The Undecidability of Probabilistic Conditional Independence Implication

Cheuk Ting Li
Dept. of Information Engineering, Chinese University of Hong Kong
Email: ctli@ie.cuhk.edu.hk

Dagstuhl Seminar 24111

Random Variables

- A **random variable (RV)** $X : \Omega \to \mathcal{X}$ is a measurable function from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to a measurable space
- We focus on **discrete** random variables, i.e., the support $\mathcal{X}$ is finite or countable
  - Suffices to consider $\Omega = [0, 1]$ to be the standard probability space, i.e., $[0, 1]$ with the Lebesgue measure as the probability
- X, Y are (unconditionally) independent, denoted as $X \perp\!\!\!\perp Y$, if for all x, y,

$$\mathbb{P}((X, Y) = (x, y)) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

- X, Y are conditionally independent given Z, denoted as $X \perp\!\!\!\perp Y|Z$, if for all x, y, z,

$$\mathbb{P}((X,Y,Z) = (x,y,z))\mathbb{P}(Z = z) = \mathbb{P}((X,Z) = (x,z))\mathbb{P}((Y,Z) = (y,z))$$

- WLOG assume all random variables are positive-integer-valued, i.e., measurable functions $X : [0, 1] \to \mathbb{N}$

First-order Theory of Random Variables

- Consider first-order formulae (with logical symbols $\forall$, $\exists$, $\wedge$, $\vee$, $\neg$), with non-logical symbols $\cdot \perp\!\!\!\perp \cdot$ (unconditional independence) and $\cdot \perp\!\!\!\perp \cdot | \cdot$ (conditional independence)
- Variables in the formulae are random variables, i.e., measurable functions $[0,1] \to \mathbb{N}$
  - Just the ordinary first-order logic over the domain of measurable functions $[0,1] \to \mathbb{N}$, with the usual semantics
- Relation with probabilistic team semantics [Durand et al., 2018, Hannula et al., 2023]:
  - A probabilistic team $\mathbb{X}$ can be regarded as a joint distribution of the variables
  - Conditional independence $\mathfrak{A} \models_{\mathbb{X}} x \perp\!\!\!\perp_z y$ means $X \perp\!\!\!\perp Y | Z$ as RVs
  - Different semantics for $\vee$ and $\forall$

Undecidable Problems

- Undecidable problems are decision problems that cannot be solved by any algorithm
  - E.g., Halting problem [Turing, 1936], Diophantine equations [Matiyasevich, 1993], Wang tiles [Berger, 1966], word problem of groups [Novikov, 1955]
- We discuss the undecidability of:
  - Conditional independence implication problem
  - First-order theory of random variables with probabilistic independence relation
  - Conditional information inequalities
  - Network coding

Probabilistic Independence Implication Problem

- Determine whether a probabilistic independence relation among several random variables follows from a list of other such relations [Geiger et al., 1991, Matúš, 1994]
- E.g. $X \perp\!\!\!\perp Y \wedge XY \perp\!\!\!\perp Z \Rightarrow X \perp\!\!\!\perp YZ$
  - i.e., $\forall X, Y, Z. ((X \perp\!\!\!\perp Y \wedge XY \perp\!\!\!\perp Z) \rightarrow X \perp\!\!\!\perp YZ)$
  - In the language of probabilistic team semantics: $\mathfrak{A} \models_{\mathbb{X}} (x \perp\!\!\!\perp y \wedge xy \perp\!\!\!\perp z) \Rightarrow \mathfrak{A} \models_{\mathbb{X}} x \perp\!\!\!\perp yz$
- Geiger et al. [1991] gave a complete set of axioms:
  - (Triviality) $X \perp\!\!\!\perp \emptyset$
  - (Symmetry) $X \perp\!\!\!\perp Y \Rightarrow Y \perp\!\!\!\perp X$
  - (Decomposition) $X \perp\!\!\!\perp YZ \Rightarrow X \perp\!\!\!\perp Y$
  - (Mixing) $X \perp\!\!\!\perp Y \wedge XY \perp\!\!\!\perp Z \Rightarrow X \perp\!\!\!\perp YZ$
- Complete – all true probabilistic independence implications can be deduced from these axioms
- Hence probabilistic independence implication is **decidable**

Conditional Independence Implication Problem

- Determine whether a conditional independence relation among several random variables follows from a list of other such relations [Dawid, 1979, Spohn, 1980, Mouchart and Rolin, 1984]
- E.g. $X \perp\!\!\!\perp Y|Z \wedge X \perp\!\!\!\perp W|YZ \Rightarrow X \perp\!\!\!\perp W|Z$
- **Decidable** if all random variables have bounded cardinalities [Geiger and Meek, 1999, Niepert, 2012]
  - Follows from the decidability of the real polynomial equations
  - Hannula et al. [2019] – in EXPSPACE if all RVs are binary
- What about the case where the cardinalities of the random variables are not bounded?

