

<https://helda.helsinki.fi>

---

## Facets of Distribution Identities in Probabilistic Team Semantics

Hannula, Miika

Springer  
2019

---

Hannula , M , Hirvonen , Å , Kontinen , J , Kulikov , V & Virtema , J 2019 , Facets of Distribution Identities in Probabilistic Team Semantics . in F Calimeri , N Leone & M Manna (eds) , Logics in Artificial Intelligence : 16th European Conference, JELIA 2019, Rende, Italy, pý May 7 11 , 2019 , Proceedings . vol. 11468 , Lecture Notes in Computer , Springer , pp. 304-320 , The 16th European Conference on Logics in Artificial Intelligence (JELIA 2019) , 07/05/2019 . [https://doi.org/10.1007/978-3-030-19570-0\\_20](https://doi.org/10.1007/978-3-030-19570-0_20)

---

<http://hdl.handle.net/10138/310088>

[https://doi.org/10.1007/978-3-030-19570-0\\_20](https://doi.org/10.1007/978-3-030-19570-0_20)

---

acceptedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

# Facets of Distribution Identities in Probabilistic Team Semantics\*

Miika Hannula<sup>1</sup>[0000-0002-9637-6664], Åsa Hirvonen<sup>1</sup>[0000-0003-2149-4153], Juha Kontinen<sup>1</sup>[0000-0003-0115-5154], Vadim Kulikov<sup>1,2</sup>, and Jonni Virtema<sup>3</sup>[0000-0002-1582-3718]

<sup>1</sup> University of Helsinki, Finland {miika.hannula, asa.hirvonen, juha.kontinen}@helsinki.fi  
<sup>2</sup> Aalto University, Finland, vadim.kulikov@iki.fi  
<sup>3</sup> Hasselt University, Belgium, jonni.virtema@uhasselt.be

**Abstract.** We study probabilistic team semantics which is a semantical framework allowing the study of logical and probabilistic dependencies simultaneously. We examine and classify the expressive power of logical formalisms arising by different probabilistic atoms such as conditional independence and different variants of marginal distribution equivalences. We also relate the framework to the first-order theory of the reals and apply our methods to the open question on the complexity of the implication problem of conditional independence.

**Keywords:** team semantics · probabilistic logic · conditional independence

## 1 Introduction

Team semantics, introduced by Hodges [20] and popularised by Väänänen [25], shifts the focus of logics away from assignments as the primitive notion connected to satisfaction. In team semantics formulae are evaluated with respect to sets of assignments (i.e., teams) as opposed to single assignments of Tarskian semantics. During the last decade the research on team semantics has flourished, many logical formalisms have been defined, and surprising connections to other fields identified. In particular, several promising application areas of team semantics have been identified recently. Krebs et al. [22] developed a team based approach to linear temporal logic for the verification of information flow properties. In applications to database theory, a team corresponds exactly to a database table (see, e.g., [16]). Hannula et al. [18] introduced a framework that extends the connection of team semantics and database theory to polyrelational databases and data exchange.

The focus of this article is probabilistic team semantics which connects team based logics to probabilistic dependency notions. Probabilistic team semantics is built compositionally upon the notion of a probabilistic team, that is, a probability distribution over variable assignments. While the first ideas of probabilistic teams trace back to the works of Galliani [11] and Hyttinen et al. [21], the systematic study of the topic was initiated and further continued by Durand et al. in [8,9]. It is worth noting that in [2] so-called causal teams have been introduced to logically model causality and interventions. Probabilistic team semantics has also a close connection to the area of metafinite model theory [14]. In metafinite model theory, finite structures are extended with an another (infinite) domain sort such as the real numbers (often with arithmetic) and with weight functions that work as a bridge between the two sorts. This approach provides an elegant way to model weighted graphs and other structures that refer to infinite structures. The exact relationship between probabilistic team semantics and logics over metafinite models as well as with probabilistic databases of [6] will be a topic of future research.

The starting point of this work comes from [9] in which probabilistic team semantics was defined following the lines of [11]. The main theme in [9] was to characterize logical formalisms in this framework in terms of existential second-order logic. Two main probabilistic dependency atoms were examined. The probabilistic conditional independence atom  $\mathbf{y} \perp\!\!\!\perp_x \mathbf{z}$  states that the two

\* The first and the third author were supported by grant 308712, the fourth by grant 285203 of the Academy of Finland.

variable tuples  $\mathbf{y}$  and  $\mathbf{z}$  are independent given the third tuple  $\mathbf{x}$ . The marginal identity atom  $\mathbf{x} \approx \mathbf{y}$  states that the marginal distributions induced from the two tuples  $\mathbf{x}$  and  $\mathbf{y}$  (of the same length) are identical. The extension of first-order logic with these atoms ( $\text{FO}(\perp_c, \approx)$ ) was then shown to correspond to a two-sorted variant of existential second-order logic that allows a restricted access to arithmetical operations for numerical function terms. What was left unexamined were the relationships between different logical formalisms in probabilistic team semantics. In fact, it was unknown whether there are any meaningful probabilistic dependency notions such that the properties definable with one notion are comparable to those definable with another.

In this article we study the relative expressivity of first-order logic with probabilistic conditional independence atoms ( $\text{FO}(\perp_c)$ ) and with marginal identity atoms ( $\text{FO}(\approx)$ ). The logic  $\text{FO}(\approx)$  is a probabilistic variant of *inclusion logic* that is strictly less expressive than *independence logic*, after which  $\text{FO}(\perp_c)$  is modelled [12,15]. In addition, we examine  $\text{FO}(\approx^*)$  which is another extension defined in terms of so-called marginal distribution equivalence. The *marginal distribution equivalence atom*  $\mathbf{x} \approx^* \mathbf{y}$  for two variable tuples  $\mathbf{x}$  and  $\mathbf{y}$  (not necessarily of the same length) relaxes the truth condition of the marginal identity atom in that the two distributions induced from  $\mathbf{x}$  and  $\mathbf{y}$  are required to determine the same multisets of probabilities. The aforementioned open question is now answered in the positive. The logics mentioned above are not only comparable, but they form a linear expressivity hierarchy:  $\text{FO}(\approx) < \text{FO}(\approx^*) \leq \text{FO}(\perp_c)$ . We also show that  $\text{FO}(\approx)$  enjoys a union closure property that is a generalization of the union closure property of inclusion logic, and that conditional independence atoms  $\mathbf{y} \perp_x \mathbf{z}$  can be defined with an access to only marginal independence atoms  $\mathbf{x} \perp \mathbf{y}$  between two variable tuples. Furthermore, we show that, surprisingly,  $\text{FO}(\approx^*)$  corresponds to  $\text{FO}(\approx, =(\cdot))$ , where  $=(\cdot)$  refers to the dependence atom defined as a declaration of functional dependence over the support of the probabilistic team. The question whether  $\text{FO}(\approx, =(\cdot))$  is strictly less expressive than  $\text{FO}(\perp_c)$  is left as an open question; in team semantics the corresponding logics are known to be equivalent. The above findings look outwardly very similar to many results in team semantics. However, it is important to note that, apart perhaps from the union closure property, the results of this paper base on entirely new ideas and do not recycle old arguments from the team semantics context.

We also investigate (quantified) propositional logics with probabilistic team semantics. By connecting these logics to the arithmetic of the reals we show upper bounds for their associated computational problems. Our results suggest that the addition of probabilities to team semantics entails an increase in the complexity. Satisfiability of propositional team logic ( $\text{PL}(\sim)$ ), i.e., propositional logic with classical negation is in team semantics known to be complete for alternating exponential time with polynomially many alternations [19]. Shifting to probabilistic team semantics analogous problems are here shown to enjoy double exponential space upper bound. This is still lower than the complexity of satisfiability for modal team logic ( $\text{ML}(\sim)$ ) in team semantics, known to be complete for the non-elementary complexity class  $\text{TOWER}(\text{poly})$  which consists of problems solvable in time restricted by some tower of exponentials of polynomial height [23]. One intriguing consequence of our translation to real arithmetic is that the implication problem of conditional independence statements over binary distributions is decidable in exponential space. The decidability of this problem is open relative to all discrete probability distributions [24].

## 2 Preliminaries

First-order variables are denoted by  $x, y, z$  and tuples of first-order variables by  $\mathbf{x}, \mathbf{y}, \mathbf{z}$ . By  $\text{Var}(\mathbf{x})$  we denote the set of variables that appear in the variable sequence  $\mathbf{x}$ . The length of the tuple  $\mathbf{x}$  is denoted by  $|\mathbf{x}|$ . A *vocabulary*  $\tau$  is a set of relation symbols and function symbols with prescribed arities. We mostly denote relation symbols by  $R$  and function symbols by  $f$ , and the related arities

by  $\text{ar}(R)$  and  $\text{ar}(f)$ , respectively. The closed interval of real numbers between 0 and 1 is denoted by  $[0, 1]$ . Given a finite set  $A$ , a function  $f: A \rightarrow [0, 1]$  is called a (*probability*) *distribution* if  $\sum_{s \in A} f(s) = 1$ . In addition, the empty function is a *distribution*.

The probabilistic logics investigated in this paper are extensions of first-order logic FO over a vocabulary  $\tau$  given by the grammar rules:

$$\phi ::= x = y \mid x \neq y \mid R(\mathbf{x}) \mid \neg R(\mathbf{x}) \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid \exists x \phi \mid \forall x \phi,$$

where  $\mathbf{x}$  is a tuple of first-order variables and  $R$  a relation symbol from  $\tau$ .

Let  $D$  be a finite set of first-order variables and  $A$  be a nonempty set. A function  $s: D \rightarrow A$  is called an *assignment*. For a variable  $x$  and  $a \in A$ , the assignment  $s(a/x): D \cup \{x\} \rightarrow A$  is equal to  $s$  with the exception that  $s(a/x)(x) = a$ . A *team*  $X$  is a finite set of assignments from  $D$  to  $A$ . The set  $D$  is called the *domain* of  $X$  (written  $\text{Dom}(X)$ ) and the set  $A$  the *range* of  $X$  (written  $\text{Ran}(X)$ ). Let  $X$  be a team with range  $A$ , and let  $F: X \rightarrow \mathcal{P}(A) \setminus \{\emptyset\}$  be a function. We denote by  $X[A/x]$  the modified team  $\{s(a/x) \mid s \in X, a \in A\}$ , and by  $X[F/x]$  the team  $\{s(a/x) \mid s \in X, a \in F(s)\}$ . A *probabilistic team*  $\mathbb{X}$  is a distribution  $\mathbb{X}: X \rightarrow [0, 1]$ . Let  $\mathfrak{A}$  be a  $\tau$ -structure and  $\mathbb{X}: X \rightarrow [0, 1]$  a probabilistic team such that the domain of  $\mathfrak{A}$  is the range of  $X$ . Then we say that  $\mathbb{X}$  is a probabilistic team of  $\mathfrak{A}$ . In the following, we will define two notations  $\mathbb{X}[A/x]$  and  $\mathbb{X}[F/x]$ . Let  $\mathbb{X}: X \rightarrow [0, 1]$  be a probabilistic team,  $A$  a finite non-empty set,  $p_A$  the set of all probability distributions  $d: A \rightarrow [0, 1]$ , and  $F: X \rightarrow p_A$  a function. We denote by  $\mathbb{X}[A/x]$  the probabilistic team  $X[A/x] \rightarrow [0, 1]$  such that

$$\mathbb{X}[A/x](s(a/x)) = \sum_{\substack{t \in X \\ t(a/x) = s(a/x)}} \mathbb{X}(t) \cdot \frac{1}{|A|},$$

for each  $a \in A$  and  $s \in X$ . Note that if  $x$  does not belong to the domain of  $X$  then the righthand side of the above equation is simply  $\mathbb{X}(s) \cdot \frac{1}{|A|}$ . By  $\mathbb{X}[F/x]$  we denote the probabilistic team  $X[F/x] \rightarrow [0, 1]$  defined such that

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

for each  $a \in A$  and  $s \in X$ . Again if  $x$  does not belong to the domain of  $X$ ,  $\sum$  can be dropped from the above equation.

