On elementary logics for quantitative dependencies

Hannula, Miika

2022-12

Hannula , M , Hirvonen , M & Kontinen , J 2022 , ' On elementary logics for quantitative dependencies ' , Annals of Pure and Applied Logic , vol. 173 , no. 10 , 103104 . https://doi.org/10.1016/j.apal.2022.10

http://hdl.handle.net/10138/349864
https://doi.org/10.1016/j.apal.2022.103104

Annals of Pure and Applied Logic 173 (2022) 103104

Contents lists available at ScienceDirect

Annals of Pure and Applied Logic

www.elsevier.com/locate/apal

On elementary logics for quantitative dependencies

Miika Hannula [1], Minna Hirvonen [2,*], Juha Kontinen [3]

Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

A R T I C L E   I N F O

Article history:
Available online 15 February 2022

MSC:
03B60
03B70
03D15
68Q17

Keywords:
Probabilistic team semantics
Dependence logic
Conditional independence
Metafinite structures

A B S T R A C T

We define and study logics in the framework of probabilistic team semantics and over metafinite structures. Our work is paralleled by the recent development of novel axiomatizable and tractable logics in team semantics that are closed under the Boolean negation. Our logics employ new probabilistic atoms that resemble so-called extended atoms from the team semantics literature. We also define counterparts of our logics over metafinite structures and show that all of our logics can be translated into functional fixed point logic implying a polynomial time upper bound for data complexity with respect to BSS-computations.

1. Introduction

Team semantics is a semantical framework originally introduced by Hodges [17] and Väänänen with the introduction of *dependence logic* [22]. Soon after the introduction of dependence logic, the focus in (first-order) team semantics turned to independence logic and inclusion logic that were introduced in [11,7]. During the past decade research on logics in team semantics has flourished with interesting connections to many fields such as database theory [13], statistics [2], and temporal hyperproperties [20].

In team semantics formulas are evaluated over sets of assignments (called *teams*) rather than single assignments as in first-order logic. This feature has the effect that knowing the expressive power of a logic for sentences does not immediately give a characterization for the expressive power of the open formulas of the logic. For example, while it follows from the earlier results of [16,6,23] that dependence logic and independence logic are both equivalent to existential second-order logic (ESO) on the level of sentences, the open formulas of dependence logic are strictly less expressive compared to independence logic: The latter

* Corresponding author.
E-mail address: minna.hirvonen@helsinki.fi (M. Hirvonen).

[1] Supported by grant 322795 of the Academy of Finland.
[2] Supported by the Finnish Academy of Science and Letters (the Vilho, Yrjö and Kalle Väisälä Foundation).
[3] Supported by grant 308712 of the Academy of Finland.

https://doi.org/10.1016/j.apal.2022.103104

characterizes all ESO-definable team properties [7], whereas the former only downward closed ESO-definable properties [18].

A salient feature of (most) logics in team semantics is that their expressive power exceeds that of first-order logic. Only recently a team-based logic FOT was defined whose expressive power coincides with first-order logic both on the level of sentences and open formulas. Previously it had been observed e.g., that the extensions of FO by constancy atoms or the Boolean negation $\sim$ are both equivalent to FO over sentences but strictly less expressive than FO for open formulas when the team is represented by a relation [8,21]. The logic FOT utilizes a weaker version of disjunction and the existential quantifier in order not to go beyond the expressivity of FO (see [5] for a systematic study of this phenomenon). We will follow the same strategy when defining our new logics in the probabilistic setting.

In this paper our focus is on probabilistic team semantics that extends the area of team semantics from qualitative to quantitative dependencies such as probabilistic independence. A probabilistic team is a set of assignments with an additional function that maps each assignment to some numerical value. Usually, the function is a probability distribution, but it can also be thought of as a frequency distribution. We allow the values to be any non-negative real numbers. The systematic study of logics in probabilistic team semantics was initiated by the works [3,4] and they have already found applications, e.g., in the study of the implication problem of conditional independence [12] and the foundations of quantum mechanics [1]. In the literature, the probabilistic team semantics setting has only been considered on finite base structures. As infinite base structures pose some technical problems, we also restrict ourselves to finite base structures, and therefore only consider finite teams.

By the results of [4,14] probabilistic independence logic is equivalent to a sublogic of ESO interpreted over so-called $\mathbb{R}$-structures ($\mathsf{ESO}_{\mathbb{R}}$). In this paper our goal is to initiate a study of tractable probabilistic logics and to find their analogues over metafinite structures. We note that the tractability frontier of the previously defined logics in probabilistic team semantics has been recently charted in [15]. We introduce a new logic called $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})$, in which the disjunction and the quantifiers are similar to the ones in FOT and the atoms compare the probabilities of events defined by quantifier-free formulas. In fact, the logic $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})$ can be seen as a generalization of FOT for probabilistic team semantics. In addition to the qualitative atoms expressible in FOT, certain previously studied probabilistic atoms, i.e. marginal identity and probabilistic conditional independence, are also expressible in $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})$.

We also define two other team-based logics: $\mathsf{FOPT}(\leq^{\delta})$ which is a fragment of $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})$, and $\mathsf{FOPT}(\leq^{\delta}_{\mathrm{c}})$ in which every formula of $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})$ is expressible. The logic $\mathsf{FOPT}(\leq^{\delta}_{\mathrm{c}})$ features a new type of atom, *conditional probability inequality*, that can be used to compare conditional probabilities. With this atom, we can express both kinds of extended atoms from $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})$, i.e. the extended probabilistic inclusion and the extended probabilistic conditional independence. We also take a look at $\mathsf{FOPT}(\leq^{\delta}_{\mathrm{c}})$ from a complexity theoretic point of view and show that its satisfiability and validity problems are r.e.-complete and co-r.e.-complete, respectively.

In the second part of the article we consider logics over two-sorted (metafinite) structures which, in addition to a finite structure, come with an infinite second sort and functions that bridge the two sorts. Metafinite structures have been introduced in [9] as a way to handle objects that consist of both structures and numbers. These types of objects arise naturally, e.g., complexity theory, database theory, and optimization theory. We define a logic, $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$, which is an extension of first-order logic on metafinite structures with a numerical second sort that has access to multiplication and aggregate sums over non-negative reals. We show that $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})$ can be translated into $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$, and identify a fragment of $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ which is equi-expressive with $\mathsf{FOPT}(\leq^{\delta})$. We also give a translation from $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ to functional fixed point logic $\mathsf{FFP}_{\mathbb{R}}$ over metafinite structures and thus obtain a polynomial time upper bound for the data complexity of our new logics in the BSS-model.

2. Preliminaries

First-order variables are denoted by x, y, z and tuples of first-order variables by $\bar{x}, \bar{y}, \bar{z}$. The set of variables that appear in the tuple $\bar{x}$ is denoted by $\mathrm{Var}(\bar{x})$, and by $|\bar{x}|$, we denote the length of the tuple $\bar{x}$. A vocabulary τ is a finite set of relation, function, and constant symbols, denoted by R, f, and c, respectively. Each relation symbol R and function symbol f has a prescribed arity which we denote by $\mathrm{ar}(R)$ and $\mathrm{ar}(f)$.

A vocabulary τ is called *relational* if it only contains relation symbols, and *functional* if it only contains function symbols. We sometimes assume that the vocabulary we are considering is relational. This assumption can be made without loss of generality since each function can be expressed by a relation that describes its graph. For some proofs, it is useful to allow the vocabulary to contain constants, and therefore we sometimes assume that the vocabulary solely consists of relation and constant symbols.

2.1. Team semantics and the logics FOT *and* FOT$^{\downarrow}$

Let τ be a finite vocabulary that only contains relation and constant symbols. We assume that $\{=\} \subseteq \tau$.[4] Let D be a finite set of variables and $\mathcal{A}$ a finite τ-structure. An *assignment* of a structure $\mathcal{A}$ for the set D is a function $s\colon D \to A$. A *team* X of $\mathcal{A}$ over the set D is a finite set of assignments $s\colon D \to A$.[5] The set D is also called the *domain* of X, or $\mathrm{Dom}(X)$ for short. For a variable x and $a \in A$, we denote by $s(a/x)$, the modified assignment $s(a/x)\colon D \cup \{x\} \to A$ such that $s(a/x)(y) = a$ if $y = x$, and $s(a/x)(y) = s(y)$ otherwise. The modified team $X(a/x)$ is defined as the set $X(a/x) := \{s(a/x) \mid s \in X\}$.

We consider two team-based logics, FOT and FOT$^{\downarrow}$, which were introduced in [19]. The expressive power of FOT coincides with first-order logic, and FOT$^{\downarrow}$ captures downward closed first-order team properties [19]. The logics that we introduce in section 3 can be seen as generalizations of these two logics.

First-order τ-terms and atomic formulas are defined in the usual way. We let δ be a quantifier- and disjunction-free[6] first-order formula, i.e. $\delta ::= \lambda \mid \neg\delta \mid (\delta \wedge \delta)$ for any first-order atomic formula λ of the vocabulary τ. Let x be a first-order variable, and let $\bar{x}$ and $\bar{y}$ be tuples of variables with $|\bar{x}| = |\bar{y}|$. The syntax for the logic FOT over a vocabulary τ is then as follows:

$$\phi ::= \lambda \mid \bar{x} \subseteq \bar{y} \mid \dot{\sim}\,\phi \mid (\phi \wedge \phi) \mid (\phi \mathbin{\dot\vee} \phi) \mid \exists^1 x\phi \mid \forall^1 x\phi,$$

and for the logic FOT$^{\downarrow}$ as follows:

$$\phi ::= \delta \mid (\phi \wedge \phi) \mid (\phi \mathbin{\dot\vee} \phi) \mid \exists^1 x\phi \mid \forall^1 x\phi.$$

The semantics for the two logics is defined as follows:

- $\mathcal{A} \models_X \delta$ iff $\mathcal{A} \models_s \delta$ for all $s \in X$.
- $\mathcal{A} \models_X \bar{x} \subseteq \bar{y}$ iff for all $s \in X$, there exists $s' \in X$ such that $s(\bar{x}) = s'(\bar{y})$.
- $\mathcal{A} \models_X \dot{\sim}\,\phi$ iff $\mathcal{A} \not\models_X \phi$ or $X = \varnothing$.
- $\mathcal{A} \models_X \phi \wedge \psi$ iff $\mathcal{A} \models_X \phi$ and $\mathcal{A} \models_X \psi$.

[4] We regard equality as a part of the vocabulary rather than a logical constant for two reasons: (1) in Section 6, we will consider metafinite structures where the infinite structure has inequality $\leq$ instead of equality, and (2) in Section 8, we will consider metafinite structures where all the relations from the finite structure have been replaced with their characteristic functions, and having the characteristic function for equality simplifies things.

[5] Note that unlike in our version of probabilistic team semantics, here X is not required to be maximal; it can be any finite set of assignments.

[6] We have ruled out disjunction at the quantifier-free first-order level for notational convenience. We could define disjunction using negation and disjunction as in first-order logic, i.e. $\delta_0 \vee \delta_1 := \neg(\neg\delta_0 \wedge \neg\delta_1)$ but we want to refrain from using the symbol $\vee$ in this meaning because in team semantics it is customary to use it for the so-called tensor disjunction. We do not want to use $\dot\vee$ in this way either: although for single assignments the meaning of $\neg(\neg\delta_0 \wedge \neg\delta_1)$ and $\delta_0 \mathbin{\dot\vee} \delta_1$ are the same, e.g., $\mathcal{A} \models_X \neg(\neg x = y \wedge \neg x \neq y)$ is true for any X, whereas $\mathcal{A} \models_X x = y \mathbin{\dot\vee} x \neq y$ might not be if X is not a singleton.

- $\mathcal{A} \models_X \phi \vee\!\!\!\vee \psi$ iff $\mathcal{A} \models_X \phi$ or $\mathcal{A} \models_X \psi$.
- $\mathcal{A} \models_X \exists^1 x\phi$ iff $\mathcal{A} \models_{X(a/x)} \phi$ for some $a \in A$.
- $\mathcal{A} \models_X \forall^1 x\phi$ iff $\mathcal{A} \models_{X(a/x)} \phi$ for all $a \in A$.

Both logics have the so-called *empty team property*: if X is empty, then $\mathcal{A} \models_X \phi$ for any $\phi \in \mathsf{FOT}[\tau]$ or $\phi \in \mathsf{FOT}^{\downarrow}[\tau]$. Note that even though FOT does not contain the negation symbol $\neg$, the formula $\neg\delta$ is expressible in FOT using $\subseteq$, $\dot\sim$, and $\vee\!\!\!\vee$, as shown in [19].

2.2. Probabilistic team semantics

Let τ, D, $\mathcal{A}$, and X be as above, with the exception that we assume that X is maximal, i.e. it contains all assignments $s\colon D \to A$. A *probabilistic team* $\mathbb{X}$ is a function $\mathbb{X}\colon X \to \mathbb{R}_{\geq 0}$, where $\mathbb{R}_{\geq 0}$ is the set of non-negative real numbers. The value $\mathbb{X}(s)$ is also called the *weight* of assignment s. We define the *support* of $\mathbb{X}$ as follows:

$$\mathrm{supp}(\mathbb{X}) := \{s \in X \mid \mathbb{X}(s) \neq 0\},$$

and say that the team $\mathbb{X}$ is *nonempty* if $\mathrm{supp}(\mathbb{X}) \neq \varnothing$. Note that even when $D = \varnothing$, the probabilistic team $\mathbb{X}$ may still be nonempty: if $D = \varnothing$, then X is the singleton containing the empty assignment whose weight can be nonzero.

