Rationally, the Universe is Infinite – Maybe*.

Jürgen Landes[1 [0000−0003−3105−6624]]

Department of Philosophy "Piero Martinetti", University of Milan, Italy
juergen_landes@yahoo.de
https://jlandes.wordpress.com/

> Two things are infinite: the universe and human stupidity; and I'm not sure about the universe.

Abstract. This note shows that an explication of rationality within Pure Inductive Logic requires us to believe that the universe has infinitely many elements with probability one. A weaker explication of rationality within Pure Inductive Logic leaves open the possibility and even the necessity of a finite universe.

Keywords: Pure Inductive Logic · Uncertain Inference · Rationality· Induction · Universe.

1 Introduction

"Two things are infinite: the universe and human stupidity; and I'm not sure about the universe." This quote, sometimes attributed to Albert Einstein [2, p. 478], captures (among other things) our desires to learn about the universe and the arising difficulties. Unlike Einstein, who developed physical theories that have been tested empirically, this short paper instead seeks to inform our beliefs about the size of the universe from the armchair relying on our rational faculties.

The Copernican Principle postulating that our place in the world is not special [1]; it has been used to reason about the expected physical size of an alien (most species are expected to exceed 300 kg in body mass) and the size of their home planet(s) [16]. However, the principle and its applications have a less than stellar standing according to some critics. The framework of Pure Inductive Logic instead is a rigorous approach to uncertain inference in the absence of background information [5,14,15].

The rest of this short note is organised as follows: I next introduce the formal framework (§ 2.1), explicate rationality in this framework (§ 2.2) and capture

* I gratefully acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) 405961989 and 432308570

what it means to rationally hold that the universe is infinite. I go on to show that depending on the explication of rationality the universe is rationally infinite (Theorem 2) or not (Theorem 3) and conclude (§ 3). The appendix contains some technicalities.

2 Formal Analysis

There are a number of approaches for drawing uncertain inferences. In this paper, I shall employ Pure Inductive Logic (PIL) [14]. This approach has three main ingredients: firstly, a fixed language of first order logic is chosen to represent the propositions we want to reason about; secondly, all sentences of this language are assigned probabilities representing our rational degrees of belief, our credences. Finally, principles of rationality constrain the choice of probabilities.

2.1 The Framework

Since I'm here interested only in sizes of domains and not in properties of the elements of these domains, I pick the language without relation symbols and without function symbols.

The Language $\mathbb{L}$ of first order logic contains countably many constant symbols $a_1, a_2, \dots$ and countably many variables $x, y, z, \dots$. The language $\mathbb{L}$ does not contain relation symbols nor function symbols but does contain a symbol for equality, $\equiv$. From these symbols the sentences of the language, $S\mathbb{L}$, and the quantifier free sentences of the language, $QFS\mathbb{L}$, can be constructed in the usual way [14, P. 9].

Probability functions w are maps $w : S\mathbb{L} \to [0, 1]$ satisfying three axioms [14, P. 11]:

P1 If $\tau \in S\mathbb{L}$ is a tautology, $\models \tau$, then $w(\tau) = 1$.
P2 If $\tau, \theta \in S\mathbb{L}$ are mutually exclusive, $\tau \models \neg\theta$, then $w(\tau \vee \theta) = w(\tau) + w(\theta)$.
P3 $w(\exists x \theta(x)) = \lim_{n\to\infty} w(\bigvee_{i=1}^{n} \theta(a_i))$.

P1 expresses the thought that sure events have probability 1. P2 is the usual condition concerning the additivity of mutually exclusive events. P3 means that for every element in an underlying domain (or universe) there is at least one constant symbol representing it. We can think of the constants as names for the elements of the underlying domain. Every name is associated with one and only one element; however, elements may have more than one name.

Three axioms of equality are imposed to ensure that the logic behaves as properly with respect to equality. The following types of sentences are added to the tautologies of the logic:

1. $a_i \equiv a_i$ for all constants a_i. (Reflexivity)
2. $a_i \equiv a_k \to a_k \equiv a_i$ for all constants a_i, a_k. (Symmetry)
3. $(a_i \equiv a_k \wedge a_k \equiv a_s) \to a_i \equiv a_s$ for all constants a_i, a_k, a_s. (Transitivity)

In the following, logical consequence is restricted to structures which satisfy these axioms of equality.

