Probabilistic team semantics

Miika Hannula
(joint work with Åsa Hirvonen, Juha Kontinen, Vadim Kulikov, Jonni Virtema)

University of Helsinki, Finland

UH-CAS workshop, October 31, 2018

Team semantics

Semantical framework for logics that describe properties of sets of objects

Team := set of assignments (e.g., a data table)

x	y	z
a	b	c
a	b	d
b	a	c
b	a	d

"y depends on x"

"z does not depend on x"

"x is not independent of y"

"x is included in y"

Team semantics

Team semantics not ideal for dealing with counting/probability properties

x	y	z
a	b	c
a	b	d
b	b	c
b	a	d

"x=a is occurs twice as many times as y=a"

"x and y are distributed differently"

"x and y are independent random variables"

Multiteam Semantics/Probabilistic team semantics

Add a multiplicity count/probability measure for each assignment.

x	y	z
a	b	c
a	b	c
a	b	d
b	a	c
b	a	d

x	y	z	#
a	b	c	2
a	b	d	1
b	a	c	1
b	a	d	1

x	y	z	P
a	b	c	$\frac{2}{5}$
a	b	d	$\frac{1}{5}$
b	a	c	$\frac{1}{5}$
b	a	d	$\frac{1}{5}$

Distributions

Consider:

- A collection of data from some repetitive science experiment.
- Data obtained from a poll.
- Any collection of data, that involves meaningful duplicates of data.

One natural way to represent the data is to use multisets (sets with duplicates).

Often the multiplicities themselves are not important; the distribution of data is:

- The locations of the electrons of an atom.
- Pre-election poll of party support.
- Distribution of a population with attributes like education, salary, and age.

Distributions

Consider:

- A collection of data from some repetitive science experiment.
- Data obtained from a poll.
- Any collection of data, that involves meaningful duplicates of data.

One natural way to represent the data is to use multisets (sets with duplicates).

Often the multiplicities themselves are not important; the distribution of data is:

- The locations of the electrons of an atom.
- Pre-election poll of party support.
- Distribution of a population with attributes like education, salary, and age.

Probabilistic team semantics

Semantical framework for logics that describe properties of probability distributions.

x	y	z	#
a	b	c	$\frac{2}{5}$
a	b	d	$\frac{1}{5}$
b	a	c	$\frac{1}{5}$
b	a	d	$\frac{1}{5}$

"x=a is more likely than y=a"
"x and y are distributed differently"
"x is not independent of y"

Distributions

Definition

A (discrete) distribution is a mapping $f : A \to [0,1]$ from a set A of values to the unit interval $[0,1]$ such that the probabilities sum to 1, i.e.,

$$\sum_{a \in A} f(a) = 1.$$

- A team is a set of first-order assignments (a database without duplicates).
- A probabilistic team is a pair (X, p), where X is a finite team and $p : X \to [0,1]$ is a distribution (distribution of data).

Distributions

Definition

A (discrete) distribution is a mapping $f : A \to [0, 1]$ from a set A of values to the unit interval $[0, 1]$ such that the probabilities sum to 1, i.e.,

$$\sum_{a \in A} f(a) = 1.$$

- A team is a set of first-order assignments (a database without duplicates).
- A probabilistic team is a pair (X, p), where X is a finite team and $p : X \to [0, 1]$ is a distribution (distribution of data).

Probabilistic teams

- Modelling of data that is inherently a probability distribution.
- Abstraction of data with duplicates.

We introduce a logic that describe properties of probabilistic teams.

We consider the expansion of first-order logic with the marginal identity atoms

$$(x_1, \ldots, x_n) \approx (y_1, \ldots, y_n)$$

and with the probabilistic conditional independence atoms

$$\vec{y} \perp\!\!\!\perp_{\vec{x}} \vec{z}.$$

Probabilistic teams

- Modelling of data that is inherently a probability distribution.
- Abstraction of data with duplicates.

We introduce a logic that describe properties of probabilistic teams.

We consider the expansion of first-order logic with the marginal identity atoms

$$(x_1, \ldots, x_n) \approx (y_1, \ldots, y_n)$$

and with the probabilistic conditional independence atoms

$$\vec{y} \perp\!\!\!\perp_{\vec{x}} \vec{z}.$$

Probabilistic atoms

Let $\mathbb{X} = (X, p)$ be a probablistic team and $\vec{x}, \vec{a}$ be tuples of variables and values of length k. We define

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

Probabilistic atoms

Let $\mathbb{X} = (X, p)$ be a probablistic team and $\vec{x}, \vec{a}$ be tuples of variables and values of length k. We define

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

We define that

$$\mathfrak{A} \models_{\mathbb{X}} \vec{x} \approx \vec{y} \text{ iff } |\mathbb{X}|_{\vec{x}=\vec{a}} = |\mathbb{X}|_{\vec{y}=\vec{a}}, \text{ for each } \vec{a} \in A^k,$$