Functions $\mathbb{X}\colon X \to \mathbb{R}_{\geq 0}$ such that $\sum_{s \in X} \mathbb{X}(s) = 1$ are called *probability distributions*. They are an important special case of probabilistic teams and originally probabilistic teams were required to be probability distributions (hence the name *probabilistic* team). If $\mathbb{X}$ is a probability distribution, we also write $\mathbb{X}\colon X \to [0,1]$. Note that from every nonempty probabilistic team $\mathbb{X}\colon X \to \mathbb{R}_{\geq 0}$ team we obtain a probability distribution $\mathrm{distr}(\mathbb{X})\colon X \to [0,1]$ by setting

$$\mathrm{distr}(\mathbb{X})(s) = \frac{1}{\sum_{t \in X} \mathbb{X}(t)} \cdot \mathbb{X}(s)$$

for all $s\colon D \to A$. It does not matter whether we evaluate formulas using the original team or the team that has been scaled in order to obtain a probability distribution (see Proposition 3.1).

By $\mathbb{X}(a/x)$, we denote the probabilistic team such that

$$\mathbb{X}(a/x)(s) = \sum_{\substack{t \in X,\\ t(a/x)=s}} \mathbb{X}(t)$$

for all $s\colon D \cup \{x\} \to A$. Note that if x is a fresh variable (i.e. $x \notin D$), then for all $s \in X$,

$$\mathbb{X}(a/x)(s(b/x)) = \begin{cases} \mathbb{X}(s), & \text{when } b = a \\ 0, & \text{when } b \neq a. \end{cases}$$

3. Logics in probabilistic team semantics

3.1. The logics $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp_{\mathrm{c}}^{\delta})$ and $\mathsf{FOPT}(\leq^{\delta})$

First-order τ-terms and atomic formulas are defined in the usual way. Let δ be a quantifier- and disjunction-free first-order formula as before. The syntax for the logic $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp_{\mathrm{c}}^{\delta})$ over a vocabulary τ is then as follows:

$$\phi ::= \delta \mid \delta \leq \delta \mid \delta \perp\!\!\!\perp_{\delta} \delta \mid \dot\sim \phi \mid (\phi \wedge \phi) \mid (\phi \mathbin{\mathpalette\make@circled\vee} \phi) \mid \exists^1 x\phi \mid \forall^1 x\phi.$$

Atoms of the form $\delta \leq \delta$ and $\delta \perp\!\!\!\perp_{\delta} \delta$ are called *extended probabilistic inclusion* and *extended probabilistic conditional independence* atoms, respectively. For better readability, we sometimes use extra parentheses around these atoms, e.g. we write $\forall^1 x(\delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2)$ instead of $\forall^1 x\delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2$. The fragment of $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})$ without extended probabilistic conditional independence atoms is denoted by $\mathsf{FOPT}(\leq^{\delta})$.

The semantics for first-order formulas δ and connectives $\dot\sim, \wedge, \mathbin{\mathpalette\make@circled\vee}$ is defined as in FOT using $\mathrm{supp}(\mathbb{X})$ as the team. The definitions for the quantifiers are analogous to FOT, except that we use the probabilistic version $\mathbb{X}(a/x)$ of the modified team. We have the following semantics for the new atoms:

- $\mathcal{A} \models_{\mathbb{X}} \delta_0 \leq \delta_1$ iff $\sum_{s\in S_0} \mathbb{X}(s) \leq \sum_{s\in S_1} \mathbb{X}(s)$, where $S_i = \{s \in X \mid \mathcal{A} \models_s \delta_i\}$ for $i = 0, 1$.
- $\mathcal{A} \models_{\mathbb{X}} \delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2$ iff

$$\sum_{s\in S_0\cap S_1} \mathbb{X}(s) \cdot \sum_{s\in S_0\cap S_2} \mathbb{X}(s) = \sum_{s\in S_0} \mathbb{X}(s) \cdot \sum_{s\in S_0\cap S_1\cap S_2} \mathbb{X}(s),$$

  where $S_i = \{s \in X \mid \mathcal{A} \models_s \delta_i\}$ for $i = 0, 1, 2$.

Note that if $\mathbb{X}$ is an empty probabilistic team, then $\mathcal{A} \models_{\mathbb{X}} \phi$ for any $\phi \in \mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})[\tau]$. The following proposition can also be verified using a simple induction:

Proposition 3.1. *Let $\mathbb{X}\colon X \to \mathbb{R}_{\geq 0}$ be a nonempty probabilistic team. Then for any formula $\phi \in \mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})[\tau]$ and any τ-structure $\mathcal{A}$*

$$\mathcal{A} \models_{\mathrm{distr}(\mathbb{X})} \phi \iff \mathcal{A} \models_{\mathbb{X}} \phi.$$

Proposition 3.1 and its proof is similar to one from [12] which considers team-based logics with several different atoms, including marginal identity and probabilistic conditional independence (see also subsection 4.2).

Next, we present a few notions that are needed to formulate the so-called *locality* property. For a formula ϕ, we denote by $\mathrm{Var}(\phi)$ the set of the free variables of ϕ. Let V be a set of variables. We write $s \upharpoonright V$ for the restriction of the assignment s to V. The *restriction of a team* X to V is defined as $X \upharpoonright V = \{s \upharpoonright V \mid s \in X\}$. The *restriction of a probabilistic team* $\mathbb{X}$ to V is defined as $\mathbb{X} \upharpoonright V\colon X \upharpoonright V \to \mathbb{R}_{\geq 0}$ where

$$(\mathbb{X} \upharpoonright V)(s) = \sum_{\substack{s'\upharpoonright V = s,\\ s'\in X}} \mathbb{X}(s').$$

Proposition 3.2 *(**Locality**). Let ϕ be any $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})[\tau]$-formula. Then for any set of variables V, any τ-structure $\mathcal{A}$, and any probabilistic team $\mathbb{X}\colon X \to \mathbb{R}_{\geq 0}$ such that* $\mathrm{Var}(\phi) \subseteq V \subseteq D$,

$$\mathcal{A} \models_{\mathbb{X}} \phi \iff \mathcal{A} \models_{\mathbb{X}\upharpoonright V} \phi.$$

Proof. By induction. If $\phi = \delta$, the claim immediately holds since $\mathcal{A} \models_s \delta \iff \mathcal{A} \models_{s\upharpoonright V} \delta$ for all $s \in X$. The cases $\phi = \theta_0 \wedge \theta_1$ and $\phi = \theta_0 \mathbin{\mathpalette\make@circled\vee} \theta_1$ directly follow from the induction hypothesis.

For the cases $\phi = \delta_0 \leq \delta_1$ and $\phi = \delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2$, we notice that

$$\sum_{s'\in S\upharpoonright V} (\mathbb{X} \upharpoonright V)(s') = \sum_{s'\in S\upharpoonright V} \left(\sum_{\substack{s\upharpoonright V = s',\\ s\in X}} \mathbb{X}(s) \right) = \sum_{s\in S} \mathbb{X}(s),$$

where $S = \{s \in X \mid \mathcal{A} \models_s \delta\}$ and $S \upharpoonright V = \{s' \in X \upharpoonright V \mid \mathcal{A} \models_{s'} \delta\}$ for any δ. Then

$$\begin{aligned}
\mathcal{A} \models_{\mathbb{X}} \delta_0 \leq \delta_1 &\iff \sum_{s \in S_0} \mathbb{X}(s) \leq \sum_{s \in S_1} \mathbb{X}(s), \text{ where } \ S_i = \{s \in X \mid \mathcal{A} \models_s \delta_i\} \text{ for } i = 0, 1 \\
&\iff \sum_{s' \in S_0 \upharpoonright V} (\mathbb{X} \upharpoonright V)(s') \leq \sum_{s' \in S_1 \upharpoonright V} (\mathbb{X} \upharpoonright V)(s'), \\
&\qquad \text{where } S_i \upharpoonright V = \{s' \in X \upharpoonright V \mid \mathcal{A} \models_{s'} \delta_i\} \text{ for } i = 0, 1 \\
&\iff \mathcal{A} \models_{\mathbb{X} \upharpoonright V} \delta_0 \leq \delta_1.
\end{aligned}$$

The proof is similar for the case $\phi = \delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2$.

If $\phi = \dot\sim \theta$, then

$$\begin{aligned}
\mathcal{A} \models_{\mathbb{X}} \dot\sim \theta &\iff \mathcal{A} \not\models_{\mathbb{X}} \theta \text{ or } \mathrm{supp}(\mathbb{X}) = \varnothing \\
&\iff \mathcal{A} \not\models_{\mathbb{X} \upharpoonright V} \theta \text{ or } \mathrm{supp}(\mathbb{X} \upharpoonright V) = \varnothing \quad \text{(by the induction hypothesis)} \\
&\iff \mathcal{A} \models_{\mathbb{X} \upharpoonright V} \dot\sim \theta.
\end{aligned}$$

If $\phi = Qx\theta$ where $Q \in \{\exists^1, \forall^1\}$, then

$$\begin{aligned}
\mathcal{A} \models_{\mathbb{X}} Qx\theta &\iff \mathcal{A} \models_{\mathbb{X}(a/x)} \theta \text{ for some/all } a \in A \\
&\iff \mathcal{A} \models_{\mathbb{X}(a/x) \upharpoonright (V \cup \{x\})} \theta \text{ for some/all } a \in A \quad \text{(by the induction hypothesis)} \\
&\iff \mathcal{A} \models_{(\mathbb{X} \upharpoonright V)(a/x)} \theta \text{ for some/all } a \in A \quad (\text{since } \mathbb{X}(a/x) \upharpoonright (V \cup \{x\}) = (\mathbb{X} \upharpoonright V)(a/x)) \\
&\iff \mathcal{A} \models_{\mathbb{X} \upharpoonright V} Qx\theta. \quad \square
\end{aligned}$$

The next proposition shows that the quantifier-induced modifications of probabilistic teams can also be viewed as substitution of quantified variables with suitable constants. We use this proposition in the proofs of Proposition 3.4 and Theorem 7.1. Let ϕ be a formula and let $\bar{a} = (a_1, \ldots, a_n)$ be a tuple of elements from A. We denote by $\phi_{(\bar{a}/\bar{x})}$ the formula obtained from ϕ by substituting the free occurrences of variables $\bar{x}$ with constant symbols $\bar{c}$ whose interpretations are the elements $\bar{a}$, i.e., for all $i = 1, \ldots, n$, the constant symbol c_i is interpreted as the element a_i. When using the notation $\phi_{(\bar{a}/\bar{x})}$, we assume that the vocabulary of the model we are considering is complemented with the constant symbols $\bar{c}$ that are interpreted as the elements $\bar{a}$.

Proposition 3.3. *Let ϕ be any $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{\mathrm{c}})[\tau]$-formula. Then for any τ-structure $\mathcal{A}$, any probabilistic team $\mathbb{X}$, any tuple of variables $\bar{x}$, and any sequence $\bar{a} \in A^{|\bar{x}|}$*

$$\mathcal{A} \models_{\mathbb{X}(\bar{a}/\bar{x})} \phi \iff \mathcal{A} \models_{\mathbb{X}} \phi_{(\bar{a}/\bar{x})}.$$

Proof. If $\phi = \delta$, then

$$\begin{aligned}
\mathcal{A} \models_{\mathbb{X}(\bar{a}/\bar{x})} \delta &\iff \text{for all } s\colon D \cup \mathrm{Var}(\bar{x}) \to A, \text{ if } s \in \mathrm{supp}(\mathbb{X}(\bar{a}/\bar{x})), \text{ then } \mathcal{A} \models_s \delta \\
&\iff \text{for all } s\colon D \cup \mathrm{Var}(\bar{x}) \to A, \text{ if } s \in \mathrm{supp}(\mathbb{X}(\bar{a}/\bar{x})), \text{ then } \mathcal{A} \models_s \delta_{(\bar{a}/\bar{x})} \\
&\qquad (\text{if } s \in \mathrm{supp}(\mathbb{X}(\bar{a}/\bar{x})), \text{ then } s(\bar{x}) = \bar{a}) \\
&\iff \mathcal{A} \models_{\mathbb{X}(\bar{a}/\bar{x})} \delta_{(\bar{a}/\bar{x})} \\
&\iff \mathcal{A} \models_{\mathbb{X}} \delta_{(\bar{a}/\bar{x})} \quad (\text{by locality since } \mathbb{X}(\bar{a}/\bar{x}) \upharpoonright (D \backslash \mathrm{Var}(\bar{x})) = \mathbb{X} \upharpoonright (D \backslash \mathrm{Var}(\bar{x}))).
\end{aligned}$$

For the cases $\phi = \delta_0 \leq \delta_1$ and $\phi = \delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2$, we notice that

$$\sum_{s\in S} \mathbb{X}(\bar{a}/\bar{x})(s) = \sum_{s'\in S'} \mathbb{X}(\bar{a}/\bar{x})(s'),$$

where $S = \{s\colon D \cup \mathrm{Var}(\bar{x}) \to A \mid \mathcal{A} \models_s \delta\}$ and $S' = \{s'\colon D \cup \mathrm{Var}(\bar{x}) \to A \mid \mathcal{A} \models_{s'} \delta_{(\bar{a}/\bar{x})}\}$ for any δ. For this, first note that if $s(\bar{x}) \neq \bar{a}$, then $\mathbb{X}(\bar{a}/\bar{x})(s) = 0$. Therefore, only those assignments s for which $s(\bar{x}) = \bar{a}$ may contribute to the sums. For those assignments s, clearly $\mathcal{A} \models_s \delta \iff \mathcal{A} \models_s \delta_{(\bar{a}/\bar{x})}$, and therefore $\sum_{s\in S} \mathbb{X}(\bar{a}/\bar{x})(s) = \sum_{s'\in S'} \mathbb{X}(\bar{a}/\bar{x})(s')$. With this, it is straightforward to check that the claim holds for the cases $\phi = \delta_0 \leq \delta_1$ and $\phi = \delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2$.