The equality symbol needs to behave in the expected way also with respect to assigned probabilities. It is hence required that every probability function w satisfies three further axioms of equality [14, Chapter 37] (probabilities need to respect logical equivalence):

1. $w(a_i \equiv a_i) = 1$ for all constants a_i. (Reflexivity)
2. $w(a_i \equiv a_k \to a_k \equiv a_i) = 1$ for all constants a_i, a_k. (Symmetry)
3. $w((a_i \equiv a_k \wedge a_k \equiv a_s) \to a_i \equiv a_s) = 1$ for all constants a_i, a_k, a_s. (Transitivity)

Gaifman's Theorem states that probability functions are determined by the probabilities they assign to quantifier free sentences [4]. Probability functions are hence uniquely determined by assigning numbers to all quantifier free sentences such that the assignment satisfies P1 and P2.

A table τ (on p) is a complete description of the world containing only $a_1, \ldots, a_p$ [14, P. 276]:

$$\tau = \bigwedge_{i,k=1}^{p} a_i \equiv^{\epsilon_{i,k}} a_k$$

where $\epsilon_{i,k} \in \{0, 1\}$ with $\equiv^1 := \equiv$ and $\equiv^0 := \not\equiv$. We denote by T_p the set of all tables on p that are reflexive, symmetry and transitive.

A simple application of the Disjunctive Normal Form Theorem shows that for all quantifier free sentences, $\varphi \in QFS\mathbb{L}$, such that no constant with a greater index than p is mentioned in φ, it holds that

$$\models \varphi \leftrightarrow \bigvee_{\substack{\tau \in T_p \\ \tau \models \varphi}} \tau \ .$$

Since probability functions assign logically equivalent sentences the same probability [14, Proposition 3.1], it follows that every probability function is completely determined by the probabilities it assigns to tables. Gaifman's Theorem and this observation hence allow us to quickly and simply define probability functions $w : S\mathbb{L} \to [0, 1]$.

2.2 Rationality

Rationality can then be explicated by principles of rationality constraining the choice of a probability function. Traditionally, principles of rationality in Pure Inductive Logic are often rooted in intuitions we have about symmetries, (ir-)relevance and/or induction. I next introduce two principles of rationality.

Constant Exchangeability (CX) is the most widely accepted principle of rationality in PIL. It brands it as irrational to treat constants differently in the absence of all information. Formally, if $a_k \notin \varphi \in S\mathbb{L}$, then $w(\varphi) = w(\varphi(\frac{a_k}{a_i}))$

where $\varphi(\frac{a_k}{a_i})$ is obtained from $\varphi \in S\mathbb{L}$ by replacing all occurrences of the constant a_i in φ by the constant a_k.

Johnson's Sufficientness Principle (JSP) is one of the oldest principle of rationality [7], it even predates [3] by more than a decade. JSP says in the present context that the probability of an unobserved constant, a_{p+1}, being indiscernible from an observed constant, $w(a_{p+1} \equiv a_i)$, only depends on the number of observations, p, and the number of constants which are indiscernible from a_{p+1}. That is, $w(\tau_{p+1}|\tau_p)$ for $\tau_p \in T_p, \tau_{p+1} \in T_{p+1}$ with $\tau_{p+1} \vDash \tau_p$ only depends on p and on $|\{1 \leq k \leq p \ : \ \tau_{p+1} \vDash a_k \equiv a_{p+1}\}|$. In other words, $w(\tau_{p+1}|\tau_p)$ is given by some function $f(p, |\{1 \leq k \leq p \ : \ \tau_{p+1} \vDash a_k \equiv a_{p+1}\}|)$.

The set of probability functions satisfying CX and JSP is given by a 1-parameter family w_λ parametrised by $\lambda \in (0, \infty)$ and two probability functions at the endpoints, $\lambda \in \{0, \infty\}$, see [8, Theorem 21] and [14, Theorem 38.1].