Probabilistic atoms

Let $\mathbb{X} = (X, p)$ be a probablistic team and $\vec{x}, \vec{a}$ be tuples of variables and values of length k. We define

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

We define that $\mathfrak{A} \models_{\mathbb{X}} \vec{y} \perp\!\!\!\perp_{\vec{x}} \vec{z}$ iff, for all assignments s for $\vec{x}, \vec{y}, \vec{z}$

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

Semantics of complex formulae

Definition

Let $\mathfrak{A}$ be a structure over a *finite* domain A, and $\mathbb{X} = (X, p)$ a probabilistic team of $\mathfrak{A}$. The satisfaction relation $\models_{\mathbb{X}}$ for first-order logic is defined as follows:

$$\mathfrak{A} \models_{\mathbb{X}} x = y \Leftrightarrow \text{for all } s \in X : \text{if } p(s) > 0, \text{ then } s(x) = s(y)$$

$$\mathfrak{A} \models_{\mathbb{X}} x \neq y \Leftrightarrow \text{for all } s \in X : \text{if } p(s) > 0, \text{ then } s(x) \neq s(y)$$

$$\mathfrak{A} \models_{\mathbb{X}} R(\vec{x}) \Leftrightarrow \text{for all } s \in X : \text{if } p(s) > 0, \text{ then } s(\vec{x}) \in R^{\mathfrak{A}}$$

$$\mathfrak{A} \models_{\mathbb{X}} \neg R(\vec{x}) \Leftrightarrow \text{for all } s \in X : \text{if } p(s) > 0, \text{ then } s(\vec{x}) \notin R^{\mathfrak{A}}$$

$$\mathfrak{A} \models_{\mathbb{X}} (\psi \wedge \theta) \Leftrightarrow \mathfrak{A} \models_{\mathbb{X}} \psi \text{ and } \mathfrak{A} \models_{\mathbb{X}} \theta$$

Semantics of complex formulae

Definition

Let $\mathfrak{A}$ be a structure over a *finite* domain A, and $\mathbb{X} = (X, p)$ a probabilistic team of $\mathfrak{A}$. The satisfaction relation $\models_{\mathbb{X}}$ for first-order logic is defined as follows:

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

$$\mathfrak{A} \models_{\mathbb{X}} \forall x \psi \Leftrightarrow \mathfrak{A} \models_{\mathbb{X}[A/x]} \psi$$

$$\mathfrak{A} \models_{\mathbb{X}} \exists x \psi \Leftrightarrow \mathfrak{A} \models_{\mathbb{X}[F/x]} \psi \text{ holds for some } F \colon X \to p_A.$$

Above p_A denote the set those of distributions that have domain A.

Intuition of the disjunction

Question: How do we split distributions?
Answer: We rescale.

Let $\mathbb{X}\colon X \to [0,1]$ and $\mathbb{Y}\colon Y \to [0,1]$ be probabilistic teams and let $k \in [0,1]$.

We denote by $\mathbb{X} \sqcup_k \mathbb{Y}$ the k-scaled union of $\mathbb{X}$ and $\mathbb{Y}$, that is, the probabilistic team $\mathbb{X} \sqcup_k \mathbb{Y}\colon X \cup Y \to [0,1]$ defined s.t. for each $s \in X \cup Y$,

$$(\mathbb{X} \sqcup_k \mathbb{Y})(s) := \begin{cases} k \cdot \mathbb{X}(s) + (1-k) \cdot \mathbb{Y}(s) & \text{if } s \in X \text{ and } s \in Y, \\ k \cdot \mathbb{X}(s) & \text{if } s \in X \text{ and } s \notin Y, \\ (1-k) \cdot \mathbb{Y}(s) & \text{if } s \in Y \text{ and } s \notin X. \end{cases}$$

We then write that $\mathbb{Z} = \mathbb{X} \sqcup \mathbb{Y}$ if $\mathbb{Z} = \mathbb{X} \sqcup_k \mathbb{Y}$, for some k.

Intuition of the disjunction

Question: How do we split distributions?
Answer: We rescale.

Let $\mathbb{X}\colon X \to [0,1]$ and $\mathbb{Y}\colon Y \to [0,1]$ be probabilistic teams and let $k \in [0,1]$.

We denote by $\mathbb{X} \sqcup_k \mathbb{Y}$ the k-scaled union of $\mathbb{X}$ and $\mathbb{Y}$, that is, the probabilistic team $\mathbb{X} \sqcup_k \mathbb{Y}\colon X \cup Y \to [0,1]$ defined s.t. for each $s \in X \cup Y$,

$$(\mathbb{X} \sqcup_k \mathbb{Y})(s) := \begin{cases} k \cdot \mathbb{X}(s) + (1-k) \cdot \mathbb{Y}(s) & \text{if } s \in X \text{ and } s \in Y, \\ k \cdot \mathbb{X}(s) & \text{if } s \in X \text{ and } s \notin Y, \\ (1-k) \cdot \mathbb{Y}(s) & \text{if } s \in Y \text{ and } s \notin X. \end{cases}$$

We then write that $\mathbb{Z} = \mathbb{X} \sqcup \mathbb{Y}$ if $\mathbb{Z} = \mathbb{X} \sqcup_k \mathbb{Y}$, for some k.

