Proof of Theorem 1

For the proof we will use the classical model-theoretic notion of an *extension axiom*; see for instance Chapter 4 of (Ebbinghaus and Flum 2006) or (Keisler and Lotfallah 2009) for an introduction.

Definitions A *complete atomic diagram in* n *variables* is a quantifier-free formula $\varphi(x_1, \ldots x_n)$ such that for every atom $R(\vec{x})$ with $\vec{x} \subset \{x_1, \ldots x_n\}$ either $\varphi(x_1, \ldots x_n) \models R(\vec{x})$ or $\varphi(x_1, \ldots x_n) \models \neg R(\vec{x})$.

An *r+1-extension-axiom* for a language L is a sentence $\chi_\Phi :=$

$$\forall_{x_1,\ldots,x_r} \left(\bigwedge_{1 \le i < j \le r} x_i \neq x_j \to \exists_{x_{r+1}} \left(\bigwedge_{1 \le i \le r} x_i \neq x_{r+1} \wedge \bigwedge_{\varphi \in \Phi} \varphi \wedge \bigwedge_{\varphi \in \Delta_{r+1} \setminus \Phi} \neg \varphi \right) \right)$$

where $r \in \mathbb{N}$ and Φ is a subset of

$$\Delta_{r+1} := \{R(\vec{x}) | R \in \mathcal{R},\ \vec{x} \text{ a tuple from } \{x_1, \ldots, x_{r+1}\} \text{ containing } x_{r+1}\}.$$

Fact The set of extension axioms is a complete first-order theory that eliminates quantifiers.

Lemma 1 Let the distribution $(\mathbb{P}_{\vec{n}})$ of $\{R_1, \ldots, R_l\}$ be asymptotically equivalent to the reduct of a distribution on $\{R_1, \ldots, R_l, P_1, \ldots, P_m\}$ where $P_1, \ldots, P_m$ are independently distributed with probabilities $p_1, \ldots, p_m$ and $R_1, \ldots, R_l$ are defined to be Boolean combinations of $P_1, \ldots, P_m$ that are neither contradictory nor tautologies. Furthermore, assume that the following hold for $(\mathbb{P}_{\vec{n}})$:
1. For all $0 \le q < l$ and all different tuples $\vec{a}$ and $\vec{b}$, $R_{q+1}(\vec{a})$ is conditionally $\mathbb{P}_{\vec{n}}$-independent of $R_{q+1}(\vec{b})$ given the interpretation of $R_1, \ldots R_q$.
2. The probabilities $\mathbb{P}_{\vec{n}}(R_i(\vec{a}))$ do not depend on $\vec{a}$.

Then $\{R_1, \ldots, R_l\}$ satisfies all extension axioms.

Proof of Lemma 1 For the sake of simplifying notation, we will only consider the single-sorted case. For the multi-sorted case, the calculation is analogous. By induction on l.

$l = 1$: Choose an arbitrary $r + 1$ extension axiom corresponding to a set Φ. Let $\varphi_\Phi(\vec{x}, y)$ be the quantifier-free formula in $r + 1$ variables expressing that y is a witness to the extension axiom for $(\vec{x})$. Let $\vec{a}$ be arbitrarily chosen. Then by asymptotic equivalence, for sufficiently large $\vec{n}$, $\mathbb{P}_n(\varphi_\Phi(\vec{a}, y)) > \delta$ for a $\delta > 0$. By conditions 1 and 2, we can conclude that

$$(\mathbb{P}_n)(\neg\chi) \le n^r (1 - \delta)^{n-r})$$

which limits to 0 as n approaches ∞.

