13th Panhellenic Logic Symposium

PLS13

Originally July 2021 – postponed to July 2022

Volos, Greece

Preface

The Panhellenic Logic Symposium was established in 1997 as a biennial scientific event. It aims to promote interaction and cross-fertilisation among different areas of logic. Originally conceived as a way of bringing together the many logicians of Hellenic descent throughout the world, the PLS has evolved into an international forum for the communication of state-of-the-art advances in logic.

The 13th Panhellenic Logic Symposium was initially planned to take place in July 2021, in the city of Volos. However, the outbreak of the pandemic and its continuing repercussions have forced us to postpone the event to July 2022, with the hope that, by then, the general situation will allow for an in-person event.

In spite of the singular circumstances, we held a regular submission and reviewing process this year. We received a total of 16 submissions, of which 13 were accepted and 11 appear in this collection; we would like to thank the authors of all submitted papers. The reviewing procedure involved assigning to each submission (at least) two referees, among the members of the Scientific Committee and some external reviewers; we would like to thank our colleagues for their diligent work: Antonis Achilleos, Costas Dimitracopoulos, Pantelis Eleftheriou, Vassilis Gregoriades, Kostas Hatzikiriakou, Antonis Kakas, Alex Kavvos, Nikolaos Papaspyrou, Thanases Pheidas, Rizos Sklinos, Ana Sokolova, Alexandra Soskova, Mariya Soskova, Yannis Stephanou, Nikos Tzevelekos, Niki Vazou, and Stathis Zachos, as well as the external reviewers: Russell Miller, Andre Nies, Frank Stephan, and Jon Williamson.

We collect here, in this volume, the accepted papers of this year's submission round. A next round of submissions will happen in 2022; hence, this volume can be considered as the first installment of the projected comprehensive PLS proceedings, expected to appear finalized after the end of the actual event, in 2022. This first volume, along with a list of accepted papers and links to pre-recorded videos of talk presentations (of those authors who chose to prepare and share with us a first version of their talk this year), appear on the event's webpage.

Finally, we would like to thank our invited speakers and our sponsors, who have expressed their willingness to renew their commitment to our event despite the postponement.

We hope to see you all next summer.

Beijing and Samos, July 2021

Giorgos Barmpalias and Kostas Tsaprounis

List of accepted papers

1. Theofanis Aravanis and Pavlos Peppas. *Types of Rational Horn Revision Operators*
2. Theofanis Aravanis. *Properties of Parametrized-Difference*
3. Riccardo Bruni and Lorenzo Rossi. *Truth meets Vagueness*
4. Matteo de Ceglie. *The V-logic Multiverse and MAXIMIZE*
5. Dimitra Chompitaki, Manos Kamarianakis, and Thanases Pheidas. *Decidability of the theory of addition and the Frobenius map in certain rings of rational functions*
6. Bruno Dinis and Bruno Jacinto. *A Nonstandard view on vagueness*
7. Konstantinos Kartas. *Decidability of local fields and their extensions*
8. Jürgen Landes. *A Triple Uniqueness of the Maximum Entropy Approach*
9. Ivo Pezlar. *The placeholder view of assumptions and the Curry-Howard correspondence*
10. Ivan Titov and Wolfgang Merkle. *A total Solovay reducibility and totalizing of the notion of speedability*
11. Jon Williamson. *Determining maximal entropy functions for objective Bayesian inductive logic*

Types of Rational Horn Revision Operators

Theofanis Aravanis and Pavlos Peppas

Department of Business Administration
School of Economics & Business
University of Patras, Greece
{taravanis,pavlos}@upatras.gr

Abstract

In this article, we identify some interesting types of rational revision operators that implement Horn revision. In particular, we first define (both axiomatically and semantically) a class of Horn revision operators based on proper set inclusion of the atoms satisfied by possible worlds. Furthermore, we show that a well-behaved type of rational revision, called uniform revision, is Horn-compliant. This is demonstrated by proving that concrete Horn revision operators implement particular uniform-revision policies.

1 Introduction

Belief revision (or revision) is the process by which a rational agent changes their beliefs, in the light of new information [7]. A prominent approach that formalizes belief revision is that proposed by Alchourrón, Gärdenfors and Makinson in [1], now known as the *AGM paradigm*. Within the AGM paradigm, the agent's belief corpus is modelled by a logical theory K, also referred to as a *belief set*, new information (alias, *epistemic input*) is represented as a logical sentence φ, and the revision of K by φ is modelled as a (revision) function $*$ that maps K and φ to the revised (new) theory $K * \varphi$. Eight postulates, called the *AGM postulates for revision*, *axiomatically* characterize any *rational* revision operator, named *AGM revision function*. It has been, also, proven that any AGM revision function can be *semantically constructed* (specified) by means of a special kind of total preorders over possible worlds, called *faithful preorders* [8].

Given the nice properties of *Horn logic* (i.e., the Horn fragment of propositional logic), the AGM paradigm was modified by Delgrande and Peppas, so that it characterizes the class of AGM revision functions that map a Horn belief set and a Horn sentence to a (new) revised Horn belief set [5]; we shall refer to such AGM revision functions as *Horn AGM revision functions*.

In this article, we first identify (both *axiomatically* and *semantically*) an interesting *proper sub-class* of the class of Horn AGM revision functions, which is based on *proper set inclusion* of the atomic propositions satisfied by possible worlds (Section 4). Furthermore, we prove that a special type of well-behaved AGM revision functions, called *uniform-revision* (UR) operators —introduced in [3] and subsequently studied in detail in [2]— is Horn-compliant, in the sense that the class of UR operators *intersects* the class of Horn AGM revision functions (Section 5). This is demonstrated by proving that a concrete "off-the-shelf" Horn AGM revision function, proposed in [5], as well as some of the aforementioned inclusion-based Horn AGM revision functions are, as a matter of fact, particular UR operators.

2 Basic Notations and Conventions

We shall be working with a propositional language $\mathcal{L}$, built over a *finite*, non-empty set $\mathcal{P}$ of atoms, using the standard Boolean connectives, and governed by *classical propositional logic*. For a set of sentences Γ of $\mathcal{L}$, $Cn(\Gamma)$ denotes the set of all logical consequences of Γ; i.e.,

$Cn(\Gamma) = \{\varphi \in \mathcal{L} : \Gamma \models \varphi\}$. A *theory*, also referred to as a *belief set*, K is any deductively closed set of sentences of $\mathcal{L}$; i.e., $K = Cn(K)$. The set of all theories is denoted by $\mathbb{K}$.

A *literal* is an atom $p \in \mathcal{P}$ or its negation. For a set of literals Q, $|Q|$ denotes the cardinality of Q, and $\overline{Q}$ denotes the set of negated elements of Q; i.e., $\overline{Q} = \{\neg l : l \in Q\}$. A *possible world* (or, simply, *world*) r is any consistent set of literals, such that, for any atom $p \in \mathcal{P}$, either $p \in r$ or $\neg p \in r$. The set of all possible worlds is denoted by $\mathbb{M}$. The set of all atoms satisfied by a world $r \in \mathbb{M}$ is denoted by r^+; i.e., $r^+ = r \cap \mathcal{P}$. For a sentence (set of sentences) φ of $\mathcal{L}$, $[\varphi]$ is the set of worlds at which φ is true. Possible worlds will, occasionally, be represented as sequences of literals, and the negation of an atom p will be represented as $\overline{p}$.

A clause (i.e., a disjunction of literals) is called a *Horn clause* iff it contains *at most* one atom; e.g., $a \vee \neg b \vee \neg c$, where a, b, c are atoms. A *Horn formula* is a conjunction of Horn clauses. The *Horn language* $\mathcal{L}_H$ is the maximal subset of $\mathcal{L}$ containing only Horn formulas. The set of all Horn theories is denoted by $\mathbb{H}$. The Horn logic generated from $\mathcal{L}_H$ is specified by the consequence operator Cn_H, such that, for any set Γ of Horn formulas, $Cn_H(\Gamma) = Cn(\Gamma) \cap \mathcal{L}_H$. The *atoms-intersection* of two worlds $r, r' \in \mathbb{M}$ is a world denoted by $r \cap^+ r'$, and defined as follows: $r \cap^+ r' = (r^+ \cap r'^+) \cup \big(\overline{\mathcal{P} - (r^+ \cap r'^+)}\big)$.[1] An arbitrary formula (or theory) φ is Horn (i.e., $\varphi \in \mathcal{L}_H$) iff $r, r' \in [\varphi]$ entails $r \cap^+ r' \in [\varphi]$.

A *preorder* $\preceq$ over a set V is called *total* iff, for all $r, r' \in V$, $r \preceq r'$ or $r' \preceq r$. The strict part of $\preceq$ is denoted by $\prec$; i.e., $r \prec r'$ iff $r \preceq r'$ and $r' \not\preceq r$. The symmetric part of $\preceq$ is denoted by $\approx$; i.e., $r \approx r'$ iff $r \preceq r'$ and $r' \preceq r$. Also, $min(V, \preceq)$ denotes the set of all $\preceq$-minimal elements of V; i.e., $min(V, \preceq) = \Big\{ r \in V : \text{for all } r' \in V, \text{ if } r' \preceq r, \text{ then } r \preceq r' \Big\}$.

3 AGM-Style Horn Revision

Within the AGM paradigm, the revision-process is modelled as a (binary) function $*$ that maps a theory K and a sentence φ to the revised (new) theory $K * \varphi$; i.e., $* : \mathbb{K} \times \mathcal{L} \mapsto \mathbb{K}$. The *AGM postulates for revision* —which are not presented herein due to space limitations— *axiomatically* characterize all *rational* revision functions, the so-called *AGM revision functions*.[2] Katsuno and Mendelzon proved that any AGM revision function can be *semantically* specified with the use of a special type of total preorders over all possible worlds, called *faithful preorders* [8].

Definition 1 (Faithful Preorder, [8]). *A total preorder $\preceq_K$ over $\mathbb{M}$ is faithful to a theory K iff the $\preceq_K$-minimal worlds are those satisfying K; i.e., $min(\mathbb{M}, \preceq_K) = [K]$.*

Intuitively, $r \preceq_K r'$ holds whenever r is at least as *plausible* (relative to K) as r'.

Theorem 1 ([8]). *For every theory K and any sentence φ of $\mathcal{L}$, an AGM revision function $*$ can be defined (specified) by means of the following condition:*

$$\textbf{(F}*\textbf{)} \quad [K * \varphi] = min([\varphi], \preceq_K).$$

3.1 The AGM Paradigm in the Horn Setting

Since satisfiability in the Horn setting —i.e., evaluating whether $\varphi \in H$ is true, where $H \in \mathbb{H}$ and $\varphi \in \mathcal{L}_H$— can be determined in *linear time* [6, 9], *Horn revision* has gained great interest. A notable approach, in this regard, constitutes the work of Delgrande and Peppas [5]. In that work, the authors *axiomatically* characterized the class of Horn AGM revision functions, by

[1] For instance, $\{a, b, c\} \cap^+ \{a, \neg b, \neg c\} = \{a, \neg b, \neg c\}$ and $\{\neg a, b, c\} \cap^+ \{a, \neg b, \neg c\} = \{\neg a, \neg b, \neg c\}$.

[2] See [7, Section 3.3] or [10, Section 8.3.1] for a detailed presentation of the AGM postulates for revision.

strictly strengthening the class of AGM revision functions, with the aid of an extra postulate that supplements the AGM postulates for revision — for details, the reader is referred to [5]. We recall that a Horn AGM revision function $*$ is an AGM revision function that maps a Horn theory and a Horn formula to a (new) revised Horn theory; i.e., $* : \mathbb{H} \times \mathcal{L}_H \mapsto \mathbb{H}$.

Now, consider the following constraint on a total preorder $\preceq_K$ over $\mathbb{M}$, which is faithful to a theory K, introduced by Zhuang and Pagnucco in [11].

(H) If $r \approx_K r'$, then $r \cap^+ r' \preceq_K r$.

Condition (H) says that whenever two worlds r and r' are equidistant from the beginning of a preorder $\preceq_K$, then the world $r \cap^+ r'$, resulting from their atoms-intersection, cannot appear later in $\preceq_K$. The results of Delgrande and Peppas [5], along with important results established by Zhuang and Pagnucco in [11], entail the following *representation theorem*.

Theorem 2 ([11, 5]). *Let $*$ be an AGM revision function, and let $\{\preceq_K\}_{K\in\mathbb{K}}$ be the family of total preorders over worlds that correspond to $*$, by means of condition (F$*$). Then, $*$ is a Horn AGM revision function iff $\{\preceq_K\}_{K\in\mathbb{K}}$ satisfies condition (H).*

3.2 Basic Horn Revision

Delgrande and Peppas not only axiomatically characterized AGM-style Horn revision, but also proposed some interesting concrete Horn AGM revision functions [5]. In this subsection, we present one of their proposals, which is inspired by the *Hamming-based* Dalal's approach [4].

The *basic Horn revision function*, denoted by $\diamond$, is defined as the (*unique*) Horn AGM revision function induced, via condition (F$*$), from a family $\{\preceq_H\}_{H\in\mathbb{H}}$ of total preorders over $\mathbb{M}$, that satisfies the following constraint.

(BH) $r \preceq_H r'$ iff $|r^+| \leqslant |r'^+|$.

Condition (BH) orders the relative plausibility of worlds according to the number of atoms they satisfy; notice that (BH) *uniquely* specifies $\preceq_H$, thus, $\diamond$ is unique.

4 Inclusion-Based Horn Revision

In this section, we introduce a new *proper sub-class* of Horn AGM revision functions, based on *proper set inclusion* of atoms of worlds. First, we introduce the next definition.

Definition 2 (Atoms-Ordered Theory). *Let K be a consistent theory of $\mathcal{L}$. We shall say that K is atoms-ordered iff the atoms of the worlds of $[K]$ are totally ordered with respect to proper set inclusion; i.e., for any two worlds $r, r' \in [K]$, either $r^+ \subset r'^+$ or $r'^+ \subset r^+$.*

On that premise, consider the following postulate (PI). The semantic condition that corresponds to (PI) is condition (PIS), also presented below, which constrains a total preorder $\preceq_K$ over $\mathbb{M}$, which is faithful to a theory K.

(PI) For any consistent φ of $\mathcal{L}$, $K * \varphi$ is atoms-ordered.

(PIS) If $r \approx_K r'$, then either $r^+ \subset r'^+$ or $r'^+ \subset r^+$.

According to (PIS), the faithful preorder $\preceq_K$ is defined so that the atoms of the non-K-worlds of every equivalent class (layer) of $\preceq_K$ are totally ordered with respect to *proper set inclusion*. An example of a faithful preorder that respects condition (PIS) is shown below, for $\mathcal{P} = \{a, b, c\}$ and the Horn belief set $H = Cn_H\big((\neg a \vee b) \wedge \neg c\big)$ — notice that $[H] = \big\{\{a, b, \neg c\}, \{\neg a, b, \neg c\}, \{\neg a, \neg b, \neg c\}\big\}$.

$$\begin{array}{ccccc} ab\overline{c} & & & & abc \\ \overline{a}b\overline{c} & \prec_H & \begin{array}{c}\overline{a}bc \\ \overline{a}\overline{b}c\end{array} & \prec_H & a\overline{b}c \\ \overline{a}\overline{b}\overline{c} & & & & a\overline{b}\overline{c} \end{array}$$

Theorem 3 establishes the connection between postulate (PI) and constraint (PIS).

Theorem 3. *Let $*$ be an AGM revision function, and let $\{\preceq_K\}_{K\in\mathbb{K}}$ be the family of total preorders over $\mathbb{M}$ that correspond to $*$, by means of condition (F$*$). Then, $*$ satisfies postulate (PI) iff $\{\preceq_K\}_{K\in\mathbb{K}}$ satisfies condition (PIS).*

Proof. For the left-to-right implication, assume that $*$ satisfies (PI). We show that $\{\preceq_K\}_{K\in\mathbb{K}}$ satisfies (PIS). Let r, r' be two worlds of $\mathbb{M}$, such that $r \approx_K r'$. Define φ to be a sentence of $\mathcal{L}$, such that $[\varphi] = \{r, r'\}$. Then, condition (F$*$) entails that $[K * \varphi] = \{r, r'\}$. Therefore, from condition (PI), we have that either $r^+ \subset r'^+$ or $r'^+ \subset r^+$, as desired.

For the right-to-left implication, assume that $\{\preceq_K\}_{K\in\mathbb{K}}$ satisfies (PIS). We show that $*$ satisfies (PI). For any consistent sentence φ of $\mathcal{L}$, it follows, from condition (F$*$), that all worlds in $[K * \varphi]$ are equally plausible, with respect to K. That is, for any two worlds $r, r' \in [K * \varphi]$, it is true that $r \approx_K r'$. Hence, from (PIS), we have that either $r^+ \subset r'^+$ or $r'^+ \subset r^+$. This again entails that $K * \varphi$ is atoms-ordered, as desired. ■

The next theorem shows that any AGM revision function that respects postulate (PI) is a Horn AGM revision function.

Theorem 4. *Let $*$ be an AGM revision function. If $*$ satisfies postulate (PI), then $*$ is a Horn AGM revision function.*

Proof. Let H be a Horn belief set, and let $\preceq_H$ be the faithful preorder that $*$ assigns at H, via (F$*$). Since $*$ satisfies (PI), $\preceq_H$ satisfies (PIS). It suffices to show that $\preceq_H$ satisfies condition (H). Let r, r' be two worlds of $\mathbb{M}$, such that $r \approx_H r'$. We will show that $r \cap^+ r' \preceq_H r$. If $r \cap^+ r' \in [H]$, this is clearly true. Assume, therefore, that $r \cap^+ r' \notin [H]$. Then, since H is a Horn theory, not both r and r' can be members of $[H]$ (for otherwise $r \cap^+ r'$ would, also, belong to $[H]$). Since one of r, r' is not in $[H]$, from $r \approx_H r'$, we derive that neither of r, r' belong to $[H]$. From $r \approx_H r'$, (PIS) entails that either $r^+ \subset r'^+$ or $r'^+ \subset r^+$. This again implies that $r \cap^+ r' \approx_H r \approx_H r'$; that is, $r \cap^+ r' \preceq_H r$, as desired. ■

Theorem 4, along with Theorem 5 shown below, prove that the family of Horn AGM revision functions identified by postulate (PI) constitutes a *proper* sub-class of the whole class of Horn AGM revision functions.

Theorem 5. *There exists a Horn AGM revision function that does not satisfy postulate (PI).*

Proof. Let $\mathcal{P} = \{a, b, c\}$, and let H be a Horn belief set such that $H = Cn_H\big((\neg a \vee b) \wedge \neg c\big)$. Clearly, $[H] = \big\{\{a, b, \neg c\}, \{\neg a, b, \neg c\}, \{\neg a, \neg b, \neg c\}\big\}$. Let $*$ be a Horn AGM revision function that assigns at H (via (F$*$)) the next total preorder $\preceq_H$ over $\mathbb{M}$, which respects condition (H).

$$\begin{array}{ccccc} ab\overline{c} & & \overline{a}bc & & \\ \overline{a}b\overline{c} & \prec_H & \overline{a}\overline{b}c & \prec_H & \begin{array}{c} abc \\ a\overline{b}\overline{c} \end{array} \\ \overline{a}\overline{b}\overline{c} & & a\overline{b}c & & \end{array}$$

Observe that $\preceq_H$ does *not* respect condition (PIS), hence, $*$ does not satisfy postulate (PI). ■

5 Uniform Revision is Horn-Compliant

This section investigates *uniform revision* in the realm of Horn logic. A uniform-revision (UR) operator is *uniquely* specified by means of a *single* total preorder $\preccurlyeq$ over all possible worlds of $\mathbb{M}$, which essentially expresses their *prior* relative plausibility, as considered by a particular rational agent. This is accomplished since a total preorder $\preccurlyeq$ suffices to *uniquely* specify the agent's revision policy, with respect to *every* belief set of the language (by means of condition (URS1), presented subsequently) [3, 2].

It proves to be the case that a UR operator is any AGM revision function $*$ satisfying the following postulate (UR); the semantic condition that corresponds to (UR) is condition (URS1), which is, in turn, equivalent to condition (URS2) (where $K, T \in \mathbb{K}$) [3, 2].

(UR) For any $\neg\varphi \in K \cap T$, $K * \varphi = T * \varphi$.

(URS1) For any $r, r' \notin [K]$, $r \preceq_K r'$ iff $r \preccurlyeq r'$.

(URS2) For any $r, r' \notin [K] \cup [T]$, $r \preceq_K r'$ iff $r \preceq_T r'$.

Against this background, Theorem 6 proves that basic Horn revision encodes, as a matter of fact, a particular uniform-revision policy.

Theorem 6. *The basic Horn revision function $\diamond$ is a UR operator.*

Proof. Let H, H' be any two Horn belief sets, and let $\preceq_H$, $\preceq_{H'}$ be the faithful preorders that $\diamond$ assigns (via (F$*$)) at H, H', respectively. It suffices to show that $\preceq_H$, $\preceq_{H'}$ satisfy condition (URS2). Since $\preceq_H$, $\preceq_{H'}$ satisfy condition (BH), it is true that, for any worlds $z, z' \notin [H]$ and any worlds $u, u' \notin [H']$, $z \preceq_H z'$ iff $|z^+| \leqslant |z'^+|$, and $u \preceq_{H'} u'$ iff $|u^+| \leqslant |u'^+|$. This again entails that, for any worlds $r, r' \notin [H] \cup [H']$, $r \preceq_H r'$ iff $r \preceq_{H'} r'$. Therefore, the faithful preorders $\preceq_H$, $\preceq_{H'}$ satisfy condition (URS2), as desired. ■

The aforementioned theorem, essentially, shows that the class of UR operators *intersects* the class of Horn AGM revision functions. Therefore, uniform revision is Horn-compliant — this is an important result that comes to extend the favourable properties of uniform revision.

Remark 1. *One can easily find UR operators that are not Horn AGM revision functions, as well as Horn AGM revision functions that are not UR operators.*

Next, Theorem 7 proves that there exist uniform-revision policies that implement inclusion-based Horn revision.

Theorem 7. *There exists a UR operator that satisfies postulate (PI).*

Proof. Let $\preccurlyeq$ be a total preorder over $\mathbb{M}$, defined as follows ($\prec\!\cdot$ denotes the strict part of $\preccurlyeq$):

$$\begin{array}{ccccc} ab\overline{c} & & & & abc \\ \overline{a}b\overline{c} & \prec\!\cdot & \begin{array}{c} \overline{a}bc \\ \overline{a}\overline{b}c \end{array} & \prec\!\cdot & a\overline{b}c \\ \overline{a}\overline{b}\overline{c} & & & & a\overline{b}\overline{c} \end{array}$$

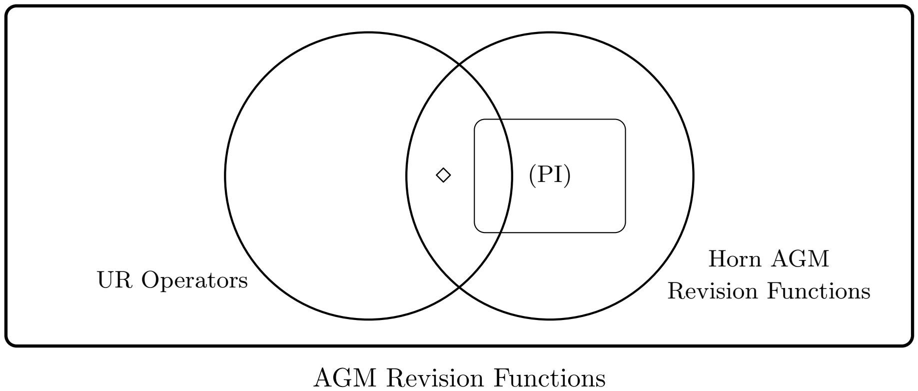

Figure 1: The types of AGM revision functions discussed herein.

The preorder $\preccurlyeq$ specifies (via (URS1)) a unique family $\{\preceq_K\}_{K\in\mathbb{K}}$ of total preorders over $\mathbb{M}$, which, in turn, induces (via (F$*$)) a unique UR operator $*$. Since $\preccurlyeq$ respects condition (PIS), it follows that $\{\preceq_K\}_{K\in\mathbb{K}}$ respects (PIS) as well. Hence, $*$ satisfies postulate (PI), as desired. ■

Remark 2. *One can easily find inclusion-based Horn AGM revision functions that are not UR operators. Moreover, there exists a Horn AGM revision function, which is, also, a UR operator —namely, the basic Horn revision function— that can be verified to violate postulate (PI).*

It can be shown that any class of AGM revision functions that are induced from faithful preorders which specify the relative plausibility of possible worlds *regardless* of the (information contained in the) respective belief set —such as the basic Horn revision function $\diamond$ and the inclusion-based Horn AGM revision functions, considered in this section— *intersects* the class of UR operators; the details are left for future research.

The results of the present work are summarized in Figure 1, which depicts the types of AGM revision functions discussed herein.

6 Conclusion

In this work, we identified some interesting types of Horn AGM revision functions. In particular, we defined (axiomatically and semantically) a proper sub-class of Horn AGM revision functions, based on proper set inclusion of the atoms of possible worlds. We, also, showed that the well-behaved uniform revision is Horn-compliant, since concrete Horn AGM revision functions are, in fact, particular UR operators. Given the critical role that Horn logic plays in belief revision, further research on other solid types of Horn AGM revision functions is quite compelling.

Acknowledgments

This research is co-financed by Greece and the European Union (European Social Fund – ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Reinforcement of Postdoctoral Researchers – 2nd Cycle" (MIS-5033021), implemented by the State Scholarships Foundation (IKY).

We thank the anonymous referee for their valuable input on this article.

References

[1] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[2] Theofanis Aravanis. On uniform belief revision. *Journal of Logic and Computation*, 30:1357–1376, 2020.

[3] Carlos Areces and Verónica Becher. Iterable AGM functions. In Mary-Anne Williams and Hans Rott, editors, *Frontiers in Belief Revision*, Lecture Notes in Computer Science, pages 261–277. Kluwer Academic Publishers, 2001.

[4] Mukesh Dalal. Investigations into theory of knowledge base revision: Preliminary report. In *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence (AAAI 1988)*, pages 475–479. The AAAI Press, Menlo Park, California, 1988.

[5] James P. Delgrande and Pavlos Peppas. Belief revision in Horn theories. *Artificial Intelligence*, 218:1–22, 2015.

[6] William F. Dowling and Jean H. Gallier. Linear-time algorithms for testing the satisfiability of propositional Horn theories. *Journal of Logic Programming*, 3:267–284, 1984.

[7] Peter Gärdenfors. *Knowledge in Flux – Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, Massachusetts, 1988.

[8] Hirofumi Katsuno and Alberto Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3):263–294, 1991.

[9] Michael Minoux. LTUR: A simplified linear-time unit resolution algorithm for Horn formulae and computer implementation. *Information Processing Letters*, 29:1–12, 1988.

[10] Pavlos Peppas. Belief revision. In Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, editors, *Handbook of Knowledge Representation*, pages 317–359. Elsevier Science, 2008.

[11] Zhiqiang Zhuang and Maurice Pagnucco. Model based horn contraction. In *Proceedings of the 13th International Conference on Principles of Knowledge Representation and Reasoning (KR 2012)*, pages 169–178, 2012.

Properties of Parametrized-Difference Revision

Theofanis Aravanis

Department of Business Administration
School of Economics & Business
University of Patras, Greece
taravanis@upatras.gr

Abstract

Parametrized-difference (PD) revision is a special type of rational belief revision, based on a fixed ranking over the atoms of the underlying language, with a plethora of appealing characteristics. The process of PD revision is encoded in the so-called PD operators, which essentially constitute a particular family of rational revision functions. In this article, we identify some new properties of PD revision. Specifically, we demonstrate how this type of belief change is tightly connected with selective revision, another type of revision according to which the new information is partially accepted in the revised state of belief. Furthermore, we show that PD operators respect a relevance-sensitive postulate, which introduces dependencies between revisions carried out on different (overlapping) belief sets.

1 Introduction

Belief revision (or, simply, *revision*) is the process by which a rational agent modifies her/his beliefs, in the light of new information [8, 11]. A prominent and versatile approach which formalizes belief revision is that proposed by Alchourrón, Gärdenfors and Makinson [1], known as the *AGM paradigm*. The AGM paradigm characterizes any *rational* revision operator, named *AGM revision function*, which, essentially, is a (binary) function that maps a belief set (theory) and an epistemic input (a sentence that represents new information) to a (new) revised belief set.

A family of well-behaved concrete AGM revision functions, called *parametrized-difference* (PD) operators, was recently introduced by Peppas and Williams [12, 13]. Each PD operator is *uniquely* defined by means of a *single* total preorder over the atoms of the underlying language, hence, it is compactly-specified. PD operators, also, have an embedded solution to the *iterated-revision* problem, and are expressive enough to cover a wide range of revision-scenarios, features that make them an ideal candidate for real-world implementations.

In this article, we identify some new interesting properties of PD revision. Specifically, we first demonstrate how PD revision is strongly connected with *selective revision*, a type of belief change according to which the new information is *partially accepted* in the revised belief set [7]. Already in [12, 4], it was shown that PD operators respect Parikh's *relevance-sensitive* axiom [10], an intuitive principle that supplements the AGM postulates for revision in dealing with *relevant* change. Herein, we show that PD operators, also, respect another relevance-sensitive postulate, which introduces dependencies between revisions carried out on different (overlapping) belief sets.

The article is structured as follows. The next section fixes the required formal preliminaries. Sections 3 and 4 introduce the AGM paradigm and PD revision, respectively. Thereafter, Section 5 discusses PD revision with respect to selective revision, and Section 6 points out some relevance-sensitive properties of PD revision. A brief conclusion closes the paper.

2 Formal Preliminaries

Herein, we work with a propositional language $\mathcal{L}$, built over a *finite*, non-empty set $\mathcal{P}$ of atoms (propositional variables), using the standard Boolean connectives $\wedge$ (conjunction), $\vee$ (disjunction), $\rightarrow$ (implication), $\leftrightarrow$ (equivalence), $\neg$ (negation), and governed by *classical propositional logic*. The classical consequence relation is denoted by $\models$.