Intuition of the disjunction

Question: How do we split distributions?
Answer: We rescale.

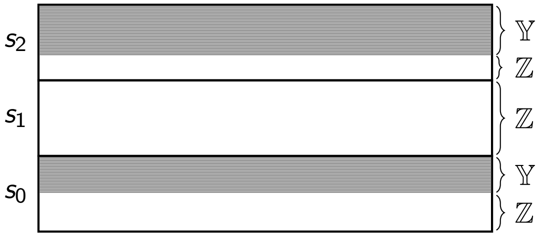
Let $\mathbb{X}\colon X \to [0,1]$ and $\mathbb{Y}\colon Y \to [0,1]$ be probabilistic teams and let $k \in [0,1]$.

We denote by $\mathbb{X} \sqcup_k \mathbb{Y}$ the k-scaled union of $\mathbb{X}$ and $\mathbb{Y}$, that is, the probabilistic team $\mathbb{X} \sqcup_k \mathbb{Y}\colon X \cup Y \to [0,1]$ defined s.t. for each $s \in X \cup Y$,

$$(\mathbb{X} \sqcup_k \mathbb{Y})(s) := \begin{cases} k \cdot \mathbb{X}(s) + (1-k) \cdot \mathbb{Y}(s) & \text{if } s \in X \text{ and } s \in Y, \\ k \cdot \mathbb{X}(s) & \text{if } s \in X \text{ and } s \notin Y, \\ (1-k) \cdot \mathbb{Y}(s) & \text{if } s \in Y \text{ and } s \notin X. \end{cases}$$

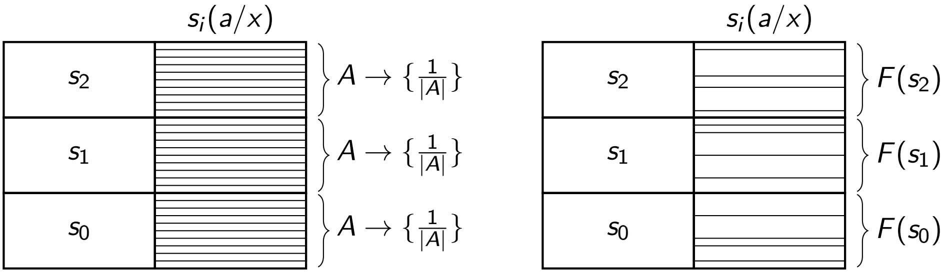
We then write that $\mathbb{Z} = \mathbb{X} \sqcup \mathbb{Y}$ if $\mathbb{Z} = \mathbb{X} \sqcup_k \mathbb{Y}$, for some k.

Intuition of the disjunction



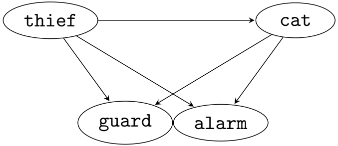
- Partition $\mathbb{X}$ to two probabilistic teams $\mathbb{Y}$ and $\mathbb{Z}$ and re-scale both back to 1.
- **NB.** The empty set is considered as a probabilistic team.

Intuition of the quantifiers



- Universal quantification (i.e., the set $\mathbb{X}[A/x]$) is depicted on left.
- Existential quantification (i.e., the set $\mathbb{X}[F/x]$) is depicted on right.
- Height of a box corresponds to the probability of an assignment.

Example



guard

thief,cat	T	F
TT	0.8	0.2
TF	0.7	0.3
FT	0	1
FF	0	1

thief

T	F
0.1	0.9

cat

thief	T	F
T	0.1	0.9
F	0.6	0.4

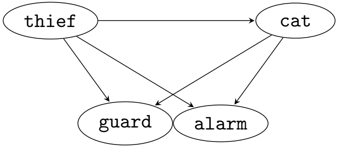
alarm

thief,cat	T	F
TT	0.9	0.1
TF	0.8	0.2
FT	0.1	0.9
FF	0	1

From the Bayesian network above we obtain that the joint probability distribution for t, c, g, a can be factorized as

$$P(t, c, g, a) = P(t) \cdot P(c \mid t) \cdot P(g \mid t, c) \cdot P(a \mid t, c)$$

Example



guard

thief,cat	T	F
TT	0.8	0.2
TF	0.7	0.3
~~FT~~	~~0~~	~~1~~
~~FF~~	~~0~~	~~1~~

thief

T	F
0.1	0.9

cat

thief	T	F
T	0.1	0.9
F	0.6	0.4

alarm

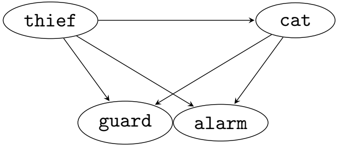
thief,cat	T	F
TT	0.9	0.1
TF	0.8	0.2
FT	0.1	0.9
FF	0	1

If additionally we have

$$t = F \rightarrow g = F$$

(i.e., `guard` never raises alert in absence of `thief`), the two bottom rows of the conditional probability distribution for `guard` become superfluous.