$\lambda = 0$ gives rise to the unique 1-heterogeneous probability function according to which all constants are equal to each other with probability one (Carnap's c_0), $w_0(a_i \equiv a_{i+1}) = 1$ for all i. All names refer to the same element. That is, the universe contains only one element with probability one.

For $\lambda = \infty$ one obtains the completely independent probability function according to which all two pairwise different constants are different with probability one. For every element in the universe there is only one name, $w_\infty(a_i \equiv a_k) = 0$ for all $i \neq k$, and thus the universe contains infinitely many elements with probability one.

In the context of PIL, these two probability functions are peculiar functions which are often considered as unsuitable for explicating inductive intuitions. The first probability function is unsuitably opinionated in the absence of background information. The second probability function does not capture inductive entailment [18, P. 58]. There are however further probability functions satisfying CX and JSP.

Theorem 1 (Theorem 21 of [8]). *All other probability functions w defined on $\mathbb{L}$ satisfying CX and JSP are given by*

$$w_\lambda(\tau) = \lambda^{t-1} \cdot \left(\prod_{i=1}^{p-1} \frac{1}{i+\lambda} \right) \cdot \left(\prod_{j=1}^{t} (x_j - 1)! \right),$$

where $\tau \in T_p$, t is the number of discernible constant symbols in τ (the number of equivalence classes) and the x_j are the size of equivalence classes according to τ.[1]

[1] There is an unfortunate slip-up in [8, Theorem 21] where the exponent of the first factor is t rather than $t-1$. This factor arises from every constant symbol which is different from all the previous ones, with the exception of the very first constant a_1 for which $w(a_1 \equiv a_1) = 1$ needs to hold by the axioms of equality.

In the terminology of PIL, the x_j are the *spectrum* of τ.[2]

To simplify notation let T_p^t and $T_p^{\leq t}$ denote the sets of tables on p which have exactly t equivalence classes, respectively, less or equal than t equivalence classes.

2.3 Rationally, the Universe contains infinitely many Elements

The number of elements in the universe can be measured by how many constant symbols are pairwise indiscernible. However, a table τ on p can only tell us about the first p constant symbols. The number of elements in the universe becomes accessible only when p is sent to infinity. This allows us to state that the universe contains exactly t elements with non-zero probability by:

$$\lim_{p\to\infty} w(T_p^t) := \lim_{p\to\infty} \sum_{\tau\in T_p^t} w(\tau) > 0$$

according to some probability function w.

That the universe contains infinitely many elements with probability zero can then be formalised by increasing the number of indiscernible elements:

$$\lim_{p\to\infty} w(T_p^{\leq t}) := \lim_{p\to\infty} \sum_{\tau\in T_p^{\leq t}} w(\tau) = 0 \quad \text{for all fixed numbers } t \in \mathbb{N} \ .$$

The rational probability that the universe contains only finitely many elements is thus equal to the following expression

$$\lim_{t\to\infty} \lim_{p\to\infty} w_\lambda(T_p^{\leq t}) \ .$$

I'm obliged to Alena Vencovská and Jeff Paris for pointing out the following result proved in the Appendix.

Theorem 2. *For all $\lambda > 0$ it holds that*

$$\lim_{t\to\infty} \lim_{p\to\infty} w_\lambda(T_p^{\leq t}) = 0 \ .$$

[2] Tangentially, we can now also formalise and answer the question asked in many bars "Don't I know you from somewhere?" Formally, the probability of having met is

$$w_\lambda(\bigvee_{i=1}^{p} a_{p+1} \equiv a_i | \tau) = 1 - \frac{w_\lambda(\tau \wedge \bigwedge_{i=1}^{p} a_{p+1} \not\equiv a_i)}{w_\lambda(\tau)}$$
$$= 1 - \frac{\lambda^{t+1}}{\lambda^t} \cdot \frac{1}{p+\lambda} \cdot \frac{1! \cdot \prod_{j=1}^{t}(x_j - 1)!}{\prod_{j=1}^{t}(x_j-1)!} = \frac{p}{p+\lambda},$$