A sentence φ of $\mathcal{L}$ is *contingent* iff $\not\models \varphi$ and $\not\models \neg\varphi$. For a set of sentences Γ of $\mathcal{L}$, $Cn(\Gamma)$ denotes the set of all logical consequences of Γ; i.e., $Cn(\Gamma) = \{\varphi \in \mathcal{L} : \Gamma \models \varphi\}$. An agent's belief corpus shall be modelled by a *theory*, also referred to as a *belief set*. A theory K is any deductively closed set of sentences of $\mathcal{L}$; i.e., $K = Cn(K)$. The set of all *consistent* theories is denoted by $\mathbb{K}$. For a theory K and a sentence φ of $\mathcal{L}$, we define $K + \varphi = Cn\big(K \cup \{\varphi\}\big)$.

A *literal* is an atom $p \in \mathcal{P}$ or its negation. For a set of literals Q, $|Q|$ denotes the cardinality of Q. A *possible world* (or, simply, *world*) r is a consistent set of literals, such that, for any atom $p \in \mathcal{P}$, either $p \in r$ or $\neg p \in r$. The set of all possible worlds is denoted by $\mathbb{M}$. For a sentence (set of sentences) φ of $\mathcal{L}$, $[\varphi]$ is the set of worlds at which φ is true.

Let Q be a subset of $\mathcal{P}$. We denote by $\mathcal{L}^Q$ the *sublanguage* of $\mathcal{L}$ defined over Q, using the standard Boolean connectives. For a sentence x of $\mathcal{L}$, $\mathcal{P}_x$ denotes the (*unique*) *minimal* subset of $\mathcal{P}$, through which a sentence that is logically equivalent to x can be formulated. If x is inconsistent or a tautology, we take $\mathcal{P}_x$ to be the empty set. Then, we define $\mathcal{L}_x$ and $\overline{\mathcal{L}_x}$ to be the propositional (sub)languages defined over $\mathcal{P}_x$ and $\mathcal{P} - \mathcal{P}_x$, respectively, using the standard Boolean connectives. Let φ be a contingent sentence of $\mathcal{L}$. For a world $r \in \mathbb{M}$, r_φ and $r_{\overline{\varphi}}$ denote the *restrictions* of r to $\mathcal{L}_\varphi$ and $\overline{\mathcal{L}_\varphi}$, respectively; i.e., $r_\varphi = r \cap \mathcal{L}_\varphi$ and $r_{\overline{\varphi}} = r \cap \overline{\mathcal{L}_\varphi}$. For a set of worlds V, V_φ and $V_{\overline{\varphi}}$ denote the sets of (restricted) worlds resulting from the restriction of all V-worlds to $\mathcal{L}_\varphi$ and $\overline{\mathcal{L}_\varphi}$, respectively; i.e., $V_\varphi = \{r_\varphi : r \in V\}$ and $V_{\overline{\varphi}} = \{r_{\overline{\varphi}} : r \in V\}$.

A *preorder* over a set V is any reflexive, transitive binary relation in V. A preorder $\preceq$ is called *total* iff, for all $r, r' \in V$, $r \preceq r'$ or $r' \preceq r$. Also, $min(V, \preceq)$ denotes the set of all $\preceq$-minimal elements of V; i.e., $min(V, \preceq) = \big\{r \in V : \text{for all } r' \in V, \text{ if } r' \preceq r, \text{ then } r \preceq r'\big\}$.

3 The AGM Paradigm

Within the AGM paradigm, the process of belief revision is modelled as a (binary) function $*$ mapping a theory K and a sentence φ to a revised (new) theory $K * \varphi$. *Rational* revision functions, the so-called *AGM revision functions*, are those constrained by a set of eight postulates, called *AGM postulates for revision*, listed below [8, 11].

- **(K ∗ 1)** $K * \varphi$ is a theory of $\mathcal{L}$.
- **(K ∗ 2)** $\varphi \in K * \varphi$.
- **(K ∗ 3)** $K * \varphi \subseteq K + \varphi$.
- **(K ∗ 4)** If $\neg\varphi \notin K$, then $K + \varphi \subseteq K * \varphi$.
- **(K ∗ 5)** $K * \varphi$ is inconsistent iff φ is inconsistent.
- **(K ∗ 6)** If $Cn\big(\{\varphi\}\big) = Cn\big(\{\psi\}\big)$, then $K * \varphi = K * \psi$.
- **(K ∗ 7)** $K * (\varphi \wedge \psi) \subseteq (K * \varphi) + \psi$.
- **(K ∗ 8)** If $\neg\psi \notin K * \varphi$, then $(K * \varphi) + \psi \subseteq K * (\varphi \wedge \psi)$.

Katsuno and Mendelzon proved that the revision functions that satisfy postulates $(K * 1)$–

$(K * 8)$ are precisely those that are induced by means of a special type of total preorders over all possible worlds, called *faithful preorders* [9].

Definition 1 (Faithful Preorder, [9])**.** *A total preorder $\preceq_K$ over $\mathbb{M}$ is faithful to a theory K iff the $\preceq_K$-minimal worlds are those satisfying K; i.e., $min(\mathbb{M}, \preceq_K) = [K]$.*

Intuitively, $r \preceq_K r'$ holds when r is at least as *plausible* (relative to K) as r'.

Definition 2 (Faithful Assignment, [9])**.** *A faithful assignment is a function that maps each theory K of $\mathcal{L}$ to a total preorder $\preceq_K$ over $\mathbb{M}$, that is faithful to K.*

The following representation theorem precisely characterizes the class of AGM revision functions, in terms of faithful preorders.

Theorem 1 ([9])**.** *A revision function $*$ satisfies $(K * 1)$–$(K * 8)$ iff there exists a faithful assignment that maps each theory K to a total preorder $\preceq_K$ over $\mathbb{M}$, such that, for any $\varphi \in \mathcal{L}$:*

$$\textbf{(F}*\textbf{)} \quad [K * \varphi] = min([\varphi], \preceq_K).$$

For ease of presentation, we shall consider, herein, only the principal case of *consistent* belief sets and *contingent* epistemic input.

4 Parametrized-Difference Revision

Peppas and Williams [12, 13], recently, introduced a *proper sub-class* of concrete AGM revision functions, well-suited for real-world implementations, called *parametrized-difference* (PD) operators. PD operators are a generalization of the *Hamming-based* Dalal's revision operator [6], as each such operator is specified by a $\trianglelefteq$-parametrization of Dalal's construction, where $\trianglelefteq$ denotes a fixed total preorder over all atoms of $\mathcal{P}$, which encodes their (prior) relative epistemic value; the more epistemic entrenched (and, thus, more resistant to change) an atom is, the higher it appears in $\trianglelefteq$. In this section, we briefly review PD operators; for details on their definition, the reader is referred to [12, 13, 4, 5].

Definition 3 (Difference between Worlds)**.** *The difference between two (possibly restricted) worlds w, r, denoted by Diff (w, r), is the set of atoms over which w and r disagree. In symbols,*

$$\mathit{Diff}\,(w, r) = \Big((w - r) \cup (r - w)\Big) \cap \mathcal{P}.$$

For a set of atoms $\mathcal{S}$ and an atom q, we define $\mathcal{S}_q = \{p \in \mathcal{S} : p \trianglelefteq q\}$. Definition 4 extends, then, the total preorder $\trianglelefteq$ to *sets* of atoms.

Definition 4 (Total Preorder over Sets of Atoms, [12])**.** *For any two sets of atoms $\mathcal{S}$, $\mathcal{S}'$, $\mathcal{S} \trianglelefteq \mathcal{S}'$ iff one of the next three conditions holds ($\triangleleft$ denotes the strict part of $\trianglelefteq$):*

(i) $|\mathcal{S}| < |\mathcal{S}'|$.

(ii) $|\mathcal{S}| = |\mathcal{S}'|$, *and for all $q \in \mathcal{P}$,* $|\mathcal{S}_q| = |\mathcal{S}'_q|$.

(iii) $|\mathcal{S}| = |\mathcal{S}'|$, *and for some $q \in \mathcal{P}$,* $|\mathcal{S}_q| > |\mathcal{S}'_q|$, *and for all $p \triangleleft q$,* $|\mathcal{S}_p| = |\mathcal{S}'_p|$.

Definition 5 (PD Operator, [12]). *Let $\trianglelefteq$ be a total preorder over $\mathcal{P}$. A PD operator is the revision function induced, via condition (F$*$), from the family of PD preorders $\{\sqsubseteq_T^{\trianglelefteq}\}_{T\in\mathbb{K}}$, where each PD preorder $\sqsubseteq_K^{\trianglelefteq}$ is uniquely defined, for any $r, r' \in \mathbb{M}$, by means of condition (PD) below.*

$$\textbf{(PD)} \quad r \sqsubseteq_K^{\trianglelefteq} r' \quad \textit{iff} \quad \textit{there is a } w \in [K]\textit{, such that, for all } w' \in [K], \; \textit{Diff}\,(w, r) \trianglelefteq \textit{Diff}\,(w', r').$$

Definition 5 implies that there is a *one-to-one correspondence* between the total preorders over atoms and the PD operators. Note, lastly, that, when $\trianglelefteq\, = \mathcal{P} \times \mathcal{P}$ (i.e., all atoms have *equal* epistemic value), the PD preorder $\sqsubseteq_K^{\trianglelefteq}$ is defined, for any $r, r' \in \mathbb{M}$, as follows: $r \sqsubseteq_K^{\trianglelefteq} r'$ iff there is a $w \in [K]$, such that, for all $w' \in [K]$, $|\textit{Diff}\,(w, r)| \leqslant |\textit{Diff}\,(w', r')|$. In this case, the family $\{\sqsubseteq_K^{\trianglelefteq}\}_{K\in\mathbb{T}}$ produces Dalal's operator [6].

5 PD and Selective Revision

In this section, we demonstrate how a total preorder $\trianglelefteq$ over the atoms of the language —which, essentially, constitutes the generative unit of PD revision— can be utilized for implementing *selective revision*, a special type of belief revision according to which *only a part* of an epistemic input is accepted in the revised state of belief [7]. Recall that, in the standard AGM paradigm, the new information is *always* accepted (due to postulate $(K * 2)$). This, however, is a rather unrealistic assumption, since real-world rational agents do not always receive information from reliable sources. The following scenario, borrowed from [7], is illustrative.

Example 1 ([7]). *You return back from work and your son tells you, as soon as you see him: "A dinosaur has broken grandma's vase in the living-room". You, probably, accept the information that the vase has been broken, and reject the part of the information that refers to the dinosaur.*

A plausible way for such information filtering would be by taking into account the relative plausibility of the "building blocks" of sentences of the language. As these "building blocks", essentially, are the atoms of the language, and the relative plausibility of the atoms is encoded in a total preorder $\trianglelefteq$, it turns out that PD revision provides a means for *information filtering*. To see this, suppose that, during revision, the following *filtering-rule* is applied:

> *"If the epistemic input is a conjunction of atoms, then only the strictly most $\trianglelefteq$-plausible atoms should be accepted in the revised belief set".*

On that premise, if $a \wedge b$ is an epistemic input (where a, b are atoms), and, moreover, we have that $a \triangleleft b$, then only the atom b should be accepted in the revised state of belief.

6 Relevance-Sensitive Properties of PD Revision

Parikh pointed out that the AGM postulates for revision are liberal in their treatment of *relevance* [10]. To remedy this weakness, he proposed an additional axiom that supplements postulates $(K*1)$–$(K*8)$, named axiom (P), according to which the revision of a theory K that can be divided in two *syntax-disjoint* compartments by an epistemic input φ that is syntax-related only to the first compartment of K should *not affect* the second compartment of K. In a subsequent work [14], two interpretations of Parikh's axiom were identified, namely, its *weak* and *strong* version.[1] Already in [12], it was shown that PD operators respect the weak version

[1] The *semantic* properties of Parikh's axiom were investigated, in detail, in [2, 3].

of axiom (P), whereas, in [4], it was shown that PD operators, also, respect the strong version of (P).

In this section, we point out further interesting *relevance-sensitive* properties of PD revision. For establishing our results, we shall use the already known fact that PD operators respect the following postulate (R), which states that the $\overline{\mathcal{L}_\varphi}$-part of the revised theory $K * \varphi$ contains *at least* every sentence of the $\overline{\mathcal{L}_\varphi}$-part of the initial theory K.

$$\textbf{(R)} \quad K \cap \overline{\mathcal{L}_\varphi} \subseteq (K * \varphi) \cap \overline{\mathcal{L}_\varphi}.$$

Postulate (R) —which is equivalent to $[K * \varphi]_{\overline{\varphi}} \subseteq [K]_{\overline{\varphi}}$ in the realm of possible worlds— implies some interesting properties of PD revision, encoded in Lemma 1. To present this lemma, let us first introduce the required notation. For an arbitrary theory K and a sentence φ, such that $\mathcal{L}_\varphi \subset \mathcal{L}$, we denote by $[\varphi]^K$ the set of φ-worlds whose $\overline{\mathcal{L}_\varphi}$-part agrees with the $\overline{\mathcal{L}_\varphi}$-part of some K-world; i.e., $[\varphi]^K = \Big\{ r \in [\varphi] : r_{\overline{\varphi}} \in [K]_{\overline{\varphi}} \Big\}$. By definition, it holds that $[\varphi]^K \subseteq [\varphi]$.

Lemma 1. *Let $\trianglelefteq$ be a total preorder over atoms, and let $*$ be a PD operator induced from the family of PD preorders $\{\sqsubseteq_T^{\trianglelefteq}\}_{T \in \mathbb{K}}$. Moreover, let K be a theory, and let φ be a sentence, such that $\mathcal{L}_\varphi \subset \mathcal{L}$. Then, the following identity is true:*

$$\begin{aligned}
&[K * \varphi] = min\big([\varphi]^K, \sqsubseteq_{K_{IC}}^{\trianglelefteq}\big) \\
&= \Big\{ u \in [\varphi]^K : \exists\, w' \in [K] \ s.t. \ \mathit{Diff}\,(w', u) \in min\Big(\big\{ \mathit{Diff}\,(w, r) : w \in [K] \ and \ r \in [\varphi]^K \big\}, \trianglelefteq \Big) \Big\} \\
&= \Big\{ u \in [\varphi]^K : \exists\, w' \in [K] \ s.t. \ \mathit{Diff}\,(w', u) \in min\Big(\big\{ \mathit{Diff}\,(w_\varphi, r_\varphi) : w \in [K] \ and \ r \in [\varphi]^K \big\}, \trianglelefteq \Big) \Big\}
\end{aligned}$$

Proof. The first equality follows directly from condition (F$*$), from which we have that $[K * \varphi] = min([\varphi], \sqsubseteq_K^{\trianglelefteq})$, and postulate (R), which entails that $[K * \varphi] \subseteq [\varphi]^K \subseteq [\varphi]$. The second equality follows from condition (PD). The last equality follows from the fact that $\big\{ w_{\overline{\varphi}} : w \in [K] \big\} = \big\{ r_{\overline{\varphi}} : r \in [\varphi]^K \big\}$, which is implied by the definition of the set of worlds $[\varphi]^K$. ■

Lemma 1, essentially, says that the specification of the $\sqsubseteq_K^{\trianglelefteq}$-minimal φ-worlds, through the differences between worlds of Definition 3, does *not* require the $\overline{\mathcal{L}_\varphi}$-part of the involved worlds.

Against this background, we will show that PD operators, also, respect the following *relevance-sensitive* postulate (C).

$$\textbf{(C)} \quad \text{If } K \cap \mathcal{L}_\varphi = H \cap \mathcal{L}_\varphi, \text{ then } (K * \varphi) \cap \mathcal{L}_\varphi = (H * \varphi) \cap \mathcal{L}_\varphi.$$

Postulate (C) makes an association between the revision policies of two *different* (overlapping) theories. In particular, it states that, if two theories K and H share the same $\mathcal{L}_\varphi$-part, then the revised theories $K * \varphi$ and $H * \varphi$ should, also, share the same $\mathcal{L}_\varphi$-part. Therefore, any beliefs of K, H that are outside the sublanguage $\mathcal{L}_\varphi$ *do not affect* the way that the $\mathcal{L}_\varphi$-parts of K, H are modified — stated otherwise, the *context* of the $\mathcal{L}_\varphi$-parts of K, H does not affect the modification of the $\mathcal{L}_\varphi$-parts of K, H themselves.

It is noteworthy that, since $K \cap \mathcal{L}_\varphi = Cn\big(K \cap \mathcal{L}_\varphi\big) \cap \mathcal{L}_\varphi$, the following identity is, immediately, derived from postulate (C):

$$(K * \varphi) \cap \mathcal{L}_\varphi = \Big(Cn\big(K \cap \mathcal{L}_\varphi\big) * \varphi \Big) \cap \mathcal{L}_\varphi.$$

Theorem 2 proves that PD operators respect postulate (C).

Theorem 2. *PD operators satisfy postulate (C).*

Proof. Let $\trianglelefteq$ be a total preorder over atoms, and let $*$ be a PD operator induced from the family of PD preorders $\{\sqsubseteq_T^{\trianglelefteq}\}_{T \in \mathbb{K}}$. Moreover, let K, H be two theories, and φ be a sentence of $\mathcal{L}$, such that $K \cap \mathcal{L}_\varphi = H \cap \mathcal{L}_\varphi$. If φ is consistent with both K and H, or $\mathcal{L}_\varphi = \mathcal{L}$, then (C) trivially holds. Assume, therefore, that φ contradicts K, H, and $\mathcal{L}_\varphi \subset \mathcal{L}$. Given that $[K]_\varphi = [H]_\varphi$ (as $K \cap \mathcal{L}_\varphi = H \cap \mathcal{L}_\varphi$), $[\varphi]_\varphi^K = [\varphi]_\varphi^H$, $\{w_{\overline{\varphi}} : w \in [K]\} = \{r_{\overline{\varphi}} : r \in [\varphi]^K\}$ and $\{w_{\overline{\varphi}} : w \in [H]\} = \{r_{\overline{\varphi}} : r \in [\varphi]^H\}$, we derive that $min\Big(\{\mathit{Diff}\,(w_\varphi, r_\varphi) : w \in [K]\,,\, r \in [\varphi]^K\}, \trianglelefteq\Big) = min\Big(\{\mathit{Diff}\,(w_\varphi, r_\varphi) : w \in [H]\,,\, r \in [\varphi]^H\}, \trianglelefteq\Big)$. Then, it is not hard to verify that Lemma 1 entails $[K * \varphi]_\varphi = [H * \varphi]_\varphi$; thus, $(K * \varphi) \cap \mathcal{L}_\varphi = (H * \varphi) \cap \mathcal{L}_\varphi$, as desired. ■

7 Conclusion

Parametrized-difference (PD) revision constitutes a well-behaved type of belief revision, which is perfectly-suited for real-world implementations. In this work, we identified some new interesting properties of this type of revision. Specifically, we demonstrated how PD revision is tightly connected with selective revision. Furthermore, we pointed out that PD operators respect a relevance-sensitive postulate, which introduces dependencies between revisions carried out on different (overlapping) belief sets.

Acknowledgments

This research is co-financed by Greece and the European Union (European Social Fund – ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Reinforcement of Postdoctoral Researchers – 2nd Cycle" (MIS-5033021), implemented by the State Scholarships Foundation (IKY).

I am grateful to Pavlos Peppas and Yannis Stamatiou for many fruitful discussions on this work, as well as to the anonymous referees for their valuable input.

References

[1] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[2] Theofanis Aravanis, Pavlos Peppas, and Mary-Anne Williams. Full characterization of Parikh's relevance-sensitive axiom for belief revision. *Journal of Artificial Intelligence Research*, 66:765–792, 2019.

[3] Theofanis Aravanis, Pavlos Peppas, and Mary-Anne Williams. A study of possible-worlds semantics of relevance-sensitive belief revision. *Journal of Logic and Computation*, 30:1125–1142, 2020.

[4] Theofanis Aravanis, Pavlos Peppas, and Mary-Anne Williams. An investigation of parametrized difference revision operators. *Annals of Mathematics and Artificial Intelligence*, 89:7–28, 2021.

[5] Theofanis I. Aravanis. Strengthening parametrized-difference belief revision. *Bulletin of the Hellenic Mathematical Society*, 64:69–91, 2020.

[6] Mukesh Dalal. Investigations into theory of knowledge base revision: Preliminary report. In *Proceedings of 7th National Conference of the American Association for Artificial Intelligence (AAAI 1988)*, pages 475–479, 1988.

[7] Eduardo Fermé and Sven O. Hansson. Selective revision. *Studia Logica*, 63:331–342, 1999.

[8] Peter Gärdenfors. *Knowledge in Flux – Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, Massachusetts, 1988.

[9] Hirofumi Katsuno and Alberto Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3):263–294, 1991.

[10] Rohit Parikh. Beliefs, belief revision, and splitting languages. In Lawrence S. Moss, Jonathan Ginzburg, and Maarten de Rijke, editors, *Logic, Language and Computation*, volume 2, pages 266–278. CSLI Publications, 1999.

[11] Pavlos Peppas. Belief revision. In Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, editors, *Handbook of Knowledge Representation*, pages 317–359. Elsevier Science, 2008.

[12] Pavlos Peppas and Mary-Anne Williams. Kinetic consistency and relevance in belief revision. In *Proceedings of the 15th European Conference on Logics in Artificial Intelligence (JELIA 2016)*, pages 401–414. Springer International Publishing, 2016.

[13] Pavlos Peppas and Mary-Anne Williams. Parametrised difference revision. In *Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning (KR 2018)*, pages 277–286, 2018.

[14] Pavlos Peppas, Mary-Anne Williams, Samir Chopra, and Norman Foo. Relevance in belief revision. *Artificial Intelligence*, 229:126–138, December 2015.

Truth Meets Vagueness

Introduction

Semantic and soritical paradoxes (the paradoxes generated by vague expressions) display remarkable family resemblances. In particular, several logics have been independently applied to both kinds of paradoxes—including many-valued, supervaluational, and non-transitive. These facts have been taken by some authors to suggest that truth and vagueness require a unified logical framework (see e.g. Field 2004). However, there currently is no identification, much less a theory, of what the common features of semantic and soritical paradoxes exactly are. This is what we set out to do in this work. We analyze semantic and soritical paradoxes, we propose a diagnosis of what is the common core of the two phenomena, which we identify in a general form of *indiscernibility*, and we then provide a *theory of paradoxicality*, which formalizes both semantic and soritical paradoxes as arguments involving specific instances of our generalized indiscernibility principle, and correctly predicts which logics can non-trivially solve them. For the sake of concreteness, we focus on four three-valued logics.

1 Truth, vagueness, and paradoxes in three-valued logics

Definition 1.1. *$\mathcal{L}_{\mathrm{t,v}}$ is a first-order language (including predicates* Tr *for truth and* P *for vague properties) that satisfies the following requirements:*

(i) It is possible to define in $\mathcal{L}_{\mathrm{t,v}}$ a function $\ulcorner\ \urcorner$ s.t. for every $\mathcal{L}_{\mathrm{t,v}}$-formula φ, $\ulcorner\varphi\urcorner$ is a closed term.

(ii) For every open $\mathcal{L}_{\mathrm{t,v}}$-formula $\varphi(x)$, there is an $\mathcal{L}_{\mathrm{t,v}}$-term t_φ s.t. $t_\varphi = \ulcorner\varphi(t_\varphi/x)\urcorner$, where $\varphi(t_\varphi/x)$ is the result of uniformly replacing every free occurrence of x with t_φ in φ.

(iii) There is at least one $\mathcal{L}_{\mathrm{t,v}}$-structure $\mathcal{M}$ with support M s.t. (a) M is countable, (b) $\mathcal{M}$ is acceptable,[1] (c) for every $a \in M$ there is an $\mathcal{L}_{\mathrm{t,v}}$-constant c_a.

The truth predicate is often argued to satisfy a property of *naïveté* or *transparency*, to the effect that, for any sentence φ, φ and $\mathsf{Tr}(\ulcorner\varphi\urcorner)$ are always intersubstitutable (in all non-opaque contexts). More precisely, it is required that from ψ one can always infer any formula ψ^{t} that results from ψ by replacing, possibly non-uniformly, a subformula φ of ψ with $\mathsf{Tr}(\ulcorner\varphi\urcorner)$ or *vice versa*.

Transparency famously gives rise to semantic paradoxes. Here is a model-theoretic presentation of the Liar Paradox. Let λ be equivalent to $\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)$ (i.e. a Liar sentence). Suppose there is a classical evaluation v s.t. for every sentence $\varphi \in \mathcal{L}$, $v(\varphi) = v(\mathsf{Tr}(\ulcorner\varphi\urcorner))$. Since v is a classical evaluation, either $v(\lambda) = \mathbf{1}$ or $v(\lambda) = \mathbf{0}$. If $v(\lambda) = \mathbf{1}$, then $v(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)) = \mathbf{1}$ (by definition of λ), but also $v(\neg\lambda) = \mathbf{1}$, which is absurd. We conclude that $v(\lambda) = \mathbf{0}$, and therefore $v(\neg\mathsf{Tr}(\ulcorner\lambda\urcorner)) = \mathbf{0}$ (by definition of λ). But the latter, by naïveté, yields $v(\neg\lambda) = \mathbf{0}$, which is also absurd.

[1] The notion of acceptable model is a standard model-theoretic notion (we do not include the definition for space reasons). For details, see Moschovakis (1974, Chapter 5).

Vague predicates (such as 'rich', 'tall', 'red', ...) are often argued to satisfy a property of *tolerance.* Let P be a vague predicate. Tolerance for P dictates that, if s is P and t is extremely similar to s as far as P is concerned ($s \sim_P t$, for short), then t is P as well. Tolerance can be formalized as follows:

(TOLERANCE$_c$) $\forall x, \forall y(P(x) \wedge x \sim_P y \rightarrow P(y))$

Tolerance also famously gives rise to paradoxes – the *soritical* paradoxes. Suppose there is a classical evaluation v s.t. $v(\mathsf{P}(c_0)) = 1$, $v(c_i \sim_\mathsf{P} c_{i+1}) = 1$ for every i, and $v(\forall x \forall y(\mathsf{P}(x) \wedge x \sim_\mathsf{P} y \rightarrow \mathsf{P}(y))) = 1$. Since v is a classical evaluation, $v(\forall x \forall y(\mathsf{P}(x) \wedge x \sim_\mathsf{P} y \rightarrow \mathsf{P}(y))) = 1$ entails that $v(\mathsf{P}(c_0) \wedge c_0 \sim_\mathsf{P} c_1 \rightarrow \mathsf{P}(c_1)) = 1$, and since $v(\mathsf{P}(c_0)) = 1$ and $v(c_0 \sim_\mathsf{P} c_1) = 1$, also $v(\mathsf{P}(c_1))) = 1$. By induction, this establishes that for every n, $v(\mathsf{P}(c_n)) = 1$.

In order to avoid both semantic and soritical paradoxes, several authors have advocated the use of some non-classical logic. Here we focus in particular on *three-valued* logics.[2]

Definition 1.2. *A* partial model $\mathcal{M}$ *is a pair* $\langle M, f\rangle$*, where* M *is a non-empty set and* f *is a multi-function from closed* $\mathcal{L}_{\mathrm{t,v}}$*-terms to* M *and from atomic* $\mathcal{L}_{\mathrm{t,v}}$*-sentences to the set* $\{0, 1/2, 1\}$.

Definition 1.3. *For every partial model* $\mathcal{M} = \langle M, f\rangle$*, the* strong Kleene evaluation *induced by* $\mathcal{M}$ *is the function* $e_\mathcal{M}$ *from sentences to* $\{0, 1/2, 1\}$ *such that:*

$$\begin{aligned}
e_\mathcal{M}(R(t_0, \dots, t_n)) &:= f(R(t_0, \dots, t_n)) \\
e_\mathcal{M}(\neg\varphi) &:= 1 - e_\mathcal{M}(\varphi) \\
e_\mathcal{M}(\varphi \wedge \psi) &:= \min(e_\mathcal{M}(\varphi), e_\mathcal{M}(\psi)) \\
e_\mathcal{M}(\forall x \varphi(x)) &:= \inf\{e_\mathcal{M}(\varphi(t)) \in \{0, 1/2, 1\} \mid t \textit{ is a closed term}\}
\end{aligned}$$

Definition 1.4. *For every set of sentences* Γ*, an evaluation* e *makes* Γ S-true *if for every* $\varphi \in \Gamma$, $e(\varphi) = 1$*, and* T-true *if for every* $\varphi \in \Gamma$, $e(\varphi) \geq 1/2$.

Definition 1.5. SS*,* TT*,* ST*, and* TS

- Γ SS*-entails* φ *(in symbols* $\Gamma \models_{\mathsf{SS}} \varphi$*) if for every partial model* $\mathcal{M} = \langle D, f\rangle$*, every evaluation* $e_\mathcal{M}$ *induced by* $\mathcal{M}$ *that makes all the sentences in* Γ S*-true, also makes* φ S*-true.*

- Γ TT*-entails* φ *(in symbols* $\Gamma \models_{\mathsf{TT}} \varphi$*) if for every partial model* $\mathcal{M} = \langle D, f\rangle$*, every evaluation* $e_\mathcal{M}$ *induced by* $\mathcal{M}$ *that makes all the sentences in* Γ T*-true, also makes* φ T*-true.*

- Γ TS*-entails* φ *(in symbols* $\Gamma \models_{\mathsf{TS}} \varphi$*) if for every partial model* $\mathcal{M} = \langle D, f\rangle$*, every evaluation* $e_\mathcal{M}$ *induced by* $\mathcal{M}$ *that makes all the sentences in* Γ T*-true, also makes* φ S*-true.*

- Γ ST*-entails* φ *(in symbols* $\Gamma \models_{\mathsf{ST}} \varphi$*) if for every partial model* $\mathcal{M} = \langle D, f\rangle$*, every evaluation* $e_\mathcal{M}$ *induced by* $\mathcal{M}$ *that makes all the sentences in* Γ S*-true, also makes* φ T*-true.*

We now use SS, TT, TS, and ST to formulate theories of truth and vagueness. In order to include a treatment of truth-theoretical sentences, we move from a starting partial model $\mathcal{M} = \langle M, f\rangle$ to a *Kripke model*: a triple $\langle M, f, S\rangle$, where S is the *extension* of the truth predicate, so that $\langle M, f, S\rangle$ satisfies the transparency requirement. The model-theoretic construction is due to Kripke (1975) (we won't reproduce it here). Let's associate a strong Kleene transparent evaluation to a Kripke model.