Example



guard

thief,cat	T	F
TT	0.8	0.2
TF	0.7	0.3
FT	0	1
FF	0	1

thief

T	F
0.1	0.9

cat

thief	T	F
T	0.1	0.9
F	0.6	0.4

alarm

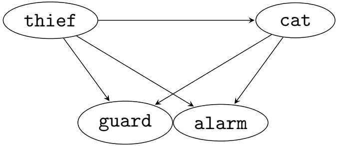
thief,cat	T	F
TT	0.8	0.2
TF	0.7	0.3
FT	0	1
FF	0	1

Given

$$\phi := tca \approx tcg$$

(i.e., `alarm` and `guard` have the same reliability for any given value of `thief` and `cat`), then the conditional distributions for `alarm` and `guard` are equal and one of them can be removed.

Example



guard		
thief, cat	T	F
TT	0.8	0.2
TF	0.7	0.3
FT	0	1
FF	0	1

thief	
T	F
0.1	0.9

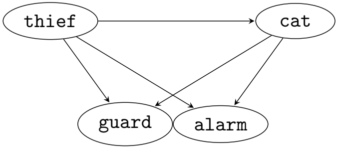
cat		
thief	T	F
T	0.1	0.9
F	0.6	0.4

Given

$$\phi := tca \approx tcg$$

(i.e., alarm and guard have the same reliability for any given value of thief and cat), then the conditional distributions for alarm and guard are equal and one of the them can be removed.

Example



guard

thief,cat	T	F
TT	0.45	0.55
TF	0.4	0.6
FT	0.05	0.95
FF	0	1

thief

T	F
0.1	0.9

cat

thief	T	F
T	0.1	0.9
F	0.6	0.4

alarm

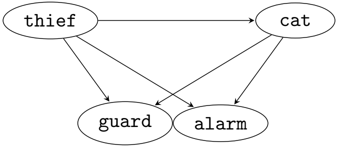
thief,cat	T	F
TT	0.9	0.1
TF	0.8	0.2
FT	0.1	0.9
FF	0	1

Given

$$\phi := \exists x(tcg \approx tcx \wedge tcga \perp\!\!\!\perp y \wedge x = T \leftrightarrow ay = TT)$$

(i.e., `guard` is of a factor $P(y = T)$ less sensitive to raise alert than `alarm` for any given `thief` and `cat`), it suffices to store the conditional probability table for `alarm` and the probability $P(y = T)$.

Example



$P(Y = T) = 0.5$

thief	
T	F
0.1	0.9

cat		
thief	T	F
T	0.1	0.9
F	0.6	0.4

alarm		
thief,cat	T	F
TT	0.9	0.1
TF	0.8	0.2
FT	0.1	0.9
FF	0	1

Given

$$\phi := \exists x(tcg \approx tcx \wedge tcga \perp\!\!\!\perp y \wedge x = T \leftrightarrow ay = TT)$$

(i.e., guard is of a factor $P(y = T)$ less sensitive to raise alert than alarm for any given thief and cat), it suffices to store the conditional probability table for alarm and the probability $P(y = T)$.

More examples

- The formula $\forall \vec{y}\, \vec{x} \approx \vec{y}$ states that the probabilities for $\vec{x}$ are *uniformly distributed* over all value sequences of length $|\vec{x}|$.
- The probability of $P(x)$ is at least twice the probability of $Q(x)$.
- Can we *characterise* the expressive power of $\mathrm{FO}(\approx, \perp\!\!\!\perp)$ in the probabilistic setting?

More examples

- The formula $\forall \vec{y}\, \vec{x} \approx \vec{y}$ states that the probabilities for $\vec{x}$ are *uniformly distributed* over all value sequences of length $|\vec{x}|$.
- The probability of $P(x)$ is at least twice the probability of $Q(x)$.
- Can we *characterise* the expressive power of $\mathrm{FO}(\approx, \perp\!\!\!\perp)$ in the probabilistic setting?

Benchmark logic

- In team semantics context fragments of *second-order logic* are captured.
- $\mathrm{FO}(\perp)$ (team semantics) is as expressive as *existential second-order logic*.
- We define a two-sorted variant of ESO in which we allow
  - quantification of distributions, which constitute the base of numerical terms,
  - sum and multiplication on numerical terms.
- This logic characterises the expressive power of $\mathrm{FO}(\perp\!\!\!\perp)$.

