Complexity of probabilistic inclusion logic and additive real arithmetics

Miika Hannula and Jonni Virtema

Workshop on Logics of Dependence and Independence (LoDE 2020V)
August 12, 2020

Qualitative vs. quantitative dependence

Team logics can reason only about qualitative (relational) dependencies.

What about quantitative (probabilistic) dependencies?

Qualitative:

Functional dependency $X \to Y$

Multivalued dependency $X \twoheadrightarrow Y$

Inclusion dependency $X \subseteq Y$

Quantitative:

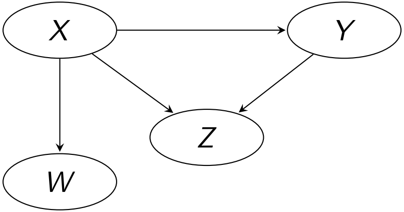
Marginal independence $X \perp\!\!\!\perp Y$

Conditional independence $X \perp\!\!\!\perp Y \mid Z$

Identical distribution of X and Y

Applications of quantitative atoms

- Sample $X_1, \ldots, X_n$ from a population distribution: $X_i \sim X_j$ and $X_i \perp\!\!\!\perp X_j$ for $i \neq j$
- Markov chain $(X_1, X_2, X_3, \ldots)$: $X_1 \ldots X_{i-1} \perp\!\!\!\perp_{X_i} X_{i+1}$ for i
- Bayesian network



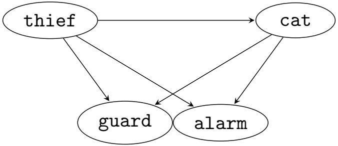
: $\quad Y \perp\!\!\!\perp_X W$ and $Z \perp\!\!\!\perp_{XY} W$

Probabilistic team semantics

Basic concepts:

- Probabilistic team = probability distribution on a finite team (FoIKS 2018)
- Quantitative atoms (e.g., conditional independence $\perp\!\!\!\perp_{\mathrm{c}}$, identical distribution $\approx$)
- $\{\forall, \exists, \wedge, \vee\}$ for complex probability statements

Example 2



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

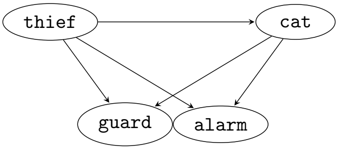
alarm

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 2



thief	
T	F
0.1	0.9

cat

thief	T	F
T	0.1	0.9
F	0.6	0.4

guard

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

alarm

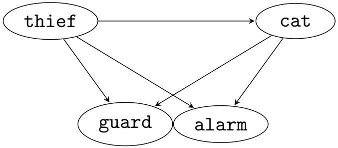
thief,cat	T	F
TT	0.8	0.2
TF	0.7	0.3
FT	0	1
FF	0	1

Given

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

(i.e., conditioned on `thief` and `cat`, `alarm` and `guard` are identically distributed), then the conditional probability tables for `alarm` and `guard` are identical and one of them can be removed.

Example 2



	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

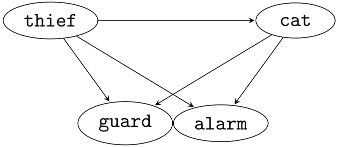
	alarm	
thief,cat	T	F
TT	0.9	0.1
TF	0.8	0.2
FT	0.1	0.9
FF	0	1

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 2



$P(Y = T) = 0.5$

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 $P(y = T)$.

Probabilistic inclusion logic $\mathrm{FO}(\approx)$

Syntax: FO (negation normal form) $+$ $\vec{x} \approx \vec{y}$ (only positively)

Semantics: Defined in terms of a finite structure $\mathfrak{A}$ and a probabilistic team $\mathbb{X}$

(1) Team $=$ a set of variable assignments with a shared domain

(2) Probabilistic team $=$ a pair $\mathbb{X} = (X, p)$, where X is a finite team and $p : X \to [0, 1]$ a probability distribution

In general: $\mathrm{FO}(A_1, \ldots, A_n)$ is FO (negation normal form) $+$ $A_1, \ldots, A_n$ (only positively)

Semantics of (probabilistic) dependencies

Let $\mathbb{X} = (X, p)$ be a probabilistic team and $\vec{x}, \vec{a}$ be tuples of variables and values.

