JÜRGEN LANDES

The Entropy-Limit (Conjecture) for Σ_2-Premisses

Abstract. The application of the maximum entropy principle to determine probabilities on finite domains is well-understood. Its application to infinite domains still lacks a well-studied comprehensive approach. There are two different strategies for applying the maximum entropy principle on first-order predicate languages: (i) applying it to finite sublanguages and taking a limit; (ii) comparing finite entropies of probability functions defined on the language as a whole. The entropy-limit conjecture roughly says that these two strategies result in the same probabilities. While the conjecture is known to hold for monadic languages as well as for premiss sentences containing only existential or only universal quantifiers, its status for premiss sentences of greater quantifier complexity is, in general, unknown. I here show that the first approach fails to provide a sensible answer for some Σ_2-premiss sentences. I discuss implications of this failure for the first strategy and consequences for the entropy-limit conjecture.

Keywords: Inductive logic, Maximum entropy, Objective Bayesianism.

1. Introduction

We can understand *Inductive logic* as a logical approach seeking to model the (un-)certainty one ought (or could) attach to a conclusion of interest (ψ) given (un-)certain premisses $\varphi_1, \ldots, \varphi_k$. Denoting an inductive entailment relation by $|\approx$ we may formalise this by

$$\varphi_1^{X_1}, \ldots, \varphi_k^{X_k} |\approx \psi^Y,$$

where $X_1, \ldots, X_k, Y$ represent (un-)certainties. Inductive logicians then develop inductive entailment relations (e.g., [7]) and aim to determine the uncertainty attaching to the conclusion ψ, Y.

Inductive logic was popularised by Carnap [2–4], his approach today lives on in the Pure Inductive Logic approach [9,10,22]. In this paper, I however pursue an alternative approach due to Jaynes applying the Maximum Entropy Principle. The Maximum Entropy Principle compels rational agents

to adopt a probability function which satisfies all the premisses. While this first step is rather uncontroversial, the principle has a second part which applies, if there is more than one probability function satisfying all the premisses. In this case, the principle compels rational agents to adopt one of these probability functions with maximum entropy (hence the name of the principle). In this paper, I investigate consequences of a accepting this principle.

The application of the Maximum Entropy Principle to finite propositional language is "merely" a matter of computational complexity, well-understood and well-behaved [13,15,16,19–21]. Applications to infinite domains are another kettle of fish altogether, because it is not clear how to explicate the notion of entropy.

I here consider a first-order predicate language $\mathcal{L}$ with infinitely many constant symbols. Intriguing questions arise for the application of the Maximum Entropy Principle. Two, possibly conflicting, ways to apply the Maximum Entropy Principle to predicate languages have been put forward. It is not clear whether the resulting inductive logics agree. Furthermore, it is not clear which approach to prefer, if they do indeed differ.

The *entropy-limit* approach, due to Jeff Paris and his co-workers [1,17,18, 23–26], proceeds as follows: (i) reinterpret the premisses as constraints on the probabilities of sentences of a *finite* predicate language $\mathcal{L}_n$ containing only the first n constants; (ii) determine the function P^n that maximises entropy on this finite language, subject to constraints imposed by the reinterpreted premisses; (iii) draw inductive inferences using the function P^∞ defined by $P^\infty(\theta) := \lim_{n\to\infty} P^n(\theta)$ for sentences θ of $\mathcal{L}$, in case the limit exists.

The *maximal-entropy* approach, due to the objective Bayesian Jon Williamson [27–30], avoids a reinterpretation of premisses: (i) consider probability functions defined on the language $\mathcal{L}$ as a whole; (ii) deem one probability function P to have greater entropy than another function Q, if and only if P has greater entropy than Q on all but finitely many finite sublanguages $\mathcal{L}_n$; (iii) draw inductive inferences using those functions $P^\dagger$, from all the probability functions on $\mathcal{L}$ that satisfy the premisses, that have maximal entropy (i.e., no other function satisfying the premisses has [in this sense] greater entropy).

The entropy-limit approach is constructive and allows for a (often much) simpler calculation of probabilities. The maximal-entropy approach determines probabilities in some cases in which the entropy-limit approach fails to determine probabilities. If these two approaches lead to different inductive probabilities, then this would provide ammunition to opponents of the Maximum Entropy Principle. On the other hand, if it can be shown, that

both approaches agree (where they are both defined), then this would point towards a unique consistent interpretation of the Maximum Entropy Principle on infinite domains providing further support for its application.

The consistency of these applications of the Maximum Entropy Principle to predicate languages with infinitely many constants has recently been conjectured:

Entropy-limit Conjecture.   If P^∞ exists and satisfies the constraints imposed by the premisses, then it is the unique function with maximal entropy from all those that satisfy the premisses, i.e., $P^\dagger = P^\infty$. [30, p. 191]

This paper is concerned with the case that all premisses are known with certainty, $X_1 = \ldots = X_k = 1$. This amounts to having a single certain premiss φ known with certainty, $\varphi^1 = \varphi_1^1 \wedge \ldots \wedge \varphi_k^1$. To simplify the notation the superscripted 1 will be dropped.

There are some results which show that the conjecture is true for relatively simple premisses. For a quantifier-free premiss sentence entropy maximisation reduces to entropy maximisation on a finite domain, where both approaches agree. Two recent results show that the conjecture holds for all satisfiable premiss sentences in $\Sigma_1 \cup \Pi_1$ [14,25]. The conjecture is also known to hold for monadic languages [1].

The situation concerning premiss sentences containing existential and universal quantifiers is less well-understood. We know of satisfiable premiss sentences $\varphi \in \Sigma_2$ such that the entropy maximiser ($P_\varphi^\dagger$) does not exist [24]. We also know of satisfiable premiss sentences $\varphi \in \Pi_2$ such that the entropy limit (P_φ^∞) does not exist [23]. As far as I'm aware, this all that is known, see Table 1 located at the end of the paper for an overview.

If there exists a $\varphi \in \Sigma_2$ premiss sentence for which the entropy maximiser ($P_\varphi^\dagger$) does not exist but the entropy limit (P_φ^∞) does exist, then the Entropy-limit Conjecture would be false.

In this paper, I show in Theorem 1 that for every polyadic language $\mathcal{L}$ there exists a satisfiable sentence $\varphi \in \Sigma_2$ of $\mathcal{L}$ such that $P_\varphi^\infty = P_=$, where $P_=$ is the (uniform) equivocator function defined in Example 2. In Theorem 2, I generalise this result to the existence of such sentences of all quantifier complexities greater than Σ_2. In these cases, the maximal entropy function $P_\varphi^\dagger$ does not exist [24, Section 2.2]. While this looks prima facie like conclusive evidence against the Entropy-limit Conjecture, I conclude that these cases are, despite appearance to the contrary, cases in which the entropy limit is not defined.

