for $\vec{x}_i$ from the K-team.
  - On the Boolean semiring the above yields the (qualitative) inclusion atoms.
  - On the probabilistic semiring we obtain the marginal distribution identity atoms.
- Consider an independence atom $\vec{x}_j \perp\!\!\!\perp_{\vec{x}_i} \vec{x}_k$.
  - Defining formula: $\forall \vec{u}\vec{v}\vec{w}\Big(\big(\theta_{\vec{i},\vec{j}}(\vec{u},\vec{v}) \wedge \theta_{\vec{i},\vec{k}}(\vec{u},\vec{w})\big) = \big(\theta_{\vec{i}}(\vec{u}) \wedge \theta_{\vec{i},\vec{j},\vec{k}}(\vec{u},\vec{v},\vec{w})\big)\Big)$
  - The formulae θ extract the relevant marginalisation from the K-team.
  - The above is similar to how the probabilistic conditional independence $y \perp\!\!\!\perp_x z$ could be written in probability theory:
    $P(xy = ab) \cdot P(xz = ac) = P(xyz = abc) \cdot P(x = a)$, for all values a,b,c
  - On the Boolean semiring the above yields the (qualitative) independence atoms.
  - On the probabilistic semiring we obtain the probabilistic independence atoms.

RECAP

- Goal: Give a general recipe for different flavours of team semantics.
- What do we need?
  - Abstraction of a team.
  - A uniform way to define semantics of connectives.
  - A uniform way to define semantics of atoms.
  - A way of obtaining team, multi team, and probabilistic team semantics as instances!
- Solution: Define notions with logical formulae that are interpreted as algebraic expressions over some semiring!

Future and ongoing work

- Database repairs
  - A logic where the values of formulae indicate how far the formula is from being true.
  - The above is used to define various repair notions logically.
  - *A fine-grained framework for database repairs*, ArXiv 2023 (with Nina Pardal) `https://doi.org/10.48550/arXiv.2306.15516`
- What does the semiring approach reveal about axiomatisations?
- Study of provenance using multivariate polynomials as annotations.
- Counting proofs in team semantics setting (initated in Haak et. al. 2019).
- Complexity theoretic issues related to BSS-machines and the existential first-order theory of K.

References I

Durand, A., Hannula, M., Kontinen, J., Meier, A., and Virtema, J. (2018a).
Approximation and dependence via multiteam semantics.
Ann. Math. Artif. Intell., 83(3-4):297–320.

Durand, A., Hannula, M., Kontinen, J., Meier, A., and Virtema, J. (2018b).
Probabilistic team semantics.
In *Foundations of Information and Knowledge Systems - 10th International Symposium, FoIKS 2018, Budapest, Hungary, May 14-18, 2018, Proceedings*, pages 186–206.

Grädel, E. and Tannen, V. (2017).
Semiring provenance for first-order model checking.
CoRR, abs/1712.01980.

Haak, A., Kontinen, J., Müller, F., Vollmer, H., and Yang, F. (2019).
Counting of teams in first-order team logics.
In Rossmanith, P., Heggernes, P., and Katoen, J., editors, *44th International Symposium on Mathematical Foundations of Computer Science, MFCS 2019, August 26-30, 2019, Aachen, Germany*, volume 138 of *LIPIcs*, pages 19:1–19:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik.

Hannula, M., Hirvonen, Å., Kontinen, J., Kulikov, V., and Virtema, J. (2019).
Facets of distribution identities in probabilistic team semantics.
In *JELIA*, volume 11468 of *Lecture Notes in Computer Science*, pages 304–320. Springer.

References II

Hannula, M., Kontinen, J., den Bussche, J. V., and Virtema, J. (2020).
Descriptive complexity of real computation and probabilistic independence logic.
In Hermanns, H., Zhang, L., Kobayashi, N., and Miller, D., editors, *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020*, pages 550–563. ACM.

Pardal, N. and Virtema, J. (2023).
A fine-grained framework for database repairs.
CoRR, abs/2306.15516.