If  $\mathbb{Y}: X \rightarrow [0, 1]$  and  $\mathbb{Z}: X \rightarrow [0, 1]$  are probabilistic teams and  $k \in [0, 1]$ , then we write  $\mathbb{Y} \sqcup_k \mathbb{Z}$  for the  $k$ -scaled union of  $\mathbb{Y}$  and  $\mathbb{Z}$ , that is, the probabilistic team  $\mathbb{Y} \sqcup_k \mathbb{Z}: X \rightarrow [0, 1]$  defined such that  $(\mathbb{Y} \sqcup_k \mathbb{Z})(s) = k \cdot \mathbb{Y}(s) + (1 - k) \cdot \mathbb{Z}(s)$  for each  $s \in X$ .

We may now define probabilistic team semantics for first-order formulae. The definition is the same as in [9]. The only exception is that it is here applied to probabilistic teams that have real probabilities, whereas in [9] rational probabilities were used.

**Definition 1.** Let  $\mathfrak{A}$  be a probabilistic  $\tau$ -structure over a finite domain  $A$ , and  $\mathbb{X}: X \rightarrow [0, 1]$  a probabilistic team of  $\mathfrak{A}$ . The satisfaction relation  $\models_{\mathbb{X}}$  for first-order logic is defined as follows:

$$\begin{aligned} \mathfrak{A} \models_{\mathbb{X}} x = y &\Leftrightarrow \text{for all } s \in X : \text{if } \mathbb{X}(s) > 0, \text{ then } s(x) = s(y) \\ \mathfrak{A} \models_{\mathbb{X}} x \neq y &\Leftrightarrow \text{for all } s \in X : \text{if } \mathbb{X}(s) > 0, \text{ then } s(x) \neq s(y) \\ \mathfrak{A} \models_{\mathbb{X}} R(\mathbf{x}) &\Leftrightarrow \text{for all } s \in X : \text{if } \mathbb{X}(s) > 0, \text{ then } s(\mathbf{x}) \in R^{\mathfrak{A}} \\ \mathfrak{A} \models_{\mathbb{X}} \neg R(\mathbf{x}) &\Leftrightarrow \text{for all } s \in X : \text{if } \mathbb{X}(s) > 0, \text{ then } s(\mathbf{x}) \notin R^{\mathfrak{A}} \\ \mathfrak{A} \models_{\mathbb{X}} (\psi \wedge \theta) &\Leftrightarrow \mathfrak{A} \models_{\mathbb{X}} \psi \text{ and } \mathfrak{A} \models_{\mathbb{X}} \theta \\ \mathfrak{A} \models_{\mathbb{X}} (\psi \vee \theta) &\Leftrightarrow \mathfrak{A} \models_{\mathbb{Y}} \psi \text{ and } \mathfrak{A} \models_{\mathbb{Z}} \theta \text{ for some } \mathbb{Y}, \mathbb{Z}, k \text{ s.t. } \mathbb{Y} \sqcup_k \mathbb{Z} = \mathbb{X} \\ \mathfrak{A} \models_{\mathbb{X}} \forall x \psi &\Leftrightarrow \mathfrak{A} \models_{\mathbb{X}[A/x]} \psi \\ \mathfrak{A} \models_{\mathbb{X}} \exists x \psi &\Leftrightarrow \mathfrak{A} \models_{\mathbb{X}[F/x]} \psi \text{ holds for some } F: X \rightarrow p_A. \end{aligned}$$

$\mathbb{X}$			
$x$	$y$	$z$	$P$
$a$	$b$	$c$	$1/2$
$b$	$c$	$b$	$1/2$

**Fig. 1.** A representation of a probabilistic team  $\mathbb{X}$ , for Example 2, with domain  $\{x, y, z\}$  that consists of two assignments whose probabilities are  $1/2$ .

Probabilistic team semantics is in line with Tarskian semantics for first-order formulae ( $\models_s$ ):

$$\mathfrak{A} \models_{\mathbb{X}} \psi \Leftrightarrow \forall s \in X \text{ such that } \mathbb{X}(s) > 0 : \mathfrak{A} \models_s \psi.$$

In particular the non-classical semantics for negation is required for the above equivalence to hold.

In this paper we consider three probabilistic atoms: marginal identity, probabilistic independence, and marginal distribution equivalence atom. The first two were first introduced in the context of multiteam semantics in [8], and they extend the notions of inclusion and independence atoms from team semantics [12].

We define  $|\mathbb{X}_{\mathbf{x}=\mathbf{a}}|$  where  $\mathbf{x}$  is a tuple of variables and  $\mathbf{a}$  a tuple of values, as

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

If  $\phi$  is some first-order formula, then  $|\mathbb{X}_{\phi}|$  is defined analogously as the total sum of weights of those assignments in  $X$  that satisfy  $\phi$ .

If  $\mathbf{x}, \mathbf{y}$  are variable sequences of length  $k$ , then  $\mathbf{x} \approx \mathbf{y}$  is a *marginal identity atom* with the following semantics:

$$\mathfrak{A} \models_{\mathbb{X}} \mathbf{x} \approx \mathbf{y} \Leftrightarrow |\mathbb{X}_{\mathbf{x}=\mathbf{a}}| = |\mathbb{X}_{\mathbf{y}=\mathbf{a}}| \text{ for each } \mathbf{a} \in A^k. \quad (1)$$

Note that the equality  $|\mathbb{X}_{\mathbf{x}=\mathbf{a}}| = |\mathbb{X}_{\mathbf{y}=\mathbf{a}}|$  in (4) can be equivalently replaced with  $|\mathbb{X}_{\mathbf{x}=\mathbf{a}}| \leq |\mathbb{X}_{\mathbf{y}=\mathbf{a}}|$  since the tuples  $\mathbf{a}$  range over  $A^k$  for a finite  $A$  (see [8, Definition 7] for details). Due to this alternative formulation, marginal identity atoms were in [8] called probabilistic inclusion atoms. Intuitively, the atom  $\mathbf{x} \approx \mathbf{y}$  states that the distributions induced from  $\mathbf{x}$  and  $\mathbf{y}$  are identical.

The marginal distribution equivalence atom is defined in terms of multisets of assignment weights. We distinguish multisets from sets by using double wave brackets, e.g.,  $\{\{a, a, b\}\}$  denotes the multiset  $(\{a, b\}, m)$  where  $a$  and  $b$  are given multiplicities  $m(a) = 2$  and  $m(b) = 1$ . If  $\mathbf{x}, \mathbf{y}$  are variable sequences, then  $\mathbf{x} \approx^* \mathbf{y}$  is a *marginal distribution equivalence atom* with the following semantics:

$$\mathfrak{A} \models_{\mathbb{X}} \mathbf{x} \approx^* \mathbf{y} \Leftrightarrow \{\{|\mathbb{X}_{\mathbf{x}=\mathbf{a}}| > 0 \mid \mathbf{a} \in A^{|\mathbf{x}|}\}\} = \{\{|\mathbb{X}_{\mathbf{y}=\mathbf{b}}| > 0 \mid \mathbf{b} \in A^{|\mathbf{y}|}\}\}. \quad (2)$$

The next example illustrates the relationships between marginal distribution equivalence atoms and marginal identity atoms; the latter implies the former, but not vice versa.

*Example 2.* Let  $\mathbb{X}$  be the probabilistic team depicted in Figure 1. The team  $\mathbb{X}$  satisfies the atoms  $xy \approx^* y$ ,  $x \approx^* y$ ,  $y \approx^* z$ , and  $y \approx z$ . The team  $\mathbb{X}$  falsifies the atom  $x \approx y$ , whereas  $xy \approx y$  is not a well formed formula.

If  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  are variable sequences, then  $\mathbf{y} \perp_{\mathbf{x}} \mathbf{z}$  is a *probabilistic conditional independence atom* with the satisfaction relation defined as

$$\mathfrak{A} \models_{\mathbb{X}} \mathbf{y} \perp_{\mathbf{x}} \mathbf{z} \quad (3)$$

if for all  $s: \text{Var}(\mathbf{x}\mathbf{y}\mathbf{z}) \rightarrow A$  it holds that

$$|\mathbb{X}_{\mathbf{x}\mathbf{y}=s(\mathbf{x}\mathbf{y})}| \cdot |\mathbb{X}_{\mathbf{x}\mathbf{z}=s(\mathbf{x}\mathbf{z})}| = |\mathbb{X}_{\mathbf{x}\mathbf{y}\mathbf{z}=s(\mathbf{x}\mathbf{y}\mathbf{z})}| \cdot |\mathbb{X}_{\mathbf{x}=s(\mathbf{x})}|.$$

Furthermore, we define *probabilistic marginal independence atom*  $\mathbf{x} \perp\!\!\!\perp \mathbf{y}$  as  $\mathbf{x} \perp\!\!\!\perp_{\emptyset} \mathbf{y}$ , i.e., probabilistic independence conditioned by the empty tuple.

In addition to atoms based on counting or arithmetic operations, we may also include all dependency atoms from the team semantics literature. Let  $\alpha$  be an atom that is interpreted in team semantics, let  $\mathfrak{A}$  be a finite structure, and  $\mathbb{X}: X \rightarrow [0, 1]$  a probabilistic team. We define  $\mathfrak{A} \models_{\mathbb{X}} \alpha$  if  $\mathfrak{A} \models_{X^+} \alpha$ , where  $X^+$  consists of those assignments of  $X$  that are given positive weight by  $\mathbb{X}$ . In this paper we will discuss dependence atoms also in the context of probabilistic team semantics. If  $\mathbf{x}, \mathbf{y}$  are two variable sequences, then  $=(\mathbf{x}, \mathbf{y})$  is a *dependence atom* with team semantics:

$$\mathfrak{A} \models_{\mathbb{X}} =(\mathbf{x}, \mathbf{y}) \Leftrightarrow s(\mathbf{x}) = s'(\mathbf{x}) \text{ implies } s(\mathbf{y}) = s'(\mathbf{y}) \text{ for all } s, s' \in X. \quad (4)$$

A dependence atom of the form  $=(\emptyset, \mathbf{x})$  is called a *constancy atom*, written  $=(\mathbf{x})$  in shorthand notation. Dependence atoms can be expressed by using probabilistic independence atoms. This has been shown for multiteams in [8], and the proof applies to probabilistic teams.

**Proposition 3 ([8]).** *Let  $\mathfrak{A}$  be a structure,  $\mathbb{X}: X \rightarrow [0, 1]$  a probabilistic team of  $\mathfrak{A}$ , and  $\mathbf{x}$  and  $\mathbf{y}$  two sequences of variables. Then  $\mathfrak{A} \models_{\mathbb{X}} =(\mathbf{x}, \mathbf{y}) \Leftrightarrow \mathfrak{A} \models_{\mathbb{X}} \mathbf{y} \perp\!\!\!\perp_{\mathbf{x}} \mathbf{y}$ .*

Given a collection  $C$  of atoms from  $\{\perp\!\!\!\perp_c, \perp\!\!\!\perp, \approx, \approx^*, =(\cdot)\}$ , we write  $\text{FO}(C)$  for the logic that extends FO with the atoms in  $C$ .