Semi-graphoid Axioms

- Pearl and Paz [1987] proposed the following 4 axioms:
  - (Symmetry) $X \perp\!\!\!\perp Y|Z \;\Rightarrow\; Y \perp\!\!\!\perp X|Z$
  - (Decomposition) $X \perp\!\!\!\perp YW|Z \;\Rightarrow\; X \perp\!\!\!\perp Y|Z$
  - (Weak union) $X \perp\!\!\!\perp YW|Z \;\Rightarrow\; X \perp\!\!\!\perp Y|ZW$
  - (Contraction) $X \perp\!\!\!\perp Y|Z \wedge X \perp\!\!\!\perp W|YZ \;\Rightarrow\; X \perp\!\!\!\perp YW|Z$
- CI implication would be decidable if semi-graphoid axioms are complete (i.e., all true CI implications can be deduced from these axioms)
  - Simply apply the axioms repeatedly on every combination of random variables until we obtain the desired CI statement
- For the special case where every CI statement involves all random variables (saturated CI), semi-graphoid axioms are complete, and hence **decidable** [Malvestuto, 1992, Geiger and Pearl, 1993]
- Unfortunately, semi-graphoid axioms are incomplete [Studený, 1989]
- Is conditional independence implication decidable in general?

Undecidability of Conditional Independence Implication

- Studený [1989]: Semi-graphoid axioms [Pearl and Paz, 1987] are incomplete
  - Is it possible to add more axioms to make it complete?
- Studený [1992]: No, conditional independence has no finite axiomization
  - Does not rule out other kinds of algorithms
- Herrmann [1995]: Embedded multivalued database dependency is undecidable
- Li [2021]: CI implication is undecidable if one of the RVs is binary
- Li [2022a]: First-order theory of random variables with probabilistic independence relation is undecidable
  - Allow any combination of $\perp\!\!\!\perp, \forall, \exists, \wedge, \vee, \neg$, not only implication
- Li [2022b]: CI implication is **undecidable**
  - Uses the ideas of Herrmann [1995]
- Kühne and Yashfe [2022]: Another concurrent proof of undecidability via matroid theory

First-order Theory of Probabilistic Independence

- Consider first-order formulae with only one non-logical symbol $\perp\!\!\!\perp$ (probabilistic independence)
  - Variables are random variables $(X, Y, \ldots)$
- How to define condition that X is constant, written as $X \overset{\iota}{=} \emptyset$?
  - $X \overset{\iota}{=} \emptyset \;\Leftrightarrow\; X \perp\!\!\!\perp X$
- How to define relation that X is a function of Y, written as $X \overset{\iota}{\leq} Y$?
  - $X \overset{\iota}{\leq} Y \;\Leftrightarrow\; \forall U. (U \perp\!\!\!\perp Y \to U \perp\!\!\!\perp X)$
  - Write $X \overset{\iota}{=} Y \;\Leftrightarrow\; X \overset{\iota}{\leq} Y \wedge Y \overset{\iota}{\leq} X$ and
    $X \overset{\iota}{<} Y \;\Leftrightarrow\; X \overset{\iota}{\leq} Y \wedge \neg(Y \overset{\iota}{\leq} X)$
- How to define the joint random variable of X, Y, written as XY?
  - $Z \overset{\iota}{=} XY \;\Leftrightarrow\; X \overset{\iota}{\leq} Z \wedge Y \overset{\iota}{\leq} Z \wedge \forall U. \left((X \overset{\iota}{\leq} U \wedge Y \overset{\iota}{\leq} U) \to Z \overset{\iota}{\leq} U\right)$
- How to define conditional independence, written as $X \perp\!\!\!\perp Y | Z$?
  - $X \perp\!\!\!\perp Y | Z \;\Leftrightarrow\; \exists U. U \perp\!\!\!\perp XZ \wedge Y \overset{\iota}{\leq} ZU$

Cardinality

- Check X is (at most) a binary random variable (i.e., $|\mathcal{X}| \leq 2$):

$$\mathrm{card}_{\leq 2}(X) \Leftrightarrow \forall U\big(U \overset{\iota}{<} X \rightarrow U \overset{\iota}{=} \emptyset\big)$$

  - Any random variable with strictly less information than X is degenerate
- The condition that $|\mathcal{X}| \leq n$:

$$\mathrm{card}_{\leq n}(X) \Leftrightarrow \forall U\big(U \overset{\iota}{<} X \rightarrow \mathrm{card}_{\leq n-1}(U)\big)$$

$$\mathrm{card}_{\leq 1}(X) \Leftrightarrow (X \overset{\iota}{=} \emptyset)$$

- Define

$$\mathrm{card}_{=n}(X) \Leftrightarrow \mathrm{card}_{\leq n}(X) \wedge \neg\mathrm{card}_{\leq n-1}(X)$$

$$\mathrm{card}_{\geq n}(X) \Leftrightarrow \neg\mathrm{card}_{\leq n-1}(X)$$

Uniformity

- If X, Y, Z are discrete random variables such that any one of them is a function of the other two, and they are pairwise independent, then they are all uniformly distributed over their supports, which have the same size [Zhang and Yeung, 1997]
- The condition that X is uniformly distributed over its support:

$$\mathrm{unif}(X) \Leftrightarrow \exists Y, Z.\, \mathrm{triple}(X, Y, Z),$$

where

$$\begin{aligned}\mathrm{triple}(X, Y, Z) \Leftrightarrow & X \overset{\iota}{\leq} YZ \wedge Y \overset{\iota}{\leq} XZ \wedge Z \overset{\iota}{\leq} XY \\ & \wedge X \perp\!\!\!\perp Y \wedge X \perp\!\!\!\perp Z \wedge Y \perp\!\!\!\perp Z\end{aligned}$$