If $\phi = \dot\sim\theta$, then

$$\begin{aligned}
\mathcal{A} \models_{\mathbb{X}(\bar{a}/\bar{x})} \dot\sim\theta &\iff \mathcal{A} \not\models_{\mathbb{X}(\bar{a}/\bar{x})} \theta \text{ or } \mathrm{supp}(\mathbb{X}(\bar{a}/\bar{x})) = \varnothing \\
&\iff \mathcal{A} \not\models_{\mathbb{X}} \theta_{(\bar{a}/\bar{x})} \text{ or } \mathrm{supp}(\mathbb{X}) = \varnothing \quad \text{(by the induction hypothesis)} \\
&\iff \mathcal{A} \models_{\mathbb{X}} \dot\sim\theta_{(\bar{a}/\bar{x})}.
\end{aligned}$$

The proofs for the cases $\phi = \theta_0 \wedge \theta_1$ and $\phi = \theta_0 \vee\!\!\!\vee\, \theta_1$ directly follow from the induction hypothesis.

If $\phi = Qy\theta$ where $Q \in \{\exists^1, \forall^1\}$, then

$$\begin{aligned}
\mathcal{A} \models_{\mathbb{X}(\bar{a}/\bar{x})} Qy\theta &\iff \mathcal{A} \models_{\mathbb{X}(\bar{a}b/\bar{x}y)} \theta \text{ for some/all } b \in A \\
&\iff \mathcal{A} \models_{\mathbb{X}} \theta_{(\bar{a}b/\bar{x}y)} \text{ for some/all } b \in A \quad \text{(by the induction hypothesis)} \\
&\iff \mathcal{A} \models_{\mathbb{X}(b/y)} \theta_{(\bar{a}/\bar{x})} \text{ for some/all } b \in A \quad \text{(by the induction hypothesis)} \\
&\iff \mathcal{A} \models_{\mathbb{X}} Qy\theta_{(\bar{a}/\bar{x})}. \quad \square
\end{aligned}$$

The next proposition shows that we can rename quantified variables in the formulas. This is used in the proofs of Theorems 5.1 and 7.1, where we assume that certain variables have no bounded occurrences in the formulas. We introduce a notation that is analogous to $\phi_{(\bar{a}/\bar{x})}$: we write $\phi_{(\bar{y}/\bar{x})}$ for the formula where, instead of the constant symbols, we substitute $\bar{x}$ with the variables $\bar{y}$.

Proposition 3.4. *Let θ be any $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp_{\mathrm{c}}^{\delta})[\tau]$-formula with free variables from $\{v_1, \ldots, v_k\}$. Suppose that x does not appear in θ. Then for any τ-structure $\mathcal{A}$, any probabilistic team $\mathbb{X}$ over $\{v_1, \ldots, v_k\}$, any $Q \in \{\exists^1, \forall^1\}$, and any $w \in \{v_1, \ldots, v_k\}$*

$$\mathcal{A} \models_{\mathbb{X}} Qw\theta \iff \mathcal{A} \models_{\mathbb{X}} Qx\theta_{(x/w)}.$$

Proof. Let $\mathbb{X}_{x/w}\colon X_{x/w} \to A$ be the probabilistic team such that $X_{x/w} = \{s' \mid s \in X\}$ is the team over $(\{v_1, \ldots, v_k\}\setminus\{w\}) \cup \{x\}$ where $s'(v_i) = s(v_i)$ when $v_i \neq w$, $s'(x) = s(w)$, and $\mathbb{X}_{x/w}(s') = \mathbb{X}(s)$. Thus the probabilistic team $\mathbb{X}_{x/w}$ is otherwise the same as the team $\mathbb{X}$ but the variable w is replaced with x. Now we have

$$\begin{aligned}
\mathcal{A} \models_{\mathbb{X}} Q\bar{w}\theta &\iff \mathcal{A} \models_{\mathbb{X}(a/w)} \theta \quad \text{for some/all } a \in A \\
&\iff \mathcal{A} \models_{\mathbb{X}(a/w)_{x/w}} \theta_{(x/w)} \quad \text{for some/all } a \in A \\
&\iff \mathcal{A} \models_{\mathbb{X}_{x/w}(a/x)} \theta_{(x/w)} \quad \text{for some/all } a \in A \\
&\iff \mathcal{A} \models_{\mathbb{X}_{x/w}} \theta_{(x/w)(a/x)} \quad \text{for some/all } a \in A \quad \text{(by Proposition 3.3)} \\
&\iff \mathcal{A} \models_{\mathbb{X}} \theta_{(x/w)(a/x)} \quad \text{for some/all } a \in A \quad \text{(by locality since} \\
&\qquad \mathbb{X}_{x/w}\upharpoonright(\mathrm{Var}(\bar{v})\setminus\{w\}) = \mathbb{X}\upharpoonright(\mathrm{Var}(\bar{v})\setminus\{w\})\ )
\end{aligned}$$

$$\iff \mathcal{A} \models_{\mathbb{X}(a/x)} \theta_{(x/w)} \quad \text{for some/all } a \in A \quad \text{(by Proposition 3.3)}$$
$$\iff \mathcal{A} \models_{\mathbb{X}} Qx\theta_{(x/w)}. \quad \square$$

3.2. The logic $\mathsf{FOPT}(\leq^{\delta}_{c})$

Next, we introduce a logic similar to $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{c})$. The difference is that, instead of the extended probabilistic inclusion and extended probabilistic conditional independence atoms, we have atoms of the form $(\delta_0|\delta_1) \leq (\delta_2|\delta_3)$, where each δ_i is as quantifier- and disjunction-free first-order formula. We call these *conditional probability inequality* atoms. The syntax for the logic $\mathsf{FOPT}(\leq^{\delta}_{c})$ over a vocabulary τ is as follows:

$$\phi ::= \delta \mid (\delta|\delta) \leq (\delta|\delta) \mid \dot\sim \phi \mid (\phi \wedge \phi) \mid (\phi \vee\!\!\!\vee \phi) \mid \exists^1 x\phi \mid \forall^1 x\phi.$$

The semantics for the atom $(\delta_0|\delta_1) \leq (\delta_2|\delta_3)$ is defined as follows:

$$\mathcal{A} \models_{\mathbb{X}} (\delta_0|\delta_1) \leq (\delta_2|\delta_3) \iff \sum_{s\in S_0\cap S_1} \mathbb{X}(s) \cdot \sum_{s\in S_3} \mathbb{X}(s) \leq \sum_{s\in S_2\cap S_3} \mathbb{X}(s) \cdot \sum_{s\in S_1} \mathbb{X}(s)$$

where $S_i = \{s \in X \mid \mathcal{A} \models_s \delta_i\}$ for $i = 0, 1, 2, 3$. Extended probabilistic inclusion and extended probabilistic conditional independence can be expressed in $\mathsf{FOPT}(\leq^{\delta}_{c})$. Suppose that $\delta_0, \delta_1, \delta_2$ are formulas with free variables from $\bar{x} = (x_1, \ldots, x_n)$. It is easy to check that

$$\delta_0 \leq \delta_1 \equiv (\delta_0|x_1 = x_1) \leq (\delta_1|x_1 = x_1)$$

and

$$\delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2 \equiv (\delta_1|\delta_0) \approx (\delta_1|\delta_0 \wedge \delta_2),$$

where $(\delta_1|\delta_0) \approx (\delta_1|\delta_0 \wedge \delta_2)$ is an abbreviation for the formula $(\delta_1|\delta_0) \leq (\delta_1|\delta_0 \wedge \delta_2) \wedge (\delta_1|\delta_0 \wedge \delta_2) \leq (\delta_1|\delta_0)$.

Note that $\mathsf{FOPT}(\leq^{\delta}_{c})$ is local since the proof of Proposition 3.2 can easily be extended to cover atoms of the form $(\delta_0|\delta_1) \leq (\delta_2|\delta_3)$. Moreover, proofs for Propositions 3.1, 3.3, and 3.4 can also be extended for $\mathsf{FOPT}(\leq^{\delta}_{c})$.

4. Comparison of logics in team semantics

4.1. Expressibility of inclusion atoms in $\mathsf{FOPT}(\leq^{\delta})$

The following proposition shows that FOT-formulas can be translated into $\mathsf{FOPT}(\leq^{\delta})$ by demonstrating how inclusion atoms can be expressed with $\neg$, $\dot\sim$, and $\vee\!\!\!\vee$.

Proposition 4.1. *Let ϕ be any $\mathsf{FOT}[\tau]$-formula. Then there exists an $\mathsf{FOPT}(\leq^{\delta})[\tau]$-formula ψ_ϕ such that for any τ-structure $\mathcal{A}$, and any probabilistic team $\mathbb{X}$*

$$\mathcal{A} \models_{\mathrm{supp}(\mathbb{X})} \phi \iff \mathcal{A} \models_{\mathbb{X}} \psi_\phi.$$

Proof. Notice that only inclusion atoms, i.e. atoms of the form $\bar{v}_0 \subseteq \bar{v}_1$ need to be translated. For each formula ϕ, we let ψ_ϕ be the same as ϕ, except that each inclusion atom θ appearing in ϕ is substituted with the formula ψ_θ as described below. Provided that we can successfully translate each θ, it is easy to check that the claim holds. If $\theta = \bar{v}_0 \subseteq \bar{v}_1$, then we let $\psi_\theta := \forall^1\bar{x}(\neg\bar{v}_0 = \bar{x} \vee\!\!\!\vee \dot\sim\neg\bar{v}_1 = \bar{x})$. We show that the claim holds for θ and ψ_θ.

If $\mathrm{supp}(\mathbb{X}) = \varnothing$, then both $\mathbb{X}$ and $\mathrm{supp}(\mathbb{X})$ satisfy every formula. Thus, without loss of generality, we may assume that $\mathrm{supp}(\mathbb{X}) \neq \varnothing$. Now

$$\begin{aligned}
\mathcal{A} \models_{\mathrm{supp}(\mathbb{X})} \bar{v}_0 \subseteq \bar{v}_1 &\iff \text{for all } s \in \mathrm{supp}(\mathbb{X}), \text{ there exists } s' \in \mathrm{supp}(\mathbb{X}) \text{ such that } s(\bar{v}_0) = s'(\bar{v}_1)\\
&\iff \text{for all } \bar{a} \in A^{|\bar{v}_0|}, \text{ if there is } s \in \mathrm{supp}(\mathbb{X}) \text{ such that } s(\bar{v}_0) = \bar{a},\\
&\qquad \text{then there exists } s' \in \mathrm{supp}(\mathbb{X}) \text{ such that } s'(\bar{v}_1) = \bar{a}\\
&\iff \text{for all } \bar{a} \in A^{|\bar{v}_0|}, \ s(\bar{v}_0) \neq \bar{a} \text{ for all } s \in \mathrm{supp}(\mathbb{X})\\
&\qquad \text{or there exists } s' \in \mathrm{supp}(\mathbb{X}) \text{ such that } s'(\bar{v}_1) = \bar{a}\\
&\iff \text{for all } \bar{a} \in A^{|\bar{v}_0|}, \ \mathcal{A} \models_{\mathbb{X}} (\neg\bar{v}_0 = \bar{x})_{(\bar{a}/\bar{x})} \text{ or } \mathcal{A} \not\models_{\mathbb{X}} (\neg\bar{v}_1 = \bar{x})_{(\bar{a}/\bar{x})}\\
&\iff \text{for all } \bar{a} \in A^{|\bar{v}_0|}, \ \mathcal{A} \models_{\mathbb{X}} (\neg\bar{v}_0 = \bar{x} \varovee \dot{\sim}\neg\bar{v}_1 = \bar{x})_{(\bar{a}/\bar{x})} \quad (\text{since } \mathrm{supp}(\mathbb{X}) \neq \varnothing)\\
&\iff \text{for all } \bar{a} \in A^{|\bar{v}_0|}, \ \mathcal{A} \models_{\mathbb{X}(\bar{a}/\bar{x})} \neg\bar{v}_0 = \bar{x} \varovee \dot{\sim}\neg\bar{v}_1 = \bar{x} \quad (\text{by Proposition 3.3})\\
&\iff \mathcal{A} \models_{\mathbb{X}} \forall^1\bar{x}(\neg\bar{v}_0 = \bar{x} \varovee \dot{\sim}\neg\bar{v}_1 = \bar{x}). \quad \square
\end{aligned}$$

4.2. Expressibility of marginal identity and probabilistic conditional independence atoms in $\mathsf{FOPT}(\leq^\delta, \perp\!\!\!\perp^\delta_{\mathrm{c}})$