Benchmark logic

- In team semantics context fragments of *second-order logic* are captured.
- $\mathrm{FO}(\perp)$ (team semantics) is as expressive as *existential second-order logic*.
- We define a two-sorted variant of ESO in which we allow
  - quantification of distributions, which constitute the base of numerical terms,
  - sum and multiplication on numerical terms.
- This logic characterises the expressive power of $\mathrm{FO}(\perp\!\!\!\perp)$.

Probabilistic structures

Definition

Let τ and σ be a relational and a functional vocabulary. A probabilistic $\tau \cup \sigma$-structure is a tuple

$$\mathfrak{A} = (A, [0,1], (R_i^{\mathfrak{A}})_{R_i \in \tau}, (f_i^{\mathfrak{A}})_{f_i \in \sigma}),$$

where

- A (i.e. the domain of $\mathfrak{A}$) is a finite nonempty set,
- $[0,1]$ is the unit interval,
- each $R_i^{\mathfrak{A}}$ is a relation on A (i.e., a subset of $A^{\mathrm{ar}(R_i)}$),
- each $f_i^{\mathfrak{A}}$ is a probability distribution from $A^{\mathrm{ar}(f_i)}$ to $[0,1]$ (i.e., a function such that $\sum_{\vec{a} \in A^{\mathrm{ar}(f_i)}} f_i(\vec{a}) = 1$).

Second-order logic for probabilistic structures

- As *first-order terms* we have first-order variables.
- The set of *numerical σ-terms* i is defined via the grammar

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

where $\vec{x}, \vec{y}$ are tuples of first-order variables, $f \in \sigma$ and σ is a set of functions.

- The *value* of a numerical term i in a structure $\mathfrak{A}$ under an assignment s is denoted by $[i]^{\mathfrak{A}}_s$ and defined as follows:

$$[f(\vec{x})]^{\mathfrak{A}}_s := f^{\mathfrak{A}}(s(\vec{x})), \qquad [i \times j]^{\mathfrak{A}}_s := [i]^{\mathfrak{A}}_s \cdot [j]^{\mathfrak{A}}_s,$$

$$[\mathrm{SUM}_{\vec{x}}\, i(\vec{x}, \vec{y})]^{\mathfrak{A}}_s := \sum_{\vec{a} \in A^{|\vec{x}|}} [i(\vec{a}, \vec{y})]^{\mathfrak{A}}_s,$$

where $\cdot$ and $\sum$ are the multiplication and sum of real numbers.

Second-order logic for probabilistic structures

- As *first-order terms* we have first-order variables.
- The set of *numerical σ-terms* i is defined via the grammar

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

  where $\vec{x}, \vec{y}$ are tuples of first-order variables, $f \in \sigma$ and σ is a set of functions.
- The *value* of a numerical term i in a structure $\mathfrak{A}$ under an assignment s is denoted by $[i]^{\mathfrak{A}}_s$ and defined as follows:

$$[f(\vec{x})]^{\mathfrak{A}}_s := f^{\mathfrak{A}}(s(\vec{x})), \qquad [i \times j]^{\mathfrak{A}}_s := [i]^{\mathfrak{A}}_s \cdot [j]^{\mathfrak{A}}_s,$$

$$[\mathrm{SUM}_{\vec{x}}\, i(\vec{x}, \vec{y})]^{\mathfrak{A}}_s := \sum_{\vec{a} \in A^{|\vec{x}|}} [i(\vec{a}, \vec{y})]^{\mathfrak{A}}_s,$$

  where $\cdot$ and $\sum$ are the multiplication and sum of real numbers.

Second-order logic for probabilistic structures

Definition

The formulae of $\mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$ is defined via the following grammar:

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

where i is a numerical term, R is a relation symbol, f is a function variable, $\vec{x}$ is a tuple of first-order variables.

Semantics of $\mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$ is defined via probabilistic structures and assignments analogous to FO. In addition to the clauses of first-order logic, we have:

$$\mathfrak{A} \models_s i = j \Leftrightarrow [i]^{\mathfrak{A}}_s = [j]^{\mathfrak{A}}_s,$$

$$\mathfrak{A} \models_s \exists f\phi \Leftrightarrow \mathfrak{A}[h/f] \models_s \phi \text{ for some probability distribution } h\colon A^{\mathrm{ar}(f)} \to [0,1],$$

where $\mathfrak{A}[h/f]$ denotes the expansion of $\mathfrak{A}$ that interprets f to h.

Second-order logic for probabilistic structures

Definition

The formulae of $\mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$ is defined via the following grammar:

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

where i is a numerical term, R is a relation symbol, f is a function variable, $\vec{x}$ is a tuple of first-order variables.