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

The semantics of marginal identity atoms (identical distribution) $\vec{x} \approx \vec{y}$:

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

Semantics of (probabilistic) dependencies

Let $\mathbb{X} = (X, p)$ be a probabilistic team and $\vec{x}, \vec{a}$ be tuples of variables and values.

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

The semantics of probabilistic conditional independence atoms $\vec{y} \perp\!\!\!\perp_{\vec{x}} \vec{z}$:

$\mathfrak{A} \models_{\mathbb{X}} \vec{y} \perp\!\!\!\perp_{\vec{x}} \vec{z}$ 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 (probabilistic) dependencies

Let $\mathbb{X} = (X, p)$ be a probabilistic team and $\vec{x}, \vec{a}$ be tuples of variables and values.

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

The semantics of team-based dependencies α:

$$\mathfrak{A} \models_{\mathbb{X}} \alpha \text{ iff } \mathfrak{A} \models_{X^+} \alpha,$$

where X^+ consists of $s \in X$ such that $\mathbb{X}(s) > 0$

Semantics of first-order part I

Definition (FoIKS 2018)

Let $\mathfrak{A}$ be a finite structure and $\mathbb{X} = (X, p)$ a probabilistic team.

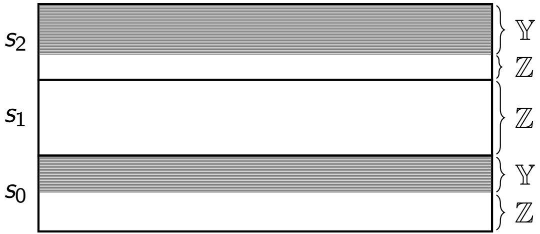
$$
\begin{aligned}
\mathfrak{A} \models_{\mathbb{X}} \ell \quad &\Leftrightarrow \quad \mathfrak{A} \models_{s} \ell \text{ for all } s \in X \text{ such that } p(s) > 0 \\
& \quad \text{(when } \ell \text{ is a first-order literal)} \\
\mathfrak{A} \models_{\mathbb{X}} (\psi \wedge \theta) \quad &\Leftrightarrow \quad \mathfrak{A} \models_{\mathbb{X}} \psi \text{ and } \mathfrak{A} \models_{\mathbb{X}} \theta
\end{aligned}
$$

Semantics of first-order part II

Disjunction via convex combinations:

$$\mathfrak{A} \models_{\mathbb{X}} (\psi \vee \theta) \quad \Leftrightarrow \quad \mathfrak{A} \models_{\mathbb{Y}} \psi \text{ and } \mathfrak{A} \models_{\mathbb{Z}} \theta,$$

where $\mathbb{X} = \alpha \cdot \mathbb{Y} + (1 - \alpha) \cdot \mathbb{Z}$, for some $\alpha \in [0, 1]$.



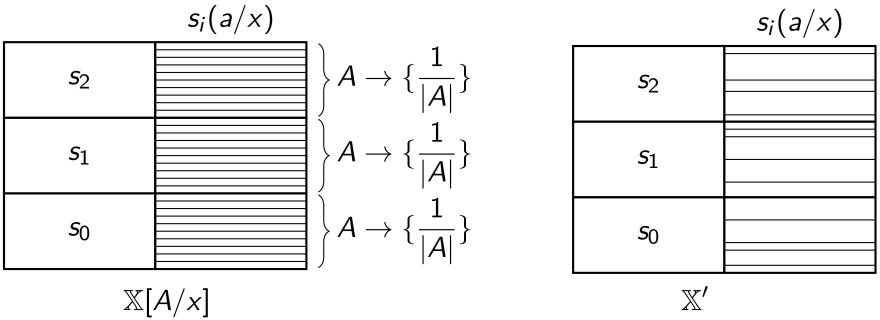
NB. The empty set is considered as a probabilistic team.