The logics in probabilistic team semantics often include the *marginal identity atom* $\bar{v}_0 \approx \bar{v}_1$ and the *probabilistic conditional independence atom* $\bar{v}_1 \perp\!\!\!\perp_{\bar{v}_0} \bar{v}_2$ where $\bar{v}_0, \bar{v}_1$ and $\bar{v}_2$ are tuples of variables, instead of formulas. (See e.g. [12].) In the case of the marginal identity atom, we additionally require that the tuples $\bar{v}_0$ and $\bar{v}_1$ are of the same length. Let $\bar{x}$ be a tuple of variables and $\bar{a} \in A^{|\bar{x}|}$, and define

$$|\mathbb{X}_{\bar{x}=\bar{a}}| := \sum_{\substack{s \in X,\\ s(\bar{x})=\bar{a}}} \mathbb{X}(s).$$

The semantics for the marginal identity atom and the probabilistic conditional independence atom is as follows:

- $\mathcal{A} \models_{\mathbb{X}} \bar{v}_0 \approx \bar{v}_1$ iff $|\mathbb{X}_{\bar{v}_0=\bar{a}}| = |\mathbb{X}_{\bar{v}_1=\bar{a}}|$ for all $\bar{a} \in A^{|\bar{v}_0|}$.
- $\mathcal{A} \models_{\mathbb{X}} \bar{v}_1 \perp\!\!\!\perp_{\bar{v}_0} \bar{v}_2$ iff for all $s\colon \mathrm{Var}(\bar{v}_0\bar{v}_1\bar{v}_2) \to A$,

$$|\mathbb{X}_{\bar{v}_0\bar{v}_1=s(\bar{v}_0\bar{v}_1)}| \cdot |\mathbb{X}_{\bar{v}_0\bar{v}_2=s(\bar{v}_0\bar{v}_2)}| = |\mathbb{X}_{\bar{v}_0=s(\bar{v}_0)}| \cdot |\mathbb{X}_{\bar{v}_0\bar{v}_1\bar{v}_2=s(\bar{v}_0\bar{v}_1\bar{v}_2)}|.$$

Note that we do not require that the tuples $\bar{v}_0$, $\bar{v}_1$, $\bar{v}_2$ are disjoint.

We show that the atoms of the form $\delta_0 \leq \delta_1$ and $\delta_1 \perp\!\!\!\perp_{\delta_0} \delta_2$ extend these in the sense that, when the weak universal quantifier $\forall^1$ is available, $\bar{v}_0 \approx \bar{v}_1$ and $\bar{v}_1 \perp\!\!\!\perp_{\bar{v}_0} \bar{v}_2$ are also expressible. For probabilistic conditional independence, the equivalent formula of $\mathsf{FOPT}(\leq^\delta, \perp\!\!\!\perp^\delta_{\mathrm{c}})$ is straightforward:

$$\bar{v}_1 \perp\!\!\!\perp_{\bar{v}_0} \bar{v}_2 \equiv \forall^1\bar{x}\bar{y}\bar{z}(\bar{v}_1 = \bar{y} \perp\!\!\!\perp_{\bar{v}_0=\bar{x}} \bar{v}_2 = \bar{z}).$$

For the marginal identity atom, it feels natural to first define a new kind of formula $\delta_0 \approx \delta_1 := \delta_0 \leq \delta_1 \wedge \delta_1 \leq \delta_0$, and use that to obtain that

$$\bar{v}_0 \approx \bar{v}_1 \equiv \forall^1\bar{x}(\bar{v}_0 = \bar{x} \approx \bar{v}_1 = \bar{x}).$$

However, there is also a shorter formula for the marginal identity atom:

$$\bar{v}_0 \approx \bar{v}_1 \equiv \forall^1\bar{x}(\bar{v}_0 = \bar{x} \leq \bar{v}_1 = \bar{x}).$$

To see that this formula suffices, note that since A is finite,

$$|\mathbb{X}_{\bar{v}_0=\bar{a}}| \leq |\mathbb{X}_{\bar{v}_1=\bar{a}}| \text{ for all } \bar{a} \in A^{|\bar{v}_0|}$$

implies that

$$|\mathbb{X}_{\bar{v}_0=\bar{a}}| = |\mathbb{X}_{\bar{v}_1=\bar{a}}| \text{ for all } \bar{a} \in A^{|\bar{v}_0|}.$$

Because of this, marginal identity atoms were originally (in [3]) called *probabilistic inclusion atoms* and denoted by $\bar{v}_0 \leq \bar{v}_1$. Instead of defining the formula $\delta_0 \approx \delta_1$ as we have done above, we could also treat it as a new kind of atomic formula. Then the atoms of the form $\delta_0 \leq \delta_1$ and $\delta_0 \approx \delta_1$ can be seen as extended probabilistic inclusion and extended marginal identity atoms, respectively. However, even though the truth definitions for $\bar{v}_0 \leq \bar{v}_1$ and $\bar{v}_0 \approx \bar{v}_1$ are equivalent, this is not the case for $\delta_0 \leq \delta_1$ and $\delta_0 \approx \delta_1$.

5. Translation from $\mathsf{FOPT}(\leq^{\delta}_{c})$ to real arithmetic

In this section, we show that the satisfiability and validity problems for $\mathsf{FOPT}(\leq^{\delta}_{c})$ are r.e.-complete and co-r.e.-complete, respectively. Note that the definitions of our logics assume that the structure $\mathcal{A}$ is finite, so the satisfiability and validity for $\mathsf{FOPT}(\leq^{\delta}_{c})$ are only considered over finite structures. The main ingredient of the proof is constructing a translation from $\mathsf{FOPT}(\leq^{\delta}_{c})$ to real arithmetic.

We say that a τ-formula $\phi \in \mathsf{FOPT}(\leq^{\delta}_{c})$ is *satisfiable in a τ-structure* $\mathcal{A}$ if there exists a nonempty probabilistic team $\mathbb{X}$ of $\mathcal{A}$ such that $\mathcal{A} \models_{\mathbb{X}} \phi$. Analogously, ϕ is *valid in* $\mathcal{A}$ if $\mathcal{A} \models_{\mathbb{X}} \phi$ for all probabilistic teams $\mathbb{X}$ of $\mathcal{A}$ over $\mathrm{Var}(\phi)$. A τ-formula $\phi \in \mathsf{FOPT}(\leq^{\delta}_{c})$ is *satisfiable* if there exists a τ-structure $\mathcal{A}$ such that ϕ is satisfiable in $\mathcal{A}$. A τ-formula $\phi \in \mathsf{FOPT}(\leq^{\delta}_{c})$ is *valid* if ϕ is valid in $\mathcal{A}$ for all a τ-structures $\mathcal{A}$.

Theorem 5.1. *Let τ be a finite relational vocabulary, and $\mathcal{A}$ a finite τ-structure.*

(i) *For each τ-formula ϕ from $\mathsf{FOPT}(\leq^{\delta})$ there exists a first-order sentence ψ over vocabulary $\{+, \leq, 0\}$ such that ϕ is satisfiable in $\mathcal{A}$ iff $(\mathbb{R}, +, \leq, 0) \models \psi$.*
(ii) *For each τ-formula ϕ from $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{c})$ or $\mathsf{FOPT}(\leq^{\delta}_{c})$ there exists a first-order sentence ψ over vocabulary $\{+, \times, \leq, 0, 1\}$ such that ϕ is satisfiable in $\mathcal{A}$ iff $(\mathbb{R}, +, \times, \leq, 0, 1) \models \psi$.*

Proof. Without loss of generality, we may assume that $A = \{1, \ldots, n\}$. Let $\bar{v} = (v_1, \ldots, v_m)$ be a tuple that consists of the first-order variables that appear free in ϕ. Since $\mathsf{FOPT}(\leq^{\delta})$, $\mathsf{FOPT}(\leq^{\delta}, \perp\!\!\!\perp^{\delta}_{c})$, and $\mathsf{FOPT}(\leq^{\delta}_{c})$ are local, it suffices to consider teams over $\{v_1, \ldots, v_m\}$. Moreover, by Proposition 3.4, it suffices to only consider formulas $\phi(\bar{v})$ in which there are no bound occurrences of the variables $\bar{v}$. For the tuple $\bar{v}$, we will need a fresh first-order variable $s_{\bar{v}=\bar{a}}$ for each $\bar{a} \in A^m$. Each variable $s_{\bar{v}=\bar{a}}$ will correspond to the weight of the assignment that interprets variables $\bar{v}$ as elements $\bar{a}$. By $\bar{s}$, we denote the tuple $(s_{\bar{v}=\bar{1}}, \ldots, s_{\bar{v}=\bar{n}})$ that contains all these variables. Now we let

$$\psi := \exists s_{\bar{v}=\bar{1}} \ldots s_{\bar{v}=\bar{n}} \left(\bigwedge_{\bar{a}} 0 \leq s_{\bar{v}=\bar{a}} \wedge \neg 0 = \sum_{\bar{a}} s_{\bar{v}=\bar{a}} \wedge \phi^*(\bar{s}) \right),$$

where $\phi^*(\bar{s})$ is defined inductively as follows:

- If $\phi(\bar{v}) = \delta$, then $\phi^*(\bar{s}) := \bigwedge_{s \in S} s = 0$, where $S = \{s \in \{s_{\bar{v}=\bar{1}}, \ldots, s_{\bar{v}=\bar{n}}\} \mid \mathcal{A} \not\models_s \delta\}$.
- If $\phi(\bar{v}) = \delta_0(\bar{v}) \leq \delta_1(\bar{v})$, then for $S_i = \{s \mid \mathcal{A} \models_s \delta_i\}$, $i = 0, 1$,

$$\phi^*(\bar{s}) := \sum_{s \in S_0} s \leq \sum_{s \in S_1} s.$$

- If $\phi(\bar{v}) = (\delta_0(\bar{v}) \mid \delta_1(\bar{v})) \leq (\delta_2(\bar{v}) \mid \delta_3(\bar{v}))$, then for $S_i = \{s \mid \mathcal{A} \models_s \delta_i\}$, $i = 0, 1, 2, 3$,

$$\phi^*(\bar{s}) := \sum_{s \in S_0 \cap S_1} s \times \sum_{s \in S_3} s \leq \sum_{s \in S_2 \cap S_3} s \times \sum_{s \in S_1} s.$$

- If $\phi(\bar{v})$ is $\dot\sim \theta_0(\bar{v})$, $\theta_0(\bar{v}) \wedge \theta_1(\bar{v})$ or $\theta_0(\bar{v}) \mathbin{\dot\vee} \theta_1(\bar{v})$, then $\phi^*(\bar{s})$ is defined as $\neg\theta_0^*(\bar{s})$, $\theta_0^*(\bar{s}) \wedge \theta_1^*(\bar{s})$ or $\theta_0^*(\bar{s}) \vee \theta_1^*(\bar{s})$, respectively.
- If $\phi(\bar{v}) = \exists x \theta_0(\bar{v}, x)$, then

$$\phi^*(\bar{s}) := \exists t_{\bar{v}x=\bar{1}1} \dots t_{\bar{v}x=\bar{n}n} \Big(\bigvee_b \bigwedge_{\bar{a}} (t_{\bar{v}x=\bar{a}b} = s_{\bar{v}=\bar{a}} \wedge \bigwedge_{c \neq b} t_{\bar{v}x=\bar{a}c} = 0) \wedge \theta_0^*(\bar{t}) \Big).$$

- If $\phi(\bar{v}) = \forall x \theta_0(\bar{v}, x)$, then

$$\phi^*(\bar{s}) := \bigwedge_b \Big(\exists t_{\bar{v}x=\bar{1}1} \dots t_{\bar{v}x=\bar{n}n} \Big(\bigwedge_{\bar{a}} (t_{\bar{v}x=\bar{a}b} = s_{\bar{v}=\bar{a}} \wedge \bigwedge_{c \neq b} t_{\bar{v}x=\bar{a}c} = 0) \wedge \theta_0^*(\bar{t}) \Big) \Big). \quad \square$$

Theorem 5.2. *The satisfiability problem for* $\mathsf{FOPT}(\leq_c^\delta)$ *is r.e.-complete.*

Proof. Inclusion: Suppose that $\phi \in \mathsf{FOPT}(\leq_c^\delta)[\tau]$ is satisfiable. Let $\mathcal{A}$ be any finite τ-structure. By Theorem 5.1, we can construct a sentence $\psi_{\mathcal{A},\phi}$ such that ϕ is satisfiable in $\mathcal{A}$ iff $(\mathbb{R}, +, \times, \leq, 0, 1) \models \psi_{\mathcal{A},\phi}$. Note that the sentence $\psi_{\mathcal{A},\phi}$ is computable since $\mathcal{A} \models_s \delta$ is decidable when structure $\mathcal{A}$, assignment s, and formula δ are given. Since truth in real arithmetic is decidable, given a structure $\mathcal{A}$, we can also decide whether ϕ is satisfiable in $\mathcal{A}$. Thus we can verify that ϕ is satisfiable by going through all finite τ-structures until we find a structure $\mathcal{A}$ such that ϕ is satisfiable in $\mathcal{A}$.