Semantics of $\mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$ is defined via probabilistic structures and assignments analogous to FO. In addition to the clauses of first-order logic, we have:

$$\mathfrak{A} \models_s i = j \Leftrightarrow [i]^{\mathfrak{A}}_s = [j]^{\mathfrak{A}}_s,$$

$$\mathfrak{A} \models_s \exists f \phi \Leftrightarrow \mathfrak{A}[h/f] \models_s \phi \text{ for some probability distribution } h \colon A^{\mathrm{ar}(f)} \to [0,1],$$

where $\mathfrak{A}[h/f]$ denotes the expansion of $\mathfrak{A}$ that interprets f to h.

Examples

- Uniformity of a distribution f can be expressed with

$$\phi(f) := \forall \vec{x}\vec{y}(f(\vec{x}) = 0 \lor f(\vec{y}) = 0 \lor f(\vec{x}) = f(\vec{y})).$$

- For a numerical term i and rational number $\frac{p}{q}$, the property

$$i(\vec{x}) = \frac{p}{q}$$

  can be expressed in $\mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$.

Correspondence between $\mathrm{FO}(\perp\!\!\!\perp)$ and $\mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$

For a probabilistic team $\mathbb{X}\colon X \to [0,1]$, we let $f_{\mathbb{X}} : A^n \to [0,1]$ be the probability distribution such that $f_{\mathbb{X}}(s(\vec{x})) = \mathbb{X}(s)$ for all $s \in X$.

Theorem

For every $\phi(\vec{x}) \in \mathrm{FO}(\perp\!\!\!\perp)$ there is a formula $\phi^*(f) \in \mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$ with one free function variable f s.t. for all structures $\mathfrak{A}$ and nonempty probabilistic teams $\mathbb{X}$

$$\mathfrak{A} \models_{\mathbb{X}} \phi(\vec{x}) \iff (\mathfrak{A}, f_{\mathbb{X}}) \models \phi^*(f).$$

Theorem

Let $\phi(p) \in \mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$ be a sentence with exactly one free function symbol p. Then there is a formula $\Phi \in \mathrm{FO}(\perp\!\!\!\perp)$ such that for all structures $\mathfrak{A}$ and probabilistic teams $\mathbb{X} := p^{\mathfrak{A}}$,

$$\mathfrak{A} \models_{\mathbb{X}} \Phi \iff (\mathfrak{A}, p) \models \phi.$$

Correspondence between $\mathrm{FO}(\perp\!\!\!\perp)$ and $\mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$

For a probabilistic team $\mathbb{X}\colon X \to [0,1]$, we let $f_{\mathbb{X}} : A^n \to [0,1]$ be the probability distribution such that $f_{\mathbb{X}}(s(\vec{x})) = \mathbb{X}(s)$ for all $s \in X$.

Theorem

For every $\phi(\vec{x}) \in \mathrm{FO}(\perp\!\!\!\perp)$ there is a formula $\phi^*(f) \in \mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$ with one free function variable f s.t. for all structures $\mathfrak{A}$ and nonempty probabilistic teams $\mathbb{X}$

$$\mathfrak{A} \models_{\mathbb{X}} \phi(\vec{x}) \iff (\mathfrak{A}, f_{\mathbb{X}}) \models \phi^*(f).$$

Theorem

Let $\phi(p) \in \mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$ be a sentence with exactly one free function symbol p. Then there is a formula $\Phi \in \mathrm{FO}(\perp\!\!\!\perp)$ such that for all structures $\mathfrak{A}$ and probabilistic teams $\mathbb{X} := p^{\mathfrak{A}}$,

$$\mathfrak{A} \models_{\mathbb{X}} \Phi \iff (\mathfrak{A}, p) \models \phi.$$

Relation to earlier works

Probabilistic structures are closely related to metafinite structures (Grädel, Gurevich '98), such as $\mathbb{R}$-structures (Grädel, Meer '95) that consist of a finite structure $\mathfrak{A}$ together with an ordered fiels of reals and a finite set of weight functions from $\mathfrak{A}$ to $\mathbb{R}$.

$\mathbb{R}$-structures can be analyzed in terms of $\mathrm{ESO}_{\mathbb{R}}(+, \times, <, (c_k)_{k\in\mathbb{R}})$, i.e., a two-sorted variant of ESO with existential quantification over functions from $\mathfrak{A}$ to reals.

Expressivity of $\mathrm{ESO}_{\mathbb{R}}(+, \times, <, (c_k)_{k\in\mathbb{R}})$ can be characterized in terms of Blum–Shub–Smale machines, i.e., a model of computation which treats real numbers as basic entities and performs arithmetic operations on reals in a single step.

Theorem (Grädel, Meer '95)

$\mathrm{ESO}_{\mathbb{R}}(+, \times, <, (c_k)_{k\in\mathbb{R}}) \equiv \mathsf{NP}_{\mathbb{R}}$, *where* $\mathsf{NP}_{\mathbb{R}}$ *is non-deterministic polynomial time over BSS machines.*

Expressivity of $\mathrm{FO}(\approx)$

The union closure property of inclusion logic in team semantics extends to $\mathrm{FO}(\approx)$ in probabilistic team semantics.