which converges to one for ever greater sample sizes, $\tau \in T_p$ with growing p, for $0 < \lambda < \infty$. As we get older (increasing p), we become more and more sure to have already met the person at the bar. Note that $w_\lambda(\tau) > 0$ for all $p \geq 1$ and all $\tau \in T_p$; the conditional probability considered is hence well-defined. $w_0(\bigvee_{i=1}^{p} a_{p+1} \equiv a_i|\tau) = 1$ and $w_1(\bigvee_{i=1}^{p} a_{p+1} \equiv a_i|\tau) = 0$.

Thus, CX and JSP jointly entail that either there is only one single thing in the universe ($w = w_0$) or the universe contains infinitely many elements ($w = w_\lambda$ with $\lambda > 0$). Having rejected the former answer, we conclude with probability one that there are infinitely many elements in the universe.

2.4 A Weaker Notion of Rationality

The axiom JSP is arguably too strong of a demand to explicate rationality. Could the conditional probability $w(\tau_{p+1}|\tau_p)$ not also depend on, say, the number of equivalence classes in τ_p? The question arises what we ought to believe about the make up of the universe under a weaker construal of rationality, i.e., how much pressure can Theorem 2 bear before breaking down?

Alternatively, one may think that our inductive assumptions ought not to deductively entail that the universe is infinite with probability one. Instead, a finite universe should remain an open possibility.[3] Within the PIL framework one may then be willing to give up on CX, JSP or both these axioms. Since the axiom CX is more widely accepted (and studied) than JSP, I will here drop JSP and keep CX.

Just assuming the axiom CX, [14, Corollary 37.2] demonstrates that for every probability function w satisfying Ex on the language containing only the equality symbol there has to exist an associated probability function w' defined on the language containing only one binary relation symbol but not the equality symbol. Furthermore, w' satisfies the axiom of Spectrum Exchangeability (Sx) [6,9,10,11,12,13] and [14, Chapters 27-38]. This axiom roughly states that probabilities of worlds only depend on how the world distinguishes constants but not on the actual properties of the constants at this world.

Probability functions on this predicate language L (and all other predicate languages with at least one non-unary relation symbol) come in two different shapes. They either satisfy Li with Sx or they don't (Li stands for *L*anguage *I*nvariance). Li with Sx harkens back to Carnap [17, P. 974].

Li with Sx requires that there exists a mutually consistent family of probability functions $w^{\mathcal{L}}$ that all satisfy Sx defined on all predicate languages such that for all predicate languages L, L' it holds that w^L and $w^{L'}$ agree on all sentences φ that are sentences of L and L'. In words, whenever w^L is a probability function on L satisfying Sx and L' is a language containing all the relation symbols of L and further relation symbols, then there exists a probability function on L' that satisfies Sx, $w^{L'}$; $w^{L'}$ agrees with w^L on all the sentences of L.

All probability functions w **satisfying Li with Sx** have a particular form [14, Theorem 37.1]:

$$w^L = \int_{\vec{p}\in\mathbb{B}} u^{\vec{p},L} \, d\mu(\vec{p})$$

[3] Many thanks to Jon Williamson for pointing this out to me.

for some measure μ, the *de Finetti prior*, where the $u^{\vec{p},L}$ are particular probability functions defined on L that satisfy Sx. Technical details do not concern us here; some can be found in the appendix.

Applying [14, Corollaries 37.2 and 37.3] we find that all probability functions w (defined on the language with equality but without relation symbols $\mathbb{L}$) satisfying Ex can be written as:

$$w = \int_{\vec{p}\in\mathbb{B}} u^{\vec{p}}_{\mathrm{Eq}}\, d\mu(\vec{p}) \ .$$

Again, many details shall not be relevant here, see [14, § 37] for details and the appendix for main points.

