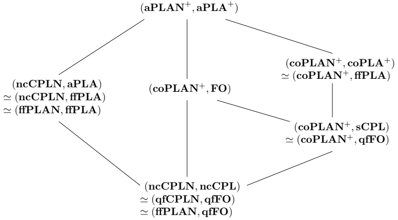
On the relative asymptotic expressivity of probabilistic inference frameworks



Vera Koponen, Uppsala University,
presenting joint work with
Felix Weitkämper, Ludwig Maximilians Universität, Munich.

Model Theoretic Logics and their Frontiers, January 14, 2022

- In the field of *Statistical Relational AI* one uses so-called *Probabilistic Graphical Models (PGM)* to model conditional probabilities between random variables.
- A PGM may determine a probability distribution $\mathbb{P}_D$ on the set $\mathbf{W}_D$ of structures with a given finite domain/universe D.
- We wish to compute/estimate the probability of an event $\mathbf{E} \subseteq \mathbf{W}_D$ which can be defined by a formula of some logic L.
- A larger variety of PGM's gives a larger, or more expressive, variety of probability distributions $\mathbb{P}_D$.
- A more expressive logic allows us to consider a larger variety of events.
- But high expressivity tends to be coupled with computational inefficiency for large domains D, at least by "brute force" methods.
- Thus we seek better than "brute force" methods as well as a suitable "trade-off" between expressivity and efficiency.
- For this it may be useful to compare the relative asymptotic (as $|D| \to \infty$) expressivity of different pairs $(\mathbb{P}_D, L)$.

Logics

Let σ be a finite and relational signature.
By a *logic (for σ)* we mean a set L of objects, called *formulas*, such that the following hold:

1. For every $\varphi \in L$ a finite set $Fv(\varphi)$ of so-called *free variables* is associated to φ. If we write $\varphi(\bar{x})$ where $\varphi \in L$ then we mean that $Fv(\varphi) \subseteq \bar{x}$ and when using this notation we assume that there are no repetitions in the sequence $\bar{x}$.
2. To every triple $(\varphi(\bar{x}), \mathcal{A}, \bar{a})$ such that $\varphi(\bar{x}) \in L$, $\mathcal{A}$ is a finite σ-structure and $\bar{a} \in A^{|\bar{x}|}$ a number $\alpha \in [0, 1]$ is associated. We write $\mathcal{A}(\varphi(\bar{a})) = \alpha$ to express that α is the number, or *(truth) value*, associated to the triple $(\varphi(\bar{x}), \mathcal{A}, \bar{a})$.

Logics

Let σ be a finite and relational signature.
By a *logic (for σ)* we mean a set L of objects, called *formulas*, such that the following hold:

1. For every $\varphi \in L$ a finite set $Fv(\varphi)$ of so-called *free variables* is associated to φ. If we write $\varphi(\bar{x})$ where $\varphi \in L$ then we mean that $Fv(\varphi) \subseteq \bar{x}$ and when using this notation we assume that there are no repetitions in the sequence $\bar{x}$.
2. To every triple $(\varphi(\bar{x}), \mathcal{A}, \bar{a})$ such that $\varphi(\bar{x}) \in L$, $\mathcal{A}$ is a finite σ-structure and $\bar{a} \in A^{|\bar{x}|}$ a number $\alpha \in [0, 1]$ is associated. We write $\mathcal{A}(\varphi(\bar{a})) = \alpha$ to express that α is the number, or *(truth) value*, associated to the triple $(\varphi(\bar{x}), \mathcal{A}, \bar{a})$.

The expressions '$\mathcal{A} \models \varphi(\bar{x})$' and '$\mathcal{A} \not\models \varphi(\bar{x})$' mean the same as $\mathcal{A}(\varphi(\bar{a})) = 1$ and $\mathcal{A}(\varphi(\bar{a})) = 0$, respectively.

For logics L and L', $L \leq L'$ means that if for every $\varphi(\bar{x}) \in L$ there is $\varphi'(\bar{x}) \in L'$ such that for every finite σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{x}|}$, $\mathcal{A}(\varphi(\bar{a})) = \mathcal{A}(\varphi'(\bar{a}))$.

Sequences of probability distributions

For every $n \in \mathbb{N}^+ = \{1, 2, 3, \ldots\}$, let $\mathbf{W}_n$ be the set of all σ-structures with domain $[n] = \{1, \ldots, n\}$.

Let $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$ and $\mathbb{P}' = (\mathbb{P}'_n : n \in \mathbb{N}^+)$ where, for each n, $\mathbb{P}_n$ and $\mathbb{P}'_n$ are probability distributions on $\mathbf{W}_n$.