Since the Entropy-limit Conjecture only applies to cases in which the entropy limit is defined, the conjecture is not troubled by my results. The Entropy-limit Conjecture even emerges strengthened, since it applies to fewer cases, and can hence also fail in fewer cases.

2.   The Formal Framework

The formal framework and the notation are taken from [14]. I now reproduce this framework with the simplification that the premiss here is a single sentence known with certainty rather than multiple premisses which may be uncertain.

Predicate languages. Throughout this paper, I consider first-order predicate languages $\mathcal{L}$, with countably many constant symbols $t_1, t_2, \ldots$ and finitely many relation symbols, $U_1, \ldots, U_n$. The atomic sentences, i.e., sentences of the form $U_i t_{i_1} \ldots t_{i_k}$ where k is the arity of the relation U_i, will be denoted by $a_1, a_2, \ldots$, ordered in such a way that atomic sentences involving only constants among $t_1, \ldots, t_n$ occur before those atomic sentences that also involve t_{n+1}. The set of sentences of $\mathcal{L}$ is denoted by $S\mathcal{L}$.

I also consider the finite sublanguages $\mathcal{L}_n$ of $\mathcal{L}$, where $\mathcal{L}_n$ has only the first n constant symbols $t_1, \ldots, t_n$ but the same relation symbols as $\mathcal{L}$. $\mathcal{L}_n$ has finitely many atomic sentences $a_1, \ldots, a_{r_n}$. We call the state descriptions of $\mathcal{L}_n$ (i.e., the sentences of the form $\pm a_1 \wedge \ldots \wedge \pm a_{r_n}$), n-*states*. We let Ω_n be the set of n-states for each n. Note that $|\Omega_n| = 2^{r_n}$, and every n-state $\omega_n \in \Omega_n$ has $|\Omega_{n+1}|/|\Omega_n| = 2^{r_{n+1}-r_n}$ many $n+1$-states, ω_{n+1} which *extend* it (i.e., $\omega_{n+1} \models \omega_n$). Denote the sentences of $\mathcal{L}_n$ by $S\mathcal{L}_n$.

Reinterpretation. I use N to refer to the largest number n such that the constant t_n appears in $\varphi \in S\mathcal{L}$. For a sentence $\varphi \in S\mathcal{L}$ and fixed $n \geq N$, one can reinterpret φ as a sentence of $\mathcal{L}_n$, by substituting every sub formula of the form $\exists x \theta(x)$ by $\theta(t_1) \vee \ldots \vee \theta(t_n)$ and substituting every sub formula of the form $\forall x \theta(x)$ by $\theta(t_1) \wedge \ldots \wedge \theta(t_n)$. To deal with multiple quantifications over the same variable, substitutions begin at the innermost quantifiers and proceed outwards.

I write $(\varphi)_n$, or if there is no ambiguity, simply φ_n, to denote this reinterpretation of φ, which thus becomes a quantifier free sentence of $\mathcal{L}_n$. For any sentence $\varphi \in S\mathcal{L}$ I denote by $[\varphi]_n$ the set of n-states that satisfy φ. The number of n-states in $[\varphi]_n$ is denoted by $|\varphi|_n$.

Example 1. If $\varphi = \forall x \theta(x)$ with quantifier-free $\theta(x)$, then $[\varphi_n]_n = \{\omega \in \Omega_n : \omega \models \bigwedge_{i=1}^n \theta(t_i)\}$. For quantification over l variables with quantifier-free $\theta(\vec{x})$ we have

$$(\forall \vec{x} \theta(\vec{x}))_n = \bigwedge_{t_{i_1}, \ldots, t_{i_k} \in \{t_1, \ldots, t_n\}} \theta(t_{i_1}, \ldots, t_{i_k})$$

$$(\exists \vec{x} \theta(\vec{x}))_n = \bigvee_{t_{i_1}, \ldots, t_{i_k} \in \{t_1, \ldots, t_n\}} \theta(t_{i_1}, \ldots, t_{i_k}).$$

Probability. A *probability function P* on $\mathcal{L}$ is a function $P : S\mathcal{L} \longrightarrow \mathbb{R}_{\geq 0}$ such that:

P1: If τ is a tautology, i.e., $\models \tau$, then $P(\tau) = 1$.

P2: If θ and φ are mutually exclusive, i.e., $\models \neg(\theta \wedge \varphi)$, then $P(\theta \vee \varphi) = P(\theta) + P(\varphi)$.

P3: $P(\exists x \theta(x)) = \sup_m P(\bigvee_{i=1}^m \theta(t_i))$.

A probability function on $\mathcal{L}_n$ is defined similarly (the supremum in P3 is dropped and m is equal to n). We shall use the notation $\mathbb{P}$ and $\mathbb{P}_n$ to denote the set of all probability functions on $\mathcal{L}$ and $\mathcal{L}_n$ respectively.

A probability function is determined by the values it gives to the n-states, for each n. Consequently, a probability function is determined by the values it gives to the quantifier-free sentences, a result known as *Gaifman's Theorem* [5].

Example 2. [Equivocator function] The *equivocator function $P_=$* is defined by:

$$P_=(\omega_n) := \frac{1}{|\Omega_n|} = \frac{1}{2^{r_n}} \text{ for each } n-\text{state } \omega_n \in \Omega_n \text{ and each } n = 1, 2, \ldots.$$

The equivocator function is a probability function on $\mathcal{L}$. It is called the equivocator function because it equivocates between n-states: it is the only probability function that gives each n-state the same probability, for each n. The restriction $P_{=|\mathcal{L}_n}$ of $P_=$ to $\mathcal{L}_n$ is a probability function on $\mathcal{L}_n$, for any n. To simplify notation, I use $P_=$ to refer to these restrictions, as well as to the function on $\mathcal{L}$ itself.

Entropy. The *n-entropy* of a probability function P (which is defined on either $\mathcal{L}$ or $\mathcal{L}_n$) is defined as:

$$H_n(P) := - \sum_{\omega \in \Omega_n} P(\omega) \log P(\omega).$$

The usual conventions are $0 \log 0 := 0$ and the logarithm to be the natural logarithm. The second convention is inconsequential.

The two key approaches are the entropy-limit approach and the maximal-entropy approach.

The entropy-limit approach. For fixed $n \geq N$, reinterpret φ as a statement of $\mathcal{L}_n$. Let $\mathbb{E}_n$ be the set of probability functions on $\mathcal{L}_n$ such that $P(\varphi_n) = 1$. If $\mathbb{E}_n \neq \emptyset$ consider the entropy maximiser:

$$P^n \stackrel{\mathrm{df}}{=} \arg \max_{P \in \mathbb{E}_n} H_n(P).$$

$\mathbb{E}_n$ is closed and convex and P^n is uniquely determined, if φ_n is satisfiable.