[2] See e.g. (Kripke 1975, Priest 1979, Field 2008, Beall 2009, Smith 2008, Cobreros et al. 2012).

Definition 1.6. *For every Kripke model $\mathcal{M} = \langle M, f, S\rangle$ for $\mathcal{L}_{\text{t,v}}$, the* Kripke (strong Kleene) evaluation *induced by $\mathcal{M}$ is the function e from sentences to $\{0, 1/2, 1\}$ s.t.:*

$$e_{\mathcal{M}}(\varphi) := \begin{cases} 1, & \textit{if } \varphi \in S \\ 0, & \textit{if } \neg\varphi \in S \\ 1/2, & \textit{otherwise} \end{cases}$$

Lemma 1.7. *For every Kripke model $\mathcal{M}$, the evaluation $e_{\mathcal{M}}$ is a strong Kleene evaluation, and it validates a form of naïveté, i.e. for every $\varphi \in \mathcal{L}_{\text{t,v}}$ and every truth-theoretic substitution φ^{t}:*

$$e_{\mathcal{M}}(\varphi) = e_{\mathcal{M}}(\varphi^{\text{t}})$$

Definition 1.8. SSTT, TTTT, STTT, *and* TSTT

- Γ SSTT-*entails φ (in symbols $\Gamma \models_{\mathsf{SSTT}} \varphi$) if for every Kripke model $\mathcal{M}$, if the Kripke evaluation $e_{\mathcal{M}}$ induced by $\mathcal{M}$ makes all the sentences in Γ* S*-true, it also makes φ* S*-true.*
- Γ TTTT-*entails φ (in symbols $\Gamma \models_{\mathsf{TTTT}} \varphi$) if for every Kripke model $\mathcal{M}$, if the Kripke evaluation $e_{\mathcal{M}}$ induced by $\mathcal{M}$ makes all the sentences in Γ* T*-true, it also makes φ* T*-true.*
- Γ TSTT-*entails φ (in symbols $\Gamma \models_{\mathsf{TSTT}} \varphi$) if for every Kripke model $\mathcal{M}$, if the Kripke evaluation $e_{\mathcal{M}}$ induced by $\mathcal{M}$ makes all the sentences in Γ* T*-true, it also makes φ* S*-true.*
- Γ STTT-*entails φ (in symbols $\Gamma \models_{\mathsf{STTT}} \varphi$) if for every Kripke model $\mathcal{M}$, if the Kripke evaluation $e_{\mathcal{M}}$ induced by $\mathcal{M}$ makes all the sentences in Γ* S*-true, it also makes φ* T*-true.*

Proposition 1.9. *Letting φ^{t} be as above (a result of a truth-theoretical substitution within φ):*

- *For every $\varphi \in \mathsf{Sent}_{\mathcal{L}_{\text{t,v}}}$, $\varphi \models_{\mathsf{SSTT}} \varphi^{\text{t}}$, $\varphi \models_{\mathsf{TTTT}} \varphi^{\text{t}}$, and $\varphi \models_{\mathsf{STTT}} \varphi^{\text{t}}$.*
- *For every $\Gamma \cup \{\varphi\} \subseteq \mathsf{Sent}_{\mathcal{L}_{\text{t,v}}}$, if $\Gamma \models_{\mathsf{TSTT}} \varphi$, then $\Gamma \models_{\mathsf{TSTT}} \varphi^{\text{t}}$.*

We now consider the applications of strong Kleene semantics and of the four resulting logics (SS, TT, TS, and ST) to vague predicates. A three-valued model $\mathcal{M} = \langle M, f\rangle$ is called *soritical* if:

(a) $e_{\mathcal{M}}(\mathsf{P}(c_0)) = 1$.

(b) There is an individual c_j s.t. $e_{\mathcal{M}}(\mathsf{P}(c_j)) = 1/2$.

(c) There is an individual c_n s.t. $e_{\mathcal{M}}(\mathsf{P}(c_n)) = 0$.

(d) For every q, $e_{\mathcal{M}}(c_q \sim_{\mathsf{P}} c_{q+1}) = 1$.

(e) $e_{\mathcal{M}}(\mathsf{P}(c_q)) \geq e_{\mathcal{M}}(\mathsf{P}(c_r))$ just in case $q \leq r$.

Definition 1.10. SSV, TTV, STV, *and* TSV

- Γ SSV-*entails φ (in symbols $\Gamma \models_{\mathsf{SSV}} \varphi$) if for every soritical model $\mathcal{M}$ and every induced evaluation $e_{\mathcal{M}}$, if $e_{\mathcal{M}}$ makes all the sentences in Γ* S*-true, it also makes φ* S*-true.*
- Γ TTV-*entails φ (in symbols $\Gamma \models_{\mathsf{TTV}} \varphi$) if for every soritical model $\mathcal{M}$ and every induced evaluation $e_{\mathcal{M}}$, if $e_{\mathcal{M}}$ makes all the sentences in Γ* T*-true, it also makes φ* T*-true.*

- Γ TSV-*entails* φ *(in symbols* $\Gamma \models_{\mathsf{TSV}} \varphi$*) if for every soritical model* $\mathcal{M}$ *and every induced evaluation* $e_{\mathcal{M}}$*, if* $e_{\mathcal{M}}$ *makes all the sentences in* Γ T-*true, it also makes* φ S-*true.*

- Γ STV-*entails* φ *(in symbols* $\Gamma \models_{\mathsf{STV}} \varphi$*) if for every soritical model* $\mathcal{M}$ *and every induced evaluation* $e_{\mathcal{M}}$*, if* $e_{\mathcal{M}}$ *makes all the sentences in* Γ S-*true, it also makes* φ T-*true.*

Proposition 1.11.

- TTV *and* STV *are tolerant logics. For every vague predicate* P*:*

$$\models_{\mathsf{TTV}} \forall x \forall y(\mathsf{P}(x) \land x \sim_{\mathsf{P}} y \to \mathsf{P}(y)) \qquad \models_{\mathsf{STV}} \forall x \forall y(\mathsf{P}(x) \land x \sim_{\mathsf{P}} y \to \mathsf{P}(y))$$

- SSV *and* TSV *are intolerant logics. For every vague predicate* P*:*

$$\not\models_{\mathsf{SSV}} \forall x \forall y(\mathsf{P}(x) \land x \sim_{\mathsf{P}} y \to \mathsf{P}(y)) \qquad \not\models_{\mathsf{TSV}} \forall x \forall y(\mathsf{P}(x) \land x \sim_{\mathsf{P}} y \to \mathsf{P}(y))$$

2 Unifying the paradoxes

In order to 'unify' the paradoxes in the sense described above, we expand the theory developed in Rossi (2019). The basic idea of Rossi (2019), in a nutshell, is the following. Each sentence φ is analyzed individually. The components of φ are identified, and used to define an equation system. The possible solutions to the equation system yields the possible values of φ, given (i) a base model (for the base vocabulary), (ii) an evaluation for logically complex sentences, and (iii) an evaluation for the truth-theoretic (and, here) also the vague vocabulary. A sentence φ is *paradoxical* if it cannot receive a unique classical value, and *classical* otherwise.

Let $\mathsf{N}_3 = \{0, 1/2, 1\}$. Let's fix the language we will use to assign equations to formulas of $\mathcal{L}_{\mathrm{t,v}}$.

Definition 2.1. *Let* $\mathcal{L}_3$ *be the language whose alphabet comprises the following sets of symbols:*

- *a countable set* Var_3 *of variables* $\{v_{\varphi_1}, \ldots, v_{\varphi_n}, \ldots\}$*, where each* φ_k *is the k-th element in a non-redundant enumeration of sentences of* $\mathcal{L}_{\mathrm{t,v}}$*;*

- *a set of constants* Con_3 *containing an individual constant for every element in* N_3*;*

- *a binary relation = for equality.*

Let the set of terms and the set of atomic formulas of $\mathcal{L}_3$ *be defined by the following clauses:*

- *the set of terms of* $\mathcal{L}_3$ *is built by recursively closing off* $\mathsf{Var}_3 \cup \mathsf{Con}_3$ *under the operations* $(1 - x)$, $\min(x, y)$*, and* $\inf\{x_1, x_2, \ldots, x_n, \ldots\}$ *employed in Definition 1.3;*

- *atomic formulas of* $\mathcal{L}_3$ *are* $s = t$ *where s and t are* $\mathcal{L}_3$*-terms; we denote their set as* $\mathbb{E}_3$*.*

Let boldface lower-case letters $\mathbf{e}$ vary over elements of $\mathbb{E}_3$, while capital letters $\mathbf{E}$ vary over elements of $\mathcal{P}(\mathbb{E}_3)$. For $\mathbf{E} \subseteq \mathbb{E}_3$, let $\mathsf{Var}(\mathbf{E})$ indicate the collection of $\mathcal{L}_3$-variables of formulas in $\mathbf{E}$.

Definition 2.2. *A* 3-valued semantics for $\mathcal{L}_{\mathrm{t,v}}$ *is a structure* S_3 *given by* $S_3 = \langle \mathsf{N}_3, \mathbb{E}_3, e, A \rangle$*, where* N_3 *and* $\mathbb{E}_3$ *are as above, and*

- $e : \mathsf{Form}_{\mathcal{L}_{\mathrm{t,v}}} \longmapsto \mathcal{P}(\mathbb{E}_3)$ *obeys the clauses from Definition 1.3;*[3]

[3]We refer the interested reader to Rossi (2019, §4) for the construction of the function e.

- *A is a (possibly infinite) set of functions α which are* assignments of values *in* N_3 *to variables in any set* $\mathsf{Var}(\{\mathbf{e}\})$ *for* $\mathbf{e} \in \mathbb{E}_3$*; that is,* $\alpha : \{Var(\{\mathbf{e}\}) \mid \mathbf{e} \in \mathbb{E}_3\} \longmapsto \mathsf{N}_3$.

For every assignment α and for every $\mathbf{e}$ in $\mathbb{E}_3$, let $\models^{\alpha} \mathbf{e}$ indicate that $\mathbf{e}$ is a true arithmetical equation under the assignment α of values in N_3 to its variables. So, $\models^{\alpha} \mathbf{e}$ holds if $\alpha(\mathbf{e})$ is a true arithmetical identity. Let also put, for every assignment α and for every $\mathbf{E} \subseteq \mathbb{E}_3$, $\alpha(\mathbf{E}) = \{\alpha(\mathbf{e}) \mid \mathbf{e} \in \mathbf{E}\}$, and put $\models^{\alpha} \mathbf{E}$ if and only if $\models^{\alpha} \mathbf{e}$ for every $\mathbf{e} \in \mathbf{E}$. We can now use the existence of solutions to $\mathcal{L}_3$-equations to provide a generalized notion of satisfiability, which we will use to model paradoxical arguments

Definition 2.3. *Let $S_3 = \langle N_3, \mathbb{E}_3, e, A\rangle$ be a semantic structure.*

- *A set $\mathbf{E} \subseteq \mathbb{E}_3$ is* solvable in S_3 *if and only if there exists an assignment $\alpha \in A$ such that $\models^{\alpha} \mathbf{E}$.*
- *An $\mathcal{L}_{\mathrm{t,v}}$-sentence φ is* satisfiable in S_3 *if and only if $e(\varphi)$ is solvable.*
- *A set Γ of $\mathcal{L}_{\mathrm{t,v}}$-sentences is* satisfiable in S_3 *if and only if all sentences φ of Γ are.*

Definition 2.4.

- *An $\mathcal{L}_{\mathrm{t,v}}$-sentence φ is* $\mathsf{S}(\mathsf{T})$*-true in S_3 in short, if $\alpha(v_\varphi) = 1$ ($\alpha(v_\varphi) \geq 1/2$).*
- *A set Γ of $\mathcal{L}_{\mathrm{t,v}}$-sentences is* $\mathsf{S}(\mathsf{T})$*-true in S_3 in short, if for every $\varphi \in \Gamma$, $\alpha(v_\varphi) = 1$ ($\alpha(v_\varphi) \geq 1/2$).*

Definition 2.5. *Let Γ be a set of formulas of $\mathcal{L}_{\mathrm{t,v}}$, and let φ be also a formula of it. Let $\vdash$ be the symbol we use for logical consequence in some given theory. Then, we say that:*

- $\Gamma \vdash \varphi$ *is* SS-satisfiable in S_3 *if for every assignment α that makes Γ S-true, α makes φ S-true;*
- $\Gamma \vdash \varphi$ *is* ST-satisfiable in S_3 *if for every assignment α that makes Γ S-true, α makes φ T-true;*
- $\Gamma \vdash \varphi$ *is* TS-satisfiable in S_3 *if for every assignment α that makes Γ T-true, α makes φ S-true;*
- $\Gamma \vdash \varphi$ *is* TT-satisfiable in S_3 *if for every assignment α that makes Γ T-true, α makes φ T-true;*

Proposition 2.6. *Let $\Gamma_\lambda := \{\lambda, \lambda \leftrightarrow \neg\mathsf{Tr}(\lambda), \lambda \leftrightarrow \mathsf{Tr}(\lambda)\}$. $\Gamma_\lambda \vdash \bot$ is NM-unsatisfiable for* $\mathsf{N}, \mathsf{M} \in \{\mathsf{S}, \mathsf{T}\}$.

Proposition 2.7. *Let Γ_σ be the following set of sentences of $\mathcal{L}_{\mathrm{t,v}}$:*

$$\Gamma_\sigma = \left\{ \begin{array}{l} \mathsf{P}(a_0), \\ (\mathsf{P}(a_0) \wedge a_0 \sim_{\mathsf{P}} a_1) \rightarrow \mathsf{P}(a_1), \ldots, (\mathsf{P}(a_{n-2}) \wedge a_{n-2} \sim_{\mathsf{P}} a_{n-1}) \rightarrow \mathsf{P}(a_{n-1}), \\ a_0 \sim_{\mathsf{P}} a_1, \ldots, a_{n-2} \sim_{\mathsf{P}} a_{n-1} \end{array} \right\}$$

$\Gamma_\sigma \vdash \mathsf{P}(a_n)$ *is* TT- *and* ST-*unsatisfiable, and is only vacuously* SS- *and* TS-*satisfiable, i.e. is only* SS- *and* TS-*satisfiable by strong Kleene evaluations which assign value* 1 *to every* $\mathsf{P}(a_i)$.

We have obtained a unification of the paradoxes of truth and vagueness in the sense that both semantic and soritical paradoxes become satisfiable or unsatisfiable arguments in the same logics. But can we further reduce the 'engines' of the paradoxes – that is, transparency and tolerance – to a common source? Consider the following principle:

(∼-Ind) $\forall x \forall y (x \sim y \rightarrow (\varphi(x) \leftrightarrow \varphi(y)))$

where $\sim$ is a binary relation on terms, and $\varphi(s)$ is a schematic formula featuring at least one occurrence of s. It seems to us that ($\sim$-IND) should be understood as a necessary, albeit not sufficient, condition for a pair $\langle\sim,\varphi\rangle$ to be a similarity relation, and a formula about that similarity relation, respectively.

Our claim, simply put, is that both naïveté and tolerance can be understood as forms of indiscernibility. More precisely:

(i) one can find a relation $\sim_{To}$ and a formula $\varphi_{To}(x)$ of $\mathcal{L}_{\mathrm{t,v}}$ such that tolerance follows from ($\sim_{To}$-IND$^{\varphi To}$), which is the instance of ($\sim$-IND) obtained by substituing $\sim_{To}$ for $\sim$, and φ_{To} for φ;

(ii) that one can find a relation $\sim_{Tr}$ and a formula φ_{Tr} of $\mathcal{L}_{\mathrm{t,v}}$ such that transparency follows from ($\sim_{Tr}$-IND$^{\varphi Tr}$), which is the instance of ($\sim$-IND) obtained by substituing $\sim_{Tr}$ for $\sim$, and φ_{Tr} for φ.

Claim (i) is immediate, while claim (ii) is much less evident. First, it is not clear that a similarity relation somehow connected to the use of the truth predicate can be found. Second, it is even less clear that this relation might turn out to allow to deduce tolerance from indiscernibility. Despite the appearances though, such a relation can be found.

Let $\sim_{Tr}$ be a relation such that $s \sim_{Tr} t$ holds if and only if either $s = \ulcorner\varphi\urcorner$ and $t = \ulcorner\mathsf{Tr}(\ulcorner\varphi\urcorner)\urcorner$, or $t = \ulcorner\psi\urcorner$ and $s = \ulcorner\mathsf{Tr}(\ulcorner\psi\urcorner)\urcorner$. Let also, for every acceptable model $\mathcal{M}$, $\mathfrak{r}(x)$ be the $\mathcal{M}$-definable function such that $\mathfrak{r}(\ulcorner\psi\urcorner) = \psi$ for every formula ψ of $\mathcal{L}_{\mathrm{t,v}}$. Finally, let, for every term s of $\mathcal{L}_{\mathrm{t,v}}$,

$$\varphi_{Tr}(x/s) :\equiv \Big((s = \ulcorner\underset{\cdot}{\mathfrak{r}}(s)\urcorner) \wedge \mathfrak{r}(s)\Big)$$

where $\underset{\cdot}{\mathfrak{r}}$ denotes the $\mathcal{L}_{\mathrm{t,v}}$-term that strongly represents the coding function $\mathfrak{r}$.[4] In all cases in which s is the code of a sentence, i.e., such that $s = \ulcorner\psi\urcorner$ for some $\psi \in \mathcal{L}_{\mathrm{t,v}}$

$$\varphi_{Tr}(x/s) :\equiv \Big((\ulcorner\psi\urcorner = \ulcorner\underset{\cdot}{\mathfrak{r}}(\ulcorner\psi\urcorner)\urcorner) \wedge \psi\Big)$$

Notice that $\varphi_{Tr}(x/s)$ is then equivalent to ψ. Take the instance ($\sim_{Tr}$-IND$^{\varphi Tr}$). It is clear that, for every pair of terms s and t of $\mathcal{L}_{\mathrm{t,v}}$, the following formula is deducible by logic from ($\sim_{Tr}$-IND$^{\varphi Tr}$):

$$s \sim_{Tr} t \rightarrow (\varphi_{Tr}(x/s) \leftrightarrow \varphi_{Tr}(y/t))$$

Assuming that, for some $\psi \in \mathcal{L}_{\mathrm{t,v}}$, $s = \ulcorner\psi\urcorner$ and $t = \mathsf{Tr}(\ulcorner\psi\urcorner)$, then:

$$\begin{aligned}&\ulcorner\psi\urcorner \sim_{\mathsf{Tr}} \ulcorner\mathsf{Tr}(\ulcorner\psi\urcorner)\urcorner \rightarrow \\ &\qquad\Big(\underbrace{((\ulcorner\underset{\cdot}{\mathfrak{r}}(\ulcorner\psi\urcorner)\urcorner = \ulcorner\psi\urcorner) \wedge \psi)}_{\varphi_{\mathsf{Tr}}(s/x)} \leftrightarrow \underbrace{((\ulcorner\underset{\cdot}{\mathfrak{r}}(\ulcorner\mathsf{Tr}(\ulcorner\psi\urcorner)\urcorner)\urcorner = \ulcorner\mathsf{Tr}(\ulcorner\psi\urcorner)\urcorner) \wedge \mathsf{Tr}(\ulcorner\psi\urcorner))}_{\varphi_{\mathsf{Tr}}(t/x)}\Big)\end{aligned}$$

$\varphi_{Tr}(x/s)$ is then equivalent to ψ, and $\varphi_{Tr}(x/t)$ is then equivalent to $\mathsf{Tr}(\ulcorner\psi\urcorner)$. Thus, for every s and t s.t. $s \sim_{Tr} t$, the corresponding instance of transparency is provable, as wanted.

References

Beall, J.: 2009, *Spandrels of Truth*, Oxford University Press, Oxford.

Cobreros, P., Egré, P., Ripley, D. and van Rooij, R.: 2012, Tolerant, classical, strict, *Journal of Philosophical Logic* **41**(2), 347–85.

[4] For details on the dot notation, see Feferman (1991).

Feferman, S.: 1991, Reflecting on Incompleteness, *Journal of Symbolic Logic* **56**(1), 1–49.

Field, H.: 2004, The semantic paradoxes and the paradoxes of vagueness, *in* J. Beall (ed.), *Liars and heaps: New essays on paradox*, Oxford University Press.

Field, H.: 2008, *Saving Truth from Paradox*, Oxford University Press, Oxford.

Kripke, S.: 1975, Outline of a theory of truth, *Journal of Philosophy* **72**, 690–716.

Moschovakis, Y.: 1974, *Elementary Induction on Abstract Structures*, North-Holland and Elsevier, Amsterdam, London and New York.

Priest, G.: 1979, The logic of paradox, *Journal of Philosophical Logic* **8**, 219–241.

Rossi, L.: 2019, A unified theory of truth and paradox, *Review of Symbolic Logic* **12**(2), 209–254.

Smith, N. J. J.: 2008, *Vagueness and Degrees of Truth*, Oxford University Press, Oxford.

The V-logic Multiverse and MAXIMIZE

Matteo de Ceglie

Universität Salzburg
Salzburg, Österreich
decegliematteo@gmail.com

Abstract

In this paper, I show that $ZFC + LCs$ is restrictive compared to the V-logic multiverse, characterised as $ZFC + LCs$+multiverse axioms. This means showing that the V-logic multiverse proves the existence of an extra object that it is unavailable in $ZFC + LCs$ and that, in turn, this object realises a new isomorphism type. I argue that such an object is a class-iterable sharp, that can only be found if there are proper, uncountable, width extensions of V. Such extensions are present in the V-logic multiverse, but not in classical set theory.

1 Introduction

Classical set theory (ZFC), as instantiated by the cumulative hierarchy V, has been a very successful foundation of mathematics for over a century. Nonetheless, there are some problems with it. The first, and foremost, problem is Gödel's Incompleteness: in the context of set theory, it entails that some set theoretic statements are *independent* from ZFC. This means that we cannot prove neither that they are true, nor that they are false. The main example of such statements is the Continuum Hypothesis (CH), as proved by Cohen in the 1960s using forcing. For some time it was thought that adding new axioms to ZFC would solve this problem (Gödel's program), with large cardinals axioms (LCs) being the main candidate for addition. However, it was proved that they do not settle CH and, moreover, that they are incompatible with the Axiom of Constructibility.[1] In the end, the impossibility of solving independent questions by the simple addition of new axioms, coupled with the multiplication of incompatible models generated through forcing, led to several possible expansions of ZFC, each one giving rise to interesting mathematics. But which one should be chosen as the new foundations of mathematics? The set theoretic multiverse (as introduced by J. D. Hamkins (2012)) side-steps the question: there is no need to choose, we can integrate all these different set theories in just one *multiverse conception*. Such a solution is not appealing to the advocates of *universism*, that instead defend the idea that there is only one set theoretic universe, V, that contains all the possible sets and cannot be further expanded. Moreover, they point out that everything that can be done in the multiverse can actually already be done in the Single Universe. For example, they argue that the main tool used to generated new universes, i.e. forcing, can be interpreted as taking place entirely within the Single Universe.[2] The only possible argument against this kind of universism and in favour of multiversism entails proving that in the set theoretic multiverse we can some object that we cannot have in the Single Universe at all.

In this paper, I argue exactly that. In particular, I contend that classical set theory, $ZFC(+LCs)$, is *restrictive* compared to the V-logic multiverse (a novel set theoretic multiverse conception developed by the present author and Claudio Ternullo). This multiverse conception is based upon Friedman's Hyperuniverse[3] and Steel's set-generic multiverse[4]: like the Hyperuniverse, it uses the infinitary V-logic as background logic (this logic admits formulas of length less than the first successor of the least inaccessible cardinal, but only a finite block of quantifiers in front of them) and admits all kinds of outer models of V (produced by set-generic, class-generic, hyperclass forcing, etc.). Like

[1] The Axiom of Constructibility says that $V = L$, i.e. that all sets are constructible from simpler ones.

[2] Here and throughout the paper I will refer to models of set theory and universes interchangeably, as done in the literature.

[3] See S. Friedman (2016).

[4] See Steel (2014).

Steel's set-generic multiverse, it is recursively axiomatisable and is rooted on a ground universe that satisfies ZFC. For this proof, I compare $ZFC + LCs$ and the V-logic multiverse, characterised as $ZFC + LCs+$ the multiverse axioms, following Maddy's methodological principle MAXIMIZE (as introduced in Maddy (1997)). According to this principle, when choosing between two theories T and S we should prefer the one that can prove more isomorphism types. I claim that the V-logic multiverse, as opposed to $ZFC + LCs$, does exactly that. This is because in the V-logic multiverse theory we can prove the existence of proper, uncountable, extensions of V, that we cannot have in $ZFC + LCs$ (see Neil Barton (2019)). In turn, this extra object means we can realise more isomorphism types, since in the V-logic multiverse we can prove the existence of iterable class sharps and, more importantly, maps between them (see Antos, N. Barton, and S.-D. Friedman (nd)). Moreover, when moving from $ZFC + LCs$ to the V-logic Multiverse we are not losing anything: ZFC, all the large cardinals, inner models and V are still there. On the other hand, when moving from the V-logic multiverse to $ZFC + LCs$ we lose the actual outer models of V, iterable class sharps and iterable class sharp generated models. Thus, this latter theory is restrictive compared to the V-logic multiverse theory.

This paper is structured as follows. First, I describe the infinitary V-logic and the V-logic multiverse (section 2). After that, I show that classical set theory is restrictive compared to the V-logic multiverse (section 3). Finally, some concluding remarks sketching the road ahead end the paper (section 4).

2 The V-logic Multiverse

I now proceed to mathematically describe the V-logic multiverse. The system to be adopted allows to address universes arising from extending V in width and height.[5] More specifically, it is able to:

1. code representations of the "canonical" relationship between V and its outer models;
2. incorporate all kinds of outer-model constructs (e.g. extensions produced by various kinds of forcing);
3. formulate what one should easily acknowledge as multiverse axioms.

This multiverse conception is, philosophically, a refinement of Hamkins' broad multiverse. The key difference is that, instead of admitting all the possible universes, without any hierarchy (as done in Hamkins' multiverse), the V-logic Multiverse only admits the universes that can be defined and described in a certain, uniform way. While this is philosophical starting point aims at restricting Hamkins' philosophical conception (goal shared with other multiverses), the mathematical implementation ends up being more open.[6] In order to satisfy these requirements, I adopt the infinitary V-logic, i.e. a logic whose language $\mathcal{L}_{\kappa^+,\omega}$ is that of first-order logic, admitting formulas of length less than κ^+ (the first successor of the least inaccessible cardinal) and with a finite number (less than ω) of quantifiers in front, and supplemented with the membership relation symbol $\in$ and the following constant symbols:

- $\bar{a}$, one for each $a \in V$;
- $\bar{V}$, denoting the *ground universe* (that is, our initial V).

Proofs in V-logic are infinitary, because of the addition of the following inference rules::

Set-rule $\{\varphi(\bar{b}) | b \in a\} \vdash (\forall x \in \bar{a})\varphi(x)$

[5] An *height* extension of V is produced by adding new ordinals, while a *width* extension by adding new subsets.

[6] Hamkins's multiverse, as implemented by the axioms introduced in Gitman and Joel David Hamkins (2011), is composed by only the countable computably saturated models of ZFC, while the V-logic multiverse axioms admits any kind of model.

V**-rule** $\{\varphi(\bar{a})|a \in V\} \vdash (\forall x)\varphi(x)$

A sentence of V-logic may also use additional symbols.[7] For example, a case of special interest is when $\bar{W}$ is introduced as a new predicate symbol (variable) ranging over "generic *outer* models of V", and one considers sentences of the form "$\bar{W} \models ZFC + \psi$" for some sentence ψ, possibly containing constants $\bar{a}$ for $a \in V$. The following fact is fundamental for my purposes:

Fact 1 (Barwise)**.** *If $\bar{V}$ is countable then a theory T of $\mathcal{L}_V$ has a $\bar{W}$-structure for a model iff T is consistent in V-logic.*

This means that, in V-logic, one can produce a sentence about *any* outer model W of V which, if consistent in V-logic, then *really* expresses a property of an outer model W of V.[8] As Barwise has shown, it turns out that structures satisfying the axioms of $\mathfrak{M}$-logic, that is, $\mathfrak{M}$-structures (V-structures, in the case of V-logic), are models of Kripke-Platek set theory (KP), a weak fragment of ZFC.[9] In turn, models of KP are called "admissible sets" (as these models are related to admissible ordinals[10] in recursion theory). The least such model, which contains $\mathfrak{M}$ as an object, is called $\mathfrak{M}^+$.[11] If we turn to consider V-logic, a V-structure is the *least admissible set* beyond V, that is the least model of KP containing V as an object, which is called V^+. In V^+, we finally have codes for proofs in V-logic, which allows one to express *syntactic* facts that are essential for axiomatising the V-logic multiverse.

I start by considering a set-theoretic sentence φ, which expresses, in V-logic, that "$\bar{W} \models \psi$". Now, consider the theory $T = ZFC +$ "$\bar{W} \models \psi$" and let Con(T) be the statement "T is consistent" (in V-logic).[12] Then, by Fact 1 above, if $Con(T)$ holds, then $\bar{W}$ really is an outer model W of V enjoying the property ψ. This outer model, identified by Con(T), is, thus, a member of the V-logic multiverse. The procedure may be generalised to *all* kinds of ψ and all kinds of outer models W,[13] which leads to the formulation of the first, and key, "multiverse axiom" of the new theory $MZFC$ – more precisely, an *axiom schema*:

Axiom 1 (Multiverse Axiom Schema)**.** *For any first-order φ with parameters from V, if the sentence of V-logic expressing "$\bar{W}$ is an outer model of V satisfying φ" is consistent in V-logic, then there is a universe W which is an outer model of V that satisfies φ.*[14]

In addition to this axiom schema, and in analogy with ZFC, $MZFC$ also features axioms describing how the *sets* and *universes* of the multiverse behave.[15] To this end, $MZFC$ contains all of ZFC, together with the following axiom:

Axiom 2 (Core Axiom)**.** *Every universe of the multiverse models ZFC.*

Thus, $MZFC$, as of now, consists of:

1. ZFC;
2. the *Multiverse Axiom Schema*;
3. additional V-logic "axioms" (the "V-rule" and the "set rule");

[7] The general features of infinitary logics, and their relationships with admissible sets (structures) are discussed in, among other works, Keisler (1974), Barwise (1975), and Dickmann (1975).