Definition. $\mathbb{P}$ and $\mathbb{P}'$ are *asymptotically total variation equivalent*, denoted $\mathbb{P} \sim_{tv} \mathbb{P}'$, if there is a function $\delta : \mathbb{N}^+ \to \mathbb{R}$ such that $\lim_{n\to\infty} \delta(n) = 0$ and for all sufficiently large n and every $\mathbf{X} \subseteq \mathbf{W}_n$, $|\mathbb{P}_n(\mathbf{X}) - \mathbb{P}'_n(\mathbf{X})| \leq \delta(n)$.

Sequences of probability distributions

For every $n \in \mathbb{N}^+ = \{1, 2, 3, \dots\}$, let $\mathbf{W}_n$ be the set of all σ-structures with domain $[n] = \{1, \dots, n\}$.

Let $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$ and $\mathbb{P}' = (\mathbb{P}'_n : n \in \mathbb{N}^+)$ where, for each n, $\mathbb{P}_n$ and $\mathbb{P}'_n$ are probability distributions on $\mathbf{W}_n$.

Definition. $\mathbb{P}$ and $\mathbb{P}'$ are *asymptotically total variation equivalent*, denoted $\mathbb{P} \sim_{tv} \mathbb{P}'$, if there is a function $\delta : \mathbb{N}^+ \to \mathbb{R}$ such that $\lim_{n\to\infty} \delta(n) = 0$ and for all sufficiently large n and every $\mathbf{X} \subseteq \mathbf{W}_n$, $|\mathbb{P}_n(\mathbf{X}) - \mathbb{P}'_n(\mathbf{X})| \leq \delta(n)$.

Let L be a logic and let $\varphi(\bar{x}), \psi(\bar{x}) \in L$.
Definition. $\varphi(\bar{x})$ and $\psi(\bar{x})$ are *asymptotically equivalent with respect to* $\mathbb{P}$, denoted $\varphi(\bar{x}) \sim_{\mathbb{P}} \psi(\bar{x})$, if for all $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}_n\Big(\big\{\mathcal{A} \in \mathbf{W}_n : \exists \bar{a} \in A^{|\bar{x}|},\ |\mathcal{A}(\varphi(\bar{a})) - \mathcal{A}(\psi(\bar{a}))| > \varepsilon\big\}\Big) = 0.$$

Note that if φ and ψ are 0/1-valued then $\varphi \sim_{\mathbb{P}} \psi$ if and only if they are almost surely equivalent w.r.t. $\mathbb{P}$.

Inference frameworks

Definition. An *inference framework (for σ)* is a set $\mathbf{F}$ of pairs $(\mathbb{P}, L)$ where L is a logic (for σ) and $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$ where each $\mathbb{P}_n$ is a probability distribution on $\mathbf{W}_n$.

Definition. $\mathbf{F}'$ is *asymptotically at least as expressive as* $\mathbf{F}$, denoted $\mathbf{F} \preccurlyeq \mathbf{F}'$, if for every $(\mathbb{P}, L) \in \mathbf{F}$ there is $(\mathbb{P}', L') \in \mathbf{F}'$ such that $\mathbb{P} \sim_{tv} \mathbb{P}'$ and for every $\varphi(\bar{x}) \in L$ there is $\varphi'(\bar{x}) \in L'$ such that $\varphi(\bar{x}) \sim_{\mathbb{P}} \varphi'(\bar{x})$.

$\mathbf{F}$ and $\mathbf{F}'$ are *asymptotically equally expressive*, denoted $\mathbf{F} \simeq \mathbf{F}'$, if $\mathbf{F} \preccurlyeq \mathbf{F}'$ and $\mathbf{F}' \preccurlyeq \mathbf{F}$.

Lemma. $\preccurlyeq$ is reflexive and transitive.

Convergence law in the context of (possibly) many valued logic

Definition. Let L be a logic and $\mathbb{P}$ a sequence of probability distributions (on $\mathbf{W}_n, n \in \mathbb{N}^+$). We say that $(\mathbb{P}, L)$ *has a convergence law* if for every $\varphi(\bar{x}) \in L$ there are k and $c_1, \dots, c_k \in [0, 1]$ such that for all $m \in \mathbb{N}^+$, all $\bar{a} \in [m]^{|\bar{x}|}$ and all $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n\Big(\big\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) \in \bigcup_{i=1}^{k} [c_i - \varepsilon, c_i + \varepsilon]\big\}\Big) = 1$$

and, for all $i = 1, \dots, k$,

$$\lim_{n\to\infty} \mathbb{P}_n\Big(\big\{\mathcal{A} \in \mathbf{W}_n : |\mathcal{A}(\varphi(\bar{a})) - c_i| < \varepsilon\big\}\Big) \text{ exists.}$$

Remark: If L is a 0/1-valued logic then the above is equivalent to the usual definition of convergence law for L and $\mathbb{P}$.