Relevantly, since w is a "convex mixture" of $u^{\vec{p}}_{\mathrm{Eq}}$ (w is an integral), the following are logically equivalent:

1. $\lim_{t\to\infty}\lim_{p\to\infty} w(T_p^{\leq t}) = 0$ and
2. the measure $\mu(\vec{p})$ puts all mass on probability functions which satisfy $\lim_{t\to\infty}\lim_{p\to\infty} u^{\vec{p}}_{\mathrm{Eq}}(T_p^{\leq t}) = 0$.

The latter condition is known to depend only on the infinite sequence of numbers $\vec{p} = \langle p_0, p_1, p_2, \ldots\rangle$. By construction this is a sequence of, not necessarily strictly, positive numbers summing to 1 such that $p_1 \geq p_2 \geq p_3 \ldots$.[4] The following two conditions are equivalent [14, Chapters 29 and 30]:

1. $\lim_{t\to\infty}\lim_{p\to\infty} u^{\vec{p}}_{\mathrm{Eq}}(T_p^{\leq t}) = 0$ and
2. $p_0 > 0$ or there are infinitely many $p_i > 0$.

Letting

$$\mathbb{B}_\infty := \{\vec{p} \in \mathbb{B} \mid p_0 > 0 \text{ or there are infinitely many } p_i > 0\} \ ,$$

I can now compactly state main result of this section.

Theorem 3. *For all probability functions w on $S\mathbb{L}$ satisfying Ex that are associated with a family of probability functions defined predicate languages satisfying Li with Sx it holds that*

$$w = \int_{\vec{p}\in\mathbb{B}} u^{\vec{p}}_{Eq}\, d\mu(\vec{p})$$

for some measure $\mu(\vec{p})$. Furthermore, the following three conditions are equivalent

1. $\lim_{t\to\infty}\lim_{p\to\infty} w(T_p^{\leq t}) = 0$,
2. $\mu(\mathbb{B}_\infty) = 1$,
3. $w = \int_{\vec{p}\in\mathbb{B}_\infty} u^{\vec{p}}_{Eq} d\mu(\vec{p})$.

[4] Note that there is no requirement that $p_0 \geq p_1$.

As a result, the probability that the universe contains infinitely many elements is equal to the measure the de Finetti prior μ assigns to B_∞, $\int_{p\in B_\infty} d\mu(\vec{p})$. In other words, Theorem 2 can take some heat but breaks down when stretched widely.

In particular, if $\int_{p\in B_\infty} d\mu(\vec{p}) = 0$, then the universe contains infinitely many elements with probability zero. In that case, we are rational to fully believe that the universe contains only finitely many elements.

3 Conclusions

We have seen that the framework of PIL allows us to reason about the (size of the) universe. Depending on how exactly rationality is explicated, PIL gives different answers as to how many different elements we should rationally expect to see in the universe – in the absence of all evidence. Assuming that elements are of equal size (or at least that there is a strictly positive lower bound on their size), believing that there are infinitely many different elements entails a belief in an infinitely large universe.

It is tempting to capture our uncertainty about the size of the universe by the philosophical puzzle of adopting *the* correct de Finetti prior, which has posed a formidable challenge to many great thinkers. A convincing solution has yet to be found.

Reversing our train of thought, this analysis can also be read as a way to inform our thinking about rational probabilities via science, cf. [19] for whether one should in general read $A \to B$ as "A entails B" or as "$\neg B$ entails $\neg A$". Knowing that the universe is infinite and reasonably homogeneous entails that we know of the existence of infinitely many elements. Subsequently, assuming Li with Sx rules out certain de Finetti priors [(science + Li with Sx) constrains the set of rational priors]. However, until science has determined the size of the universe, I suggest to read my argument in the here presented direction and worry about the reverse when that day comes.

Acknowledgements Many thanks to Alena Vencovská and Jeff Paris for pointing out Theorem 2 and further useful comments. Thanks also to Jon Williamson for the suggestion in § 2.4 and further helpful comments.