[8] This was proved in Barwise (1975).

[9] KP is the theory which results from removing the Power-Set and Infinity Axioms from ZF, and admitting restricted forms of the Separation Axiom (Δ_0-Separation) and of the Replacement Axiom (Δ_0-Replacement).

[10] An ordinal α is admissible iff the corresponding constructible universe L_α is a transitive model of KP.

[11] Barwise's original notation is $Hyp(\mathfrak{M})$, but to avoid confusion, the alternative notation $\mathfrak{M}^+$ is preferable.

[12] Technically, Con(T) is the V-logic statement: "$T \nvdash_V \varphi \wedge \neg\varphi$", where '$\nvdash_V$' is the V-logic provability relation.

[13] For instance, Con(T) above may further specify that $\bar{W}$ contains a filter $\bar{G}^C \subseteq \mathbb{P}^C$, where $\mathbb{P}^C$ is a class-poset, which would mean that $\bar{W}$ is a class-forcing extension of V.

[14] This also means that each φ consistent in V-logic has a model in the multiverse. Note that the consistency of such a φ is Π_1-expressible in a first-order way, not over V but over V^+.

[15] Of course, there are distinct variables for sets and universes in our multiverse theory.

4. the *Core Axiom*.

Although the multiverse axioms clearly describe semantic constructs, what we have at this stage is just a collection of theories. In particular, if we take an incremental approach to $MZFC$, we may informally view it as a tree made up of *branches* corresponding to alternative set-theoretic statements, and of *nodes* where alternative V-logic theories extending ZFC appear. Thus, the V-logic multiverse may be seen as *the collection of all the combinatorially conceivable consistent V-logic theories of outer models.*

Note that the V-logic multiverse maximises over outer-model constructs, as no constraint upon the nature of the outer models has been placed in the formulation of the Multiverse Axiom Schema. I argue that this represents a significant improvement over the set-generic multiverse conception, which only allows for *certain kinds* of outer models.[16]

3 $ZFC+LCs$ is restrictive compared to the V-logic multiverse

In this section I present the result that I have hinted to in the Introduction: classical set theory (ZFC) is *restrictive* compared to the V-logic Multiverse. To do so I use the methodological principle MAXIMIZE as discussed by Maddy (1997). This principle states that when comparing two theories, the one that proves more isomorphism types is preferable. The theory that proves more isomorphism types *maximizes* over the other (or, equivalently, the theory that proves less isomorphism types is *restrictive* compared to the other). Maddy (1997) applies this principle to argue against the addition of $V = L$ to ZFC, and I plan to apply the same line of reasoning to the V-logic Multiverse.

The argument consists of the following steps:

1. first of all, prove that one theory proves the existence of an extra object that cannot exists in the (claimed) restrictive one;

2. prove that this extra object realises a new isomorphism type;

3. if the two above steps are done, then we can conclude that one theory *maximizes* (in Maddy's sense) over the other (or, equivalently, that one theory is *restrictive* over the other).

I contend that this is true in the case of the V-logic multiverse and classical set theory.

First of all I need to precise the terms of this comparison. On the one hand, for the Single Universe framework I am taking classical set theory in its usual axiomatization ZFC plus the addition of large cardinals axioms, as instantiated by the cumulative hierarchy V. According to universists, this, together with the restriction of set-generic forcing to countable transitive models, is enough for set theoretic practice.[17] On the other hand, the V-logic Multiverse is characterised as $ZFC + LCs+$ the Multiverse Axioms. Note that, as usually argued by the universist, the addition of the Multiverse Axioms do not add any "real" power to $ZFC + LCs$, since everything we need is already in the latter theory, at least according to universists.

We can now proceed to the first step of my argument, i.e. showing that the V-logic multiverse can prove the existence of an extra object that it is unavailable in $ZFC + LCs$. This object is a proper, uncountable, outer model of V. Such an object cannot exists in the universist's framework of $ZFC + LCs$: indeed, the application of forcing in that usual setting is done only to countable transitive models.[18] This is because to do it we need the existence of generic filters, and for the universist there are no V-generic filters.

However, in the V-logic multiverse framework we can prove the following theorem:

Theorem 1. *Let φ be a V-logic sentence (for instance, a sentence which says "$Con(T)$" for some V-logic theory T). The following are equivalent:*

[16] In particular, models obtained through *set-forcing*.
[17] This point is argued by N. Barton (2019).
[18] See Nik (2014), Antos, N. Barton, and S.-D. Friedman (nd) and N. Barton (2019).

1. *φ is consistent in V-logic.*
2. *φ is consistent in v-logic (this is essentially the V-logic build upon a transitive countable model v, instead of the full uncountable V).*
3. *$\bar{V}$ has an outer model, $\bar{W}$, such that $\bar{W} \models \varphi$.*
4. *There exists a $\bar{W}^*$, elementarily equivalent to $\bar{W}$, such that $\bar{W}^* \models \varphi$.*[19]

This theorem implies that, in the V-logic multiverse, even if we start with a countable model of ZFC inside V, we can then end up with a proper, uncountable outer model of an *uncountable* V[20]

Consequently we have, in the V-logic Multiverse, an object that cannot be found in the universist's framework. We now need to prove that this new object realises a new isomorphism type. And this is exactly my claim.

To see this, consider the technique of #-generation.[21] As stated by Antos, Barton and Friedman, this method is very useful in encapsulating several large cardinals consequences of reflection properties.[22] It is based upon the existence of *class-iterable sharps*: these are transitive structures that are amenable (i.e. $x \cap U \in N$ for any $x \in N$), with a normal measure and iterable in the sense that all successive ultrapower iterations along class well-orders are well-founded.[23] If such an object exists, then we can have *class iterated sharp generated* models, i.e. models that arise through collecting together each level indexed by the largest cardinal of the model that result from the iteration of a class-iterable sharp.[24] Finally, we can claim that V is such class iterably sharp generated, and enjoy all advantages of this fact (the main advantage is that any satisfaction obtainable in height extensions of V adding ordinals is already reflected to an initial segment of V itself). However, in V we cannot find such a class-iterable sharp, since, if it were the case, then we would be able to prove the existence of a cardinal that is both regular *and* singular[25], but this is impossible.[26] So in the classical set theoretic framework V is not a class iterably sharp generated model, and all of the above is unattainable.

This situation is fundamentally different in the V-logic multiverse. Indeed, since in the V-logic multiverse we can have proper, uncountable, extensions of V, we can also have, in these extensions, a class-iterable sharp. And thus, in the V-logic multiverse, we can claim that V is, in fact, class iterably sharp generated. This result opens a new realm of isomorphisms types between all the various iterated ultrapowers, and models of different heights that are provided by #-generation. Thus, we can claim that $ZFC + LCs$ is restrictive compared to the V-logic multiverse, since in the latter we can find a new object that realises a new isomorphism type.

4 Concluding remarks

I have shown that the V-logic multiverse, characterised as $ZFC + LCs+$ Multiverse Axioms, and with V-logic as the background logic, proves more isomorphism types than classical set theory ($ZFC + LCs$), and thus we can say that classical set theory is *restrictive* compared to the V-logic multiverse.

[19]This theorem has been proved by the present author and Claudio Ternullo in the paper *Outer Models, V-logic and the Multiverse*, currently in preparation, and based on related results from Antos, N. Barton, and S.-D. Friedman (nd) and N. Barton (2019).

[20]The V-logic multiverse is not the only multiverse conception that claims the existence of proper outer models of V, the other being the Hyperuniverse. However, the latter assume that V is countable, thus simplifying the setting by a lot.

[21]See Antos, N. Barton, and S.-D. Friedman (nd) for a discussion of it.

[22]A reflection property is a property of a model that can be proved to be true already in an initial segment of that model.

[23]Here I am following the definition from Antos, N. Barton, and S.-D. Friedman (nd). The original definition in S. Friedman (2016) is slightly different.

[24]Again, the precise definition can be found in Antos, N. Barton, and S.-D. Friedman (nd).

[25]A regular cardinal is a cardinal which cofinality is equal to the cardinal itself, otherwise it is singular.

[26]See Antos, N. Barton, and S.-D. Friedman (nd) for the details.

The argument I presented is compelling, but it is only one step of a much wider research program. Other than the already mentioned UNIFY principle, it must be noted that my argument uses an intuitive definition of restrictiveness and isomorphism type that can both be refined. This can be done by following first and foremost the definitions present in Maddy (1997), and then the subsequent work done by Benedikt Löwe, Luca Incurvati and especially Albert Visser.[27]

In conclusion, showing that the V-logic multiverse is better than classical set theory concerning the principle MAXIMIZE is the first, necessary step to a better understanding of the set theoretic multiverse and the requirements for a good foundational framework for mathematics.

References

Antos, C., N. Barton, and S.-D. Friedman (nd). "Universism and Extensions of V". Pre-print.

Barton, N. (2019). "Forcing and the Universe of Sets: Must We Lose Insight?" In: *Journal of Philosophical Logic*. Forthcoming.

Barton, Neil (2019). "Forcing and the Universe of Sets: Must we lose insight?" In: *Journal of Philosophical Logic*, pp. 1–38.

Barwise, Jon (1975). *Admissible Sets and Structures*. Springer Verlag, Berlin.

Dickmann, M. (1975). *Large Infinitary Languages*. North Holland, Amsterdam.

Friedman, S. (2016). "Evidence for Set-Theoretic Truth and the Hyperuniverse Programme". In: *IfCoLog Journal of Logics and their Applications* 3.4, pp. 517–555.

Gitman, Victoria and Joel David Hamkins (2011). "A natural model of the multiverse axioms". In: *arXiv preprint arXiv:1104.4450*.

Hamkins, J. D. (2012). "The Set-Theoretic Multiverse". In: *Review of Symbolic Logic* 5.3, pp. 416–449.

Incurvati, Luca and Benedikt Löwe (2016). "Restrictiveness relative to notions of interpretation". In: *The Review of Symbolic Logic* 9.2, pp. 238–250.

Keisler, H. J. (1974). *Model Theory for Infinitary Logic*. North Holland, Amsterdam.

Löwe, Benedikt (2001). "A first glance at non-restrictiveness". In: *Philosophia Mathematica* 9.3, pp. 347–354.

— (2003). "A Second Glance at Non-restrictiveness†." In: *Philosophia Mathematica* 11.3.

Maddy, P. (1997). *Naturalism in Mathematics*. Oxford University Press, Oxford.

Nik, Weaver (2014). *Forcing for mathematicians*. World Scientific.

Steel, J.R. (2014). "Gödel's Program". In: *Interpreting Gödel. Critical Essays*. Ed. by J. Kennedy. Cambridge University Press, Cambridge, pp. 153–179.

Visser, Albert et al. (2006). "Categories of theories and interpretations". In: *Logic in Tehran* 26, pp. 284–341.

[27]See Löwe (2001), Löwe (2003), Incurvati and Löwe (2016), and Visser et al. (2006).

Decidability of the theory of addition and the Frobenius map in certain rings of rational functions

Dimitra Chompitaki[1], Manos Kamarianakis[1]*, and Thanases Pheidas[1]

[1]Department of Mathematics & Applied Mathematics, University of Crete, Heraklion, Greece

Abstract

Let p be a prime number, $\mathbb{F}_p$ a finite field with p elements, z a variable, $\mathbb{F}_p[z]$ the ring of polynomials in z with coefficients in $\mathbb{F}_p$ and $\mathbb{F}_p(z)$ the field of rational functions of z over $\mathbb{F}_p$. We consider the existential theory of addition and the Frobenius map of a ring $R \subset \mathbb{F}_p(z)$, where R is generated over $\mathbb{F}_p[z]$ by inverting finitely many irreducible polynomials of $\mathbb{F}_p[z]$. We prove that it is model-complete and hence decidable. We also prove that, if the existential theory, in the same language, of $\mathbb{F}_p(z)$ is decidable, then its first-order theory is also decidable.

1 Introduction

Our work is part of the on-going research revolving around the fact that the ring-theory and even the existential ring-theory of any field of rational functions $\mathbb{F}_p(z)$ over a finite field $\mathbb{F}_p$ with p elements is undecidable (see [Phe91] and [Vid94]). So, proving decidability of structures weaker than the ring-structure of such a field and its subrings is desirable (cf. [PZ00]).

Consider R as a structure (model) of the language $\mathcal{L} := \{+, =, x \mapsto x^p, x \mapsto zx, 0, 1\}$, where $=$ and $+$ denote regular equality and addition, $x \mapsto zx$ denotes the multiplication-by-z map and $x \mapsto x^p$ is the *Frobenius map*. In [PZ04], the authors proved that the $\mathcal{L}$-theory of $\mathbb{F}_p[z]$ is model-complete (meaning that every formula is equivalent in $\mathbb{F}_p[z]$ to an existential $\mathcal{L}$-formula), and, hence, decidable. A similar (model completeness) result has been proved for the $\mathcal{L}$-structure of the ring of power series $\mathbb{F}_p[[z]]$ of z (see [Ona18]). It is a natural question to ask whether the $\mathcal{L}$-theory of $\mathbb{F}_p(z)$ is model-complete. For the moment, this problem seems inapproachable with current means and may demand the use of novel tools. Those that we use here suffice to prove model completeness for subrings of $\mathbb{F}_p(z)$, generated over $\mathbb{F}_p[z]$ by the inverses of finitely many irreducible polynomials.

The structure of addition and the Frobenius map is interesting, not only for its own sake, but also because it is connected to various important mathematical and logical domains and problems. For example, the derivative of a function (polynomial, rational or power series) is positive-existentially definable in $\mathcal{L}$ (see [PZ04]). So, the structure of $\mathbb{F}_p(z)$ as a model of addition and differentiation is encodable in its $\mathcal{L}$-structure. It is also interesting to study this structure with p as a parameter, cf., the open *Gröthendieck-Katz conjecture* [Ber91].

In another direction, it is a long-standing (and famous) problem whether there is *resolution of singularities* of algebraic varieties in positive characteristic. In the zero characteristic case, it has been proved to always exist, in the algebraic and the analytic sense, by Hironaka [Hir64]. But in positive characteristic it is an open problem. Although it has been thought, for a long time, that such resolution always exists (no counter-example is known), all efforts to prove it have failed so far and, recently, experts have expressed doubts. From the investigations so far, it seems that the main obstacle in characteristic $p > 0$ is the existence of the Frobenius map .

*Corresponding Author, Email: kamarianakis@uoc.gr

2 Our results

2.1 The main theorems

The existential $\mathcal{L}$-theory of the ring $R \subset \mathbb{F}_p(z)$, where R is generated over $\mathbb{F}_p[z]$ by inverting finitely many irreducible polynomials of $\mathbb{F}_p[z]$, is the set of *existential sentences* of $\mathcal{L}$, i.e., the first-order sentences of the language $\mathcal{L}$ which are of the form $\exists x_1 \ldots \exists x_m \in R : \phi$, where ϕ is a boolean combination of equations that may be written in the language $\mathcal{L}$.

We construct an algorithm which, given an existential formula σ of $\mathcal{L}$, finds an equivalent universal formula, thus we prove model-completeness.

Theorem 2.1. *Let R be the subring of $\mathbb{F}_p(z)$ generated over $\mathbb{F}_p[z]$ by inverting finitely many irreducibles polynomials of $\mathbb{F}_p[z]$. The $\mathcal{L}$-theory of R is model-complete.*

For the case $R = \mathbb{F}_p(z)$, we observe that our structure is a module over the non-commutative ring $\mathbb{F}_p(z)[P]$, where P is defined by $Px := x^P$, but with constant symbols that are not contained in the language of modules that is used in the existing bibliography. We make a variant of the well-known theorem of Baur and Monk [Bau76, Mon75], using work of Van den Dries and Holly [VdDH92], and obtain the following theorem.

Theorem 2.2. *Assume that the existential $\mathcal{L}$-theory of $\mathbb{F}_p(z)$ is decidable. Then the $\mathcal{L}$-theory of $\mathbb{F}_p(z)$ is decidable.*

But the decidability (or not) of the existential $\mathcal{L}$-theory remains an open problem.

2.2 The main technical theorem

We present an outline of a new method that we introduced in order to prove Theorem 2.1.

Definition 2.3. *An* additive polynomial *is a polynomial of the form*

$$f(\bar{x}) = \sum_{i=1}^{n} f_i(x_i),$$

where $\bar{x} = (x_1, \ldots, x_n)$ and, for each i,

$$f_i(x_i) := b_i x_i^{p^{s(i)}} + \sum_{j=1}^{s(i)-1} c_{i,j} x_i^{p^{s(i)-j}},$$

with $b_i, c_{i,j} \in \mathbb{F}_p(z)$.

For $s \in \mathbb{N}$, let $\mathcal{V}_s$ be $\mathbb{F}_p(z)$, considered as a vector space over the field $\mathbb{F}_p(z^s)$. An additive polynomial f as above is called *normalized* if all degrees $s(i)$ are equal to some s and the set of leading coefficients $\{b_i \mid i = 1, \ldots, n\}$ is linearly independent over $\mathcal{V}_s$.

Theorem 2.4. *Let f be a normalized additive polynomial of the variables $\bar{x} = (x_1, \ldots, x_n)$, which has positive degree in all the variables. Let $u \in \mathbb{F}_p(z)$. Then the set $\{\bar{x} \in R^n \mid f(\bar{x}) = u\}$ is finite.*

The method of proof involves diverse tools, such as the *Hasse derivative* D_i [Has36], where i denotes the order of derivation. This "hyperderivative" is used to create a matrix operator W which generalizes the concept of the Wronskian operator in characteristic zero.

Note that Theorem 2.4 is indicative of the usefulness of our methodology; the inverse image of a rational function through a multi-variate polynomial is, in general, infinite. We also show that Theorem 2.4 is not true if one replaces R by $\mathbb{F}_p(z)$ - and this shows the limits of our method: We can construct a normalized additive polynomial f and choose a $u \in \mathbb{F}_p(z)$ such that the set $\{\bar{x} \in \mathbb{F}_p(z)^n \mid f(\bar{x}) = u\}$ is infinite.

Acknowledgments

This research work was supported from Greek and European Union resources, through the National Strategic Reference Framework (NSRF 2014-2020), under the call "Support for researchers with emphasis on young researchers (EDBM103)" and the funded project "Problems of Diophantine Nature in Logic and Number Theory" with code MIS 5048407.

References

[Bau76] Walter Baur. Elimination of quantifiers of modules. *Israel J. Math.*, 25:64–70, 1976.

[Ber91] Daniel Bertrand. Groupes algébriques et équations différentielles linéaires. *Astérisque*, 206:183–204, 1991.

[Has36] Helmut Hasse. Theorie der höheren Differentiale in einem algebraischen Funktionenkörper mit vollkommenem Konstantenkörper bei beliebiger Charakteristik. *J. Reine. Ang. Math.*, 1936(175):50–54, January 1936.

[Hir64] Heisuke Hironaka. Resolution of singularities of an algebraic variety over a field of characteristic zero. I. *Ann. of Math.*, 79(1):109–203, 1964.

[Mon75] Leonard Gaines Monk. Elementary-recursive decision procedures. *PhD Dissertation, Berkeley*, 1975.

[Ona18] Gönenç Onay. $\mathbf{F}_p((\mathrm{X}))$ is decidable as a module over the ring of additive polynomials. *arXiv.org*, June 2018.

[Phe91] Thanases Pheidas. Hilbert's tenth problem for fields of rational functions over finite fields. *Invent. Math.*, 103(1):1–8, 1991.

[PZ00] Thanases Pheidas and Karim Zahidi. Undecidability of existential theories of rings and fields: a survey. *Contemp. Math.*, 270:49–106, 2000.

[PZ04] Thanases Pheidas and Karim Zahidi. Elimination theory for addition and the Frobenius map in polynomial rings. *J. Symb. Log.*, 69(4):1006–1026, 2004.

[VdDH92] Lou Van den Dries and Jan Holly. Quantifier elimination for modules with scalar variables. *Ann. Pure Appl. Logic*, 57:161–179, 1992.

[Vid94] Carlos Videla. Hilbert's tenth problem for rational function fields in characteristic 2. *Proc. Amer. Math. Soc.*, 120(1):249–253, 1994.

A nonstandard view on vagueness

Bruno Dinis and Bruno Jacinto

Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal
Universidade de Lisboa
bmdinis@fc.ul.pt
bmjacinto@fc.ul.pt

Abstract

We propose a new theory of vagueness based on the notions of marginal and large differences in the context of nonstandard mathematics. We apply this theory to an explanation of the seductiveness of the Sorites paradox by coupling it with Fara's interest-relative theory of vagueness.

1 Introduction

Vague predicates give rise to *soritical reasoning.* The following ingredients are thought to give rise to Sorites paradoxes. A (weak) order R and a finite series D of objects ordered by R such that: i) the first member of D (definitely) has property P; ii) The last member of D (definitely) fails to have P; and iii) the following principle is plausible.

Inductive premise: $\forall x, y \in D((P(x) \wedge S(x,y)) \Rightarrow P(y))$.

From the inductive premise and the fact that the first member of D has property P it follows that the second member of D also has P. Via a novel application of the inductive premise it is concluded that the third member of D has property P, and so on, until one reaches the *absurd* conclusion that the last member of D both has and lacks property P.

A typical example of the paradox consists of a finite sequence of people ordered from shortest to tallest, with clear cases of, respectively, shortness and tallness, and such that adjacent members in the series differ by no more than $0.5\,cm$ in height. The weak order R consists in the *shorter than* relation and the property P is the property of not being *tall.* Since the first member of the series isn't tall, the second member of the series also isn't tall, by the inductive premise. And so on, until one reaches the absurd conclusion that the last member of the sequence is both tall and non-tall.

Our particular view on the Sorites paradox is that the inductive premise is false and so that there is a *sharp boundary* between those objects in the series that fall under P and those that do not. Among the questions for those that accept the existence of sharp boundaries for vague predicates is the following **psychological question** [4]: Why are we inclined to accept the inductive premise, if it is false? Why are we inclined to think that the cut-off point between the Ps and the not-Ps is not at any particular point in the series?

The present paper has three aims: i) is to formulate and defend what we call the ML theory of the notions of *marginal difference* and *large difference*; ii) to show that the ML theory determines a space of vague magnitudes with the structure of nonstandard models of arithmetic and analysis;[1] iii) to apply the ML theory to an account of why soritical reasoning is seductive and yet fallacious, i.e., to give an answer to the psychological question. While in our view the theory of vague magnitudes here developed seems to be a part of the correct answers to other important questions concerning the phenomenon of vagueness, addressing those other questions is beyond the paper's scope.

[1] See [1] for a related idea.

2 The ML Theory

We start by assuming the existence of a (strict) weak order R and two binary predicates M and L. Their intended interpretation is the following: $M(x, y)$ states that x is *marginally smaller* than y with respect to order R, while $L(x, y)$ states that x is *largely smaller* than y with respect to R. As such, the ML Theory is a theory for "marginal" and "large" differences. We now present the theory's axioms:[2]

The first axiom postulates that the theory is not trivial in the sense that there are indeed objects with "large" differences. The existence of an object with "marginal" differences is a direct consequence of this together with Axiom 2.10 below.

Axiom 2.1. $\exists x, y\, L(x, y)$

The second and third axioms are that whenever x is marginally/largely smaller than y, then these objects are ordered by R.

Axiom 2.2. $\forall x, y\, (M(x, y) \Rightarrow R(x, y))$

Axiom 2.3. $\forall x, y\, (L(x, y) \Rightarrow R(x, y))$

We assume that marginal differences are transitive. This is in line with one of the so-called *Leibniz rules* [3, Chapter 1], corresponding to the intuition that the sum of two infinitesimals is still infinitesimal.

Axiom 2.4. $\forall x, y, z\, (M(x, y) \wedge M(y, z) \Rightarrow M(x, z))$

The next axiom states that marginal differences are not large.

Axiom 2.5. $\forall x, y\, (M(x, y) \Rightarrow \neg L(x, y))$

The essence of vagueness is that marginal differences make no difference with respect to large differences. The next axiom captures that idea.

Axiom 2.6. $\forall x, y, z\, (M(x, y) \Rightarrow ((L(z, y) \Rightarrow L(z, x)) \wedge (L(x, z) \Rightarrow L(y, z))))$

The next two axioms capture the intuition that increasing large differences only results in large differences.

Axiom 2.7. $\forall x, y, z\, ((R(x, y) \wedge L(y, z)) \Rightarrow L(x, z))$

Axiom 2.8. $\forall x, y, z\, ((L(x, y) \wedge R(y, z)) \Rightarrow L(x, z))$

So far it would seem plausible that there are only two orders of magnitude. The next axiom gives a tameness condition for possible elements between "marginal" and "large": if x is smaller than y, then either x is marginally smaller than y, or x is largely smaller than y, or there exists some z, between x and y such that x is neither marginally nor largely smaller than z.

Axiom 2.9. $\forall x, y\, (R(x, y) \Rightarrow (M(x, y) \vee L(x, y) \vee \exists z(R(x, z) \wedge R(z, y) \wedge \neg M(x, z) \wedge \neg L(x, z))))$

[2]While the Sorites paradox (typically) involves only finitely many objects, the ML theory implies that there are infinitely many objects. Such a commitment is a common idealization of several theories of measurement. For this and related reasons some authors [7, 12] have proposed that theories of measurement should be formulated as theories about relations between *magnitudes* rather than about relations between objects. For simplicity, we here proceed in the standard manner by treating the ML Theory as a first-order theory.

Any large difference can be decomposed into a marginal difference and a large difference. This is of course another crucial property of vague predicates: removing a grain of sand from a heap still leaves a heap.

Axiom 2.10. $\forall x, y\,(L(x,y) \Rightarrow (\exists z(M(x,z) \wedge L(z,y)) \wedge \exists w(M(w,y) \wedge L(x,w))))$

Finally, we have an axiom of "density" for marginal differences: if x is marginally smaller than y then, for every z between x and y, it must be the case that x is marginally smaller than z and z is marginally smaller than y.

Axiom 2.11. $\forall x, y(M(x,y) \Rightarrow \forall z((R(x,z) \wedge R(z,y)) \Rightarrow (M(x,z) \wedge M(z,y))))$

3 Nonstandard analysis

'Nonstandard analysis' was the name given by Abraham Robinson's [10, 11] to any theory giving a formal and consistent treatment of infinitesimals. There are many formulations of nonstandard analysis (cf. e.g. [2, 3, 5]). However, for the purposes of this paper, a very "economical" version due to Edward Nelson [8, Chapter 4], which we shall denote by ENA (for Elementary Nonstandard Analysis), is sufficient.[3] Indeed, the main requirement that we need from nonstandard analysis is the possibility of defining different orders of magnitude, which can be achieved in any theory that permits the existence of infinitesimals.

Mathematics is usually formalized through the axioms of Zermelo-Fraenkel Set Theory (with or without the Axiom of Choice) and in a language which only contains one undefined non-logical symbol, '$\in$', for set membership (cf. e.g. [6, 9]). The theory ENA (cf. Figure 1) is a conservative extension of Zermelo-Fraenkel Set Theory which is governed by a simple set of extra axioms after adding to the language a new predicate, 'st'. One should read $\mathrm{st}(x)$ as 'x is standard'. Formulas which involve the predicate 'st' are called *external* and formulas which do not, i.e. formulas in the language of classical mathematics, are called *internal*.

(1) $$\mathrm{st}(0)$$

(2) $$\forall n \in \mathbb{N}\,(\mathrm{st}(n) \Rightarrow \mathrm{st}(n+1))$$

(3) $$\exists \omega \in \mathbb{N}\,(\neg \mathrm{st}(\omega))$$

(4) $$(\Phi(0) \wedge \forall^{\mathrm{st}} n(\Phi(n) \Rightarrow \Phi(n+1))) \Rightarrow \forall^{\mathrm{st}} n\, \Phi(n),$$

where Φ is an arbitrary formula (internal or external) and $\forall^{\mathrm{st}} n\, \Phi(n)$ is an abbreviation of $\forall n(\mathrm{st}(n) \Rightarrow \Phi(n))$.

Figure 1: The axioms of ENA

Let us briefly comment on the axioms of ENA. The first two axioms state that the usual natural numbers are standard and the third axiom postulates the existence of nonstandard natural numbers. Finally, we have an axiom scheme which is a form of induction that allows us to conclude that some property is true for all standard natural numbers given that it is true for zero and that whenever it is true for some standard n, then it is also true for its successor $n+1$. Since ENA is a conservative extension of classical mathematics, the usual form of induction is

[3] This theory is dubbed ENA^- in [3].

still valid, albeit only for internal properties. To see why such restriction is required, consider the formula $\Phi(n) :\equiv \mathrm{st}(n)$. If one could apply internal induction to Φ the conclusion would be that every natural number is standard, in contradiction with the third axiom.

One defines different orders of magnitude as follows. A real number x is said to be *infinitesimal* if its absolute value is smaller than the inverse of any positive standard natural number; *limited*, if it is, in absolute value, bounded by some standard natural number and *unlimited*, or *infinitely large* if it is not limited.

4 Models and Representation

In the following we show that the ML Theory is consistent by providing two different instantiations $\mathfrak{I}_1, \mathfrak{I}_2$ of the predicates M and L such that all the axioms of the ML Theory are satisfied in the context of ENA. The choice of nonstandard models is guided by the observation that the simplest standard structures are not models of our theory.

The instantiation $\mathfrak{I}_1$ is *arithmetical* in nature because it only involves natural numbers:

1. $R(x,y) := x < y$, where $<$ is the usual order in the natural numbers
2. $M(x,y) := \exists n(\mathrm{st}(n) \wedge y = x + n)$
3. $L(x,y) := x \leq y \wedge \neg \mathrm{st}(y - x)$

Instantiation $\mathfrak{I}_2$ is *analytic* in nature as it requires (nonstandard) real numbers:

1. $R(x,y) := x < y$, where $<$ is the usual order in the real numbers
2. $M(x,y) :=$ "$x \leq y$ and their difference is infinitesimal"
3. $L(x,y) :=$ "$x \leq y$ and their difference is infinitely large"

Theorem 4.1. ENA *together with any of the instantiations $\mathfrak{I}_1$ and $\mathfrak{I}_2$ satisfies the* ML *axioms.*

The verifications of the previous theorem are all immediate. We observe that in Axiom 9 the existentially quantified z is not realized under the instantiation $\mathfrak{I}_1$ but is realized under the instantiation $\mathfrak{I}_2$.

