

Journal of applied logics (Print) ISSN 2631-9810
Journal of applied logics (Online) ISSN 2631-9829

Journal of Applied Logics

The IfCoLog Journal of Logics and their Applications

Volume 8 • Issue 7 • August 2021

Available online at
www.collegepublications.co.uk/journals/ifcolog/
Free open access

JOURNAL OF APPLIED LOGICS - IFCoLOG
JOURNAL OF LOGICS AND THEIR APPLICATIONS

Volume 8, Number 7

August 2021

Disclaimer

Statements of fact and opinion in the articles in Journal of Applied Logics - IfCoLog Journal of Logics and their Applications (JALs-FLAP) are those of the respective authors and contributors and not of the JALs-FLAP. Neither College Publications nor the JALs-FLAP make any representation, express or implied, in respect of the accuracy of the material in this journal and cannot accept any legal responsibility or liability for any errors or omissions that may be made. The reader should make his/her own evaluation as to the appropriateness or otherwise of any experimental technique described.

© Individual authors and College Publications 2021
All rights reserved.

ISBN 978-1-84890-374-6

ISSN (E) 2631-9829

ISSN (P) 2631-9810

College Publications

Scientific Director: Dov Gabbay

Managing Director: Jane Spurr

<http://www.collegepublications.co.uk>

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form, or by any means, electronic, mechanical, photocopying, recording or otherwise without prior permission, in writing, from the publisher.

EDITORIAL BOARD

Editors-in-Chief
Dov M. Gabbay and Jörg Siekmann

Marcello D'Agostino
Natasha Alechina
Sandra Alves
Arnon Avron
Jan Broersen
Martin Caminada
Balder ten Cate
Agata Ciabattoni
Robin Cooper
Luis Farinas del Cerro
Esther David
Didier Dubois
PM Dung
David Fernandez Duque
Jan van Eijck
Marcelo Falappa
Amy Felty
Eduaro Fermé

Melvin Fitting
Michael Gabbay
Murdoch Gabbay
Thomas F. Gordon
Wesley H. Holliday
Sara Kalvala
Shalom Lappin
Beishui Liao
David Makinson
George Metcalfe
Claudia Nalon
Valeria de Paiva
Jeff Paris
David Pearce
Pavlos Peppas
Brigitte Pientka
Elaine Pimentel

Henri Prade
David Pym
Ruy de Queiroz
Ram Ramanujam
Chrtian Retoré
Ulrike Sattler
Jörg Siekmann
Jane Spurr
Kaile Su
Leon van der Torre
Yde Venema
Rineke Verbrugge
Heinrich Wansing
Jef Wijsen
John Woods
Michael Wooldridge
Anna Zamansky

SCOPE AND SUBMISSIONS

This journal considers submission in all areas of pure and applied logic, including:

pure logical systems	dynamic logic
proof theory	quantum logic
constructive logic	algebraic logic
categorical logic	logic and cognition
modal and temporal logic	probabilistic logic
model theory	logic and networks
recursion theory	neuro-logical systems
type theory	complexity
nominal theory	argumentation theory
nonclassical logics	logic and computation
nonmonotonic logic	logic and language
numerical and uncertainty reasoning	logic engineering
logic and AI	knowledge-based systems
foundations of logic programming	automated reasoning
belief change/revision	knowledge representation
systems of knowledge and belief	logic in hardware and VLSI
logics and semantics of programming	natural language
specification and verification	concurrent computation
agent theory	planning
databases	

This journal will also consider papers on the application of logic in other subject areas: philosophy, cognitive science, physics etc. provided they have some formal content.

Submissions should be sent to Jane Spurr (jane@janespurr.net) as a pdf file, preferably compiled in \LaTeX using the IFCoLog class file.

CONTENTS

ARTICLES

- Squeezing Arguments and the Plurality of Informal Notions 1899**
Edson Bezerra and Giorgio Venturi
- A Framework for Modelling Molecular Interaction Maps 1917**
*Jean-Marc Alliot, Marta Cialdea Mayer, Robert Demolombe,
Martín Diéguez and Luis Fariñas del Cerro*
- Variation and Pattern in the 3-term Syllogisms and the Biliteral
Sorites of Types Barbara, Darapti and Darii 1953**
Dan Constantin Radulescu
- What is Negation in a System 2020? 1977**
Dov Gabbay
- The Class of all 3-valued Implicative Expansions of Kleene's
Strong Logic Containing Anderson and Belnap's
First Degree Entailment Logic 2035**
Gemma Robles
- Enayat Theories 2073**
Albert Visser

How Paradox Fares in Inconsistency Robust Logic and Beyond: Computational and Naturalized Approaches	2119
<i>John Woods</i>	
Two Generalizations of Bounded Involutive Lattices and of Ortholattices	2173
<i>Afrodita Iorgulescu and Michael Kinyon</i>	
A Science Like Any Other	2221
<i>Andrew David Irvine</i>	
Groups Definable in o-minimal Structures: Various Properties and a Diagram	2237
<i>Annalisa Conversano</i>	
Addenda: Note on ‘Normalisation for Bilateral Classical Logic with some Philosophical Remarks’	2259
<i>Nils Kürbis</i>	

SQUEEZING ARGUMENTS AND THE PLURALITY OF INFORMAL NOTIONS

EDSON BEZERRA*
University of Campinas
edson.vinber92@gmail.com

GIORGIO VENTURI†
University of Campinas
gio.venturi@gmail.com

Abstract

In this paper we argue that squeezing arguments *à la* Kreisel fail to univocally capture an informal or intuitive notion of validity. This suggests a form of logical pluralism, at a conceptual level, not only among but also within logical systems.

1 Introduction

Logical practice, understood as what logicians do in their everyday work, attests that there are many different logical systems. This comes as a practical confirmation of logical pluralism: the thesis according to which there is more than one legitimate consequence relation. But, of course, what counts as legitimate and in which sense different logics can be put on a par motivates different forms of logical pluralism. Some formulations of logical pluralism are more permissive than others, in the sense of accepting more or less different notions of logical consequence. In their famous book, Beall & Restall [3] proposed a general framework able to accommodate a plurality of logical systems, by describing logical consequence as a relation which preserves truth from premises to conclusion. In [3], they base their version of logical pluralism on the following definition.

We would like to thank the referees, Bruno Mendonça, João Schmidt and Santiago Jockwich for their valuable comments. We also thank the audience of the *Logic Supergroup* seminar, held online in October 16 2020, for the discussion about some ideas contained in this paper.

*The author acknowledges support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES) - Finance Code 001.

†The author acknowledges support from the FAPESP Jovem Pesquisador grant n. 2016/25891-3.

Definition 1.1. φ is a logical consequence of Σ if and only if in every case that the sentences of Σ are true, φ is true.

Definition 1.1 expresses an informal notion of logical consequence, since it leaves open the interpretation of the word ‘case’. Some specifications of ‘case’ give rise to a formal notion of consequence relation. For example, by considering the concept of “formal construction” (or “tarskian model”), one can obtain intuitionistic logic (or classical first-order logic). Therefore, given an informal concept of validity, one can obtain different notions of formal system. Notice that Definition 1.1 rests on an extensional notion of consequence relation, in the sense that the relation of logical consequence between Σ and φ is completely determined by the notion of preservation of truth between relevant cases. What Definition 1.1 captures is *when* the relation of logical consequence attains and not *what it means* for Σ and φ to be in a relation of logical consequence. Thus, Definition 1.1 is not suitable for capturing the nature of logical consequence and its intensional aspects. This is clearly a limitation of such an account, since this logical relation can be seen to preserve other notions than just truth – as it is the case for intuitionistic logic, which is designed to preserve provability. As a matter of fact, we can find different interpretations of what it means to be logically valid. But then, how is it possible to capture the meaning of a notion of logical consequence? In this paper we argue that there is no obvious answer to this question, showing that logical validity is underdetermined with respect to its intended interpretation, even in the presence of (a variant of) a so called squeezing argument.

Squeezing arguments originate from Kreisel’s original formulation [17], which was meant to show that first-order classical logic is able to capture a notion of informal validity that is stronger than the Tarskian model-theoretic one, since it is also able to account for validity in the universe of all sets. Kreisel’s argument made essential use of the possibility of bridging syntax and semantics by means of a completeness theorem and, for this reason, this argument has been extended to other logical systems displaying the same degree of completeness. Therefore, the different squeezing arguments we find in the literature are able to match informal notions of logical validity (other than the classical one) with their corresponding formal definitions.

One might wonder whether squeezing argument, thus, can help to mitigate the plurality of notions of validity we find in the literature. In this paper we reckon that this is not the case and we argue for an even stronger form of pluralism, which we call *informal pluralism*. Informal pluralism consists in reckoning that once we fix a notion of logical validity, it is hard to tell what is the corresponding notion of informal validity that the formal notion captures. Then, we show that squeezing

arguments cannot squeeze in the uniqueness of the corresponding informal notion, since a complete logical system can be made compatible with different notions of informal validity. In this sense, the perspective we adopt here goes in the opposite direction of Beal & Restall’s proposal: while they initially consider an informal notion to obtain different formal systems, we first consider formal systems, and then we look for informal notions that correspond to each of these formal systems.

The paper is organized as follows. In §2, we recall the main structure of Kreisel’s squeezing argument, and we advance a first criticism in the context of classical logic. In §3, we review a variant of Kreisel’s argument, by [23], and the criticism that it received by [10]. Then in §4, we extend the criticism of §3 to Intuitionistic Logic. We end with a few remarks in §5.

2 Kreisel’s squeezing argument

Let us briefly recall the main points of Kreisel’s squeezing argument. Its main goal is to capture an informal notion of logical validity by squeezing it between two formal notions. To this aim, let us fix a first-order formula φ . By $Val(\varphi)$ we denote *informal validity*, which is defined as follows.

$Val(\varphi)$: φ is true in every structure.

We say that $Val(\varphi)$ is informal because it is, intentionally, theoretically vague. The notion of structure present in Val includes the standard set-theoretic notion as well as class-structures. Besides the informal definition of validity, Kreisel presents two formal counterparts of this notion.

$V(\varphi)$: φ is true in all set-theoretic structures.

$D(\varphi)$: φ is deducible by a given set of formal rules (Hilbert Calculus, Natural Deduction, Sequent Calculi) for First-Order Logic (FOL).

V and D are *formal* because they are theoretically precise, in the sense that each one is presented in a well-structured conceptual framework. Then, *informal* notions are defined as non-formal.

Although $Val(\varphi)$ cannot be reduced, in principle, to the other two notions, nonetheless it is directly connected to them as follows.

$$(1) \quad D(\varphi) \Rightarrow Val(\varphi)$$

$$(2) \quad Val(\varphi) \Rightarrow V(\varphi)$$

To justify (1), one could provide an argument, similar to an induction on the length of proofs, showing that the axioms of FOL are informally valid and that the rules of FOL preserve informal validity. This means that we use classical principles and inference rules to justify the axioms and rules of FOL. This move is not circular, once Kreisel argues that D is itself a codification of the deductive practice of mathematical reasoning. For this reason, we can assert that (1) holds for Val .¹ To justify (2), the argument runs as follows: the definition of Val encompasses structures whose domains are sets as well as structures whose domains are not sets. So, if φ is valid in the informal sense, then it is so when only set-sized structures are taken into consideration.

Argument 2.1. *Kreisel's squeezing argument for Val .*

- (1) $D(\varphi) \Rightarrow Val(\varphi)$
- (2) $Val(\varphi) \Rightarrow D(\varphi)$
- (3) $V(\varphi) \Rightarrow D(\varphi)$ Completeness
- (4) $D(\varphi) \Leftrightarrow Val(\varphi) \Leftrightarrow V(\varphi)$ from (1)-(3)

This squeezing argument offers some philosophical content to the completeness theorem, since, as [1] argues, the informal notion Val can be seen as bridging the gap between the two formal notions of validity.

There are a few important remarks about Kreisel's argument. First, Argument 2.1 does not constitute a formal proof, because Val is itself informal, whereas V and D are formal.² Second, as Smith [24] notes, informal validity Val is not, properly speaking, an intuitive notion of validity, but the result of a necessary process of idealization, without which the Argument 2.1 could not work. Instead, it is a rigorously defined notion, still informal, which is close to the model-theoretical construct. This constitutes what Kreisel calls *informal rigour*, the activity of providing a precise analysis of intuitive notions in order "to eliminate the doubtful properties of the intuitive notions when drawing conclusions about them" [17, p.138]. Last, but not least, is the role of the completeness theorem for FOL. Its role is essential,

¹In Kreisel's original paper [17], D_F stands for Frege's axioms for first-order logic. As Kennedy & Väänänen [16] argue, Kreisel considers D_F as an adequate formalisation of informal validity. So, D_F is correct with respect to Val .

²Halbach [11] provides a substitutional analysis of logical validity, defending that such an approach is closer to an informal understanding logical validity. Interestingly, such analysis provides an informal notion for which it is possible, as he argues, to present a formal proof that connects it to the formal definitions of logical validity.

to the extent that in its absence the argument would not work, as it does not for (full) Second-Order Logic (SOL).³

That *Val* may fail to capture our intuitive/pre-theoretical notion of logical validity (which we barely grasp) does not constitute a crucial problem for Kreisel's original argument. As argued by Kennedy & Väänänen [16], the informal notion adopted in natural mathematical language is semantical and it is close to a model-theoretical approach. Then, from this perspective, Kreisel's informal notion seems to capture a notion of validity from mathematical practice. And indeed this seemed to be Kreisel's goal in [17].

A possible response to [16] and to its retraction to a formal interpretation of *Val*, consists in noting that, even at the formal level, the notion of validity is not completely transparent. The reason being, that a formal notion of validity depends on the choice of the structures that are considered relevant for its definition. But this choice is not always univocal. Indeed, we can easily come up with a distinction between standard and non-standard structures, for example, with respect to the way they interpret equality. For the sake of concreteness, we now analyse a specific instance of this phenomenon, in the case of classical set theory. This example will not only help us to undermine the transparency of a formal notion of validity, but will also show that, even in the classical case, we can find a plurality of notions of validity. The latter, therefore, will be an important hint of a larger phenomenon, that we will later analyse in the broader context of non-classical logic.

To give an example of non-standard models, which, however, are standardly used in mathematical practice, one can think of Boolean-valued models for set theory. These are class structures of the form $\mathbf{V}^{\mathbb{B}}$, where \mathbb{B} is a Boolean algebra and where any element $u \in \mathbf{V}^{\mathbb{B}}$ is a function $u : \mathbf{V}^{\mathbb{B}} \rightarrow \mathbb{B}$. Although every $\mathbf{V}^{\mathbb{B}}$ is a model of ZFC, whenever \mathbb{B} is a complete⁴ Boolean algebra, nonetheless equality receives an *ad hoc* interpretation, recursively intertwined with that of \in . By $\llbracket \cdot \rrbracket$ we indicate the interpretation function $\llbracket \cdot \rrbracket : \mathcal{L}_{\mathbb{B}} \rightarrow \mathbb{B}$, where $\mathcal{L}_{\mathbb{B}}$ consists of the language of set theory extended with constants⁵ for every element of $\mathbf{V}^{\mathbb{B}}$.

³In [16], the authors show that it is possible to provide a squeezing argument for SOL, if we look for its fragment characterized by Henkin's models.

⁴[30]: A Boolean algebra is called *complete* whenever each of its subsets possesses a supremum and an infimum.

⁵In order to simplify notation we will use the same symbol for an element u of $\mathbf{V}^{\mathbb{B}}$ and the constant which represents it in the language $\mathcal{L}_{\mathbb{B}}$.

Definition 2.2. For any two elements $u, v \in \mathbf{V}^{\mathbb{B}}$,

$$\begin{aligned} \llbracket u \in v \rrbracket &= \bigvee_{x \in \text{dom}(v)} (v(x) \wedge \llbracket x = u \rrbracket), \\ \llbracket u = v \rrbracket &= \bigwedge_{x \in \text{dom}(u)} (u(x) \Rightarrow \llbracket x \in v \rrbracket) \wedge \bigwedge_{y \in \text{dom}(v)} (v(y) \Rightarrow \llbracket y \in u \rrbracket). \end{aligned}$$

Even if Boolean-valued models are non-standard, we can stretch the definition of informal validity so to include such structures. But what about when the algebra \mathbb{B} is not complete, or if we replace it by a Heyting algebra \mathbb{H} ? And what about when we give the same construction with an algebra \mathbb{P} which models a paraconsistent logic?⁶

We do not want to take a stance here on what counts as a proper class-structure, but we notice that the shift from standard to non-standard is not discrete but continuous. And wherever the dividing line is, there will always be two disjoint classes of structures which can give rise to two different notions of informal validity.

One might object that we are here considering a too wide range of structures, to the extent that even classicality results undermined. But this seems to be exactly the point of logical pluralism: considering different classes of structures we end up changing the notion of logical consequence that we consider. Indeed, if we allow such a comprehensive semantics as the algebraic one, we find different classes of structures that determine opposite and incompatible logical notions.

We can distinguish here between two complementary forms of pluralism. We call the first *formal pluralism*. It consists in the absence of a purely logical reason for deciding which is the notion of logical validity that captures the informal notion of validity that we consider appropriate (where thus the appropriateness depends on pre-theoretical reasons).⁷ The second may be called *informal pluralism* and consists in reckoning that once we fix a notion of logical validity, it is hard to tell what is the corresponding notion of informal validity that the formal notion captures. These two forms of pluralism point at the same phenomenon; what changes is the perspective one takes: to look at the informal notions from a formal perspective or vice versa.

In the concrete case of the example discussed (i.e. non-standard models of set theory), we see formal pluralism arising when making a choice of the algebra-valued models which correctly capture a pre-theoretic notion of set-theoretic validity. Depending on the choice, we end up with classical FOL, or intuitionistic FOL, or a

⁶This is not just a mental experiment since such models are well-studied in the literature: [4] and [19]

⁷In [20] we can find a similar problem: in the presence of an informal proof, it is hard to tell what it is the best formalization of it.

paraconsistent FOL, or a combination of them. For what concerns informal pluralism, we can witness it in deciding which class-structures we should count for the definition of informal validity, once we have fixed a classical notion of validity: just the universe of all sets \mathbf{V} , or also the Boolean-valued models?

While formal pluralism has been already recognised and discussed in [29],⁸ in connection with a sufficiently expressive semantics, this is not the case for the informal one. The connection between the latter and Kreisel's squeezing argument is clear. If any squeezing argument is meant to capture an informal notion validity, informal pluralism says that there are many informal notions and the corresponding formal system is not able to uniquely point to one of them. Further, we will argue that this form of pluralism does not necessarily depend on a sufficiently expressive formal semantics, like the algebraic one. Indeed, we will argue that even at the level of an intuitive notion of validity, a squeezing argument is not able to deliver uniqueness.

From the observations above, we can only conclude that Argument 2.1 establishes the extensional equivalence between notions V , D and Val , a highly theorized informal notion, when first-order formulas are considered. It does not mean that Val is the only informal notion of validity. Even if it was not Kreisel's objective in determining whether Val is unique or not, we think it to be relevant to ask whether it is the case. If such notion is not the only notion captured by V and D , there may be other informal notions, still theorized, which may provide a more intuitive understanding of the formal notions of validity. And, interestingly, it may suggest that the formal notions of validity are underdetermined by its informal counterparts. In what follows, we develop this idea by investigating variations of Kreisel's argument presented in the literature.

3 Variants of Kreisel's argument

Kreisel's squeezing argument is not meant to establish that $Val(\varphi)$ is *the* intuitive notion of validity. The point of the argument is only to show that an informal notion corresponds to a formal one. However, because of its simple form, the squeezing argument has been proposed to capture the intuitive notions of validity from natural language. In order to see how this reduction works, let F and $Prem$ be the counterparts of the formula φ and the set of formulas Γ in natural language. Shapiro [23] defines a notion of consequence as follows.

⁸In [29], it is argued that this is an instance of relativism. We do not take a stance on this matter.