Theorem

Let $\mathfrak{A}$ be a model, let $\phi \in \mathrm{FO}(\approx)$, let $k \in [0,1]$, and let $\mathbb{X}$ and $\mathbb{Y}$ be two probabilistic teams. Then

$$\mathfrak{A} \models_{\mathbb{X}} \phi \wedge \mathfrak{A} \models_{\mathbb{Y}} \phi \implies \mathfrak{A} \models_{\mathbb{X} \sqcup_k \mathbb{Y}} \phi.^*$$

Thus $\mathrm{FO}(\approx)$ cannot express that a variable x is constant in $\mathbb{X} = (X, p)$, i.e., $p(x) = 1$.

*Recall $\mathbb{X} \sqcup_k \mathbb{Y}$ denotes the k-scaled union of $\mathbb{X}$ and $\mathbb{Y}$.

Other probabilistic atoms

We consider the expansion of first-order logic with the entropy atoms

$$\vec{x} \approx^* \vec{y}$$

with the semantics

$$\mathfrak{A} \models_{\mathbb{X}} \vec{x} \approx^* \vec{y} \text{ iff } \{\{|\mathbb{X}|_{\vec{x}=\vec{a}} : \vec{a} \in A^{|\vec{x}|}\}\} = \{\{|\mathbb{X}|_{\vec{y}=\vec{b}} : \vec{b} \in A^{|\vec{y}|}\}\},$$

where $\{\{\cdot\}\}$ refer to multisets, e.g., $\{\{0, 0, 1\}\} \neq \{\{0, 1, 1\}\}$.

Other probabilistic atoms

Dependence atoms

$$\mathrm{dep}(\vec{x}, \vec{y})$$

can be also incorporated in probabilistic team semantics:

$\mathfrak{A} \models_{\mathbb{X}} \mathrm{dep}(\vec{x}, y)$ iff for all $s, s' \in X : s(\vec{x}) = s'(\vec{x}) \wedge p(s) > 0, p(s') > 0 \to s(\vec{y}) = s'(\vec{y})$,

where $\mathbb{X} = (X, p)$.

Relationships between probabilistic atoms

(1) $\mathrm{dep}(\cdot)$-atoms can be expressed in $\mathrm{FO}(\approx^*)$ and $\mathrm{FO}(\perp\!\!\!\perp)$,
(2) $\approx^*$-atoms in $\mathrm{FO}(\approx, \mathrm{dep}(\cdot))$,
(3) $\approx$-atoms in $\mathrm{FO}(\approx^*)$ and $\mathrm{FO}(\perp\!\!\!\perp)$:

(1) $\mathrm{dep}(\vec{x}, \vec{y}) \equiv \vec{x} \approx^* \vec{x}\vec{y} \equiv \vec{y} \perp\!\!\!\perp_{\vec{x}} \vec{y}$,

(2) $\vec{x} \approx^* \vec{y} \equiv \exists \vec{z}(\mathrm{dep}(\vec{y}, \vec{z}) \wedge \mathrm{dep}(\vec{z}, \vec{y}) \wedge \vec{y} \approx^* \vec{z})$.

(3) $\vec{x} \approx \vec{y} \equiv \forall \vec{z}\big[(z \neq x \wedge z \neq y) \vee \big((z = x \vee z = y) \wedge \vec{z} \approx^* \vec{x} \wedge \vec{z} \approx^* y\big)\big]$

$$\equiv \exists ww' \forall \vec{z} z' \forall u \exists u' \Big\{ ww' \perp\!\!\!\perp ww' \wedge w \neq w' \wedge \Big[x \neq y \vee \Big(x = y \wedge$$
$$\{(\vec{z} \neq \vec{x} \wedge \vec{z} \neq \vec{y}) \vee [(\vec{z} = \vec{x} \vee \vec{z} = \vec{y}) \wedge \vec{z} \perp\!\!\!\perp z' \wedge$$
$$((u \neq w \wedge u \neq w') \vee \{(u = w \vee u = w') \wedge$$
$$[(z' = w \wedge u = u') \vee (z' = w' \wedge u \neq u')] \wedge u \perp\!\!\!\perp u'\})]\}\Big)\Big]\Big\}$$

Expressiveness of probabilistic logics

We obtain the following classification:

Theorem

- $\mathrm{FO}(\approx) < \mathrm{FO}(\approx, \mathrm{dep}(\cdot)) \equiv \mathrm{FO}(\approx^*) \leq \mathrm{FO}(\perp\!\!\!\perp)$
- $\mathrm{FO}(\perp\!\!\!\perp) \equiv \mathrm{ESO}_{[0,1]}(\mathrm{SUM}, \times)$
- $\mathrm{FO}(\approx^*) \equiv \mathrm{ESO}_{[0,1]}(\mathrm{SUM})$

Complexity of probabilistic team semantics

Consider propositional logic PL with probabilistic team semantics. Denote classical negation by "$\sim$", i.e., $\mathfrak{A} \models_{\mathbb{X}} \sim \phi$ iff $\mathfrak{A} \not\models_{\mathbb{X}} \phi$.