However, the premiss is intended as a statement on $\mathcal{L}$, not $\mathcal{L}_n$, and the question arises as to what would be the most appropriate probability function for drawing inferences from this premiss when it is interpreted as a statement about an infinite domain. If it exists, one can consider the function P^∞ defined on $\mathcal{L}$ as a pointwise limit of maximum entropy functions [1]:

$$P^\infty(\theta) := \lim_{n \to \infty} P^n(\theta).$$

The entropy-limit approach takes P^∞ for inference, attaching probability $Y = P^\infty(\psi)$ to sentence ψ.

There is one complication about the definition of P^∞. While [1] define P^∞ in terms of a pointwise limit where the limit is taken independently for each sentence of $\mathcal{L}$, [18,23,25] define P^∞ in a slightly different way: take the pointwise limit on quantifier-free sentences and extend this to the (unique) probability function on $\mathcal{L}$ as a whole which agrees with the values obtained on the quantifier-free sentences, assuming that the pointwise limit exists and satisfies the axioms of probability on quantifier-free sentences of $\mathcal{L}$ [5]. The Rad-Paris definition [1] circumvents a problem that can arise with the Barnett-Paris definition [18,23,25], namely that the pointwise limit on $\mathcal{L}$ as whole may exist but may fail to be a probability function. This detail will be a crucial for interpreting the main result of this paper (Theorem 1).

Note that P^n and $\mathbb{E}_n$ are defined on $\mathcal{L}_n$ not $\mathcal{L}$. To simplify notation, when P is defined on $\mathcal{L}$, I say $P \in \mathbb{E}_n$ meaning that the restriction of P to $\mathcal{L}_n$ is in $\mathbb{E}_n$, $P_{|\mathcal{L}_n} \in \mathbb{E}_n$.

The maximal-entropy approach. This alternative approach avoids appealing to the finite sublanguages. For probability functions P and Q defined on $\mathcal{L}$, P is deemed to have greater entropy than function Q, if and only if it has greater n-entropy for sufficiently large n, i.e., if and only if there is

some natural number N such that for all $n \geq N$ it holds that $H_n(P) > H_n(Q)$. Then one can consider the set of probability functions in $\mathbb{E}$, the set of probability functions on $\mathcal{L}$ with $P(\varphi) = 1$, with maximal entropy:

maxent$\mathbb{E} := \{P \in \mathbb{E} : \text{ there is no } Q \in \mathbb{E} \text{ that has greater entropy than } P\}$.

If maxent$\mathbb{E} \neq \emptyset$, one can draw inferences using the maximal entropy functions $P^\dagger$.[1] Thus, the maximal-entropy approach attaches the set of probabilities $Y = \{P^\dagger(\psi) : P^\dagger \in \text{maxent}\mathbb{E}\}$ to ψ. If maxent$\mathbb{E}$ contains only a single element, so does Y and the maximal entropy function is said to exist and be unique; it is then unambiguously denoted by $P^\dagger$.

See [11,12,30] for one kind of justification of this approach.

Arithmetic Hierarchy of Formulae. Δ_0 is the set of sentences of $S\mathcal{L}$ which are logically equivalent to a quantifier free sentence. The level of the hierarchy Σ_n consists of those sentence which are logically equivalent to a sentence of the form $\exists\vec{x}_1\forall\vec{x}_2 \ldots \varphi(\vec{x}_1, \vec{x}_2, \ldots)$ (with n-many quantifier blocks) and Π_n consists of those sentence which are logically equivalent to a sentence of the form $\forall\vec{x}_1\exists\vec{x}_2 \ldots \varphi(\vec{x}_1, \vec{x}_2, \ldots)$. The level Δ_n is the set of sentences which are logically equivalent to a sentence of the form $\exists\vec{x}_1\forall\vec{x}_2 \ldots \varphi(\vec{x}_1, \vec{x}_2, \ldots)$ and to a sentence of the form $\forall\vec{x}_1\exists\vec{x}_2 \ldots \psi(\vec{x}_1, \vec{x}_2, \ldots)$. Let Ψ be one of these classes, I then denote by Ψ^* the set of sentences which are in Ψ but not in a class of lower complexity. In other words, sentences are classified according to their minimal quantifier complexity.

3. The Σ_2^* Case

I now prove and explain the main technical result of this paper.

3.1. Proof of the Σ_2^* Case

This result concerns the most simple Σ_2^* sentence of a polyadic language there is, $\varphi = \exists x\forall y U xy$. φ only mentions two variables, one binary relation symbol, no connective and no constant.

---

[1]Note that the maximal-entropy approach may license inferences with *multiple* probability functions in case maxent$\mathbb{E}$ contains multiple probability functions [28, p. 91]. It is not obvious how this approach is to be implemented in such a case. Fortunately, this complication of maxent$\mathbb{E}$ containing multiple probability functions does not to concern us here.

PROPOSITION 1. *[Σ_2^*] Let $\mathcal{L}$ contain only one binary relation symbol U and $\varphi = \exists x \forall y \, U xy$, then*

$$P_\varphi^\infty = P_=. \tag{1}$$

PROOF. The proof is a rather simple but not short exercise in counting states and their extensions.

There are $2^{(n^2)}$-many n-states, $|\Omega_n| = 2^{(n^2)}$. There are 2^n-many conjunctions of the form $\bigwedge_{i=1}^n \pm U tt_i$, only one of them is a witness for φ_n (satisfying φ_n), $\bigwedge_{i=1}^n U tt_i$. So, there are $(2^n - 1)^n$-many n-states which are not in $[\varphi_n]_n$. Hence, there are $2^{n^2} - (2^n - 1)^n = 2^{n \cdot n} - (2^n - 1)^n$-many n-states in $[\varphi_n]_n$. P^n equivocates on these n-states to which it assigns probability $(2^{n \cdot n} - (2^n - 1)^n)^{-1}$.

Let us now consider an arbitrary k-state $\omega_k \notin [\varphi_k]_k$ and count the number of n-states with $n > k$ which extend ω_k such that $\omega_n \in [\varphi_n]_n$. An n-state extending ω_k is in $[\varphi_n]_n$, if and only if there exists some $c \in \{k+1, k+2, \ldots, n\}$ such that $\bigwedge_{l=1}^n U t_c t_l$. Since all those k-states are assigned the same probability by P^n, we may consider $\omega_k = \bigwedge_{i,j=1}^k \neg U t_i t_j$ to make things more concrete.

There are $k \cdot (n-k)$-many sentences of the form $U t_i t_l$ with $1 \le i \le k$ and $k+1 \le l \le n$ which do not affect whether $\omega_n \in [\varphi_n]_n$ or not (see below for an illustration).