Hardness: Denote by $\mathrm{SAT}_{\mathrm{fin}}(\mathsf{FO})$ the finite satisfiability problem for FO. Notice that every first-order sentence is also expressible in $\mathsf{FOPT}(\leq_c^\delta)$, and therefore $\mathrm{SAT}_{\mathrm{fin}}(\mathsf{FO})$ is reducible to the satisfiability problem for $\mathsf{FOPT}(\leq_c^\delta)$. By Trahtenbrot's Theorem, $\mathrm{SAT}_{\mathrm{fin}}(\mathsf{FO})$ is r.e.-hard, and thus the satisfiability problem for $\mathsf{FOPT}(\leq_c^\delta)$ is also r.e.-hard. $\square$

The validity problem for $\mathsf{FOPT}(\leq_c^\delta)$ is co-r.e.-complete because its complement is reducible to the satisfiability problem, and vice versa. Note that these reductions rely on the fact that $\mathsf{FOPT}(\leq_c^\delta)$ has the Boolean negation $\dot\sim$ for which $\mathcal{A} \models_{\mathbb{X}} \dot\sim\phi$ if and only if $\mathcal{A} \not\models_{\mathbb{X}} \phi$. In team semantic setting, the negation $\neg$ behaves differently, e.g., $\mathcal{A} \not\models_{\mathbb{X}} x = y$ does not necessarily imply $\mathcal{A} \models_{\mathbb{X}} \neg x = y$ since $\mathrm{supp}(\mathbb{X})$ may contain both assignment s for with $s(x) = s(y)$ and assignment s' for with $s'(x) \neq s'(y)$.

6. Counterparts of logics in probabilistic team semantics over metafinite structures

In this section, we introduce the logic $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ and its fragment $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$. Later, in Section 7, we will show that $\mathsf{FOPT}(\leq^\delta, \perp\!\!\!\perp_c^\delta)$ can be translated into $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$, but there is no full translation from $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ to $\mathsf{FOPT}(\leq^\delta, \perp\!\!\!\perp_c^\delta)$. However, we will also see that $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$ is equi-expressive with the logic $\mathsf{FOPT}(\leq^\delta)$.

6.1. The logic $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$

Let τ_0, τ_1, and σ be vocabularies such that σ is functional, and $\tau_0 \cap \sigma = \tau_1 \cap \sigma = \varnothing$. A two-sorted structure of vocabulary $\tau_0 \cup \tau_1 \cup \sigma$ is a tuple $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, F)$ where $\mathcal{A}_i$ is a τ_i-structure of domain A_i for $i = 0, 1$, and F is a set that contains functions $f^{\mathcal{A}} \colon A_0^{\mathrm{ar}(f)} \to A_1$ for each function symbol $f \in \sigma$. In this paper, we always assume that the structure $\mathcal{A}_0$ is finite, and both σ and F are finite. For simplicity, τ_0 can only contain relation and constant symbols. Note that $\mathcal{A}_1$ is not assumed to be finite, on the contrary, we consider metafinite structures where $A_1 = \mathbb{R}_{\geq 0}$ or $A_1 = \mathbb{R}$.

Let $\{=\} \subseteq \tau_0$, $\tau_1 = \{\leq\}$, and $\sigma = \{f\}$. Consider structures $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, F)$ where $\mathcal{A}_0$ is a finite τ_0-structure, $\mathcal{A}_1 = (\mathbb{R}_{\geq 0}, \leq)$, and $F = \{f^{\mathcal{A}}\}$ for some $f^{\mathcal{A}} \colon A_0 \to \mathbb{R}_{\geq 0}$. These structures are called $\mathbb{R}_{\geq 0}$-structures.

First-order τ_0-terms and atomic formulas are constructed in the usual way. In addition to the usual τ_0-terms, there are *numerical* $\tau_0 \cup \sigma$-terms i which are as follows:

$$i ::= f(\bar{y}) \mid i \times i \mid \mathrm{SUM}_{\bar{x}}(i, \gamma),$$

where $\bar{x}$ and $\bar{y}$ are tuples of variables, $|\bar{y}| = \mathrm{ar}(f)$, and γ is a quantifier-free first-order formula. If $|\bar{x}| = 0$, we denote $\mathrm{SUM}_{\bar{x}}(i, \gamma) = \mathrm{SUM}_{\varnothing}(i, \gamma)$. The syntax for the logic $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ over a vocabulary $\tau_0 \cup \tau_1 \cup \sigma$ is then as follows:

$$\phi ::= \lambda \mid i \leq i \mid \neg\phi \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid \exists x\phi \mid \forall x\phi,$$

where x is a first-order variable.

We now present the semantics for $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$. Let $\mathcal{A}$ be an $\mathbb{R}_{\geq 0}$-structure of a vocabulary $\tau_0 \cup \tau_1 \cup \sigma$. The interpretations of τ_0-terms are defined in the usual way. Note that first-order terms only range over A_0; they cannot take values from $\mathbb{R}_{\geq 0}$. For the numerical terms we have the interpretations $[f(\bar{x})]^{\mathcal{A}}_s := f^{\mathcal{A}}(s(\bar{x}))$,

$$[i \times j]^{\mathcal{A}}_s := [i]^{\mathcal{A}}_s \cdot [j]^{\mathcal{A}}_s,$$

and

$$[\mathrm{SUM}_{\bar{x}}(i, \gamma)]^{\mathcal{A}}_s := \sum_{\bar{a} \in B} [i]^{\mathcal{A}}_{s(\bar{a}/\bar{x})},$$

where $B = \{\bar{a} \in A_0^{|\bar{x}|} \mid \mathcal{A}_0 \models_s \gamma(\bar{a}/\bar{x})\}$. The semantics for $\leq$ is defined in the obvious way, i.e.

$$\mathcal{A} \models_s i \leq j \iff [i]^{\mathcal{A}}_s \leq [j]^{\mathcal{A}}_s.$$

For atomic τ_0-formulas and connectives $\neg$, $\wedge$, $\vee$, $\exists x$, and $\forall x$, we define semantics as in first-order logic.

6.2. *The fragment* $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$

We denote by $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$ the fragment of $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ with the following syntax:

$$\phi ::= \lambda \mid \neg\phi \mid \mathrm{SUM}_{\bar{x}}(f(\bar{y}), \gamma) \leq \mathrm{SUM}_{\bar{x}}(f(\bar{y}), \gamma) \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid \exists x\phi \mid \forall x\phi$$

where λ and γ are as before, and $\bar{x}$ and $\bar{y}$ are tuples of distinct variables such that $\mathrm{Var}(\bar{x}) \subseteq \mathrm{Var}(\bar{y})$ and $|\bar{y}| = \mathrm{ar}(f)$. Note that despite the restricted syntax of the fragment, we can still refer to $f^{\mathcal{A}}(s(\bar{y}))$ (and also the constant 0). For this, we notice that the set $A_0^{|\varnothing|} = A_0^0$ is the singleton containing only the empty tuple, and therefore

$$[\mathrm{SUM}_{\varnothing}(f(\bar{y}), \gamma)]^{\mathcal{A}}_s = \begin{cases} f^{\mathcal{A}}(s(\bar{y})), & \text{when } \mathcal{A} \models_s \gamma \\ 0, & \text{when } \mathcal{A} \not\models_s \gamma. \end{cases}$$

Additionally, we introduce a useful abbreviation

$$i = j := i \leq j \wedge j \leq i,$$

and write $f(\bar{u}) = 0$ for the formula

$$\mathrm{SUM}_{\varnothing}(f(\bar{u}), u_1 = u_1) = \mathrm{SUM}_{\varnothing}(f(\bar{u}), \neg u_1 = u_1),$$

where $\bar{u} = (u_1, \ldots, u_k)$. Note that $[\mathrm{SUM}_{\varnothing}(f(\bar{u}), u_1 = u_1)]_s^{\mathcal{A}} = f^{\mathcal{A}}(s(\bar{u}))$ and $[\mathrm{SUM}_{\varnothing}(f(\bar{u}), \neg u_1 = u_1)]_s^{\mathcal{A}} = 0$, and thus

$$\mathcal{A} \models_s f(\bar{u}) = 0 \iff f^{\mathcal{A}}(s(\bar{u})) = 0$$

as one would expect.

7. Translations and the equi-expressivity result

In this section we show that $\mathsf{FOPT}(\leq_c^{\delta})$ can be translated to $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$, and that $\mathsf{FOPT}(\leq^{\delta})$ and $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$ are equi-expressive. The main idea is to use a function $f_{\mathbb{X}}$ in the metafinite structure to express the weights given by the probabilistic team $\mathbb{X}$.

7.1. Translation from $\mathsf{FOPT}(\leq_c^{\delta})$ *to* $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$

Theorem 7.1. *Let $\phi(v_1, \ldots, v_k)$ be any $\mathsf{FOPT}(\leq_c^{\delta})[\tau_0]$-formula and f a k-ary function symbol. Then there exists an $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})[\tau_0 \cup \{\leq\} \cup \{f\}]$-sentence $\psi_\phi(f)$ such that for any $\mathbb{R}_{\geq 0}$-structure $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \{f_{\mathbb{X}}\})$ and any probabilistic team $\mathbb{X}$ over $\{v_1, \ldots, v_k\}$*

$$\mathcal{A}_0 \models_{\mathbb{X}} \phi(\bar{v}) \iff \mathcal{A} \models \psi_\phi(f),$$

where $f_{\mathbb{X}} \colon A_0^k \to \mathbb{R}_{\geq 0}$ is a function such that $f_{\mathbb{X}}(s(\bar{v})) = \mathbb{X}(s)$ for all $s \in X$.

Proof. We show by induction that for any subformula $\theta(\bar{v}, \bar{x})$ of $\phi(\bar{v})$, there exists an $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})[\tau_0 \cup \{\leq\} \cup \{f\}]$-formula $\psi_\theta(f, \bar{x})$ such that for any $\mathbb{R}_{\geq 0}$-structure $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \{f_{\mathbb{X}}\})$, any probabilistic team $\mathbb{X}$ over $\{v_1, \ldots, v_k\}$, and any sequence $\bar{a} \in A_0^{|\bar{x}|}$

$$\mathcal{A}_0 \models_{\mathbb{X}(\bar{a}/\bar{x})} \theta(\bar{v}, \bar{x}) \iff \mathcal{A} \models \psi_\theta(f, \bar{x})(\bar{a}/\bar{x}),$$

where $f_{\mathbb{X}} \colon A_0^k \to \mathbb{R}_{\geq 0}$ is a function defined as above. Note that by Proposition 3.4, it suffices to only consider formulas $\phi(\bar{v})$ in which there are no bound occurrences of the variables $\bar{v}$. We will also use Proposition 3.3, so that we can evaluate the formulas in the original team $\mathbb{X}$ instead of the modified team $\mathbb{X}(\bar{a}/\bar{x})$.

(1) Suppose that $\theta(\bar{v}, \bar{x}) = \delta(\bar{v}, \bar{x})$. Then let $\psi_\theta(f, \bar{x}) := \forall \bar{u}(f(\bar{u}) = 0 \vee \delta(\bar{u}/\bar{v}, \bar{x}))$. Now

$$\begin{aligned}
\mathcal{A}_0 \models_{\mathbb{X}(\bar{a}/\bar{x})} \delta(\bar{v}, \bar{x}) &\iff \mathcal{A}_0 \models_{\mathbb{X}} \delta(\bar{v}, \bar{x})_{(\bar{a}/\bar{x})} \quad \text{(by Proposition 3.3)} \\
&\iff \text{for all } s \in X,\ \text{if } s \in \mathrm{supp}(\mathbb{X}),\ \text{then } \mathcal{A}_0 \models_s \delta(\bar{v}, \bar{x})_{(\bar{a}/\bar{x})} \\
&\iff \text{for all } \bar{b} \in A_0^k,\ f_{\mathbb{X}}(\bar{b}) = 0 \text{ or } \mathcal{A}_0 \models \delta(\bar{b}/\bar{v}, \bar{x})(\bar{a}/\bar{x}) \\
&\iff \mathcal{A} \models \forall \bar{u}(f(\bar{u}) = 0 \vee \delta(\bar{u}/\bar{v}, \bar{x}))(\bar{a}/\bar{x}).
\end{aligned}$$

(2) Suppose that $\theta(\bar{v},\bar{x}) = (\delta_0|\delta_1) \leq (\delta_2|\delta_3)$. Then let

$$\begin{aligned}\psi_\theta(f,\bar{x}) :=& \mathrm{SUM}_{\bar{u}}(f(\bar{u}),(\delta_0\wedge\delta_1)(\bar{u}/\bar{v},\bar{x})) \times \mathrm{SUM}_{\bar{u}}(f(\bar{u}),\delta_3(\bar{u}/\bar{v},\bar{x})) \leq \\ & \mathrm{SUM}_{\bar{u}}(f(\bar{u}),(\delta_2\wedge\delta_3)(\bar{u}/\bar{v},\bar{x})) \times \mathrm{SUM}_{\bar{u}}(f(\bar{u}),\delta_1(\bar{u}/\bar{v},\bar{x})).\end{aligned}$$