Corollary 4.2. *The* ML *axiomatics is consistent, if* ZF *is consistent.*

Finally, we note the following *representation theorem* revealing the conditions under which countable models of the ML Theory are represented by countable nonstandard models of arithmetic.

Definition 4.3. *For any nonstandard arithmetical model $\mathbf{M} = \langle D_{\mathbf{M}}, \mathrm{st}_{\mathbf{M}}, <_{\mathbf{M}}, +_{\mathbf{M}}, \cdot_{\mathbf{M}} \rangle$, let $\mathbf{M}^* = \langle D_{\mathbf{M}}, <_{\mathbf{M}}, M_{\mathbf{M}^*}, L_{\mathbf{M}^*} \rangle$, where: i) $M_{\mathbf{M}^*}(x,y)$ iff $\mathbf{M} \vDash \exists n(\mathrm{st}(n) \wedge y = x + n)$; and ii) $L_{\mathbf{M}^*}(x,y)$ iff $x <_{\mathbf{M}} y$ and $\mathbf{M} \vDash \neg \exists n(\mathrm{st}(n) \wedge y = x + n)$.*

Consider the following condition on models of ML:

Definition 4.4 (Local linearity)**.** *A model* $\mathbf{M}$ *of* ML *is* locally linear *if and only if, for every $x \in D_{\mathbf{M}}$, the restriction of $\lesssim$ to $\{y \in \mathbf{D}_M : M(x,y)$ or $(\neg(x <_{\mathbf{M}} y)$ and $\neg(y <_{\mathbf{M}} x))\}$ is a linear order.*

Let the ML$^+$ THEORY be that theory whose axioms are those of the ML THEORY except that axiom 2.9 is strengthened to

$$\forall x, y\,(R(x,y) \Rightarrow (M(x,y) \vee L(x,y)))$$

The following result, whose proof we omit for lack of space, reveals the nonstandard character of the ML$^+$ THEORY.

Theorem 4.5. *For every countable and locally linear model* $\mathbf{M}$ *of the* ML$^+$ THEORY *there is an homomorphism from* $\mathbf{M}$ *to* $\mathfrak{M}^*$, *for any countable nonstandard model* $\mathfrak{M}$ *of arithmetic.*

5 The ML THEORY and the psychological question

Consider the following constraint on *tallness*.

TALLNESS CONSTRAINT: $\forall y(S(y) \Rightarrow \forall x(T(x) \Leftrightarrow L(y,x)))$

Once conjoined with the TALLNESS CONSTRAINT the ML THEORY implies the following tolerance principle.

TOLERANCE: $\forall x \forall y((T(x) \wedge M(y,x)) \Rightarrow T(y))$

TOLERANCE is akin to the Sorites paradox's INDUCTIVE PREMISE. In our view the *de facto* truth of TOLERANCE partially accounts for the seductiveness of the INDUCTIVE PREMISE. To see why, note that there are two ideas encapsulated in the notion of *succession*: $i)$ y is adjacent to x is the Sorites series; and $ii)$ y is marginally taller than x. Part of the explanation for the seductiveness of the INDUCTIVE PREMISE is thus a *confusion* between marginal difference and adjacency. While some pairs of adjacent elements in a Sorites series are only marginally different, others are appreciably different.

Yet, this is an insufficient explanation for the principle's seductiveness. It remains to be explained *why* we are prone to confuse marginal difference with adjacency. Still, and while an explanation must perforce go beyond the scope of the ML THEORY, the theory does encapsulate a conception of vagueness that may be conjoined with specific theories of extant phenomena involving vagueness to provide prima facie satisfactory replies to the *psychological question*.

Here we'll focus on Delia Graff Fara's [4] *interest-relative* theory of vagueness. According to this theory whether two objects of different heights are only marginally different with respect to their height depends on whether they are *saliently* similar in this respect, and whether two objects are saliently similar is a context-dependent matter. Furthermore, two objects are *similar* in a situation whenever, relative to the agent's interests in the situation, the costs of discriminating the objects outweigh the benefits of doing so, and *salient* in a situation when they are being actively compared as live options for the agent's purposes in that situation.

Now, consider a particular Sorites series for tallness, assuming the existence of a cut-off point between the tall and the non-tall. Suppose that before you start going through the objects arranged in the Sorites series o_5 is the last of the non-tall and o_6 is the first of the tall ones. Now, imagine that you start going through the objects in the series, starting with the shortest one. Assuming that you have no particular interest in discriminating the objects with respect to height, o_5 and o_6 will be similar in this respect. Moreover, once you actively compare o_5 with o_6 they become salient. So, even if o_5 was the last of the non-tall ones and o_6 was the first of the tall ones before you actively compared them, once you do so they differ solely marginally with respect to their height, and therefore are either both tall or both non-tall.

The answer to the psychological question afforded by Fara's interest-relative theory of vagueness is thus that we are prone to confuse marginal difference with adjacency in a Sorites series because adjacent objects in a Sorites series *are* marginally different provided that one is actively considering them and the costs of distinguishing them with respect to their height outweigh the benefits of doing so. Interestingly, Fara's theory perfectly aligns with the judgment prompted by the ML THEORY that the Sorites paradox arises from a confusion between adjacency in a Sorites series and marginal difference. Furthermore, a theory of vague magnitudes must arguably have a structure of the sort characterized by the ML THEORY if it is to be compatible with Fara's answer to the psychological question. Suppose once more that you are going through the previous Sorites series for tallness, and that you assign degrees of tallness to the objects in the series according to some scale. Initially, o_5 has a degree t of tallness, which places it among the non-tall, and o_6 has a degree t' of tallness, which places it among the tall.

Assume, without loss of generality, that once the similarity in height between o_5 and o_6 becomes salient, you place o_6 together with o_5 among the non-tall. Then, in order for o_6 to also be assigned a non-tallness magnitude, there must be a non-tallness magnitude t'' greater than t that gets assigned to objects taller than o_5. By continuing through the Sorites series, it is easy to realize that there must also be a non-tallness magnitude t''' greater than t' (and t). And so on. Thus, the non-tallness magnitudes must be open ended. Similar reasoning shows that the tallness magnitudes cannot have a smallest magnitude.

This structure of the vague magnitudes is precisely of the sort determined by the ML THEORY. Yet, other theories of vague magnitudes would not have this prediction. For instance, no theory that treated marginal difference as amounting to nothing more, nothing less, than difference of exactly $n\,cm$, for a fixed real number n, would possess this sort of structure. So, in this sense, the interest-relative theory of vagueness's answer to the psychological question presupposes the ML THEORY, or at least theories quite close to it.

It is worth emphasizing that this particular structure of vague magnitudes arises from the fact that the ML THEORY is *non-reductionist* about marginal difference and large difference. That is, marginal and large difference are not appropriately captured by equating them with specific (non-infinitesimal) differences with respect to the precise quantity in question. Arguably, what is required to pin down such a non-reductionist conception of marginal and large difference is an axiomatization of the sort we have offered, with the consequence that vague magnitudes possess a nonstandard structure.

6 Conclusion

In this paper we have offered a new theory of vagueness – the ML THEORY –, as well as a representation of the structure of vague magnitudes determined by the ML THEORY in terms of models of nonstandard mathematics.

Furthermore, we have shown that the ML THEORY and Fara's *interest-relative* theory of vagueness jointly offer a reasonable answer to the psychological question about vagueness. Not only is it the case that TOLERANCE is a prediction of the ML THEORY, it is also the case that in order for contextual shifts in what counts as a cut-off point to be admissible the space of vague magnitudes must be *open-ended* in the way predicted by the ML THEORY.

7 Acknowledgments

Both authors acknowledge the support of FCT — Fundação para a Ciência e Tecnologia under the projects UIDB/04561/2020 and UIDP/04561/2020 (first author) and UIDB/00678/2020, UIDP/00678/2020 and CEECIND/04761/2017/CP1387/CT0008,(second author). The first author also acknowledges the research center Centro de Matemática, Aplicações Fundamentais e Investigação Operacional, Universidade de Lisboa and the second author the research center Centro de Filosofia das Ciências, Universidade de Lisboa.

References

[1] W. Dean. Strict finitism, feasibility, and the sorites. *The Review of Symbolic Logic*, 11(2):295–346, 2018.

[2] M. di Nasso. On the foundations of nonstandard mathematics. *Math. Japon.*, 50(1):131–160, 1999.

[3] B. Dinis and I. van den Berg. *Neutrices and external numbers: a flexible number system.* Monographs and Research Notes in Mathematics. CRC Press, Boca Raton, FL, 2019. With a foreword by Claude Lobry.

[4] D. Graff. Shifting sands. *Philosophical Topics*, 28(1):45–81, 2000.

[5] V. Kanovei and M. Reeken. *Nonstandard Analysis, Axiomatically.* Springer Monographs in Mathematics, 2014.

[6] K. Kunen. *Set theory*, volume 102 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam-New York, 1980. An introduction to independence proofs.

[7] B. Mundy. The metaphysics of quantity. *Philosophical Studies*, 51(1):29–54, 1987.

[8] E. Nelson. *Radically Elementary Probability Theory.* Annals of Mathematical Studies, vol. 117. Princeton University Press, Princeton, N. J., 1987.

[9] M. Potter. *Set theory and its philosophy.* Oxford University Press, New York, 2004. A critical introduction.

[10] A. Robinson. Non-standard analysis. *Nederl. Akad. Wetensch. Proc. Ser. A 64*, 23:432–440, 1961.

[11] A. Robinson. *Non-standard analysis.* North-Holland Publishing Co., Amsterdam, 1966.

[12] C. Swoyer. The metaphysics of measurement. In John Forge, editor, *Measurement, Realism and Objectivity Essays on Measurement in the Social and Physical Sciences*, pages 235–290. Reidel, 1987.

Decidability of local fields and their extensions

Konstantinos Kartas

University of Oxford, Mathematical Institute, Woodstock Road, Oxford
kartas@maths.ox.ac.uk

Abstract

We provide a survey of classical decidability results for local fields and then present some new results for various infinite extensions of local fields which are of arithmetic interest.

1 Introduction

The decidability of the p-adic numbers $\mathbb{Q}_p$, established by Ax-Kochen [AK65] and Ershov [Ers65], still remains one of the highlights of model theory. It motivated several decidability results both in mixed and positive characteristic:

- In mixed characteristic, Kochen [Koc74] showed that $\mathbb{Q}_p^{ur}$, the maximal unramified extension of $\mathbb{Q}_p$, is decidable. More generally, by work of [Zie72], [Ers65], [Bas78], [Bél99] and more recently [AJ19], [Lee20] and [LL21], we have a good understanding of the model theory of unramified and finitely ramified mixed characteristic henselian fields.

- In positive characteristic, our understanding is much more limited. Nevertheless, by work of Denef-Schoutens [DS03], we know that $\mathbb{F}_p((t))$ is *existentially* decidable in $L_t = \{+, \cdot, -, 0, 1, t\}$, modulo resolution of singularities. In fact, Theorem 4.3 [DS03] applies to show that any finitely ramified extension of $\mathbb{F}_p((t))$ is existentially decidable relative to its residue field.

Note that all of the above results are restricted to *finitely* ramified extensions of $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$. The situation is less clear for *infinitely* ramified fields and there are many such algebraic extensions of $\mathbb{Q}_p$, whose decidability problem is still open. This is the content of the following sections. Our results are divided into two categories, the wildly ramified extensions and the tamely ramified extensions.

2 Wildly ramified extensions

Recall the definition:

Definition 2.0.1. *A finite extension $(L, w)/(K, v)$ of valued fields is said to be wildly ramified if the ramification degree $e(L/K)$ is p-divisible, where p is the residue characteristic of (K, v). An algebraic extension is said to be wildly ramified if any finite subextension is wildly ramified.*

In practice, one refers to wildly ramified extensions when the ramification degree is highly p-divisible. Important wildly ramified extensions of $\mathbb{Q}_p$ include:

Example. *(a) $\mathbb{Q}_p^{ab}$, the maximal abelian extension of $\mathbb{Q}_p$.*
(b) $\mathbb{Q}_p(\zeta_{p^\infty})$, the totally ramified extension obtained by adjoining all p^n-th roots of unity.

These extensions have been discussed in Macintyre's survey on pg.140 [Mac86] and a conjectural axiomatization of $\mathbb{Q}_p^{ab}$ was given by Koenigsmann on pg.55 in [Koe18]. Another interesting extension is $\mathbb{Q}_p(p^{1/p^\infty})$, a totally ramified extension of $\mathbb{Q}_p$ obtained by adjoining a compatible system of p-power roots of p.
The p-adic completions of the above fields are typical examples of perfectoid fields (see [Sch12]). For any such field K, one can define its tilt, which intuitively is its *local* function field analogue and serves as a characteristic p approximation of K. For our fields of interest, one has that $\mathbb{Q}_p(p^{1/p^\infty})$ and $\mathbb{Q}_p(\zeta_{p^\infty})$ are approximated by $\mathbb{F}_p((t))^{1/p^\infty}$, the perfect hull of $\mathbb{F}_p((t))$, while $\mathbb{Q}_p^{ab}$ is approximated by $\bar{\mathbb{F}}_p((t))^{1/p^\infty}$, the perfect hull of $\bar{\mathbb{F}}_p((t))$. The fields $\mathbb{F}_p((t))^{1/p^\infty}$ and $\bar{\mathbb{F}}_p((t))^{1/p^\infty}$ are typical examples of wildly ramified extensions of $\mathbb{F}_p((t))$.
In [Kar20], the following is established:

Theorem A (Corollary A [Kar20])**.** (*a*) *Assume $\mathbb{F}_p((t))^{1/p^\infty}$ is decidable (resp. $\exists$-decidable) in L_t. Then $\mathbb{Q}_p(p^{1/p^\infty})$ and $\mathbb{Q}_p(\zeta_{p^\infty})$ are decidable (resp. $\exists$-decidable) in L_{val}.*
(*b*) *Assume $\bar{\mathbb{F}}_p((t))^{1/p^\infty}$ is decidable (resp. $\exists$-decidable) in L_t. Then $\mathbb{Q}_p^{ab}$ is decidable (resp. $\exists$-decidable) in L_{val}.*

In the above result, the language L_t is the language of valued fields together with a constant symbol for t. This is essentially a special case of the main result of [Kar20], which is a relative decidability result for perfectoid fields. The proof uses Fontaine's period rings, which are relevant in the construction of the Fargues-Fontaine curve.
One may also prove the following unconditional decidability result:

Theorem B. *There is an algorithm that decides whether a system of polynomial equations and inequations, defined over $\mathbb{Z}$, has a solution modulo p over each of the valuation rings of $\mathbb{Q}_p(p^{1/p^\infty}), \mathbb{Q}_p(\zeta_{p^\infty})$ and $\mathbb{Q}_p^{ab}$.*

The proof of Theorem B goes via reduction to characteristic p, but unlike Theorem A only existential decidability in L_{val} is needed on the characteristic p side. The latter is known by work of Anscombe-Fehm [AF16].
Relative decidability results in the reverse direction are also established in [Kar20]. For example:

Proposition. *If $\mathbb{Q}_p(p^{1/p^\infty})$ is $\forall^1\exists$-decidable in L_{val}, then $\mathbb{F}_p[[t]]^{1/p^\infty}$ is $\exists^+$-decidable in L_t.*

The above Proposition is not exactly a converse of the existential version of Theorem A but still suggests that if we eventually want to understand the theories of $\mathbb{Q}_p(p^{1/p^\infty})$, $\mathbb{Q}_p(\zeta_{p^\infty})$ and $\mathbb{Q}_p^{ab}$ (even modest parts of their theories), we have to face the diophantine problem over the perfect hull of $\mathbb{F}_p((t))$ and $\bar{\mathbb{F}}_p((t))$.

3 Tamely ramified extensions

We now discuss some new results for tamely ramified extensions that are established in [Kar21], where details and proofs may be found. Theorem C below is a general *existential* Ax-Kochen-Ershov principle for tamely ramified fields, with *no restriction* on the characteristic, but which is conditional on a certain form of resolution of singularities.
For our model-theoretic purposes, we need to extend the usual notion of a tamely ramified field extension to the context of transcendental valued field extensions:

Definition 3.0.1. *A valued field extension $(L,w)/(K,v)$ is said to be tamely ramified if l/k is separable*[1]*, the quotient group Δ/Γ has no p-torsion, where $p = char(k)$, and every finite subextension is defectless.*

Example. (a) *Every valued field extension is tamely ramified when the residue characteristic is zero.*
(b) *The valued field extension $(\mathbb{Q}_p(p^{1/n}), v_p)/(\mathbb{Q}, v_p)$ is tamely ramified if and only if $p \nmid n$.*
(c) *Let $\mathbb{F}_p((t^\Gamma))$ be the Hahn series field with residue field $\mathbb{F}_p$ and value group Γ. The valued field extension $\mathbb{F}_p((t^\Gamma))/\mathbb{F}_p(t)$ is tamely ramified if and only if 1 is not p-divisible in Γ.*

Our results in this section depend on a certain form of resolution of singularities. In very simple terms, resolution of singularities allows us to transform a given variety, which may have lots of singularities, to one which is non-singular. Moreover, the latter variety is in some sense close to the former, so that anything useful that can be said about the latter variety can often be translated into something useful about the former. The advantage of resolving the singularities of a variety lies in the fact that it is usually much easier to deal with non-singular varieties for all sorts of problems.
We now state the precise form that is assumed in [Kar21]:

Conjecture R (Log-Resolution)**.** *Let X be a reduced, flat scheme of finite type over an excellent discrete valuation ring R. Then there exists a blow-up morphism $f : \tilde{X} \to X$ in a nowhere dense center $Z \subset X$ such that*

1. *$\tilde{X}$ is a regular scheme.*
2. *$\tilde{X}_s = \tilde{X} \times_{SpecR} Spec(R/\mathfrak{m}_R)$ is a strict normal crossings divisor.*

The notion of an excellent ring, introduced by Grothendieck(see §7.9 [Gro65]), is quite technical to define here. However, for the case of discrete valuation rings, this simply means that $\widehat{K}/K$ is a separable (not necessarily algebraic), where $K = Frac(R)$ and $\widehat{K}$ denotes the completion of K. A divisor is said to be strict normal crossings if its reduced underlying scheme locally looks like a union of smooth varieties crossing transversely. In [Kar21], the following general existential Ax-Kochen-Ershov is obtained:

Theorem C (Theorem A [Kar21])**.** *Assume Conjecture R. Suppose (K,v) and (L,w) are henselian and tamely ramified over a discrete valued field (F,v_0) with $\mathcal{O}_F$ excellent. If $RV(K) \equiv_{\exists,RV(F)} RV(L)$, then $K \equiv_{\exists,F} L$ in L_r.*

Theorem C specializes to well-known Ax-Kochen-Ershov results in residue characteristic 0 and in the mixed characteristic *unramified* setting. Moreover, these Ax-Kochen-Ershov principles are known not only for the existential theories but also for the full-first order theories. The case of *finite* tame ramification in mixed characteristic and with perfect residue fields was proven recently in Corollary 5.9 [Lee20].
At the same time, Theorem C implies conditional existential decidability results for $\mathbb{F}_p((t))$ and its finite extensions, which were already known by the work of Denef-Schoutens [DS03]. Our proof does not use Greenberg's approximation theorem, which is an essential ingredient in [DS03].
On the other hand, Theorem C applies also to the setting of *infinite* ramification, providing us with an abundance of examples of *infinitely* ramified extensions of $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ whose theory is existentially decidable. This is the content of the next section.

[1]A field extension l/k (not necessarily algebraic) is said to be separable if l is linearly disjoint from k^{1/p^∞} (see §2.6 [FJ04]).

3.1 Decidability

In Remark 7.6 [AF16], the authors write:
"*At present, we do not know of an example of a mixed characteristic henselian valued field (K, v) for which k and (Γ, vp) are $\exists$-decidable but (K, v) is $\exists$-undecidable.*"
The existence of such an example is proved in Observation 1.2.2 [Kar20]. However, if we restrict ourselves to the tamely ramified setting, we indeed get such an Ax-Kochen style statement:

Corollary (Mixed characteristic)**.** *Assume Conjecture R. Suppose (K, v) and (L, w) are henselian and tamely ramified over $(\mathbb{Q}, v_p)$, admitting cross-sections that extend a given cross-section of $(\mathbb{Q}, v_p)$. If $k \equiv_\exists l$ in L_r and $(\Gamma, vp) \equiv_\exists (\Delta, wp)$ in L_{oag}, then $K \equiv_\exists L$ in L_r.*

In particular, if (K, v) is henselian and tamely ramified over $(\mathbb{Q}, v_p)$, admitting a cross-section extending one of $(\mathbb{Q}, v_p)$, then K is existentially decidable in L_r relative to k in L_r and (Γ, vp) in L_{oag} (see Corollary 4.1.4 [Kar21]). Similarly, we obtain a positive characteristic analogue:

Corollary (Positive characteristic)**.** *Assume Conjecture R. Suppose (K, v) and (L, w) are henselian and tamely ramified over $(\mathbb{F}_p(t), v_t)$, admitting cross-sections that extend a given cross-section of $(\mathbb{F}_p(t), v_t)$. If $k \equiv_\exists l$ in L_r and $(\Gamma, vt) \equiv_\exists (\Delta, wt)$ in L_{oag}, then $K \equiv_\exists L$ in L_t.*

Among the fields that are existentially decidable, the maximal tamely ramified extensions of $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ are of arithmetic significance.

Corollary. *Assume Conjecture R. Then the fields $\mathbb{Q}_p^{tr}$ and $\mathbb{F}_p((t))^{tr}$ are existentially decidable in L_r.*

3.2 Tweaking Abhyankar's example

Finally, we discuss a tame variant of the following famous example, essentially due to Abhyankar [Abh56]. It is also presented by Kuhlmann in a model-theoretic context in Example 3.13 [Kuh11]:

Example. *Let $(K, v) = (\mathbb{F}_p((t))^{1/p^\infty}, v_t)$ and $(L, w) = (\mathbb{F}_p((t^{1/p^\infty})), v_t)$ be the Hahn series field with value group $\frac{1}{p^\infty}\mathbb{Z}$ and residue field $\mathbb{F}_p$. We observe that $RV(K) \cong_{RV(\mathbb{F}_p((t)))} RV(L)$ but $(K, v) \not\equiv_{\exists, \mathbb{F}_p((t))} (L, w)$ since the Artin-Schreier equation $x^p - x - \frac{1}{t} = 0$ has a solution in L but not in K.*

Our version of Abhyankar's example is obtained by replacing p-power roots of t with l-power roots and exhibits a totally different behaviour:

Example. *Fix any prime $l \neq p$. Consider the valued fields $(K, v) = (\mathbb{F}_p((t))(t^{1/l^\infty}), v_t)$ and $(L, w) = (\mathbb{F}_p((t^{1/l^\infty})), v_t)$, with the latter being the Hahn series field with value group $\frac{1}{l^\infty}\mathbb{Z}$ and residue field $\mathbb{F}_p$. We observe that $RV(K) \cong RV(L)$ and by Theorem C we get that $(K, v) \equiv_{\mathbb{F}_p((t^{1/l^n})), \exists} (L, w)$, for all $n \in \mathbb{N}$. It follows that $\mathbb{F}_p((t))(t^{1/l^\infty}) \prec_\exists \mathbb{F}_p((t^{1/l^\infty}))$ in L_r.*

Acknowledgements

I would like to thank E. Hrushovski and J. Koenigsmann for invaluable guidance. Also many thanks to T. Scanlon for his hospitality while I was a visitor in UC Berkeley, where part of this work was completed.

References

[Abh56] Shreeram Abhyankar. Two notes on formal power series. *Proceedings of the American Mathematical Society Vol. 7, No. 5, pp. 903-905*, 1956.

[AF16] Sylvy Anscombe and Arno Fehm. The existential theory of equicharacteristic henselian valued fields. *Algebra & Number Theory Volume 10, Number 3 (2016), 665-683.*, 2016.

[AJ19] Sylvy Anscombe and Franziska Jahnke. The model theory of Cohen rings. *arXiv https://arxiv.org/abs/1904.08297*, (2019).

[AK65] James Ax and Simon Kochen. Diophantine problems over local fields II. *Amer. J. Math. 87*, 1965.

[Bas78] Serban A. Basarab. Some model theoery for henselian valued fields. *Journal of Algebra*, 1978.

[Bél99] Luc Bélair. Types dans les corps valués munis d'applications coefficients. *Illinois J. Math. 43(2):410- 425*, 1999.

[DS03] Jan Denef and Hans Schoutens. On the decidability of the existential theory of $\mathbb{F}_p[[t]]$. *Fields Institute Communications Volume 00, 0000*, (2003).

[Ers65] Ju.L. Ershov. On elementary theories of local fields. *Algebra i Logika 4, No. 2, 5-30*, 1965.

[FJ04] Michael D. Fried and Moshe Jarden. *Field Arithmetic.* Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. 11 (2nd revised and enlarged ed.). Springer-Verlag, 2004.

[Gro65] Alexander Grothendieck. *Éléments de géométrie algébrique : IV. Étude locale des schémas et des morphismes de schémas, Seconde partie.* Publications mathématiques de l' I.H.É.S., tome 24, p. 5-231, 1965.

[Kar20] Konstantinos Kartas. Decidability via the tilting correspondence. *arXiv https://arxiv.org/abs/2001.04424*, (2020).

[Kar21] Konstantinos Kartas. Diophantine problems over tamely ramified fields. *arXiv https://arxiv.org/abs/2103.14646*, (2021).

[Koc74] Simon Kochen. The model theory of local fields. *ISILC Logic Conference (Proc. Internat. Summer Inst. and Logic Colloq., Kiel, 1974), pp. 384-425. Lecture Notes in Math., Vol. 499, Springer, Berlin, 1975.*, 1974.

[Koe18] Jochen Koenigsmann. Decidability in local and global fields. *Proc. Int. Cong. of Math. Rio de Janeiro, Vol. 1 (45-60)*, 2018.

[Kuh11] Franz-Viktor Kuhlmann. The Defect. *Commutative algebra - Noetherian and non-Noetherian perspectives, 277-318, Springer, New York*, 2011.

[Lee20] Junguk Lee. Hyperfields, truncated DVRs, and valued fields. *Journal of Number Theory, Volume 212, July 2020, Pages 40-71*, 2020.

[LL21] Junguk Lee and Wan Lee. On the structure of certain valued fields. *Annals of Pure and Applied Logic Volume 172, Issue 4*, 2021.

[Mac86] Angus Macintyre. Twenty years of p-adic model theory. *Logic Colloquium '84 Elsevier Science Publishers B. K (North-Holland)*, 1986.

[Sch12] Peter Scholze. Perfectoid spaces. *Publ. Math. Inst. Hautes Études Sei. 116*, November 2012.

[Zie72] Martin Ziegler. *Die elementare Theorie der henselschen Körper.* Dissertation thesis, 1972.

A Triple Uniqueness of the Maximum Entropy Approach

Jürgen Landes*

Munich Center for Mathematical Philosophy
LMU Munich Munich Germany
juergen_landes@yahoo.de

Abstract

Inductive logic is concerned with assigning probabilities to sentences given probabilistic constraints. The maximum entropy approach to inductive logic assigns probabilities to all sentences of a first order predicate logic. This assignment is built on an application of the Maximum Entropy Principle. This paper puts forward two different modified applications of this principle and shows that the original and the modified applications agree in many cases. A third promising modification is studied and rejected.

1 Introduction

Inductive logic is a formal approach to model uncertain inferences. It seeks to analyse the degree to which premisses entail putative conclusions. Given uncertain premisses $\varphi_1, \ldots, \varphi_k$ with attached uncertainties $X_1, \ldots, X_k$ an inductive logic provides means to attach an uncertainty Y to a conclusion ψ, where the X_i and Y are non-empty subsets of the unit interval. Using $\approx$ to denote an inductive entailment relation this can be represented as

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

where $\approx$ denotes an inductive entailment relation [4].

The main early proponent of inductive logic was Rudolf Carnap [2]. Nowadays, the spirit of his approach today continues in the Pure Inductive Logic approach [7, 8, 14]. In this paper, I however consider uncertain inference within the Maximum Entropy Principle, which goes back to Edwin Jaynes [5]. Roughly speaking, the Maximum Entropy Principle compels rational agents to use a probability function consistent with the evidence for drawing uncertain inferences. In case there is more than one such probability function, a rational agent ought to use one of those probability functions that has maximal entropy.

If the underlying domain is finite, then applying the Maximum Entropy Principle for inductive entailment is straight-forward and well-understood due to the work of Alena Vencovská & Jeff Paris [11, 12, 13]. Matters change dramatically for infinite domains. Naively replacing the sum by an integral in the definition of Shannon Entropy produces probability functions with infinite entropy. But then there is no way to pick a probability function with maximal entropy out of a set in which all functions have infinite entropy.

There are two different suggestions for inductive logic on an infinite first order predicate logic explicating the Maximum Entropy Principle. The entropy limit approach [1] precedes the maximum entropy approach [17, 18]. It has been conjectured, that both approaches agree in cases in which the former approach is-well defined [18, p. 191]. This conjecture has been shown to hold in a number of cases of evidence bases with relatively low quantifier-complexity [6, 9, 16].

This paper introduces modifications of the maximum entropy approach and studies their relationships. I next properly introduce this approach, the modifications and investigate their relationships.

*Many thanks to Soroush Rafiee Rad and Jon Williamson for continued collaboration on these topics.

2 The Maximum Entropy Approach and Modifications

The formal framework and notation is adapted from [9].

Given is a fixed first-order predicate language $\mathcal{L}$ with countably many constant symbols $t_1, t_2, \ldots$ and finitely many relation symbols, $U_1, \ldots, U_n$. The atomic sentences are 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 $S\mathcal{L}$.