*Example 4.* Let  $f_1, \dots, f_n, g$  be univariate distributions. Then  $g$  is a *finite mixture* of  $f_1, \dots, f_n$  if it can be expressed as a convex combination of  $f_1, \dots, f_n$ , i.e., if there are non-negative real numbers  $r_1, \dots, r_n$  such that  $r_1 + \dots + r_n = 1$  and  $g(a) = \sum_{i=1}^n r_i f_i(a)$ . A probabilistic team  $\mathbb{X}: X \rightarrow [0, 1]$  gives rise to a univariate distribution  $f_x(a) := |\mathbb{X}_{x=a}|$  for each variable  $x$  from the domain of  $X$ . The next formula expresses that the distribution  $f_y$  is a finite mixture of the distributions  $f_{x_1}, \dots, f_{x_n}$ :

$$\exists qr [x_1 \dots x_n \perp\!\!\!\perp r \wedge \bigvee_{i=1}^n r = i \wedge \bigwedge_{i=1}^n \exists x'r' (x_i r \approx x'r' \wedge [(q = i \vee r' = i) \rightarrow yq = x'r'])],$$

where the indices  $1, \dots, n$  are also thought of as distinct constants, and  $(q = i \vee r' = i) \rightarrow yq = x'r'$  stands for  $\neg(q \neq i \wedge r' \neq i) \vee yq = x'r'$ . The non-negative real numbers  $r_i$  are represented by the weights of  $r = i$  where  $r$  is distributed independently of each  $x_i$ . The summand  $r_i f_{x_i}(a)$  is then represented by the weight of  $x_i r = ai$  and  $f_y(a)$  by the weight of  $y = a$ . The quantified subformula expresses that the former weight matches the weight of  $yq = ai$ , which implies that  $f_y(a)$  is  $r_1 f_{x_1}(a) + \dots + r_n f_{x_n}(a)$ .

*Example 5.* Probabilistic team semantics can be also used to model properties of data obtained from a quantum experiment (adapting the approach of [1]). Consider a probabilistic team  $\mathbb{X}$  over variables  $m_1, \dots, m_n, o_1, \dots, o_n$ . The intended interpretation of  $\mathbb{X}(s) = r$  is that the joint probability that  $s(m_i)$  was measured with outcome  $s(o_i)$ , for  $1 \leq i \leq n$ , is  $r$ . In this setting many important properties of the experiment can be expressed using our formalism. For example the formula

$$o_i \perp\!\!\!\perp_{\mathbf{m}} (o_1, \dots, o_{i-1}, o_{i+1}, \dots, o_n)$$

expresses a property called *Outcome-Independence*; given the measurements  $\mathbf{m}$ , the outcome at  $i$  is independent of the outcomes at other positions. The dependence atom  $=(\mathbf{m}, \mathbf{o})$  on the other hand corresponds to a property called *Weak-Determinism*. Moreover, if  $\phi$  describes some property of hidden-variable models (Outcome-Independence, etc.), then the formula  $\exists \lambda \phi$  expresses that the experiment can be explained by a hidden-variable model satisfying that property.

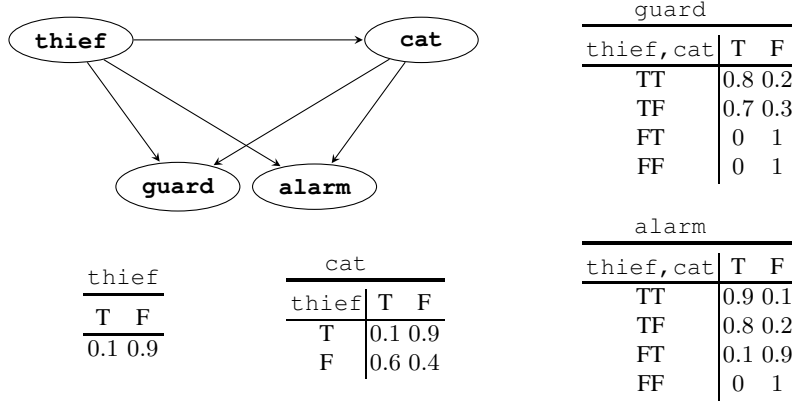


Fig. 2. Bayesian network  $\mathbb{G}$  and its related conditional distributions

The next example relates probabilistic team semantics to Bayesian networks. The example is an adaptation of an example discussed also in [8].

*Example 6.* Consider the Bayesian network  $\mathbb{G}$  in Fig. 2 that models beliefs about house safety using four Boolean random variables `thief`, `cat`, `guard` and `alarm`. We refer to these variables by  $t, c, g, a$ . The dependence structure of a Bayesian network is characterized by the so-called local directed Markov property stating that each variable is conditionally independent of its non-descendants given its parents. For our network  $\mathbb{G}$  the only non-trivial independence given by this property is  $g \perp_{tc} a$ . Hence a joint distribution  $P$  over  $t, c, g, a$  factorizes according to  $\mathbb{G}$  if  $\mathbb{X}$  satisfies  $g \perp_{tc} a$ . In this case  $P$  can be factorized by

$$P(t, c, g, a) = P(t) \cdot P(c | t) \cdot P(g | t, c) \cdot P(a | t, c) \quad (5)$$

where, for instance,  $t$  abbreviates either `thief` =  $T$  or `thief` =  $F$ , and  $P(c | t)$  is the probability of  $c$  given  $t$ . The joint probability distribution (i.e., the team  $\mathbb{X}$ ) can hence be stored as in Fig. 2. Note that while  $\mathbb{G}$  expresses the independence statement  $g \perp_{tc} a$ ,  $\text{FO}(\perp_c, \approx)$ -formulas can be used to further refine the joint probability distribution as follows. Assume we have information suggesting that we may safely assume an  $\text{FO}(\perp_c, \approx)$  formula  $\phi$  on  $\mathbb{X}$ :

- $\phi := t = F \rightarrow g = F$  indicates that `guard` never raises alarm in absence of `thief`. In this case the two bottom rows of the conditional probability distribution for `guard` become superfluous.
- the assumption that  $\phi$  is satisfied also exemplifies an interesting form of *context-specific* independence (CSI) that cannot be formalized by the usual Bayesian networks (see, e.g., [7]). Namely,  $\phi$  implies that `guard` is independent of `cat` in the context `thief` =  $F$ . Interestingly such CSI statements can be formalized utilizing the disjunction of  $\text{FO}(\perp_c, \approx)$ :

$$t = T \vee (t = F \wedge g \perp c).$$

- satisfaction of  $\phi := tca \approx tcg$  would imply that `alarm` and `guard` have the same reliability for any given value of `thief` and `cat`. Consequently, the conditional distributions for `alarm` and `guard` are equal and one of the them could be removed.

The following locality property dictates that satisfaction of a formula  $\phi$  in probabilistic team semantics depends only on the free variables of  $\phi$ . For this, we define the *restriction* of a team  $X$  to  $V$  as  $X \upharpoonright V = \{s \upharpoonright V \mid s \in X\}$  where  $s \upharpoonright V$  denotes the restriction of the assignment  $s$  to  $V$ . The restriction of a probabilistic team  $\mathbb{X} : X \rightarrow [0, 1]$  to  $V$  is then defined as the probabilistic team  $\mathbb{Y} : X \upharpoonright V \rightarrow [0, 1]$  where  $\mathbb{Y}(s) = \sum_{s' \upharpoonright V = s} \mathbb{X}(s')$ . The set of *free variables*  $\text{Fr}(\phi)$  of a formula over probabilistic team semantics is defined recursively as in first-order logic; note that for any atom  $\phi$ ,  $\text{Fr}(\phi)$  consists of all variables that appear in  $\phi$ .

**Proposition 7 (Locality, [9]).** Let  $\phi(\mathbf{x}) \in \text{FO}(\perp\!\!\!\perp_c, \approx, \approx^*, =(\cdot))$  be a formula with free variables from  $\mathbf{x} = (x_1, \dots, x_n)$ . Then for all structures  $\mathfrak{A}$  and probabilistic teams  $\mathbb{X} : X \rightarrow [0, 1]$  where  $\{x_1, \dots, x_n\} \subseteq V \subseteq \text{Dom}(X)$ ,  $\mathfrak{A} \models_{\mathbb{X}} \phi \iff \mathfrak{A} \models_{\mathbb{X}|_V} \phi$ .

Given two logics  $\mathcal{L}$  and  $\mathcal{L}'$  over probabilistic team semantics, we write  $\mathcal{L} \leq \mathcal{L}'$  if for all open formulae  $\phi(\mathbf{x}) \in \mathcal{L}$  there is a formula  $\psi(\mathbf{x}) \in \mathcal{L}'$  such that  $\mathfrak{A} \models_{\mathbb{X}} \phi \iff \mathfrak{A} \models_{\mathbb{X}} \psi$ , for all structures  $\mathfrak{A}$  and probabilistic teams  $\mathbb{X}$ . The equality “ $\equiv$ ” and strict inequality “ $<$ ” relations between  $\mathcal{L}$  and  $\mathcal{L}'$  are defined from “ $\leq$ ” in the standard way.

**Alternative Definition.** Probabilistic teams can also be defined as mappings  $\mathbb{X} : X \rightarrow \mathbb{R}_{\geq 0}$  that have no restriction for the total sum of assignment weights,  $\mathbb{R}_{\geq 0}$  being the set of all non-negative reals. Probabilistic team semantics with respect to such real weighted teams is then given exactly as in Definition 1, except that we define disjunction without scaling:

$$\mathfrak{A} \models_{\mathbb{X}} (\psi \vee \theta) \iff \mathfrak{A} \models_{\mathbb{Y}} \psi \text{ and } \mathfrak{A} \models_{\mathbb{Z}} \theta \text{ for some } \mathbb{Y}, \mathbb{Z} \text{ s.t. } \mathbb{Y} \sqcup \mathbb{Z} = \mathbb{X},$$

where the union  $\mathbb{Y} \sqcup \mathbb{Z}$  is defined such that  $(\mathbb{Y} \sqcup \mathbb{Z})(s) = \mathbb{Y}(s) + \mathbb{Z}(s)$  for each  $s$ . Whether interpreting probabilistic teams as probability distributions or just mappings from assignments to non-negative reals does not make any difference in our framework. Hence we write  $\mathbb{X} : X \rightarrow [0, 1]$  for a probabilistic team that is a distribution such that  $\sum_{s \in X} \mathbb{X}(s) = 1$ , and  $\mathbb{X} : X \rightarrow \mathbb{R}_{\geq 0}$  for a probabilistic team that is any mapping from assignments to non-negative reals. A probabilistic team of the former type is then a special case of that of the latter. We will use both notions and their associated semantics interchangeably. If we need to distinguish between the two semantics, we write  $\models^{[0,1]}$  and  $\models^{\geq 0}$  respectively for the scaled (i.e., Definition 1) and non-scaled variants. Given  $\mathbb{X} : X \rightarrow \mathbb{R}_{\geq 0}$  and  $r \in \mathbb{R}_{\geq 0}$ , we write  $|\mathbb{X}|$  for the total weight  $\sum_{s \in X} \mathbb{X}(s)$  of  $\mathbb{X}$ , and  $r \cdot \mathbb{X}$  for the probabilistic team  $\mathbb{Y} : X \rightarrow \mathbb{R}_{\geq 0}$  such that  $\mathbb{Y}(s) = r \cdot \mathbb{X}(s)$  for all  $s \in X$ . The proposition below follows from a straightforward induction (see Appendix A).