  - Satisfied when $X, Y \sim \mathrm{Unif}\{0, \ldots, k-1\}$, $Z = X + Y \bmod k$

- Represent $k \in \mathbb{Z}_{>0}$ as a uniform random variable X with $|\mathcal{X}| = k$
- **Equality.** Formula for checking $|\mathcal{X}| = |\mathcal{Y}|$ for uniform X, Y [Li, 2021]:

$$\mathrm{ueq}(X, Y) \Leftrightarrow \exists U_1, U_2, U_3. \quad \mathrm{triple}(X, U_1, U_2) \wedge \mathrm{triple}(Y, U_1, U_3)$$

  - To check for equality against constants:

$$\mathrm{ueq}_n(X) \Leftrightarrow \mathrm{unif}(X) \wedge \mathrm{card}_{=n}(X)$$

- **Multiplication.** Formula for $|\mathcal{X}||\mathcal{Y}| = |\mathcal{Z}|$ for uniform X, Y, Z:

$$\mathrm{uprod}(X, Y, Z) \Leftrightarrow \exists \tilde{X}, \tilde{Y}. \left(\mathrm{ueq}(X, \tilde{X}) \wedge \mathrm{ueq}(Y, \tilde{Y}) \wedge \tilde{X} \perp\!\!\!\perp \tilde{Y} \wedge \tilde{X}\tilde{Y} \overset{\iota}{=} Z\right)$$

- **Comparison.** Formula for $|\mathcal{X}| \le |\mathcal{Y}|$ for uniform X, Y [Li, 2021]:

$$\mathrm{ule}(X, Y) \Leftrightarrow \exists G, \tilde{Y}. \left(\mathrm{uprod}(X, Y, G) \wedge \mathrm{ueq}(Y, \tilde{Y}) \wedge G \overset{\iota}{\le} Y\tilde{Y}\right)$$

  - "$\Leftarrow$": $G \overset{\iota}{\le} Y\tilde{Y} \Rightarrow |\mathcal{G}| \le |\mathcal{Y}||\tilde{\mathcal{Y}}| \Rightarrow |\mathcal{X}||\mathcal{Y}| \le |\mathcal{Y}|^2$
  - "$\Rightarrow$": $X \sim \mathrm{Unif}\{0, \ldots, a-1\}$, $Y \sim \mathrm{Unif}\{0, \ldots, b-1\}$, $G = (X, Y)$, $\tilde{Y} = X + Y \bmod b$

Addition between Integers

- To define addition, the main idea is that if Z is uniform with $|\mathcal{Z}| = |\mathcal{X}| + |\mathcal{Y}|$, then we can partition $\mathcal{Z}$ into two sets with sizes $|\mathcal{X}|, |\mathcal{Y}|$ respectively
  - If $U \in \{0,1\}$ is the indicator of whether Z is in the first set, then $U \sim \mathrm{Bern}(|\mathcal{X}|/(|\mathcal{X}| + |\mathcal{Y}|))$
- The following checks that X, Y, Z are uniform, $|\mathcal{Z}| = |\mathcal{X}| + |\mathcal{Y}|$, and $U \sim \mathrm{Bern}(|\mathcal{X}|/(|\mathcal{X}| + |\mathcal{Y}|))$:

$$\begin{aligned}\mathrm{frac}(X,Y,Z,U) \Leftrightarrow\ & \big(\mathrm{ueq}_2(U) \wedge \mathrm{uprod}(X,U,Z) \wedge \mathrm{uprod}(Y,U,Z)\big) \\ & \vee \exists \tilde{X}, \tilde{Y}. \big(\mathrm{ueq}(X,\tilde{X}) \wedge \mathrm{ueq}(Y,\tilde{Y}) \wedge \mathrm{unif}(Z) \\ & \wedge \mathrm{card}_{=2}(U) \wedge \neg\mathrm{unif}(U) \wedge U \overset{\iota}{\leq} Z \wedge \tilde{X} \perp\!\!\!\perp \tilde{Y} \perp\!\!\!\perp U \wedge Z \overset{\iota}{\leq} \tilde{X}\tilde{Y}U \\ & \wedge \forall V. \big(\mathrm{smi}(Z,V) \rightarrow \mathrm{smi}(\tilde{X}U, V) \vee \mathrm{smi}(\tilde{Y}U, V)\big)\big)\end{aligned}$$

where

$$\begin{aligned}\mathrm{smi}(X,Y) \Leftrightarrow\ & (X \overset{\iota}{=} Y \overset{\iota}{=} \emptyset) \vee \big(Y \overset{\iota}{\leq} X \wedge \mathrm{card}_{=2}(Y) \\ & \wedge \forall U. (U \overset{\iota}{\leq} X \wedge \mathrm{card}_{=4}(U) \rightarrow \neg\exists V. (\mathrm{card}_{\leq 2}(V) \wedge U \overset{\iota}{\leq} YV))\big)\end{aligned}$$