The finite sublanguages $\mathcal{L}_n$ of $\mathcal{L}$ are those languages, which only contain the first n constant symbols $t_1, \ldots, t_n$ and the same relation symbols as $\mathcal{L}$. The sentences of the form $\pm a_1 \wedge \ldots \wedge \pm a_{r_n}$ are called the *n-states*. Let Ω_n be the set of n-states for each n. Denote the sentences of $\mathcal{L}_n$ by $S\mathcal{L}_n$.

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

A probability function on $\mathcal{L}_n$ *is defined similarly (the supremum in P3 is dropped and* m *is equal to* n*).* $\mathbb{P}$ *denotes the set of all probability functions on* $\mathcal{L}$*. The set of probability functions consistent with all premisses is denoted by* $\mathbb{E}$*,* $\mathbb{E} := \{P \in \mathbb{P} \ : \ P(\varphi_i) \in X_i \text{ for all } 1 \leq i \leq k\}$*.*

A probability function $P \in \mathbb{P}$ is determined by the values it gives to the quantifier-free sentences, a result known as *Gaifman's Theorem* [3]. Consequently, a probability function is determined by the values it gives to the n-states, for each n. It is thus sensible to measure entropy of P via n-states with varying n.

Definition 2 (n-entropy)**.** *The* n-entropy *of a probability function* P *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* $\log$ *denoting the natural logarithm. The second convention is inconsequential for current purposes.* $H_n(\cdot)$ *is a strictly concave function.*

The key idea is to combine the n-entropies defined on finite sublanguages into an overall notion of comparative entropy comparing probability functions P and Q defined on the entire first order language. So far, the literature has only studied such inductive logics with respect to the first binary relation in the following definition.

Definition 3 (Comparative Notions of Entropy)**.** *That a probability function* $P \in \mathbb{P}$ has greater (or equal) entropy than *a probability function* $Q \in \mathbb{P}$ *could be defined in the following three ways.*

1. *If and only if there is some natural number* N *such that for all* $n \geq N$ *it holds that* $H_n(P) > H_n(Q)$*, denoted by* $P \succ Q$*.*

2. *If and only if there is some natural number* N *such that for all* $n \geq N$ *it holds that* $H_n(P) \geq H_n(Q)$ *and there are infinitely many* n *such that* $H_n(P) > H_n(Q)$*, denoted by* $P]Q$*.*

3. *If and only if there is some natural number N such that for all $n \geq N$ it holds that $H_n(P) \geq H_n(Q)$, denoted by $P)Q$.*

The lower two definitions are alternative ways in which one could explicate the intuitive idea of comparative entropy.

Definition 4 (Maximum Entropy)**.** *The set of probability functions on $\mathcal{L}$ with maximal entropy relative to a notion of comparative entropy $>$ defined on $\mathbb{P}^2$ can then be defined as*

$$\text{maxent}_> \mathbb{E} :=\{P \in \mathbb{E} : \text{ there is no } Q \in \mathbb{E} \setminus \{P\} \textit{ with } Q > P\} \ . \tag{1}$$

Definition 5 (Maximum Entropy Inductive Logics)**.** *An* inductive logic *with respect to $>$ is induced by attaching uncertainty $Y_>(\psi) \subseteq [0,1]$ to the sentences ψ of $\mathcal{L}$ via*

$$Y_>(\psi) := \{r \in [0,1] \mid \textit{ there exists } P \in \text{maxent}_> \mathbb{E} \textit{ with } P(\psi) = r\} \ .$$

In case there are two or more different probability functions in $\text{maxent}_> \mathbb{E}$, there are some sentences of ψ of $\mathcal{L}$ to which multiple different probabilities attach.

In the next section, I study (the relationships of) these binary relations and the arising inductive logics. Particular attention is paid to the case of a unique probability function for inference, $|\text{maxent}_> \mathbb{E}| = 1$.

3 Maximal (Modified) Entropy

I first consider notions of refinement relating these three binary relations.

Definition 6 (Strong Refinement)**.** *$>$ is called a* strong refinement *of $\gg$, if and only if the following hold*

- *$>$ is a refinement of $\gg$, for all $P, Q \in \mathbb{P}$ it holds that $P \gg Q$ entails $P > Q$,*
- *for all $R, P, Q \in \mathbb{P}$ it holds that, if $R \gg P$ and $P > Q$, then $R \gg Q$ and $R \neq Q$.*

Definition 7 (Centric Refinement)**.** *I call a refinement $>$ of $\gg$* centric*, if and only if for all different $R, P \in \mathbb{P}$ with $R > P$ it holds that $(R + P)/2 \gg P$.*

Clearly, not all binary relations possess strong refinements; not all binary relations possess centric refinements.

Proposition 1. *$]$ is a strong and centric refinement of $\succ$. $)$ is a strong and centric refinement of $]$ and of $\succ$.*

Proof. I now display the three notions of comparative entropy line by line. The second conjunct in the first definition is superfluous as is the second conjunct in the third definition:

$$\begin{aligned}
P \succ Q :&\Longleftrightarrow (H_n(P) \leq H_n(Q) \text{ not infinitely often} \,\&\, H_n(P) > H_n(Q) \text{ infinitely often}) \\
P]Q :&\Longleftrightarrow (H_n(P) < H_n(Q) \text{ not infinitely often} \,\&\, H_n(P) > H_n(Q) \text{ infinitely often}) \\
P)Q :&\Longleftrightarrow (H_n(P) < H_n(Q) \text{ not infinitely often} \,\&\, H_n(P) \geq H_n(Q) \text{ infinitely often}) \ .
\end{aligned}$$

By thusly spelling out both comparative notions of entropy one observes that $P \succ Q$ entails $P]Q$, and that $P]Q$ entails $P)Q$. This establishes the refinement relationships.

Strong Refinements Next note that, if $R \succ Q$ or if $R]Q$, then $R \neq Q$.

$]$ is a strong refinement of $\succ$: Let $R \succ P$ and $P]Q$. Then $R \neq Q$. Furthermore, $H_n(R) \leq H_n(Q)$ is true for at most finitely many n, since from some N onwards P has always greater or equal n-entropy than Q. So, $R \succ Q$.

$)$ is a strong refinement of $]$: Let $R]P$ and $P)Q$. Then $R \neq Q$. From some N onwards P has always greater or equal n-entropy than Q. There are also infinitely many $n \in \mathbb{N}$ such that $H_n(R) > H_n(P)$. So, $R]Q$.

$)$ is a strong refinement of $\succ$: Let $R \gg P$ and $P)Q$. Then $R \neq Q$. From some N onwards P has always greater or equal n-entropy than Q. From some N' onwards R has always greater n-entropy than P. Hence, $H_n(R) \leq H_n(Q)$ can only be the case for finitely many $n \in \mathbb{N}$. So, $R \succ Q$.

Centric Refinement First, note that different probability functions disagree on some quantifier free sentence $\varphi \in L_N$ (*Gaifman's Theorem* [3]). Since $\varphi \in L_{n+N}$ for all $n \geq 1$, these probability functions also disagree on all more expressive sub-languages L_{n+N}.

$]$ is a centric refinement of $\succ$: Fix arbitrary probability functions R, P defined on $\mathcal{L}$ with $R]P$. $R \neq P$. From the concavity of the function H_n it follows that $H_n(\frac{R+P}{2}) > H_n(P)$, whenever $H_n(R) \geq H_n(P)$. By definition of $]$, there are only finitely many n for which $H_n(R) \geq H_n(P)$ fails to hold. Hence, $\frac{R+P}{2} \succ P$ by definition of $\succ$.

$)$ is a centric refinement of $\succ$: Fix arbitrary probability functions R, P defined on $\mathcal{L}$ with $R)P$. Note that R may be equal to P. From the concavity of the function H_n it follows that $H_n(\frac{R+P}{2}) > H_n(P)$, whenever $H_n(R) \geq H_n(P)$. By definition of $)$, there are only finitely many n for which $H_n(R) \geq H_n(P)$ fails to hold. Hence, $\frac{R+P}{2} \succ P$ by definition of $\succ$.

$)$ is a centric refinement of $]$: Fix arbitrary probability functions R, P defined on $\mathcal{L}$ with $R)P$. Note that R may be equal to P. Since $\frac{R+P}{2} \succ P$ (see above case) and since $]$ is a refinement of $\succ$, it holds that $\frac{R+P}{2}]P$. $\square$

Remark 1 (Properties of Comparative Entropies)**.** *If $H_n(P) = H_n(Q)$ for all even n and $H_n(P) > H_n(Q)$ for all odd n, then $P]Q$ and $P \not\succ Q$. Hence, $]$ is a proper refinement of $\succ$.*

For $P = Q$ it holds that $P)Q$ and $Q)P$. Hence, $)$ is a proper refinement of $]$ and thus a proper refinement of $\succ$.

$]$ is transitive, irreflexive, acyclic and asymmetric. $)$ is transitive, reflexive and has non-trivial cycles, e.g, for all probability functions P, Q with zero-entropy, $H_n(P) = 0$ for all $n \in \mathbb{N}$, it holds that $P)Q$.

I now turn to entropy maximisation and the induced inductive logics.

Proposition 2. *Let $>$ be a strong refinement of $\gg$. If $\{Q\} = \mathrm{maxent}_{\gg}\mathbb{E}$, then $\{Q\} = \mathrm{maxent}_{\gg}\mathbb{E} = \mathrm{maxent}_{>}\mathbb{E}$.*

Proof. Note at first that since $>$ is a refinement of $\gg$ it holds that

$$\mathrm{maxent}_{>}\mathbb{E} \subseteq \mathrm{maxent}_{\gg}\mathbb{E} \ . \tag{2}$$

Maximal elements according to $\gg$ may not be maximal according to $>$ and all maximal elements according to $>$ are also maximal according to $\gg$.

Assume for the purpose of deriving a contradiction that $Q \notin \mathrm{maxent}_{>}\mathbb{E}$. Then, there has to exist a $P \in \mathbb{E} \setminus \{Q\}$ such that $P > Q$ but $P \gg Q$ fails to hold ($\{Q\} = \mathrm{maxent}_{\gg}\mathbb{E}$).

However, since $\{Q\} = \mathrm{maxent}_{\gg}\mathbb{E}$ and $Q \notin \mathrm{maxent}_{>}\mathbb{E}$ hold, there has to exist some $R \in \mathbb{E} \setminus \{P\}$ such that $R \gg P$, P cannot have maximal $\gg$-entropy. We hence have $R \gg P$ and $P > Q$. Since $>$ is a strong refinement of $\gg$, we obtain $R \gg Q$ and $R \neq Q$. Since $R \in \mathbb{E}$ it follows from the definition of $\mathrm{maxent}_{\gg}$ that $Q \notin \mathrm{maxent}_{\gg}\mathbb{E}$. Contradiction. So, $Q \in \mathrm{maxent}_{>}\mathbb{E}$.

Since $\{Q\} = \mathrm{maxent}_{\gg}\,\mathbb{E} \overset{(2)}{\supseteq} \mathrm{maxent}_{>}\,\mathbb{E} \ni Q$, it follows that $\mathrm{maxent}_{>}\,\mathbb{E} = \{Q\}$. □

Proposition 3. *If $\mathbb{E}$ is convex, $>$ is a centric refinement of $\gg$ and $\{Q\} = \mathrm{maxent}_{>}\,\mathbb{E}$, then $\{Q\} = \mathrm{maxent}_{\gg}\,\mathbb{E} = \mathrm{maxent}_{>}\,\mathbb{E}$.*

Proof. Assume for contradiction that there exists a $P \in \mathbb{E}\setminus\{Q\}$ such that P is not $\gg$-dominated by the probability functions in $\mathbb{E}$ but $>$-dominated by some $R \in \mathbb{E} \setminus \{P\}$, $R > P$. Now define $S = \frac{1}{2}(P + R)$ and note that $S \in \mathbb{E}$ (convexity) and that S, P, R are pairwise different, $|\{S, P, R\}| = 3$.

Since $>$ is a centric refinement of $\gg$, conclude that $S \gg P$, which contradicts that $P \in \mathrm{maxent}_{\gg}\,\mathbb{E}$ and $P \neq Q$. So, only Q can be in $\mathrm{maxent}_{\gg}\,\mathbb{E}$.

Since $Q \in \mathrm{maxent}_{>}\,\mathbb{E}$ and $\mathrm{maxent}_{>}\,\mathbb{E} \overset{(2)}{\subseteq} \mathrm{maxent}_{\gg}\,\mathbb{E}$ it follows that $\{Q\} = \mathrm{maxent}_{\gg}\,\mathbb{E}$. □

Theorem 1 (Triple Uniqueness)**.** *If $\mathbb{E}$ is convex and at least one of $\mathrm{maxent}_{)}\,\mathbb{E}$, $\mathrm{maxent}_{]}\,\mathbb{E}$ or $\mathrm{maxent}_{\succ}\,\mathbb{E}$ is a singleton, then*

$$\mathrm{maxent}\,\mathbb{E}_{)} = \mathrm{maxent}_{]}\,\mathbb{E} = \mathrm{maxent}_{\succ}\,\mathbb{E}\ \ .$$

Proof. Simply apply the above three propositions. □

Having studied refinements of $\succ$, I now briefly consider how $\succ$ could refine a binary relation. Closest to the spirit of Definition 3 would be to consider $P\}Q$, if and only if $H_n(P) > H_n(Q)$ for all $n \in \mathbb{N}$. Clearly, the other three notions of comparative entropy are refinements of $\}$.

Neither of these three binary relations is a strong refinement and neither is a centric refinement. To see this, consider three pairwise different probability functions P, Q, R with i) $H_n(P) > H_n(Q)$ for all n , ii) $H_n(P)/H_n(Q) \approx 1$, iii) $H_1(Q) = H_1(R) - \delta$ for large $\delta > 0$ and iv) $H_n(Q) > H_n(R)$ for all $n \geq 1$. Then $P\}Q$ and $Q \succ R, Q]R, Q)R$. Now note that $H_1(P) < H_1(R)$ and thus $P\}R$ fails to hold. None of $\succ,],)$ is a strong refinement of $\}$. Finally, observe that $\frac{Q+R}{2}\}R$ fails to hold. None of $\succ,],)$ is a centric refinement of $\}$.

The binary relation $\}$ induces a different inductive logic than $\succ,],)$:

Example 1. *Let U be the only and unary relation symbol of $\mathcal{L}$. Suppose there is no evidence, $\mathbb{E} = \mathbb{P}$. Then every $P \in \mathbb{P}$ with $P(Ut_1) = P(\neg Ut_1) = 0.5$ has maximal 1-entropy. Hence, all such P are members of $\mathrm{maxent}_{\}}\,\mathbb{E}$. For $\square \in \{\succ,],)\}$ it holds that $\mathrm{maxent}_{\square}\,\mathbb{E} = P_{=}$, where $P_{=}$ denotes the equivocator function, which for all n assigns all n-states the same probability of $1/|\Omega_n|$. So, $\mathrm{maxent}_{\square}\,\mathbb{E} \neq \mathrm{maxent}_{\}}\,\mathbb{E}$.*

This leads to the following more general observation:

Proposition 4. *If there exists an $n \in \mathbb{N}$ such $H_n(P) = \max\{H_n(Q) : Q \in \mathbb{E}\}$, then $P \in \mathrm{maxent}_{\}}\,\mathbb{E}$.*

This strong focus on single sublanguages $\mathcal{L}_n$ makes $\mathrm{maxent}_{\}}$ unsuitable as an inductive logic for infinite predicate languages.

4 Conclusions

Maximum entropy inductive logic on infinite domains lacks a paradigm approach. The entropy limit approach, the maximum entropy approach as well as the here studied modified maximum entropy approaches induce a unique inductive logic in a number of natural cases. This points towards a, perhaps surprisingly, unified picture of maximum entropy inductive logics – in spite of the number possible ways to define such inductive logics.

The Maximum Entropy Approach fails to provide probabilities for uncertain inference for certain evidence bases of quantifier complexity Σ_2 [15, § 2.2]. In these cases, for all $P \in \mathbb{E}$ there exists a $Q \in \mathbb{E}$ such that $Q \succ P$ and maxent $\mathbb{E}$ is hence empty [10]. One way to sensibly define an inductive logic could be to consider a binary relation which is refined by $\succ$. Unfortunately, the most obvious way fails to deliver a sensible inductive logic (Proposition 4). Finding a way to define such a sensible inductive logic must be left to further study.

References

[1] Owen Barnett and Jeff B. Paris. Maximum Entropy Inference with Quantified Knowledge. *Logic Journal of IGPL*, 16(1):85–98, 2008.

[2] Rudolf Carnap. The Two Concepts of Probability: The Problem of Probability. *Philosophy and Phenomenological Research*, 5(4):513–532, 1945.

[3] Haim Gaifman. Concerning measures in first order calculi. *Israel Journal of Mathematics*, 2(1):1–18, 1964.

[4] Rolf Haenni, Jan-Willem Romeijn, Gregory Wheeler, and Jon Williamson. *Probabilistic Argumentation*, volume 350 of *Synthese Library*. Springer, 2011.

[5] Edwin T Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, 2003.

[6] Jürgen Landes. The Entropy-limit (Conjecture) for Σ_2-Premisses. *Studia Logica*, 109:423–442, 2021.

[7] Jürgen Landes, Jeff B. Paris, and Alena Vencovská. Some Aspects of Polyadic Inductive Logic. *Studia Logica*, 90(1):3–16, 2008.

[8] Jürgen Landes, Jeff B. Paris, and Alena Vencovská. A survey of some recent results on Spectrum Exchangeability in Polyadic Inductive Logic. *Synthese*, 181:19–47, 2011.

[9] Jürgen Landes, Soroush Rafiee Rad, and Jon Williamson. Towards the Entropy-Limit Conjecture. *Annals of Pure and Applied Logic*, 172, 2021.

[10] Landes, Jürgen and Rafiee Rad, Soroush and Williamson, Jon. Determining maximal entropy functions for objective Bayesian inductive logic. Forthcoming.

[11] Jeff B. Paris. Common Sense and Maximum Entropy. *Synthese*, 117:75–93, 1998.

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

[13] Jeff B. Paris and Alena Vencovská. In Defense of the Maximum Entropy Inference Process. *International Journal of Approximate Reasoning*, 17(1):77–103, 1997.

[14] Jeff B. Paris and Alena Vencovská. *Pure Inductive Logic*. Cambridge University Press, 2015.

[15] Soroush Rafiee Rad. Equivocation axiom on first order languages. *Studia Logica*, 105(1):121–152, 2017.

[16] Soroush Rafiee Rad. Maximum Entropy Models for Σ_1 Sentences. *Journal of Applied Logics - IfCoLoG Journal of Logics and their Applications*, 5(1):287–300, 2018.

[17] Jon Williamson. Objective Bayesianism with predicate languages. *Synthese*, 163(3):341–356, 2008.

[18] Jon Williamson. *Lectures on Inductive Logic*. Oxford University Press, Oxford, 2017.

The placeholder view of assumptions and the Curry–Howard correspondence

Ivo Pezlar[1]

Czech Academy of Sciences, Institute of Philosophy, Prague, Czech Republic
pezlar@flu.cas.cz

Abstract

Proofs from assumptions are amongst the most fundamental reasoning techniques. Yet the precise nature of assumptions is still an open topic. One of the most prominent conceptions is the placeholder view of assumptions generally associated with natural deduction for intuitionistic propositional logic. It views assumptions essentially as holes in proofs (either to be filled with closed proofs of the corresponding propositions via substitution or withdrawn as a side effect of some rule), thus in effect making them an auxiliary notion subservient to proper propositions. The Curry-Howard correspondence is typically viewed as a formal counterpart of this conception. In this talk, based on my paper of the same name (*Synthese*, 2020), I will argue against this position and show that even though the Curry-Howard correspondence typically accommodates the placeholder view of assumptions, it is rather a matter of choice, not a necessity, and that another more assumption-friendly view can be adopted.

Introduction. Proofs from assumptions are amongst the most fundamental reasoning techniques. Yet the precise nature of assumptions is still an open topic. One of the most prominent conceptions is the placeholder view of assumptions generally associated with natural deduction for intuitionistic propositional logic. It views assumptions essentially as holes in proofs (either to be filled with closed proofs of the corresponding propositions via substitution or withdrawn as a side effect of some rule), thus in effect making them an auxiliary notion subservient to proper propositions (see, e.g., [14], p. 5). The Curry-Howard correspondence is typically viewed as a formal counterpart of this conception (recently, see, e.g., [12]). I this talk, based on my paper [8], I will argue against this position and show that even though the Curry-Howard correspondence typically accommodates the placeholder view of assumptions, it is rather a matter of choice, not a necessity, and that another more assumption-friendly view can be adopted.

Assumption withdrawing. The rule for implication introduction from natural deduction for intuitionistic propositional logic is arguably the best-known example of the assumption withdrawing rule:

$$\dfrac{\begin{array}{c}[A]\\ \vdots \\ B\end{array}}{A \supset B}$$

It prescribes the following inference step: if we can derive B from assumption A, then we can derive $A \supset B$ and withdraw the initial assumption A (it is worth noting that other assumptions than A may be used in deriving B and those remain open after discharging A). Note that this rule effectively embodies the deduction theorem from standard axiomatic systems. In other words, the implication introduction rule is internalizing structural information from the proof level ("B is derivable from A") to the propositional level ("A implies B").[1]

The problematic aspect of this and other assumption withdrawing rules stems from the fact that it behaves differently from the non-assumption withdrawing rules. More specifically, with implication introduction rule we are deriving the proposition $A \supset B$ not from other propositions as with other standard rules (e.g., conjunction introduction), but from a hypothetical proof. To put it differently, the inference step validated by the implication introduction takes us from a *derivation* starting with a hypothesis to a proposition, not just from propositions to another proposition as do rules without assumptions.[2]

For example, consider the following simple proof of the theorem $A \supset ((A \supset B) \supset B)$ of propositional logic:

$$\dfrac{\dfrac{\dfrac{[A \supset B]^1 \qquad [A]^2}{B}\,\supset\mathrm{E}}{(A \supset B) \supset B}\,\supset\mathrm{I}_1}{A \supset ((A \supset B) \supset B)}\,\supset\mathrm{I}_2$$

We start by making two assumptions $A \supset B$ and A. Applying the implication elimination rule (modus ponens) we derive B. What follows are two consecutive applications of implication introduction rule, first withdrawing the assumption $A \supset B$, the second withdrawing the assumption B. Note that it is the fact that B is derivable from $A \supset B$ together with A that warrants the application of the implication introduction rule and the derivation of the corresponding proposition $(A \supset B) \supset B$, at that moment still depending on the assumption A. Analogously with the second application of the implication introduction rule that withdraws this remaining assumption.

A proof that relies on no assumptions is called a closed proof. If a proof depends on some assumptions that are yet to be withdrawn (i.e., open/active assumptions) it is called an open proof. For example, our derivation of $A \supset ((A \supset B) \supset B)$ constitutes a closed proof, since both assumption were withdrawn in the course of the derivation. Assuming we would not have carried out the last inference step, we would get an open proof:

$$\dfrac{\dfrac{[A \supset B]^1 \qquad A}{B}\,\supset\mathrm{E}}{(A \supset B) \supset B}\,\supset\mathrm{I}_1$$

since the assumption A, upon which the derivation of $((A \supset B) \supset B)$ depends, is still active.

Closed proofs are usually preferred to open ones for the simple reason that closed proofs are generally viewed as the fundamental notion in standard proof-theoretic systems. From this perspective, assumptions are just temporary holes in the proof that are preventing us from reaching a closed proof. These open holes can be are either completely discarded via assumption withdrawing rules or filled in with other already closed proofs via substitution. This is the reason why [12] and others[3] call this the placeholder view of assumptions: active assumptions are just auxiliary artefacts of the employed proof system that behave differently than proper propositions, i.e., propositions that do not appear as assumptions.

[1] [12] describes this as a two-layer system. Note that, strictly speaking, the assumptions are not really withdrawn, they are rather incorporated into the propositional level in the form of an antecedent.

[2] This non-standard behaviour is also the reason why [9] describes assumption withdrawing rules as improper rules and introduces the distinction between inference rules and deductions rules. For more, see [9], [7].

[3] See, e.g., [1]

The Curry-Howard correspondence. The placeholder view of assumptions is also supported to a large extent by the Curry-Howard correspondence in its basic form which links typed lambda calculus and implicational fragment of intuitionistic propositional logic.[4] Under this correspondence, natural deduction assumptions correspond to free variables of lambda calculus, which fits well with the interpretation of assumptions as open holes in the proof.

For example, assuming only the implicational fragment of intuitionistic propositional natural deduction, we get the following correspondences between the propositional and functional dimensions of the Curry-Howard correspondence:

Natural Deduction	Lambda Calculus
assumption	free variable
implication introduction	function abstraction
implication elimination	function application

Under this correspondence, the implication introduction rule will then look as follows:

$$\dfrac{\begin{array}{c}[x : A]\\ \vdots \\ b(x) : B\end{array}}{\lambda x.b(x) : A \supset B}$$

Note that the act of withdrawing the assumption A corresponds to λ-binding of the free variable x. The whole proof of the theorem $A \supset ((A \supset B) \supset B)$ would then proceed in the following way:

$$\dfrac{\dfrac{\dfrac{[x : A \supset B]^1 \qquad [y : A]^2}{xy : B}\,{\supset}\mathrm{E}}{\lambda x.xy : (A \supset B) \supset B}\,{\supset}\mathrm{I}_1}{\lambda y.\lambda x.xy : A \supset ((A \supset B) \supset B)}\,{\supset}\mathrm{I}_2$$

with the concluding proof object (closed term) $\lambda y.\lambda x.xy$ with no free variables representing the final closed proof with no active assumptions. In contrast, the open proof discussed earlier:

$$\dfrac{\dfrac{[x : A \supset B]^1 \qquad y : A}{xy : B}\,{\supset}\mathrm{E}}{\lambda x.xy : (A \supset B) \supset B}\,{\supset}\mathrm{I}_1$$

concludes with the proof object $\lambda x.xy$ that still contains the free variable y corresponding to the yet to be withdrawn assumption A.

The placeholder view of assumptions and consequence statements. The Curry-Howard correspondence is generally viewed as incorporating the placeholder view of assumptions. Probably most recently, this point was explicitly made in [12]. Furthermore, in the same paper Schroeder-Heister advocates for a more general concept of inference that takes us not from propositions to other propositions, but from (inferential) consequence statements $A \models B$ to other consequence statements in order to, amongst other things, equalize the status of assumptions and assertions. The general form of inference rules he discusses is the following:

[4] See, e.g., [13].

$$\frac{A_1 \models B_1 \quad \ldots \quad A_n \models B_n}{C \models D}$$

where the antecedents can be empty. As he explains:

> This corresponds to the idea that in natural deduction, derivations can depend on assumptions. Here this dependency is expressed by non-empty antecedents, as is the procedure of the sequent calculus. Our model of inference is the sequent-calculus model. . . ([11], p. 938)

To show that this rule is correct, we have demonstrate that given the grounds for the premises (denoted as $g : A \models B$) we can construct grounds for the conclusion. In other words, the grounds of the conclusion have to contain some operation f transforming the grounds for the premises to the grounds for the conclusion. Schematically:

$$\frac{g_1 : A_1 \models B_1 \quad \ldots \quad g_n : A_n \models B_n}{f(g_1, \ldots g_n) : C \models D}$$

Schroeder-Heister comments on this rule as follows:

> . . . [H]andling of grounds in the sense described is different from that of terms in the typed lambda calculus. When generating grounds from grounds according to [the rule immediately above], we consider grounds for whole sequents, whereas in the typed lambda calculus terms representing such grounds are handled within sequents. So the notation $g : A \models B$ we used above, which is understood as $g : (A \models B)$, differs from the lambda calculus notation $x : A \vdash t : B$, where t represents a proof of B from A and the declaration $x : A$ on the left side represents the assumption A. ([11], p. 939)

However, it should be mentioned that he left it "open how to formalize grounds and their handling." (ibid., p. 938) I will argue that even though lambda calculus with the Curry-Howard interpretation can be seen as embodying the placeholder view of assumptions in the intuitionistic propositional logic, within the family of Curry-Howard correspondence based systems we can consider a generalized approach that is free of this view. This generalized approach will treat consequence statements $A \models B$ as higher-order functions $A \Rightarrow B$ that can be naturally captured in Martin-Löf's constructive type theory ([4]), specifically in its higher-order presentation (see [5], [6]).

Function-based approach to assumptions. Let us return to the implication introduction rule. Adopting the sequent-style notation for natural deduction,[5] we can rewrite this rule as follows:

$$\frac{x : A \vdash b(x) : B}{\vdash \lambda x.b(x) : A \supset B}$$

where the symbol $\vdash$ is used to separate assumptions from (derived) propositions.

Notice that the derivation of B from A is coded with an abstraction term from lambda calculus, which means it captures some sort of a function. Reasoning backwards, this should mean that between the assumption (context) and the conclusion (asserted proposition) has to be a relationship that can be understood functionally, otherwise, we would have nothing to code via lambda terms. To put it differently, there has to be some more fundamental notion of a function at play that we are coding through the concrete abstraction term.

We can try to capture this observation via the following rule:

[5]See, e.g., Gentzen's system NLK, discussed in [15].

$$\frac{x : A \vdash b(x) : B}{f : A \Rightarrow B}$$

where f is to be understood as exemplifying the more fundamental notion of a function that takes us from A to B.

Note that this rule can be roughly understood as the opposite of the implication introduction rule that goes in the other direction: while the implication introduction rule makes the hypothetical derivation "from A is derivable B" in its premise more concrete in the form of implication proposition $A \supset B$ and the corresponding lambda term $\lambda x.b(x)$, this rule makes the derivation more general in the sense that it is now considered as a function f (not specifically a lambda term) from A to B. Also notice that assumptions are no longer placeholders or contexts, but types of arguments for the function f capturing the corresponding derivation. In other words, assumptions now stand equal to proper propositions, they are not just an auxiliary notion captured via free variables.