Semantics of first-order part III

Quantification introduces a new column:

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

$$\mathfrak{A} \models_{\mathbb{X}} \exists x \psi \quad \Leftrightarrow \quad \mathfrak{A} \models_{\mathbb{X}'} \psi, \text{ for some } \mathbb{X}' \text{ such that } \mathbb{X}' \upharpoonright \mathrm{Dom}(\mathbb{X}) = \mathbb{X}$$



Formulae vs. sentences

This paper: logical, computational, axiomatic properties of $\mathrm{FO}(\approx)$ and $\mathrm{FO}(\approx, \mathrm{dep}(\cdots))$

Two levels of analysis:

- Sentences $\sim$ finite structures $\sim$ strings of Booleans
- Formulae $\sim$ probabilistic teams $\sim$ strings of reals

Formulae vs. sentences

This paper: logical, computational, axiomatic properties of $\mathrm{FO}(\approx)$ and $\mathrm{FO}(\approx, \mathrm{dep}(\cdots))$

Two levels of analysis:

- **Sentences $\sim$ finite structures $\sim$ strings of Booleans**
- Formulae $\sim$ probabilistic teams $\sim$ strings of reals

Expressivity over sentences – background

$$\mathrm{FO}(\subseteq) \overset{\mathrm{ord}}{=} \mathsf{P} \subseteq \mathsf{NP} \subseteq \mathsf{PSPACE}$$

$$\mathrm{FO}(\subseteq, \mathrm{dep}(\cdots)) = \mathsf{NP} = \mathrm{FO}(\perp_{\mathrm{c}})$$

Table: Team semantics

$$\mathsf{P} \subseteq \mathsf{NP} \subseteq \exists[0,1]^{\leq} \subseteq \exists\mathbb{R} \subseteq \mathsf{PSPACE}$$

$$\mathrm{FO}(\perp\!\!\!\perp_{\mathrm{c}}) = \exists[0,1]^{\leq}$$

Table: Probabilistic team semantics

Probabilistic independence and existential theory of the reals

- The existential theory of the reals consists of all true sentences of the form

$$\exists x_1 \dots \exists x_n \psi(x_1, \dots x_n)$$

  where ψ is a quantifier-free formula of the real arithmetic

- Gives rise to the Boolean complexity class $\exists\mathbb{R}$:
  the closure of the existential theory of the reals under polynomial-time reductions
- $\exists[0,1]^{\leq}$ defined as $\exists\mathbb{R}$ but in terms of sentences of the form

$$\exists x_1 \dots \exists x_n \Big(\bigwedge_{1 \leq i \leq n} 0 \leq x_i \wedge x_i \leq 1 \wedge \psi \Big),$$

  where ψ does not contain $\neg$ nor $<$.

Theorem (LiCS 2020)

Over finite structures, $\mathrm{FO}(\perp\!\!\!\perp_{\mathrm{c}}) \equiv \exists[0,1]^{\leq}$.

Probabilistic independence and existential theory of the reals

- The existential theory of the reals consists of all true sentences of the form

$$\exists x_1 \dots \exists x_n \psi(x_1, \dots x_n)$$

  where ψ is a quantifier-free formula of the real arithmetic

- Gives rise to the Boolean complexity class $\exists\mathbb{R}$:
  the closure of the existential theory of the reals under polynomial-time reductions
- $\exists[0,1]^{\leq}$ defined as $\exists\mathbb{R}$ but in terms of sentences of the form

$$\exists x_1 \dots \exists x_n \Big(\bigwedge_{1 \leq i \leq n} 0 \leq x_i \wedge x_i \leq 1 \wedge \psi \Big),$$

  where ψ does not contain $\neg$ nor $<$.