There are $n \cdot (n-k)$-many sentences of the form $U t_i t_l$ with $k+1 \le i \le n$ and $1 \le l \le n$ which do affect whether $\omega_n \in [\varphi_n]_n$ or not. There are hence $2^{n \cdot (n-k)}$-many possible conjunctions which decide whether the extension ω_n of ω_k is in $[\varphi_n]_n$.
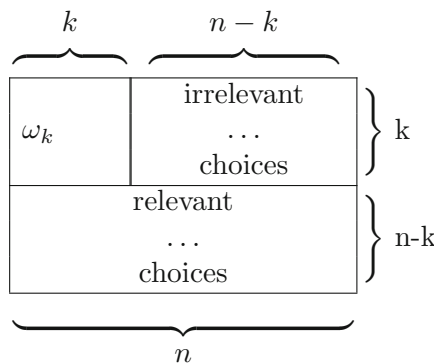
For an extending n-state ω_n not to be in $[\varphi_n]_n$ it must hold that for all $k+1 \le i \le n$ at least one of the literals of the form $\bigwedge_{l=1}^n U t_i t_l$ is negated.

For every fixed i, there are 2^n many conjunctions of the form $\bigwedge_{l=1}^{n} Ut_it_l$, of which $2^n - 1$-many such conjunctions mentioning at least one negation symbol.

There are hence $(2^n - 1)^{n-k}$-many conjunctions of the $2^{n \cdot (n-k)}$-many conjunctions which entail that ω_n is not in $[\varphi_n]_n$. Hence, for all $1 \leq k < n$ and all $\omega_k \notin [\varphi_k]_k$

$$P^n(\omega_k) = P^n(\bigwedge_{i,j=1}^{k} \neg Ut_it_j) = \frac{|\{\omega_n \in [\varphi_n]_n \mid \omega_n \models \omega_k\}|}{|\varphi_n|_n}$$

$$= (2 \overbrace{ n \cdot (n - k)}^{\text{total relevant choices}} - \overbrace{(2^n - 1)^{n-k}}^{\text{no new witness}}) \cdot 2 \overbrace{ k \cdot (n - k)}^{\text{irrelevant choices}} \cdot P^n(\bigwedge_{i,j=1}^{n} Ut_it_j)$$

$$= \frac{(2^{n \cdot (n-k)} - (2^n - 1)^{n-k}) \cdot 2^{k \cdot (n-k)}}{2^{n \cdot n} - (2^n - 1)^n}$$

$$= \frac{2^{(n+k) \cdot (n-k)} - (2^n - 1)^{n-k} \cdot 2^{k \cdot (n-k)}}{2^{n \cdot n} - (2^n - 1)^n}. \tag{2}$$

I now turn to computing the limit as n goes to infinity and find

$$2^{k \cdot k}\left((P^n(\bigwedge_{i,j=1}^{k} \neg Ut_it_j) - \frac{1}{2^{k \cdot k}}\right) = \frac{2^{(n^2)} - 2^{(k^2)} \cdot (2^n - 1)^{n-k} \cdot 2^{kn-k^2}}{2^{n \cdot n} - (2^n - 1)^n} - 1$$

$$= \frac{2^{(n^2)} - (2^n - 1)^{n-k} \cdot 2^{kn}}{2^{n \cdot n} - (2^n - 1)^n} - 1$$

$$= \frac{-(2^n - 1)^{n-k} \cdot 2^{kn} + (2^n - 1)^n}{2^{n \cdot n} - (2^n - 1)^n}$$

$$= \frac{1 - \left(\frac{2^n}{2^n - 1}\right)^k}{\left(\frac{2^n}{2^n - 1}\right)^n - 1}$$

$$= -\frac{\left(\frac{2^n}{2^n - 1}\right)^k - 1}{\left(\frac{2^n}{2^n - 1}\right)^n - 1}.$$

The numerator and the denominator converge to zero and the limit as n approaches infinity cannot simply be read off.

Let s greatest natural number such that $k \cdot s \leq n$, $s := \lfloor \frac{n}{k} \rfloor$. Define a sequence $f_n := \frac{2^n}{2^n - 1}$ and put $\delta_n := (f_n)^k - 1 > 0$. Then

$$\frac{\left(\frac{2^n}{2^n - 1}\right)^k - 1}{\left(\frac{2^n}{2^n - 1}\right)^n - 1} = \frac{\delta_n}{\sqrt[k]{((f_n)^k)^n} - 1}$$

$$= \frac{\delta_n}{\sqrt[k]{(\delta_n + 1)^n} - 1}$$

$$\leq \frac{\delta_n}{\sqrt[k]{(\delta_n + 1)^{s \cdot k}} - 1}$$

$$= \frac{\delta_n}{(\delta_n + 1)^s - 1}$$

$$\leq \frac{\delta_n}{\binom{s}{1}\delta_n + 1 - 1} = \frac{1}{s}.$$

So,

$$2^{k \cdot k}\left((P^n(\bigwedge_{i,j=1}^{k} \neg U t_i t_j) - \frac{1}{2^{k \cdot k}}\right) \geq -\frac{1}{s}. \tag{3}$$

But since $\omega_k = \bigwedge_{i,j=1}^{k} \neg U t_i t_j$ has the fewest n-extensions in $[\varphi_n]_n$, it follows that $P^n(\bigwedge_{i,j=1}^{k} \neg U t_i t_j) \leq \frac{1}{2^{k \cdot k}}$. Thus, the sequence in (3) is always negative. But since this sequence is greater or equal than $-\frac{1}{s}$ and s grows with growing n, this sequence is a null-sequence.

This in turn entails that

$$\lim_{n \to \infty} P^n(\bigwedge_{i,j=1}^{k} \neg U t_i t_j) = \frac{1}{2^{k \cdot k}} = P_=(\bigwedge_{i,j=1}^{k} \neg U t_i t_j).$$

Consider an extension $\omega_n \in [\varphi_n]_n$ of $\bigwedge_{i,j=1}^{k} \neg U t_i t_j$ and denote by λ the conjunction of literals in $[\varphi_n]_n$ which do not appear in $\bigwedge_{i,j=1}^{k} \neg U t_i t_j$, $\omega_n = \lambda \wedge \bigwedge_{i,j=1}^{k} \neg U t_i t_j$. Note that for every other k-state ν_k it holds that $\nu_k \wedge \lambda \in [\varphi_n]_n$. Hence, the set of n-extensions in $[\varphi_n]_n$ of every ν_k is either equal to or a proper superset of the extensions of $\bigwedge_{i,j=1}^{k} \neg U t_i t_j$ in $[\varphi_n]_n$

$$\{\omega_n \in [\varphi_n]_n | \omega_n \models \bigwedge_{i,j=1}^{k} \neg U t_i t_j\} \subseteq \{\omega_n \in [\varphi_n]_n | \omega_n \models \nu_k\}.$$

This means that

$$P^n(\bigwedge_{i,j=1}^{k} \neg U t_i t_j) \leq P^n(\nu_k)$$

for all k, all $\nu_k \in \Omega_k \setminus \{\bigwedge_{i,j=1}^{k} \neg U t_i t_j\}$ and all $n > k$.