Now

$$\begin{aligned}\mathcal{A}_0 \models_{\mathbb{X}(\bar{a}/\bar{x})} (\delta_0|\delta_1) \leq (\delta_2|\delta_3) &\iff \mathcal{A}_0 \models_{\mathbb{X}} ((\delta_0|\delta_1) \leq (\delta_2|\delta_3))_{(\bar{a}/\bar{x})} \quad \text{(by Proposition 3.3)} \\ &\iff \sum_{s\in S_0\cap S_1}\mathbb{X}(s)\cdot\sum_{s\in S_3}\mathbb{X}(s) \leq \sum_{s\in S_2\cap S_3}\mathbb{X}(s)\cdot\sum_{s\in S_1}\mathbb{X}(s), \\ &\qquad \text{where } S_i = \{s\in X \mid \mathcal{A}_0 \models_s \delta_{i(\bar{a}/\bar{x})}\} \text{ for } i=0,1,2,3 \\ &\iff \sum_{\bar{b}\in B_0\cap B_1} f_{\mathbb{X}}(\bar{b})\cdot\sum_{\bar{b}\in B_3} f_{\mathbb{X}}(\bar{b}) \leq \sum_{\bar{b}\in B_2\cap B_3} f_{\mathbb{X}}(\bar{b})\cdot\sum_{\bar{b}\in B_1} f_{\mathbb{X}}(\bar{b}), \\ &\qquad \text{where } B_i = \{\bar{b}\in A_0^k \mid \mathcal{A}_0 \models \delta_i(\bar{b}/\bar{v},\bar{a}/\bar{x})\} \text{ for } i=0,1,2,3 \\ &\iff \mathcal{A} \models \psi_\theta(f,\bar{x})(\bar{a}/\bar{x}).\end{aligned}$$

(3) Suppose that $\theta(\bar{v},\bar{x}) = \dot{\sim}\theta_0(\bar{v},\bar{x})$, $\theta(\bar{v},\bar{x}) = \theta_0(\bar{v},\bar{x}) \wedge \theta_1(\bar{v},\bar{x})$ or $\theta(\bar{v},\bar{x}) = \theta_0(\bar{v},\bar{x}) \vee \theta_1(\bar{v},\bar{x})$. Then let $\psi_\theta(f,\bar{x}) := \neg\psi_{\theta_0}(f,\bar{x}) \vee \forall\bar{u} f(\bar{u}) = 0$, $\psi_\theta(f,\bar{x}) := \psi_{\theta_0}(f,\bar{x}) \wedge \psi_{\theta_1}(f,\bar{x})$ or $\psi_\theta(f,\bar{x}) := \psi_{\theta_0}(f,\bar{x}) \vee \psi_{\theta_1}(f,\bar{x})$, respectively. The claims directly follow from Proposition 3.3 and the induction hypothesis.
(4) Suppose that $\theta(\bar{v},\bar{x}) = Q^1 y\theta_0(\bar{v},\bar{x}y)$ where $Q \in \{\exists,\forall\}$. Then let $\psi_\theta(f,\bar{x}) := Qy\psi_{\theta_0}(f,\bar{x}y)$. Now

$$\begin{aligned}\mathcal{A}_0 \models_{\mathbb{X}(\bar{a}/\bar{x})} Q^1 y\theta_0(\bar{v},\bar{x}y) &\iff \mathcal{A}_0 \models_{\mathbb{X}(\bar{a}b/\bar{x}y)} \theta_0(\bar{v},\bar{x}y) \text{ for some/all } b\in A_0 \\ &\iff \mathcal{A} \models \psi_{\theta_0}(f,\bar{x}y)(\bar{a}b/\bar{x}y) \text{ for some/all } b\in A_0 \quad \text{(by the induction hypothesis)} \\ &\iff \mathcal{A} \models Qy\psi_{\theta_0}(f,\bar{x}y)(\bar{a}/\bar{x}). \quad \square\end{aligned}$$

The next theorem shows that the converse does not hold in full generality. We will show that the scaling property of $\mathsf{FOPT}(\leq^\delta_c)$, i.e. Proposition 3.1, fails for $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times,\mathrm{SUM})$.

Theorem 7.2. *Let f be a k-ary function symbol. There exists a sentence $\psi \in \mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times,\mathrm{SUM})[\tau_0 \cup \{\leq\} \cup \{f\}]$ for which there is no formula $\phi_\psi(v_1,\dots,v_k) \in \mathsf{FOPT}(\leq^\delta_c)[\tau_0]$ such that for any $\mathbb{R}_{\geq 0}$-structure $\mathcal{A} = (\mathcal{A}_0,\mathcal{A}_1,\{f_{\mathbb{X}}\})$ and any nonempty probabilistic team $\mathbb{X}$ over $\{v_1,\dots,v_k\}$*

$$\mathcal{A}_0 \models_{\mathbb{X}} \phi_\psi \iff \mathcal{A} \models \psi,$$

where $f_{\mathbb{X}}\colon A_0^k \to \mathbb{R}_{\geq 0}$ is a function such that $f_{\mathbb{X}}(s(\bar{v})) = \mathbb{X}(s)$ for all $s \in X$.

Proof. Let $x, y_1, \cdots, y_k$ be variables such that $k = \mathrm{ar}(f)$, $\bar{y} = (y_1,\cdots,y_k)$, and $x \notin \mathrm{Var}(\bar{y})$. Define

$$i_0 := \mathrm{SUM}_{\bar{y}}(f(\bar{y}), y_1 = y_1) \qquad \text{and} \qquad i_1 := \mathrm{SUM}_x(i_0, x = x).$$

Let $\psi := i_0 \times i_0 \leq i_1$. We show that ψ is as wanted. For a contradiction, suppose that there is an equivalent formula ϕ_ψ. We notice that

$$\begin{aligned}[i_0 \times i_0]^{\mathcal{A}}_s &= [\mathrm{SUM}_{\bar{y}}(f(\bar{y}), y_1 = y_1) \times \mathrm{SUM}_{\bar{y}}(f(\bar{y}), y_1 = y_1)]^{\mathcal{A}}_s \\ &= [\mathrm{SUM}_{\bar{y}}(f(\bar{y}), y_1 = y_1)]^{\mathcal{A}}_s \cdot [\mathrm{SUM}_{\bar{y}}(f(\bar{y}), y_1 = y_1)]^{\mathcal{A}}_s\end{aligned}$$

$$= \sum_{\bar{b}\in A_0^k} f_{\mathbb{X}}(\bar{b}) \cdot \sum_{\bar{b}\in A_0^k} f_{\mathbb{X}}(\bar{b})$$
$$= \sum_{s} \mathbb{X}(s) \cdot \sum_{s} \mathbb{X}(s),$$

and

$$[i_1]^{\mathcal{A}}_s = [\mathrm{SUM}_x(i_0, x = x)]^{\mathcal{A}}_s = \sum_{a\in A_0} [i_0]^{\mathcal{A}}_{s(a/x)} = \sum_{a\in A_0}\sum_{\bar{b}\in A_0^k} f_{\mathbb{X}}(\bar{b}) = |A_0| \cdot \sum_{s} \mathbb{X}(s).$$

Now $\mathcal{A} \models \psi$ if and only if $\sum_s \mathbb{X}(s) \cdot \sum_s \mathbb{X}(s) \leq |A_0| \cdot \sum_s \mathbb{X}(s)$. Since $\mathbb{X}$ is nonempty, we have $\sum_s \mathbb{X}(s) > 0$, and therefore $\mathcal{A} \models \psi$ iff $\sum_s \mathbb{X}(s) \leq |A_0|$. Let $\mathbb{X}$ and A_0 be such that $\sum_s \mathbb{X}(s) > |A_0|$. Then $\mathcal{A} \not\models \psi$, which implies that $\mathcal{A}_0 \not\models_{\mathbb{X}} \phi_\psi$. By Proposition 3.1, we have $\mathcal{A}_0 \not\models_{\mathrm{distr}(\mathbb{X})} \phi_\psi$. Let $\mathcal{A}' = (\mathcal{A}_0, \mathcal{A}_1, \{f_{\mathrm{distr}(\mathbb{X})}\})$. Then also $\mathcal{A}' \not\models \psi$. But now $\sum_s \mathrm{distr}(\mathbb{X})(s) = 1 \leq |A_0|$, which is a contradiction. □

7.2. Equi-expressivity of $\mathsf{FOPT}(\leq^\delta)$ *and* $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$

In this subsection, we show that the logics $\mathsf{FOPT}(\leq^\delta)$ and $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$ are equi-expressive on $\mathbb{R}_{\geq 0}$-structures. Since $\mathsf{FOPT}(\leq^\delta_c)$ subsumes $\mathsf{FOPT}(\leq^\delta)$, the first part, the translation from $\mathsf{FOPT}(\leq^\delta)$ to $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$, almost already follows from the result of the previous subsection; we just have to show that extended probabilistic inclusion atoms can be translated.

Theorem 7.3. *Let $\phi(v_1, \ldots, v_k)$ be any $\mathsf{FOPT}(\leq^\delta)[\tau_0]$-formula and f a k-ary function symbol. Then there exists an $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)[\tau_0 \cup \{\leq\} \cup \{f\}]$-sentence $\psi_\phi(f)$ such that for any $\mathbb{R}_{\geq 0}$-structure $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \{f_{\mathbb{X}}\})$ and any probabilistic team $\mathbb{X}$ over $\{v_1, \ldots, v_k\}$*

$$\mathcal{A}_0 \models_{\mathbb{X}} \phi(\bar{v}) \iff \mathcal{A} \models \psi_\phi(f),$$

where $f_{\mathbb{X}}\colon A_0^k \to \mathbb{R}_{\geq 0}$ is a function such that $f_{\mathbb{X}}(s(\bar{v})) = \mathbb{X}(s)$ for all $s \in X$.

Proof. It suffices to complement the proof of Theorem 7.1 with the case $\theta(\bar{v}, \bar{x}) = \delta_0(\bar{v}, \bar{x}) \leq \delta_1(\bar{v}, \bar{x})$ since the translations of all subformulas, except for the conditional probability inequality, are $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)[\tau_0 \cup \{\leq\} \cup \{f\}]$-sentences.

Suppose that $\theta(\bar{v}, \bar{x}) = \delta_0(\bar{v}, \bar{x}) \leq \delta_1(\bar{v}, \bar{x})$. Then let

$$\psi_\theta(f, \bar{x}) := \mathrm{SUM}_{\bar{u}}(f(\bar{u}), \delta_0(\bar{u}/\bar{v}, \bar{x})) \leq \mathrm{SUM}_{\bar{u}}(f(\bar{u}), \delta_1(\bar{u}/\bar{v}, \bar{x})).$$

Now

$$\begin{aligned}
\mathcal{A}_0 \models_{\mathbb{X}(\bar{a}/\bar{x})} \delta_0 \leq \delta_1 &\iff \mathcal{A}_0 \models_{\mathbb{X}} (\delta_0 \leq \delta_1)_{(\bar{a}/\bar{x})} \\
&\iff \sum_{s\in S_0} \mathbb{X}(s) \leq \sum_{s\in S_1} \mathbb{X}(s), \text{ where } S_i = \{s \in X \mid \mathcal{A}_0 \models_s \delta_{i(\bar{a}/\bar{x})}\} \text{ for } i = 0, 1 \\
&\iff \sum_{\bar{b}\in B_0} f_{\mathbb{X}}(\bar{b}) \leq \sum_{\bar{b}\in B_1} f_{\mathbb{X}}(\bar{b}), \text{ where } B_i = \{\bar{b} \in A_0^k \mid \mathcal{A}_0 \models \delta_i(\bar{b}/\bar{v}, \bar{a}/\bar{x})\} \text{ for } i = 0, 1 \\
&\iff \mathcal{A} \models (\mathrm{SUM}_{\bar{u}}(f(\bar{u}), \delta_0(\bar{u}/\bar{v}, \bar{x})) \leq \mathrm{SUM}_{\bar{u}}(f(\bar{u}), \delta_1(\bar{u}/\bar{v}, \bar{x})))(\bar{a}/\bar{x}). \quad \square
\end{aligned}$$

For the second part, the translation from $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$ to $\mathsf{FOPT}(\leq^\delta)$, we need the following lemma which shows that it suffices to only consider certain kinds of aggregate sums:

Lemma 7.4. *Every aggregate sum term of the logic* $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$ *can be expressed by a term of the form*

$$\mathrm{SUM}_{\bar{u}}(f(\bar{u}), \delta(\bar{u}, \bar{x})),$$

where $\bar{u} = (u_1, \ldots, u_k)$*, and* δ *is a disjunction-free and quantifier-free formula, i.e.* $\delta ::= \lambda \mid \neg\delta \mid \delta \wedge \delta$.