- We then have $\mathrm{usum}(X,Y,Z) \Leftrightarrow \exists U. \mathrm{frac}(X,Y,Z,U)$

Undecidability of FOTPI

Theorem (Li [2022a])

The first-order theory of probabilistic independence is undecidable, i.e., no algorithm can determine whether a statement in FOTPI holds

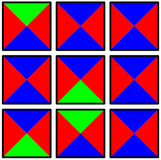
- Direct consequence of the fact that true arithmetic (over natural numbers) is interpretable in FOTPI, and that true arithmetic is undecidable [Tarski, 1933]

Undecidability of CI when one RV is Binary [Li, 2021]

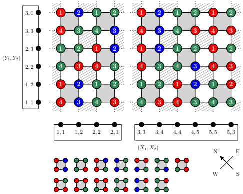
- It is **undecidable** to determine whether

$$|\mathcal{X}_1| \le 2 \wedge \bigwedge_{i=1}^{k} X_{A_i} \perp\!\!\!\perp X_{B_i} | X_{C_i} \Rightarrow X_{A_0} \perp\!\!\!\perp X_{B_0} | X_{C_0}$$

- Use $\mathrm{unif}(X_1)$ to force X_1 to be uniform, and make independent copies
- Use comparison to force any RV to have any cardinality
  - E.g. $a = 5$ is the only solution to $2^9 \le a^4 \le 2^{10}$
- Reduction from periodic tiling problem [Gurevich and Koryakov, 1972]: deciding whether a set of square tiles can tile a torus
- Use uniform RVs to represent coordinates and colors



$\Rightarrow$



Undecidability of CI Implication [Li, 2022b]

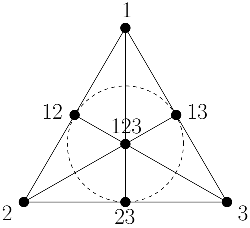
- It is **undecidable** to determine whether

$$\bigwedge_{i=1}^{k} X_{A_i} \perp\!\!\!\perp X_{B_i} | X_{C_i} \Rightarrow X_{A_0} \perp\!\!\!\perp X_{B_0} | X_{C_0}$$

  for given $(A_i)_i, (B_i)_i, (C_i)_i$
- Use the strategy in undecidability of embedded multivalued dependency [Herrmann, 1995]
- Show undecidability by reduction from uniform word problem for finite monoids [Gurevich, 1966]
- Problem – there is no algebraic structure in the RVs X_i!
- Have to impose some algebraic structure using conditional independence

Undecidability of CI Implication [Li, 2022b]

- RVs $A_1, A_2, A_3, A_{12}, A_{13}, A_{23}, A_{123}$
- Impose the "Fano-non-Fano condition":
  - For any three RVs on same solid line, any one is a function of other two
  - Any three RVs not on same solid/dotted line are independent



Lemma

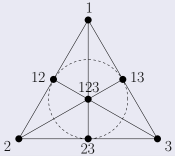
Fano-non-Fano condition holds iff A_1, A_2, A_3 are uniform elements in abelian group, and $A_{12} = A_1 + A_2$, $A_{13} = A_1 + A_3$, $A_{23} = A_2 + A_3$, $A_{123} = A_1 + A_2 + A_3$, up to relabeling

- Equivalent form used in [Herrmann, 1995] for undecidability of EMVD
- Appeared in [Dougherty et al., 2006a] to show unachievability of network coding capacity

Fano-non-Fano Condition

Lemma

Fano-non-Fano condition holds iff A_1, A_2, A_3 are uniform elements in abelian group, $A_{12} = A_1 + A_2$, $A_{13} = A_1 + A_3$, $A_{23} = A_2 + A_3$, $A_{123} = A_1 + A_2 + A_3$, up to relabeling



- A_k is a function of A_i, A_j, let this function be $f_k^{i,j}(a_i, a_j)$
- Bijection between independent (A_i, A_j, A_k) and $(A_i)_i$, let function from (A_i, A_j, A_k) to A_l be $f_l^{i,j,k}(a_i, a_j, a_k)$

Lemma

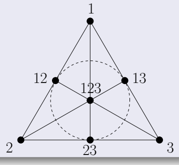
We have

- $f_k^{i,j}(a, b) = f_k^{j,i}(b, a)$, *and* $f_l^{i,j,k}(a, b, c) = f_l^{j,k,i}(b, c, a)$
- $f_i^{k,j}(f_k^{i,j}(a, b),\ b) = a$
- $f_l^{i,j,k}(a, b, c) = f_l^{m,k}(f_m^{i,j}(a, b), c)$

Fano-non-Fano Condition

Lemma

Fano-non-Fano condition holds iff A_1, A_2, A_3 *are uniform elements in abelian group,* $A_{12} = A_1 + A_2$, $A_{13} = A_1 + A_3$, $A_{23} = A_2 + A_3$, $A_{123} = A_1 + A_2 + A_3$, *up to relabeling*