Since for all k $\lim_{n \to \infty} P^n(\bigwedge_{i,j=1}^{k} \neg U t_i t_j) = \frac{1}{2^{k \cdot k}}$, all other k-states ν_k and all $\delta > 0$ it holds for all large enough n that $\frac{1}{2^{k \cdot k}} - \delta \leq P^n(\nu_k) \leq$

1 and all n it holds that $\sum_{\omega_n \in \Omega_n} P^n(\omega_n) = 1$, it must be the case that $\lim_{n \to \infty} P^n(\nu_k) \in [\frac{1}{2^{k \cdot k}}, 1)$, if the limit exists.

If a single such limit does not exists, then there must exist an infinite set of natural numbers $I \subseteq \mathbb{N}$ and an $\epsilon > 0$ such that $P^n(\nu_k) \geq \frac{1}{2^{k \cdot k}} + \epsilon$ for all $n \in I$. This however contradicts the fact that the least probable k-state has a probability arbitrarily close to $\frac{1}{2^{k \cdot k}}$ which can only happen, if all other k-states have a probability very close to $\frac{1}{2^{k \cdot k}}$ (probabilities add up to 1), too. Hence, all these other limits exist, too.

Since probabilities add up to 1, it must be the case that for all $\omega_k \in \Omega_k$ it holds that $\lim_{n \to \infty} P^n(\omega_k) = \frac{1}{2^{k \cdot k}} = P_=(\bigwedge_{i,j=1}^{k} \neg U t_i t_j)$, too. This means that for all $k \in \mathbb{N}$ that $\lim_{n \to \infty} P^n$ and $P_=$ agree on all k states. And so,

$$P^\infty = P_=.$$

∎

Mutatis mutandis, this proof goes through for more expressive languages and more complicated premiss sentences, see the Appendix for details:

THEOREM 1. *[More Expressive Languages] For every polyadic language $\mathcal{L}$ the satisfiable sentence $\varphi = \exists x \forall y U x @ r - 1 y \in \Sigma_2^*$ of $\mathcal{L}$ is such that $P_\varphi^\infty = P_=$.*

THEOREM 2. *[Beyond a Σ_2 Premiss] For every polyadic language $\mathcal{L}$ containing at least two relation symbols and every set of sentences $\Lambda^* \in \Psi^*$ which is one of $\Sigma_i^*, \Pi_i^*, \Delta_i^*$, there exists a contingent sentence $\chi = \psi \wedge \exists x \forall y U_1 x @ r - 1 y$ in Λ^* such that $P_\chi^\infty = P_=$.*

3.2. Explanation of the Σ_2^* Case

Since Proposition 1 is the main result of this paper and the proof does not transmit much understanding, I now explain why Proposition 1 holds.

To help our understanding I find it helpful to think of a k-state as the *presence* and its extensions as *futures* in which currently contingent facts (literals mentioning constants t_c with $c \geq k + 1$) have been decided. A future is called *possible*, if and only if it is $[\varphi_n]_n$. The name is inspired by the observation that only n-states in $[\varphi_n]_n$ are assigned a strictly positive probability by P^n.

Let us consider the probability of a presence, $P^n(\omega_k)$ for a fixed k-state ω_k for some fixed point in time ($k \in \mathbb{N}$).

This probability is given as the ratio of the number n-states in which φ_n is true and which extend ω_k divided by the number n-states in which φ_n is true (2). That is just the ratio of the number of all *possible* futures over the

number of all possible and all alternative *possible* futures (possible futures of alternative presences).

Let n be very large, then this probability (measured as a ratio of such conjunctions) that there exists some $k < c \leq n$ such that $\bigwedge_{j=1}^{k} Rt_c t_j$ is close to 1. In fact, for large enough n, the probability that there exist many more such c than k is close to 1. But since every constant t such that $\bigwedge_{j=1}^{k} Rtt_j$ has the same probability to satisfy $\bigwedge_{j=1}^{n} Rtt_j$, it follows that a witness for φ is most likely among the constants $t_{k+1}, \ldots, t_n$. In almost all possible futures there will be a witness, which is not yet present.

Denote by λ a conjunction of literals such that $\lambda \wedge \omega_k \in [\varphi_n]_n$, that is a completion of the k-world to a φ_n-world. This may be understood as time evolving from now to the future. With probability approximating 1 (as n goes to infinity), a witness for φ is to be found in the completion λ, the witness is among $t_{k+1}, \ldots, t_n$. For every way the presence could be, in almost all possible futures there will be a witness which is not yet present.

Since probabilities assigned by P^{∞}, if P^{∞} exists, are according to n-extensions in $[\varphi_n]_n$, it follows that ω_k and ν_k have the same probability. The number k and the two k-states ω_k and ν_k were chosen arbitrarily, so for all fixed k all k-states are assigned the same probability. This means that $P^{\infty} = P_{=}$. Since the probability of a presence is approximated by the ratio of possible futures over the number of all ways the future could have been, all presences are equally likely.

To make a long story very short: all presences have [approximately] the same number of possible futures, which are all equally likely. Hence, all presences are [approximately] equally likely.

4. Conclusions

4.1. The Entropy-Limit Conjecture

Recall that Theorem 1 shows that for every polyadic language $\mathcal{L}$ there exists a satisfiable sentence $\varphi = \exists x \forall y U x @ r - 1 y \in \Sigma_2$ of $\mathcal{L}$ such that $P_{\varphi}^{\infty} = P_{=}$. Prima facie, this seems like evidence against the Entropy-limit Conjecture because it provides a plethora of cases, in which the entropy limit is seemingly well-defined and the maximal entropy function does not exists (easy corollary from [24, Section 2.2]).