Definition 3.1. *The relation $Val_B(Prem, F)$ holds whenever F is logical consequence of $Prem$ in the blended sense; that is, it is not possible to every member of $Prem$ to be true and F be false, and this impossibility holds in virtue of the meaning of the logical terms.*

By $Val_B(F)$ we mean that F is informally valid in the blended sense (i.e. considering an empty set of premises). Shapiro argues that this blended notion captures the formality and necessity of the model-theoretical consequence relation. Now, it is clear that every φ provable in FOL is valid in the sense of Val_B . So, we say that the deductive system of FOL is *faithful* with respect to Val_B . It is also clear that every valid F , which is a natural language correspondent of φ , has a valid formalization φ in FOL. In this sense, we say that V is *adequate* to Val_B . Indeed, assuming that F is the natural language counterpart of a FOL formula φ , it is possible to apply a squeezing argument to Val_B as follows.⁹

Argument 3.2. *Shapiro’s squeezing argument for Val_B .*

- | | |
|--|--------------|
| (1) $D(\varphi) \Rightarrow Val_B(F)$ | Faithfulness |
| (2) $Val_B(F) \Rightarrow V(\varphi)$ | Adequacy |
| (3) $V(\varphi) \Rightarrow D(\varphi)$ | Completeness |
| (4) $D(\varphi) \Leftrightarrow Val_B(F) \Leftrightarrow V(\varphi)$ | from (1)-(3) |

The argument is meant to show that blended validity in natural language extensionally coincides both with the Tarskian model-theoretic notion of validity and with its proof-theoretical counterpart.

Interestingly, a similar argument can be applied to a more syntactic notion of informal validity.

Definition 3.3. *The relation $Val_{Ded}(Prem, F)$ holds whenever F is logical consequence of $Prem$ in the deductive sense; that is, there is a deduction of F from $Prem$ by a chain of legitimate gap-free (self-evident) rules of inference.*

By $Val_{Ded}(F)$ we mean that F is informally valid in the deductive sense (i.e. considering an empty set of premises). The arguments for faithfulness and adequacy of V and D with respect to Val_{Ded} are similar to the arguments for Val_B . Then we also have a corresponding squeezing argument for Val_{Ded} .

⁹In [23], one finds a version of the Argument 3.2. But we follow here Griffiths’s version [10] for the sake of simplicity.

Argument 3.4. *Shapiro's squeezing argument for Val_{Ded} .*

- | | | |
|-----|---|--------------|
| (1) | $D(\varphi) \Rightarrow Val_{Ded}(F)$ | Faithfulness |
| (2) | $Val_{Ded}(F) \Rightarrow V(\varphi)$ | Adequacy |
| (3) | $V(\varphi) \Rightarrow D(\varphi)$ | Completeness |
| (4) | $D(\varphi) \Leftrightarrow Val_{Ded}(F) \Leftrightarrow V(\varphi)$ from (1)-(3) | |

Before discussing the relevance of these arguments for the present discussion, let us briefly review the main criticism that these arguments received.

Following Griffiths's argument [10] against Shapiro's squeezing argument, we can argue that Val_B and Val_{Ded} , although defined for natural language, do not account for the totality of such language, but just for a non-ambiguous fragment of it: the one that can be formalized in FOL. Indeed, F is obtained as the *reading* of a first-order formula where reading is understood as the reverse process of formalization. Then, Griffiths objects to Shapiro's argument by arguing that (1) and (2) of the Argument 3.2 hold only in virtue of the connection between F and φ and not because of the definition of Val_B . Therefore there is nothing special about the blended notion of validity because Val_B is coextensive with the Val relation used in Kreisel's argument. And of course the same works for Val_{Ded} .

We can now advance a further objection to Shapiro's argument(s). We agree with Griffiths that the three notions Val , Val_B , and Val_{Ded} , are coextensive. But then, if the formal notions of classical FOL capture the three of them, which one can be seen as the informal or the intuitive content of the formal notions? This is a relevant question due to the fact that the notions Val , Val_B , and Val_{Ded} are meant to be intensional objects: properties of formulas.¹⁰ If these squeezing arguments are able to show that we can capture these notions by means of extensional concepts (V and D), however, we are left in the dark with respect to which one of these represents the intensional concept we associate to logical (classical FOL) validity. In this sense, logical validity is therefore underdetermined by its formal counterparts, even if these manifest a perfect correspondence between syntax and (formal) semantics.

To counter our point, one could simply accept that FOL captures informal notions of validity which have semantic or syntactic aspects. Then, the completeness theorem shows that these notions are extensionally equivalent, despite their intensional difference. This response, however, misses our main point: Argument 3.2 and

¹⁰One could say that the objections also work against Kreisel. It would do so if Kreisel's interests were natural language. But, as we highlighted before, his interests were only mathematical.

Argument 3.4 only hold in virtue of the correspondence between F and φ , not in virtue of the intrinsic characteristics of Val_B and Val_{Ded} . To make our point clear, consider the following notion presented in [10]:

Definition 3.5. *The relation $Val_{Nec}(Prem, F)$ holds whenever F is logical consequence of $Prem$ in the modal sense if and only if necessarily, every member of $Prem$ is true, F is true.*

Griffiths presents a squeezing argument for Val_{Nec} in order to show that there is nothing distinctive about Val_B , since Val_{Nec} holds by the same reason as the validity in the blended sense. Given this abundance of options, and since formal logic is mute on this topic, any judgement about which notion is more appropriated is moved by pre-theoretical reasons, which therefore suggests a form of informal pluralism with respect to our pre-formal notion of validity.

A consequence of this phenomenon is that the formal consequence relation of FOL is not able to capture the intuitive notion of logical consequence. Kreisel's argument and its variants neither capture such intuitive notion nor do they capture a unique one, even if they capture relevant informal notions, which regulate our inferential practice. Moreover, and following Griffiths's analysis of Shapiro's argument, we cannot hold that these informal notions capture the whole of our inferential practices, but only a small, formalisable fragment of natural language inferences. Probably, to capture all inferences of natural language in a system like FOL, we should extend this system to the point of doubting that it remains formal.¹¹

So far, we maintain that all that is safe to infer from Kreisel's and Shapiro's arguments is that formal validity is able to capture a fragment of the intuitive validity of natural language which deals with preservation of truth from premises to the conclusion. From this perspective, we can say that the axioms and rules of FOL capture general principles for correct truth-preserving reasoning. Therefore, we can follow a pragmatic vindication of the logical principles of FOL (as [7] does) in arguing that FOL is the correct logic to adopt in the case we want to yield true conclusions from true premises.

However, given the plurality of the logical systems used in our formal practice, one can ask whether the pluralism found with respect to the informal notions captured by classical FOL can be extended to other formal notions of logical validity. In other terms, we wonder whether the underdetermination of the informal classical notions is a weakness of classical FOL, or else an intrinsic limitation of any squeezing argument. It is, therefore, to a pluralist perspective on logic that we now turn.

¹¹It has been argued by [8] that logical consequence is not determined by natural language.

4 Informal notions and Logical pluralism

We want here to explore squeezing arguments for propositional Intuitionistic Logic (IL). As in the classical case, we will show that the squeezing arguments for IL do not determine a unique informal validity. Then we will argue that the multiplicity of informal notions corresponding to a particular formal system suggests that formal systems, in general, do not have a canonical interpretation.

The squeezing arguments for FOL presented in §3 show that V and D do not manage to squeeze in a unique notion of classical informal validity. We now investigate whether a similar phenomenon occurs in other logical contexts.

4.1 Intuitionistic Logic and BHK

Intuitionists generally agree that the classical notion of truth is not as tractable as the notion of proof, because it validates some principles which are not constructively valid, such as the *Principle of Excluded Middle* (PEM). Consider, for example, a mathematical conjecture such as Goldbach conjecture. Since we do not have an available proof of either it or its negation, therefore PEM cannot be considered as a valid logical principle. From this perspective, a notion that seems to harmonize better with intuitionism is offered by the *Brouwer-Heyting-Kolmogorov interpretation* (for short, BHK interpretation), which captures a well-justified notion of *constructibility*. Moreover, in [27], one can find an informal justification of all the axioms and rules of IL with respect to BHK. Thus we can define a notion of informal validity with respect to BHK as follows; where φ is a propositional formula.

$Val_I(\varphi)$: φ is constructively provable.

The informality of Val_I stems from the absence of a specification of the methods of construction. And indeed different interpretations of constructibility may lead us to different conceptions of constructivism. For example, under Markov's [28] interpretation of constructivism, every algorithm must terminate, whereas Brouwer's intuitionism allows the construction of infinite sequences of objects.

(...) the ideal mathematician may construct longer and longer initial segments $\alpha(0), \dots, \alpha(n)$ of an infinite sequence of natural numbers a where a is not a priori determined by some fixed process of producing the values, so the construction of a is never finished: a is an example of a choice sequence. ([28, p.5])

Let D_I and V_I now stand for deductibility in an intuitionistic proof system and

V_I for structures whose internal logic is IL. We now argue for the coextensivity of D_I , Val_I and V_I .

Although Val_I is theoretically irreducible to both D_I and V_I , nonetheless Val_I is sound with respect to IL (as argued in [27]), and also the constructions allowed by Val_I can clearly be carried out in structures whose internal logic is IL. Indeed, the methods of constructions codified by IL represent a qualification of the (in principle) more general notion of constructibility. Therefore, we can argue that D_I captures a more restricted version of constructibility than Val_I . We can, therefore, run an analogous version of Kreisel's squeezing argument.

Argument 4.1. *First version of squeezing argument for IL.*

- | | | |
|-----|--|--------------|
| (1) | $D_I(\varphi) \Rightarrow Val_I(\varphi)$ | Soundness |
| (2) | $Val_I(\varphi) \Rightarrow V_I(\varphi)$ | Adequacy |
| (3) | $V_I(\varphi) \Rightarrow D_I(\varphi)$ | Completeness |
| (4) | $D_I(\varphi) \Leftrightarrow Val_I(\varphi) \Leftrightarrow V_I(\varphi)$ | from (1)-(3) |

As for the classical notion of validity, we therefore seem to have squeezed Val_I within the syntax and semantics of Intuitionistic Logic.

4.2 Intuitionistic Logic and $S4$

The philosophical debate about logics and their interpretations has its origins in the proper development of mathematical logic, especially with respect to non-classical logics. In [2], we can find a distinction between *pure logics* and their *philosophical interpretations*. Pure logics, understood as languages equipped with consequence relations, have no intrinsic connections with their possible philosophical interpretations. Of course, there are interpretations more interesting than others and any judgement about the adequacy of a notion to the detriment of others is moved by pre-theoretical reasons and not only by the formal aspect of a pure logic.

For example, although the BHK interpretation is considered the intended informal notion captured by IL, nothing prevents us from proposing alternative conceptions associated to an intuitionistic notion of logical validity.

Let us consider a concrete case. It is a well-known fact that IL is modally characterized as the logic $S4$ ([15]), whose modality \Box has a well-justified epistemic interpretation as true introspective (and deductively closed) knowledge; let us call this specific notion *knowledge**. The logic $S4$ is normally presented as follows.

Taut: all axioms and inference rules of classical propositional logic;

K: $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$;

T: $\Box\varphi \rightarrow \varphi$;

4: $\Box\varphi \rightarrow \Box\Box\varphi$;

Nec: From $\vdash \varphi$ we obtain $\vdash \Box\varphi$.

Under such epistemological interpretation, the axiom *K* says that knowledge is closed under logical consequence. That is, if an agent knows* φ , then they know* all the consequences of φ .¹² The axiom *T* says that knowledge* is factual. The axiom 4 is an introspection principle. And *Nec* says that the agent knows* all logical validities.

Theorem 4.2. *There is a translation $T : \mathcal{L} \rightarrow \mathcal{L}^\Box$ from the language of propositional logic to the language of modal propositional logic, such that $IL \vdash \varphi$ iff $S4 \vdash T(\varphi)$.*

Given the specificity of the notion of knowledge we defined, knowledge*, by definition, is such that (A) every theorem of *S4* is a thesis about knowability* and (B) every thesis about knowability* is *S4*-valid. We can, therefore, state the second intuitive notion we can associate to *IL*.

$Val_I^*(\varphi)$: $T(\varphi)$ is known*

Despite knowledge and provability being different concepts, nonetheless we can provide an analogous version of Argument 4.1, using the informal notion Val_I^* . Therefore, we can squeeze Val_I^* in between the syntax and the semantics of *IL*.

By D_{S4} we mean the predicate of derivability in *S4* and by V_{S4} the predicate of validity for *S4*, for example in reflexive and transitive Kripke frames.

Argument 4.3. *Second version of squeezing argument for IL.*

¹²The principle *K* is usually taken as implausible due to the problem of *logical omniscience* ([25],[22]). Here, the logical omniscience is not a problem because we are supposing an idealized notion of knowledge.

- (1) $D_I(\varphi) \Rightarrow D_{S4}(T(\varphi))$ Theorem 4.2
- (2) $D_{S4}(T(\varphi)) \Rightarrow Val_I^*(\varphi)$ (A)
- (3) $Val_I^*(\varphi) \Rightarrow V_{S4}(T(\varphi))$ (B)
- (4) $V_{S4}(T(\varphi)) \Rightarrow V_I(\varphi)$ Theorem 4.2
- (5) $V_I(\varphi) \Rightarrow D_I(\varphi)$ Completeness of IL
- (6) $D_I(\varphi) \Leftrightarrow Val_I^*(\varphi) \Leftrightarrow V_I(\varphi)$ from (1)-(5)

One might object that the choice of this informal notion: knowledge*, is too *ad hoc* or even that *S4* does not capture the notion of knowledge.¹³

To counter this objection there are two possible moves. On the one hand, we can argue that the issues one can raise on the epistemic completeness of *S4* are the result of an intrinsic and non-eliminable gap between formality and informality; a gap that is also present in Kreisel’s original argument. That is to say that any formalization of an informal notion has non-trivial elements that, similar to a Carnapian explication, have the effect of normatively modifying the (necessarily vague) informal notion.

On the other hand, we can provide another example of an informal notion that we can associate to *S4*; one for which the identification of the informal notion and its formalization is less controversial. This second move does not eliminate the qualms raised by the objection, but it only advances a dialectical strategy: for any well-justified informal notions we can associate to IL, it is the task of the proponent of the objection to show that this association fails.

We can then define a third informal notion for IL.

$Val'_I(\varphi)$: $T(\varphi)$ is informally provable.

By informally provable we mean provability by any correct mathematical/logical means, not being tied to a particular formal system. Indeed this was Gödel’s original interpretation [9] of the *S4* operator \Box . According to this interpretation of *S4*, the axiom *K* says that provability is preserved under modus ponens. The axiom *T* says that whatever is provable is true. The axiom 4 says that if φ is provable, then it is provable that φ is provable. The latter axiom is a kind of introspection principle. The necessitation rule says that all logical validities are informally provable. Thus, the characteristic modal axioms of *S4* suggest that this logic captures the concept of informal provability.¹⁴

¹³See for example the discussion in [26]. In this paper, *S4* is presented as the logic which captures knowledge without doxastic elements.

¹⁴Gödel’s interpretation of *S4* was well-received in the literature. In [21], [5] and [12] one can find

Since T is an $S4$ axiom, \Box cannot stand for provability in a consistent formal system which contains arithmetic, on pain of contradicting the second incompleteness theorem. Then, the argument runs as follows.

Similarly to Val_I^* , (A') stands for the assumption that every theorem of $S4$ is a thesis about informal provability and (B') that every thesis about informal provability is $S4$ -valid.

Argument 4.4. *Third version of squeezing argument for IL.*

- | | |
|---|--------------------|
| (1) $D_I(\varphi) \Rightarrow D_{S4}(T(\varphi))$ | Theorem 4.2 |
| (2) $D_{S4}(T(\varphi)) \Rightarrow Val'_I(\varphi)$ | (A') |
| (3) $Val'_I(\varphi) \Rightarrow V_{S4}(T(\varphi))$ | (B') |
| (4) $V_{S4}(T(\varphi)) \Rightarrow V_I(\varphi)$ | Theorem 4.2 |
| (5) $V_I(\varphi) \Rightarrow D_I(\varphi)$ | Completeness of IL |
| (6) $D_I(\varphi) \Leftrightarrow Val'_I(\varphi) \Leftrightarrow V_I(\varphi)$ | from (1)-(5) |

We are therefore in the position to argue that IL can equally capture, by means of squeezing arguments, three informal notions: being constructively provable, being knowable*, and being informally provable. Therefore we notice that also in the case of IL a squeezing argument is not able to univocally capture an informal notion that can thus be presented as the informal content of a formal notion of logical validity.

5 Conclusion

The underdetermination of informal notions with respect to logical systems has the effect of suggesting that logical systems have no canonical interpretations. As formal systems, logics allow different interpretations. Therefore, IL cannot be seen only as a system that captures constructive reasoning. Indeed, because of Arguments 4.3 and 4.4 we can interpret IL in epistemological terms or as capturing a notion of informal provability. Modal logic displays a paradigmatic example of this phenomenon. Not

arguments in defense of the informal provability interpretation of $S4$. On the other hand, Leitgeb [18] argues that the completeness of $S4$ with respect to informal provability is still an open problem because there is the question whether statements about unprovability should be taken as axioms of informal provability. Given that this suspicion is an open problem, we take Gödel's original interpretation because it is not controversial and it seems to capture the minimal principles of such notion.

only $S4$, but also a logic like $S4.2$ can be seen as an epistemic logic ([26]) or one that codifies set-theoretical notions ([13]). This is, of course, coherent with the views that a formal notion of logical consequence stands with respect to an informal notion of validity in a relation akin to a Carnapian explication (as suggested in [10]) and that this kind of relation is imprecise in nature (as argued in [6]).

To realise that a logical system have no canonical interpretations, thus, amount to recognise that a squeezing arguments does not really help in capturing an informal or intuitive notion of validity. This should not come as a surprise, given the modern-Hilbertian-axiomatic perspective widely accepted on formal systems. What on the other hand is interesting to notice is that Griffiths's criticism to squeezing arguments extends from the informal to the formal context. Not only Kreisel's squeezing argument is not able to capture an intuitive notion of validity, but it also fails to univocally capture an informal idealized notion of validity definable in a mathematical context. Moreover, this phenomenon extends to other logical systems, thus showing a form of pluralism not only between different logics, but also within a fixed one.

The informal pluralism discussed here also suggests that formal systems cannot really conflict over a given interpretation. Indeed, if we undermine the link between a formal system and its interpretation, we realise that any competition between different logical systems is only apparent, since none can really claim to fully capture a given notion of informal validity. Different logics preserve validity in different cases, which in turn refer to a multiplicity of informal notions. Consider again the case of classical and intuitionistic logics. While Val talks about truth in all structures, Val_I talks about constructive provability. Thus, the rivalry between classical and intuitionistic logics is apparent, once one recognizes that their informal interpretations talk about different notions and that these cannot be completely reduced to their formal counterparts. Heyting seemed to have already held a similar position ([14]). For him, a classical mathematician can maintain that mathematical entities exist autonomously, while, at the same time, recognizing that the notion of existence do not play any role, when dealing with proofs. Thus, the plurality of informal notions may offer a more tolerant perspective on logical pluralism, thus vindicating the plurality of formal systems in mathematical practice.