Furthermore, capturing derivations in this way allows us to consider grounds for the whole consequence statements as Schroeder-Heister required, not just grounds for the conclusions under some assumptions. More specifically, treating consequence statement $A \models B$ as a function type $A \Rightarrow B$ (in accord with the Curry-Howard correspondence) and a ground g as an object f of this type, we can reformulate the general rule as follows:

$$\frac{g_1 : A_1 \Rightarrow B_1 \quad \ldots \quad g_n : A_n \Rightarrow B_n}{f(g_1, \ldots g_n) : C \Rightarrow D}$$

Formalization. So far, I have treated $f : A \Rightarrow B$ informally to mean "f is a function from A to B". Utilizing Martin-Löf's constructive type theory ([4]), specifically its higher-order presentation ([5], [6]), we can capture it more rigorously as a higher-order judgment of the form $(x)b : (A)B$. To explain why, let us return to the hypothetical judgment $x : A \vdash b(x) : B$ that appears as the sole premise of the implication introduction rule. It tells us that we know $b(a)$ to be a proof of the proposition B assuming we know a to be a proof of the proposition A. In other words, the hypothetical judgment $x : A \vdash b(x) : B$ can be seen as stating that $b(x)$ is a function with domain A and range B.[6] This fact, however, cannot be stated directly in the lower-order presentation of constructive type theory. Thus we move towards the higher-order presentation, which is as a generalization of the lower-order presentation using a more primitive notion of a type. The higher-order variant of constructive type theory allows us to form a higher-order notion of a function which can be used to capture the function hidden behind the hypothetical judgment $x : A \vdash b(x) : B$ as an object $(x)b$ of type $(A)B$. Consequently, $(x)b : (A)B$ can then be used to interpret our statement $f : A \Rightarrow B$, as was required. In other words, $(x)b : (A)B$ can be understood as a higher-order judgment declaring that we have (potentially open) derivation of B from A captured by the function $(x)b$.

It is important to emphasize that the higher-order function type $(A)B$ cannot be conflated with the lower-order function type $A \supset B$. The most basic reason is that they are inhabited by different objects: the former by functions, the latter by elements specified by $\supset$-introduction rule, i.e., objects of the form $\lambda x.b(x)$ that are used to code functions. More generally, the notion of a function behind the type $A \supset B$ is parasitic on a more fundamental notion of a function behind the type $(A)B$.[7] From the logical point of view, the main reason we should avoid merging $(A)B$ and $A \supset B$ is that A in $(A)B$ is an assumption of derivation, while A in $A \supset B$ is an antecedent of implication, hence they are objects of different inferential roles. This

[6]See [4].

[7]See [2], [3].

is perhaps best illustrated by the fact that assuming some function f of type $(A)B$ essentially corresponds to assuming a rule $\frac{\mathrm{A}}{\mathrm{B}}$ in Schroeder-Heister's natural deduction with higher-level rules ([10]).

Conclusion. In this talk, I have argued that the Curry-Howard correspondence is not necessarily connected with the placeholder view of assumptions generally associated with natural deduction systems for intuitionistic propositional logic. Although in the basic form of this correspondence, assumptions, which correspond to free variables, can indeed be thought of as just holes to be filled, we can consider also a functional approach where derivations from assumptions are regarded as functions (see [8]). On this account, assumptions are no longer just placeholders but domains of the corresponding functions. From the logical point of view, this move corresponds to the shift from reasoning with propositions to reasoning with consequence statements.

References

[1] Nissim Francez. *Proof-theoretic Semantics*. College Publications, 2015.

[2] Ansten Klev. A comparison of type theory with set theory. In Stefania Centrone, Deborah Kant, and Deniz Sarikaya, editors, *Reflections on the Foundations of Mathematics*. Springer, Cham, 2019.

[3] Ansten Klev. Name of the Sinus Function. In Igor Sedlár and Martin Blicha, editors, *The Logica Yearbook 2018*, London, 2019. College Publications.

[4] Per Martin-Löf. *Intuitionistic type theory: Notes by Giovanni Sambin of a series of lectures given in Padua, June 1980*. Bibliopolis, Napoli, 1984.

[5] Bengt Nordström, Kent Petersson, and Jan M Smith. *Programming in Martin-Löf's type theory: an introduction*. Clarendon Press, 1990.

[6] Bengt Nordström, Kent Petersson, and Jan M Smith. *Martin-Löf's type theory, Handbook of logic in computer science: Volume 5: Logic and algebraic methods*. Oxford University Press, Oxford, 2001.

[7] Ivo Pezlar. Towards a More General Concept of Inference. *Logica Universalis*, 8(1), 2014.

[8] Ivo Pezlar. The Placeholder View of Assumptions and the Curry–Howard Correspondence. *Synthese*, 2020.

[9] Dag Prawitz. *Natural Deduction: A Proof-theoretical Study*. Almqvist & Wiksell, Stockholm, reprinted edition, 1965.

[10] Peter Schroeder-Heister. A natural extension of natural deduction. *Journal of Symbolic Logic*, 49(4):1284–1300, 12 1984.

[11] Peter Schroeder-Heister. The categorical and the hypothetical: a critique of some fundamental assumptions of standard semantics. *Synthese*, 187(3):925–942, 8 2012.

[12] Peter Schroeder-Heister. Open Problems in Proof-Theoretic Semantics. In Thomas Piecha and Peter Schroeder-Heister, editors, *Advances in Proof-Theoretic Semantics*, pages 253–283. Springer, Cham, 2016.

[13] Morten H Sørensen and Pawel Urzyczyn. *Lectures on the Curry-Howard Isomorphism, Volume 149 (Studies in Logic and the Foundations of Mathematics)*. Elsevier Science Inc., New York, NY, USA, 2006.

[14] A. S. Troelstra and H. Schwichtenberg. *Basic Proof Theory*. Cambridge University Press, Cambridge, 2000.

[15] Jan von Plato. Gentzen's proof systems: byproducts in a work of genius. *Bull. Symbolic Logic*, 18(3):313–367, 2012.

A total Solovay reducibility and totalizing of the notion of speedability

Ivan Titov, Wolfgang Merkle

March 2021

Abstract

While the set of Martin-Löf random left-c.e. reals is equal to the maximum degree of Solovay reducibility, Miyabe, Nies and Stephan [5] have shown that the left-c.e. Schnorr random reals are not closed upwards under Solovay reducibility. Recall that for two left-c.e. reals α and β, the former is Solovay reducible to the latter in case there is a partial computable function f and constant c such that for all rational numbers $q < \alpha$ we have

$$\alpha - f(q) < c(\beta - q).$$

By requiring the translation function f to be total, we introduce a total version of Solovay reducibility that implies Schnorr reducibility.Accordingly, by Downey and Griffiths [1], the set of Schnorr random left-c.e. reals is closed upwards relative to total Solovay reducibility.

Furthermore, we observe that the the notion of speedability introduced by Merkle and Titov [4] can be equivalently characterized via partial computable translation functions in a way that resembles Solovay reducibility. By requiring the translation function to be total, we obtain the concept of total speedability. Like for speedability, this notion does not dependent on the choice of the speeding constant.

1 A total version of Solovay reducibility

We first review the usual definition of Solovay reducibility in terms of a partial recursive function [3].

Definition 1.1 (Solovay reducibility, $\leq_{S,c}$). *Let α and β be reals and let $c > 0$ be a rational number. Then α is* Solovay reducible *to β* with respect to a constant *c, written $\alpha \leq_{S,c} \beta$, if there is a partial computable function $\varphi\colon \mathbb{Q} \to \mathbb{Q}$ such that for all $q < \beta$ it holds that $\varphi(q) \downarrow < \alpha$ and $\alpha - \varphi(q) < c(\beta - q)$. The real α is* Solovay reducible *to β, written $\alpha \leq_S \beta$, if α is Solovay reducible to β with respect to some c.*

In case $\alpha \leq_S \beta$, we will also say that α is S-reducible to β, and similarly notation will be used for other reducibilities introduced in what follows.

Definition 1.2 (total Solovay reducibility, $\leq_{S,c}^{tot}$). *A real α is* total Solovay reducible *to a real β* with respect to a constant *c, written $\alpha \leq_{S,c}^{tot} \beta$, if there is a computable function $f\colon \mathbb{Q} \to \mathbb{Q}$ such that for all $q < \beta$ it holds that $f(q) < \alpha$ and $\alpha - f(q) < c(\beta - q)$. The real α is* total Solovay reducible *to β, written $\alpha \leq_S^{tot} \beta$, if α is total Solovay reducible to β with respect to some c.*

The total Solovay reducibility obviously implies the normal one, thus, the Martin-Löf random left c.e. reals are closed upwards relative to the total Solovay reducibility.

2 The structural properties of the $\leq_S^{tot}$ lattice of left-c.e. reals

In this section, we argue that total Solovay reducibility is in Σ_3^0 but is not a standard reducibility in the sense of Downey and Hirschfeldt [3] because neither is addition a join operator nor is there a least degree.

Proposition 2.1. *Total Solovay reducibility is in Σ_3^0.*

Proof. Let $\alpha^0, \alpha^1, \ldots$ be an effective enumeration of left-c.e. reals, where we can assume that for given n on can compute a recursive index for a nondecreasing approximation $a_0^n, a_1^n, \ldots$ to α^n from below. Then we have

$$\begin{aligned}\alpha^a \leq_S^{tot} \alpha^b :\Longleftrightarrow\; & \exists\langle e,c\rangle\forall\langle q,s\rangle\exists\langle r,t\rangle\colon (\varphi_e(q)[t]\downarrow \\ & \wedge (q < b_s \implies (a_r - \varphi_e(q) > 0 \wedge a_r - \varphi_e(q) < c(b_s - q)))).\end{aligned}$$

□

Proposition 2.2. *Let α be a left-c.e. real and let $r > 0$ be a rational number. Then it holds that $r\alpha \equiv_S^{tot} \alpha$.*

Proof. It holds that $r\alpha \leq_S^{tot} \alpha$ via the identify function and constant r, and similarly for a reduction in the reverse direction with constant $1/r$.

□

Next we review the notion of a hyperimmune set.

Definition 2.3. *Let A be an infinite set. By p_A, we denote the principal function of A, i.e., the members of A are $p_A(0) < p_A(1) < \cdots$. Let $k_A(n)$ be the least member of $A \setminus \{0, \ldots, n-1\}$.*

Recall that a set A is HYPERIMMUNE if p_A is not majorized by a computable function, i.e., for no computable function g we have $p_A(n) \leq g(n)$ for all n.

Lemma 2.4. *For any set A, the following assertions are equivalent.*

(i) p_A is not majorized by any computable function

(ii) k_A is not majorized by any computable function

Proof. In case the computable function $g(n)$ majorizes $k_A(n)$, where we can assume that h in nondecreasing, then a computable function that majorizes $p_A(n)$ is given by

$$n \mapsto \underbrace{g(g(\ldots(g(0)\ldots))}_{n\text{-fold application of } g}$$

Conversely, in case the computable function $g(n)$ majorizes $p_A(n)$, then the function $n \mapsto g(n+1)$ majorizes $k_A(n)$. □

Proposition 2.5. *There exists no least degree in the total Solovay degrees.*

Every least set with respect to total Solovay reducibility is also a least set with respect to Solovay reducibility. Since the sets of the latter type are exactly the computable sets, the proposition is immediate from the following lemma.

Lemma 2.6. *Let $\alpha = 0.A(0)\dots$ and $\beta = 0.B(0)\dots$ be reals where the set A is computable and infinite. Then α is total Solovay reducible to β if and only if the set B is not hyperimmune.*

Proof. First assume that B is not hyperimmune. For a dyadic rational q that can be written as $q = 0.\sigma$ where the last letter of σ is equal to 1, define $|q| = |\sigma|$. Then for any such q and σ where $q < \beta$, we have

$$2^{-k_B(|q|)} \leq \beta - 0.\sigma = \beta - q.$$

By Lemma 2.4, we can fix a computable function g that majorizes k_B. We obtain a computable function f witnessing $\alpha \leq_S^{tot} \beta$ by choosing $f(q) < \alpha$ such that we have

$$\alpha - f(q) < 2^{-g(|q|)}.$$

Next assume that α is total Solovay reducible to β via some function f and constant c. Then for every n and for $q_n = 0.B(0)\dots B(n-1)$, we have $q < \beta$, thus, for some appropriate constant d one holds that

$$g(n) := \min_{|\sigma_n|=n} \{\alpha - f(0.\sigma_n) : \alpha - f(0.\sigma_n) > 0\} \leq \alpha - f(q) < c(\beta - q) \leq 2^{-k_B(n)+d}.$$

Consequently, the function $n \mapsto d + \lceil \log g(n) \rceil$ is a computable upper bound for k_B, hence B is not hyperimmune. $\square$

Indeed, the total Solovay-lattice satisfies the following stronger property, which we state here without proof.

Proposition 2.7. *There exists a countably infinite antichain of mutually $\leq_S^{tot}$-incomparable left-c.e. reals such that each of them is incomparable with every computable real.*

Before proving that addition is not a join operator, we recall the notion of a Schnorr reducibility, namely, the uniform version of it.

Definition 2.8 (UNIFORM SCHNORR REDUCIBILITY, $\leq_{uSch,c}$). *A real α is* UNIFORM SCHNORR REDUCIBLE, *or uSch*-REDUCIBLE, *to a real β* WITH RESPECT TO A CONSTANT *c, written $\alpha \leq_{uSch,c} \beta$, if there is a computable functional φ that, given a description of a computable measure machine (or, shortly, cmm) B, returns a description of another computable measure machine $\varphi(B)$, so that*

$$K_{\varphi(B)}(\alpha \upharpoonright n) \leq K_B(\beta \upharpoonright n) + c.$$

The real α is UNIFORM SCHNORR REDUCIBLE *to β, written $\alpha \leq_{uSch} \beta$, if α is uniform Schnorr reducible to β with respect to some c.*

Obviously, the uniform Schnorr reducibility implies the Schnorr reducibility, with respect to which the Schnorr random reals are closed upwards.

Proposition 2.9. *For all left-c.e. α, β, $\alpha \leq_S^{tot} \beta$ implies $\alpha \leq_{uSch} \beta$*

Corollary 2.10. *The Schnorr random left-c.e. reals are closed upwards relative to the total Solovay reducibility.*

Proof. Let f be a total computable function, such that $\alpha \leq_{S,c}^{tot} \beta$ via f. Given a cmm machine B computing β, we construct a cmm machine A computing α in the following uniform way:

Input:$(x \in \mathbb{Q}, w \in \{0,1\}^{\lceil log(c)+1 \rceil})$

- compute $\sigma := B(x)$ (the computation halts iff $x \in dom(B)$)
- compute τ, so that $0.\tau := (f(0.\sigma) \upharpoonright n)$

  If $0.\sigma < \beta$, then on holds

  $$\alpha - f(0.\sigma) < c(\beta - 0.\sigma)$$

  In particular, if $0.\sigma = \beta \upharpoonright n$, then $\beta - 0.\sigma < 2^{-n}$, so $\alpha - f(0.\sigma) < c2^{-n} = 2^{\lceil log(c) \rceil - n}$. Thus,

  $$\alpha \upharpoonright n - 0.\tau < \alpha - f(0.\sigma) + f(0.\sigma) - (f(0.\sigma) \upharpoonright n) < c2^{-n} + 2^{-n} = 2^{\lceil log(c+1) \rceil - n}$$

- return $y \in \{0,1\}^n$, so that $0.y = 0.\tau + 2^{-n} \cdot 0, w$

The constructed machine A has the following properties:

- prefix-freeness (since B is prefix-free)
- computable measure of the domain (the following relation:

  $$B(x) \downarrow \implies A((x,w) \downarrow \forall w \in \{0,1\}^{\lceil log(c+1) \rceil}$$

  implies, that $\mu(dom(A)) = \mu(dom(B))$
- $K_A(\alpha \upharpoonright n) \leq K_B(\beta \upharpoonright n) + log(c+1) + O(1)$ (since there always exists a word $w \in \{0,1\}^{\lceil log(c+1) \rceil}$ such that

  $$\alpha \upharpoonright n - 0.\tau = 2^{-n} \cdot 0.w$$

  For that w, on holds $A(x,w) = y$, such that

  $$0.y = 0.\tau + 2^{-n} \cdot 0, w = \alpha \upharpoonright n$$

  that implies $K_A(\alpha \upharpoonright n) \leq |x| + |w|$, where x may be the shortest code of $\beta \upharpoonright n$.

□

Proposition 2.11. *There is a pair of left-c.e.reals α, β where $\alpha \not\leq_S^{tot} \alpha + \beta$.*

Proof. Miyabe, Nies and Stephan [5, Paragraph 3] demonstrated that there exists a pair of left-c.e. reals α and β such that $\alpha \not\leq_{Sch} \alpha + \beta$. Thus we also have $\alpha \not\leq_S^{tot} \alpha + \beta$ because total Solovay reducibility implies Schnorr reducibility. □

Remark. The uniform Schnorr-reducibility is, due to the similar argumentation, also implied by the weaken version of the total Solovay reducibility, whose requirement for f differs from the original one in the additional term:

$$\alpha - f(q) < c(\beta - q) + 2^{-|q|}.$$

The motivation of this weakening is that now its lattice on the field of left-c.e. reals has a minimal degree containing all the computable reals.

3 Speedability of left-c.e. numbers.

Definition 3.1. *A function $f\colon \mathbb{N} \to \mathbb{N}$ is a* SPEED-UP FUNCTION *if it is nondecreasing and $n \le f(n)$ holds for all n. A left-c.e. number α is* ρ-SPEEDABLE WITH RESPECT TO ITS GIVEN LEFT APPROXIMATION $a_0, a_1, \ldots \nearrow \alpha$ *for some real number $\rho \in (0,1)$ if there is a computable speed-up function f such that we have*

$$\liminf_{n\to\infty} \frac{\alpha - a_{f(n)}}{\alpha - a_n} \le \rho, \tag{1}$$

and SPEEDABLE *if it is ρ-speedable with respect to some its left-c.e. approximation for some $\rho \in (0,1)$. Otherwise we call α nonspeedable.*

Whether a real is speedable depends neither on the left-c.e. approximation nor on the constant ρ one considers.

Theorem 3.2 (Merkle and Titov [4])**.** *Every speedable left-c.e. real number is ρ-speedable for any $\rho > 0$ with respect to any of its left approximations.*

The following theorem is immediate from the main result of Barmpalias and Lewis-Pye [2].

Theorem 3.3 (Barmpalias and Lewis-Pye [2])**.** *Martin-Löf random left-c.e. real numbers are never speedable.*

By the following proposition, the notion of speedability can be equivalently characterized as a Solovay reduction of a real number to itself via a special partial computable functions on the rational numbers. By applying the same characterization to computable functions, in what follows we obtain a variant of speedability, similar to the introduction of total Solovay reducibility.

Proposition 3.4. *Let α be a left-c.e. real and let ρ be a real number such that $0 < \rho < 1$. Then α is speedable if and only if there is a partial computable function $g\colon \mathbb{Q} \to \mathbb{Q}$ that is defined and nondecreasing on the interval $(-\infty, \alpha)$, maps this interval to itself and satisfies*

$$\liminf_{q\nearrow\alpha} \frac{\alpha - g(q)}{\alpha - q} \le \rho. \tag{2}$$

Proof. Fix some left approximation $a_0, a_1, \ldots$ of α. First assume that α is speedable. By Theorem 3.2 there is then a computable speed-up function f that witnesses that α is ρ-speedable with respect to its left approximation $a_0, a_1, \ldots$. Let n be the partial computable function on the set of rational numbers that maps every $q < \alpha$ to the least index i such that $q \le a_i$, and is undefined for all other q. Here we assume that rational numbers are represented in a form such that equality is a computable predicate. Then the partial function g defined by

$$g(q) = a_{f(n(q))},$$

by choice of n and f, is partial computable, is defined and nondecreasing on the interval $(-\infty, \alpha)$ and maps this interval to itself. Furthermore, the sequence $a_0, a_1, \ldots$ witnesses that (2) holds, because we have $g(a_i) = a_{f(i)}$.

Next assume that there is a function g as stated in the proposition. Then there is a not necessarily computable left approximation $q_0, q_1, \dots$ of α such that we have

$$\liminf_{j \to \infty} \frac{\alpha - g(q_j)}{\alpha - q_j} \leq \rho.$$

Let f be the computable speed up function that maps i to the least index $n > i$ such that $g(a_{i+1}) < a_n$. Then for all q and i such that q is an element of the half-open interval $[a_i, a_{i+1})$, we have

$$\frac{\alpha - a_{f(i)}}{\alpha - a_i} \leq \frac{\alpha - g(a_{i+1})}{\alpha - q} \leq \frac{\alpha - g(q)}{\alpha - q}.$$

In particular, this chain of inequalities holds true with q replaced by any of the q_j, which by choice of the q_j implies that α is ρ-speedable via its left approximation $a_0, a_1, \dots$ and the speed-up function f. $\square$

From Proposition 3.4 it is immediate that the the equivalent characterization of speedability stated there does not depend on the choice of ρ in the interval $(0, 1)$. In particular, the characterization holds for some ρ in this interval if and only if it holds for all ρ in this interval.

In a same way as the totalizing of translation function for the Solovay reducibility, we can totalize the concept of speedability by requiring the function g from the latter definition to be total.

Definition 3.5. *Let ρ be a real number such that $0 < \rho < 1$. A left-c.e. real α is called* TOTAL ρ-SPEEDABLE *if there exists a nondecreasing computable function $g \colon \mathbb{Q} \mapsto \mathbb{Q}$ that maps every q in the interval $(-\infty, \alpha)$ to a value $g(q) > q$ in this interval and satisfies*

$$\liminf_{q \nearrow \alpha} \frac{\alpha - g(q)}{\alpha - q} \leq \rho. \tag{3}$$

Such a function g is called TOTAL SPEED-UP FUNCTION.

By the following proposition, the total version of speedability does again not depend on the choice of the constant. The proof is omitted due to space considerations.

Proposition 3.6. *Whether a left-c.e. real is total ρ-speedable does not depend on the choice of $\rho \in (0, 1)$.*

Barmpalias and Lewis-Pye [2] have shown that speedability implies Martin-Löf nonrandomness. We currently research the characteristics of the total speed-up function via which the total speedability will imply Schnorr nonrandomness.

References

[1] Rodney Downey and Evan Griffiths, *Schnorr randomness*, Journal of Symbolic Logic 69(2):533–554 (2004).

[2] George Barmpalias and Andrew Lewis-Pye, *Differences of halting probabilities*, Journal of Computer and System Sciences 89:349–360 (2017).

[3] Rodney Downey and Denis Hirschfeldt, *Algorithmic Randomness and Complexity*, Springer, Berlin (2010).

[4] Wolfgang Merkle and Ivan Titov, *Speedable left-c.e. numbers*, CSR 2020: Computer Science – Theory and Applications pp 303–313 (2020).

[5] Kenshi Miyabe, André Nies and Frank Stephan, *Randomness and Solovay degrees*, Journal of Logic and Analysis 10(3):1–13 (2018).

[6] Joseph Miller, *On work of Barmpalias and Lewis-Pye: A derivation on the d.c.e. reals*, Lecture Notes in Computer Science 10010:644—659 (2016).

[7] André Nies, *Computability and Randomness*, Oxford University Press (2012).

★

Determining maximal entropy functions for objective Bayesian inductive logic

★

Jon Williamson[*]

Draft of June 9, 2021

Abstract

According to the objective Bayesian approach to inductive logic, on a first-order predicate language, premisses inductively entail a conclusion just when every probability function with maximal entropy, from all those that satisfy the premisses, satisfies the conclusion. But it is by no means obvious as to how to determine these maximal entropy functions. This paper makes progress on the problem by introducing the concept of an entropy limit point and showing that, if the set of probability functions satisfying the premisses contains an entropy limit point then this limit point is unique and is the maximal entropy probability function. The paper goes on to show that, in various circumstances, the maximal entropy function is the uniform distribution conditionalised on the premisses.

§1 Objective Bayesian probabilistic logic

An important class of probabilistic logics consider entailment relationships of the following form (Haenni et al., 2011):

$$\varphi_1^{X_1}, \dots, \varphi_k^{X_k} \approx \psi^Y.$$

Here, $\varphi_1, \dots, \varphi_k, \psi$ are sentences of a logical language $\mathscr{L}$ and $X_1, \dots, X_k, Y$ are sets of probabilities. This should be interpreted as saying: $\varphi_1, \dots, \varphi_k$ having probabilities in $X_1, \dots, X_k$ respectively inductively entails that ψ has probability in Y.

One particular approach to inductive logic, the objective Bayesian approach, interprets probabilities as rational degrees of belief and uses Jaynes' Maximum Entropy Principle to determine a rational belief function (Jaynes, 1957). Thus if $\mathscr{L}$ is a finite propositional language, $X_1, \dots, X_k$ are closed convex sets of probabilities, and the premisses are consistent, an entailment relationship holds just when the probability function with maximum entropy, from all those that satisfy the premisses, gives a probability in Y to ψ (Williamson, 2010, Chapter 7).

This approach has been extended to the case in which $\mathscr{L}$ is a first-order predicate language in the following way. Suppose $\mathscr{L}$ has countably many constant symbols $t_1, t_2, \dots$ and finitely many relation symbols $U_1, \dots, U_l$. Let $a_1, a_2, \dots$ run through the atomic sentences of the form $U_i t_{i_1} \cdots t_{i_k}$ in such a way that those atomic sentences involving only $t_1, \dots, t_n$ occur before those involving t_{n+1}, for each n. We will sometimes consider the finite sublanguages $\mathscr{L}_n$, containing only constant symbols $t_1, \dots, t_n$. Let Ω_n be the set of n-states, i.e., state descriptions $\pm a_1 \wedge \cdots \wedge \pm a_{r_n}$ involving the atomic sentences $a_1, \dots, a_{r_n}$ of $\mathscr{L}_n$, which only feature the constants $t_1, \dots, t_n$. Let $S\mathscr{L}, S\mathscr{L}_n$ be the sets of sentences of $\mathscr{L}, \mathscr{L}_n$ respectively. A *probability function* P on $\mathscr{L}$ is a function $P : S\mathscr{L} \longrightarrow \mathbb{R}_{\geq 0}$ 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: $P(\exists x \theta(x)) = \sup_m P\left(\bigvee_{i=1}^m \theta(t_i)\right)$.

A probability function is determined by the values it gives to the n-states (see, e.g., Williamson, 2017, §2.6.3). Of particular importance will be the *equivocator* function, $P_=$, which gives the same probability to each n-state, for each n. We denote the set of probability functions by $\mathbb{P}$. The n-entropy of a probability function P is defined as $H_n(P) \stackrel{\text{df}}{=} -\sum_{\omega \in \Omega_n} P(\omega) \log P(\omega)$. We say that function P has greater entropy than Q if the n-entropy of P dominates that of Q for sufficiently large n, i.e., if there is an $N \in \mathbb{N}$ such that for all $n \geq N$, $H_n(P) > H_n(Q)$. Let $\mathbb{E}$ be the set of probability functions that satisfy the premisses $\varphi_1^{X_1}, \dots, \varphi_k^{X_k}$, i.e., $\mathbb{E} \stackrel{\text{df}}{=} \{P \in \mathbb{P} : P(\varphi_1) \in X_1, \dots, P(\varphi_k) \in X_k\}$. Let $\text{maxent}\,\mathbb{E}$ be the set of maximal entropy functions in $\mathbb{E}$, i.e., the set of probability functions in $\mathbb{E}$ that are not dominated in entropy by any other probability function. (If there are no premisses, $\text{maxent}\,\mathbb{E} = \text{maxent}\,\mathbb{P} = \{P_=\}$.) Then we deem an inductive entailment relationship to hold if $P(\psi) \in Y$ for any $P \in \text{maxent}\,\mathbb{E}$ (Williamson, 2017, §5.3). The entailment relation under this objective Bayesian interpretation is written $\mathrel{\approx}^{\circ}$.

While the objective Bayesian approach provides coherent semantics for inductive logic, it is not obvious how to determine the maximal entropy functions in order to ascertain whether a given entailment relationship holds. This is because the definition of $\text{maxent}\,\mathbb{E}$ seems to require a sort through all members of $\mathbb{E}$ in order to find those with undominated entropy—a process that would be unfeasible in practice. This paper aims to address this question.

§2 introduces the concept of an entropy limit point in order to characterise $\text{maxent}\,\mathbb{E}$ in terms of certain limits of n-entropy maximisers. This gives a constructive procedure for determining $\text{maxent}\,\mathbb{E}$ when it contains an entropy limit point.

In §3 and §4 we consider an important special case—that in which the premisses are categorical sentences $\varphi_1, \dots, \varphi_k$ (without attached probabilities) and where the maximal entropy function can be obtained simply by conditionalising on the equivocator function.

§2 Entropy limit points

This section adapts the techniques of Landes et al. (2021b, §5) in order to characterise $\text{maxent}\,\mathbb{E}$ in terms of certain limits of n-entropy maximisers. Landes et al. (2021b) were concerned with a very different question: that of showing that the above objective Bayesian semantics for inductive logic yields the same inferences

*Department of Philosophy and Centre for Reasoning, University of Kent, Canterbury, UK, j.williamson@kent.ac.uk

as those produced by another method for extending the Maximum Entropy Principle from a finite language to a first-order predicate language. Nevertheless, the results of Landes et al. (2021b, §5) can be straightforwardly adapted to the present problem. Proofs of the two results in this section are very close to those of Landes et al. (2021b, Proposition 36) and Landes et al. (2021b, Theorem 39), but have been included in Appendix 1 for completeness.

In what follows we suppose that $X_1,\dots,X_k$ are convex and that the premisses are satisfiable, i.e., $\mathbb{E} \stackrel{\mathrm{df}}{=} \{P \in \mathbb{P} : P(\varphi_1) \in X_1,\dots,P(\varphi_k) \in X_k\} \neq \emptyset$. We will consider the set of n-entropy maximisers for each n:

$$\mathbb{H}_n = \{P \in \mathbb{E} : H_n(P) \text{ is maximised }\}.$$

We now introduce the key concept of the paper:

Definition 1 (*Entropy Limit Point*). $P \in \mathbb{P}$ is an *entropy limit point* of $\mathbb{P}_1, \mathbb{P}_2, \ldots \subseteq \mathbb{P}$ if for each n there is some $Q_n \in \mathbb{P}_n$ such that $|H_n(Q_n) - H_n(P)| \longrightarrow 0$ as $n \longrightarrow \infty$. $P \in \mathbb{P}$ will be called an *entropy limit point* of $\mathbb{E}$ if it is an entropy limit point of $\mathbb{H}_1, \mathbb{H}_2, \ldots$.