Let us take a closer look. $P_\varphi^\infty = P_=$ entails that $P_\varphi^\infty(\varphi) = P_=(\varphi)$ but observe that

$$
\begin{aligned}
P_=(\varphi) &= \lim_{n\to\infty} \lim_{k\to\infty} P_=(\bigvee_{i=1}^n \bigwedge_{j=1}^k Ut_i@r - 1t_j) \\
&\leq \lim_{n\to\infty} \lim_{k\to\infty} \sum_{i=1}^n P_=(\bigwedge_{j=1}^k Ut_i@r - 1t_j) \\
&= \lim_{n\to\infty} \sum_{i=1}^n \lim_{k\to\infty} P_=(\bigwedge_{j=1}^k Ut_i@r - 1t_j) \\
&= \lim_{n\to\infty} \sum_{i=1}^n 0 \\
&= 0.
\end{aligned}
$$

So, the limit of the P_φ^n assigns the certain premiss φ probability 0, $\lim_{n\to\infty} P_\varphi^n(\varphi) = P_\varphi^\infty(\varphi) = 0$. This is absurd. The entropy limit was designed to inform our inductive inferences on the basis that we are sure about the premiss φ. How can the rational probability of φ be zero? It cannot!

It was, presumably, for these reasons that the entropy limit was originally only defined, if $P_\varphi^\infty(\varphi) = 1$. The definition evolved in [18,23,25] into defining the entropy limit by first considering the limits of $P^n(\omega_k)$ for all k and all $\omega_k \in \Omega_k$ and *then* define the probability of sentences containing quantifiers via P3 (Gaifman's Theorem). In previous work, both definitions of the entropy limit agreed. Hence, the entropy limit always assigned probability 1 to certain premisses. I speculate that this is caused by previous work (with the exception of [23, Section 3.2]) not considering premiss sentences mentioning a universal and an existential quantifier.

It might somehow be argued, that it is in (exceptional) cases acceptable to have an entropy limit which assigns an uncertain premiss, φ^X, a non-extreme probability outside of X in the open unit interval. For example, if X is an open interval, it might not appear criminal for P^∞ to assign φ the infimum or supremum of X. The case here is Different. We are given a certain premiss φ and yet the "entropy limit" assigns it probability 0.

Let me briefly consider premiss sentences of greater quantifier complexity. As Theorem 2 shows, we again have cases with $P_\chi^\infty = P_=(\chi)$. Again we see that that $P_\chi^\infty(\chi) = P_=(\chi) = P_=(\psi \wedge \varphi) \leq P_=(\varphi) = 0$. This shows that for all greater quantifier complexities there exists a premiss sentence χ such that $P_\chi^\infty(\chi) = 0$, too.

So, according to the original definition of the entropy limit, the entropy limit is not defined for the cases in Theorems 1 and 2. A sensible interpretation of the later definition of the entropy limit ought to come to the same conclusion—and I hope you agree with my assessment.

This then entails that these cases do not trouble the Entropy-limit Conjecture (which was stated for the original definition of the entropy limit). Rather than troubling the Entropy-limit Conjecture, discovering cases in which the entropy limit is not defined *strengthens* the case for the Entropy-limit Conjecture.

But what about the maximal entropy functions $P_\varphi^\dagger$ and $P_\chi^\dagger$? If they are well-defined, then $P_\varphi^\dagger(\varphi) = 1 = P_\chi^\dagger(\chi)$ by the definition of a maximum entropy function which must satisfy the constraints imposed by the evidence. The absurd situation of the entropy limit ($P_\varphi^\infty(\varphi) = 0 = P_\chi^\infty(\chi)$) cannot arise in the maximal-entropy approach. Unfortunately, they are not well-defined in these cases [24, Section 2.2], and both entropy maximisation methods do not provide sensible beliefs for these premises.

Developing a comprehensive approach to the Maximum Entropy Principle yielding sensible beliefs for these premises, too, is pressing future work.

4.2.   The Entropy Limit

What have we learned about the entropy limit? Its definition only makes sense as long as the probability of the premiss sentence is computed from a *single* limit of probabilities of n-states.

This is, obviously, the case for Π_1, Σ_1 sentences (Gaifman) and for unary languages (every unary sentence is logically equivalent to a disjunction of mutually exclusive Π_1 and Σ_1 sentences ([6, Theorem 35, pp. 68], restated in [1, Lemma 2.2, p. 89], and so the probability is equal to the sum of the probabilities of the disjuncts). Furthermore, the recipe of using the entropy limit to (more simply) compute the maximum entropy function only seems to produce palatable results in these cases.

In case the premiss is of greater quantifier complexity, the entropy limit may fail to provide a (sensible) answer: for $\varphi_2 \in \Pi_2^*$ and beyond the entropy limit may fail to exist [see (6) below], for $\varphi \in \Sigma_2^* \cup \Psi^*$ the entropy limit may fail to provide a sensible answer (Theorems 1 and 2).

Finally, note that the stated assumptions of Theorems 1 and 2 are all crucial (Table 1):

1. For every purely unary language $\mathcal{L}$ and every satisfiable sentence φ of $\mathcal{L}$ it holds that $P_\varphi^\infty(\cdot) = P_\varphi^\dagger(\cdot)$ and $P_\varphi^\infty(\varphi) = 1$ [1, Theorem 2.4] and [23, Theorem 29].

Table 1. Summary of entropy maximisers for categorical premisses for polyadic languages

Premiss	$P^\dagger$	P^∞	Entropy-limit conjecture
$\varphi \in \Delta_0$	Exists and is unique	exists	✓
$\varphi \in \Sigma_1$	Eexists and is unique (2)	Exists (2)	✓ (2)
$\varphi \in \Pi_1$ with $P_=(\varphi) > 0$	Exists and is unique (3)	Exists (3)	✓ (3)
$\varphi \in \Pi_1$ with $P_=(\varphi) = 0$	Exists and is unique, if P^∞ Exists (4)	?	✓ (4)
$\varphi \in \Sigma_2$	not always well-defined (5)	Theorem 1	?
$\varphi \in \Pi_2$	[14, Example 46]	Not always well-defined (5)	?
$\varphi \in \Psi^*$	(5)	Theorem 2	?

Since P^∞ is defined as the limit of a sequence, its existence entails its uniqueness. While we do not know the general answer for the Π_2^* case for $P^\dagger$, it follows from [14, Example 46], that for all polyadic languages and $\varphi = \forall x \exists y U x @ r - 1 y$ it holds that $P_\varphi^\dagger = P_= = P_\varphi^\infty$

2. For every language $\mathcal{L}$ and every satisfiable sentence $\varphi \in \Sigma_1$ of $\mathcal{L}$ it holds that $P_\varphi^\infty(\cdot) = P_=(\cdot|\varphi) = P_\varphi^\dagger(\cdot)$ and $P_\varphi^\infty(\varphi) = 1$ [25, Theorem 4].

3. For every language $\mathcal{L}$ and every satisfiable sentence $\varphi \in \Pi_1$ of $\mathcal{L}$ such that $P_=(\varphi) > 0$ it holds that $P_\varphi^\infty(\cdot) = P_=(\cdot|\varphi) = P_\varphi^\dagger(\cdot)$ and $P_\varphi^\infty(\varphi) = 1$ [14, Theorem 12].