Theorem (LiCS 2020)

Over finite structures, $\mathrm{FO}(\perp\!\!\!\perp_{\mathrm{c}}) \equiv \exists[0,1]^{\leq}$.

Expressivity over sentences – new results

$$\begin{array}{ccccccc} \mathrm{FO}(\subseteq) & & \mathrm{FO}(\subseteq, \mathrm{dep}(\cdots)) & & \\ \|_{\mathrm{ord}} & & \| & & \\ \mathsf{P} & \subseteq & \mathsf{NP} & \subseteq & \mathsf{PSPACE} \\ & & \| & & \\ & & \mathrm{FO}(\perp_{\mathrm{c}}) & & \end{array}$$

Table: Team semantics

$$\begin{array}{ccccccccc} & & & & \mathrm{FO}(\perp\!\!\!\perp_{\mathrm{c}}) & & & & \\ & & & & \| & & & & \\ \mathsf{P} & \subseteq & \mathsf{NP} & \subseteq & \exists[0,1]^{\leq} & \subseteq & \exists\mathbb{R} & \subseteq & \mathsf{PSPACE} \end{array}$$

Table: Probabilistic team semantics.

Expressivity over sentences – new results

$$
\begin{array}{ccccccc}
\mathrm{FO}(\subseteq) & & \mathrm{FO}(\subseteq, \mathrm{dep}(\cdots)) & & \\
\rotatebox{90}{$=$}_{\mathrm{ord}} & & \rotatebox{90}{$=$} & & \\
\mathsf{P} & \subseteq & \mathsf{NP} & \subseteq & \mathsf{PSPACE} \\
 & & \rotatebox{90}{$=$} & & \\
 & & \mathrm{FO}(\perp_{\mathrm{c}}) & &
\end{array}
$$

Table: Team semantics

$$
\begin{array}{ccccccccccc}
\mathrm{FO}(\approx) & & \mathrm{FO}(\approx, \mathrm{dep}(\cdots)) & & \mathrm{FO}(\perp\!\!\!\perp_{\mathrm{c}}) & & & & \\
\rotatebox{90}{$=$}_{\mathrm{ord}}^{*} & & \rotatebox{90}{$=$}_{*} & & \rotatebox{90}{$=$} & & & & \\
\mathsf{P} & \subseteq & \mathsf{NP} & \subseteq & \exists[0,1]^{\leq} & \subseteq & \exists\mathbb{R} & \subseteq & \mathsf{PSPACE}
\end{array}
$$

Table: Probabilistic team semantics. * New results

Probabilistic inclusion logic over sentences

Lemma

Let $\phi \in \mathrm{FO}(\approx)$ be a sentence. There is a polynomial-time reduction from finite structures $\mathfrak{A}$ to systems of linear inequations $\mathcal{S}$ such that $\mathfrak{A} \models \phi$ if and only if $\mathcal{S}$ has a solution.

Proof.

Sketch. Add a variable $x_{s,\psi}$, for any partial assignment s and any subformula ψ of ϕ. Initialize $\mathcal{S}$ with $x_{\emptyset,\phi} = 1$, and $x_{\psi,s} \geq 0$ for all s and ψ. For each ψ add a set of equations to describe its corresponding team operation. E.g., for disjunction weights of assignments are split to two:

- If ψ is $\theta \vee \theta'$, add $x_{s,\theta} + x_{s,\theta'} = x_{s,\psi}$ for all s.

□

Deciding whether a system of linear inequalities has solutions is in polynomial time

Theorem

Let $\phi \in \mathrm{FO}(\approx)$ be a sentence. The problem of determining whether $\mathfrak{A} \models \phi$ for a given finite structure $\mathfrak{A}$ is in P.

From inclusion to probabilistic inclusion logic

Theorem

Every sentence of $\mathrm{FO}(\subseteq)$ *is equivalent to a sentence of* $\mathrm{FO}(\approx)$.

Proof.