**Proposition 8.** Let  $\mathfrak{A}$  be a structure,  $\mathbb{X} : X \rightarrow \mathbb{R}_{\geq 0}$  a probabilistic team of  $\mathfrak{A}$ , and  $\phi \in \text{FO}(\perp\!\!\!\perp_c, \approx, \approx^*, =(\cdot))$ . Then  $\mathfrak{A} \models_{\mathbb{X}}^{\geq 0} \phi \iff \mathfrak{A} \models_{\frac{1}{|\mathbb{X}|} \cdot \mathbb{X}}^{[0,1]} \phi$ .

### 3 Expressiveness of $\text{FO}(\perp\!\!\!\perp)$

Let  $\mathbb{X} : X \rightarrow [0, 1]$  be a probabilistic team where  $X$  is a finite set of assignments from a finite set  $D$  of variables. A variable  $x \in D$  is *uniformly distributed* in  $\mathbb{X}$  over a set of values  $S$ , if

$$\mathbb{X}_{x=a} = \frac{1}{|S|} \text{ for all } a \in S \text{ and } \mathbb{X}_{x=a} = 0 \text{ otherwise.}$$

The following lemma says essentially that if we can express constancy and independence for a uniform distribution, then we can express  $\approx$ . Note that it may happen that we can express “ $\mathbf{x}$  uniformly distributed and independent of  $\mathbf{y}$ ” even when we cannot express “ $\mathbf{x}$  is independent of  $\mathbf{y}$ ” in general. For a proof of the lemma, see Appendix B.

**Lemma 9.** Let  $\mathfrak{A}$  be structure with at least two elements and  $\mathbf{z}$  an  $n$ -tuple of variables. Let  $\phi(\mathbf{z}, d, c_1, c_2)$  be a formula such that for all probabilistic teams  $\mathbb{X}$ , whose variable domain includes  $\mathbf{z}, d, c_1, c_2$  and for which  $\mathfrak{A} \models_{\mathbb{X}} c_1 \neq c_2$  and  $\mathfrak{A} \models_{\mathbb{X}} =(\mathbf{z}) \wedge =(\mathbf{z})$ , it holds that

$$\mathfrak{A} \models_{\mathbb{X}} \phi \iff d \text{ is uniformly distributed over the two values of } c_1, c_2 \text{ and } d \text{ is independent of } \mathbf{z}. \quad (6)$$

Then  $\mathbf{x} \approx \mathbf{y}$  can be expressed for  $n$ -tuples  $\mathbf{x}$  and  $\mathbf{y}$  using  $\phi$  and the constancy atom.



**Theorem 10.**  $\text{FO}(\approx) \leq \text{FO}(\perp\!\!\!\perp)$ .

*Proof.* Proposition 3 established that the constancy atom  $\text{=}(x)$  can be equivalently expressed by the independence atom  $x \perp\!\!\!\perp x$ . Hence it is enough to show that we can define the formula  $\phi$  of Lemma 9 by using  $\perp\!\!\!\perp$ .

Let  $\mathfrak{A}$  and  $\mathfrak{X}$  be as assumed in Lemma 9. We use below  $\exists b \in \{c_1, c_2\} \theta$  as an abbreviation for  $\exists b(b = c_1 \vee b = c_2) \wedge \theta$ , and  $\forall b \in \{c_1, c_2\} \theta$  for  $\forall b(b \neq c_1 \wedge b \neq c_2) \vee ((b = c_1 \vee b = c_2) \wedge \theta)$ . Define  $\phi(\mathbf{z}, d, c_1, c_2)$  as

$$(\mathbf{z} \perp\!\!\!\perp d) \wedge \forall a \in \{c_1, c_2\} \exists b \in \{c_1, c_2\} [(a \perp\!\!\!\perp b) \wedge ((a = b \wedge d = c_1) \vee (a \neq b \wedge d = c_2))].$$

It suffices to prove (5). The formula  $\phi$  clearly states that  $\mathbf{z}$  and  $d$  are independent. The formula also states that the values of  $d$  range over the values of  $c_1$  and  $c_2$ . It remains to be shown, conditioned on that  $\mathbf{z}$  and  $d$  are independent, that

$$\mathfrak{A} \models_{\mathfrak{X}} \phi \text{ if and only if } d \text{ is uniformly distributed over } c_1 \text{ and } c_2.$$

Note that, by assumption of Lemma 9,  $c_1$  and  $c_2$  are distinct constants. Let  $\mathfrak{X}_1$  be a team obtained from  $\mathfrak{X}$  by the quantification of  $a$  and  $b$ . By the definition of universal quantification, in  $\mathfrak{X}_1$   $a$  is uniformly distributed and independent of everything else except maybe  $b$ . Note that  $d$  is uniformly distributed over the values of  $c_1$  and  $c_2$  in  $X$  if and only if it is in  $X_1$ .

If  $d$  is uniformly distributed over the values of  $c_1$  and  $c_2$ , then picking values of  $b$  with a uniform probability such that the right conjunct in

$$[(a \perp\!\!\!\perp b) \wedge ((a = b \wedge d = c_1) \vee (a \neq b \wedge d = c_2))] \tag{7}$$

holds clearly yields a team in which the left conjunct also holds. However, if  $d$  is not uniformly distributed over  $c_1$  and  $c_2$ , then picking values for  $b$  such that the right conjunct of (7) holds will yield  $b$  that is not independent on  $a$ .  $\square$

We also note that conditional independence is definable using marginal independence. The proof applies ideas from [9] and can be found in Appendix C.

**Theorem 11.**  $\text{FO}(\perp\!\!\!\perp) \equiv \text{FO}(\perp\!\!\!\perp_c)$ .

## 4 Expressiveness of $\text{FO}(\approx^*)$ and $\text{FO}(\approx)$

Initially it may seem that first-order logic with marginal distribution equivalence atoms is less expressive than that with marginal identity atoms, as the former atoms are given a strictly weaker truth condition. Contrary to this intuition, however, we will in this section show that  $\text{FO}(\approx^*)$  is actually strictly more expressive than  $\text{FO}(\approx)$ . The result is proven in two phases. First, in Sect. 4.1 we show that dependence and marginal identity can be defined in  $\text{FO}(\approx^*)$ , the former by a single marginal distribution equivalence atom and the latter by a more complex formula. Second, in Sect. 4.2 we show that the expressiveness of  $\text{FO}(\approx)$  is restricted by a union closure property which is similar to that of inclusion logic in team semantics. Since dependence atoms lack this property, the strict inequality between  $\text{FO}(\approx)$  and  $\text{FO}(\approx^*)$  follows.

### 4.1 Translations of Dependence and Marginal Identity to $\text{FO}(\approx^*)$

We observe first that dependence atoms can be expressed in terms of marginal distribution equivalence atoms, which in turn are definable using marginal identity and dependence atoms.

**Proposition 12.** *The following equivalences hold:*

1.  $=(\mathbf{x}, \mathbf{y}) \equiv \mathbf{x}\mathbf{y} \approx^* \mathbf{x}$ ,
2.  $\mathbf{x} \approx^* \mathbf{y} \equiv \exists \mathbf{z} (=(\mathbf{y}, \mathbf{z}) \wedge =(\mathbf{z}, \mathbf{y}) \wedge \mathbf{x} \approx \mathbf{z})$ .

Defining marginal identity atoms in  $\text{FO}(\approx^*)$  is more cumbersome. Let  $\mathbb{X} : X \rightarrow \mathbb{R}_{\geq 0}$  be a probabilistic team, and  $\phi$  a quantifier-free first-order formula over the empty vocabulary (i.e., such that its satisfaction depends only on the variable assignment). We define  $\mathbb{X}_\phi : X \rightarrow \mathbb{R}_{\geq 0}$  as the probabilistic team such that  $\mathbb{X}_\phi(s) = \mathbb{X}(s)$  if  $s$  satisfies  $\phi$ , and  $\mathbb{X}_\phi(s) = 0$  otherwise. Given two sequences of variables  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{y} = (y_1, \dots, y_n)$ , we write  $\mathbf{x} \neq \mathbf{y}$  as a shorthand for  $\bigvee_{i=1}^n \neg x_i = y_i$ .

**Theorem 13.**  $\mathbf{x} \approx \mathbf{y}$  is equivalent to  $\phi \in \text{FO}(\approx^*)$  where

$$\phi := \forall \mathbf{z} ((\mathbf{z} \neq \mathbf{x} \wedge \mathbf{z} \neq \mathbf{y}) \vee ((\mathbf{z} = \mathbf{x} \vee \mathbf{z} = \mathbf{y}) \wedge \mathbf{z} \approx^* \mathbf{x} \wedge \mathbf{z} \approx^* \mathbf{y})).$$

*Proof.* Assume that  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  are all  $m$ -ary. Let  $\mathfrak{A}$  be a structure with domain  $A = \{1, \dots, n\}$ , and let  $\mathbb{X} : X \rightarrow \mathbb{R}_{\geq 0}$  a probabilistic team. Assume first that  $\mathfrak{A} \models_{\mathbb{X}} \mathbf{x} \approx \mathbf{y}$ , that is, for all  $i \in A^m$ , the weights  $|\mathbb{X}_{\mathbf{x}=i}|$  and  $|\mathbb{X}_{\mathbf{y}=i}|$  coincide. It suffices to show that  $\mathfrak{A} \models_{\mathbb{Y}} \mathbf{z} \approx^* \mathbf{x} \wedge \mathbf{z} \approx^* \mathbf{y}$  for  $\mathbb{Y} := \mathbb{X}'_\theta$  where  $\theta$  is  $\mathbf{z} = \mathbf{x} \vee \mathbf{z} = \mathbf{y}$  and  $\mathbb{X}' = \mathbb{X}[A^m/\mathbf{z}]$  is the probabilistic team obtained from  $\mathbb{X}$  by distributing  $A^m$  to  $\mathbf{z}$  uniformly. For each  $i \in A^m$  we consider three weight measures, obtained by dividing assignments associated with  $i$  into three parts,  $l_i := |\mathbb{X}_{\mathbf{x}=i \wedge \mathbf{x} \neq \mathbf{y}}|$ ,  $r_i := |\mathbb{X}_{\mathbf{y}=i \wedge \mathbf{x} \neq \mathbf{y}}|$ , and  $c_i := |\mathbb{X}_{\mathbf{x}=i \wedge \mathbf{y}=i}|$ . Then

$$|\mathbb{Y}_{\mathbf{x}=i}| = |\mathbb{X}'_{\theta \wedge \mathbf{x}=i}| = |\mathbb{X}'_{\theta \wedge \mathbf{x}=i \wedge \mathbf{x} \neq \mathbf{y}}| + |\mathbb{X}'_{\theta \wedge \mathbf{x}=i \wedge \mathbf{y}=i}| = \frac{2l_i + c_i}{n^m}.$$

Observe that for  $\mathbb{X}'_{\theta \wedge \mathbf{x}=i \wedge \mathbf{x} \neq \mathbf{y}}$  we first partition each assignment in  $\mathbb{X}_{\mathbf{x}=i \wedge \mathbf{x} \neq \mathbf{y}}$  uniformly to  $n^m$  parts in terms of the value of  $\mathbf{z}$  and then keep only those parts where  $\theta$  holds. Since  $\mathbf{x}$  and  $\mathbf{y}$  disagree for every assignment in  $\mathbb{X}'_{\mathbf{x}=i \wedge \mathbf{x} \neq \mathbf{y}}$ , the total weight of  $\mathbb{X}'_{\theta \wedge \mathbf{x}=i \wedge \mathbf{x} \neq \mathbf{y}}$  is obtained by multiplying  $l_i$  with  $\frac{2}{n^m}$ . For  $\mathbb{X}'_{\theta \wedge \mathbf{x}=i \wedge \mathbf{y}=i}$  we have identical  $\mathbf{x}$  and  $\mathbf{y}$ , and hence its weight is obtained by multiplying  $c_i$  with  $\frac{1}{n^m}$ . By analogous reasoning we obtain that

$$|\mathbb{Y}_{\mathbf{y}=i}| = \frac{2r_i + c_i}{n^m} \text{ and } |\mathbb{Y}_{\mathbf{z}=i}| = \frac{r_i + l_i + c_i}{n^m}.$$

Since our assumption implies  $l_i = r_i$  for all  $i$ , the claim now follows from the observation that  $\{|\mathbb{Y}_{\mathbf{u}=i}| \mid i \in A^m\}$  are identical multisets for  $\mathbf{u} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ .

Vice versa, assuming  $\mathfrak{A} \models_{\mathbb{X}} \phi$  we show  $\mathfrak{A} \models_{\mathbb{X}} \mathbf{x} \approx \mathbf{y}$ . Let the weights  $l_i, r_i, c_i$  and the probabilistic team  $\mathbb{Y}$  be as above. By assumption we have  $\mathfrak{A} \models_{\mathbb{Y}} \mathbf{z} \approx^* \mathbf{x} \wedge \mathbf{z} \approx^* \mathbf{y}$ , and thus the following multisets are identical:

$$\begin{aligned} W_{\mathbf{x}} &:= \{ \{2l_1 + c_1, \dots, 2l_n + c_n\} \}, \\ W_{\mathbf{y}} &:= \{ \{2r_1 + c_1, \dots, 2r_n + c_n\} \}, \\ W_{\mathbf{z}} &:= \{ \{l_1 + r_1 + c_1, \dots, l_n + r_n + c_n\} \}, \end{aligned}$$

where  $\mathbf{1} = (1, \dots, 1)$  and  $\mathbf{n} = (n, \dots, n)$ . Assume to the contrary that  $\mathfrak{A} \not\models_{\mathbb{X}} \mathbf{x} \approx \mathbf{y}$ , that is,  $l_i \neq r_i$  for some  $i$ . Observe that whenever  $l_j = r_j$  agree, then  $j$  contributes the same weight to all  $W_{\mathbf{x}}, W_{\mathbf{y}}$ , and  $W_{\mathbf{z}}$ . Therefore, we may assume without loss of generality that  $l_i \neq r_i$  for all  $i$ . Assume that  $2l_j + c_j$  is the smallest element from  $W_{\mathbf{x}}$ . Since  $W_{\mathbf{x}} = W_{\mathbf{z}}$ , it follows that  $2l_j + c_j = l_k + r_k + c_k$  for some  $k$ . If  $l_k < r_k$ , then  $2l_k + c_k < l_k + r_k + c_k$  which contradicts the assumption that  $2l_j + c_j$  is smallest. Since  $W_{\mathbf{x}} = W_{\mathbf{y}}$ , similar contradiction follows from  $r_k < l_k$ , too. Hence,  $\mathfrak{A} \models_{\mathbb{X}} \mathbf{x} \approx \mathbf{y}$  which concludes the proof.  $\square$

The following theorem now combines the results of this section. Note that the translations to both directions are of linear size.

**Theorem 14.**  $\text{FO}(\approx^*) \equiv \text{FO}(\approx, =(\cdot))$ .

## 4.2 Scaled Union Closure of $\text{FO}(\approx)$

Inclusion logic is known to be union closed over teams. This means that for all structures  $\mathfrak{A}$ , teams  $X$ , and inclusion logic formulae  $\phi$ : if  $\mathfrak{A} \models_X \phi$  and  $\mathfrak{A} \models_Y \phi$ , then  $\mathfrak{A} \models_{X \cup Y} \phi$ . The following proposition, proven in Appendix D, demonstrates that  $\text{FO}(\approx)$  is endowed with an analogous closure property, namely, that all formulae of  $\text{FO}(\approx)$  are closed under all  $k$ -scaled unions of probabilistic teams.

**Proposition 15.** *Let  $\mathfrak{A}$  be a model,  $\phi \in \text{FO}(\approx)$  a formula, and  $\mathbb{X} : X \rightarrow [0, 1]$  and  $\mathbb{Y} : X \rightarrow [0, 1]$  two probabilistic teams. Then for all  $k \in [0, 1]$ :*

$$\text{if } \mathfrak{A} \models_{\mathbb{X}} \phi \text{ and } \mathfrak{A} \models_{\mathbb{Y}} \phi, \text{ then } \mathfrak{A} \models_{\mathbb{X} \sqcup_k \mathbb{Y}} \phi.$$

As a corollary we observe that  $\text{FO}(\approx)$  is strictly weaker than  $\text{FO}(\approx^*)$ . Recall from Proposition 12 that the constancy atom  $=(x)$  is definable in  $\text{FO}(\approx^*)$ . However, constancy is clearly not preserved under  $k$ -scaled unions, therefore falling outside the scope of  $\text{FO}(\approx)$ . Furthermore, by Theorem 13  $\text{FO}(\approx^*)$  is at least as expressive as  $\text{FO}(\approx)$ .

**Corollary 16.**  $\text{FO}(\approx) < \text{FO}(\approx^*)$ .

## 5 Binary Probabilistic Teams

In this section we restrict attention to binary probabilistic teams and propositional logic extended with quantifiers (see [17] for related work). We define the syntax of *quantified propositional logic* QPL by the following grammar

$$\phi ::= p \mid \neg p \mid \phi \vee \phi \mid \phi \wedge \phi \mid \exists p \phi \mid \forall p \phi, \quad (8)$$

where  $p$  is a proposition variable. The probabilistic team semantics of QPL is defined analogously to that of first-order formulae. We say that a probabilistic team  $\mathbb{X} : X \rightarrow [0, 1]$  is *binary* if  $X$  assigns variables into  $\{0, 1\}$ . For a QPL formula  $\phi$  and a binary probabilistic team  $\mathbb{X} : X \rightarrow [0, 1]$ , we write  $\mathbb{X} \models \phi$  iff  $\mathfrak{A} \models_{\mathbb{X}} \phi^*$ , where  $\phi^*$  is the first-order formula obtained from  $\phi$  by substituting  $P(p)$  for  $p$  and  $\neg P(p)$  for  $\neg p$ , and letting  $\mathfrak{A} := (\{0, 1\}, P^{\mathfrak{A}} := \{1\})$ . Furthermore, we denote classical negation by " $\sim$ ". That is, we write  $\mathbb{X} \models_{\sim} \phi$  if  $\mathbb{X} \not\models \phi$ . We let  $\text{QPL}(\sim)$  denote the logic obtained by the grammar (8) extended with  $\sim \phi$ , and denote by  $\text{QPL}(\sim, C)$  the extension of  $\text{QPL}(\sim)$  by any collection of dependencies  $C$ .

We observe that  $\text{QPL}(\sim, \perp_c, \approx)$  can be interpreted as statements of real arithmetic. As truth in real arithmetic is decidable, this gives us some fairly conservative upper bounds with respect to the complexity of satisfiability and validity of  $\text{QPL}(\sim, \perp_c, \approx)$ . We say that  $\phi \in \text{QPL}(\sim, \perp_c, \approx)$  is *satisfiable* if  $\phi$  is satisfied by some non-empty binary probabilistic team.<sup>4</sup> Also,  $\phi$  is *valid* if  $\phi$  is satisfied by all binary probabilistic teams. Note that the *free variables* of a  $\text{QPL}(\sim, C)$  formula are defined analogously to the first-order case.

**Theorem 17.** *For each  $\phi \in \text{QPL}(\sim, \perp_c)$  ( $\phi \in \text{QPL}(\sim, \approx)$ , resp.) there exists a first-order sentence  $\psi$  over vocabulary  $\{+, \times, \leq, 0, 1\}$  ( $\{+, \leq, 0\}$ , resp.) such that  $\phi$  is satisfiable iff  $(\mathbb{R}, +, \times, \leq, 0, 1) \models \psi$  ( $(\mathbb{R}, +, \leq, 0) \models \psi$ , resp.).*

*Proof.* We show that satisfiability of a formula  $\phi \in \text{QPL}(\sim, \perp_c)$  is definable in real arithmetic in terms of the non-scaled variant of probabilistic team semantics. For a given tuple  $p = (p_1, \dots, p_n)$  of proposition variables, we introduce fresh first-order variables  $s_{p=i}$  for each

<sup>4</sup> Empty team satisfies every formula without  $\sim$ ; with  $\sim$  it is a non-interesting special case [19].

propositional assignment  $s(\mathbf{p}) = \mathbf{i}$ , where  $\mathbf{i}$  is a binary string of length  $n$ . We write  $\mathbf{s}$  to denote the complete tuple of these variables. For a  $\mathbf{p}$  listing the free variables of  $\phi$ , we define

$$\psi := \exists s_{\mathbf{p}=0} \dots s_{\mathbf{p}=1} \left( \bigwedge_{\mathbf{i}} 0 \leq s_{\mathbf{p}=\mathbf{i}} \wedge \neg 0 = \sum_{\mathbf{i}} s_{\mathbf{p}=\mathbf{i}} \wedge \phi^*(\mathbf{s}) \right)$$

where the mapping  $\phi(\mathbf{p}) \mapsto \phi^*(\mathbf{s})$  is defined recursively as follows:

- If  $\phi(\mathbf{p})$  is a propositional literal, then  $\phi^*(\mathbf{s}) := \bigwedge_{s \neq \phi} s = 0$ .
- If  $\phi(\mathbf{p})$  is  $\mathbf{b} \perp_a \mathbf{c}$ , where  $\mathbf{p} = \mathbf{abcd}$  for some  $\mathbf{d}$ , then  $\phi^*(\mathbf{s})$  is defined as

$$\bigwedge_{ijk} \left( \sum_{l'} s_{abcd=ijkl'} \times \sum_{j'k'l'} s_{abcd=ij'k'l'} = \sum_{k'l'} s_{abcd=ijk'l'} \times \sum_{j'l'} s_{abcd=ij'k'l'} \right).$$

- If  $\phi(\mathbf{p})$  is  $\mathbf{a} \approx \mathbf{b}$ , where  $\mathbf{p} = \mathbf{abc}$  for some  $\mathbf{c}$ , then

$$\phi^*(\mathbf{s}) := \bigwedge_{\mathbf{i}} \sum_{j'k'} s_{abc=ij'k'} = \sum_{j'k'} s_{abc=j'ik'}.$$