- A_k is a function of A_i, A_j, let this function be $f_k^{i,j}(a_i, a_j)$
- Bijection between independent (A_i, A_j, A_k) and $(A_i)_i$, let function from (A_i, A_j, A_k) to A_l be $f_l^{i,j,k}(a_i, a_j, a_k)$
- $f_{12}^{1,2} = f_{12}^{2,1} = f_{13}^{1,3} = f_{13}^{3,1} = f_{23}^{2,3} = f_{23}^{3,2} = f_{123}^{1,23} = f_{123}^{2,13} = f_{123}^{3,12} = f_{123}^{23,1} = f_{123}^{13,2} = f_{123}^{12,3}$
- Define abelian group over $\mathcal{A}$ by $a + b := f_{12}^{1,2}(a, b)$, $-a := f_2^{1,12}(a, 0)$

Undecidability of CI Implication

- Strategy proposed by Dougherty [2009] – reduction from the identity problem for finite groups
  - Identity – equality that holds **for all** values of the variables, e.g., $\forall x, y.\, xy = yx$ (iff group is abelian)
  - Identity problem – whether a list of identities implies another identity

$$\bigwedge_{i=1}^{l} \left(\forall x_{1..k}.P_i(x_{1..k})\right) \;\rightarrow\; \forall x_{1..k}.P_0(x_{1..k})$$

  - Uniform RVs act as the universally-quantified variables
  - However, identity problem for finite groups is not known to be decidable or undecidable [Albert et al., 1992]!
- Herrmann [1995], Li [2022b]: instead use uniform word problem for finite monoids [Gurevich, 1966]
  - Whether a list of equalities implies another equality

$$\forall x_{1..k}.\left(\bigwedge_{i=1}^{l} P_i(x_{1..k}) \;\rightarrow\; P_0(x_{1..k})\right)$$

  - Need to use uniform RVs to represent specific monoid elements

Word Problem and Endomorphism Monoid

- Uniform word problem for finite monoids [Gurevich, 1966] – Given $a_i, b_i, c_i \in \{1, \ldots, k\}$ for $i = 0, \ldots, l$, determine whether the implication

$$\bigwedge_{i=1}^{l} (x_{a_i} \cdot x_{b_i} = x_{c_i}) \ \rightarrow \ (x_{a_0} = x_{c_0})$$

  holds for all finite monoids $\mathcal{M}$ and all k-tuples $x_1, \ldots, x_k \in \mathcal{M}$
- Consider endomorphism monoid of abelian group
  - Homomorphism $g : \mathcal{A} \rightarrow \mathcal{B}$ between abelian groups $\mathcal{A}, \mathcal{B}$ is a function satisfying $g(a + b) = g(a) + g(b)$
  - Endomorphism in $\mathcal{A}$ is a homomorphism $g : \mathcal{A} \rightarrow \mathcal{A}$
  - The *endomorphism monoid* $\mathrm{End}(\mathcal{A})$ is the set of endomorphisms in $\mathcal{A}$, equipped with the operation $g \cdot h : \mathcal{A} \rightarrow \mathcal{A}$ where $g \cdot h(a) = g(h(a))$
- Kurosh [1963] – For any finite monoid, there exists embedding from that monoid into $\mathrm{End}(\mathcal{A})$ for some finite abelian group $\mathcal{A}$
  - No loss of generality of considering only endomorphism monoids

Representing Endomorphism by RV

- A_1, A_2, A_3 are uniform elements in abelian group $\mathcal{A}$, and $A_{12} = A_1 + A_2$, $A_{13} = A_1 + A_3$, $A_{23} = A_2 + A_3$, $A_{123} = A_1 + A_2 + A_3$
- Represent an endomorphism $g : \mathcal{A} \to \mathcal{A}$ by $U = A_1 - g(A_2)$
- Check whether U corresponds to an endomorphism [Herrmann, 1995, Li, 2023]:

$$\begin{aligned} \mathrm{end}_{1,2}((A_i)_i, U) \;\Leftrightarrow\; \exists V, W :\; & \mathrm{FanoNonFano}((A_i)_i) \wedge \mathrm{ueq}(U, A_1) \\ & \wedge\ \mathrm{ueq}(V, A_1) \wedge \mathrm{ueq}(W, A_1) \wedge U \overset{\iota}{=} A_1 | A_2 \\ & \wedge\ V \overset{\iota}{=} A_1 | A_{23} \wedge U \overset{\iota}{=} V | A_3 \wedge W \overset{\iota}{=} A_{13} | A_2 \wedge U \overset{\iota}{=} W | A_3, \end{aligned}$$

  where $X \overset{\iota}{=} Y|Z \;\Leftrightarrow\; X \overset{\iota}{\leq} ZY \wedge Y \overset{\iota}{\leq} ZX$, i.e., if we are given Z, then X has the same information as Y, and $\mathrm{ueq}(X, Y)$ checks whether X, Y are both uniform and have the same cardinality
  - "$\Rightarrow$": $V = A_1 - g(A_2 + A_3)$, $W = A_1 - g(A_2) + A_3$
- Representing composition – if $\mathrm{end}_{1,2}((A_i)_i, U_1)$, $\mathrm{end}_{2,3}((A_i)_i, U_2)$, $\mathrm{end}_{1,3}((A_i)_i, U_3)$, we have $U_3 \overset{\iota}{\leq} U_1 U_2$ iff $g_3 = g_1 \cdot g_2$
  - "$\Leftarrow$": $U_3 = A_1 - g_1 \cdot g_2(A_3) = A_1 - g_1(A_2) + g_1(A_2 - g_2(A_3))$