Proof. Consider an aggregate sum of the form $\mathrm{SUM}_{\bar{u}_0}(f(\bar{u}_0\bar{x}_0), \gamma(\bar{u}_0, \bar{x}))$, where $\bar{x}_0$ are among $\bar{x}$, and γ may contain disjunctions. The sum can be expressed by the term

$$\mathrm{SUM}_{\bar{u}_0\bar{u}_1}(f(\bar{u}_0\bar{u}_1), (\gamma^*(\bar{u}_0, \bar{x}) \wedge \bar{u}_1 = \bar{x}_0)),$$

where γ^* is the formula obtained from γ by expressing each disjunction with negation and conjunction in the usual way, i.e. for example, formula $\gamma_0 \vee \gamma_1$ is expressed as $\neg(\neg\gamma_0 \wedge \neg\gamma_1)$. To see this, notice that

$$[\mathrm{SUM}_{\bar{u}_0}(f(\bar{u}_0\bar{x}_0), \gamma(\bar{u}_0, \bar{x}))]_s^{\mathcal{A}} = \sum_{\bar{a}_0 \in B_0} f^{\mathcal{A}}(s(\bar{a}_0/\bar{u}_0)(\bar{u}_0\bar{x}_0)),$$

where $B_0 = \{\bar{a}_0 \in A_0^{|\bar{u}_0|} \mid \mathcal{A}_0 \models_s \gamma(\bar{a}_0/\bar{u}_0)\}$, and

$$[\mathrm{SUM}_{\bar{u}_0\bar{u}_1}(f(\bar{u}_0\bar{u}_1), \gamma^* \wedge \bar{u}_1 = \bar{x}_0)]_s^{\mathcal{A}} = \sum_{\bar{a}_0\bar{a}_1 \in B_{01}} f^{\mathcal{A}}(s(\bar{a}_0\bar{a}_1/\bar{u}_0\bar{u}_1)(\bar{u}_0\bar{u}_1)),$$

where $B_{01} = \{\bar{a}_0\bar{a}_1 \in A_0^{|\bar{u}_0\bar{u}_1|} \mid \mathcal{A}_0 \models_s (\gamma^* \wedge \bar{u}_1 = \bar{x}_0)(\bar{a}_0\bar{a}_1/\bar{u}_0\bar{u}_1)\}$. We then have

$$B_{01} = \{\bar{a}_0 s(\bar{x}_0) \in A_0^{|\bar{u}_0\bar{u}_1|} \mid \mathcal{A}_0 \models_s \gamma(\bar{a}_0/\bar{u}_0)\},$$

from which it follows that

$$\begin{aligned}
[\mathrm{SUM}_{\bar{u}_0\bar{u}_1}(f(\bar{u}_0\bar{u}_1), \gamma^* \wedge \bar{u}_1 = \bar{x}_0)]_s^{\mathcal{A}} &= \sum_{\bar{a}_0\bar{a}_1 \in B_{01}} f^{\mathcal{A}}(s(\bar{a}_0\bar{a}_1/\bar{u}_0\bar{u}_1)(\bar{u}_0\bar{u}_1)) \\
&= \sum_{\bar{a}_0 s(\bar{x}_0) \in B_{01}} f^{\mathcal{A}}(s(\bar{a}_0 s(\bar{x}_0)/\bar{u}_0\bar{u}_1)(\bar{u}_0\bar{u}_1)) \\
&= \sum_{\bar{a}_0 \in B_0} f^{\mathcal{A}}(s(\bar{a}_0/\bar{u}_0)(\bar{u}_0\bar{x}_0)) \\
&= [\mathrm{SUM}_{\bar{u}_0}(f(\bar{u}_0\bar{x}_0), \gamma)]_s^{\mathcal{A}}. \quad \square
\end{aligned}$$

Next, we give a translation from $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$ to $\mathsf{FOPT}(\leq^\delta)$. This is similar to the translation of Theorem 7.1, but simpler, since now we assume that the team $\mathbb{X}$ is nonempty. The assumption is necessary: if the team was empty, all formulas would be satisfied in the team side.

Theorem 7.5. *Let* $\psi(f)$ *be any* $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)[\tau_0 \cup \{\leq\} \cup \{f\}]$*-sentence, where* f *is a* k*-ary function symbol. Then there exists an* $\mathsf{FOPT}(\leq^\delta)[\tau_0]$*-formula* $\phi_\psi(v_1, \ldots, v_k)$ *such that for any* $\mathbb{R}_{\geq 0}$*-structure* $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \{f_{\mathbb{X}}\})$ *and any nonempty probabilistic team* $\mathbb{X}$ *over* $\{v_1, \ldots, v_k\}$

$$\mathcal{A}_0 \models_{\mathbb{X}} \phi_\psi(\bar{v}) \iff \mathcal{A} \models \psi(f),$$

where $f_{\mathbb{X}} \colon A_0^k \to \mathbb{R}_{\geq 0}$ *is a function such that* $f_{\mathbb{X}}(s(\bar{v})) = \mathbb{X}(s)$ *for all* $s \in X$.

Proof. Without loss of generality, we may assume that $\psi(f)$ is in prenex normal form, i.e.

$$\psi(f) = Q_1 x_1 \dots Q_n x_n \theta(f, \bar{x}),$$

where $Q_i \in \{\exists, \forall\}$, $1 \leq i \leq n$, and θ is quantifier free.

We then let $\phi_\psi(\bar{v}) := Q_1^1 x_1 \dots Q_n^1 x_n \phi_\theta(\bar{v}, \bar{x})$, where $\phi_\theta(\bar{v}, \bar{x})$ is defined inductively as follows:

(1) Suppose that $\theta(\bar{x}) = \lambda(\bar{x})$, where λ is a first-order atomic formula of vocabulary τ_0. Then let $\phi_\theta(\bar{v}, \bar{x}) := \lambda(\bar{x})$. The claim follows from Proposition 3.3 and the fact that f does not appear in λ.
(2) By Lemma 7.4, it suffices to consider the case $\theta(f, \bar{x}) = \mathrm{SUM}_{\bar{u}}(f(\bar{u}), \delta_0(\bar{u}, \bar{x})) \leq \mathrm{SUM}_{\bar{u}}(f(\bar{u}), \delta_1(\bar{u}, \bar{x}))$, where δ_i for $i = 0, 1$ is a disjunction-free and quantifier-free formula. Then let $\phi_\theta(\bar{v}, \bar{x}) := \delta_0(\bar{v}/\bar{u}, \bar{x}) \leq \delta_1(\bar{v}/\bar{u}, \bar{x})$. This is similar to the proof of Theorem 7.3, and therefore the proof is not shown here again.
(3) Suppose that $\theta(f, \bar{x}) = \neg\theta_0(f, \bar{x})$, $\theta(\bar{x}) = \theta_0(\bar{x}) \wedge \theta_0(\bar{x})$ or $\theta(\bar{x}) = \theta_0(\bar{x}) \vee \theta_0(\bar{x})$. Then let $\phi_\theta(\bar{v}, \bar{x}) := \dot{\sim}\, \phi_{\theta_0}(\bar{v}, \bar{x})$, $\phi_\theta(\bar{v}, \bar{x}) := \phi_{\theta_0}(\bar{v}, \bar{x}) \wedge \phi_{\theta_1}(\bar{v}, \bar{x})$ or $\phi_\theta(\bar{v}, \bar{x}) := \phi_{\theta_0}(\bar{v}, \bar{x}) \vee\!\!\!\vee\, \phi_{\theta_1}(\bar{v}, \bar{x})$, respectively. The claims directly follow from the induction hypothesis and the fact that $\mathbb{X}(\bar{a}/\bar{x})$ is nonempty.

Now

$$\begin{aligned} \mathcal{A}_0 \models_{\mathbb{X}} Q_1^1 x_1 \dots Q_n^1 x_n \phi_\theta(\bar{v}, \bar{x}) &\iff Q_1 a_1, \dots, Q_n a_n \in A_0,\ \mathcal{A}_0 \models_{\mathbb{X}(\bar{a}/\bar{x})} \phi_\theta(\bar{v}, \bar{x}) \\ &\iff Q_1 a_1, \dots, Q_n a_n \in A_0,\ \mathcal{A} \models \theta(\bar{a}/\bar{x}) \\ &\iff \mathcal{A} \models Q_1 x_1 \dots Q_n x_n \theta(\bar{x}). \quad \square \end{aligned}$$

By combining Theorems 7.3 and 7.5, we obtain that $\mathsf{FOPT}(\leq^\delta)$ and $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\mathrm{SUM}^*)$ are equi-expressive on $\mathbb{R}_{\geq 0}$-structures.

8. Translation from $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ to $\mathsf{FFP}_{\mathbb{R}}$

In this section, we present a translation from $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ to a fragment of $\mathsf{FFP}_{\mathbb{R}}$. The logic $\mathsf{FFP}_{\mathbb{R}}$ was introduced in [10] as a logic for PTIME over reals (w.r.t. ordered structures). It is a fixed point logic with constants for every real number. In the fragment that we consider, the constants are restricted to 0 and 1, and therefore the data complexity of the fragment corresponds to the class $\mathrm{P}^0_{\mathbb{R}}$, i.e., the class of languages over $\mathbb{R}$ decidable in polynomial time by a BSS-machine with restriction to machine constants 0 and 1. The translation gives us an upper bound for the data complexity of $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$. We summarize those definitions from [10] which are needed for the translation; for further details on $\mathsf{FFP}_{\mathbb{R}}$, see [10].

A two-sorted structure $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, F)$ is called an $\mathbb{R}$-structure if

$$\mathcal{A}_1 = \mathcal{R} := (\mathbb{R}, +, -, \times, /, \mathrm{sign}, =, <, 0, 1).$$

We also denote $\tau_{\mathcal{R}} = \{+, -, \times, /, \mathrm{sign}, =, <, 0, 1\}$. In the following, we restrict to functional $\mathbb{R}$-structures or $\mathbb{R}$-*algebras*. These are $\mathbb{R}$-structures $(\mathcal{A}_0, \mathcal{R}, F)$ such that structure $\mathcal{A}_0$ is a plain set A_0, i.e. $\tau_0 = \emptyset$.

We consider a fragment of the *functional fixed-point logic for* $\mathbb{R}$-*algebras*, or $\mathsf{FFP}_{\mathbb{R}}$. First-order τ_0-terms are defined in the usual way. Note that since $\tau_0 = \emptyset$, we only have variables as first-order terms. The syntax of the fragment over a vocabulary $\tau_0 \cup \tau_{\mathcal{R}} \cup \sigma = \tau_{\mathcal{R}} \cup \sigma$ is the set of numerical terms, defined as follows:

$$i ::= c \mid f(\bar{x}) \mid i + i \mid i - i \mid i \times i \mid i/i \mid \mathrm{sgn}(i) \mid \max_{\bar{x}} i(\bar{y}) \mid \mathbf{fp}[Z(\bar{z}) \leftarrow i(Z, \bar{z})](\bar{y})$$

where $c \in \{0, 1\}$, f and Z are function symbols such that $f \in \sigma$ and $Z \notin \sigma$, $\bar{x}, \bar{y}, \bar{z}$ are tuples of distinct variables with $|\bar{x}| = \mathrm{ar}(f)$, $\mathrm{Var}(\bar{x}) \subseteq \mathrm{Var}(\bar{y})$, and $|\bar{y}| = |\bar{z}| = \mathrm{ar}(Z)$.

First-order terms are interpreted in the usual way. Intended interpretations for most of the numerical terms are clear. The problem of division by 0 is handled by letting $[i/j]^{\mathcal{A}}_s := 0$ when $[j]^{\mathcal{A}}_s = 0$. We give interpretations for the non-obvious ones: $\mathrm{sgn}(i)$, $\max_{\bar{x}} i(\bar{y})$, and $\mathbf{fp}[Z(\bar{z}) \leftarrow i(Z,\bar{z})](\bar{y})$. We let

$$[\mathrm{sgn}(i)]^{\mathcal{A}}_s := \begin{cases} 1, & \text{when } [i]^{\mathcal{A}}_s > 0 \\ 0, & \text{when } [i]^{\mathcal{A}}_s = 0 \\ -1, & \text{when } [i]^{\mathcal{A}}_s < 0, \end{cases}$$

and

$$[\max_{\bar{x}} i(\bar{y})]^{\mathcal{A}}_s := \max\{[i(\bar{y})]^{\mathcal{A}}_{s(\bar{a}/\bar{x})} \mid \bar{a} \in A_0^{|\bar{x}|}\}.$$

Because of the terms of the form $\mathbf{fp}[Z(\bar{z}) \leftarrow i(Z,\bar{z})](\bar{y})$, we also allow partially defined functions Z that map tuples from A_0 to $\mathbb{R}$. We define a *partial* $\mathbb{R}$*-algebra* as an $\mathbb{R} \cup \{\text{undef}\}$-algebra obtained by extending the basic operations on $\mathbb{R}$ as follows: if $[j]^{\mathcal{A}}_s = \text{undef}$, then

$$[i+j]^{\mathcal{A}}_s = [i-j]^{\mathcal{A}}_s = \text{undef}, \qquad [\mathrm{sign}(j)]^{\mathcal{A}}_s = \text{undef},$$

and

$$[i \times j]^{\mathcal{A}}_s = [i/j]^{\mathcal{A}}_s = \begin{cases} 0, & \text{when } [i]^{\mathcal{A}}_s = 0 \\ \text{undef}, & \text{when } [i]^{\mathcal{A}}_s \neq 0. \end{cases}$$

Additionally, $[\max_{\bar{x}} i(\bar{y})]^{\mathcal{A}}_s = \text{undef}$, when $[i(\bar{y})]^{\mathcal{A}}_{s(\bar{a}/\bar{x})} = \text{undef}$ for some $\bar{a} \in A_0^{|\bar{x}|}$.