Consequences of $\preccurlyeq$

Let $\mathbf{F}$ and $\mathbf{F}'$ be inference frameworks.

Lemma. (Transfer of a convergence law) If $\mathbf{F} \preccurlyeq \mathbf{F}'$ and every $(\mathbb{P}', L') \in \mathbf{F}'$ has a convergence law, then every $(\mathbb{P}, L) \in \mathbf{F}$ has a convergence law.

Lemma. (Transfer of asymptotic elimination to a smaller logic) Suppose that L_0 is a logic such that, for every $(\mathbb{P}, L) \in \mathbf{F}$ and every $(\mathbb{P}', L') \in \mathbf{F}'$, $L_0 \subseteq L$ and $L_0 \subseteq L'$. Furthermore suppose that

- $\mathbf{F} \preccurlyeq \mathbf{F}'$, and
- for every $(\mathbb{P}', L') \in \mathbf{F}'$ and every $\varphi'(\bar{x}) \in L'$ there is $\psi(\bar{x}) \in L_0$ such that $\varphi'(\bar{x}) \sim_{\mathbb{P}'} \psi(\bar{x})$.

Then, for every $(\mathbb{P}, L) \in \mathbf{F}$ and every $\varphi(\bar{x}) \in L$, there is $\psi(\bar{x}) \in L_0$ such that $\varphi(\bar{x}) \sim_{\mathbb{P}} \psi(\bar{x})$.

Conditional probability logic (CPL)

Let *FO* denote the set of all first-order formulas (formed from a fixed finite relational signature/vocubulary σ).

Conditional probability logic (CPL) is an extension of *FO* where, in addition to the constructions of FO, the following construction is allowed:

If $r \geq 0$ is a real number, $\varphi, \psi, \theta, \tau \in CPL$ and $\bar{y}$ is a sequence of distinct variables, then

$$\Big(r + \|\varphi \mid \psi\|_{\bar{y}} \ \geq \ \|\theta \mid \tau\|_{\bar{y}} \Big) \in CPL \quad \text{and}$$
$$\Big(\|\varphi \mid \psi\|_{\bar{y}} \ \geq \ \|\theta \mid \tau\|_{\bar{y}} + r \Big) \in CPL.$$

In both these new formulas all variables of φ, ψ, θ and τ that appear in the sequence $\bar{y}$ become *bound*.

We write $CPL(\sigma)$ if we want to emphasize that the signature used is σ.

Semantics of CPL

Notation: For a formula $\varphi(\bar{x}, \bar{y})$, a structure $\mathcal{A}$ and $\bar{a} \in A^{|x|}$, we let $\varphi(\bar{a}, \mathcal{A}) = \{\bar{b} \in A^{|\bar{y}|} : \mathcal{A} \models \varphi(\bar{a}, \bar{b})\}$.

Let $\varphi(\bar{x}, \bar{y}), \psi(\bar{x}, \bar{y}), \theta(\bar{x}, \bar{y}), \tau(\bar{x}, \bar{y}) \in CPL$, let $\mathcal{A}$ be a finite structure and let $\bar{a} \in A^{|\bar{x}|}$.

If $\psi(\bar{a}, \mathcal{A}) \neq \emptyset$, $\tau(\bar{a}, \mathcal{A}) \neq \emptyset$ and

$$r + \frac{\left|\varphi(\bar{a}, \mathcal{A}) \cap \psi(\bar{a}, \mathcal{A})\right|}{\left|\psi(\bar{a}, \mathcal{A})\right|} \;\geq\; \frac{\left|\theta(\bar{a}, \mathcal{A}) \cap \tau(\bar{a}, \mathcal{A})\right|}{\left|\tau(\bar{a}, \mathcal{A})\right|}$$

then

$$\mathcal{A}\Big(r + \|\varphi(\bar{a}, \bar{y}) \mid \psi(\bar{a}, \bar{y})\|_{\bar{y}} \;\geq\; \|\theta(\bar{a}, \bar{y}) \mid \tau(\bar{a}, \bar{y})\|_{\bar{y}}\Big) = 1.$$

Otherwise

$$\mathcal{A}\Big(r + \|\varphi(\bar{a}, \bar{y}) \mid \psi(\bar{a}, \bar{y})\|_{\bar{y}} \;\geq\; \|\theta(\bar{a}, \bar{y}) \mid \tau(\bar{a}, \bar{y})\|_{\bar{y}}\Big) = 0.$$

CPL-network

If G is a directed acyclic graph (DAG) and v is a vertex of G, then $par(v)$ denotes the set of parents of v.