$l \to l+1$: Choose an Choose an arbitrary $r+1$ extension axiom corresponding to a set Δ. Let $\varphi_\Phi(\vec{x}, y)$ be the quantifier-free formula in $r+1$ variables expressing that y is a witness to the extension axiom for $(\vec{x})$. By the induction hypothesis, all extension axioms hold for $R_1, \ldots, R_l$. Therefore, for any natural number k, there are $(\mathbb{P}_n)$-almost-everywhere more than k witnesses of $\varphi_{\Phi\cap\{R_1,\ldots,R_l\}}(\vec{a}, y)$ for every $\vec{a}$, for arbitrary k. Just as in the base case above, we can now apply asymptotic equivalence and conditions 1 and 2 to conclude that conditioned on there being at least k witnesses $\varphi_{\Phi\cap\{R_1,\ldots,R_l\}}(\vec{a}, y)$,

$$(\mathbb{P}_n)(\neg\chi) \leq k^r(1-\delta)^{k-r})$$

which limits to 0 as k approaches ∞. □

Lemma 2 Let ϕ be a complete atomic diagram in r variables. Then the relative frequency of ϕ with respect to an independent distribution is almost surely convergent to $p_\phi \in (0,1)$. This also holds when conditioning on a complete atomic diagram in $m < r$ variables.

Proof of Lemma 2 Induction on r.

In the case $r = 1$, $\phi(x)$ is asymptotically equivalent to a sequence of independent random variables, and the strong law of large numbers gives the result.

So assume true for r. Then $\phi(\vec{x}, y)$ is given by $\phi_r(\vec{x}) \wedge \phi'(\vec{x}, y)$, where $\phi'(\vec{x}, y)$ is a sequence of independent random variables and $\phi_r(\vec{x})$ is a complete atomic diagram in r variables. We can conclude with the strong law of large numbers again. □

Proof of Theorem 1 We perform a parallel induction by height on the following statements:

The family of distributions $(\mathbb{P}_{\vec{n},T})$ is asymptotically equivalent to a quantifier-free LBN, in which all aggregation formulas $\chi_{R,i}$ for which neither $\forall_x(\chi_{R,i}(\vec{x}) \to R(\vec{x})$ nor $\forall_x\neg(\chi_{R,i}(\vec{x}) \wedge R(\vec{x}))$ are true $(\mathbb{P}_{\vec{n},T})$-almost-everywhere, have probability $p \in (0,1)$.

Every extension axiom for the language with those relations that have asymptotic probability $p \in (0,1)$ is valid $(\mathbb{P}_{\vec{n},T})$-almost-everywhere. Let T_h be the fragment of the FLBN of height not exceeding h.

Base step: An FLBN of height 0 is an independent distribution, showing the first statement. The extension axioms are well-known to hold almost everywhere in such an independent distribution.

Induction step: Since the extension axioms form a complete theory with quantifier elimination, every $\chi_{R,i}$ is $(\mathbb{P}_{\vec{n},T_h})$-almost-everywhere equivalent to a quantifier-free $\chi'_{R,i}$. Since we can replace those relations that are almost everywhere true for all or no elements by $\top$ and $\bot$ respectively, we can assume no such relations to occur in ϕ'_R. Let $\varphi_1, \ldots, \varphi_m$ list the complete atomic diagrams in the variables $x_1, \ldots x_r$, where r is the arity of R. Then by Lemma 2 above, for every φ_j, the relative frequency of $\chi'_{R,i}$ given φ_j is almost surely convergent to a number $p_{\chi'_{R,i},\varphi_j}$, which lies in $(0,1)$ if and only if $\chi'_{R,i}$ is neither logically implied nor contradictory to φ_j. Therefore, the family of distributions $(\mathbb{P}_{\vec{n},T})$ is asymptotically equivalent to a quantifier-free LBN, which is obtained by listing $\varphi_1, \ldots, \varphi_m$ and annotating them with the probability $f(p_{\chi'_{R,1},\varphi_j}, \ldots, p_{\chi'_{R,n},\varphi_j})$, which is in $(0,1)$ by the assumption on f.

It remains to show that the extension axioms are valid almost surely. We will verify the statement of Lemma 1. By asymptotic equivalence to a quantifier-free LBN and Proposition 1, we can find an independent distribution as required by Lemma 1. The two additional assumptions are clearly satisfied for any FLBN by the Markov condition for the underlying Bayesian network. □

References

Ebbinghaus, H.-D. and Flum, J.: Finite Model Theory: Second enlarged edition. Springer, 2006.

Keisler, H. J. and Lofallah, W. B.: Almost everywhere elimination of probability quantifiers. J. Symbolic Logic 74(4): 1121-1142 (2009).