References

- [1] Edgar Andrade-Lotero and Catarina Dutilh Novaes. Validity, the squeezing argument and alternative semantic systems: the case of aristotelian syllogistic. *Journal of philosophical logic*, 41(2):387–418, 2012.
- [2] Eduardo Alejandro Barrio. Models & proofs: Lfis without a canonical inter-

- pretations. *Principia: an international journal of epistemology*, 22(1):87–112, 2018.
- [3] JC Beall and Greg Restall. *Logical pluralism*. Oxford University Press, 2006.
- [4] J. L. Bell. *Intuitionistic set theory*. College Publications, 2014.
- [5] John P Burgess. Which modal logic is the right one? *Notre Dame Journal of Formal Logic*, 40(1):81–93, 1999.
- [6] Catarina Dutilh Novaes and Erich Reck. Carnapian explication, formalisms as cognitive tools, and the paradox of adequate formalization. *Synthese*, 194(1):195–215, 2017.
- [7] Herbert Feigl. De Principiis Non Disputandum...? In *Inquiries and provocations*, pages 237–268. Springer, 1981.
- [8] Michael Glanzberg. Logical consequence and natural language. *Foundations of logical consequence*, pages 71–120, 2015.
- [9] Kurt Gödel. Eine Interpretation des intuitionistischen Aussagenkalküls. *Ergebnisse eines mathematisches Kolloquiums*, 4:39–40, 1933.
- [10] Owen Griffiths. Formal and informal consequence. *Thought: A Journal of Philosophy*, 3(1):9–20, 2014.
- [11] Volker Halbach. The substitutional analysis of logical consequence. *Noûs*, 54(2):431–450, 2020.
- [12] Sören Halldén. A pragmatic approach to modal theory. *Acta Philosophica Fennica*, 16:53–63, 1963.
- [13] Joel Hamkins and Benedikt Löwe. The modal logic of forcing. *Transactions of the American Mathematical Society*, 360(4):1793–1817, 2008.
- [14] Arend Heyting. On intuitionism. In Paolo Mancosu, editor, *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920s*. Oxford University Press, 1998.
- [15] George Edward Hughes and Max J Cresswell. *A new introduction to modal logic*. Psychology Press, 1996.
- [16] Juliette Kennedy and Jouko Väänänen. Squeezing arguments and strong logics. In *15th International Congress of Logic, Methodology and Philosophy of Science*. College Publications, 2017.

- [17] Georg Kreisel. Informal rigour and completeness proofs. In *Studies in Logic and the Foundations of Mathematics*, volume 47, pages 138–186. Elsevier, 1967.
- [18] Hannes Leitgeb. On formal and informal provability. In *New waves in philosophy of mathematics*, pages 263–299. Springer, 2009.
- [19] Benedikt Löwe and Sourav Tarafder. Generalized algebra-valued models of set theory. *Review of Symbolic Logic*, 8(1):192–205, 2015.
- [20] Marianna Antonutti Marfori. Informal proofs and mathematical rigour. *Studia Logica*, 96(2):261–272, 2010.
- [21] John Myhill. Some remarks on the notion of proof. *The Journal of Philosophy*, 57(14):461–471, 1960.
- [22] Eric Pacuit. *Neighborhood semantics for modal logic*. Springer, 2017.
- [23] Stewart Shapiro. Logical Consequence, Proof Theory, and Model Theory. In *The Oxford Handbook of Philosophy of Mathematics and Logic*, pages 651–670. Oxford University Press, 2005.
- [24] Peter Smith. Squeezing arguments. *Analysis*, 71(1):22–30, 2011.
- [25] Robert Stalnaker. The problem of logical omniscience, I. *Synthese*, pages 425–440, 1991.
- [26] Robert Stalnaker. On logics of knowledge and belief. *Philosophical studies*, 128(1):169–199, 2006.
- [27] Dirk van Dalen. Intuitionistic logic. In *Handbook of philosophical logic*, pages 225–339. Springer, 1986.
- [28] Dirk van Dalen and Anne S Troelstra. *Constructivism in Mathematics: an introduction*, volume I. North-Holland, 1988.
- [29] Achille C. Varzi. On logical relativity. *Philosophical Issues*, 12(1):197–219, 2002.
- [30] Denis A. Vladimirov. *Boolean algebras in analysis*, volume 540. Springer Science & Business Media, 2013.

A FRAMEWORK FOR MODELLING MOLECULAR INTERACTION MAPS

JEAN-MARC ALLIOT

*Institut de Recherche en Informatique de Toulouse. Toulouse University, Toulouse,
France*

jean-marc.alliot@irit.fr

MARTA CIALDEA MAYER

Dipartimento di Ingegneria, Università degli Studi Roma Tre, Rome, Italy

cialdea@ing.uniroma3.it

ROBERT DEMOLOMBE

*Institut de Recherche en Informatique de Toulouse. Toulouse University, Toulouse,
France*

demolombe@irit.fr

MARTÍN DIÉGUEZ

*Institut de Recherche en Informatique de Toulouse. Toulouse University, Toulouse,
France*

University of Pau. Pau, France

martin.dieguez@irit.fr

LUIS FARIÑAS DEL CERRO

*Institut de Recherche en Informatique de Toulouse. Toulouse University, Toulouse,
France*

farinas@irit.fr

Abstract

Metabolic networks, formed by a series of metabolic pathways, are made of intracellular and extracellular reactions that determine the biochemical properties of a cell, and by a set of interactions that guide and regulate the activity of these reactions. Most of these pathways are formed by an intricate and complex

network of chain reactions, and can be represented in a human readable form using graphs which describe the cell cycle checkpoint pathways.

This paper proposes a method to represent Molecular Interaction Maps (graphical representations of complex metabolic networks) in Linear Temporal Logic. The logical representation of such networks allows one to reason about them, in order to check, for instance, whether a graph satisfies a given property ϕ , as well as to find out which initial conditions would guarantee ϕ , or else how can the graph be updated in order to satisfy ϕ .

Both the translation and resolution methods have been implemented in a tool capable of addressing such questions thanks to a reduction to propositional logic which allows exploiting classical SAT solvers.

1 Introduction

Metabolic networks, formed by a series of metabolic pathways, are made of intracellular and extracellular reactions that determine the biochemical properties of a cell by consuming and producing proteins, and by a set of interactions that guide and regulate the activity of such reactions. Cancer, for example, can sometimes appear in a cell as a result of some pathology in a metabolic pathway. These reactions are at the center of a cell's existence, and are regulated by other proteins, which can either activate these reactions or inhibit them. These pathways form an intricate and complex network of chain reactions, and can be represented in a human readable form using graphs, called Molecular Interaction Maps (MIMs) [26, 33] which describe the cell cycle checkpoint pathways (see for instance Figure 1).

Although capital for Knowledge Representation (KR) in biology, MIMs are difficult to use due to the very large number of elements they may involve and the intrinsic expertise needed to understand them. Moreover, the lack of a formal semantics for MIMs makes it difficult to support reasoning tasks commonly carried out by experts, such as checking properties on MIMs, determining how a MIM can explain a given property or how a MIM can be updated in order to describe empirically obtained evidences.

This contribution carries on the research undertaken by the authors aiming at providing a formal background to study MIMs. A first set of works proposed a formalisation of MIMs based on a decidable fragment of first-order logic [14, 15, 16]. In an attempt to find a simpler representation, without resorting to the expressivity of first-order logic, other works [1, 2, 3] proposed an ad-hoc defined non-monotonic logic, called Molecular Interaction Logic (MIL), allowing one to formalize the notions of *production* and *consumption* of reactives. In order to formalise the “temporal evolution” of a biological system, MIL formulae are then mapped into Linear Temporal

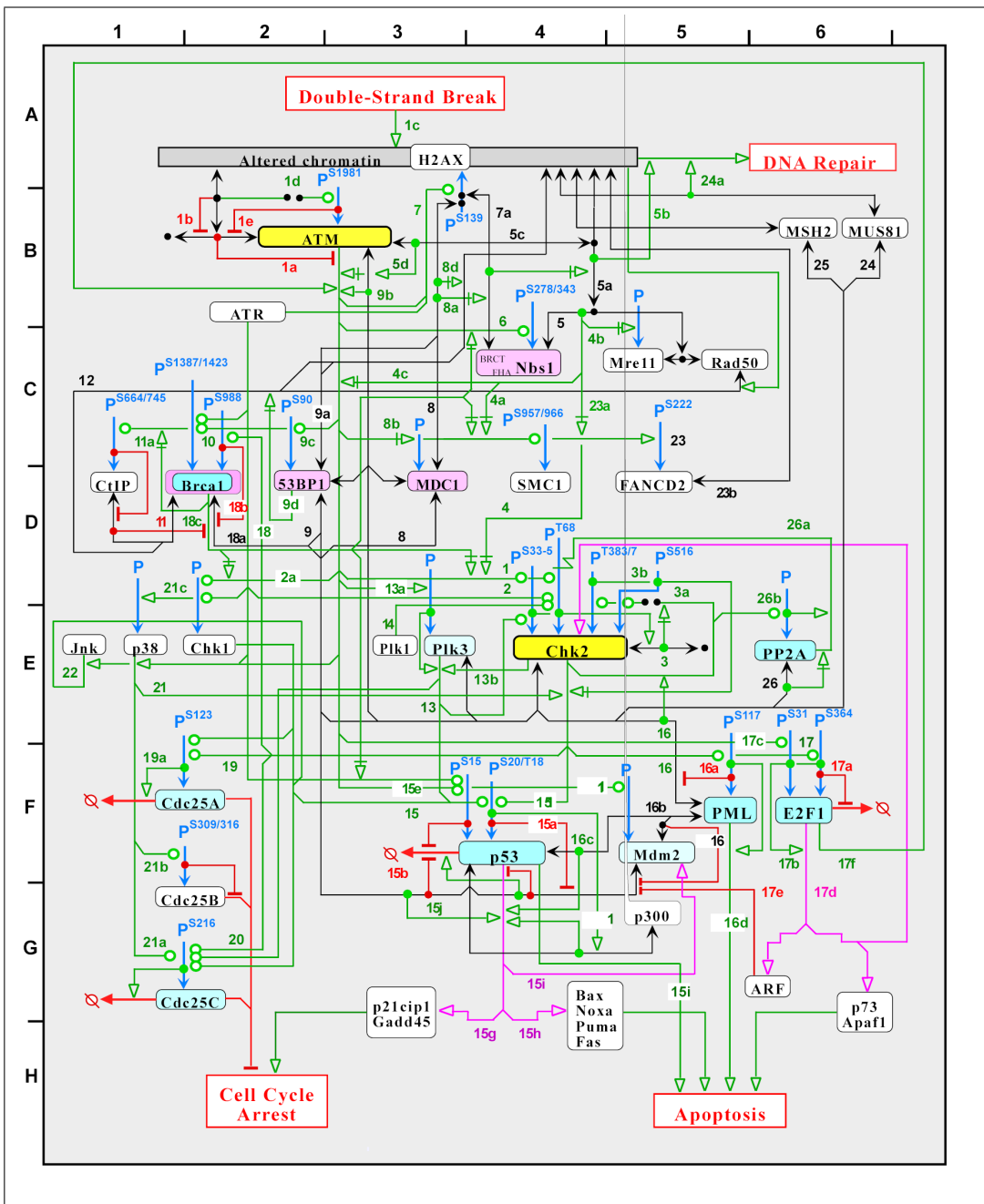


Figure 1: atm-chk2/atr-chk1 molecular interaction map.

Logic (LTL) [32].

This paper embraces the idea, proposed by the above mentioned works, that LTL is a suitable framework for modelling biological systems due to its ability to describe the interaction between components (represented by propositional variables) and their presence/absence in different time instants. Beyond giving a formal definition of graphs representing MIMs, the paper shows how they can be modeled as an LTL theory, by means of a direct “encoding”, without resorting to intermediate (and cumbersome) ad-hoc logics. The logical encoding allows one to formally address reasoning tasks, such as, for instance, checking whether a graph satisfies some given property ϕ , as well as finding out which initial conditions would guarantee ϕ , or else how the graph can be updated in order to satisfy ϕ . A first prototypal system has been implemented on the basis of the theoretical work, allowing one to automatically accomplish reasoning tasks on MIMs.

It is worth pointing out that the adequacy of LTL to model MIMs is due to the fact that the latter are qualitative representations of biological processes. In other terms, they model the interactions among the different components of a biological system without resorting, for instance, to differential equations like the Systems Biology Markup Language (SBML) [22] does.

The rest of this paper is organized as follows. Section 2 gives a brief overview of modelling approaches for networks of biological entities. Section 3 presents the lac operon that will be used as a leading example to introduce all the concepts dealt with by our approach. Section 4 describes the fundamental elements and concepts of the modelling approach. Section 5 presents Molecular Interaction Graphs (MIGs), which formalize Molecular Interaction Maps capable of describing and reasoning about general pathways. Section 6 explains how MIGs can be represented by use of Linear Temporal Logic and Section 7 how reasoning tasks on such representations can be reduced to classical propositional logic by assuming boundet time, so that SAT tools can be exploited. Section 8 describes the current state of the operational implementation of the software tool and Section 9 presents some examples on larger problems. Finally, Section 10 concludes this paper and discusses possible future work.

2 Logical Approaches to Biological Systems

The typical objects to be modelled in the framework of systems biology are *networks* of interacting elements that evolve in time. According to the features of the network and its properties, various approaches can be followed, which can describe the dynamics of the system taking the following elements into consideration:

- **Components:** they are represented by variables, which can be either *discrete* or *continuous* depending on the requirements of the model.
- **Interactions:** they are represented by rules that specify the dynamical changes in the variables values. These interactions can in their turn be classified according to the adopted representation of time (*discrete* or *continuous*). Finally, the execution of an action can be either *stochastic* or not, if a certain degree of uncertainty is considered, reflecting the assumption of a noisy environment.

According to the different possible semantics, the various modelling approaches may be classified as follows [20]:

- Models that involve component quantities and deterministic interactions: such models are mathematical, inherently quantitative and usually based on ordinary differential equations. Tools like Timed Automata representations or Continuous-Time Markov Chains are used in the construction of models of this category.
- Discrete-value models: they are characterised by the use of discrete time. Approaches like executable models based on Finite State Machines representations or stochastic models such as Discrete-Time Markov Chains belong to this category.

Other hybrid models such as Hybrid Automata or Process Algebraic Techniques, mix discrete and continuous representation for both variables and time dynamics. Biological properties can be distinguished between *qualitative* and *quantitative*: in the former case, time has an implicit consideration while the latter involves reasoning on the dynamics of the system along time. To give an example, *reachability* and *temporal ordering of events* are considered qualitative properties while *equilibrium states* and *matabolite* dynamics are quantitative properties.

Gene Regulatory Networks (GRNs) have been very well studied in the temporal context because the interaction between components may be easily represented by their presence/absence, i.e., components are represented by boolean variables and interactions are represented by logical rules on their values. Following this approach Chabrier et al. [10] successfully modeled a very large network, involving more than 500 genes. They resorted to Concurrent Transition Systems (CTS), allowing one to model modular systems, and CTSs can then be translated into the NuSMV language. They checked reachability, stability and temporal ordering properties by the use of CTL. A similar study on a much smaller (although real) biological system has been

performed in [6]. Here the LTL specification syntax and the Spin model checker are used to verify stability properties.

When a quantitative approach is chosen, the model dimensions drop drastically. This is essentially due to lack of knowledge on the parameter values for all the interactions, and to the increased computational complexity deriving from a large model. In this kind of settings, logical approaches have been used to verify temporal properties on the representations. For instance, [4, 18, 17] use CTL to verify, among other properties, reachability and stability on different types of biological networks, and in [7] such properties are checked by using LTL. All these approaches are supported not only by theoretical results but also by tools and frameworks that allow biologists to describe a biological network and then verify whether such representations satisfy some desired properties. Among others, the systems BIOCHAM [8], Bio-PEPA [12, 30] and ANIMO [36] are very popular in the community. We refer the reader to [35, 19] for an overview on this topic.

Some considerations can be made from the study of the aforementioned contributions:

- the size of the modelled systems is generally very small, and a great degree of abstraction and suitable tools are needed to deal with large models;
- qualitative approaches are generally enough to analyse a large variety of interesting biological properties;
- temporal logic plays an important role in the representation and verification of biological systems.

Contrary to approaches incorporating quantitative information into the temporal formalisation [11], our contribution belongs to the category of qualitative approaches, since quantitative information in biological relations, such as the quantity of reactives and their speed of consumption in a reaction, are not formalised. MIMs in fact represent the interaction among the different components of the system and how they evolve in time according to the different reactions. To the best of our knowledge, there is not any contribution where MIMs are used to model quantitative biological information.

3 A simple example: the lac operon

This section describes a simple example, which represents the regulation of the *lac* operon (lactose operon),¹ already used in [1, 3]. The *lac* operon is an operon required for the transport and metabolism of lactose in many bacteria. Although glucose is the preferred carbon source for most bacteria, the *lac* operon allows for the effective digestion of lactose when glucose is not available. The *lac* operon is a sequence of three genes (*lacZ*, *lacY* and *lacA*) which encodes 3 enzymes which in turn carry the transformation of lactose into glucose. We will concentrate here on *lacZ* which encodes β -galactosidase which cleaves lactose into glucose and galactose.

The *lac* operon uses a two-part control mechanism to ensure that the cell expends energy producing the enzymes encoded by the *lac* operon only when necessary. First, in the absence of lactose, the *lac* repressor halts production of the enzymes encoded by the *lac* operon. Second, in the presence of glucose, the catabolite activator protein (CAP), required for production of the enzymes, remains inactive.

Figure 2 describes this regulatory mechanism. The expression of *lacZ* gene is only possible when RNA polymerase (pink) can bind to a promotor site (marked P, black) upstream the gene. This binding is aided by the cyclic adenosine monophosphate (CAMP protein, in blue) which binds before the promotor on the CAP site (dark blue).

The *lacI* gene (yellow) encodes the repressor protein *LacI* (yellow) which binds to the promotor site of the RNA polymerase when lactose is not available, preventing the RNA polymerase to bind to the promotor and thus blocking the expression of the following genes (*lacZ*, *lacY* and *lacA*): this is a *negative regulation*, or *inhibition*, as it blocks the production of the proteins. When lactose is present, one of its isomer, allolactose, binds with repressor protein *LacI* which is no longer able to bind to the promotor site, thus enabling RNA polymerase to bind to the promotor site and to start expressing the *lacZ* gene if CAMP is bound to CAP.

The CAMP molecule is on the opposite a *positive regulation* molecule, or an *activation* molecule, as its presence is necessary to express the *lacZ* gene. However, the concentration of CAMP is itself regulated negatively by glucose: when glucose is present, the concentration of CAMP becomes low, and thus CAMP does not bind to the CAP site, blocking the expression of *lacZ*. Thus glucose prevents the activation by CAMP of the expression of galactosidase from *lacZ*.

¹The Nobel prize was awarded to Monod, Jacob and Lwoff in 1965 partly for the discovery of the *lac* operon by Monod and Jacob [24], which was the first genetic regulatory mechanism to be understood clearly, and is now a “standard” introductory example in molecular biology classes. See also [37].

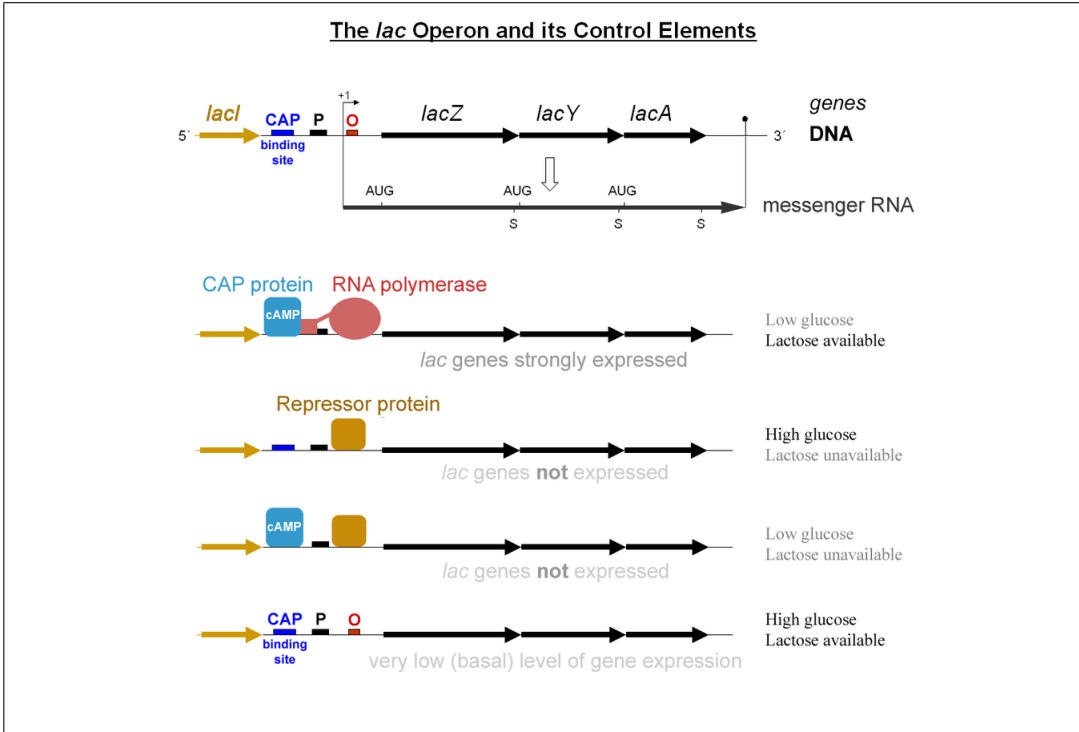


Figure 2: The lac operon

4 Molecular Interaction Maps (MIMs)

The mechanism described in the previous section is represented in Figure 3, which is an example of MIM.² This example contains all the relations and all the categories of entities (i.e. the nodes of the graph) that we use in our modelling. They are presented below.