Let σ be a finite and relational signature.

A $CPL(\sigma)$-network $\mathbb{G}$ is determined by the following items:

(a) A DAG, also denoted $\mathbb{G}$, with vertex set σ.

(b) For each $R \in \sigma$, a number $k_R \in \mathbb{N}^+$, formulas $\chi_{R,i}(\bar{x}) \in CPL(par(R))$, for $i = 1, \ldots, k_R$, where $|\bar{x}|$ equals the arity of R, such that $\forall \bar{x}\big(\bigvee_{i=1}^{k_R} \chi_{R,i}(\bar{x})\big)$ is valid and if $i \neq j$ then $\exists \bar{x}\big(\chi_{R,i}(\bar{x}) \wedge \chi_{R,j}(\bar{x})\big)$ is unsatisfiable. Each $\chi_{R,i}$ will be called an *aggregation formula (of* $\mathbb{G}$*)*.

(c) For each $R \in \sigma$ and each $1 \leq i \leq k_R$, a number denoted $\mu(R \mid \chi_{R,i})$, or $\mu(R(\bar{x}) \mid \chi_{R,i}(\bar{x}))$, in the interval $[0, 1]$.

Note that if $par(R) = \emptyset$, then χ_{R_i} is a formula such that all of its atomic subformulas have the form $x = y$.

The probability distribution induced by a *CPL*-network

Let $\mathbb{G}$ be a $CPL(\sigma)$-network.

For every finite σ-structure $\mathcal{A}$, every $R \in \sigma$, of arity r say, every $\bar{a} \in A^r$ and every $1 \leq i \leq k_R$, let

$$\lambda(\mathcal{A}, R, i, \bar{a}) = \begin{cases} \mu(R \mid \chi_{R,i}) & \text{if } \mathcal{A} \models \chi_{R,i}(\bar{a}) \wedge R(\bar{a}), \\ 1 - \mu(R \mid \chi_{R,i}) & \text{if } \mathcal{A} \models \chi_{R,i}(\bar{a}) \wedge \neg R(\bar{a}), \\ 0 & \text{otherwise.} \end{cases}$$

Recall that $\mathbf{W}_n$ is the set of all σ-structures with domain $[n]$.

For every $\mathcal{A} \in \mathbf{W}_n$, define

$$\mathbb{P}_n(\mathcal{A}) = \prod_{R \in \sigma} \prod_{i=1}^{k_R} \prod_{\bar{a} \in \chi_{R,i}(\mathcal{A})} \lambda(\mathcal{A}, R, i, \bar{a}).$$

In the sequel we fix some finite and relational signature σ and often omit it from the notation.

Noncritical numbers, formulas and *CPL*-networks

Let $\mathbb{G}$ be a *CPL*-network. For every $l \in \mathbb{N}$ there are a *finite* set of real numbers which we call *l-critical w.r.t.* $\mathbb{G}$. *The l-critical numbers depend only on the conditional probabilities associated to* $\mathbb{G}$.

A *CPL*-formula $\varphi(\bar{x})$ is *noncritical w.r.t.* $\mathbb{G}$ if, for every subformula of φ of the form

$$\big(r + \|\chi \mid \psi\|_{\bar{y}} \ \geq \ \|\theta \mid \tau\|_{\bar{y}}\big) \text{ or } \big(\|\chi \mid \psi\|_{\bar{y}} \ \geq \ \|\theta \mid \tau\|_{\bar{y}} + r\big),$$

r is not the difference of two l-critical numbers where $l = |\bar{x}| +$ the quantifier rank of $\varphi(\bar{x})$.

A *CPL*-network $\mathbb{G}$ is called *noncritical/quantifier-free*, or a *ncCPL*-network/*qfCPL*-network, if every aggregation formula of $\mathbb{G}$ is noncritical w.r.t. $\mathbb{G}$, respectively quantifier-free.

Reduction of (**ncCPLN**, **ncCPL**)

By (**ncCPLN**, **ncCPL**) we denote the inference framework consisting of all $(\mathbb{P}, L)$ where $\mathbb{P}$ is induced by a *ncCPL*-network $\mathbb{G}$ and L consists of all *CPL*-formulas which are *noncritical with respect to* $\mathbb{G}$.