Inclusion atoms definable in terms of *equiextension* atoms $\vec{x}_1 \bowtie \vec{x}_2 := \vec{x}_1 \subseteq \vec{x}_2 \wedge \vec{x}_2 \subseteq \vec{x}_1$ [Galliani, 2012]. However, $\vec{x}_1 \approx \vec{x}_2 \not\equiv \vec{x}_1 \bowtie \vec{x}_2$ as equiextension may hold even if the weights are not in balance.

Proof idea. First balance all positive weights, then apply $\approx$:

$$\forall c \forall \vec{u} \exists v_1 v_2 \forall z'_1 \ldots \forall z'_k \exists z_1 \ldots \exists z_k \big(\bigwedge_{i=1,2} \vec{x}_i = \vec{u} \leftrightarrow v_i = c \wedge \tag{1}$$

$$\bigwedge_{i=1}^{k} z'_i = c \rightarrow z_i = c \wedge (\neg \vec{z} = \vec{c} \vee \vec{u} v_1 \approx \vec{u} v_2)\big),$$

where k is the number of "splits" (from quantification, disjunction) in the underlying sentence.

Probabilistic inclusion logic over sentences cont.

Theorem

$\mathrm{FO}(\approx)$ *corresponds to* P *over finite ordered structures.*

Proof.

1. Over finite structures: $\mathrm{FO}(\subseteq) \subseteq \mathrm{FO}(\approx) \subseteq \mathrm{P}$
2. Over finite ordered structures: $\mathrm{P} \equiv \mathrm{FO}(\subseteq)$ [Galliani and Hella, 2013]

□

Future work:

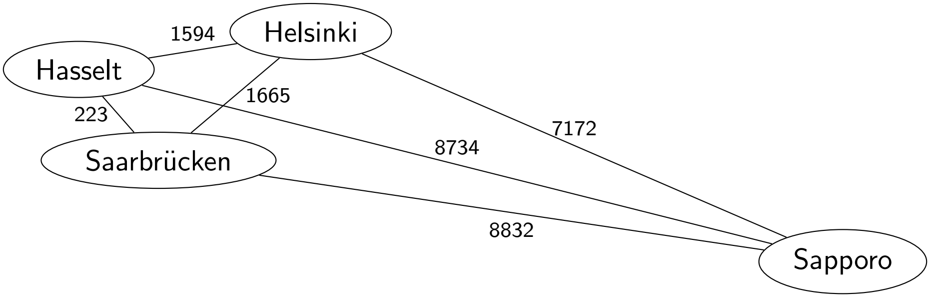
1. $\mathrm{FO}(\subseteq)$ strictly subsumed by $\mathrm{FO}(\approx)$ (over arbitrary finite structures)?
2. Relationship between $\mathrm{FO}(\approx)$ and fixed-point logic/inclusion logic with counting? Cf. [Grädel and Hegselmann, 2016]

Probabilistic inclusion/dependence logic over sentences

Results via $\mathbb{R}$-structures and Blum-Shub-Smale machines

$\mathbb{R}$-structures [Grädel and Meer, 1995] consist of a finite structure $\mathfrak{A}$ together with an ordered field of reals and a finite set of weight functions from $\mathfrak{A}$ to $\mathbb{R}$

(particular case of metafinite structures [Grädel and Gurevich, 1998])

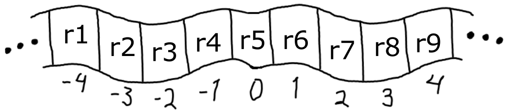


Blum-Shub-Smale machines

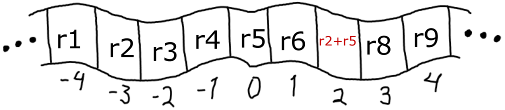
Input: finite string of reals
Output: 0 or 1 (decision problems)

A program is a finite list of instructions:

- Arithmetic instructions:
  $x_i \leftarrow (x_j + x_k)$, $x_i \leftarrow (x_j - x_k)$,
  $x_i \leftarrow (x_j \times x_k)$, $x_i \leftarrow c$.
- Shift left or right.
- Branch on inequality
  if $x_0 \leq 0$ **then** go to α; **else** go to β.