Representing Endomorphism by RV

- A_1, A_2, A_3 are uniform elements in abelian group $\mathcal{A}$, and $A_{12} = A_1 + A_2$, $A_{13} = A_1 + A_3$, $A_{23} = A_2 + A_3$, $A_{123} = A_1 + A_2 + A_3$
- Represent an endomorphism $g : \mathcal{A} \to \mathcal{A}$ by $U = A_1 - g(A_2)$
- Check whether U corresponds to an endomorphism [Herrmann, 1995, Li, 2023]: $\mathrm{end}_{1,2}((A_i)_{i\in\mathcal{E}}, U)$
- Representing composition – if $\mathrm{end}_{1,2}((A_i)_i, U_1)$, $\mathrm{end}_{2,3}((A_i)_i, U_2)$, $\mathrm{end}_{1,3}((A_i)_i, U_3)$, we have $U_3 \overset{\iota}{\le} U_1 U_2$ iff $g_3 = g_1 \cdot g_2$
- Need to convert $\mathrm{end}_{2,3}, \mathrm{end}_{1,3}$ to $\mathrm{end}_{1,2}$
- Convert $\mathrm{end}_{i,j}$ for different i, j:

$$\begin{aligned}\mathrm{conv}^{1,2}_{1,3}((A_i)_i, U, V) \Leftrightarrow \exists W :\ & \mathrm{end}_{1,2}((A_i)_i, U) \\ & \wedge\ \mathrm{end}_{1,3}((A_i)_i, V) \wedge \mathrm{end}_{2,3}((A_i)_i, W) \\ & \wedge\ A_{13} \overset{\iota}{\le} A_{12} W \wedge V \overset{\iota}{\le} UW\end{aligned}$$

Representing Endomorphism by RV

- A_1, A_2, A_3 are uniform elements in abelian group $\mathcal{A}$, and $A_{12} = A_1 + A_2$, $A_{13} = A_1 + A_3$, $A_{23} = A_2 + A_3$, $A_{123} = A_1 + A_2 + A_3$
- Represent an endomorphism $g : \mathcal{A} \to \mathcal{A}$ by $U = A_1 - g(A_2)$
- Check whether U corresponds to an endomorphism [Herrmann, 1995, Li, 2023]: $\mathrm{end}_{1,2}((A_i)_{i\in\mathcal{E}}, U)$
- Representing composition – if $\mathrm{end}_{1,2}((A_i)_i, U_1)$, $\mathrm{end}_{2,3}((A_i)_i, U_2)$, $\mathrm{end}_{1,3}((A_i)_i, U_3)$, we have $U_3 \overset{\iota}{\le} U_1 U_2$ iff $g_3 = g_1 \cdot g_2$
- Convert $\mathrm{end}_{i,j}$ for different i, j: $\mathrm{conv}_{1,3}^{1,2}((A_i)_i, U, V)$
- Check whether U_1, U_2, U_3 with $\mathrm{end}_{1,2}((A_i)_i, U_j)$ satisfy $g_3 = g_1 \cdot g_2$:

$$
\begin{aligned}
&\mathrm{comp}_{1,2}((A_i)_i, U_1, U_2, U_3) \;\Leftrightarrow\\
&\exists V_1, V_2 : \bigwedge_{j=1}^{3} \mathrm{end}_{1,2}((A_i)_i, U_j) \wedge \mathrm{conv}_{1,3}^{1,2}((A_i)_i, U_1, V_1)\\
&\wedge \mathrm{conv}_{3,2}^{1,2}((A_i)_i, U_2, V_2) \wedge U_3 \overset{\iota}{\le} V_1 V_2
\end{aligned}
$$

Undecidability of CI Implication [Li, 2022b]

- Uniform word problem for finite monoids [Gurevich, 1966]

$$\bigwedge_{i=1}^{l} (x_{a_i} \cdot x_{b_i} = x_{c_i}) \;\rightarrow\; (x_{a_0} = x_{c_0})$$

holds for all finite monoid $\mathcal{M}$ and all k-tuples $x_1, \ldots, x_k \in \mathcal{M}$

is true iff...