By (**qfCPLN**, **qfFO**) we denote the inference framework consisting of all $(\mathbb{P}, L)$ where $\mathbb{P}$ is induced by a *qfCPL*-network and L consists of all *quantifier-free FO*-formulas.

Theorem. [K] For every $(\mathbb{P}, L) \in$ (**ncCPLN**, **ncCPL**) and every $\varphi(\bar{x}) \in L$ there is a quantifier-free $\psi(\bar{x}) \in L$ such that $\varphi(\bar{x}) \sim_{\mathbb{P}} \psi(\bar{x})$ (i.e. $\varphi(\bar{x})$ and $\psi(\bar{x})$ are almost surely equivalent).

Corollary. Every $(\mathbb{P}, L) \in$ (**ncCPLN**, **ncCPL**) has a convergence law.

Corollary. (**ncCPLN**, **ncCPL**) $\simeq$ (**qfCPLN**, **qfFO**).

Logics with aggregation functions

We now consider logics which use so-called aggregation functions instead of (generalized) quantifiers.

(For a discussion on the relationship between aggregation functions and generalized quantifiers, see the last section of [KW1].)

Aggregation functions

Let $[0,1]^{<\omega}$ denote the set of all finite sequences of reals in the unit interval $[0,1]$.

Let $F : \left([0,1]^{<\omega}\right)^k \to [0,1]$.
We call F an *aggregation function* if F is symmetric in the sense that if $\bar{r}_1, \ldots, \bar{r}_k \in [0,1]^{<\omega}$ and for each $i = 1, \ldots, k$, $\bar{\rho}_i$ is an arbitrary reordering of the entries of $\bar{r}_i$, then $F(\bar{\rho}_1, \ldots, \bar{\rho}_k) = F(\bar{r}_1, \ldots, \bar{r}_k)$.

Common aggregation functions: For $\bar{r} = (r_1, \ldots, r_n) \in [0,1]^{<\omega}$, define

1. $\max(\bar{r})$ to be the *maximum* of all r_i,
2. $\min(\bar{r})$ to be the *minimum* of all r_i,
3. $am(\bar{r}) = (r_1 + \ldots + r_n)/n$, so 'am' is the *arithmetic mean*.
4. $gm(\bar{r}) = \left(\prod_{i=1}^n r_i\right)^{(1/n)}$, so 'gm' is the *geometric mean*.
5. $\text{noisy-or}(\bar{r}) = 1 - \prod_{i=1}^n (1 - r_i)$.

Remark: max and min can play the role of $\exists$ and $\forall$.

Convergence testing sequences

A sequence $\bar{r}_n \in [0,1]^{<\omega}$, $n \in \mathbb{N}$, is called *convergence testing* with parameters $c_1, \ldots, c_k \in [0,1]$ and $\alpha_1, \ldots \alpha_k \in [0,1]$ if the following hold, where $r_{n,i}$ denotes the ith entry of $\bar{r}_n$:

1. $|\bar{r}_n| < |\bar{r}_{n+1}|$ for all $n \in \mathbb{N}$.
2. For every disjoint family of open intervals $I_1, \ldots I_k \subseteq [0,1]$ such that $c_i \in I_i$ for each i, there is an $N \in \mathbb{N}$ such that $\operatorname{rng}(\bar{r}_n) \subseteq \bigcup_{j=1}^{k} I_j$ for all $n \geq N$, and for every $j \in \{1, \ldots, k\}$,

$$\lim_{n \to \infty} \frac{|\{i \leq |\bar{r}_n| : r_{n,i} \in I_j\}|}{|\bar{r}_n|} = \alpha_j$$

Admissible aggregation functions

For simplicity we only give the definition of *(strong) admissibility* for unary aggregation functions $F : [0,1]^{<\omega} \to [0,1]$.

Definition. (i) An aggregation function $F : [0,1]^{<\omega} \to [0,1]$ is called *strongly admissible* if the following two conditions hold:

1. For all $n \in \mathbb{N}^+$, F is continuous on the set $[0,1]^n$.
2. For all convergence testing sequences of tuples $\bar{r}_n \in [0,1]^{<\omega}$, $n \in \mathbb{N}$, and $\bar{\rho}_n \in [0,1]^{<\omega}$, $n \in \mathbb{N}$, with the same parameters $c_1, \dots, c_k \in [0,1]$ and $\alpha_1, \dots, \alpha_k \in [0,1]$, $\lim_{n\to\infty} |F(\bar{r}_n) - F(\bar{\rho}_n)| = 0$.

(ii) An aggregation function $F : [0,1]^{<\omega} \to [0,1]$ is called *admissible* if condition (1) above holds and condition (2) above holds whenever the parameters α_i are *positive* for all i.