- If  $\phi(\mathbf{p})$  is  $\sim \eta(\mathbf{p})$ , then  $\phi^*(\mathbf{s}) := \neg \eta^*(\mathbf{s})$ .
- If  $\phi(\mathbf{p})$  is  $\eta(\mathbf{p}) \wedge \chi(\mathbf{p})$ , then  $\phi^*(\mathbf{s}) := \eta^*(\mathbf{s}) \wedge \chi^*(\mathbf{s})$ .
- If  $\phi(\mathbf{p})$  is  $\eta(\mathbf{p}) \vee \chi(\mathbf{p})$ , then

$$\phi^*(\mathbf{s}) := \exists t_{\mathbf{p}=0} r_{\mathbf{p}=0} \dots t_{\mathbf{p}=1} r_{\mathbf{p}=1} \left( \bigwedge_{\mathbf{i}} (0 \leq t_{\mathbf{p}=\mathbf{i}} \wedge 0 \leq r_{\mathbf{p}=\mathbf{i}} \wedge s_{\mathbf{p}=\mathbf{i}} = t_{\mathbf{p}=\mathbf{i}} + r_{\mathbf{p}=\mathbf{i}}) \wedge \eta^*(\mathbf{t}) \wedge \chi^*(\mathbf{r}) \right).$$

- If  $\phi(\mathbf{p})$  is  $\exists q \eta(\mathbf{p}, q)$ , then

$$\phi^*(\mathbf{s}) := \exists t_{\mathbf{p}q=00} \dots t_{\mathbf{p}q=11} \left( \bigwedge_{ij} (0 \leq t_{\mathbf{p}q=ij} \wedge s_{\mathbf{p}=\mathbf{i}} = t_{\mathbf{p}=\mathbf{i}0} + t_{\mathbf{p}=\mathbf{i}1}) \wedge \eta(\mathbf{t}) \right).$$

- If  $\phi(\mathbf{p})$  is  $\forall y \eta(\mathbf{p}, q)$ , then

$$\phi^*(\mathbf{s}) := \exists t_{\mathbf{p}q=00} \dots t_{\mathbf{p}q=11} \left( \bigwedge_{ij} (0 \leq t_{\mathbf{p}q=ij} \wedge s_{\mathbf{p}=\mathbf{i}} = t_{\mathbf{p}=\mathbf{i}0} + t_{\mathbf{p}=\mathbf{i}1} \wedge t_{\mathbf{p}=\mathbf{i}0} = t_{\mathbf{p}=\mathbf{i}1}) \wedge \eta(\mathbf{t}) \right).$$

It is straightforward to check that the claim follows.  $\square$

From the translation above we immediately obtain some complexity bounds for the satisfiability and validity problems of quantified propositional logics over probabilistic team semantics. We write 2-EXPSPACE for the class of problems solvable in space  $O(2^{2^{p(n)}})$ , and AEXPTIME( $f(n)$ ) (2-AEXPTIME( $f(n)$ ), resp.) for the class of problems solvable by alternating Turing machine in time  $O(2^{p(n)})$  ( $O(2^{2^{p(n)}})$ , resp.) with  $f(n)$  many alternations, where  $p$  is a polynomial.

**Theorem 18.** *The satisfiability/validity problems of the logics  $\text{QPL}(\perp_c, \sim)$  and  $\text{QPL}(\approx, \sim)$  are in 2-EXPSPACE and 2-AEXPTIME( $2^{O(n)}$ ), respectively.*

*Proof.* By the proof of Theorem 17, satisfiability and validity of quantified propositional formulae can be reduced to truth of a real arithmetic sentence of size  $2^{O(n)}$ . The stated upper bounds for  $\text{QPL}(\sim, \perp_c)$  and  $\text{QPL}(\sim, \approx)$  then follow because the theory of real-closed fields,  $\text{Th}(\mathbb{R}, +, \times, \leq, 0, 1)$ , is in EXPSPACE [3], and the theory of real addition,  $\text{Th}(\mathbb{R}, +, \leq, 0)$ , is in AEXPTIME( $n$ ) [4,10].  $\square$

PTS: $\text{FO}(\approx) < \text{FO}(\approx, =(\cdot)) \equiv \text{FO}(\approx^*) \leq \text{FO}(\perp\!\!\!\perp) \equiv \text{FO}(\perp\!\!\!\perp_c)$
TS: $\text{FO}(\subseteq) < \text{FO}(\subseteq, =(\cdot)) \equiv \text{FO}(\perp) \equiv \text{FO}(\perp_c)$ [12,13]

**Table 1.** Relative expressivity in probabilistic team semantics (PTS) and team semantics (TS)

We also obtain an upper bound for the implication problem of conditional independence over binary probability distributions. The *implication problem* for conditional independence is given as a finite set  $\Sigma \cup \{\sigma\}$  of conditional independence statements, and the problem is to decide whether all probability distributions that satisfy  $\Sigma$  satisfy also  $\sigma$ . It is a famous open problem to determine whether implication of conditional independence is decidable over discrete distributions. Since binary probabilistic teams can be interpreted as discrete distributions of binary random variables, we obtain that the implication problem for conditional independence statements is decidable in exponential space over binary distributions. The result follows since any instance of such an implication problem can be expressed as an existential formula of exponential size (Theorem 17), and since the existential theory of real-closed fields is in PSPACE [5].

**Corollary 19.** *The implication problem for conditional independence over binary probability distributions is in EXPSPACE.*

It may be conjectured that the obtained complexity bounds are not optimal. The first-order translations provide only access to a very restricted type of arithmetic expressions. For instance, real multiplication is only available between sums of reals from the unit interval. We leave it as an open problem to determine whether the results of this section can be optimized using more refined arguments.

## 6 Conclusions and further directions

We have studied probabilistic team semantics in association with three notions of dependency atoms: probabilistic independence, marginal identity, and marginal distribution equivalence atoms. Our investigations give rise to an overall classification that is already familiar from the team semantics context (see Table 1). Similar to inclusion logic ( $\text{FO}(\subseteq)$ ) in team semantics, we observed that  $\text{FO}(\approx)$  enjoys a union closure property which renders it strictly less expressive than  $\text{FO}(\approx, =(\cdot))$ . A further analogous fact is that both dependence and marginal identity are definable with conditional independence, which in turn is definable using only marginal independence. An interesting open question is to determine the relationship between  $\text{FO}(\approx, =(\cdot))$  (or equivalently  $\text{FO}(\approx^*)$ ) and  $\text{FO}(\perp\!\!\!\perp_c)$ . Contrary to the picture arising from team semantics, we conjecture that the latter is strictly more expressive.

One motivation behind our marginal distribution equivalence atom was that it seemed to be weaker than marginal identity but still enough to guarantee the same entropy of two distributions. A natural next step would be to consider some form of entropy atom/atoms and study the expressive power of the resulting logics. The exact formulation of such atoms will make all the difference, as one can detect both functional dependencies and marginal independence if one has full access to the conditional entropy as a function.

We also studied (quantified) propositional logics with probabilistic team semantics. By connecting real-valued probabilistic teams to real arithmetic we showed upper bounds for computational problems associated with these logics. As a consequence of our translation to real arithmetic we also obtained an EXPSPACE upper bound for the implication problem of conditional independence statements over binary distributions.

## References

1. Abramsky, S.: Relational hidden variables and non-locality. *Studia Logica* **101**(2), 411–452 (2013)
2. Barbero, F., Sandu, G.: Interventionist counterfactuals on causal teams. In: Finkbeiner, B., Kleinberg, S. (eds.) *Proceedings 3rd Workshop on formal reasoning about Causation, Responsibility, and Explanations in Science and Technology*, Thessaloniki, Greece, 21st April 2018. *Electronic Proceedings in Theoretical Computer Science*, vol. 286, pp. 16–30. Open Publishing Association (2019). <https://doi.org/10.4204/EPTCS.286.2>
3. Ben-Or, M., Kozen, D., Reif, J.: The complexity of elementary algebra and geometry. *Journal of Computer and System Sciences* **32**(2), 251 – 264 (1986)
4. Berman, L.: The complexity of logical theories. *Theoretical Computer Science* **11**(1), 71 – 77 (1980)
5. Canny, J.: Some algebraic and geometric computations in pspace. In: *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*. pp. 460–467. STOC '88, ACM, New York, NY, USA (1988)
6. Cavallo, R., Pittarelli, M.: The theory of probabilistic databases. In: *Proceedings of the 13th International Conference on Very Large Data Bases*. pp. 71–81. VLDB '87, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1987)
7. Corander, J., Hyttinen, A., Kontinen, J., Pensar, J., Väänänen, J.: A logical approach to context-specific independence. In: Väänänen, J.A., Hirvonen, Å., de Queiroz, R.J.G.B. (eds.) *Logic, Language, Information, and Computation - 23rd International Workshop, WoLLIC 2016, Puebla, Mexico, August 16-19th, 2016*. *Proceedings. Lecture Notes in Computer Science*, vol. 9803, pp. 165–182. Springer (2016). [https://doi.org/10.1007/978-3-662-52921-8\\_11](https://doi.org/10.1007/978-3-662-52921-8_11)
8. Durand, A., Hannula, M., Kontinen, J., Meier, A., Virtema, J.: Approximation and dependence via multiteam semantics. *Ann. Math. Artif. Intell.* **83**(3-4), 297–320 (2018), <https://doi.org/10.1007/s10472-017-9568-4>
9. Durand, A., Hannula, M., Kontinen, J., Meier, A., Virtema, J.: Probabilistic team semantics. In: *FoIKS. Lecture Notes in Computer Science*, vol. 10833, pp. 186–206. Springer (2018). [https://doi.org/10.1007/978-3-319-90050-6\\_11](https://doi.org/10.1007/978-3-319-90050-6_11)
10. Ferrante, J., Rackoff, C.: A decision procedure for the first order theory of real addition with order. *SIAM J. Comput.* **4**(1), 69–76 (1975). <https://doi.org/10.1137/0204006>
11. Galliani, P.: *Game Values and Equilibria for Undetermined Sentences of Dependence Logic* (2008), MSc Thesis. ILLC Publications, MoL–2008–08
12. Galliani, P.: Inclusion and exclusion dependencies in team semantics: On some logics of imperfect information. *Annals of Pure and Applied Logic* **163**(1), 68 – 84 (2012)
13. Galliani, P., Väänänen, J.: On dependence logic. In: Baltag, A., Smets, S. (eds.) *Johan van Benthem on Logic and Information Dynamics*, pp. 101–119. Springer (2014). [https://doi.org/10.1007/978-3-319-06025-5\\_4](https://doi.org/10.1007/978-3-319-06025-5_4)
14. Grädel, E., Gurevich, Y.: Metafinite model theory. *Inf. Comput.* **140**(1), 26–81 (1998). <https://doi.org/10.1006/inco.1997.2675>
15. Grädel, E., Väänänen, J.: Dependence and independence. *Studia Logica* **101**(2), 399–410 (2013). <https://doi.org/10.1007/s11225-013-9479-2>
16. Hannula, M., Kontinen, J.: A finite axiomatization of conditional independence and inclusion dependencies. *Inf. Comput.* **249**, 121–137 (2016). <https://doi.org/10.1016/j.ic.2016.04.001>
17. Hannula, M., Kontinen, J., Lück, M., Virtema, J.: On quantified propositional logics and the exponential time hierarchy. In: *GandALF. EPTCS*, vol. 226, pp. 198–212 (2016)
18. Hannula, M., Kontinen, J., Virtema, J.: Polyteam semantics. In: *Logical Foundations of Computer Science - International Symposium, LFCS 2018, Deerfield Beach, FL, USA, January 8-11, 2018*, *Proceedings*. pp. 190–210 (2018). [https://doi.org/10.1007/978-3-319-72056-2\\_12](https://doi.org/10.1007/978-3-319-72056-2_12)
19. Hannula, M., Kontinen, J., Virtema, J., Vollmer, H.: Complexity of propositional logics in team semantic. *ACM Trans. Comput. Log.* **19**(1), 2:1–2:14 (2018). <https://doi.org/10.1145/3157054>
20. Hodges, W.: *Compositional Semantics for a Language of Imperfect Information*. *Journal of the Interest Group in Pure and Applied Logics* **5** (4), 539–563 (1997)
21. Hyttinen, T., Paolini, G., Väänänen, J.: A Logic for Arguing About Probabilities in Measure Teams. *Arch. Math. Logic* **56**(5-6), 475–489 (2017). <https://doi.org/10.1007/s00153-017-0535-x>
22. Krebs, A., Meier, A., Virtema, J., Zimmermann, M.: Team Semantics for the Specification and Verification of Hyperproperties. In: Potapov, I., Spirakis, P., Worrell, J. (eds.) *43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018)*. *Leibniz International Proceedings in Informatics (LIPIcs)*, vol. 117, pp. 10:1–10:16. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany (2018). <https://doi.org/10.4230/LIPIcs.MFCS.2018.10>
23. Lück, M.: Canonical models and the complexity of modal team logic. In: *27th EACSL Annual Conference on Computer Science Logic, CSL 2018, September 4-7, 2018, Birmingham, UK*. pp. 30:1–30:23 (2018). <https://doi.org/10.4230/LIPIcs.CSL.2018.30>
24. Niepert, M., Gyssens, M., Sayrafi, B., Gucht, D.V.: On the conditional independence implication problem: A lattice-theoretic approach. *Artif. Intell.* **202**, 29–51 (2013). <https://doi.org/10.1016/j.artint.2013.06.005>
25. Väänänen, J.: *Dependence Logic*. Cambridge University Press (2007)