4.1 Relations

The relations among the entities (represented by links in the graphs) represent reactions and can be of different types:

Productions can take two different forms, depending on whether the reactants are consumed by the reactions or not:

²Technically, the generation of cAMP from Adenosine Tri Phosphate (ATP) is blocked by the presence of glucose, but we have simplified the graph by simply writing that the presence of glucose prevents the activation by cAMP of the expression of galactosidase from *lacZ*.

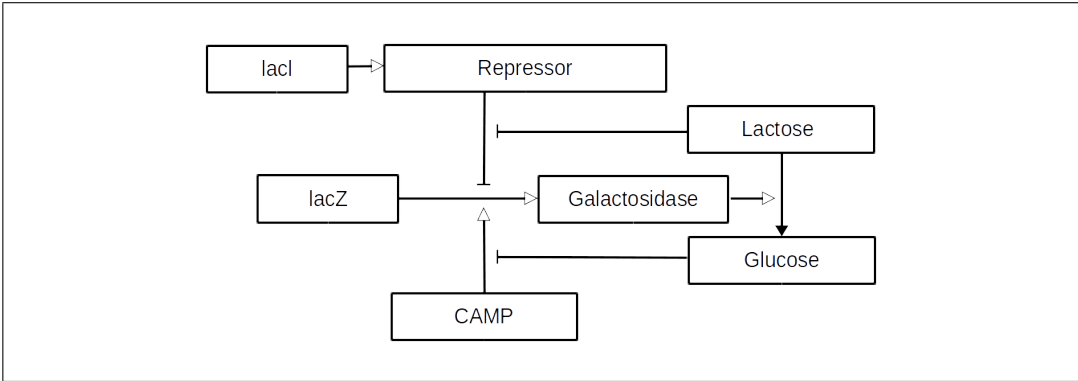


Figure 3: Functional representation of the lac operon

- 1) The graphical notation used when a reaction consumes completely the reactant(s) is $a_1, \dots, a_n \rightarrow b$, meaning that the production of b completely consumes a_1, \dots, a_n .

For instance, in Figure 3, lactose, when activated by galactosidase produces glucose, and is consumed while doing so, which is thus noted by $lactose \rightarrow glucose$.

- 2) If the reactants are not completely consumed by the reaction, the used notation is $a_1, \dots, a_n \rightarrow b$. Here b is produced but a_1, \dots, a_n are still present after the production of b .

For example, the expression of the *lacZ* gene to produce galactosidase (or of the *lacI* gene to produce the repressor protein) does not consume the gene, and we thus have $lacZ \rightarrow galactosidase$.

Regulations are also of two types: every reaction can be either *inhibited* or *activated* by other proteins or conditions.

1. The notation of the type $a_1, \dots, a_n \rightarrow \dots$ means that the simultaneous presence of a_1, \dots, a_n activates a production or another regulation.

In the example of Figure 3 the production of galactosidase from the expression of the *lacZ* gene is activated by CAMP ($CAMP \rightarrow (lacZ \rightarrow Galactosidase)$ expresses activation).

2. The notation $a_1, \dots, a_n \dashrightarrow \dots$ represents the fact that simultaneous presence of a_1, \dots, a_n inhibits a production or another regulation.

In Figure 3, $Repressor \dashrightarrow (lacZ \rightarrow Galactosidase)$ represents the fact that production of galactosidase is blocked (or inhibited) by the repressor protein.

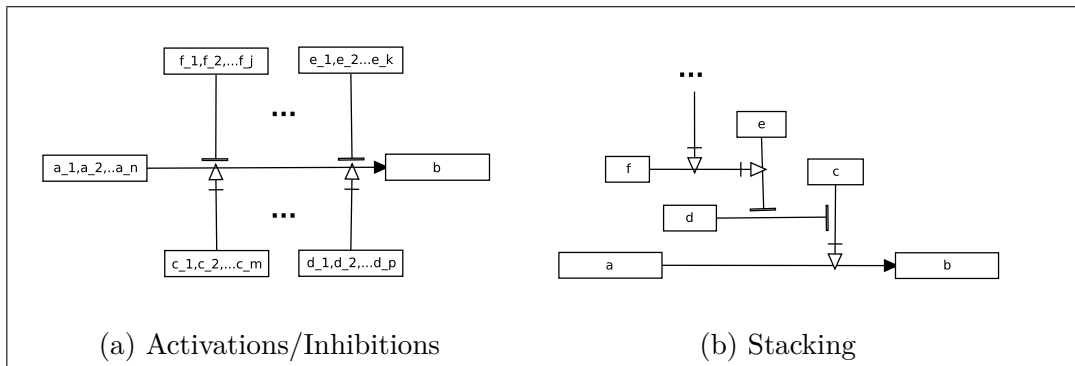


Figure 4: Activations/Inhibitions and Stacking

Figure 4.(a) shows the basic inhibitions/activations on a reaction. The production of b from a_1, \dots, a_n is activated by the simultaneous presence of **both** c_1, \dots, c_m **and** the simultaneous presence of d_1, \dots, d_p . It is inhibited by **either** the simultaneous presence of e_1, \dots, e_k **or** the simultaneous presence of f_1, \dots, f_j .

These regulations are often “stacked”, on many levels, like shown in Figure 4.(b). For example in Figure 3, the inhibition by the repressor protein of the production of galactosidase can itself be inhibited by the presence of lactose, while the activation of the same production by CAMP is inhibited by the presence of glucose.

4.2 Types of Entities

Entities occurring in node labels can be of two different types:

Exogenous: the value of an exogenous variable is set once and for all by the environment or by the experimenter at the start of the simulation and *never* changes through time; if the entity is set as present and used in a reaction, the environment will always provide “enough” of it and it will remain present.

Endogenous: an endogenous entity can either be present or absent at the beginning of the process, as set by the experimenter, and its value after the start of the process is set only by the dynamics of the graph.

These distinctions are fundamental, because the dynamics of entities are different and they must be formalized differently. In practice, the type of an entity is something which is set by the biologist, according to his professional understanding of the biological process described by the map. For instance, in Figure 3, the type of the different entities could be set as follows in order to describe the real behaviour of the lac operon: *lacI*, *lacZ*, CAMP and lactose are initial external conditions of the

model and they do not evolve in time, and are thus exogenous. Note, in particular, that lactose can be set as an exogenous entity, even if the graph “says” that it is consumed when producing glucose. Conversely, galactosidase, the repressor protein and glucose can be produced inside the graph, and are thus endogenous.

It is important to notice that glucose could be set as an exogenous variable if the experimenter is interested in testing an environment where glucose is provided externally. Reciprocally, in a more accurate representation of the lac operon, CAMP would be an endogenous variable, produced by ATP and regulated by glucose. These graphs are only a representation and an approximation of the real process, designed to fit the particular level of description that the experimenter wants to model.

Although MIMs may contain also other kinds of entities or links, the two kinds of entities and four kinds of interactions presented above are all that is needed to build the Molecular Interactions Maps we are using in this paper.

4.3 Temporal evolution

A MIM can be considered as an automaton which produces sequences of states of its entities and Linear Temporal Logic formulas can well describe such sequences of states. Time is supposed to be discrete, and all relations (productions/consumptions) that *can* be executed *are* executed simultaneously at each time step. An entity can have two states (or values): absent (0) or present (1). When an entity is consumed, it becomes absent and when it is produced it becomes present. In other terms, since quantities are not taken into account, due also to the lack of reliable data thereon, reactions do not contend to get use of given resources: if an entity is present, its quantity is assumed to be enough to be used by all reactions needing it.

This behaviour might look simplistic, as it does not take into account the kinetic of reactions, but it reflects a choice underlying MIMs representation framework and, as a matter of fact, it is nevertheless adequate to handle many problems.

The software tool that will be described in Section 8 provides default values both for the variables and for their classification as exogenous or endogenous. However, the user can modify such default settings through the graphical interface.

5 Molecular Interaction Graphs

This section is devoted to define *Molecular Interaction Graphs* (MIGs), the graph structures which are the formal representations of MIMs. The concept of *trace* will also be defined, with the aim of characterising the dynamic behaviour of a MIM.

A MIG is essentially a graph whose vertices are identified with finite sets of atoms, each of which represents a molecule. Productions are represented by links

connecting vertices, while regulations are links whose origin is a vertex and whose target is another link.

Definition 1 (Molecular Interaction Graph). *A Molecular Interaction Graph (MIG) \mathcal{G} is a tuple $\langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ where*

- *Ex and Ed are two disjoint finite sets of atoms. Their union $At = Ex \cup Ed$ constitutes the sets of atoms of the graph language. The atoms in Ex are called exogenous and those in Ed are the endogenous ones.*
- *\mathcal{B} is a consistent set of literals from atoms in At , i.e., $\mathcal{B} \subset At \cup \{\neg p \mid p \in At\}$. It represents the initial conditions on the graph: for all $p \in At$, if $p \in \mathcal{B}$ then p is initially true, and if $\neg p \in \mathcal{B}$, then p is initially false; otherwise, p is unset (or a free atom).*
- *\mathcal{P} and \mathcal{C} are sets of productions:*

$$\mathcal{P} \subseteq \{(P \rightarrow Q) \mid P, Q \subseteq At\} \quad \mathcal{C} \subseteq \{(P \rightarrow Q) \mid P, Q \subseteq At\}$$

- *\mathcal{A} and \mathcal{I} are sets of regulations, such that for some $n \in \mathbb{N}$:*

$$\mathcal{A} = \bigcup_{i=0}^n \mathcal{A}_i \quad \mathcal{I} = \bigcup_{i=0}^n \mathcal{I}_i$$

where \mathcal{A}_i and \mathcal{I}_i are inductively defined as follows:

$$\begin{aligned} \mathcal{A}_0 &\subseteq \{(P \rightarrow X) \mid X \in \mathcal{P} \cup \mathcal{C}\} & \mathcal{I}_0 &\subseteq \{(P \rightarrow X) \mid X \in \mathcal{P} \cup \mathcal{C}\} \\ \mathcal{A}_{i+1} &\subseteq \{(P \rightarrow X) \mid X \in \mathcal{A}_i \cup \mathcal{I}_i\} & \mathcal{I}_{i+1} &\subseteq \{(P \rightarrow X) \mid X \in \mathcal{A}_i \cup \mathcal{I}_i\} \end{aligned}$$

A link is either a production (i.e. an element of $\mathcal{P} \cup \mathcal{C}$) or a regulation (an element of $\mathcal{A} \cup \mathcal{I}$). The depth of a regulation X is the integer k such that $X \in \mathcal{A}_k \cup \mathcal{I}_k$.

The “stratified” definition of regulations rules out the possibility of circular chains of activations and inhibitions. Furthermore, it is worth pointing out that, since At is a finite set of atoms, then also the sets \mathcal{P} and \mathcal{C} are finite. Consequently, \mathcal{A} and \mathcal{I} are finite sets too, since the depth of their elements is bounded by a given fixed $n \in \mathbb{N}$.

Note that Definition 1 is a generalization w.r.t. the presentation of productions given in Section 4, in that it allows for multiple entities on the right-hand side of a production. This extension can be considered as an abbreviation: $a_1, \dots, a_n \multimap b_1, \dots, b_k$ (where \multimap is either \rightarrow or \rightarrow) stands for the set of productions $a_1, \dots, a_n \multimap b_1, \dots, a_1, \dots, a_n \multimap b_k$.

Example 1. *The MIG representing the MIM shown in Figure 3, ignoring the initial conditions and the partitioning of At into exogenous and endogenous atoms, is constituted by*

$$\begin{aligned}
At &= \{Lactose, Galactosidase, Glucose, CAMP, lacZ, Galactosidase, \\
&\quad Lactose, Repressor, lacI\} \\
\mathcal{P} &= \{(lacI \rightarrow Repressor), (lacZ \rightarrow Galactosidase)\} \\
\mathcal{C} &= \{(Lactose \rightarrow Glucose)\} \\
\mathcal{A} &= \{(CAMP \rightarrow (lacZ \rightarrow Galactosidase)), \\
&\quad (Galactosidase \rightarrow ((Lactose \rightarrow Glucose)))\} \\
\mathcal{I} &= \{(Repressor \rightarrow (lacZ \rightarrow Galactosidase)), \\
&\quad (Lactose \rightarrow (Repressor \rightarrow (lacZ \rightarrow Galactosidase))), \\
&\quad (Glucose \rightarrow (CAMP \rightarrow (lacZ \rightarrow Galactosidase)))\}
\end{aligned}$$

Having defined the structure of MIGs, we now need to provide some machinery that allows one to determine the set of substances that trigger an activation (resp. inhibition) in a MIG. The next definition introduces functions whose values are the regulations directly activating/inhibiting a link X in a MIG.

Definition 2 (Direct regulations of a link). *For every link $X \in \mathcal{P} \cup \mathcal{C} \cup \mathcal{A} \cup \mathcal{I}$:*

$$\begin{aligned}
\gamma_a(X) &= \{Y \in \mathcal{A} \mid Y \text{ has the form } (P \rightarrow X)\} \\
\gamma_i(X) &= \{Y \in \mathcal{I} \mid Y \text{ has the form } (P \rightarrow X)\}
\end{aligned}$$

Similarly to transition systems, a MIG constitutes a compact representation of a set of infinite sequences of states, where every state is determined by the set of proteins, genes, enzymes, metabolites, etc that are present in the cell at a given time. A sequence of such states thus represents the temporal evolution of the cell, and will be called a *trace*. Differently from transition systems, however, the evolution of a MIG is deterministic: each possible initial configuration determines a single trace. The reason for this is that the representation abstract from quantities, hence entities are not considered as resources over which reactions may compete (see the remark at the end of Section 4).

Before formally defining the concept of trace, we introduce some preliminary concepts such as the notion of *active* and *inhibited* links. These two concepts, that are relative to a given situation (i.e. a given set of atoms assumed to be true), will provide the temporal conditions under which a production can be triggered.

Definition 3 (Active and inhibited links). *Let $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ be a MIG and $At = Ex \cup Ed$. Given $D \subseteq At$ and $\rightarrow \in \{\rightarrow, \rightarrow, \rightarrow\}$, a link $X = (P \rightarrow Y) \in \mathcal{P} \cup \mathcal{C} \cup \mathcal{A} \cup \mathcal{I}$ (where Y is either a set of atoms or a link) is said to be active in D if the following conditions hold:*

1. $P \subseteq D$;
2. every $Z \in \gamma_a(X)$ is active in D – i.e. every regulation of the form $Q \rightarrow X \in \mathcal{A}$ is active in D ;
3. for all $Z \in \gamma_i(X)$, Z is not active in D – i.e. there are no regulations of the form $(Q \rightarrow X) \in \mathcal{I}$ that are active in D .

A link $X \in \mathcal{P} \cup \mathcal{C} \cup \mathcal{A} \cup \mathcal{I}$ is inhibited in D iff X is not active in D .

Before formalising the concepts of *production* and *consumption* of substances inside a cell, it is worth pointing out that:

- 1) a substance is produced in a cell as a result of a reaction, which is triggered whenever the *reactants* are present and the regulation conditions allow its execution.
- 2) A substance is consumed in a cell if it acts as a *reactive* in a reaction which has been triggered.
- 3) We do not consider quantitative information like concentrations or reaction times: if a substance is involved in several reactions at a time, its concentration does not matter, all reactions will be triggered. Conversely, if a substance belongs to the consumed *reactants* of a triggered reaction, it will be completely consumed.
- 4) It might be the case that a substance is consumed in a reaction while produced by a different one, at the same time. This possibility, that will be further commented below, will however raise no inconsistency in the definition of traces.

Definition 4 (Produced and consumed atoms). Let $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ be a MIG and $D \subseteq At = Ex \cup Ed$. An atom $p \in At$ is produced in D iff $p \in Ed$ and there exists $(P \multimap Q) \in \mathcal{P} \cup \mathcal{C}$, for $\multimap \in \{\rightarrow, \rightarrow\}$, such that:

- (i) $p \in Q$ and
- (ii) $(P \multimap Q)$ is active in D .

An atom p is consumed in D iff $p \in Ed$ and there exists $(P \rightarrow Q) \in \mathcal{C}$ such that

- (i) $p \in P$ and
- (ii) $(P \rightarrow Q)$ is active in D .

Remark 1. It may happen that an atom p is both produced and consumed in a given $D \subseteq At$. Consider, for instance, a MIG with $Ed = \{p, q, r\}$, $Ex = \emptyset$, $\mathcal{P} = \{(p \rightarrow q)\}$,

$\mathcal{C} = \{(q \rightarrow r)\}$, $\mathcal{I} = \mathcal{A} = \emptyset$. If $D = \{p, q\}$, the atom q is produced by $(p \rightarrow q)$ and consumed by $(q \rightarrow r)$ in D , since $\{p\} \subseteq D$, $q \in \{q\} \subseteq D$ and there are no regulations governing these two productions. An even simpler example is given by the (unrealistic) MIG with $Ed = \{p\}$, $Ex = \emptyset$, $\mathcal{C} = \{(p \rightarrow p)\}$, $\mathcal{P} = \mathcal{I} = \mathcal{A} = \emptyset$ and $D = \{p\}$.

The behaviour of a MIG can be finally formally defined in terms of its *trace*, taking into account activations, inhibitions, productions and consumptions.

Definition 5 (Trace). A trace T on a set At of atoms is an infinite sequence of subsets of At , T_0, T_1, \dots , called states. If $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ is a MIG, a trace for \mathcal{G} is a trace T on $At = Ex \cup Ed$ such that:

1. $p \in T_0$ for every $p \in \mathcal{B}$ and $p \notin T_0$ for every $\neg p \in \mathcal{B}$;
2. for all $k \geq 0$ and every atom $p \in At$:
 - if $p \in Ex$, then $p \in T_{k+1}$ iff $p \in T_k$;
 - if $p \in Ed$, then $p \in T_{k+1}$ if and only if either p is produced in T_k or $p \in T_k$ and p is not consumed in T_k .

It is worth pointing out that the condition on traces for a given MIG \mathcal{G} ensures that every change in a state of the trace affecting endogenous atoms has a justification in \mathcal{G} . Consequently, given the initial state T_0 of a trace for \mathcal{G} , all the others are deterministically determined by the productions and regulations of \mathcal{G} .

As a final observation we remark that, when an atom p is both produced and consumed in a given T_k , production prevails over consumption. For instance, in a trace for the MIG of Remark 1 with $T_0 = \{p, q\}$, where q is both produced and consumed, $T_k = \{p, q, r\}$ for all $k \geq 1$.

6 Representing MIGs in Linear Temporal Logic

This section considers the connection between traces and LTL and describes how to represent a MIG \mathcal{G} by means of an LTL theory whose models are exactly the traces for \mathcal{G} .

LTL formulae with only unary future time operators are built from the grammar

$$\varphi ::= \perp \mid p \mid \neg\varphi \mid \varphi \vee \varphi \mid \bigcirc\varphi \mid \square\varphi$$

where p is an atom (the other propositional connectives and the “eventually” operator can be defined as usual).

An LTL *interpretation* T is a trace, i.e. an infinite sequence T_0, T_1, \dots of *states*, where a state is a set of atoms. The satisfaction relation $T_k \models \varphi$, where T_k is a state and φ a formula built from a set of atoms At , is defined as follows:

1. $T_k \models p$ iff $p \in T_k$, for any $p \in At$;
2. $T_k \not\models \perp$;
3. $T_k \models \neg\varphi$ iff $T_k \not\models \varphi$;
4. $T_k \models \varphi \vee \psi$; iff $T_k \models \varphi$ or $T_k \models \psi$;
5. $T_k \models \bigcirc\varphi$ iff $T_{k+1} \models \varphi$;
6. $T_k \models \Box\varphi$ iff for all $j \geq k$, $T_j \models \varphi$;

A formula φ is true in an interpretation T if and only if $T_0 \models \varphi$.