Addition: [2] := [-3]+[0]



Descriptive complexity over the reals

Theorem ([Grädel and Meer, 1995])

$\mathrm{ESO}_{\mathbb{R}}[+, \times, \leq, (r)_{r \in \mathbb{R}}] \equiv \mathsf{NP}_{\mathbb{R}}$

Two-sorted variant of ESO with

1. first-order logic on the finite structure $\mathfrak{A}$
2. existential quantification of functions from $\mathfrak{A}$ to reals
3. constants r for each real
4. complex numerical terms by $\{+, \times\}$
5. (negated) inequality $\leq$ between numerical terms

Too strong for $\mathrm{FO}(\approx, \mathrm{dep}(\cdots))$**:** 1) Lacks $\neg$, $\times$, and real constants 2) Quantification over $[0, 1]$

Descriptive complexity over the reals

Theorem ([Grädel and Meer, 1995])

$\mathrm{ESO}_{\mathbb{R}}[+, \times, \leq, (r)_{r \in \mathbb{R}}] \equiv \mathsf{NP}_{\mathbb{R}}$

Two-sorted variant of ESO with

1. first-order logic on the finite structure $\mathfrak{A}$
2. existential quantification of functions from $\mathfrak{A}$ to reals
3. constants r for each real
4. complex numerical terms by $\{+, \times\}$
5. (negated) inequality $\leq$ between numerical terms

Too strong for $\mathrm{FO}(\approx, \mathrm{dep}(\cdots))$: 1) Lacks $\neg$, $\times$, and real constants 2) Quantification over $[0, 1]$

Probabilistic inclusion/dependence vs. additive ESO over the reals

We show (adapting techniques from [FoIKS 2018]):

Theorem

$\mathrm{FO}(\approx, \mathrm{dep}(\cdots)) \equiv \mathrm{L}\text{-}\mathrm{ESO}_{[0,1]}[+, \leq, 0, 1]$

- "Loose fragment": no negated atoms $\neg i \leq j$ between two numerical terms
- Existential second-order quantification over functions from $\mathrm{Dom}(\mathfrak{A})$ to $[0, 1]$
- Only constants $0, 1$ allowed

NB. The result holds for formulae of $\mathrm{FO}(\approx, \mathrm{dep}(\cdots))$

Probabilistic inclusion/dependence logic over sentences cont.

Theorem

Over finite structures, $\mathrm{FO}(\approx, \mathrm{dep}(\cdots)) \equiv \mathsf{NP}$.

Proof.

$\boxed{\supseteq}$ Over finite structures: $\mathsf{NP} \subseteq \mathrm{FO}(\mathrm{dep}(\cdots)) \subseteq \mathrm{FO}(\approx, \mathrm{dep}(\cdots))$

$\boxed{\subseteq}$ It is easy to show that over formulae:

$$\mathrm{FO}(\approx, \mathrm{dep}(\cdots)) \subseteq \mathrm{ESO}_{\mathbb{R}}[\leq, +, 0, 1] \subseteq \mathsf{NP}^0_{\mathrm{add}}.$$

$\mathsf{NP}^0_{\mathrm{add}}$ allows guessing a string of reals and then verifying in polynomial time in the **additive** Blum-Shub-Smale model of computation (with machine constants $0, 1$).

It suffices to show that $\mathsf{NP}^0_{\mathrm{add}}$ collapses to NP over Boolean inputs. □

Collapse of additive NP over the reals

Theorem

Over Boolean inputs, $\mathrm{NP}^0_{\mathrm{add}} = \mathrm{NP}$

Proof.