## A Proof of Proposition 8

**Proposition 8.** *Let  $\mathfrak{A}$  be a structure,  $\mathbb{X} : X \rightarrow \mathbb{R}_{\geq 0}$  a probabilistic team of  $\mathfrak{A}$ , and  $\phi \in \text{FO}(\perp\!\!\!\perp_c, \approx, \approx^*, =(\cdot))$ . Then  $\mathfrak{A} \models_{\mathbb{X}}^{\geq 0} \phi \Leftrightarrow \mathfrak{A} \models_{\frac{1}{|\mathbb{X}|} \cdot \mathbb{X}}^{[0,1]} \phi$ .*

*Proof.* The cases for first-order literals,  $\approx$ ,  $\approx^*$ ,  $=(\cdot)$  and the conjunction are immediate. The claim for the independence atom  $\mathbf{y} \perp\!\!\!\perp_x \mathbf{z}$  follows from the equivalence below together with the observation that the former is the definition of the atom in the unscaled team  $\mathbb{X}$  whereas the latter is equivalent to that of the scaled team  $\frac{1}{|\mathbb{X}|} \cdot \mathbb{X}$ .

$$\begin{aligned} |\mathbb{X}_{xy=s(xy)}| \cdot |\mathbb{X}_{xz=s(xz)}| &= |\mathbb{X}_{xyz=s(xyz)}| \cdot |\mathbb{X}_{x=s(x)}| \text{ if and only if} \\ \frac{1}{|\mathbb{X}|} \cdot |\mathbb{X}_{xy=s(xy)}| \cdot \frac{1}{|\mathbb{X}|} \cdot |\mathbb{X}_{xz=s(xz)}| &= \frac{1}{|\mathbb{X}|} \cdot |\mathbb{X}_{xyz=s(xyz)}| \cdot \frac{1}{|\mathbb{X}|} \cdot |\mathbb{X}_{x=s(x)}|. \end{aligned}$$

The case for disjunction follows from the following chain of equivalences

$$\begin{aligned} \mathfrak{A} \models_{\mathbb{X}}^{\geq 0} \phi \vee \psi &\Leftrightarrow \mathfrak{A} \models_{\mathbb{Y}}^{\geq 0} \phi \text{ and } \mathfrak{A} \models_{\mathbb{Z}}^{\geq 0} \psi \text{ for some } \mathbb{Y} \text{ and } \mathbb{Z} \text{ s.t. } \mathbb{Y} \sqcup \mathbb{Z} = \mathbb{X} \\ &\Leftrightarrow \mathfrak{A} \models_{\frac{1}{|\mathbb{Y}|} \cdot \mathbb{Y}}^{[0,1]} \phi \text{ and } \mathfrak{A} \models_{\frac{1}{|\mathbb{Z}|} \cdot \mathbb{Z}}^{[0,1]} \psi \text{ for some } \mathbb{Y} \text{ and } \mathbb{Z} \text{ s.t. } \mathbb{Y} \sqcup \mathbb{Z} = \mathbb{X} \\ &\Leftrightarrow \mathfrak{A} \models_{\frac{1}{|\mathbb{X}|} \cdot \mathbb{X}}^{[0,1]} \phi \vee \psi, \end{aligned}$$

where the last equivalence follows from the definition of the disjunction for  $k = \frac{|\mathbb{Y}|}{|\mathbb{X}|}$  and  $1 - k = \frac{|\mathbb{Z}|}{|\mathbb{X}|}$ , since

$$\frac{|\mathbb{Y}|}{|\mathbb{X}|} \cdot \frac{1}{|\mathbb{Y}|} \cdot \mathbb{Y} + \frac{|\mathbb{Z}|}{|\mathbb{X}|} \cdot \frac{1}{|\mathbb{Z}|} \cdot \mathbb{Z} = \frac{1}{|\mathbb{X}|} \cdot \mathbb{Y} + \frac{1}{|\mathbb{X}|} \cdot \mathbb{Z} = \frac{1}{|\mathbb{X}|} \cdot \mathbb{X}.$$

The cases for the quantifiers are similar; we show the case for the universal quantifier

$$\begin{aligned} \mathfrak{A} \models_{\mathbb{X}}^{\geq 0} \forall x \phi &\Leftrightarrow \mathfrak{A} \models_{\mathbb{X}[A/x]}^{\geq 0} \phi \Leftrightarrow \mathfrak{A} \models_{\frac{1}{|\mathbb{X}[A/x]|} \cdot \mathbb{X}[A/x]}^{[0,1]} \phi \Leftrightarrow \mathfrak{A} \models_{\left(\frac{1}{|\mathbb{X}|} \cdot \mathbb{X}\right)[A/x]}^{[0,1]} \phi \\ &\Leftrightarrow \mathfrak{A} \models_{\frac{1}{|\mathbb{X}|} \cdot \mathbb{X}}^{[0,1]} \forall x \phi, \end{aligned}$$

where the second last equivalence follows, since  $|\mathbb{X}[A/x]| = |\mathbb{X}|$  and  $\left(\frac{1}{|\mathbb{X}|} \cdot \mathbb{X}\right)[A/x] = \frac{1}{|\mathbb{X}|} \cdot \mathbb{X}[A/x]$ .  $\square$

## B Proof of Lemma 9

**Lemma 9.** *Let  $\mathfrak{A}$  be structure with at least two elements and  $\mathbf{z}$  an  $n$ -tuple of variables. Let  $\phi(\mathbf{z}, d, c_1, c_2)$  be a formula such that for all probabilistic teams  $\mathbb{X}$ , whose variable domain includes  $\mathbf{z}, d, c_1, c_2$  and for which  $\mathfrak{A} \models_{\mathbb{X}} c_1 \neq c_2$  and  $\mathfrak{A} \models_{\mathbb{X}} =(c_1) \wedge =(c_2)$ , it holds that*

$$\mathcal{M} \models_{\mathbb{X}} \phi \quad \Leftrightarrow \quad d \text{ is uniformly distributed over the two values of } c_1, c_2 \quad (9)$$

and  $d$  is independent of  $\mathbf{z}$ .

Then  $\mathbf{x} \approx \mathbf{y}$  can be expressed for  $n$ -tuples  $\mathbf{x}$  and  $\mathbf{y}$  using  $\phi$  and the constancy atom.

*Proof.* We will write a formula  $\psi(\mathbf{x}, \mathbf{y})$  which is to be equivalent with  $\mathbf{x} \approx \mathbf{y}$ . But first we need to define an auxiliary formula  $\theta$ . Define

$$\theta := (d = c_1 \wedge \mathbf{z} = \mathbf{x}) \vee (d = c_2 \wedge \mathbf{z} = \mathbf{y}).$$

This formula says that  $\mathbf{z}$  always equals either  $\mathbf{x}$  or  $\mathbf{y}$  and  $d$  is a “detector” for which one it is. We use the abbreviation  $\exists^c c_1 c_2$  below to denote  $\exists c_1 \exists c_2 (=(c_1) \wedge =(c_2) \wedge c_1 \neq c_2)$ . Now define

$$\psi(\mathbf{x}, \mathbf{y}) := \exists^c c_1 c_2 \left[ \forall \mathbf{z} \exists d \left( (\mathbf{x} = \mathbf{y}) \vee [(\mathbf{x} \neq \mathbf{y}) \wedge ((\mathbf{z} \neq \mathbf{x} \wedge \mathbf{z} \neq \mathbf{y}) \vee [\theta \wedge \phi])] \right) \right].$$

Suppose  $\mathbf{x} \approx \mathbf{y}$  holds in a team  $\mathbb{X}$  over variables  $\mathbf{x}$  and  $\mathbf{y}$ . We want to show that  $\psi(\mathbf{x}, \mathbf{y})$  is satisfied by  $\mathbb{X}$ . Let  $\mathbb{X}_1$  be the expansion of  $\mathbb{X}$  obtained by the quantification of  $c_1, c_2$ , and  $z$ . We may assume that  $c_1, c_2$  were picked such that they attain constant but distinct values. Also note that  $z$  is independent of all other variables and uniformly distributed over the domain of  $\mathfrak{A}$ . Now let  $d$  be a variable that takes its values from the values of  $c_1$  and  $c_2$  such that it “detects” whether  $z$  equals  $\mathbf{x}$  or not (value of  $d$  is the value of  $c_1$  iff  $z$  and  $\mathbf{x}$  have the same value). Let  $\mathbb{X}_2$  be the expansion of  $\mathbb{X}_1$  by this  $d$ . We need to check that  $\mathbb{X}_2$  satisfies

$$(\mathbf{x} = \mathbf{y}) \vee [(\mathbf{x} \neq \mathbf{y}) \wedge ((z \neq \mathbf{x} \wedge z \neq \mathbf{y}) \vee [\theta \wedge \phi])].$$