Theorem

Let $\phi(\vec{x}) \in \mathrm{PL}(\sim, \perp\!\!\!\perp, \approx)$. Then there is $\psi \in \mathrm{FO}[\times, +, \leq, 0]$ s.t. ϕ is satisfiable iff $(\mathbb{R}, \times, +, \leq, 0) \models \psi$.

Proof.

Sketch. Let $s_{\vec{x}=\vec{b}}$ be a fresh variable for each assignment $s(\vec{x}) = \vec{b}$ where $\vec{b}$ is a binary sequence of length $\vec{x}$. Then $\psi := \exists s_{\vec{x}=\vec{0}} \ldots s_{\vec{x}=\vec{1}}(\Sigma_s 0 \leq s \wedge \phi^*)$ where the mapping $\phi \mapsto \phi^*$ is given recursively as follows:

- Assume $\phi(\vec{x})$ is $\eta(\vec{x}) \vee \chi(\vec{x})$. Then

$$\phi^*(\vec{s}) := \exists t_{\vec{x}=\vec{0}} r_{\vec{x}=\vec{0}} \ldots t_{\vec{x}=\vec{1}} r_{\vec{x}=\vec{1}} \Big(\bigwedge_{\vec{i}} (0 \leq t_{\vec{x}=\vec{i}} \wedge 0 \leq r_{\vec{x}=\vec{i}} \wedge \\ s_{\vec{x}=\vec{i}} = t_{\vec{x}=\vec{i}} + r_{\vec{x}=\vec{i}}) \wedge \eta^*(\vec{t}) \wedge \chi^*(\vec{r}) \Big).$$

- etc.

Complexity of probabilistic team semantics

Since the (existential) first-order theory of the ordered field of reals is decidable in EXPSPACE (PSPACE), and the previous translation was exponential:

Theorem

Let $\mathcal{C} \subseteq \{\mathrm{dep}(\cdot), \perp, \subseteq\}$. *Satisfiability and validity for*

(1) $\mathrm{PL}(\sim, \mathcal{C})$ *(in team semantics) is* AEXPTIME(*poly*)*-complete (H., Kontinen, Virtema, Vollmer, '16)*

(2) $\mathrm{PL}(\sim, \perp\!\!\!\perp)$ *(in probabilistic team semantics) is in* 2EXPSPACE

(3) $\mathrm{ML}(\sim, \mathcal{C})$ *(in team semantics) is* TOWER(*poly*)*-complete (Lück, '18)*

Consequence of (2): implication of conditional independence over discrete binary distributions is in EXPSPACE*.*

TOWER(*poly*): *computation time* $2^{\underbrace{n^{n^{n^{\cdot^{\cdot^{\cdot^{n}}}}}}}_{poly}}$ *with a polynomial upper bound for the exponent tower height.*

AEXPTIME(*poly*): *alternating exponential time with polynomially many alternations.*

Conclusion

- Probabilistic team semantics extends team semantics by adding a probability measure over assignments.
- This makes possible to introduce logics for probabilistic dependencies such as $\perp\!\!\!\perp$ and $\approx$.
- The logics obtained can be compared to each other and characterized in terms of a two-sorted variant of ESO.
- Open problems:
  - Can we axiomatize $\mathrm{PL}(\perp\!\!\!\perp, \approx)$, or $\perp\!\!\!\perp$-atoms within $\mathrm{FO}(\perp\!\!\!\perp)$?
  - Data complexity of $\mathrm{FO}(\perp\!\!\!\perp), \mathrm{FO}(\approx)$? Can we logically characterize e.g. $\mathsf{P}_{\mathbb{R}}/\mathsf{NP}_{\mathbb{R}}$ classes of probability distributions in probabilistic team semantics?

Thanks!

Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, and Jonni Virtema.
Approximation and dependence via multiteam semantics.
Ann. Math. Artif. Intell., 83(3-4):297–320, 2018.

Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, and Jonni Virtema.
Probabilistic team semantics.
In *FoIKS 2018*, pages 186–206, 2018.

Erich Grädel and Yuri Gurevich.
Metafinite model theory.
Inf. Comput., 140(1):26–81, 1998.

Erich Grädel and Klaus Meer.
Descriptive complexity theory over the real numbers.
In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing, 29 May-1 June 1995, Las Vegas, Nevada, USA*, pages 315–324, 1995.

Miika Hannula, Juha Kontinen, Jonni Virtema, and Heribert Vollmer.
Complexity of propositional logics in team semantic.
ACM Trans. Comput. Log., 19(1):2:1–2:14, 2018.

Martin Lück.
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*, pages 30:1–30:23, 2018.