Sketch. $\boxed{\supseteq}$ trivial. $\boxed{\subseteq}$ Suppose $L \subseteq \{0,1\}^* \cap \mathrm{NP}^0_{\mathrm{add}}$ is decided non-deterministically by a BSS machine M whose running is bounded by some polynomial p. Let $x \in \{0,1\}^n$ be an input. First, guess the outcome of each comparison of the BSS computation; the outcome is a Boolean string z of length $p(n)$. During a computation the value of each coordinate x_i is a linear function on the constants 0 and 1, the input x, and the real guess y of length $p(n)$. Thus it is possible to construct in polynomial time a system:

$$\sum_{j=1}^{p(n)} a_{ij} y_j \leq 0 \quad (1 \leq i \leq m), \quad \sum_{j=1}^{p(n)} b_{ij} y_j < 0 \quad (1 \leq i \leq l), \quad a_{ij}, b_{ij} \in \mathbb{Z} \qquad (2)$$

such that y is a (real-valued) solution iff M accepts (x, y) wrt. z. □

Formulae vs. sentences

This paper: logical, computational, axiomatic properties of $\mathrm{FO}(\approx)$ and $\mathrm{FO}(\approx, \mathrm{dep}(\cdots))$

Two levels of analysis:

- **Sentences $\sim$ finite structures $\sim$ strings of Booleans**
- Formulae $\sim$ probabilistic teams $\sim$ strings of reals

Formulae vs. sentences

This paper: logical, computational, axiomatic properties of $\mathrm{FO}(\approx)$ and $\mathrm{FO}(\approx, \mathrm{dep}(\cdots))$

Two levels of analysis:

- Sentences $\sim$ finite structures $\sim$ strings of Booleans
- **Formulae $\sim$ probabilistic teams $\sim$ strings of reals**

Expressivity over formulae

$$\mathrm{FO}(\subseteq) \quad \subsetneq \quad \mathrm{FO}(\subseteq, \mathrm{dep}(\cdots)) \equiv \mathrm{FO}(\perp_{\mathrm{c}})$$

Table: Team semantics

$$\mathrm{FO}(\approx) \quad \subsetneq \quad \mathrm{FO}(\approx, \mathrm{dep}(\cdots)) \quad \subsetneq^{*} \quad \mathrm{FO}(\perp\!\!\!\perp_{\mathrm{c}})$$

Table: Probabilistic team semantics. * New results

Probabilistic inclusion/dependence vs. independence

Both $\approx$ and $\mathrm{dep}(\cdots)$ expressible in $\mathrm{FO}(\perp\!\!\!\perp_{\mathrm{c}})$ (JELIA 2019)

Example

Define $\phi(x) = \exists c \exists y \forall z \theta$ where θ is defined as

$$\mathrm{dep}(c) \wedge x \perp\!\!\!\perp y \wedge x \approx y \wedge ((x = c \wedge y = c) \leftrightarrow z = c). \tag{3}$$

Suppose $\{0,1\} \models_{\mathbb{X}} \phi$. Then

1. $c \in \{0,1\}$ is a constant;
2. z is uniformly distributed, so $z = c$ holds for weights $1/2$;
3. $x = c \wedge y = c$ holds for weight $1/2$;
4. x and y are independent and identically distributed, so $x = c$ holds for weight $1/\sqrt{2}$.

NB. Irrational weights not definable in $\mathrm{FO}(\approx, \mathrm{dep}(\cdots))$.

Theorem

Over formulae, $\mathrm{FO}(\approx, \mathrm{dep}(\cdots)) \subsetneq \mathrm{FO}(\perp\!\!\!\perp_{\mathrm{c}})$

Axioms for quantitative dependence

- Marginal independence ✓ [Geiger et al., 1991]
- Conditional independence ✗ [Studený, 1992]
- Marginal identity ?