References

1. Bondi, H.: Cosmology. Cambridge University Press, Cambridge (1952)
2. Calaprice, A. (ed.): The ultimate quotable Einstein. Princeton University Press, Princeton, 2 edn. (2011)
3. Carnap, R.: On Inductive Logic. Philosophy of Science **12**(2), 72–97 (1945), https://doi.org/10.1086/286851
4. Gaifman, H.: Concerning measures in first order calculi. Israel Journal of Mathematics **2**(1), 1–18 (1964), https://doi.org/10.1007/BF02759729

5. Hill, A., Paris, J.B.: An Analogy Principle in Inductive Logic. Annals of Pure and Applied Logic **164**(12), 1293–1321 (2013), https://doi.org/10.1016/j.apal.2013.06.013
6. Howarth, E., Paris, J.B.: The Theory of Spectrum Exchangeability. The Review of Symbolic Logic **8**, 108–130 (3 2015), https://doi.org/10.1017/S1755020314000331
7. Johnson, W.E.: Probability: The Deductive and Inductive Problems. Mind **41**(164), 409–423 (1932)
8. Landes, J.: The Principle of Spectrum Exchangeability within Inductive Logic. Ph.D. thesis, Manchester Institute for Mathematical Sciences (2009), https://jlandes.files.wordpress.com/2015/10/phdthesis.pdf
9. Landes, J., Paris, J.B., Vencovská, A.: Language Invariance and Spectrum Exchangeability in Inductive Logic. In: Mellouli, K. (ed.) Proceedings of ECSQARU. LNAI, vol. 4724, pp. 151–160. Springer (2007), https://doi.org/10.1007/978-3-540-75256-1_16
10. Landes, J., Paris, J.B., Vencovská, A.: Representation Theorems for probability functions satisfying Spectrum Exchangeability in Inductive Logic. International Journal of Approximate Reasoning **51**(1), 35–55 (2009), https://doi.org/10.1016/j.ijar.2009.07.001
11. Landes, J., Paris, J.B., Vencovská, A.: A characterization of the Language Invariant families satisfying Spectrum Exchangeability in Polyadic Inductive Logic. Annals of Pure and Applied Logic - Proceedings of IPM 2007 **161**(6), 800–811 (2010), https://doi.org/10.1016/j.apal.2009.06.010
12. Landes, J., Paris, J.B., Vencovská, A.: 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
13. Nix, C.J., Paris, J.B.: A Note on Binary Inductive Logic. Journal of Philosophical Logic **36**(6), 735–771 (2007), https://doi.org/10.1007/s10992-007-9066-y
14. Paris, J.B., Vencovská, A.: Pure Inductive Logic. Cambridge University Press, Cambridge (2015)
15. Paris, J.B., Vencovská, A.: Six Problems in Pure Inductive Logic. Journal of Philosophical Logic **48**, 731–747 (2019), https://doi.org/10.1007/s10992-018-9492-z
16. Ruhmkorff, S., Jiang, T.: Copernican Reasoning About Intelligent Extraterrestrials: A Reply to Simpson. Journal for General Philosophy of Science **50**(4), 561–571 (2019), https://doi.org/10.1007/s10838-019-09465-7
17. Schlipp, P.A. (ed.): The Philosophy of Rudolf Carnap. Open Court, La Salle (1963)
18. Williamson, J.: Lectures on Inductive Logic. Oxford University Press, Oxford (2017), https://doi.org/10.1093/acprof:oso/9780199666478.001.0001
19. Williamson, J.: One philosopher's modus ponens is another's modus tollens: Pantomemes and nisowir. Metaphilosophy (2022), https://doi.org/10.1111/meta.12546

Appendix

Proof of Theorem 2

Theorem 2. *For all* $\lambda > 0$ *it holds that*

$$\lim_{t\to\infty}\lim_{p\to\infty} w_\lambda(T_p^{\leq t}) = 0 \ .$$

Proof. $\lambda = \infty$ is a trivial case.