A MIG $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ is represented by means of a set of LTL formulae on the set of atoms At . First of all, classical formulae representing the fact that a given link is active (or inhibited) are defined. Below, \multimap stands for any of \rightarrow , \rightarrow or \rightarrow

Definition 6. Let $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ be a MIG. If $X = (P \multimap Y) \in \mathcal{P} \cup \mathcal{C} \cup \mathcal{A} \cup \mathcal{I}$ (where Y is either a set of atoms or a link), then:

$$\mathbf{A}(P \multimap Y) \stackrel{\text{def}}{=} \bigwedge_{p \in P} p \wedge \bigwedge_{\rho \in \gamma_a(X)} \mathbf{A}(\rho) \wedge \bigwedge_{\rho \in \gamma_i(X)} \mathbf{I}(\rho)$$

where $\mathbf{I}(X)$ is an abbreviation for the negation normal form of $\neg\mathbf{A}(X)$.

It is worth pointing out that both $\mathbf{A}(X)$ and $\mathbf{I}(X)$ are classical propositional formulae.

Example 2. Let us consider, for instance, the links of the MIG \mathcal{G} of Example 1:

- 1) ($lacI \rightarrow Repressor$)
- 2) ($lacZ \rightarrow Galactosidase$)
- 3) ($Lactose \rightarrow Glucose$)
- 4) ($CAMP \rightarrow (lacZ \rightarrow Galactosidase)$)
- 5) ($Galactosidase \rightarrow (Lactose \rightarrow Glucose)$)
- 6) ($Repressor \rightarrow (lacZ \rightarrow Galactosidase)$)
- 7) ($Lactose \rightarrow (Repressor \rightarrow (lacZ \rightarrow Galactosidase))$)
- 8) ($Glucose \rightarrow (CAMP \rightarrow (lacZ \rightarrow Galactosidase))$)

For each of them, $\mathbf{A}(X)$ and $\mathbf{I}(X)$ can be computed as follows:

$$\begin{aligned}
 \mathbf{A}(1) &= \mathbf{A}(lacI \rightarrow Repressor) = lacI; \\
 \mathbf{A}(7) &= \mathbf{A}(Lactose \rightarrow (Repressor \rightarrow (lacZ \rightarrow Galactosidase))) = Lactose; \\
 \mathbf{A}(8) &= \mathbf{A}(Glucose \rightarrow (CAMP \rightarrow (lacZ \rightarrow Galactosidase))) = Glucose; \\
 \mathbf{A}(4) &= \mathbf{A}(CAMP \rightarrow (lacZ \rightarrow Galactosidase)) = CAMP \wedge \mathbf{I}(8) = \\
 &\quad CAMP \wedge \neg Glucose; \\
 \mathbf{A}(5) &= \mathbf{A}(Galactosidase \rightarrow (Lactose \rightarrow Glucose)) = Galactosidase; \\
 \mathbf{A}(3) &= \mathbf{A}(Lactose \rightarrow Glucose) = Lactose \wedge \mathbf{A}(5) = Lactose \wedge Galactosidase; \\
 \mathbf{A}(6) &= \mathbf{A}(Repressor \rightarrow (lacZ \rightarrow Galactosidase)) = Repressor \wedge \mathbf{I}(7) = \\
 &\quad Repressor \wedge \neg Lactose; \\
 \mathbf{A}(2) &= \mathbf{A}(lacZ \rightarrow Galactosidase) = lacZ \wedge \mathbf{A}(4) \wedge \mathbf{I}(6) = \\
 &\quad lacZ \wedge CAMP \wedge \neg Glucose \wedge (\neg Repressor \vee Lactose);
 \end{aligned}$$

The next result establishes that $\mathbf{A}(X)$ is an adequate representation of the property of being active for the link X .

Lemma 1. *Let $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ be a MIG and $D \subseteq Ex \cup Ed$. For every link $X \in \mathcal{P} \cup \mathcal{C} \cup \mathcal{A} \cup \mathcal{I}$: $D \models \mathbf{A}(X)$ if and only if X is active in D .*

Proof. Let the *size* of a link X , $size(X)$, be defined as the number of arrows $\rightarrow \in \{\rightarrow, \rightarrow, \rightarrow\}$ occurring in X , and let M be the maximal size of a link in \mathcal{G} . If $X = P \rightarrow Y$ is any link in \mathcal{G} , the proof is by induction on $k = M - size(X)$.

- If $k = 0$, then \mathcal{G} does not have any link of size greater than $size(X)$, hence $\gamma_a(X) = \gamma_i(X) = \emptyset$, $\mathbf{A}(P \rightarrow Y) = \bigwedge_{p \in P} p$, and X is active in D iff $P \subseteq D$.

Clearly, $D \models \bigwedge_{p \in P} p$ iff $P \subseteq D$.

- If $k > 0$, then, for every $Z \in \gamma_a(X) \cup \gamma_i(X)$, $size(Z) = size(X) + 1$, hence $M - (k + 1) < k$. By the induction hypothesis, $D \models \mathbf{A}(Z)$ iff Z is active in D . Then the thesis follows from the facts that: (i) $D \models \bigwedge_{p \in P} p$ iff $P \subseteq D$; (ii) for all $Z \in \gamma_a(X)$, $D \models \mathbf{A}(Z)$ iff Z is active in D (by the induction hypothesis), and (iii) for all $Z \in \gamma_i(X)$, $D \models \neg \mathbf{A}(Z)$ iff Z is not active in D (by the induction hypothesis).

□

In order to give a more compact presentation of the LTL theory representing a MIG, we define, for each atom $p \in Ex \cup Ed$, classical formulae representing the fact that p is produced or consumed.

Definition 7. Let $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ be a MIG, $\mathcal{P}rod = \mathcal{P} \cup \mathcal{C}$, and $p \in Ed \cup Ex$. Then:

$$\mathbf{Pr}(p) \stackrel{\text{def}}{=} \begin{cases} \perp & \text{if } p \in Ex \\ \bigvee_{(P \multimap Q) \in \mathcal{P}rod, p \in Q} \mathbf{A}(P \multimap Q) & \text{if } p \in Ed \end{cases}$$

$$\mathbf{Cn}(p) \stackrel{\text{def}}{=} \begin{cases} \perp & \text{if } p \in Ex \\ \bigvee_{(P \rightarrow Q) \in \mathcal{C}, p \in P} \mathbf{A}(P \rightarrow Q) & \text{if } p \in Ed \end{cases}$$

Example 3. Let us consider the simple MIG \mathcal{G} of Example 1, where atoms are partitioned into $Ex = \{lacl, lacZ, CAMP\}$ and $Ed = \{Repressor, Lactose, Galactosidase, Glucose\}$.³ The abbreviations $\mathbf{Pr}(p)$ and $\mathbf{Cn}(p)$ for the endogenous atoms are the following:

$$\begin{aligned} \mathbf{Pr}(Repressor) & \stackrel{\text{def}}{=} \mathbf{A}(lacl \rightarrow Repressor) \\ & \stackrel{\text{def}}{=} lacl \\ \mathbf{Pr}(Lactose) & \stackrel{\text{def}}{=} \perp \\ \mathbf{Pr}(Galactosidase) & \stackrel{\text{def}}{=} \mathbf{A}(lacZ \rightarrow Galactosidase) \\ & \stackrel{\text{def}}{=} lacZ \wedge CAMP \wedge \neg Glucose \\ & \quad \wedge (\neg Repressor \vee Lactose) \\ \mathbf{Pr}(Glucose) & \stackrel{\text{def}}{=} \mathbf{A}(Lactose \rightarrow Glucose) \\ & \stackrel{\text{def}}{=} Lactose \wedge Galactosidase \\ \mathbf{Cn}(Lactose) & \stackrel{\text{def}}{=} \mathbf{A}((Lactose \rightarrow Glucose) \\ & \stackrel{\text{def}}{=} Lactose \wedge Galactosidase \\ \mathbf{Cn}(p) & \stackrel{\text{def}}{=} \perp \text{ for } p \in \{Repressor, Galactosidase, Glucose\} \end{aligned}$$

Finally, the set of LTL formulae ruling the overall behaviour of a MIG can be defined.

Definition 8. If $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ is a MIG, the LTL encoding of \mathcal{G} is the set of formulae containing all the literals in \mathcal{B} and, for every $p \in Ex \cup Ed$, the formula

$$\Box(\bigcirc p \leftrightarrow \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p)))$$

³In this example we assume that lactose is endogenous, because it is the only consumed entity in the simple MIM of figure 3.

It is worth pointing out that, if $p \in Ex$, then the formula encoding its behaviour is equivalent to $\Box(\bigcirc p \leftrightarrow p)$. For endogenous atoms, the encoding captures the (negative and positive) effects produced by a reaction on the environment at any time. This encoding has some similarities with the *successor state axioms* of the *Situation Calculus* [34].

Example 4. *If \mathcal{G} is the MIG of Example 1, the LTL encoding of \mathcal{G} contains (formulae equivalent to) $\Box(\bigcirc lacI \leftrightarrow lacI)$, and similar ones for $lacZ$ and $CAMP$.*

Furthermore, it contains the following formulae, ruling the behaviour of endogenous atoms:

$$\begin{aligned}
& \Box(\bigcirc Repressor \leftrightarrow \mathbf{Pr}(Repressor) \vee (Repressor \wedge \neg \mathbf{Cn}(Repressor))) \\
& \quad \equiv \Box(\bigcirc Repressor \leftrightarrow lacI \vee Repressor) \\
& \Box(\bigcirc Lactose \leftrightarrow \mathbf{Pr}(Lactose) \vee (Lactose \wedge \neg \mathbf{Cn}(Lactose))) \\
& \quad \equiv \Box(\bigcirc Lactose \leftrightarrow Lactose \wedge \neg(Lactose \wedge Galactosidase)) \\
& \Box(\bigcirc Galactosidase \leftrightarrow \mathbf{Pr}(Galactosidase) \vee \\
& \quad (Galactosidase \wedge \neg \mathbf{Cn}(Galactosidase))) \\
& \quad \equiv \Box(\bigcirc Galactosidase \leftrightarrow (lacZ \wedge CAMP \wedge \neg Glucose \\
& \quad \wedge (\neg Repressor \vee Lactose) \vee Galactosidase)) \\
& \Box(\bigcirc Glucose \leftrightarrow \mathbf{Pr}(Glucose) \vee (Glucose \wedge \neg \mathbf{Cn}(Glucose))) \\
& \quad \equiv \Box(\bigcirc Glucose \leftrightarrow (Lactose \wedge Galactosidase) \vee Glucose)
\end{aligned}$$

The rest of this section is devoted to show that the LTL encoding of a MIG correctly and completely represents its behaviour. First of all, we prove that the truth of $\mathbf{Pr}(p)$ and $\mathbf{Cn}(p)$ in a state coincides with the atom p being produced/consumed at that state.

Lemma 2. *If T is a model of the LTL encoding of a MIG, then for every k and every atom $p \in At$, p is produced in T_k iff $T_k \models \mathbf{Pr}(p)$ and p is consumed in T_k iff $T_k \models \mathbf{Cn}(p)$.*

Proof. Let T be a model of $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$, $\mathcal{P}rod = \mathcal{P} \cup \mathcal{C}$, $k \in \mathbb{N}$ and $p \in Ex \cup Ed$.

1. If $T_k \models \mathbf{Pr}(p)$ then $\mathbf{Pr}(p) \neq \perp$ and there exists some $(P \multimap Q) \in \mathcal{P}rod$ such that $p \in Q$ and $T_k \models \mathbf{A}(P \multimap Q)$. By Lemma 1, $(P \multimap Q)$ is active in T_k . Moreover, since $\mathbf{Pr}(p) \neq \perp$, $p \in Ed$. Therefore, from Definition 4 it follows that p is produced in T_k .
2. If p is produced in T_k , then $p \in Ed$ and there exists some $(P \multimap Q) \in \mathcal{P}rod$, such that $p \in Q$ and $(P \multimap Q)$ is active in T_k . By Lemma 1, $T_k \models \mathbf{A}(P \multimap Q)$, hence $T_k \models \mathbf{Pr}(p)$ by Definition 7, since $p \in Ed$.

3. If $T_k \models \mathbf{Cn}(p)$ then $\mathbf{Cn}(p) \neq \perp$ and there exists some $(P \rightarrow Q) \in \mathcal{C}$ such that $p \in P$ and $T_k \models \mathbf{A}(P \rightarrow Q)$. By Lemma 1, $P \rightarrow Q$ is active in T_k . Moreover, since $\mathbf{Cn}(p) \neq \perp$, $p \in Ed$. Therefore, from Definition 4 it follows that p is consumed in T_k .
4. If p is consumed in T_k , then $p \in Ed$ and there exists some $(P \rightarrow Q) \in \mathcal{C}$ such that $p \in P$ and $(P \rightarrow Q)$ is active in T_k . By Lemma 1, $T_k \models \mathbf{A}(P \rightarrow Q)$, therefore $T_k \models \mathbf{Cn}(p)$ by Definition 7, since $p \in Ed$.

□

The adequacy of the LTL encoding of a MIG can finally be proved.

Theorem 1 (Main result). *If \mathcal{G} is a MIG, then:*

1. every trace for \mathcal{G} is a model of the LTL encoding of \mathcal{G} ;
2. every model of the LTL encoding of \mathcal{G} is a trace for \mathcal{G} .

Proof. Let us assume that T is a trace for $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$. Clearly, for every literal $\ell \in \mathcal{B}$, $T_0 \models \ell$, since ℓ belongs to the encoding of \mathcal{G} . Moreover, for all $k \geq 0$ and every atom $p \in At = Ex \cup Ed$:

- if $p \in Ex$, then $p \in T_{k+1}$ if and only if $p \in T_k$. Hence, $T_k \models \bigcirc p \leftrightarrow p$, i.e. $T_k \models \bigcirc p \leftrightarrow \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p))$.
- If $p \in Ed$, then $p \in T_{k+1}$ if and only if either p is produced in T_k or $p \in T_k$ and p is not consumed in T_k . By Lemma 2, this amounts to saying that $p \in T_{k+1}$ if and only if either $T_k \models \mathbf{Pr}(p)$ or $p \in T_k$ and $T_k \not\models \mathbf{Cn}(p)$. Consequently, $T_k \models \bigcirc p \leftrightarrow \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p))$.

Since these properties hold for all k , it follows that for all $p \in At$, $T \models \square(\bigcirc p \leftrightarrow \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p)))$.

For the other direction, let us assume that T is a model of the LTL encoding of \mathcal{G} . Then, in particular, $T_0 \models \mathcal{B}$, hence $p \in T_0$ for every $p \in \mathcal{B}$, and $p \notin T_0$ for every $\neg p \in \mathcal{B}$. Moreover, for all $k \geq 0$ and every atom $p \in At$:

- if $p \in Ex$, then $T_k \models \bigcirc p \leftrightarrow p$, hence $p \in T_{k+1}$ if and only if $p \in T_k$.
- If $p \in Ed$, then $T_k \models \bigcirc p \leftrightarrow \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p))$, hence $p \in T_{k+1}$ if and only if either $T_k \models \mathbf{Pr}(p)$ or $p \in T_k$ and $T_k \not\models \mathbf{Cn}(p)$. By Lemma 2, this amounts to saying that $p \in T_{k+1}$ if and only if either p is produced in T_k or $p \in T_k$ and p is not consumed in T_k .

Consequently, T is a trace for \mathcal{G} .

□

7 Bounding Time and Reduction to SAT

The use of an LTL formalization allows us to consider solutions with infinite length when performing reasoning tasks such as abduction⁴ or satisfiability checks. However, LTL tools for abduction are not as developed as in the case of propositional logic, since the abductive task is in general very complex.⁵ In order to take advantage of the highly efficient tools for propositional reasoning such as SAT-solvers, abduction algorithms, etc, the solver that will be presented in Section 8 reduces the problem to propositional logic by assuming bounded time. In essence, the reduction simulates the truth value of an LTL propositional variable p along time by a finite set of n fresh atoms, one per time instant. Moreover, the behaviour of the “always” temporal operator is approximated by use of finite conjunctions. Exogenous variables are not grounded, since it is useless and expensive to consider different variables in this case.

In detail, the grounding to a given time $k \in \mathbb{N}$ of a propositional formula φ built from a set of atoms partitioned into exogenous and endogenous is first of all defined.

Definition 9 (Grounding of propositional formulae). *Let φ be a propositional formula built from the set of atoms $Ex \cup Ed$. The grounding of φ to time k , $\langle \varphi \rangle_k$, is defined as follows:*

- if $p \in Ex$, then $\langle p \rangle_k \stackrel{\text{def}}{=} p$;
- if $p \in Ed$, then $\langle p \rangle_k \stackrel{\text{def}}{=} p_k$, where p_k is a new propositional variable;
- $\langle \neg \varphi \rangle_k \stackrel{\text{def}}{=} \neg \langle \varphi \rangle_k$;
- $\langle \varphi \vee \psi \rangle_k \stackrel{\text{def}}{=} \langle \varphi \rangle_k \vee \langle \psi \rangle_k$.

If S is a set of propositional formulae, then $\langle S \rangle_k = \{ \langle \varphi \rangle_k \mid \varphi \in S \}$.

Next, the grounding of the encoding of a MIG is defined.

⁴Abduction is, in general, the reasoning aiming at explaining some observation O which is not logically implied by the background theory T , i.e., finding a set of formulae E such that $T \cup E \models O$. “Interesting” explanations are those which additionally satisfy some other requirements. In the present context, abductive reasoning is meant to look for sets of literals (atoms and negated atoms) as explanations, satisfying a minimality requirement. For instance, if $T = \emptyset$ and $O = \{p \vee q\}$, the sets $\{p\}$ and $\{q\}$ are the only minimal explanations for O . Other explanations, $\{p, q\}$, $\{p, \neg q\}$ and $\{q, \neg p\}$ are not minimal.

⁵A method to perform abduction for a fragment of LTL sufficient to represent problems on MIMs has been proposed in [9], but it has not been implemented.

Definition 10 (Grounding of the encoding of a MIG). *Let $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ be a MIG, S its LTL encoding and $k \in \mathbb{N}$.*

For all $p \in Ed$, if SSA_p is the formula $\Box(\bigcirc p \leftrightarrow \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p)))$ belonging to S , we define

$$\langle SSA_p \rangle_k = p_{k+1} \leftrightarrow \langle \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p)) \rangle_k$$

The grounding $\langle S \rangle_k$ of S up to time k is defined as follows:

$$\langle S \rangle_k = \{ \langle \ell \rangle_0 \mid \ell \in \mathcal{B} \} \cup \{ \langle SSA_p \rangle_i \mid p \in Ed \text{ and } 0 \leq i < k \}$$

The grounding $\langle SSA_p \rangle_k$ is well defined, since $\mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p))$ is a classical formula. Note that “successor state axioms” SSA_p in the LTL encoding of \mathcal{G} are grounded only for endogenous variables and only as far as the “ $\bigcirc p$ ” refers to a state that “exists” in the bounded timed model.

The next definition formalizes the notion of a temporal interpretation T and a classical one M being models of the same initial state.

Definition 11. *Let $At = Ex \cup Ed$ be a set of atoms, $T = T_0, T_1, \dots$ an LTL interpretation of the language At and $k \in \mathbb{N}$. A classical interpretation M is said to correspond to T up to time limit k if M is an interpretation of the language $Ex \cup \{p_i \mid p \in Ed \text{ and } 0 \leq i \leq k\}$ and for all $p \in At$, $M \models \langle p \rangle_0$ iff $T_0 \models p$.*

The next result establishes a kind of “model correspondence” property.