-

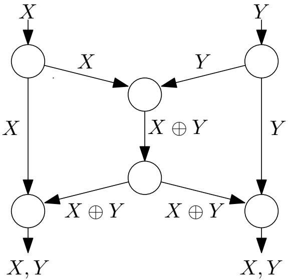
$$\left(\bigwedge_{j=1}^{k} \mathrm{end}_{1,2}((A_i)_i, U_j) \;\wedge\; \bigwedge_{j=1}^{l} \mathrm{comp}_{1,2}((A_i)_i, U_{a_j}, U_{b_j}, U_{c_j}) \right)$$
$$\rightarrow\; (U_{a_0} \overset{\iota}{\leq} U_{c_0})$$

holds for all finite random variables $(A_i)_i$, $U_1, \ldots, U_k$

- Since uniform word problem for finite monoids is undecidable, CI implication is undecidable as well

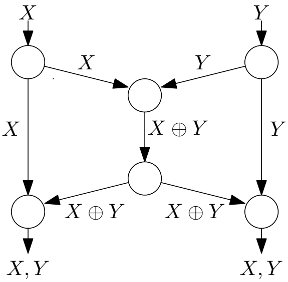
- Sequence of random variables $X^n = (X_1, \ldots, X_n)$
- Entropic vector [Zhang and Yeung, 1997] $\mathbf{h}(X^n) = \mathbf{h} \in \mathbb{R}^{2^n-1}$, where entries of $\mathbf{h}$ are indexed by nonempty subsets of $[n]$, and $\mathbf{h}_S := H(X_S)$
- Entropic region $\Gamma_n^* := \bigcup_{p_{X^n}} \{\mathbf{h}(X^n)\}$ [Zhang and Yeung, 1997]
- Non-Shannon inequalities (cannot be deduced from $I(X; Y|Z) \geq 0$) were given in [Zhang and Yeung, 1998, Makarychev et al., 2002, Dougherty et al., 2006b]
- Matúš [2007] showed that $\overline{\Gamma_n^*}$ is not polyhedral
- Conditional information inequalities: whether a linear inequality follows from a list of inequalities
  - Can encode conditional independence implication, and hence **undecidable** [Li, 2022b]
- Decidability of unconditional information inequalities is open

Related Problems: Network Coding



- Network coding [Ahlswede et al., 2000, Li et al., 2003]
  - Network of nodes connected by noiseless links with same capacity
  - Each source node has a message, and each destination node desires a set of messages
  - Each node is capable of performing coding, not only routing
- If there is one source and multiple destinations, the capacity (number of message bits per link capacity) is given by the maximum network flow [Ahlswede et al., 2000]
  - Single-source multicast network coding is **decidable**
- Significantly harder if there are multiple sources

Related Problems: Network Coding



- NP-hardness results: Lehman [2005], Langberg et al. [2006], Langberg and Sprintson [2011]
- Is network coding decidable?
  - Given a network, if the message size and the link capacity are the same, does there exist a valid coding scheme?
  - Partial result: whether a network admits a vector linear network code is undecidable [Kühne and Yashfe, 2019]
  - Shown to be **undecidable** in [Li, 2022b]
- Decidability of whether the capacity can be approached is unknown

Acknowledgements

The author acknowledges support from the Hong Kong Research Grant Council Grant ECS No. CUHK 24205621, and the Direct Grant for Research, The Chinese University of Hong Kong (Project ID: 4055133). The author would like to thank Chandra Nair and Raymond W. Yeung for their invaluable comments.

References I

Rudolf Ahlswede, Ning Cai, S-YR Li, and Raymond W Yeung. Network information flow. *IEEE Transactions on information theory*, 46(4):1204–1216, 2000.

Douglas Albert, Robert Baldinger, and John Rhodes. Undecidability of the identity problem for finite semigroups. *The Journal of symbolic logic*, 57(1):179–192, 1992.

Robert Berger. *The undecidability of the domino problem*. Number 66. American Mathematical Soc., 1966.

A Philip Dawid. Conditional independence in statistical theory. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(1):1–15, 1979.

R. Dougherty. Is network coding undecidable? In *Applications of Matroid Theory and Combinatorial Optimization to Information and Coding Theory*. The Banff International Research Station, 2009.

Randall Dougherty, Chris Freiling, and Kenneth Zeger. Unachievability of network coding capacity. *IEEE Transactions on Information Theory*, 52(6):2365–2372, 2006a.

Randall Dougherty, Christopher Freiling, and Kenneth Zeger. Six new non-Shannon information inequalities. In *2006 IEEE ISIT*, pages 233–236. IEEE, Jul 2006b.

Arnaud Durand, Miika Hannula, Juha Kontinen, Arne Meier, and Jonni Virtema. Probabilistic team semantics. In *Foundations of Information and Knowledge Systems: 10th International Symposium, FoIKS 2018, Budapest, Hungary, May 14–18, 2018, Proceedings 10*, pages 186–206. Springer, 2018.

References II

Dan Geiger and Christopher Meek. Quantifier elimination for statistical problems. In *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*, pages 226–235, 1999.

Dan Geiger and Judea Pearl. Logical and algorithmic properties of conditional independence and graphical models. *The Annals of Statistics*, pages 2001–2021, 1993.

Dan Geiger, Azaria Paz, and Judea Pearl. Axioms and algorithms for inferences involving probabilistic independence. *Information and Computation*, 91(1):128–141, 1991.

Yu Sh Gurevich and IO Koryakov. Remarks on Berger's paper on the domino problem. *Siberian Mathematical Journal*, 13(2):319–321, 1972.