Now consider $0 < \lambda < \infty$. Let us first note that the probability that any given table $\tau \in T_p$ with any number of equality classes, the conditional probability (given τ) that no extension by n more constants introduces a new class is

$$g(p,n) := \frac{p}{p+\lambda} \cdot \frac{p+1}{p+1+\lambda} \cdot \ldots \cdot \frac{p+n-1}{p+n-1+\lambda} \ .$$

Since $\lambda > 0$, there exists some $M \in \mathbb{N}$ such that $\lambda \geq \frac{1}{M}$. We thus note

$$g(p,n) \leq \frac{p}{p+\frac{1}{M}} \cdot \frac{p+1}{p+1+\frac{1}{M}} \cdot \ldots \cdot \frac{p+n-1}{p+n-1+\frac{1}{M}} \ .$$

We next observe that for all the factors on the right ($0 \leq i \leq n-1$) it holds that

$$\begin{aligned}\frac{p+i}{p+i+\frac{i}{M}} = \frac{M}{M} \cdot \frac{p+i}{p+i+\frac{1}{M}} &= \frac{pM+iM}{pM+iM+1} \\ &\leq \sqrt[M]{\frac{pM+iM}{pM+iM+1} \cdot \frac{pM+iM+1}{pM+iM+2} \cdot \ldots \cdot \frac{pM+iM+M-1}{pM+iM+M}} \ .\end{aligned}$$

Inserting these inequalities back in, we obtain

$$g(p,n) \leq \sqrt[M]{\frac{pM}{pM+nM}} = \sqrt[M]{\frac{p}{p+n}} \ . \tag{1}$$

Applying this observation to $\tau = a_1 \equiv a_1$ (the only table in T_1), we see that the probability of the first $n+1$ constants are all the same is

$$\frac{1}{1+\lambda} \cdot \frac{2}{2+\lambda} \cdot \ldots \cdot \frac{n}{n+\lambda} \leq \sqrt[M]{\frac{1}{1+n}} < \sqrt[M]{\frac{1}{n}} \ .$$

Hence, the probability of the first $n+1$ constants falling into at least two classes is greater or equal to $1 - \sqrt[M]{\frac{1}{n}}$.

Similarly, for every table $\tau \in T_{n+1}$ the conditional probability that the next n^2 constants do not introduce a new class is

$$\begin{aligned}\frac{n+1}{n+1+\lambda} \cdot \frac{n+2}{n+2+\lambda} \cdot \ldots \cdot \frac{n^2+n}{n^2+n+\lambda} &\overset{(1)}{\leq} \sqrt[M]{\frac{n+1}{n^2+n+1}} \\ &= \sqrt[M]{\frac{n}{n} \cdot \frac{1+\frac{1}{n}}{n+1+\frac{1}{n}}} \\ &\leq \sqrt[M]{\frac{2}{n}} \ .\end{aligned}$$

With ever greater n this expression vanishes, too. Thus, overall it holds that $\lim_{p\to\infty} w_\lambda(T_p^{\leq 1}) = 0$.

Continuing in this manner adding p^t constants to a table $\tau \in T_{p^{t-1}+1}$ we have

$$\begin{aligned} g(p^{t-1}+1, p^t) &= \frac{p^{t-1}+1}{p^{t-1}+1+\lambda} \cdot \ldots \cdot \frac{p^{t-1}+p^t}{p^{t-1}+p^t+\lambda} \\ &\overset{(1)}{\leq} \sqrt[M]{\frac{p^{t-1}+1}{p^{t-1}+p^t+1}} \\ &= \sqrt[M]{\frac{1+\frac{1}{p^{t-1}}}{1+p+\frac{1}{p^{t-1}}}} \\ &\leq \sqrt[M]{\frac{2}{p}} \ . \end{aligned}$$

We hence find that for all large enough p greater some fixed $t \in \mathbb{N}$ it holds that $\lim_{p\to\infty} w_\lambda(T_p^{\leq t}) = 0$.

We thus conclude that for all $\lambda > 0$ it holds that

$$\lim_{t\to\infty} \lim_{p\to\infty} w_\lambda(T_p^{\leq t}) = 0 \ .$$

Some further Technicalities

The probability functions $u^{\vec{p},L}$ and $u^{\vec{p}}_{\text{Eq}}$ are generated by drawing balls from an urn. Consider an urn containing a black ball and at most countably-many balls with unique colours (black is not a colour here). We draw balls from the urn with the following probabilities: $p_0 \geq 0$ for the black ball and $p_i \geq 0$ for the colours. We hence need to have that $\sum_{i=0}^{\infty} p_i = 1$. Let $\mathbb{B}$ be the set of all such urns.