Theorem 2 (Model correspondence). *Let $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ be a MIG, S its LTL encoding, and $\langle S \rangle_n$ the grounding of S up to time n . If $T = T_0, T_1, \dots$ is any model of S and M a model of $\langle S \rangle_n$ corresponding to T , then for every classical propositional formula φ and every $k = 0, \dots, n$: $M \models \langle \varphi \rangle_k$ iff $T_k \models \varphi$.*

Proof. By double induction on k and φ .

1. If $k = 0$, the thesis is proved by induction on φ .
 - (a) If φ is an atom, then the thesis follows immediately from the fact that M corresponds to T .
 - (b) If $\varphi = \neg \varphi_0$ or $\varphi = \varphi_0 \vee \varphi_1$, the thesis follows from the induction hypothesis, the definition of $\langle \varphi \rangle_k$ (Definition 9) and the definition of \models for classical logic.
2. $0 < k \leq n$: By the induction hypothesis $T_{k-1} \models \varphi$ iff $M \models \langle \varphi \rangle_{k-1}$ for every propositional formula φ . The thesis is proved by induction on φ :

(a) If φ is an atom, we consider two cases:

- i. $p \in Ex$: since $T_{k-1} \models \bigcirc p \leftrightarrow p$, then $T_k \models p$ iff $T_{k-1} \models p$. By the induction hypothesis, $T_{k-1} \models p$ iff $M \models \langle p \rangle_{k-1}$. Since $\langle p \rangle_{k-1} = p = \langle p \rangle_k$, $T_k \models p$ iff $M \models \langle p \rangle_k$.
- ii. $p \in Ed$: since $T_{k-1} \models \bigcirc p \leftrightarrow \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p))$, $T_k \models p$ iff $T_{k-1} \models \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p))$. By the induction hypothesis, the latter assertion holds iff $M \models \langle \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p)) \rangle_{k-1}$. By Definition 10, $\langle S \rangle_n$ contains $p_k \leftrightarrow \langle \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p)) \rangle_{k-1}$, and, since $M \models \langle S \rangle_n$, $M \models \langle \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p)) \rangle_{k-1}$ iff $M \models p_k$. Therefore, $T_k \models p$ iff $M \models p_k$.

(b) If $\varphi = \neg\varphi_0$ or $\varphi = \varphi_0 \vee \varphi_1$, the thesis follows from the induction hypothesis, Definition 9 and the definition of \models for classical logic, like in the base case.

□

The rest of this section is devoted to establish the complexity of grounding for the encoding of a MIG. Let the size of a formula be measured in terms of the number of its logical operators: if φ is a formula, $\|\varphi\|$ is the number of logical operators in φ . If S is a set of formulae, then $\|S\| = \sum_{\varphi \in S} \|\varphi\|$.

Theorem 3 (Complexity of the encoding). *Let \mathcal{G} be a MIG, S its LTL encoding and $\langle S \rangle_n$ the grounding of S up to time n . Then $\|\langle S \rangle_n\| \leq n \times \|S\|$.*

Proof. First of all we note that if φ is a classical formula, then $\|\varphi\| = \|\langle \varphi \rangle_k\|$ for any k . Consequently,

$$\|p_k \leftrightarrow \langle \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p)) \rangle_{k-1}\| = \|\bigcirc p \leftrightarrow \mathbf{Pr}(p) \vee (p \wedge \neg \mathbf{Cn}(p))\| - 1$$

and $\|\langle SSA_p \rangle_k\| = \|SSA_p\| - 2$.

Let S be the LTL encoding of a MIG $\mathcal{G} = \langle Ex, Ed, \mathcal{P}, \mathcal{C}, \mathcal{A}, \mathcal{I}, \mathcal{B} \rangle$ and $\langle S \rangle_n$ its grounding up to time n .

1. For each $\langle \ell \rangle_0 \in \langle S \rangle_n$ such that $\ell \in \mathcal{B}$, $\|\langle \ell \rangle_0\| = \|\ell\|$. Therefore $\|\langle \mathcal{B} \rangle_0\| = \|\mathcal{B}\|$.
2. Beyond the literals in $\langle \mathcal{B} \rangle_0$, $\langle S \rangle_n$ contains $\langle SSA_p \rangle_k$ for all $p \in Ed$ and $0 \leq k < n$. Hence, for every $SSA_p \in S$, $\langle S \rangle_n$ contains $n - 1$ formulae, the size of each of them being smaller than the size of SSA_p . Therefore

$$\|\{\langle SSA_p \rangle_k \mid p \in Ed \text{ and } 0 \leq k < n\}\| < n \times \|\{SSA_p \mid p \in Ed\}\|$$

Therefore, $\|\langle S \rangle_n\| \leq n \times \|S\|$. □

It is worth pointing out that exogenous variables are not grounded. Consequently, for instance, if *Lactose* is assumed to be exogenous, the grounding up to time k of the LTL formula $\Box(\bigcirc Glucose \leftrightarrow (Lactose \wedge Galactosidase) \vee Glucose)$ is the conjunction of all the formulae of the form $Glucose_{i+1} \leftrightarrow (Lactose \wedge Galactosidase_i) \vee Glucose_i$ for $0 \leq i < k$.

8 The P3M tool: a software platform for modelling and manipulating MIMs

In this section we present P3M (**P**latform for **M**anipulating **M**olecular **I**nteraction **M**aps), a prototypal system implementing the representation mechanism outlined in the previous sections and able to solve the following problems, that will be discussed further on: graph validation, graph querying and graph updating. The system is written in Objective Caml [29], and interfaces with the C implementation of the PicoSAT solver library [5]. A graphical user interface has been developed to help biologists to interact with the system in a user-friendly way. The general architecture of the system is represented in Figure 5, and will be further explained below. P3M can be downloaded at <http://www.alliot.fr/P3M/>.

8.1 Setting of types and values of variables

The system takes as input files representing MIMs as created by PathVisio⁶, a free open-source biological pathway analysis software tool that allows one to draw biological pathways. The graph is displayed to the user, using colors and typefaces to distinguish the types and initial values of atoms, which are given a default value by the software tool based on “commonsense” rules: entities that can be produced by some reaction are set as endogenous and initialized as absent, the others (those with no incoming edge) are considered as exogenous and initially unset (i.e., free atoms). Figure 6 shows how the software has set the variable types: *lacl*, *lacZ*, *CAMP* and *Lactose* are in *bold* typeface, as they are set as *exogenous* variables, *glucose*, *galactosidase* and *repressor boxes* are in *normal* typeface, as they are *endogenous*.

The initial values of variables are shown by use of different colors: by default, the initial values of all variables are *unset* and their names are shown in black. We recall that atoms whose initial value is not set are called *free*. The user is allowed to change both types and initial values of atoms. Figure 7 shows the graph when the

⁶<https://github.com/PathVisio/pathvisio>

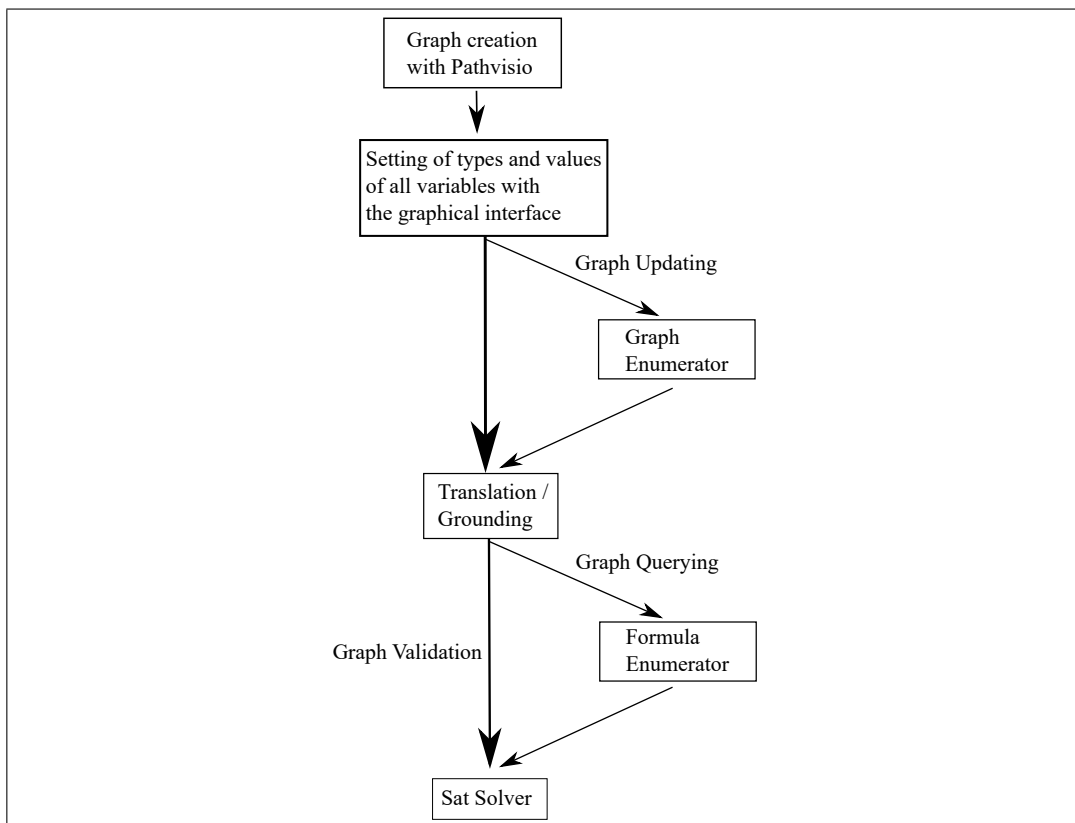


Figure 5: Implementation

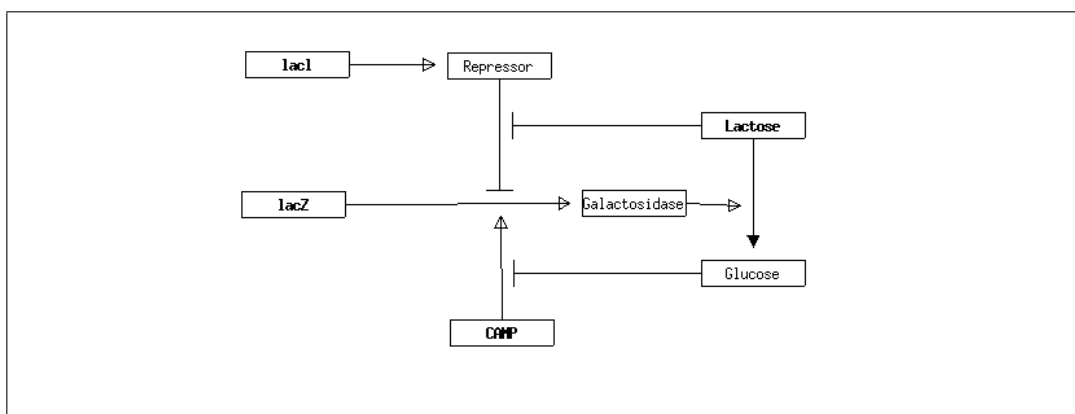


Figure 6: The lac operon after the default initialization of variables types and values

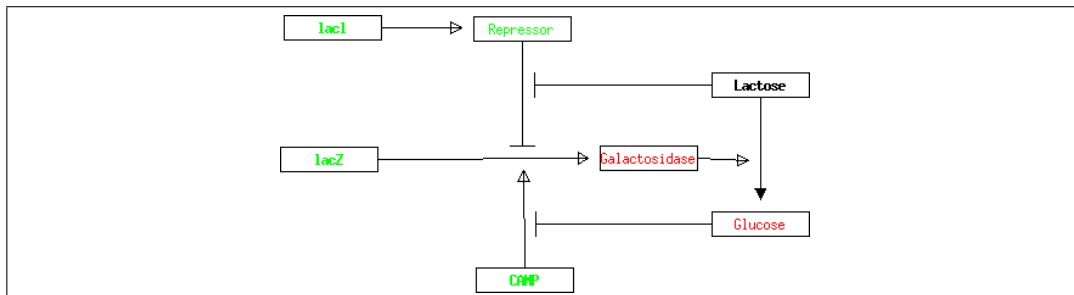


Figure 7: Lac operon after the modifications

user has modified the values of some variables: `lacI`, `lacZ` and `CAMP` are *green*, to indicate that they are *present* at the start of the process (they will remain present since they are exogenous atoms). `Repressor` is *green*, as the repressor protein is supposed to be in the cell at the start of the process. `Lactose` remains *black* since it is a free atom, about which the user is going to query the system. Initially absent variables (`Galactosidase` and `Glucose`) are shown in *red*.

Other parameters, such as the number of time steps, the number of modifications to make for graph updating, queries etc. are set via the command line.

8.2 Resolution engine

The resolution engine is able to perform the following reasoning tasks.

Graph validation. This task consists in checking whether the graph \mathcal{G} is consistent. The temporal encoding of \mathcal{G} is grounded to the specified time and the SAT solver PicoSAT is used in a straightforward way in order to check the consistence of the grounded theory.

Graph querying. This task consists in finding which initial values of the free atoms make \mathcal{G} satisfy some temporal property φ . It is an abductive reasoning task [23], that could be solved by use of classical algorithms for computing prime implicants. But we have checked that, for instance, the Kean and Tsiknis algorithm [25] results to be very slow even when the total number of atoms is small. However, biologists are usually only interested in the values of the free atoms. Since their number is often quite small, it is usually faster to use PicoSAT to solve iteratively all possible models. In other terms, all the possible combinations of initial values for free atoms are generated (by the *formula enumerator* of Figure 5) and the SAT solver is run on each of the so-obtained initial conditions. The system, tested on graphs with up to 22 nodes and 41

relations, showed to be effective up to roughly 16 to 20 free atoms depending on the complexity of the map.

In performing this task, exogenous and endogenous atoms can be treated differently: the user can either ask which values of all the free variables imply the given property, or else to find out which values of the free exogenous atoms guarantee that *for all values of the free endogenous ones* the query holds at the given time.

Graph updating. Given a graph \mathcal{G} for which a given property φ does not hold, this task consists in turning \mathcal{G} into a new graph \mathcal{G}' satisfying φ . This is the most complex task, since there might be a very large number of possible graphs solving the problem. Currently, the system computes all graphs \mathcal{G}' that can be obtained from \mathcal{G} by adding, removing or modifying a single relation (this step is called the *graph enumerator* in Figure 5). Then for each \mathcal{G}' , graph querying on \mathcal{G}' and φ is performed, in order to filter out the graphs which do not satisfy φ .

9 Examples

The software tool has been tested on graphs with up to 20 atoms, 22 nodes and 41 links. In this section we show some examples of the two most complex tasks: graph querying and graph updating.

9.1 Graph querying

A more complex example will be considered here, i.e., a meaningful part of the map presented in Figure 1, the *atm-chk2* metabolic pathway, which leads to cellular apoptosis when the DNA double strand breaks. DNA double strand break (*dsb*) is a major cause of cancers, and medical and pharmaceutical research [26, 21] has shown that *dsb* can occur in a cell as the result of a pathology in a metabolic pathway. This kind of map is used to find the molecular determinants of tumoral response to cancers. Molecular parameters included the metabolic pathways for repairing DNA, the metabolic pathways for apoptosis, and the metabolic pathways of cellular cycle control [33, 26, 21, 28, 31]. When DNA is damaged, cellular cycle control points are activated and can quickly kill the cell by apoptosis, or stop the cellular cycle to enable DNA repair before reproduction of cellular division. Two of these control points are the metabolic pathways *atm-chk2* and *atr-chk2* [33].

The graph of Figure 8 (built from the map in Figure 1) represents the metabolic pathway *atm-chk2* which can lead to apoptosis in three different ways. This map

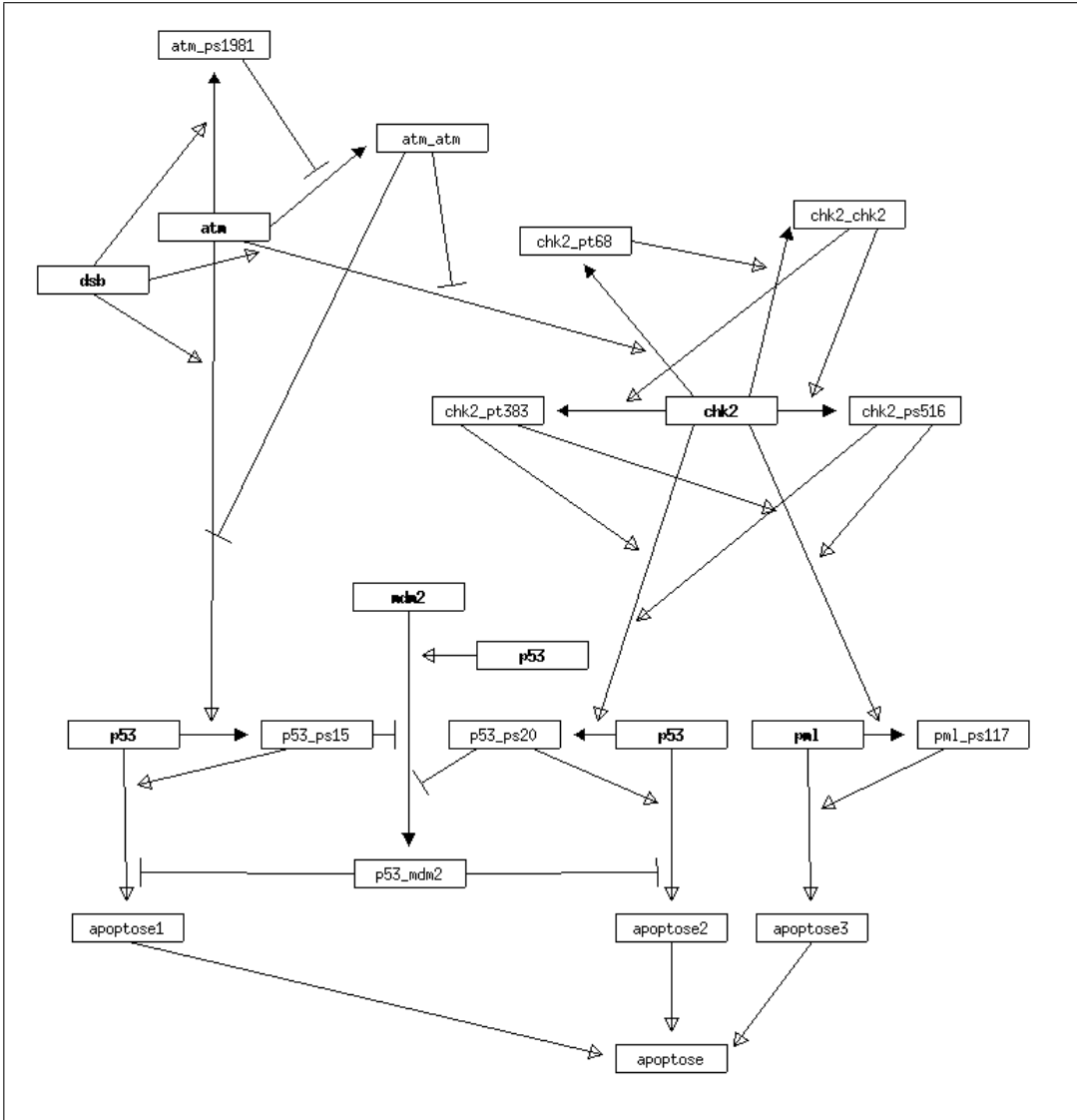


Figure 8: The Molecular Interaction Map *atm-chk2*

involves 20 variables, six of which (*atm*, *dsb*, *chk2*, *mdm2*, *pml* and *p53*) are exogenous and the rest endogenous. Some of these variables are proteins, others, such as *dsb* or *apoptose*, representing cell death, are conditions or states.