4. For every language $\mathcal{L}$ and every satisfiable sentence $\varphi \in \Pi_1$ of $\mathcal{L}$ such that $P_=(\varphi) = 0$: if P_φ^∞ exists, then $P_\varphi^\infty(\cdot) = P_\varphi^\dagger(\cdot)$ and $P_\varphi^\infty(\varphi) = 1$ [14, Theorem 14].

5. For every polyadic language $\mathcal{L}$ there exists a contingent sentence $\varphi = \exists x \forall y U_1 x @ r - 1y \in \Sigma_2$ of $\mathcal{L}$ such that $P_\varphi^\dagger$ does not exist. This is a simple corollary from [24, Section 2.2]. Similarly to Theorem 2, this fact generalises to Ψ^* by considering a suitably contingent $\chi = \psi \wedge \varphi$ with $P_\psi^\dagger = P_=$ (details omitted since out of scope of current interests), where ψ does not mention the relation symbols in φ.

6. For every polyadic language $\mathcal{L}$ with at least three relation symbols U_1, U_2, U_3 such that the arity of U_i is greater or equal than i there exists a contingent sentence $\varphi \in \Pi_2^*$ of $\mathcal{L}$ such that P_φ^∞ does not exist. This is a simple corollary from [23, Section 3.2]. Again, this fact generalises to Ψ^*.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A. Beyond a Single Binary Relation Symbol and Beyond Σ_2^*

The above focussed on a language with a single binary relation symbol. I now show how the above generalises to sentences of greater quantifier complexity and to all other languages containing at least one relation symbols that is not unary.

PROPOSITION 2. *[Greater Arities]* *Let $\mathcal{L}$ contain only one relation symbol U, which as an arity $r \geq 2$ and denote by $x@r-1$ the $r-1$ tuple with all elements equal to x. For $\varphi = \exists x \forall y U x @ r - 1 y$ it holds that*

$$P_\varphi^\infty = P_=. \tag{4}$$

PROOF. This follows immediately from the proof of Proposition 1 by only counting extensions of the form $\bigwedge_{i,j=1}^n \pm U t_i @ r - 1 t_j$. ∎

The above results generalise to languages containing further relation symbols via the following well-known fact which follows from the Irrelevance Principle and the Independence Principle [20, Principles 3 and 6, p. 186].

PROPOSITION 3. *[Multiple Relation Symbols] Let ψ be a sentence of the language that does not mention U where the arity of U is at least two. Furthermore, for a state ω of the language let ω_{-U} be conjunction of literals of ω which do not mention U and let ω_U be the conjunction of the literals that do mention U, $\omega = \omega_{-U} \wedge \omega_U$. Then*

$$P_{\psi \wedge \exists x \forall y U x @ r - 1 y}^\infty(\omega) = P_\psi^\infty(\omega_{-U}) \cdot P_{\exists x \forall y U x @ r - 1 y}^\infty(\omega_U),$$

if P_ψ^∞ exists.

We can now formulate the results into a global lesson. Let us call a language $\mathcal{L}$ *polyadic*, if and only if $\mathcal{L}$ contains at least one non-unary relation symbol. Let us call the other languages *purely unary*.

THEOREM 3. *[More Expressive Languages] For every polyadic language $\mathcal{L}$ the satisfiable sentence $\varphi = \exists x \forall y U x @ r - 1 y \in \Sigma_2^*$ of $\mathcal{L}$ is such that $P_\varphi^\infty = P_=$.*

PROOF. If $\mathcal{L}$ contains multiple relations symbols simply apply Proposition 3, where ψ is any tautology that does not mention U.

If $\mathcal{L}$ only contains one relation symbol, then this has already been shown in Proposition 2. ∎

DEFINITION 1. Denote by Ψ^* the set of all sentences which are at least Δ_3^* in the hierarchy, $\Psi^* := \bigcup_{i=3}^\infty \Delta_i^* \cup \bigcup_{i=3}^\infty \Sigma_i^* \cup \bigcup_{i=3}^\infty \Pi_i^*$.

THEOREM 4. *[Beyond a Σ_2 Premiss] For every polyadic language $\mathcal{L}$ containing at least two relation symbols and every set of sentences $\Lambda^* \in \Psi^*$ which is one of $\Sigma_i^*, \Pi_i^*, \Delta_i^*$, there exists a contingent sentence $\chi = \psi \wedge \exists x \forall y U_1 x @ r - 1 y$ in Λ^* such that $P_\chi^\infty = P_=$.*

PROOF. First recall that for the premiss $\exists x U_2 x$ in Σ_1^* the entropy limit (and the entropy maximiser) is $P_=$.

For every fixed level of the hierarchy in $\Lambda^* \in \Psi^*$ pick a contingent sentence ξ such that: i) ξ only mentions the relation symbol U_2 and ii) $\xi \vee \exists x U_2 x$ is contingent and is in Λ^*. Clearly, such sentences ξ exist for all $\Lambda^* \in \Psi^*$.

Next note that $P_{\xi \vee \exists x U_2 x}^\infty = P_=$ since $P_{\xi \vee \exists x U_2 x}^n$ assigns at most one n-state probability 0. All other n-states, for fixed n, are assigned the same probability by $P_{\xi \vee \exists x U_2 x}^n$. Let us put $\psi := \xi \vee \exists x U_2 x$.

Let us next consider the sentence $\chi := (\xi \vee \exists x U_2 x) \wedge \exists x \forall y U_1 x @ r - 1 y$ and note that it is in the same level of the hierarchy as ξ. We can simply absorb the two last quantifiers in χ into the quantifier blocks in ξ due to the assumption contains at least three alternating quantifier (blocks), and hence at least once an existential quantifier is followed by a universal quantifier in ξ.

It is well-known that given two premiss sentences (here ψ and $\varphi := \exists x \forall y U_1 x @ r - 1 y$), which do not have a relation symbol in common, the entropy maximisers of $\psi \wedge \varphi$ on $\mathcal{L}_n$ are the product of the entropy maximisers of ψ and φ. For all $n \in \mathbb{N}$ and all $\omega \in \Omega_n$ it holds that $P_{\psi \wedge \varphi}^n(\omega) = P_\psi^n(\omega_{-U_1}) \cdot P_\varphi^n(\omega_{U_1})$. Hence, $P_{\psi \wedge \varphi}^\infty(\omega) = P_\psi^\infty(\omega_{-U_1}) \cdot P_\varphi^\infty(\omega_{U_1})$, if the entropy limits exist.