Admissible aggregation functions

For simplicity we only give the definition of *(strong) admissibility* for unary aggregation functions $F : [0,1]^{<\omega} \to [0,1]$.

Definition. (i) An aggregation function $F : [0,1]^{<\omega} \to [0,1]$ is called *strongly admissible* if the following two conditions hold:

1. For all $n \in \mathbb{N}^+$, F is continuous on the set $[0,1]^n$.
2. For all convergence testing sequences of tuples $\bar{r}_n \in [0,1]^{<\omega}$, $n \in \mathbb{N}$, and $\bar{\rho}_n \in [0,1]^{<\omega}$, $n \in \mathbb{N}$, with the same parameters $c_1, \ldots, c_k \in [0,1]$ and $\alpha_1, \ldots, \alpha_k \in [0,1]$, $\lim_{n\to\infty} |F(\bar{r}_n) - F(\bar{\rho}_n)| = 0$.

(ii) An aggregation function $F : [0,1]^{<\omega} \to [0,1]$ is called *admissible* if condition (1) above holds and condition (2) above holds whenever the parameters α_i are *positive* for all i.

Proposition. [KW1] (i) The functions am (*arithmetic mean*) and gm (*geometric mean*) are strongly admissible.
(ii) The functions max and min are admissible but not strongly admissible.

Syntax of PLA^+ (PLA = Probability Logic with Aggregation functions)

1. For each $c \in [0,1]$, $c \in PLA^+$.
2. For all variables x and y, '$x = y$' belongs to PLA^+.
3. For every relation symbol in the signature, say of arity r, and any choice of variables $x_1, \dots, x_r$, $R(x_1, \dots, x_r)$ belongs to PLA^+.
4. If $n \in \mathbb{N}^+$, $\varphi_1(\bar{x}), \dots, \varphi_n(\bar{x})_n \in PLA^+$ and $\mathsf{C} : [0,1]^n \to [0,1]$ is a continuous, then $\mathsf{C}(\varphi_1, \dots, \varphi_n)$ belongs to PLA^+.
5. If $k \in \mathbb{N}^+$, $\varphi_1(\bar{x},\bar{y}), \dots, \varphi_k(\bar{x},\bar{y}) \in PLA^+$, $p^=(\bar{x},\bar{y})$ is a complete specification of the equalities and nonequalities between the involved variables, where $\bar{x}$ and $\bar{y}$ are disjoint sequences of distinct variables and $F : \left([0,1]^{<\omega}\right)^k \to [0,1]$ is an aggregation function, then
$$F(\varphi_1(\bar{x},\bar{y}), \dots, \varphi_k(\bar{x},\bar{y}) : \bar{y} : p^=(\bar{x},\bar{y}))$$
is a formula of PLA^+ and this construction binds the variables in $\bar{y}$.
6. If $k \in \mathbb{N}^+$, $\varphi_1(\bar{x},\bar{y}), \dots, \varphi_k(\bar{x},\bar{y}) \in PLA^+$ and $F : \left([0,1]^{<\omega}\right)^k \to [0,1]$ is an aggregation function, then
$$F(\varphi_1(\bar{x},\bar{y}), \dots, \varphi_k(\bar{x},\bar{y}) : \bar{y})$$
is a formula of PLA^+ and this construction binds the variables in $\bar{y}$.

Semantics of PLA^+

Let $\mathcal{A}$ be a finite structure and $\bar{a} \in A^r$ for a suitable r.