The time required for solving graph querying problems depends on the number of free variables and time steps. The P3M solver has been called on this graph to find

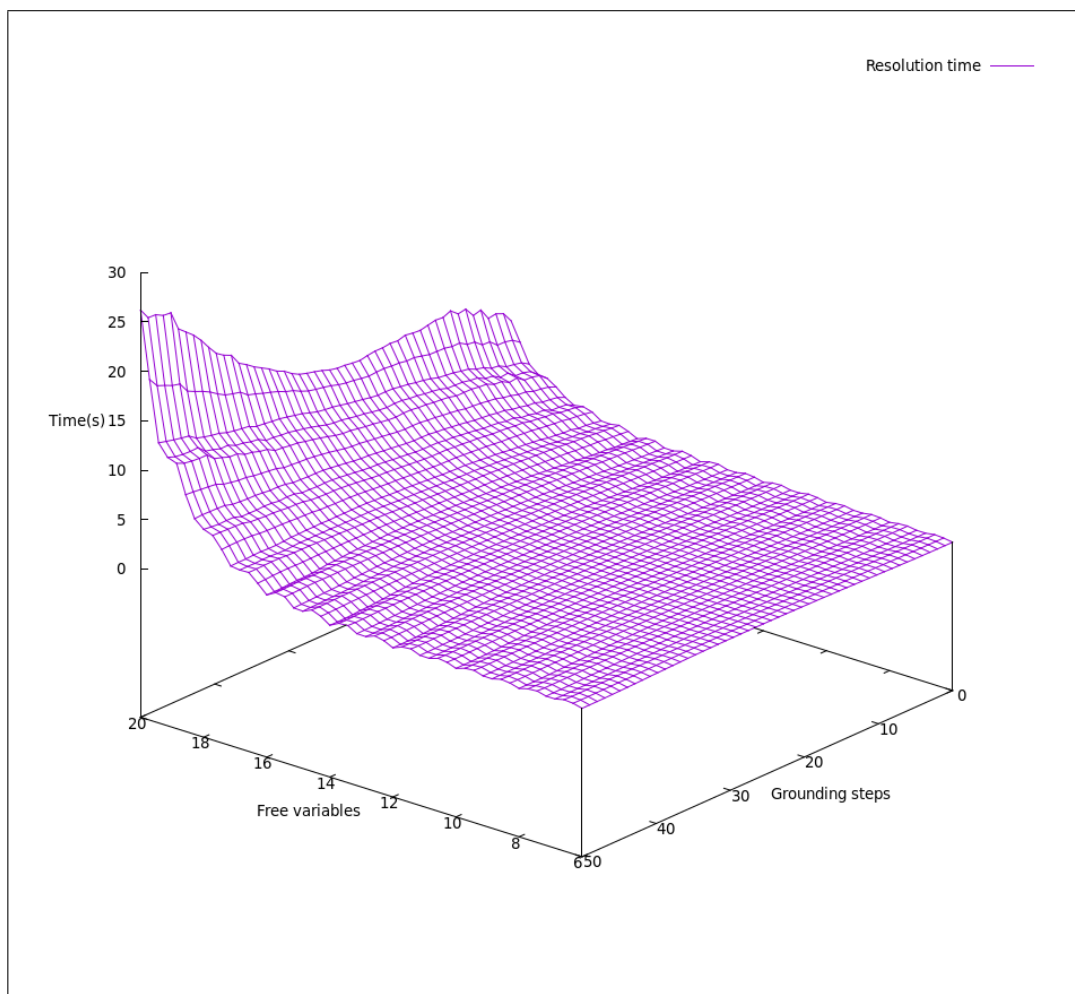


Figure 9: Time as a function of the number of grounding steps and free variables

out what would cause the cell apoptosis. It has been tested with different grounding values g , ranging from 1 to 50, and queries to find out the initial conditions that make the atom *apoptose*(g) derivable, i.e., the conditions causing cell apoptosis at time g . The system has been tested with a number of free variables ranging from 6 (only exogenous variables are free) to 20 (all variables are set to free, thus asking the system to find also their initial values).

The 3D diagram in Figure 9 plots the grounding values and the number of free variables against the time taken by the system to solve the problem, by calling PicoSAT (the time taken to encode the graph into propositional logic is negligible).

From the diagram, it is clear that the number of free variables is the bottleneck, as it was actually expected since the time required to solve the problem is exponential in the number of free variables. Moreover, 50 time steps are overkill, most systems reaching a stable state in less than 10 time steps.

The questions asked to the system can be refined, in order to find out, for instance, how much time is required to reach apoptosis on each of the three possible ways, and which are the initial conditions which lead to each of them. The questions to ask are *apoptose1(i)*, *apoptose2(i)* and *apoptose3(i)*, for different values of *i*, where a query of the form *p(i)* means that one looks for an explanation of *p* being true at time step *i*. The answers given by the system show that:

- *apoptose1* can be obtained in the fastest way: *apoptose1(2)* (*apoptose1* holding at the second time step) is true if *atm*, *dsb* and *p53* are present, and *mdm2* is absent (the values of *pml* and *chk2* do not matter). For $i \geq 3$, the answer to *apoptose1(i)* is the same, but *mdm2* does not matter any longer (*p53_mdm2* is dissociated at step 2).
- obtaining *apoptose2* requires 5 time steps; *atm*, *chk2*, *dsb*, *p53* have to be present, and *mdm2* and *pml* do not matter.
- *apoptose3* requires the same number of steps as *apoptose2* but the initial conditions are different: *atm*, *chk2*, *dsb*, and *pml* have to be present, while *mdm2* and *p53* do not matter.

9.2 Graph updating

Figure 10 shows the map of the lac operon where the inhibition of lactose on the negative regulation of the repressor to the production of galactosidase has been suppressed. So here, glucose is not produced anymore when lactose is present. The user can ask the system what modifications could be done in order to produce glucose when lactose is present. The “correct” solution is found immediately (Figure 11), along with others. Some of these other generated solutions are not interesting, such as the direct production of glucose by genes *lacZ* or *lacI*. But the system also proposes reasonable solutions, such as that shown in Figure 12, where glucose is used to provide the inhibiting action for the repressor protein. When glucose is present, the production of galactosidase is stopped, while galactosidase is produced when glucose is absent. However nature has chosen the more economical solution, because here galactosidase would be produced as soon as glucose is absent, which is useless if there is no lactose.

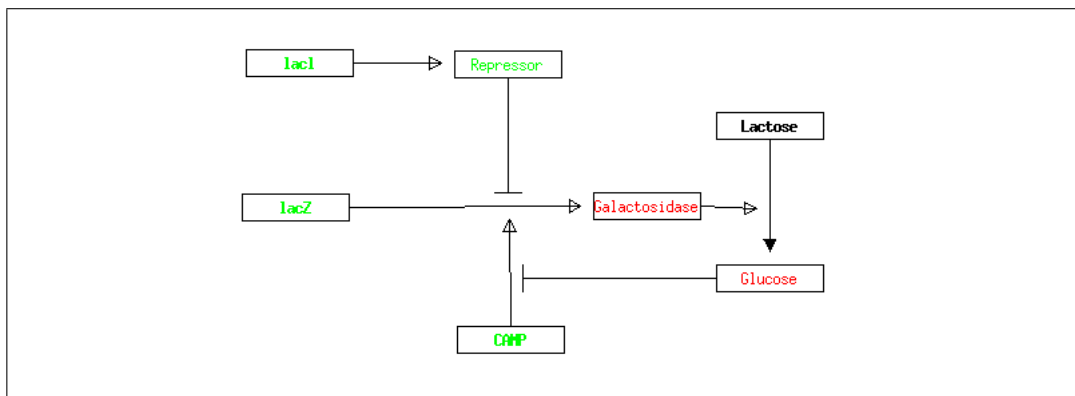


Figure 10: Lac operon without inhibition by lactose

Finding which solution is correct can only be done by biologists; it is sometimes trivial for them, but sometimes different solutions proposed by the system are plausible and have to be further tested with real biological experiments.

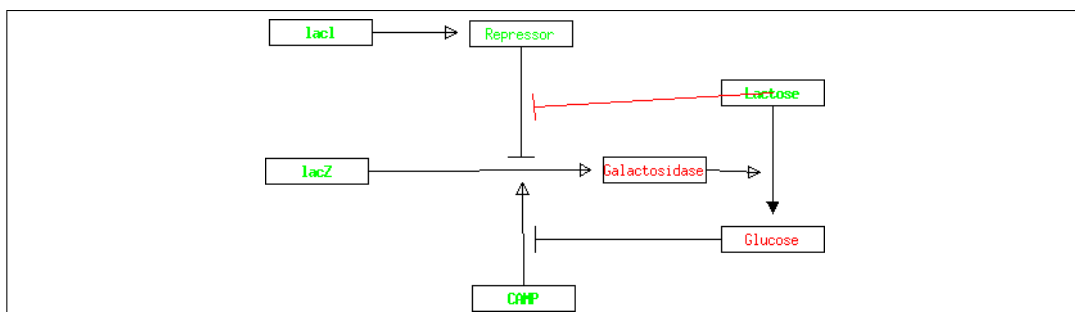


Figure 11: Correct solution

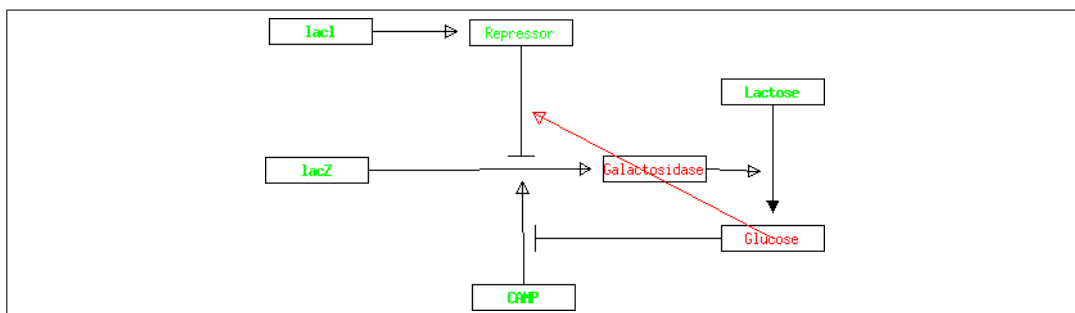


Figure 12: Another interesting solution

10 Conclusion

This paper presents a method to translate MIMs, representing biological systems, into Linear Temporal Logic, and a software tool able to solve complex questions about these graphs. The system, though still a prototype, is able to solve quite realistic examples of a large size.

The proposed approach can be improved in different directions. On the theoretical side, it is worth remarking that the speed of reactions is not taken into account. This limitation could be overcome by using the dual of speed (duration) and by using a logic that represents the duration of reactions. Moreover, the system relies on the “all or nothing” hypothesis: we do not represent quantities other than “absent” or “present”. As a consequence, all productions that are enabled at a given time are fired simultaneously, since they do not compete on the use of resources. We have been able to efficiently model complex graphs with this constraint, but an important step forward to be planned is modelling a more realistic evolution of networks by taking quantities into account.

On the practical point of view, the possibility should be explored to avoid grounding and replacing the *formula enumerator* procedure of P3M by implementing a direct abduction algorithm for (a suitable fragment of) LTL, as proposed in [9], or else by directly using temporal model checkers [13], or tools like RECAR (Recursive Explore and Check Abstraction Refinement) [27] which allows one to solve modal satisfiability problems.

Moreover, the software tool can be improved in several respects like, for instance, improving the graphical interface by enriching the number of parameters the user can choose and making it more user friendly.

Acknowledgments

The authors wish to thank Gilles Favre, Jean-Charles Faye and Olivier Sordet of the Centre de Recherches en Cancérologie de Toulouse (CRCT): without their precious collaboration this interdisciplinary project would not have been possible.

References

- [1] Jean-Marc Alliot, Robert Demolombe, Martín Diéguez, Luis Fariñas del Cerro, Gilles Favre, Jean-Charles Faye, Naji Obeid, and Olivier Sordet. Temporal logic modeling of biological systems. In *Towards Paraconsistent Engineering*, pages 205–226. Springer International Publishing, 2016.
- [2] Jean-Marc Alliot, Robert Demolombe, Luis Fariñas del Cerro, Martín Diéguez, and Naji Obeid. Abductive reasoning on molecular interaction maps. In *Interactions Be-*

- tween *Computational Intelligence and Mathematics*, pages 43–56. Springer International Publishing, 2018.
- [3] Jean-Marc Alliot, Martín Diéguez, and Luis Fariñas del Cerro. Metabolic pathways as temporal logic programs. In Loizos Michael and Antonis Kakas, editors, *Logics in Artificial Intelligence*, pages 3–17. Springer International Publishing, 2016.
 - [4] Grégory Batt, Delphine Ropers, Hidde de Jong, Johannes Geiselmann, Radu Mateescu, Michel Page, and Dominique Schneider. Validation of qualitative models of genetic regulatory networks by model checking: analysis of the nutritional stress response in *Escherichia coli*. In *Proceedings Thirteenth International Conference on Intelligent Systems for Molecular Biology 2005, Detroit, MI, USA, 25-29 June 2005*, pages 19–28, 2005.
 - [5] Armin Biere. Picosat essentials. *Journal on Satisfiability, Boolean Modeling and Computation (JSAT)*, 4:75–97, 2008.
 - [6] D. Bošnački, P.A.J. Hilbers, R.S. Mans, and E.P. de Vink. Chapter 39: Modeling and analysis of biological networks with model checking. In M. Elloumi and A.Y. Zomaya, editors, *Algorithms in Computational Molecular Biology: Techniques, Approaches and Applications*, volume 1 of *Wiley Series in Bioinformatics*, pages 915–940. Wiley, 2011.
 - [7] Luboš Brim, Milan Češka, and David Šafránek. Model Checking of Biological Systems. In *Formal Methods for Dynamical Systems: 13th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2013, Bertinoro, Italy, June 17-22, 2013. Advanced Lectures*, pages 63–112. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
 - [8] Laurence Calzone, François Fages, and Sylvain Soliman. BIOCHAM: an environment for modeling biological systems and formalizing experimental knowledge. *Bioinformatics*, 22(14):1805–1807, 2006.
 - [9] Serenella Cerrito, Marta Cialdea Mayer, and Robert Demolombe. Temporal abductive reasoning about biochemical reactions. *Journal of Applied Non-Classical Logics*, 27(3-4):269–291, 2017.
 - [10] Nathalie Chabrier-Rivier, Marc Chiaverini, Vincent Danos, François Fages, and Vincent Schächter. Modeling and querying biomolecular interaction networks. *Theoretical Computer Science*, 325(1):25–44, 2004.
 - [11] Nathalie Chabrier-Rivier, Francois Fages, and Sylvain Soliman. The Biochemical Abstract Machine BIOCHAM. In Vincent Danos and Vincent Schächter, editors, *CMSB’04: Proceedings of the second Workshop on Computational Methods in Systems Biology*, volume 3082, pages 172–191, Paris, 2004. Springer-Verlag.
 - [12] Federica Ciocchetta and Jane Hillston. Bio-pepa: An extension of the process algebra pepa for biochemical networks. *Electron. Notes Theor. Comput. Sci.*, 194(3):103–117, 2008.
 - [13] Edmund Clarke, Orna Grumberg, Somesh Jha, Yuan Lu, and Helmut Veith. Counterexample-guided abstraction refinement for symbolic model checking. *Journal of the ACM*, 50(5):752–794, 2003.
 - [14] R. Demolombe, L. Fariñas del Cerro, and N. Obeid. Automated reasoning in metabolic

- networks with inhibition. In *13th International Conference of the Italian Association for Artificial Intelligence, (AI*IA'13)*, pages 37–47, Turin, Italy, 2013.
- [15] R. Demolombe, L. Fariñas del Cerro, and N. Obeid. Translation of first order formulas into ground formulas via a completion theory. *Journal of Applied Logic*, 15:130–149, 2016.
- [16] Robert Demolombe, Luis Fariñas del Cerro, and Naji Obeid. A logical model for molecular interaction maps. In Fariñas and Inoue [19], chapter 3, pages 93–123.
- [17] Francois Fages and Sylvain Soliman. Formal Cell Biology in Biocham. In *Formal Methods for Computational Systems Biology: 8th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2008 Bertinoro, Italy, June 2-7, 2008 Advanced Lectures*, pages 54–80. Springer Verlag, Berlin, Heidelberg, 2008.
- [18] Francois Fages, Sylvain Soliman, and Nathalie Chabrier-rivier. Modelling and querying interaction networks in the biochemical abstract machine biocham. *Journal of Biological Physics and Chemistry*, 4:64–73, 2004.
- [19] L. Fariñas and K. Inoue, editors. *Logical Modeling of Biological Systems*. John Wiley & Sons, 2014.
- [20] Fisher Jasmin and Henzinger Thomas A. Executable cell biology. *Nat Biotech*, 25(11):1239–1249, nov 2007.
- [21] V. Glorian, G. Maillot, S. Poles, J. S. Iacovoni, G. Favre, and S. Vagner. HuR-dependent loading of miRNA RISC to the mRNA encoding the Ras-related small GTPase RhoB controls its translation during UV-induced apoptosis. *CCell Death and Differentiation*, 18(11):1692–1701, 2011.
- [22] M. Hucka, H. Bolouri, A. Finney, H. M. Sauro, J. C. Doyle, and H. Kitano. The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19:524–531, 2003.
- [23] Katsumi Inoue. Linear resolution for consequence finding. *Artificial Intelligence*, 56(2):301 – 353, 1992.
- [24] F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3:318–356, 1961.
- [25] A. Kean and G. Tsiknis. An incremental method for generating prime implicants/implicates. *Journal of Symbolic Computing*, 9:185–206, 1990.
- [26] K. W. Kohn and Y. Pommier. Molecular interaction map of the p53 and Mdm2 logic elements, which control the off-on switch of p53 response to DNA damage. *Biochemical and Biophysical Research Communications*, 331(3):816–827, 2005.
- [27] Jean-Marie Lagniez, Daniel Le Berre, Tiago de Lima, and Valentin Montmirail. A recursive shortcut for CEGAR: Application to the modal logic K satisfiability problem. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)*, pages 674–680, 2017.
- [28] W. J. Lee, D. U. Kim, M. Y. Lee, and K. Y. Choi. Identification of proteins interacting with the catalytic subunit of PP2A by proteomics. *Proteomics*, 7(2):206–214, 2007.

- [29] Xavier Leroy, Damien Doligez, Alain Frisch, Jacques Garrigue, Didier Rémy, and Jérôme Vouillon. *The OCaml System, Documentation and user's manual*. Institut National de Recherche en Informatique et en Automatique, 2017.
- [30] Alida Palmisano and Corrado Priami. Bio-PEPA. In *Encyclopedia of Systems Biology*, pages 145–146. Springer New York, New York, NY, 2013.
- [31] H. Pei, L. Zhang, K. Luo, Y. Qin, M. Chesi, F. Fei, P. L. Bergsagel, Wang L., Z. You, and Z. Lou. MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. *Nature*, 470(7332):124–128, 2011.
- [32] A. Pnueli. The temporal logic of programs. In *Proc. of the 18th Annual Symposium on Foundations of Computer Science*, pages 46–57, Providence, Rhode Island, USA, 1977.
- [33] Y. Pommier, O. Sordet, V. A. Rao, H. Zhang, and K.W. Kohn. Targeting chk2 kinase: molecular interaction maps and therapeutic rationale. *Current Pharmaceutical Design*, 11(22):2855–2872, 2005.
- [34] Raymond Reiter. *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems*. MIT Press, 2001.
- [35] Mara Sangiovanni. *Model Checking of Metabolic Networks: Application to Metabolic Diseases*. PhD thesis, Federico II University of Naples., 2014.
- [36] Jetse Scholma, Stefano Schivo, Ricardo A. Urquidi Camacho, Jaco van de Pol, Marcel Karperien, and Janine N. Post. Biological networks 101: Computational modeling for molecular biologists. *Gene*, 533(1):379–384, 2014.
- [37] Wikipedia. The lac operon. https://en.wikipedia.org/wiki/Lac_operon, 2015.