In our case we have $P_\psi^\infty = P_= = P_\varphi^\infty$, the entropy limits thus exist. This entails that for all n and all n-states $\omega \in \Omega_n$ that

$$\begin{aligned} P_\chi^\infty(\omega) &= P_\psi^\infty(\omega_{-U_1}) \cdot P_\varphi^\infty(\omega_{U_1}) \\ &= P_=(\omega_{-U_1}) \cdot P_=(\omega_{U_1}) \\ &= P_=(\omega). \end{aligned}$$

Since $P_\chi^\infty(\omega)$ and $P_=$ agree on all states it follows that these two probability functions have to also agree on all sentences containing quantifiers (Gaifman's Theorem). So, $P_\chi^\infty = P_=$. ∎

References

[1] BARNETT, O., and J. B. PARIS, Maximum Entropy Inference with Quantified Knowledge. *Logic Journal of IGPL* 16 (1): 85–98, 2008. https://doi.org/10.1093/jigpal/jzm028.

[2] CARNAP, R., The Two Concepts of Probability: The Problem of Probability. *Philosophy and Phenomenological Research* 5 (4): 513–532, 1945. https://doi.org/10.2307/2102817.

[3] CARNAP, R., *Studies in Inductive Logic and Probability*, volume 1. University of California Press, 1971.

[4] CARNAP, R., A basic system of inductive logic part 2. In R. Carnap and R. C. Jeffrey, (eds.), *Studies in inductive logic and probability*, volume 2, University of California Press, Berkeley CA, 1980, pp. 7–155.

[5] GAIFMAN, H., Concerning measures in first order calculi. *Israel Journal of Mathematics* 2 (1): 1–18, 1964. https://doi.org/10.1007/BF02759729.

[6] GROVE, A. J., J. Y. HALPERN, and D. KOLLER, Random worlds and maximum entropy. *Journal of Artificial Intelligence Research* 2 (1): 33–88, 1994.

[7] HAENNI, R., J.-W. ROMEIJN, G. WHEELER, and J. WILLIAMSON, *Probabilistic Argumentation*, volume 350 of *Synthese Library*. Springer, 2011. https://doi.org/10.1007/978-94-007-0008-6_3.

[8] JAYNES, E. T., *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, 2003.

[9] LANDES, J., J. B. PARIS, and A. VENCOVSKÁ, Some Aspects of Polyadic Inductive Logic. *Studia Logica* 90 (1): 3–16, 2008. https://doi.org/10.1007/s11225-008-9140-7.

[10] LANDES, J., J. B. PARIS, and A. VENCOVSKÁ, A survey of some recent results on Spectrum Exchangeability in Polyadic Inductive Logic. *Synthese* 181: 19–47, 2011. https://doi.org/10.1007/s11229-009-9711-9.

[11] LANDES, J., and J. WILLIAMSON, Objective Bayesianism and the maximum entropy principle. *Entropy* 15 (9): 3528–3591, 2013. https://doi.org/10.3390/e15093528.

[12] LANDES, J., and J. WILLIAMSON, Justifying objective Bayesianism on predicate languages. *Entropy* 17 (4): 2459–2543, 2015. https://doi.org/10.3390/e17042459.

[13] LANDES, J., and J. WILLIAMSON, Objective Bayesian nets from consistent datasets. In A. Giffin and K. H. Knuth, (eds.), *Proceedings of MaxEnt*, volume 1757, pp. 1–8. 2016. https://doi.org/10.1063/1.4959048.

[14] LANDES, J., J. WILLIAMSON, and S. R. RAD, The Entropy-Limit Conjecture for non-categorical quantified Constraints. *Annals of Pure and Applied Logic* 2020. 43 pp., accepted.

[15] PARIS, J. B., Common Sense and Maximum Entropy. *Synthese* 117: 75–93, 1998. https://doi.org/10.1023/A:1005081609010.

[16] PARIS, J. B., *The Uncertain Reasoner's Companion: A Mathematical Perspective*, volume 39 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, 2 edition, 2006.

[17] PARIS, J. B., and S. R. RAD, Inference Processes for Quantified Predicate Knowledge. In W. Hodges and R. de Queiroz, (eds.), *Proceedings of WoLLIC*, volume 5110 of *LNCS*, Springer, 2008, pp. 249–259.

[18] PARIS, J. B., and S. R. RAD, *A Note on the Least Informative Model of a Theory*, Springer Berlin Heidelberg, 2010, pp. 342–351. https://doi.org/10.1007/978-3-642-13962-8_38.

[19] PARIS, J. B., and A. VENCOVSKÁ, On the applicability of maximum entropy to inexact reasoning. *International Journal of Approximate Reasoning* 3 (1): 1–34, 1989. https://doi.org/10.1016/0888-613X(89)90012-1.

[20] PARIS, J. B., and A. VENCOVSKÁ, A note on the inevitability of maximum entropy. *International Journal of Approximate Reasoning* 4 (3): 183–223, 1990. https://doi.org/10.1016/0888-613X(90)90020-3.

[21] PARIS, J. B., and A. VENCOVSKÁ, In Defense of the Maximum Entropy Inference Process. *International Journal of Approximate Reasoning* 17 (1): 77–103, 1997. https://doi.org/10.1016/S0888-613X(97)00014-5.

[22] PARIS, J. B., and A. VENCOVSKÁ, *Pure Inductive Logic*. Cambridge University Press, 2015.

[23] RAD, S. R., *Inference Processes for on First Order Probabilistic Languages*. PhD thesis, Manchester Institute for Mathematical Sciences, 2009. http://www.rafieerad.org/manthe.pdf.

[24] RAD, S. R., Equivocation axiom on first order languages. *Studia Logica* 105 (1): 121–152, 2017. https://doi.org/10.1007/s11225-016-9684-x.

[25] RAD, S. R., Maximum Entropy Models for σ_1 Sentences. *Journal of Applied Logics - IfCoLoG Journal of Logics and their Applications* 5 (1): 287–300, 2018. URL http://www.collegepublications.co.uk/admin/download.php?ID=ifcolog00021.

[26] RAD, S. R., Probabilistic models of first order theories: A symmetry axiom to rule them all. 2020. http://www.rafieerad.org/PrCh.pdf. forthcoming.

[27] WILLIAMSON, J., Objective Bayesianism with predicate languages. *Synthese* 163 (3): 341–356, 2008. https://doi.org/10.1007/s11229-007-9298-y.

[28] WILLIAMSON, J., *In Defence of Objective Bayesianism*. Oxford University Press, Oxford, 2010.

[29] WILLIAMSON, J., From Bayesian epistemology to inductive logic. *Journal of Applied Logic* 11: 468–486, 2013. https://doi.org/10.1016/j.jal.2013.03.006.

[30] WILLIAMSON, J., *Lectures on Inductive Logic*. Oxford University Press, Oxford, 2017.

J. LANDES
Ludwig-Maximilians-Universität München
Munich
Germany
juergen_landes@yahoo.de