1. For every $c \in [0,1]$, $\mathcal{A}(c) = c$.
2. For all $a, b \in A$, $\mathcal{A}(a = b) = 1$ if $a = b$, otherwise $\mathcal{A}(a = b) = 0$.
3. For every relation symbol R, $\mathcal{A}(R(\bar{a})) = 1$ if $\mathcal{A} \models R(\bar{a})$, otherwise $\mathcal{A}(R(\bar{a})) = 0$.
4. If $\mathsf{C} : [0,1]^k \to [0,1]$ then $\mathcal{A}\big(\mathsf{C}(\varphi_1(\bar{a}), \dots, \varphi_k(\bar{a}))\big) \ = \ \mathsf{C}\big(\mathcal{A}_k(\varphi(\bar{a})), \dots, \mathcal{A}(\varphi_k(\bar{a}))\big)$.
5. If $F : \big([0,1]^{<\omega}\big)^k \to [0,1]$ is an aggregation function then
$$\mathcal{A}\big(F(\varphi_1(\bar{a}, \bar{y}), \dots, \varphi_k(\bar{a}, \bar{y}) : \bar{y} : p^{=}(\bar{a}, \bar{y}))\big) = F(\bar{r}_1, \dots, \bar{r}_k)$$
where, for $i = 1, \dots, k$,
$$\bar{r}_i = \big(\mathcal{A}(\varphi_i(\bar{a}, \bar{b})) : \bar{b} \in A^{|\bar{y}|} \text{ and } p^{=}(\bar{a}, \bar{b}) \text{ holds}\big).$$
6. If $F : \big([0,1]^{<\omega}\big)^k \to [0,1]$ is an aggregation function then
$$\mathcal{A}\big(F(\varphi_1(\bar{a}, \bar{y}), \dots, \varphi_k(\bar{a}, \bar{y}) : \bar{y}\big) = F(\bar{r}_1, \dots, \bar{r}_k)$$
where, for $i = 1, \dots, k$, $\bar{r}_i = \big(\mathcal{A}(\varphi_i(\bar{a}, \bar{b})) : \bar{b} \in A^{|\bar{y}|}\big)$.

Some special continuous connectives

A function $C : [0,1]^n \to [0,1]$ (as in item (4) in the syntax of PLA^+) will also be called a *connective*.

1. Let $\neg : [0,1] \to [0,1]$ be defined by $\neg(x) = 1 - x$.
2. Let $\wedge : [0,1]^2 \to [0,1]$ be defined by $\wedge(x,y) = \min(x,y)$.
3. Let $\vee : [0,1]^2 \to [0,1]$ be defined by $\vee(x,y) = \max(x,y)$.
4. Let $\to: [0,1]^2 \to [0,1]$ be defined by $\to(x,y) = \min(1,\ 1 - x + y)$.
5. Let $wm : [0,1]^3 \to [0,1]$ (where *wm* stands for *weighted mean*) be defined by $wm(x,y,z) = x \cdot y + (1-x)z$.

The first four of these (which use Lukasiewicz semantics) have the usual meaning when restricted to the truth values 0 and 1.

PLA

PLA is subset of PLA^+ obtained by

- omitting (6) in the syntax of PLA^+ and
- restricting (4) in the syntax of PLA^+ to the connectives on the previous slide.

We write $PLA(\sigma)$, $PLA^+(\sigma)$ etc if we want to emphasize that the signature used is σ.

PLA^+-networks

Let σ be a finite relational signature.

Definition. A $PLA^+(\sigma)$*-network* $\mathbb{G}$ is determined by the following components:

1. A DAG (also denoted $\mathbb{G}$) with vertex set σ.
2. To each relation symbol $R \in \sigma$ a formula $\theta_R(\bar{x}) \in PLA^+(par(R))$ is associated where $|\bar{x}|$ equals the arity of R. Every such θ_R is called an *aggregation formula* of the $PLA^+(\sigma)$-network.

Definition of probability distribution on $\mathbf{W}_n$ induced by a $PLA^+(\sigma)$-network $\mathbb{G}$:

For every $\mathcal{A} \in \mathbf{W}_n$,

$$\mathbb{P}_n(\mathcal{A}) = \prod_{R\in\sigma} \prod_{\bar{a}\in R^{\mathcal{A}}} \mathcal{A}(\theta_R(\bar{a})) \prod_{\bar{a}\in[n]^{k_R} \setminus R^{\mathcal{A}}} \big(1 - \mathcal{A}(\theta_R(\bar{a}))\big)$$

where k_R denotes the arity of R.

Admissible/function-free *PLA*-formulas and PLA^+-networks

A *PLA*-formula is called *function-free* if it contains no aggregation function.

A *PLA*-formula is called *admissible* if every aggregation function in it is admissible.

Let *aPLA*/*ffPLA* denote the set of *admissible/function-free* *PLA*-formulas, respectively.

We call a PLA^+-network $\mathbb{G}$ *function-free*, or an *ffPLA*-network, if every aggregation formula of $\mathbb{G}$ is function-free.

By

$$(\mathbf{ncCPLN}, \mathbf{aPLA}) \ / \ (\mathbf{ncCPLN}, \mathbf{ffPLA}) \ / \ (\mathbf{ffPLAN}, \mathbf{ffPLA})$$

we denote the inference framework consisting of all $(\mathbb{P}, L)$ where $\mathbb{P}$ is induced by a *ncCPL*/*ncCPL*/*ffPLA*-network and $L = aPLA/ffPLA/ffPLA$, respectively.