Let $i(Z,\bar{z})$ be a numerical term of vocabulary $\tau_{\mathcal{R}} \cup \{Z\}$. We write $[i(Z,\bar{z})]^{\mathcal{A},Z}_s$ for the interpretation of the term $i(Z,\bar{z})$ in the structure obtained from $\mathcal{A}$ by adding a suitable partial function $Z\colon A_0^{\mathrm{ar}(Z)} \to \mathbb{R}$. The term $i(Z,\bar{z})$ induces an operator $F^{\mathcal{A}}_i$ that updates partially defined functions Z as follows:

$$F^{\mathcal{A}}_i Z(s(\bar{z})) = \begin{cases} [i(Z,\bar{z})]^{\mathcal{A},Z}_s, & \text{when } Z(s(\bar{z})) = \text{undef} \\ Z(s(\bar{z})), & \text{otherwise.} \end{cases}$$

This defines a sequence of partial functions $Z^j\colon A_0^{\mathrm{ar}(Z)} \to \mathbb{R}$ such that

$$\begin{aligned} &Z^0(\bar{a}) = \text{undef} \quad \text{for all } \bar{a} \in A_0^{\mathrm{ar}(Z)} \\ &Z^{j+1} = F^{\mathcal{A}}_i Z^j. \end{aligned}$$

Note that $Z^{j+1} = Z^j$ for some $j \leq |A_0|^{\mathrm{ar}(Z)}$, and after this j, any further iterations do not update the function. We call this Z^j the *fixed point of* $F^{\mathcal{A}}_i$. We let

$$[\mathbf{fp}[Z(\bar{z}) \leftarrow i(Z,\bar{z})](\bar{y})]^{\mathcal{A}}_s = Z^{\infty}(s(\bar{y}))$$

where Z^{∞} is the fixed point of $F^{\mathcal{A}}_i Z$.

A function $E\colon A_0 \to \mathbb{R}$ that is a bijection from A_0 to $\{0,\dots,|A_0|-1\}$ is called a *ranking*. We say that a structure $\mathcal{A}$ is *ranked* if the set F contains a ranking. A given ranking E induces a ranking E_k of k-tuples for any $k > 0$. The ranking E_k is definable, and we will use the abbreviation $\underline{x}$ for $E_k(\bar{x})$ where $\bar{x}$ is a k-tuple of first-order variables.

Let τ_0 be a finite relational vocabulary, and $\mathcal{A}_0$ a finite τ_0-structure. We define the structure $\mathcal{A}_0^*$ as the plain set A_0, and construct the $\mathbb{R}$-algebra $\mathcal{A}^* = (\mathcal{A}_0^*, \mathcal{R}, F)$ of vocabulary $\tau_{\mathcal{R}} \cup \sigma$ by adding to σ

characteristic functions χ_R for all relation symbols $R \in \tau_0$. Let ϕ be a first-order formula of vocabulary τ_0. Then the characteristic function of ϕ, denoted by $\chi[\phi]$, is definable in $\mathsf{FFP}_{\mathbb{R}}[\tau_{\mathcal{R}} \cup \sigma]$. Moreover, if i, j are numerical $\tau_{\mathcal{R}} \cup \sigma$-terms, then functions $\chi[i = j]$ and $\chi[i \leq j]$ are also definable in $\mathsf{FFP}_{\mathbb{R}}[\tau_{\mathcal{R}} \cup \sigma]$. (See [10] or the proof of Theorem 8.1 below.)

The next theorem shows that $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})[\tau_0 \cup \{\leq\} \cup \{f\}]$-formulas can be viewed as functions of $\mathsf{FFP}_{\mathbb{R}}$. Note that the corresponding $\mathsf{FFP}_{\mathbb{R}}$-term will be over $\tau_{\mathcal{R}} \cup \sigma$, a different vocabulary since in $\mathbb{R}$-algebras $\mathcal{A}^*$ each relation $R^{\mathcal{A}} \subseteq A_0^{\mathrm{ar}(R)}$ is replaced with its characteristic function $\chi_R \colon A_0^{\mathrm{ar}(R)} \to \mathbb{R}$.

Theorem 8.1. *Let ϕ be any $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})[\tau_0 \cup \{\leq\} \cup \{f\}]$-formula, and let σ be a vocabulary that contains function symbols E and f, as well as χ_R for all relation symbols $R \in \tau_0$. Then there exists an $\mathsf{FFP}_{\mathbb{R}}[\tau_{\mathcal{R}} \cup \sigma]$-term i_ϕ such that for any $\mathbb{R}_{\geq 0}$-structure $\mathcal{A} = (\mathcal{A}_0, \mathcal{A}_1, \{f^{\mathcal{A}}\})$ and any assignment s*

$$\mathcal{A} \models_s \phi \iff [i_\phi]_s^{\mathcal{A}^*} = 1$$

where $\mathcal{A}^* = (\mathcal{A}_0^*, \mathcal{R}, F)$ is an $\mathbb{R}$-algebra such that structure $\mathcal{A}_0^*$ is the plain set A_0, and F contains a ranking E, the function $f^{\mathcal{A}}$, and the characteristic functions χ_R for all relations $R \in \tau_0$.

Proof. We begin by showing how to translate any numerical $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$-term i of vocabulary $\tau_0 \cup \{\leq\} \cup \{f\}$. We denote by i^* the translation which is a numerical $\mathsf{FFP}_{\mathbb{R}}$-term of vocabulary $\tau_{\mathcal{R}} \cup \sigma$.

(1) If $i = f(\bar{x})$, then $i^* := f(\bar{x})$.
(2) If $i = i_0 \times i_1$, then $i^* := i_0^* \times i_1^*$.
(3) If $i = \mathrm{SUM}_{\bar{x}}(i_0(\bar{y}), \gamma(\bar{y}))$ where $\mathrm{Var}(\bar{x}) \subseteq \mathrm{Var}(\bar{y})$, then

$$i^* := \max_{\bar{x}} \mathbf{fp}[Z(\bar{y}) \leftarrow j(Z, \bar{y})](\bar{y}),$$

where

$$j(Z, \bar{y}) = \chi[\underline{x} = 0] \times i_0^*(\bar{y}) \times \chi[\gamma(\bar{y})] + \max_{\bar{u}} \left(\chi[\underline{x} = \underline{u} + 1] \times (Z(\bar{y}(\bar{u}/\bar{x})) + i_0^*(\bar{y}) \times \chi[\gamma(\bar{y})]) \right).$$

(In the above, $\bar{y}(\bar{u}/\bar{x})$ denotes the tuple obtained from $\bar{y}$ by replacing $\bar{x}$ with $\bar{u}$.)

We continue by defining the corresponding $\mathsf{FFP}_{\mathbb{R}}[\tau_{\mathcal{R}} \cup \sigma]$-terms for formulas ϕ.

(4) Let $\phi = \lambda$, where λ is a first-order atomic formula of vocabulary τ_0. Then $\lambda = R(\bar{x})$ for some $R \in \tau_0$. Now, we let $i_\lambda := \chi_R(\bar{x})$. (Note that R may be the equality relation, so this also covers the case $\lambda = x_0 = x_1$.)
(5) If $\phi = i_0 \leq i_1$, then

$$\begin{aligned} i_\phi &:= \chi[i_0^* = i_1^* \vee i_0^* < i_1^*] \\ &= \chi[i_0^* = i_1^*] + \chi[i_0^* < i_1^*] - \chi[i_0^* = i_1^*] \times \chi[i_0^* < i_1^*] \end{aligned}$$

where

$$\chi[i_0^* = i_1^*] = 1 - [\mathrm{sign}(i_0^* - i_1^*)]^2 \qquad \text{and} \qquad \chi[i_0^* < i_1^*] = ([\mathrm{sign}(i_1^* - i_0^*)]^2 + \mathrm{sign}(i_1^* - i_0^*))/2.$$

(6) If $\phi = \neg\theta_0$, then $i_\phi := 1 - i_{\theta_0}$.
(7) If $\phi = \theta_0 \wedge \theta_1$, then $i_\phi := i_{\theta_0} \times i_{\theta_1}$.
(8) If $\phi = \theta_0 \vee \theta_1$, then $i_\phi := i_{\theta_0} + i_{\theta_1} - i_{\theta_0} \times i_{\theta_1}$.

(9) If $\phi = \exists x\theta_0$, then $i_\phi := \max_x i_{\theta_0}$.
(10) If $\phi = \forall x\theta_0$, then $i_\phi := 1 - \max_x(1 - i_{\theta_0})$. $\square$

The theorem shows that each formula of $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ has a corresponding characteristic function in $\mathsf{FFP}_{\mathbb{R}}$, and therefore the data complexity of $\mathsf{FO}_{\mathbb{R}_{\geq 0}}(\times, \mathrm{SUM})$ is in polynomial time with respect to BSS-computations.

9. Conclusion

We have defined new tractable logics for the framework of probabilistic team semantics that generalize the recently defined logic FOT that is expressively complete for first-order team properties. Our logics employ new probabilistic atoms that resemble so-called extended atoms from the team semantics literature. We also defined counterparts of our logics over metafinite structures and showed that all of our logics can be translated into functional fixed point logic giving a deterministic polynomial-time upper bound for data complexity with respect to BSS-computations.

The following questions remain open:

- What is the exact data complexity of our logics in the BSS-model?
- Is it possible to axiomatize (fragments) of our new logics?

Note that since the validity problem for $\mathsf{FOPT}(\leq^\delta_c)$ is undecidable (see Section 5), the logic cannot be fully axiomatized but, e.g., the axiomatizability of mere probabilistic independence atoms has been studied in several works (see [2] for references).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the anonymous referee for valuable comments.

References

[1] Rafael Albert, Erich Grädel, Unifying hidden-variable problems from quantum mechanics by logics of dependence and independence, CoRR, arXiv:2102.10931 [abs], 2021.
[2] Jukka Corander, Antti Hyttinen, Juha Kontinen, Johan Pensar, Jouko Väänänen, A logical approach to context-specific independence, Ann. Pure Appl. Log. 170 (9) (2019) 975–992.
[3] Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, Jonni Virtema, Approximation and dependence via multi-team semantics, Ann. Math. Artif. Intell. 83 (3–4) (2018) 297–320.
[4] Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, Jonni Virtema, Probabilistic team semantics, in: Foundations of Information and Knowledge Systems - 10th International Symposium, FoIKS 2018, Budapest, Hungary, May 14-18, 2018, Proceedings, 2018, pp. 186–206.
[5] Arnaud Durand, Juha Kontinen, Nicolas de Rugy-Altherre, Jouko Väänänen, Tractability frontier of data complexity in team semantics, in: Javier Esparza, Enrico Tronci (Eds.), Proceedings Sixth International Symposium on Games, Automata, Logics and Formal Verification, GandALF 2015, Genoa, Italy, 21-22nd September 2015, in: EPTCS, vol. 193, 2015, pp. 73–85.
[6] H.B. Enderton, Finite partially-ordered quantifiers, Z. Math. Log. Grundl. Math. 16 (1970) 393–397.
[7] Pietro Galliani, Inclusion and exclusion in team semantics: on some logics of imperfect information, Ann. Pure Appl. Log. 163 (1) (January 2012) 68–84.
[8] Pietro Galliani, On strongly first-order dependencies, Springer International Publishing, Cham, 2016, pp. 53–71.
[9] Erich Grädel, Yuri Gurevich, Metafinite model theory, Inf. Comput. 140 (1) (1998) 26–81.

[10] Erich Grädel, Klaus Meer, Descriptive complexity theory over the real numbers, in: Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing, STOC '95, Association for Computing Machinery, New York, NY, USA, 1995, pp. 315–324.

[11] Erich Grädel, Jouko Väänänen, Dependence and independence, Stud. Log. 101 (2) (April 2013) 399–410.

[12] Miika Hannula, Åsa Hirvonen, Juha Kontinen, Vadim Kulikov, Jonni Virtema, Facets of distribution identities in probabilistic team semantics, in: JELIA, in: Lecture Notes in Computer Science, vol. 11468, Springer, 2019, pp. 304–320.

[13] Miika Hannula, Juha Kontinen, A finite axiomatization of conditional independence and inclusion dependencies, Inf. Comput. 249 (2016) 121–137.

[14] Miika Hannula, Juha Kontinen, Jan Van den Bussche, Jonni Virtema, Descriptive complexity of real computation and probabilistic independence logic, in: Proceedings of the Thirty-Fifth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), IEEE Computer Society, United States, jul 2020, pp. 550–563. Thirty-Fifth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), LICS 2020; Conference date: 08-07-2020 Through 11-07-2020.

[15] Miika Hannula, Jonni Virtema, Tractability frontiers in probabilistic team semantics and existential second-order logic over the reals, CoRR, arXiv:2012.12830 [abs], 2020.

[16] L. Henkin, Some remarks on infinitely long formulas, in: Infinitistic Methods, Proceedings Symposium Foundations of Mathematics, Pergamon, Warsaw, 1961, pp. 167–183.

[17] W. Hodges, Compositional semantics for a language of imperfect information, Log. J. IGPL 5 (1997) 539–563.

[18] Juha Kontinen, Jouko Väänänen, On definability in dependence logic, J. Log. Lang. Inf. 18 (3) (2009) 317–332, Erratum: J. Log. Lang. Inf. 20 (1) (2011) 133–134.

[19] Juha Kontinen, Fan Yang, Logics for first-order team properties, in: Rosalie Iemhoff, Michael Moortgat, Ruy de Queiroz (Eds.), Logic, Language, Information, and Computation, Springer Berlin Heidelberg, Berlin, Heidelberg, 2019, pp. 392–414.

[20] Andreas Krebs, Arne Meier, Jonni Virtema, Martin Zimmermann, Team semantics for the specification and verification of hyperproperties, in: MFCS, in: LIPIcs, vol. 117, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018, 10.

[21] Martin Lück, Axiomatizations of team logics, Ann. Pure Appl. Log. 169 (9) (2018) 928–969.

[22] Jouko Väänänen, Dependence Logic: A New Approach to Independence Friendly Logic, Cambridge University Press, Cambridge, 2007.

[23] W.J. Walkoe, Finite partially-ordered quantification, J. Symb. Log. 35 (1970) 535–555.