Entropy limit points of $\mathbb{E}$ are of special interest because they are also limit points in terms of the L_1 distance,

$$\|P - Q\|_n \stackrel{\mathrm{df}}{=} \sum_{\omega\in\Omega_n} |P(\omega) - Q(\omega)|.$$

Proposition 2. *If P is an entropy limit point of $\mathbb{E}$ then there are $Q_n \in \mathbb{H}_n$ such that $\|Q_n - P\|_n \longrightarrow 0$ as $n \longrightarrow \infty$.*

This property enables us to characterise the set of maximal entropy functions more constructively, in terms of a limit of n-entropy maximisers:

Theorem 3. *If $\mathbb{E}$ contains an entropy limit point P then*

$$\mathrm{maxent}\,\mathbb{E} = \{P\}.$$

Note that there can be at most one entropy limit point P. This is because $\mathbb{E}$ is convex (by the convexity of $X_1,\dots,X_k$) and the n-entropy maximiser of a convex set is uniquely determined on $\mathscr{L}_n$. Thus the $\mathbb{H}_n$ can have at most one L_1 limit point.

Theorem 3 provides a simple procedure for showing that a hypothesised function P is in fact a maximal entropy function: show that it is an entropy limit point of n-entropy maximisers. (Note that this is only a sufficient condition: if P is not an entropy limit point of $\mathbb{E}$, then we cannot infer that $P \notin \mathrm{maxent}\,\mathbb{E}$.)

Example 4. Suppose we have a single premiss $\forall x U x^{\{c\}}$ where $\mathscr{L}$ has a single unary predicate U. In this case, the number r_n of atomic sentences of $\mathscr{L}_n$ is n. Any n-entropy maximiser gives probability c to the n-state $Ut_1 \wedge \cdots \wedge Ut_n$, which we abbreviate by θ_n, and divides probability $1-c$ amongst all other n-states:

$$P^n(\omega_n) = \begin{cases} c & : \quad \omega_n = \theta_n \\ \frac{1-c}{2^n-1} & : \quad \omega_n \models \neg\theta_n \end{cases}.$$

By the argument of Landes et al. (2021b, Example 42), the following probability function is an entropy limit point:

$$P(\omega_n) = \begin{cases} c + x_n & : \quad \omega_n = \theta_n \\ \frac{1-c-x_n}{2^n-1} & : \quad \omega_n \models \neg\theta_n \end{cases}$$

where $x_n = \frac{1-c}{2^n}$. Hence by Theorem 3, $\mathrm{maxent}\,\mathbb{E} = \{P\}$.

§3
Conditionalisation and entropy limit points

We now consider an important special case of Theorem 3, which links the maximal entropy approach to Bayesian conditionalisation. Suppose the premisses are categorical sentences $\varphi_1, \ldots \varphi_k$ of $\mathscr{L}$, i.e., there are no attached sets of probabilities $X_1,\dots,X_k$, or equivalently, $X_1 = \cdots = X_k = \{1\}$. Let φ be the sentence $\varphi_1 \wedge \cdots \wedge \varphi_k$. In the remainder of the paper, then, we consider $\mathbb{E} = \mathbb{E}_\varphi \stackrel{\mathrm{df}}{=} \{P \in \mathbb{P} : P(\varphi) = 1\}$.

Corollary 5. *If $P_=(\cdot|\varphi)$ is an entropy limit point of $\mathbb{E}_\varphi$ then*

$$\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}.$$

Proof: $P_=(\cdot|\varphi)$ is contained in $\mathbb{E}_\varphi$ because $P_=(\varphi_i|\varphi) = 1$ for each $i = 1,\dots,k$. Hence Theorem 3 applies. □

Note that the condition that $P_=(\cdot|\varphi)$ is an entropy limit point of $\mathbb{E}_\varphi$ presupposes that the probability function $P_=(\cdot|\varphi)$ is well-defined, i.e., that $P_=(\varphi) > 0$. We say that sentence θ has *positive measure* if $P_=(\theta) > 0$.

Corollary 6. *If $\mathbb{H}_n$ contains $P_=(\cdot|\varphi)$ for sufficiently large n then*

$$\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}.$$

Proof: If $P_=(\cdot|\varphi) \in \mathbb{H}_n$ for sufficiently large n then $P_=(\cdot|\varphi)$ is an entropy limit point of $\mathbb{E}$. Hence Corollary 5 applies. □

Corollary 6 is useful because where it applies it provides a particularly simple procedure for determining $\mathrm{maxent}\,\mathbb{E}_\varphi$. Also, it shows that the move to the infinite does not disrupt agreement between the Maximum Entropy Principle and conditionalisation: as long as conditionalising on φ maximises n-entropy for each sufficiently large n, it maximises entropy on the language as a whole. Because of its interest, we provide an alternative, more direct proof of Corollary 6 in Appendix 2.

Example 7. Suppose we have a single categorical premiss $\exists x U x$, where $\mathscr{L}$ has a single unary predicate symbol U. $P_=(\exists x U x) = 1$, so $P_=(\cdot|\exists x U x) = P_=(\cdot)$. $P_= \in \mathbb{H}_1, \mathbb{H}_2, \ldots$, so Corollary 6 applies and $\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=\}$.

Example 8. Suppose we have categorical premisses $Ut_2 \rightarrow Vt_3, \forall x \exists y W x y$, where $\mathscr{L}$ has unary predicate symbols U and V and a binary relation symbol W. Now $P_=((Ut_2 \rightarrow Vt_3) = 0.75)$ and $P_=(\forall x \exists y W x y) = 1$. So $P_=((Ut_2 \rightarrow Vt_3) \wedge \forall x \exists y W x y) = 0.75$, and $P_=(\cdot|(Ut_2 \rightarrow Vt_3) \wedge \forall x \exists y W x y) = P_=(\cdot|Ut_2 \rightarrow Vt_3)$. This latter function is in $\mathbb{H}_3, \mathbb{H}_4, \ldots$, so Corollary 6 applies and $\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|Ut_2 \rightarrow Vt_3)\}$.

§4
An alternative route to conditionalisation

This section demonstrates agreement between the maximal entropy approach and conditionalisation without appeal to entropy limit points. m As above we consider categorical sentences $\varphi_1,\dots,\varphi_k$ and abbreviate $\varphi_1 \wedge \cdots \wedge \varphi_k$ by φ. Let sentence φ_n be the disjunction of those n-states ω such that $\varphi \wedge \omega$ has positive measure:

$$\varphi_n \stackrel{\mathrm{df}}{=} \bigvee\{\omega \in \Omega_n : P_=(\omega \wedge \varphi) > 0\}.$$

The main result will appeal to the following technical lemma, of Landes et al. (2021a), which is stated here without proof:

Lemma 9. *If φ has positive measure then $P_=(\varphi) = P_=(\varphi_n)$ and $P_=(\cdot|\varphi) = P_=(\cdot|\varphi_n)$ for all $n \geq N_\varphi$, the maximum index of the constant symbols that occur in φ.*

Theorem 10. *If φ has positive measure and $\varphi \overset{\circ}{\approx} \varphi_n$ non-trivially for $n \geq N_\varphi$ then*

$$\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\} = \{P_=(\cdot|\varphi_n)\}$$

for any $n \geq N_\varphi$.

Proof: The condition that φ has positive measure ensures that $P_=(\cdot|\varphi)$ exists. That $\varphi \overset{\circ}{\approx} \varphi_n$ ensures that $P(\varphi_n) = 1$ for any $P \in \mathrm{maxent}\,\mathbb{E}_\varphi$ and $n \geq N_\varphi$; that this holds non-trivially ensures that $\mathrm{maxent}\,\mathbb{E}_\varphi \neq \emptyset$.

If $P_=(\varphi) = 1$ then $P_= \in \mathbb{E}_\varphi$ and $P_=(\cdot|\varphi) = P_=(\cdot)$. $P_=$ is the unique member of $\mathrm{maxent}\,\mathbb{E}_\varphi$, so $\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}$, as required.

If $P_=(\varphi) < 1$ then $P_=(\cdot|\neg\varphi)$ is well defined, and we can proceed as follows.

Since $P_=(\varphi) > 0$, $P_=(\cdot|\varphi)$ is well defined. $P_=(\varphi|\varphi) = 1$ so $P_=(\cdot|\varphi) \in \mathbb{E}_\varphi$.

Suppose for contradiction that $\mathrm{maxent}\,\mathbb{E}_\varphi \neq \{P_=(\cdot|\varphi)\}$. Then in $\mathrm{maxent}\,\mathbb{E}_\varphi$ there must be some $P^\dagger \neq P_=(\cdot|\varphi)$ that is not eventually dominated in entropy by $P_=(\cdot|\varphi)$. That is, there is some infinite $J \subseteq \mathbb{N}$ such that $H_n(P^\dagger) \geq H_n(P_=(\cdot|\varphi))$ for all $n \in J$.

Define a probability function Q by:

$$Q(\omega) \overset{\mathrm{df}}{=} P^\dagger(\omega|\varphi)P_=(\varphi) + P_=(\omega|\neg\varphi)P_=(\neg\varphi),$$

for all $\omega \in \Omega_n$. Q is a probability function because it is a convex combination of two probability functions, $P^\dagger(\omega|\varphi)$ and $P_=(\omega|\neg\varphi)$.

$P^\dagger \in \mathbb{E}_\varphi$, so $P^\dagger(\varphi) = 1$ and $P^\dagger(\omega|\varphi) = P^\dagger(\omega)$ for each $\omega \in \Omega_n$ and $n \in \mathbb{N}$. Taking this together with Lemma 9, we have:

$$Q(\omega) = P^\dagger(\omega)P_=(\varphi_n) + P_=(\omega|\neg\varphi_n)P_=(\neg\varphi_n),$$

for all $\omega \in \Omega_n$ and $n \geq N_\varphi$.

Now for $n \geq N_\varphi$,

$$\begin{aligned}
Q(\omega \wedge \varphi_n) &= \sum_{\nu \models \omega \wedge \varphi_n} Q(\nu) \\
&= \sum_{\nu \models \omega \wedge \varphi_n} P^\dagger(\nu)P_=(\varphi_n) + 0P_=(\neg\varphi_n) \\
&= P^\dagger(\omega \wedge \varphi_n)P_=(\varphi_n) \\
&= P^\dagger(\omega)P_=(\varphi_n)
\end{aligned}$$

since by assumption $P^\dagger(\varphi_n) = 1$. Moreover for $n \geq N_\varphi$,

$$\begin{aligned}
Q(\omega \wedge \neg\varphi_n) &= \sum_{\nu \models \omega \wedge \neg\varphi_n} Q(\nu) \\
&= \sum_{\nu \models \omega \wedge \neg\varphi_n} \left(P^\dagger(\nu)P_=(\varphi_n) + P_=(\nu|\neg\varphi_n)P_=(\neg\varphi_n)\right) \\
&= P^\dagger(\omega \wedge \neg\varphi_n)P_=(\varphi_n) + P_=(\omega \wedge \neg\varphi_n|\neg\varphi_n)P_=(\neg\varphi_n) \\
&= 0P_=(\varphi_n) + P_=(\omega|\neg\varphi_n)P_=(\neg\varphi_n) \\
&= P_=(\omega \wedge \neg\varphi_n)
\end{aligned}$$

since by assumption $P^\dagger(\neg\varphi_n) = 0$.

We next see that, for any $n \in J$ where $n \geq N_\varphi$, the equivocator function fails to dominate the Q in n-entropy:

$$\begin{aligned}
H_n(Q) &= -\sum_{\omega \in \Omega_n, \omega \models \varphi_n} Q(\omega)\log Q(\omega) - \sum_{\omega \in \Omega_n, \omega \models \neg\varphi_n} Q(\omega)\log Q(\omega) \\
&= -\sum_{\omega \in \Omega_n} Q(\omega \wedge \varphi_n)\log Q(\omega \wedge \varphi_n) \\
&\quad -\sum_{\omega \in \Omega_n} Q(\omega \wedge \neg\varphi_n)\log Q(\omega \wedge \neg\varphi_n) \\
&= -\sum_{\omega \in \Omega_n} P^\dagger(\omega)P_=(\varphi_n)\log\left(P^\dagger(\omega)P_=(\varphi_n)\right) \\
&\quad -\sum_{\omega \in \Omega_n} P_=(\omega \wedge \neg\varphi_n)\log P_=(\omega \wedge \neg\varphi_n) \\
&= -P_=(\varphi_n)\log P_=(\varphi_n) - P_=(\varphi_n)\sum_{\omega \in \Omega_n} P^\dagger(\omega)\log P^\dagger(\omega) \\
&\quad -\sum_{\omega \in \Omega_n} P_=(\omega \wedge \neg\varphi_n)\log P_=(\omega \wedge \neg\varphi_n) \\
&\geq -P_=(\varphi_n)\log P_=(\varphi_n) - P_=(\varphi_n)\sum_{\omega \in \Omega_n} P_=(\omega|\varphi)\log P_=(\omega|\varphi_n) \\
&\quad -\sum_{\omega \in \Omega_n} P_=(\omega \wedge \neg\varphi_n)\log P_=(\omega \wedge \neg\varphi_n) \\
&= -P_=(\varphi_n)\log P_=(\varphi_n) - P_=(\varphi_n)\sum_{\omega \in \Omega_n} P_=(\omega|\varphi_n)\log P_=(\omega|\varphi_n) \\
&\quad -\sum_{\omega \in \Omega_n} P_=(\omega \wedge \neg\varphi_n)\log P_=(\omega \wedge \neg\varphi_n) \\
&= -\sum_{\omega \in \Omega_n} P_=(\omega \wedge \varphi_n)\log P_=(\omega \wedge \varphi_n) \\
&\quad -\sum_{\omega \in \Omega_n} P_=(\omega \wedge \neg\varphi_n)\log P_=(\omega \wedge \neg\varphi_n) \\
&= H_n(P_=),
\end{aligned}$$

for $n \in J$ such that $n \geq N_\varphi$. The inequality holds since by assumption $H_n(P^\dagger) \geq H_n(P_=(\cdot|\varphi))$ for $n \in J$.

From the definition of Q we see that if Q were the equivocator function then $P_=(\omega|\varphi) = P^\dagger(\omega|\varphi) = P^\dagger(\omega)$ for all n and $\omega \in \Omega_n$. But by assumption, $P^\dagger \neq P_=(\cdot|\varphi)$, so Q differs from $P_=$ on Ω_n for sufficiently large n. Hence, that $H_n(Q) \geq H_n(P_=)$ for any sufficiently large $n \in J$ contradicts the fact that, for each n, $P_{=\restriction\Omega_n}$ is the unique probability function on $\mathscr{L}_n$ that maximises n-entropy. Thus $\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}$.

Moreover, $P_=(\cdot|\varphi) = P_=(\cdot|\varphi_n)$ for $n \geq N_\varphi$ by Lemma 9. □

Finally, we note that Theorem 10 implies agreement between the Maximum Entropy Principle and conditionalisation when φ is (logically equivalent to) some quantifier-free sentence:

Corollary 11. *If φ is satisfiable and logically equivalent to a quantifier-free sentence then*

$$\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}.$$

Proof: Because φ is satisfiable and equivalent to some quantifier free φ', $P_=(\varphi) = P_=(\varphi') > 0$ and so $P_=(\cdot|\varphi)$ exists. For any $P \in \mathbb{E}_\varphi$ and $n \geq N_{\varphi'}$,

$$P(\varphi_n) = \sum_{\omega \in \Omega_n, P_=(\omega \wedge \varphi) > 0} P(\omega) = \sum_{\omega \in \Omega_n, \omega \models \varphi'} P(\omega) = P(\varphi') = P(\varphi) = 1.$$

Hence $P(\varphi_n) = 1$ for any $P \in \mathrm{maxent}\,\mathbb{E}_\varphi$, i.e., $\varphi \overset{\circ}{\approx} \varphi_n$. This holds non-trivially, i.e., $\mathrm{maxent}\,\mathbb{E}_\varphi \neq \emptyset$, by Williamson (2017, Theorem 5.15), which shows that a member of $\mathbb{H}_{N_{\varphi'}}$ can be extended to yield a member of $\mathrm{maxent}\,\mathbb{E}_{\varphi'} = \mathrm{maxent}\,\mathbb{E}_\varphi$. Thus Theorem 10 implies that $\mathrm{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}$, as required. □

This result, which could also have been arrived at via Corollary 6, can be thought of as an analogue of Seidenfeld (1986, Result 1), which demonstrates agreement between the Maximum Entropy Principle and conditionalisation in the finite case.

§5
Conclusion

This paper introduced the concept of an entropy limit point. This concept provides a means of showing that a maximal entropy function P does indeed have maximal entropy. It also forges a link between maximal entropy functions and Bayesian conditionalisation.

Theorem 10 shows that this link also holds under the condition that $\varphi \overset{\circ}{\models} \varphi_n$. It is an open question as to how restrictive this condition is—are there any situations in which it fails?

Acknowledgements

Jon Williamson's research was supported by funding from the Leverhulme Trust (grant RPG-2019-059) and the Deutsche Forschungsgemeinschaft (DFG, grant LA 4093/3-1).

Appendix 1. Proofs of Proposition 2 and Theorem 3

First let us recount some basic information-theoretic facts.

The *n-divergence* of two probability functions P and Q is defined as the Kullback-Leibler divergence of P from Q on $\mathscr{L}_n$:

$$d_n(P,Q) \stackrel{\text{df}}{=} \sum_{\omega\in\Omega_n} P(\omega)\log\frac{P(\omega)}{Q(\omega)}.$$

A Pythagorean theorem holds for the n-divergence d_n (Cover and Thomas, 1991, Theorem 11.6.1):

$$d_n(P,Q) \geq d_n(P,R_n) + d_n(R_n,Q),$$

for any convex $\mathbb{F} \subseteq \mathbb{P}$, if $P \in \mathbb{F}$ and $Q \notin \mathbb{F}$, where $R_n \in \arg\inf_{S\in\mathbb{F}} d_n(S,Q)$.

Consequently, for any $P \in \mathbb{E}$ and $Q_n \in \mathbb{H}_n$ (Landes et al., 2021b, corollary 32):

$$H_n(Q_n) - H_n(P) \geq d_n(P,Q_n).$$

Pinsker's inequality connects the L_1 distance to n-divergence (see, e.g., Cover and Thomas, 1991, Lemma 11.6.1):

$$d_n(P,Q) \geq \frac{1}{2}\|P-Q\|_n^2.$$

Proposition 2. If P is an entropy limit point of $\mathbb{E}$ then there are $Q_n \in \mathbb{H}_n$ such that $\|Q_n - P\|_n \longrightarrow 0$ as $n \longrightarrow \infty$.

Proof: Putting our last two information-theoretic facts together we have that

$$\begin{aligned} H_n(Q_n) - H_n(P) &\geq d_n(P,Q_n) \\ &\geq \frac{1}{2}\|P-Q_n\|_n^2, \end{aligned}$$

for $Q_n \in \mathbb{H}_n$ and $P \in \mathbb{E}$.

Now, if P is an entropy limit point of $\mathbb{E}$ then there are $Q_n \in \mathbb{H}_n$ such that $|H_n(Q_n) - H_n(P)| \longrightarrow 0$ as $n \longrightarrow \infty$. Hence $\|P-Q_n\|_n^2$ also converge to zero, as required. □

Theorem 3. If $\mathbb{E}$ contains an entropy limit point P then

$$\text{maxent}\,\mathbb{E} = \{P\}.$$

Proof: First we shall show that $P \in \text{maxent}\,\mathbb{E}$; later we shall see that there is no other member of $\text{maxent}\,\mathbb{E}$.

First, then, assume for contradiction that $P \notin \text{maxent}\,\mathbb{E}$. Then there is some $Q \in \mathbb{E}$ such that Q has greater entropy than P. That is, for sufficiently large n, $H_n(Q_n) \geq H_n(Q) > H_n(P)$, where the $Q_n \in \mathbb{H}_n$ converge in entropy (and, by Proposition 2, in L_1) to P. N.b., $Q \neq P$. Hence, for sufficiently large n,

$$\begin{aligned} H_n(Q_n) - H_n(P) &> H_n(Q_n) - H_n(Q) \\ &\geq d_n(Q,Q_n) \\ &\geq \frac{1}{2}\|Q-Q_n\|_n^2. \end{aligned}$$

Since the Q_n converge in entropy to P, they converge in L_1 to Q. By the uniqueness of L_1 limit points, $Q = P$: a contradiction. Hence $P \in \text{maxent}\,\mathbb{E}$, as required.

Next we shall see that P is the unique member of $\text{maxent}\,\mathbb{E}$. Suppose for contradiction that there is some $P^\dagger \in \text{maxent}\,\mathbb{E}$ such that $P^\dagger \neq P$. Then P cannot eventually dominate $P^\dagger$ in n-entropy—i.e., there is some infinite set $J \subseteq \mathbb{N}$ such that for $n \in J$,

$$H_n(P^\dagger) \geq H_n(P).$$

Let $R \stackrel{\text{df}}{=} \lambda P^\dagger + (1-\lambda)P$ for some $\lambda \in (0,1)$. Now by the log-sum inequality (Cover and Thomas, 1991, Theorem 2.7.1), for all $n \in J$ large enough that $P^\dagger(\omega_n) \neq P(\omega_n)$ for some $\omega_n \in \Omega_n$,

$$\begin{aligned} H_n(R) &> \lambda H_n(P^\dagger) + (1-\lambda)H_n(P) \\ &\geq \lambda H_n(P) + (1-\lambda)H_n(P) \\ &= H_n(P). \end{aligned}$$

Hence,

$$\begin{aligned} H_n(Q_n) - H_n(P) &> H_n(Q_n) - H_n(R) \\ &\geq d_n(R,Q_n), \end{aligned}$$

for large enough $n \in J$.

Now by Pinsker's inequality and the definition of R,

$$\begin{aligned} d_n(R,Q_n) &\geq \frac{1}{2}\|R-Q_n\|_n^2 \\ &= \frac{1}{2}\left\|P-Q_n+\lambda(P^\dagger-P)\right\|_n^2 \\ &= \frac{1}{2}\left(\sum_{\omega_n\in\Omega_n}\left|P(\omega_n)-Q_n(\omega_n)+\lambda(P^\dagger(\omega_n)-P(\omega_n))\right|\right)^2. \end{aligned}$$

Let $f_n(\varphi) \stackrel{\text{df}}{=} P(\varphi) - Q_n(\varphi) + \lambda(P^\dagger(\varphi) - P(\varphi))$ and $\rho_n \stackrel{\text{df}}{=} \bigvee_{f_n(\omega_n)>0}\omega_n$. Then,

$$\begin{aligned} \sum_{\omega_n\in\Omega_n}|f_n(\omega_n)| &= \sum_{\omega_n:f_n(\omega_n)>0} f_n(\omega_n) - \sum_{\omega_n:f_n(\omega_n)\leq 0} f_n(\omega_n) \\ &= \sum_{\omega_n:f_n(\omega_n)>0} f_n(\omega_n) - \sum_{\omega_n:f_n(\omega_n)\not>0} f_n(\omega_n) \\ &= f_n(\rho_n) - f_n(\neg\rho_n) \\ &= 2f_n(\rho_n) \end{aligned}$$

after substituting $P(\neg\rho_n) = 1 - P(\rho_n)$ etc.

Let us consider the behaviour of

$$f_n(\rho_n) = P(\rho_n) - Q_n(\rho_n) + \lambda(P^\dagger(\rho_n) - P(\rho_n))$$

as $n \longrightarrow \infty$. Now, $P(\rho_n) - Q_n(\rho_n) \longrightarrow 0$ as $n \longrightarrow \infty$, because Q_n converges in L_1 to P. However, $\lambda(P^\dagger(\rho_n) - P(\rho_n)) \not\longrightarrow 0$ as $n \longrightarrow \infty$, as we shall now see. $P^\dagger \neq P$ by assumption, so they must differ on some quantifier-free sentence ψ, a sentence of $\mathscr{L}_m$, say. Suppose without loss of generality that $P^\dagger(\psi) > P(\psi)$ (otherwise take $\neg\psi$ instead) and let $\delta = P^\dagger(\psi) - P(\psi) > 0$. Now for $n \geq m$,

$$f_n(\rho_n) = \sum_{\omega_n:f_n(\omega_n)>0} f_n(\omega_n) \geq \sum_{\omega_n\models\psi} f_n(\omega_n) = f_n(\psi).$$

Since Q_n converges in L_1 to P we can consider $n > m$ large enough that (Cover and Thomas, 1991, Equation 11.137):

$$\|Q_n - P\|_n = 2 \max_{\varphi \in S\mathscr{L}_n} (Q_n(\varphi) - P(\varphi)) < \lambda\delta \ .$$

In particular, since ψ is quantifier-free, $Q_n(\psi) - P(\psi) \leq \max_{\varphi \in S\mathscr{L}_n}(Q_n(\varphi) - P(\varphi)) < \lambda\delta/2$. For any such n,

$$\begin{array}{rcl} f_n(\rho_n) & \geq & f_n(\psi) \\ & = & P(\psi) - Q_n(\psi) + \lambda(P^\dagger(\psi) - P(\psi)) \\ & > & -\dfrac{\lambda\delta}{2} + \lambda\delta \\ & = & \dfrac{\lambda\delta}{2} \ . \end{array}$$

Putting the above parts together, we have that for sufficiently large $n \in J$,

$$H_n(Q_n) - H_n(P) > d_n(R, Q_n) \geq \frac{(2 f_n(\rho_n))^2}{2} > \frac{\lambda^2\delta^2}{2} > 0 \ .$$

However, that these $H_n(Q_n) - H_n(P)$ are bounded away from zero contradicts the assumption that the Q_n converge in entropy to P. Hence, P is the unique member of maxent$\mathbb{E}$, as required. □

Appendix 2. Alternative proof of Corollary 6

This appendix provides a more direct proof of Corollary 6, which identifies an important scenario in which the equivocator function conditioned on a categorical constraint is the maximal entropy function.

Corollary 6. If $\mathbb{H}_n$ contains $P_=(\cdot|\varphi)$ for sufficiently large n then

$$\text{maxent}\,\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}.$$

Proof: There are two cases: either $P_=(\varphi) = 1$ or $P_=(\varphi) < 1$.

If $P_=(\varphi) = 1$ then $P_= \in \mathbb{E}_\varphi$ and $P_=(\cdot|\varphi) = P_=(\cdot)$. $P_=$ is the unique member of maxent$\mathbb{E}_\varphi$ because the equivocator function has greater entropy than any other probability function, so maxent$\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}$, as required.

If $P_=(\varphi) < 1$ then we can proceed as follows.

Since $P_=(\varphi) > 0$, $P_=(\cdot|\varphi)$ is well defined. $P_=(\varphi|\varphi) = 1$ so $P_=(\cdot|\varphi) \in \mathbb{E}$. Thus $\mathbb{E}_\varphi \neq \emptyset$.

Suppose for contradiction that maxent$\mathbb{E}_\varphi \neq \{P_=(\cdot|\varphi)\}$. Then in $\mathbb{E}_\varphi$ there must be some $P^\dagger \neq P_=(\cdot|\varphi)$ that is not eventually dominated in entropy by $P_=(\cdot|\varphi)$. That is, there is some infinite $J \subseteq \mathbb{N}$ such that $H_n(P^\dagger) \geq H_n(P_=(\cdot|\varphi))$ for all $n \in J$. (To see this consider that there are three cases: (i) if maxent$\mathbb{E}_\varphi = \emptyset$ then every member of $\mathbb{E}_\varphi$ is eventually dominated by some other in entropy, so $P_=(\cdot|\varphi)$ is dominated by some $P^\dagger$ and $P^\dagger$ is not dominated by $P_=(\cdot|\varphi)$; (ii) if $P_=(\cdot|\varphi) \notin$ maxent$\mathbb{E}_\varphi = \{P^\dagger, \ldots\}$ then $P^\dagger$ is not dominated by $P_=(\cdot|\varphi)$; (iii) if maxent$\mathbb{E}_\varphi = \{P_=(\cdot|\varphi), P^\dagger, \ldots\}$ then $P^\dagger$ is not dominated by $P_=(\cdot|\varphi)$.)

Define a probability function $Q \stackrel{\text{df}}{=} \lambda P^\dagger + (1-\lambda) P_=(\cdot|\varphi)$ for some $\lambda \in (0,1)$. By the log-sum inequality (Cover and Thomas, 1991, Theorem 2.7.1), for all $n \in J$ large enough that $P^\dagger(\omega) \neq P_=(\omega|\varphi)$ for some $\omega \in \Omega_n$,

$$\begin{array}{rcl} H_n(Q) & > & \lambda H_n(P^\dagger) + (1-\lambda) H_n(P_=(\cdot|\varphi)) \\ & \geq & \lambda H_n(P_=(\cdot|\varphi)) + (1-\lambda) H_n(P_=(\cdot|\varphi)) \\ & = & H_n(P_=(\cdot|\varphi)). \end{array}$$

However, that $H_n(Q) > H_n(P_=(\cdot|\varphi))$ for sufficiently large $n \in J$ contradicts the assumption that $\mathbb{H}_n$ contains $P_=(\cdot|\varphi)$ for sufficiently large n. Hence maxent$\mathbb{E}_\varphi = \{P_=(\cdot|\varphi)\}$, as required. □

Bibliography

Cover, T. M. and Thomas, J. A. (1991). *Elements of information theory*. John Wiley and Sons, New York, second (2006) edition.

Haenni, R., Romeijn, J.-W., Wheeler, G., and Williamson, J. (2011). *Probabilistic logics and probabilistic networks*. Synthese Library. Springer, Dordrecht.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *The Physical Review*, 106(4):620–630.

Landes, J., Rad, S. R., and Williamson, J. (2021a). Determining maximal entropy functions. Manuscript.

Landes, J., Rafiee Rad, S., and Williamson, J. (2021b). Towards the entropy-limit conjecture. *Annals of Pure and Applied Logic*, 172(2):102870.

Seidenfeld, T. (1986). Entropy and uncertainty. *Philosophy of Science*, 53(4):467–491.

Williamson, J. (2010). *In defence of objective Bayesianism*. Oxford University Press, Oxford.

Williamson, J. (2017). *Lectures on inductive logic*. Oxford University Press, Oxford.