Draws are with replacement (balls are put back into the urn after pulling them out) and independent from each other. We assume without loss of generality that the colours are ordered such that $p_i \geq p_{i+1}$ for all $i \geq 1$ (again note the special role of 'black', p_0). Let $c_1, \ldots, c_p \in \mathbb{N}^p$ (to simplify notation $0 \in \mathbb{N}$ is assumed) be a sequence of drawn balls, then the probability of this draw is $\prod_{i=1}^{p} p_{c_i}$.

Given a fixed urn (a non-negative vector $\vec{p}$ of countable length, non-increasing from the second element onwards and summing to one) and a draw of p balls, $c_1, \ldots, c_p$, we then associate a unique table $\tau \in T_p$ with this draw as follows:

$$a_i \equiv a_i \text{ for all } 1 \leq i \leq p$$

and for all $i \neq k$

$$\begin{aligned} a_i &\equiv a_k, \quad \text{if } c_i = c_k \neq 0 \\ a_i &\not\equiv a_k, \quad \text{if } c_i = c_k = 0 \text{ or } c_i \neq c_k \ . \end{aligned}$$

Intuitively, if two balls have the same colour different from black, then the corresponding constants are equal. If the colours are different or at least one of them is black, then the corresponding constants are different.

For a given $\tau \in T_p$ let $C(\tau)$ be the set of draws of p balls that are associated with this table. Then the probability function $u^{\vec{p}}_{\mathrm{Eq}}$ is defined by the values it gives to all tables:

$$u^{\vec{p}}_{\mathrm{Eq}}(\tau) = \sum_{\vec{c} \in C(\tau)} \prod_{i=1}^{p} p_{c_i} \quad .$$

So, $u^{\vec{p}}_{\mathrm{Eq}}(\tau)$ is the joint probability of all draws associated with the table τ.

The probability functions $u^{\vec{p},L}$ are also defined in terms of (i.i.d.) sampling urns from balls and constructing possible worlds associated with draws of balls. The definition is however much more messy, see [14, Chapter 29] for full details.

One key difference is that the drawing of a previously unseen colour or a black ball does not entail that the constant is necessarily distinguishable from the previous constants in the associated worlds. Instead, upon drawing a new colour or a black ball all ways the new constant can behave[5] are equally likely. On the other hand, if a colour is not drawn for the first time, then there is only one way that constant can behave; its has to behave in the exact same way as all constants associated with the same colour.

One important connection between the $u^{\vec{p},L}$ and the $u^{\vec{p}}_{\mathrm{Eq}}$ is the fact that for all fixed $\vec{p}$ the latter is – in some sense – equal to the limit of $u^{\vec{p},L}$ where the limit is over ever larger languages containing finitely many relation symbols, at least one of them is not unary, but no symbol for equality [14, Corollary 37.3], $u^{\vec{p}}_{\mathrm{Eq}}$ "$=$" $\lim_{|L| \to \infty} u^{\vec{p},L}$.

[5] There is one condition here that is best explained by means of an example. Suppose that $c_2 = c_4$ and ($c_6 = 0$ or $c_6 \notin \{c_1, c_2, c_3, c_4, c_5\}$). Then a_6 cannot be such that a_2 and a_4 can be told apart. That is, $Ra_2a_6 \wedge \neg Ra_4a_6$ cannot hold in associated worlds. The condition thus is that indistinguishability due to colours (those that are not due to chance) must be preserved. By contrast, indistinguishability due to chance is broken for ever longer draws of balls with limit probability one, if the urn contains a black ball ($p_0 > 0$) or infinitely many colours. In words, indistinguishability between constants obtains in the limit, if and only if they are associated with the same non-black colour – if the urn contains a black ball ($p_0 > 0$) or if the urn contains infinitely many colours. Compare this with $u^{\vec{p}}_{\mathrm{Eq}}$ according to which constants are equal, if and only if they are associated with the same non-black colour.