Axioms for quantitative dependence

- Marginal independence ✓ [Geiger et al., 1991]
- Conditional independence ✗ [Studený, 1992]
- Marginal identity ✓

Theorem

The following axiomatization is sound and complete:

1. *reflexivity:* $x_1 \ldots x_n \approx x_1 \ldots x_n$;
2. *symmetry: if* $x_1 \ldots x_n \approx y_1 \ldots y_n$, *then* $y_1 \ldots y_n \approx x_1 \ldots x_n$;
3. *projection and permutation: if* $x_1 \ldots x_n \approx y_1 \ldots y_n$, *then* $x_{i_1} \ldots x_{i_k} \approx y_{i_1} \ldots y_{i_k}$, *where* $i_1, \ldots, i_k$ *is a sequence of distinct integers from* $\{1, \ldots, n\}$.
4. *transitivity: if* $x_1 \ldots x_n \approx y_1 \ldots y_n$ *and* $y_1 \ldots y_n \approx z_1 \ldots z_n$, *then* $x_1 \ldots x_n \approx z_1 \ldots z_n$.

Conclusion

- We studied quantitative variants of inclusion and inclusion/dependence logics
- Qualitative and quantitative variants in many ways **analogous**:
  1. Inclusion logic captures P (over ordered models)
  2. Inclusion/dependence logic captures NP
  3. Marginal identity and independence has axioms, conditional independence has not
- Where the **analogy breaks down**:
  1. Relationship between inclusion/dependence logic, independence logic, and NP
     - Qualitative: Both logics capture NP properties of teams
     - Quantitative: additive vs. multiplicative properties of prob.teams. For sentences, former captures NP, latter $\exists[0,1]^{\leq}$

Thanks!

Main sources:

- Durand, A., Hannula, M., Kontinen, J., Meier, A., and Virtema, J. Probabilistic team semantics. (2018). In *Foundations of Information and Knowledge Systems - 10th International Symposium, FoIKS 2018, Budapest, Hungary, May 14-18, 2018, Proceedings*, pages 186–206.
- Hannula, M., Hirvonen, Å., Kontinen, J., Kulikov, V., and Virtema, J. (2019). Facets of distribution identities in probabilistic team semantics. In *JELIA*, volume 11468 of *Lecture Notes in Computer Science*, pages 304–320. Springer.
- Hannula, M., Kontinen, J., Van den Bussche, J., and Virtema, J. Descriptive complexity of real computation and probabilistic independence logic. (2020). In *LICS '20: 35th Annual ACM/IEEE Symposium on Logic in Computer Science, Saarbrücken, Germany, July 8-11, 2020*, pages 550–563.

Barthe, G., Hsu, J., and Liao, K. (2020).
A probabilistic separation logic.
Proc. ACM Program. Lang., 4(POPL):55:1–55:30.

Batz, K., Kaminski, B. L., Katoen, J., Matheja, C., and Noll, T. (2019).
Quantitative separation logic: a logic for reasoning about probabilistic pointer programs.
Proc. ACM Program. Lang., 3(POPL):34:1–34:29.

Galliani, P. (2012).
Inclusion and exclusion dependencies in team semantics: On some logics of imperfect information.
Annals of Pure and Applied Logic, 163(1):68 – 84.

Galliani, P. and Hella, L. (2013).
Inclusion Logic and Fixed Point Logic.
In Rocca, S. R. D., editor, *Computer Science Logic 2013 (CSL 2013)*, volume 23 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 281–295, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Geiger, D., Paz, A., and Pearl, J. (1991).
Axioms and algorithms for inferences involving probabilistic independence.
Information and Computation, 91(1):128–141.

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

Grädel, E. and Hegselmann, S. (2016).
Counting in team semantics.
In *25th EACSL Annual Conference on Computer Science Logic, CSL 2016, August 29 - September 1, 2016, Marseille, France*, pages 35:1–35:18.

Grädel, E. and Meer, K. (1995).
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.

Khachiyan, L. G. (1979).
A polynomial algorithm in linear programming.
Dokl. Akad. Nauk SSSR, 244:1093–1096.

Studený, M. (1992).
Conditional independence relations have no finite complete characterization.
pages 377–396. Kluwer.