Asymptotic reduction of $(\mathbf{ncCPLN}, \mathbf{aPLA})$

Theorem. [KW1] For every $(\mathbb{P}, L) \in (\mathbf{ncCPLN}, \mathbf{aPLA})$ and every $\varphi(\bar{x}) \in L$ there is a function-free $\psi(\bar{x}) \in L$ such that $\varphi(\bar{x}) \sim_{\mathbb{P}} \psi(\bar{x})$.

Corollary. Every $(\mathbb{P}, L) \in (\mathbf{ncCPLN}, \mathbf{aPLA})$ has a convergence law.

Corollary.
$(\mathbf{ncCPLN}, \mathbf{aPLA}) \simeq (\mathbf{ncCPLN}, \mathbf{ffPLA}) \simeq (\mathbf{ffPLAN}, \mathbf{ffPLA})$.

Since, for example, the *aPLA*-formula '1/2' is not asymptotically equivalent to any *CPL*-formula we also have

$$(\mathbf{ncCPLN}, \mathbf{ncCPL}) \prec (\mathbf{ncCPLN}, \mathbf{aPLA}).$$

$coPLA^+$

Admissibility of aggregation function is a sort of continuity condition which makes max and min "continuous".

However, there is an "asymmetry" in the definition of admissiblity since the parameters α_i must be positive.

The notion of *strong admissibility* allows α_i to be any number in the interval $[0, 1]$ and hence feels like a more natural continuity condition.

Let $coPLA^+$ be the subset of PLA^+ obtained by allowing only *strongly admissible* aggregation functions in parts (5) and (6) of the syntax of PLA^+.

We have seen that arithmetic and geometric means are strongly admissible. Also, for every $\alpha \in (0, 1)$, the function $F_\alpha(\bar{r}) = 1/|\bar{r}|^\alpha$ is strongly admissible. (Compare with random graphs with edge probability $1/n^\alpha$ where n is the number of vertices.)

Asymptotic reduction of $(\mathbf{coPLAN}^+, \mathbf{coPLA}^+)$

Let

$$(\mathbf{coPLAN}^+, \mathbf{coPLA}^+) \text{ respectively } (\mathbf{coPLAN}^+, \mathbf{ffPLA})$$

denote the inference framework which consists of all $(\mathbb{P}, L)$ where $\mathbb{P}$ is induced by a $coPLA^+$-network and $L = coPLA^+$ respectively $L = ffPLA$.

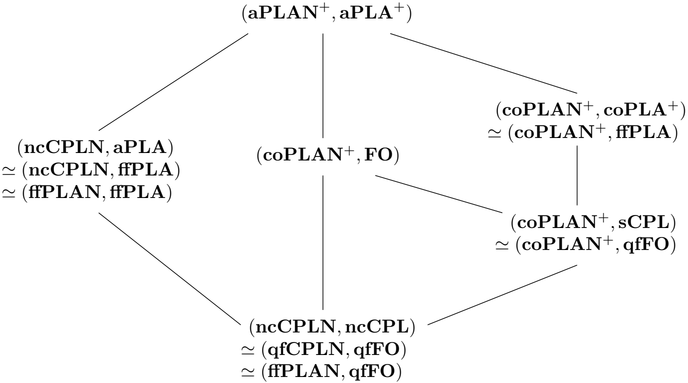
Theorem. [KW2] For every $(\mathbb{P}, L) \in (\mathbf{coPLAN}^+, \mathbf{coPLA}^+)$ and every $\varphi(\bar{x}) \in L$ there is a function-free $\psi(\bar{x}) \in L$ such that $\varphi(\bar{x}) \sim_{\mathbb{P}} \psi(\bar{x})$.

Corollary. Every $(\mathbb{P}, L) \in (\mathbf{coPLAN}^+, \mathbf{coPLA}^+)$ has a convergence law.

Corollary. $(\mathbf{coPLAN}^+, \mathbf{coPLA}^+) \simeq (\mathbf{coPLAN}^+, \mathbf{ffPLA})$.

A larger picture (see [KW2])

A path upwards means that the the upper inference framework is asymptotically more expressible (i.e. $\prec$ holds). The absence of a path upwards between two inference framewords means that the inference frameworks are incomparable with respect to $\preccurlyeq$.



References

[K] Koponen, Conditional probability logic, lifted Bayesian networks, and almost sure quantifier elimination, *Theoretical Computer Science*, Vol. 848 (2020) 1–27.

[KW1] Koponen, Weitkämper, Asymptotic elimination of partially continuous aggregation functions in directed graphical models. Available on arxiv.

[KW2] Koponen, Weitkämper, On the relative asymptotic expressivty of inference frameworks, to be published on arxiv soon.