Let  $\mathbb{X}_3$  be the maximal subteam of  $\mathbb{X}_2$  where  $\mathbf{x} \neq \mathbf{y}$ . So now we have to check that

$$(z \neq \mathbf{x} \wedge z \neq \mathbf{y}) \vee [\theta \wedge \phi] \tag{10}$$

holds in  $\mathbb{X}_3$ . Recall that  $\theta$  says in particular that  $z$  equals either  $\mathbf{x}$  or  $\mathbf{y}$ , so (10) holds in  $\mathbb{X}_3$  if and only if  $\theta \wedge \phi$  holds in the maximal subteam  $\mathbb{X}_4$  of  $\mathbb{X}_3$  in which this is the case. We also just defined  $d$  to attain the value  $c_1$  if and only if  $z = \mathbf{x}$  and the only other option is that  $z = \mathbf{y}$  in which case  $d = c_2$ , so  $\theta$  is satisfied. What about  $\phi$ ; note that  $\mathbb{X}_4$  is such that (9) holds. Now fix any value  $v$  of  $z$  in  $\mathbb{X}_4$ . Since  $\mathbf{x} \approx \mathbf{y}$  holds, we have  $|\mathbb{X}_{\mathbf{x}=v}| = |\mathbb{X}_{\mathbf{y}=v}|$ . When we expand  $\mathbb{X}$  to  $\mathbb{X}_1$  and further to  $\mathbb{X}_2$  this property is (clearly) preserved. It is also preserved when we take the subteam  $\mathbb{X}_3$ , because when we move from  $\mathbb{X}_2$  to  $\mathbb{X}_3$ , we only remove assignments  $s$  where  $s(\mathbf{x}) = s(\mathbf{y})$ , so if an assignment with  $\mathbf{x} = v$  is deleted, then also an assignment with  $\mathbf{y} = v$  is deleted (the same assignment). When we move to  $\mathbb{X}_4$  we still have  $|(\mathbb{X}_4)_{\mathbf{x}=v}| = |(\mathbb{X}_4)_{\mathbf{y}=v}|$  which follows from the fact that  $z$  is independent of  $\mathbf{x}, \mathbf{y}, c_1, c_2$ . Therefore

$$|(\mathbb{X}_4)_{\mathbf{x}z=vv}| = |(\mathbb{X}_4)_{\mathbf{y}z=vv}|.$$

But this means that conditioned on  $z = v$ ,  $d$  is uniformly distributed in  $\mathbb{X}_4$ . Since this holds for any  $v$ ,  $d$  is uniformly distributed and independent of  $z$  as desired and  $\psi(\mathbf{x}, \mathbf{y})$  is satisfied by  $\mathbb{X}$ .

Suppose now that a team  $\mathbb{X}$  satisfies  $\psi(\mathbf{x}, \mathbf{y})$ . We want to show that  $\mathbf{x} \approx \mathbf{y}$ . But the chain of reasoning above also works “backwards”. Fix a value  $v$  of  $\mathbf{x}$ . We want to show that  $|\mathbb{X}_{\mathbf{x}=v}| = |\mathbb{X}_{\mathbf{y}=v}|$ . It is clear that it is sufficient to look at  $\mathbb{X}_3$  as defined above. But because  $\theta$  says that  $d$  is a “detector” of whether  $z = \mathbf{x}$  or not, it is in fact sufficient to check  $\mathbf{x} \approx \mathbf{y}$  for the subteam  $\mathbb{X}_4$  (also as defined above). But in  $\mathbb{X}_4$ , this follows from  $\phi$ .  $\square$

## C Proof of Theorem 11

Theorem 11 follows from Lemma 22 presented below. Lemma 22 can be proven following the proof of Theorem 2 in [9]. We omit the details and instead delineate intuition behind the translation. The idea is to simulate the semantics of the probabilistic conditional independence atom using only marginal independence and marginal identity atoms. First, the universally quantified  $\mathbf{y}$  in the translation represents all possible variable assignments  $s$  of  $\mathbf{x}$ . Second,  $\psi_0$  and  $\psi_1$  indicate that the marginal distributions of  $\mathbf{x}_0, \mathbf{x}_0\mathbf{x}_1, \mathbf{x}_0\mathbf{x}_2$ , and  $\mathbf{x}_0\mathbf{x}_1\mathbf{x}_2$  are distributed respectively to  $z_0, z_1, z_2, z_3$  independently of  $\mathbf{y}$  and of each other. Third,  $\psi_2$  encodes the product of the weights of  $s(\mathbf{x}_0)$  and  $s(\mathbf{x}_0\mathbf{x}_1\mathbf{x}_2)$  by  $\alpha = 0$ , and  $\psi_3$  similarly the product of the weights of  $s(\mathbf{x}_0\mathbf{x}_1)$  and  $s(\mathbf{x}_0\mathbf{x}_2)$  by  $\bar{=}0$ . Finally, conditional independence between  $\mathbf{x}_1$  and  $\mathbf{x}_2$  given  $\mathbf{x}_0$  follows iff these products are equal relative to all assignments of  $\mathbf{y}$ . Theorem 11 then follows from this lemma since the constant 0 and the marginal identity atom are both definable in  $\text{FO}(\perp\!\!\!\perp)$ .

**Lemma 22.** *Let  $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2$  be three sequences of variables from  $\mathbf{x} = (x_1, \dots, x_n)$ , and let 0 be a constant symbol. Then  $\mathbf{x}_1 \perp\!\!\!\perp_{\mathbf{x}_0} \mathbf{x}_2$  is equivalent to*

$$\phi := \forall \mathbf{y} \exists z_0 z_1 z_2 z_3 \alpha (\psi_0 \wedge \psi_1 \wedge \psi_2 \wedge \psi_3 \wedge \psi_4)$$



where

$$\begin{aligned}
\psi_0 &:= \mathbf{y} \perp\!\!\!\perp \mathbf{z}_0 \wedge \mathbf{y}\mathbf{z}_0 \perp\!\!\!\perp \mathbf{z}_1 \wedge \mathbf{y}\mathbf{z}_0\mathbf{z}_1 \perp\!\!\!\perp \mathbf{z}_2 \wedge \mathbf{y}\mathbf{z}_0\mathbf{z}_1\mathbf{z}_2 \perp\!\!\!\perp \mathbf{z}_3, \\
\psi_1 &:= \mathbf{x}_0 \approx \mathbf{z}_0 \wedge \mathbf{x}_0\mathbf{x}_1 \approx \mathbf{z}_1 \wedge \mathbf{x}_0\mathbf{x}_2 \approx \mathbf{z}_2 \wedge \mathbf{x}_0\mathbf{x}_1\mathbf{x}_2 \approx \mathbf{z}_3, \\
\psi_2 &:= \alpha = 0 \leftrightarrow (\mathbf{z}_0 = \mathbf{y}_0 \wedge \mathbf{z}_3 = \mathbf{y}_0\mathbf{y}_1\mathbf{y}_2), \\
\psi_3 &:= \equiv 0 \leftrightarrow (\mathbf{z}_1 = \mathbf{y}_0\mathbf{y}_1 \wedge \mathbf{z}_2 = \mathbf{y}_0\mathbf{y}_2), \\
\psi_4 &:= \mathbf{y}\alpha \approx \mathbf{y}.
\end{aligned}$$

## D Proof of Proposition 15

**Proposition 15.** *Let  $\mathfrak{A}$  be a model,  $\phi \in \text{FO}(\approx)$  a formula, and  $\mathbb{X} : X \rightarrow [0, 1]$  and  $\mathbb{Y} : X \rightarrow [0, 1]$  two probabilistic teams. Then for all  $k \in [0, 1]$ :*

$$\text{if } \mathfrak{A} \models_{\mathbb{X}} \phi \text{ and } \mathfrak{A} \models_{\mathbb{Y}} \phi, \text{ then } \mathfrak{A} \models_{\mathbb{X} \sqcup_k \mathbb{Y}} \phi.$$

*Proof.* We may assume that  $\mathbb{X} = (X, f)$  and  $\mathbb{Y} = (X, g)$ . We prove the claim by structural induction on  $\phi$ . We omit the cases for atomic formulae and conjunction which are straightforward.

- Assume that  $\phi = \phi_0 \vee \phi_1$ . By the semantics of the disjunction, we find  $p, q \in [0, 1]$  and distributions  $f_0, f_1, g_0, g_1$  over  $X$  such that  $\mathfrak{A} \models_{(X, f_0)} \phi_0$ ,  $\mathfrak{A} \models_{(X, f_1)} \phi_1$ ,  $\mathfrak{A} \models_{(X, g_0)} \phi_0$ ,  $\mathfrak{A} \models_{(X, g_1)} \phi_1$ ,  $f = pf_0 + (1-p)f_1$ , and  $g = qg_0 + (1-q)g_1$ . Define  $h_0 := \frac{kpf_0 + (1-k)qg_0}{kp + (1-k)q}$  and  $h_1 := \frac{k(1-p)f_1 + (1-k)(1-q)g_1}{k(1-p) + (1-k)(1-q)}$ . By the induction hypothesis  $\mathfrak{A} \models_{(X, h_0)} \phi_0$  and  $\mathfrak{A} \models_{(X, h_1)} \phi_1$ , since  $(X, h_0) = (X, f_0) \sqcup_a (X, g_0)$  for  $a := \frac{kp}{kp + (1-k)q}$ , and  $(X, h_1) = (X, f_1) \sqcup_b (X, g_1)$  for  $b := \frac{k(1-p)}{k(1-p) + (1-k)(1-q)}$ . Then  $(X, f) \sqcup_k (X, g) = (X, h_0) \sqcup_c (X, h_1)$  for  $c := kp + (1-k)q$  because

$$\begin{aligned}
ch_0 + (1-c)h_1 &= c \frac{kpf_0 + (1-k)qg_0}{c} + (1-c) \frac{k(1-p)f_1 + (1-k)(1-q)g_1}{1-c} \\
&= k[pf_0 + (1-p)f_1] + (1-k)[qg_0 + (1-q)g_1] \\
&= kf + (1-k)g.
\end{aligned}$$

Consequently,  $\mathfrak{A} \models_{(X, f) \sqcup_k (X, g)} \phi_0 \vee \phi_1$  follows from the semantics of the disjunction which completes the disjunction step of the induction.

- Assume that  $\phi = \forall x \psi$ . Then  $\mathfrak{A} \models_{\mathbb{X}[A/x]} \psi$  and  $\mathfrak{A} \models_{\mathbb{Y}[A/x]} \psi$ , and by induction assumption  $\mathfrak{A} \models_{\mathbb{X}[A/x] \sqcup_k \mathbb{Y}[A/x]} \psi$ . The claim then follows since  $\mathbb{X}[A/x] \sqcup_k \mathbb{Y}[A/x] = (\mathbb{X} \sqcup_k \mathbb{Y})[A/x]$ .
- Assume that  $\phi = \exists x \psi$ . Then  $\mathfrak{A} \models_{\mathbb{X}[F/x]} \psi$  and  $\mathfrak{A} \models_{\mathbb{Y}[G/x]} \psi$  where  $F$  and  $G$  are functions that map each  $s \in X$  to a probability distribution  $F_s$  over  $A = \text{Dom}(\mathfrak{A})$ . We let  $H$  be a function that maps  $s \in X$  to a probability distribution  $H_s$  over  $A$  such that

$$H_s(a) := \frac{kf(s)F_s(a) + (1-k)g(s)G_s(a)}{kf(s) + (1-k)g(s)}.$$

Note that  $\sum_{a \in A} H_s(a) = 1$  follows from  $\sum_{a \in A} F_s(a) = \sum_{a \in A} G_s(a) = 1$ . By induction assumption  $\mathfrak{A} \models_{\mathbb{X}[F/x] \sqcup_k \mathbb{Y}[G/x]} \psi$ . The claim now follows from  $\mathbb{X}[F/x] \sqcup_k \mathbb{Y}[G/x] = (\mathbb{X} \sqcup_k \mathbb{Y})[H/x]$ , which holds since for all  $a \in A$ :

$$kf(s)F_s(a) + (1-k)g(s)G_s(a) = [kf(s) + (1-k)g(s)]H_s(a).$$

This concludes the case of existential quantification and the proof.  $\square$