Yurii Shlemovich Gurevich. The problem of equality of words for certain classes of semigroups. *Algebra i logika*, 5(5):25–35, 1966.

Miika Hannula, Åsa Hirvonen, Juha Kontinen, Vadim Kulikov, and Jonni Virtema. Facets of distribution identities in probabilistic team semantics. In *European Conference on Logics in Artificial Intelligence*, pages 304–320. Springer, Cham, 2019.

Miika Hannula, Minna Hirvonen, Juha Kontinen, Yasir Mahmood, Arne Meier, and Jonni Virtema. Logics with probabilistic team semantics and the boolean negation. In *European Conference on Logics in Artificial Intelligence*, pages 665–680. Springer, 2023.

References III

Christian Herrmann. On the undecidability of implications between embedded multivalued database dependencies. *Information and Computation*, 122(2):221–235, 1995.

Lukas Kühne and Geva Yashfe. Representability of matroids by c-arrangements is undecidable. *arXiv preprint arXiv:1912.06123*, 2019.

Lukas Kühne and Geva Yashfe. On entropic and almost multilinear representability of matroids. *arXiv preprint arXiv:2206.03465*, Jun 2022.

A. G. Kurosh. *Lectures on General Algebra*. Chelsea Publishing Company, New York, 1963.

Michael Langberg and Alex Sprintson. On the hardness of approximating the network coding capacity. *IEEE Transactions on Information Theory*, 57(2):1008–1014, 2011.

Michael Langberg, Alexander Sprintson, and Jehoshua Bruck. The encoding complexity of network coding. *IEEE Transactions on Information Theory*, 52(6):2386–2397, 2006.

April Rasala Lehman. *Network coding*. PhD thesis, Massachusetts Institute of Technology, 2005.

Cheuk Ting Li. The undecidability of conditional affine information inequalities and conditional independence implication with a binary constraint. In *2021 IEEE Information Theory Workshop*, 2021.

Cheuk Ting Li. First-order theory of probabilistic independence and single-letter characterizations of capacity regions. In *Proc. IEEE Int. Symp. Inf. Theory*. IEEE, 2022a.

Cheuk Ting Li. Undecidability of network coding, conditional information inequalities, and conditional independence implication, May 2022b. URL `https://arxiv.org/abs/2205.11461`.

Cheuk Ting Li. Undecidability of network coding, conditional information inequalities, and conditional independence implication. *IEEE Transactions on Information Theory*, 69(6):3493–3510, 2023. doi: 10.1109/TIT.2023.3247570.

S-YR Li, Raymond W Yeung, and Ning Cai. Linear network coding. *IEEE transactions on information theory*, 49(2):371–381, 2003.

Konstantin Makarychev, Yury Makarychev, Andrei Romashchenko, and Nikolai Vereshchagin. A new class of non-Shannon-type inequalities for entropies. *Communications in Information and Systems*, 2(2):147–166, 2002.

Francesco M Malvestuto. A unique formal system for binary decompositions of database relations, probability distributions, and graphs. *Information Sciences*, 59(1-2):21–52, 1992.

Yuri V Matiyasevich. Hilbert's tenth problem, 1993.

References V

František Matúš. Stochastic independence, algebraic independence and abstract connectedness. *Theoretical Computer Science*, 134(2):455–471, 1994.

Frantisek Matúš. Infinitely many information inequalities. In *2007 IEEE ISIT*, pages 41–44. IEEE, Jun 2007.

Michel Mouchart and Jean-Marie Rolin. A note on conditional independence. *Statistica*, 44:557, 1984.

Mathias Niepert. Logical inference algorithms and matrix representations for probabilistic conditional independence. *arXiv preprint arXiv:1205.2621*, 2012.

Petr Sergeevich Novikov. On the algorithmic unsolvability of the word problem in group theory. *Trudy Matematicheskogo Instituta imeni VA Steklova*, 44:3–143, 1955.

Judea Pearl and Azaria Paz. Graphoids: a graph-based logic for reasoning about relevance relations. *Advances in Artificial Intelligence*, pages 357–363, 1987.

Wolfgang Spohn. Stochastic independence, causal independence, and shieldability. *Journal of Philosophical logic*, 9(1):73–99, 1980.

Milan Studený. Multiinformation and the problem of characterization of conditional independence relations. *Problems of Control and Information Theory*, 18:3–16, 1989.

Milan Studený. Conditional independence relations have no finite complete characterization. *Information Theory, Statistical Decision Functions and Random Processes*, pages 377–396, 1992.

Alfred Tarski. *Pojęcie prawdy w językach nauk dedukcyjnych*. Number 34. Nakł. 'Tow. Naukowego Warszawskiego, 1933.

Alan Mathison Turing. On computable numbers, with an application to the entscheidungsproblem. *J. of Math*, 58(345-363):5, 1936.

Zhen Zhang and Raymond W Yeung. A non-Shannon-type conditional inequality of information quantities. *IEEE Trans. Inf. Theory*, 43(6):1982–1986, 1997.

Zhen Zhang and Raymond W Yeung. On characterization of entropy function via information inequalities. *IEEE Trans. Inf. Theory*, 44(4):1440–1452, 1998.