VARIATION AND PATTERN IN THE 3-TERM SYLLOGISMS AND THE BILITERAL SORITES OF TYPES BARBARA, DARAPTI AND DARI

DAN CONSTANTIN RADULESCU*

dancradulescu@yahoo.com

Abstract

One examines only the fundamental patterns of the valid syllogisms and of the biliteral sorites of Barbara, Darapti and Darii types. Instead of writing the general premises of a Barbara type sorites corresponding to a particular ordering of the terms appearing in the sorites as $M_1^* \subseteq M_2^* \subseteq M_3^* \subseteq \dots \subseteq M_n^*$, where $M_i^* \in \{M_i, M'_i\}$, $i = 1, 2, \dots, n$, and examining $n!2^n$ possible Barbara sorites, one will examine only Barbara's pattern $M_1 \subseteq M_2 \subseteq M_3 \subseteq \dots \subseteq M_n$, equivalent to the premises $M_1 \cap M'_2 = \emptyset, M_2 \cap M'_3 = \emptyset, \dots, M_{n-1} \cap M'_n = \emptyset$, where M'_i denotes the complement of M_i in the universe of discourse, U . Another reason of examining only the "pattern Barbara" sorites, is that any different Barbara type sorites being true, simultaneously with the "pattern Barbara", would impose conditions on U . For $n = 3$, if the "pattern Barbara", $S \subseteq M \subseteq P$, is valid, then none of the other seven Barbara type syllogisms could be valid without imposing conditions on U : if both Celarent, $S \subseteq M \subseteq P'$, and Barbara are simultaneously valid, then $S = M = \emptyset$. If Camestres, $S \subseteq M' \subseteq P'$, and Barbara are simultaneously valid, then $P = M$ and $S = \emptyset$. Still preserving their content, all the syllogisms or sorites of Barbara type can be re-written as the Barbara pattern syllogism or sorites. A unique partitioning subset of U characterizes the logical consequence of each of the Barbara, Darapti or Darii sorites. One argues that all the valid syllogisms are equivalent to either Barbara, Darapti or Darii, and that the moods and figures of Classical Syllogistic can be viewed as content variations of the Barbara, Darapti or Darii patterns resulting from various term substitutions in these three patterns.

*In memory of my mathematician sister Cristina Popa, 1948-2018.

1 Preliminaries

One uses the following notations and abbreviations: U for the universe of discourse made of 2^n partitioning subsets corresponding to the n terms of the sorites; M' for the complement in U of a set M . A “partitioning subset of U ” will be shortened to *subset of U* : no other set will be called a subset except a partitioning subset of U . (Boole,[1], calls the partitioning subsets “constituents”.) Juxtaposition of set names will denote set intersections: for example, M_1M_2 denotes the intersection, $M_1 \cap M_2$, of the sets M_1 and M_2 . *PCP* will stand for pair of categorical premises and *LC* for *Logical Consequence*. *Existential Import* will be shortened to *ei* and *Classical Syllogistic* to *CS*. A categorical statement connects two syllogistic terms via one of the quantifiers A (All α is ω), E (No α is ω), I (Some α is ω), O (Some α is not ω); α is called the subject of the statement and ω is called its predicate. Since George Boole [2] it is customary to interpret each term as being a class (or set of elements). If the term appears in a universal statement, A or E , the term might be, possibly, an empty set. For example, if, in All α is ω , the term ω is empty, then α is empty, too. In a particular statement, I or O , both terms are non-empty. Note that not ω means ω' — the complement of ω in the universal set U ; thus, $U = \alpha \cup \alpha' = \omega \cup \omega'$. (\cup is the union symbol; for convenience, it will be replaced by a $+$ sign). A universal set with two terms is thus partitioned into four (non-intersecting) subsets: $U = (\alpha \cap \omega) \cup (\alpha' \cap \omega) \cup (\alpha \cap \omega') \cup (\alpha' \cap \omega')$, which, after replacing the intersection symbol by a juxtaposition, will be written as $U = \alpha\omega + \alpha'\omega + \alpha\omega' + \alpha'\omega'$. The not α , or non α , i.e., α' , and non ω , i.e., ω' , are called *negative terms*, and α and ω are called *positive terms*. A syllogism contains three categorical statements — two premises and their proposed logical consequence (LC) or conclusion. Each of the two premises contains the middle term, most often denoted by M , and two other terms, P and S which will appear again in the LC. As above, the S, P, M terms are called *positive terms* and their complementary sets in U — S', P', M' , are the *negative terms*; together, S, P, M, S', P', M' , are the *indefinite terms*. *Conclusive syllogism*, (or *conclusive PCP*), stands for any PCP which entails a logical consequence (LC) of any of the eight formats $E(S^*, P^*), I(S^*, P^*)$, where $P^* \in \{P, P'\}, S^* \in \{S, S'\}$. One also defines $M^* \in \{M, M'\}$ and, from now on, one uses the P^*, S^*, M^* notations without repeating each time their definitions. For valid syllogisms, the Classical Syllogistic (CS) accepts, by definition, only these “standard” LC formats: $A(S, P), E(S, P), I(S, P), O(S, P)$.

2 Introduction

Since Aristotle created it, for more than 2300 years, the Classical Syllogistic developed without the benefit of newer mathematics. We'll use matrices and set intersections and inclusions to simplify the presentation of Classical Syllogistic — which, at the end of the 19th century, was described by Lewis Carroll [3] as follows: "...the ordinary textbooks of Formal Logic have elaborately discussed no less than nineteen different forms of Syllogisms — each with its own special and exasperating Rules, while the whole constitutes an almost useless machine, for practical purposes, many of the Conclusions being incomplete, and many quite legitimate forms being ignored", "As to syllogisms, I find that their nineteen forms, with about a score of others which the textbooks have ignored, can all be arranged under three forms, each with a very simple Rule of its own". "The theory that two negative premises prove nothing" is "another craze of 'The Logicians', fully as morbid as their dread of a negative Attribute". (Carroll refers to the "dread" of the S' , P' , M' terms, and to the Rule of Valid Syllogism which states that "Two negative premises are not allowed".) Later on, in 1952, Hans Reichenbach [12] similarly opines that "The usual exposition of the theory of the syllogism, however, whether given by the use of the familiar rules of the syllogism, or by the help of diagrams, appears clumsy and lacks the lucidity of modern chapters of logic." This paper will not discuss the Rules of Valid Syllogism.

In short, Classical Syllogistic uses premises formulable only via positive terms, uses syllogistic figures, and accepts as generating valid syllogisms only those PCPs which entail a logical consequence (LC) of one of the formats $A(S, P)$, $E(S, P)$, $I(S, P)$, $O(S, P)$. According to Burris [4], George Boole had two objections to the formulation of Classical Syllogistic: "Regarding syllogisms Boole did not care for the Aristotelian classification into Figures and Moods as it seemed rather arbitrary. In particular Boole did not like the requirement that the predicate of the conclusion had to be the major term in the premises." This last requirement which Boole did not like, is equivalent to accepting only the statements $A(S, P)$, $E(S, P)$, $I(S, P)$, $O(S, P)$ as syllogisms' LCs. All the above constraints are respected in the presentations of Classical Syllogistic found in introductory logic books. The many modern papers which address Classical Syllogistic (CS), regard it as a system without notable issues, which is worthy of syntactic and semantic studies — and of generalizations: e.g., negative terms could be allowed (Alvarez and Correia [5]). Some papers pertinently discuss CS without using complex mathematics (see, for example, Read [6] and most of the papers cited therein), while many others extend CS in various ways, and study those generalizations, using, e.g., model theory (see Moss and Kruckman [7] and van Rooij [8]). (Abstracting a modern logic system out of CS already mitigates its less

felicitous traits noticed by George Boole and Lewis Carroll.) This paper uses a set-theoretical model of CS, as started by George Boole, Stanley Jevons, John Venn, Lewis Carroll, Hans Reichenbach, etc., with the hope that mathematics unavailable to Aristotle, even if elementary, can simplify the presentation of CS and eliminate some features of CS which are unnecessarily cumbersome or superfluous:

The syllogistic terms are interpreted as sets, (the same letter will denote both a term and its interpreting set), and the syllogistic quantifiers, A, E, I, O , are interpreted as either set inclusions or set intersections. One also assumes that there is always a universe of discourse, U , generated by the sets which interpret the syllogistic terms and their complements in U . When one has three syllogistic terms, usually denoted as S, P and M , one obtains a universe of discourse that has 2^3 partitioning subsets (in short subsets). The truth-condition for the $A(M, P)$ categorical statement is: $A(M, P)$ is true iff the set, M , which interprets the syllogistic term M is included in (or identic to) the set which interprets the term P . Equivalently, and in short, $A(M, P)$ is true iff $M \cap P' = \emptyset$. The truth conditions for the other categorical statements are: $E(M, P)$ is true iff $M \cap P = \emptyset$; $O(M, P)$ is true iff $M \cap P' \neq \emptyset$; $I(M, P)$ is true iff $M \cap P \neq \emptyset$. In this set interpretation, syllogistic figures (and moods) can be replaced by a matrix, whose elements are all the distinct pairs of categorical premises (PCPs) which can be formed by adjoining a P-premise with an S-premise, where both premises contain the middle term M , (or M'), and where either positive or negative terms can appear in either premise and the LC. For example, the P – premises $A(M', P) = E(M', P')$ and $O(M', P) = I(M', P')$ do not appear in Classical Syllogistic because they cannot be expressed using only positive terms. The other six distinct P – premises can all be formulated using only positive terms: $A(M, P), E(M, P), E(M', P) = A(P, M), O(M, P), I(M, P), I(M', P) = O(P, M)$. One it is thus lead to a six-by-six PCP matrix, i.e., to 36 distinct PCPs which need to be considered by Classical Syllogistic, and to an eight-by-eight PCP matrix, or to 64 PCPs which are conclusive syllogism candidates if both positive and negative terms are accepted in the categorical statements. The eight P – premises form an M - P cube of opposition, but except from the A , (resp. E), statement being the contradictory of the O , (resp. I), statement, no other relationships are imposed between the cube's eight statements. For example, $A(M, P)$ and $E(M, P)$ are not contraries: both being true just particularizes the universe of discourse, U , to one in which $M = \emptyset$. The M - S cube of opposition of the S – premises is treated similarly. For many of the other possible relationships between a cube of opposition statements, which may be generated via various existential import (ei) conditions imposed on the syllogistic terms, see Reichenbach [12]. In the present paper, the ei is always an ad hoc supplementary premise used only as a last step of a derivation, in order to extract a particular LC out of the universal LC already entailed by the

universal PCPs of type Barbara or Darapti. This paper accounts for any PCP which entails an LC — and thus generates a conclusive syllogism, even if it is not a valid syllogism as per Classical Syllogistic definition. One shows that a PCP, if entailing an LC, does so via pinpointing a unique partitioning subset of U , and that this pinpointing happens if and only if the PCP, follows one of the Barbara, Darapti or Darii patterns. This fact is checked examining all the 64 PCPs containing indefinite terms. Then one shows, for any $n > 3$, how one can use either the pure Barbara or the pure Darapti patterns, or, how one can use the same general “recipe”, (R_k) , (see Section 4), to add premises to the Barbara, Darapti and Darii type PCPs, such that an entailed LC of a sorites with $4, 5, \dots, n$ terms will still pinpoint a unique subset of the universe of discourse, U , which now contains, respectively, $2^4, 2^5, \dots, 2^n$ (partitioning) subsets. For a set of categorical premises constructed as above, their locating, (or pinpointing), a unique partitioning subset of U is thus becoming — and intuitively so — the only LC entailment criterion. One can use two methods for finding the unique partitioning subset which constitute the LC of a sorites of type Barbara, Darapti or Darii: a graphical one, based on Karnaugh maps and induction on the number of terms in the sorites, or an analytic method, based on “Jevon’s substitution principle”. For example, if $M_1 \subseteq M_2$, (and thus $M'_2 \subseteq M'_1$), one may express the above inclusions as $M_1 = M_1M_2$, (and, resp. $M'_2 = M'_1M'_2$). Thus, via successive substitutions, the two “one subset LCs” of a Barbara type sorites, $M_1 \subseteq M_2 \subseteq M_3 \subseteq \dots \subseteq M_n$, will obviously be $M_1 = M_1M_2M_3\dots M_{n-1}M_n$ and $M'_n = M'_1M'_2M'_3\dots M'_{n-1}M'_n$. (See Sections 4 and 5.) Moreover, for $n = 3$, one argues that, once one chooses the term which will play the role of the middle term, (and one denotes it by M), Barbara, Darapti and Darii syllogisms can be chosen as pattern representatives, and all the other conclusive syllogisms represent ways of changing the patterns’ content, via either replacing some or all of the positive terms S, P, M , from the patterns, with their negative counterparts in all the possible ways, or via switching the roles which the S and P terms play in the Darii’s type syllogisms — thus obtaining the contents of the Disamis type syllogisms (see Section 5). The above “one subset of U paradigm” can replace the initial Aristotelian CS paradigm, which considers the First Figure valid syllogisms as self-evident. (In a modern logic approach, various CS treatments consider the 1st Figure syllogisms — or only Barbara and Celarent — either as axioms or as rules of inference.) Section 6 displays the structure of the eight-by-eight PCP matrix of the P and S premises with indefinite terms. The conclusive syllogisms are generated by 32 of these PCPs, (equally split between the Barbara, Darapti, Darii and Disamis type syllogisms: $8 = 2^3 =$ in how many ways one can plug the three indefinite terms into each of the above patterns or types), while the other 32 PCPs do not entail any LCs. One also argues that the elimination of the middle term out of the logical consequences (LCs), even if it

was started by Aristotle, and provides LCs presented in the same A, E, O, I formats as the premises from which they are derived, is a somewhat misplaced endeavor, because it weakens the LCs.

3 The awkwardness of Classical Syllogistic

In the Classical Syllogistic, the premises are restricted to being formulable only via the positive terms, S, P, M , and the LCs are restricted, by definition, to only the statements $A(S, P), E(S, P), I(S, P), O(S, P)$. The first task is to find all pairs of categorical premises (PCPs) in which only the three positive terms, S, P, M , do appear — even if one may say that, surreptitiously, negative terms can appear, too, because $O(S, P) = I(S, P')$, $A(M, P) = E(M, P')$, $A(P, M) = E(M', P)$, etc. But statements such as $I(S', P')$ or $E(M', P')$ are not allowed in Classical Syllogistic, since they cannot be re-written as categorical statements containing only positive terms. If, as it happens, the LC is $A(P, S)$, (as entailed by Bramantip's premises), or, if it is $O(P, S)$, then the Classical Syllogistic, which does not accept such LCs, uses the fact that a $P \leftrightarrow S$ relabeling, i.e., a switch, or a permutation, in the names of the variables, (or in the order of premises, combined with the convention to denote by P the term — different from M — appearing in the firstly listed premise), will transform such pairs of categorical premises (PCPs), without changing their content, into PCPs which have the right LC types, as an “excuse” to disregard the $A(P, S)$ LC and the PCPs whose LC is $O(P, S)$. Nevertheless, the Classical Syllogistic accepts Bramantip's premises, as they are, All P is M , All M is S , but instead of using Bramantip's universal LC, All P is S , (i.e., $A(P, S)$), it uses only the particular, existential import (ei) LC, of the same premises, after imposing a supplementary ei condition on $P, P \neq \emptyset$; thereafter, $A(P, S)$ implies the acceptable ei LC, “Some S is P ”, of the (now) valid syllogism Bramantip. Therefore Classical Syllogistic tries to find each PCP, formulable via only positive terms, which, with or without existential import (ei), validly implies one of the statements $A(S, P), E(S, P), I(S, P), O(S, P)$. According to Classical Syllogistic, the set of “PCP candidates” contains 64 PCPs: 16 pairings of the A, E, I, O statements — four per each premise, and, since in each premise one may reverse the order in which the two terms appear in the categorical statement, one multiplies with another two possibilities per premise, i.e., one finally gets $16 \cdot 4 = 64$ PCPs formulable via only positive terms. But this regards as distinct some superfluous syllogistic figures, generated by permuting the terms inside the E and I statements, even if the content of the statement changes only when permuting the arguments of the A and O statements. In reality only 36 PCPs are distinct, even if, with the help of the superfluous syllogistic figures, one arrives to a 64 “syllogistic

figure” PCP count. It turns out that only 15 PCPs out of the 64 PCPs (mistakenly counted as distinct), entail one of the required LCs, $A(S, P)$, $E(S, P)$, $I(S, P)$, or $O(S, P)$. (In fact, out of the 15 PCPs, only eight PCPs are distinct: the Ferio/Festino/Ferison/Fresison, Celarent/Cesare, Camestres/Camenes, Darii/Datisi, and Disamis/Dimaris “groupings around slashes”, homologate syllogistic figure distinctions without any PCP, nor LC, content differences. Besides the 12 valid syllogisms just mentioned, out of which just five are distinct, Barbara, Baroco and Bocardo are the other three valid syllogisms which complete the list of the 15 PCPs entailing one of the standard LCs without any ei supplementary help.) Out of these 15 PCPs, five PCPs, (of pattern Barbara: Barbara, Celarent/Cesare, Camestres/Camenes), entail the $A(S, P^*)$ conclusion — All S is P^* ; thus if one also imposes a supplementary ei condition on S , i.e., $S \neq \emptyset$, then, the statement $I(S, P^*)$ is also entailed by each one of the five PCPs (out of which only three PCPs are really distinct, since Celarent and Cesare, (resp. Camestres and Camenes), have the same PCP — modulo superfluous syllogistic figures). As already mentioned, Bramantip’s PCP entails the $A(P, S)$ logical consequence (LC) which is not on the above LC list — $A(S, P)$, $E(S, P)$, $I(S, P)$, $O(S, P)$; but, since, per $A(P, S)$, P is included in S , if one imposes the supplementary ei on P , then the $I(S, P)$ LC is also entailed — and one obtains Bramantip as another valid syllogism with a particular LC. Finally three other PCPs, (out of the 64 “figure distinguishable” PCPs formulable via only positive terms), entail these LCs, respectively: All M is $S \cap P$, All M is $S \cap P'$, All M is $S \cap P'$, which are very different from the four acceptable LCs, first of all because M is not yet eliminated from the LC — the elimination of the middle term out of the LC is another questionable requirement of the Classical Syllogistic since it weakens the LC — see below. (On epistemological grounds, Aristotle himself introduced the elimination of the middle term: he wanted a syllogism to provide “new knowledge” — by establishing a direct connection between the P and S terms — deduced from the “initial knowledge” contained in the two premises where the middle term had appeared in each of the premises.) But then, ei on M rescues these three PCPs, (out of which only two are distinct), because imposing $M \neq \emptyset$, implies $S \cap P \neq \emptyset$, (resp. $S \cap P' \neq \emptyset$), i.e., $I(S, P)$, (resp. $O(S, P)$), and thus approved LCs are again entailed. This way, after ei on M was imposed, and M was eliminated from the LC, one obtains the valid syllogisms Darapti, Felapton and Fesapo, whose PCPs are: All M is P , All M is S (Darapti), No M is P , All M is S (Felapton), and No P is M , All M is S (Fesapo). Again, one can see that there is no logical content difference between Felapton and Fesapo — only their syllogistic figures differ, while their different names homologate a distinction without a difference. (To show that out of the 64 PCPs formulable via only positive terms, only 36 PCPs are distinct when superfluous figures are removed, one just has to observe that there are six

P – premises, (resp. *S* – premises), formulable via only positive terms: they pair up to only 36 PCPs.) At last count, in Classical Syllogistic, there are 15 PCPs whose LCs may be directly expressed in one of the four LC formats homologated as correct by the Classical Syllogistic, and, via ei on *S*, five of the 15 PCPs produce a second LC, $I(S, P^*)$, of the correct type, and, there are four other PCPs which, via ei on *P* or *M*, entail the $I(S, P)$ or $O(S, P)$ LCs — and thus raise the total count to 24 valid syllogisms, six per (lucky) syllogistic figure. See Copi [10] and Hurley [11].

Instead of the Classical Syllogistic approach, one may use the eight *P* – premises containing indefinite terms, (which form a cube of opposition — see, e.g., Reichenbach [12] — $E(M^*, P^*), I(M^*, P^*)$), and pair them up with the eight *S* – premises containing indefinite terms, (which form another cube of opposition, $E(M^*, S^*), I(M^*, S^*)$), to obtain a 64 elements PCP matrix having 32 conclusive PCPs and 32 inconclusive PCPs — where $M^* \in \{M, M'\}$, $P^* \in \{P, P'\}$, $S^* \in \{S, S'\}$. Section 6 discusses this PCP matrix.

4 The “reference system” provided by the subsets of the universal set, *U*, for analyzing simple biliteral sorites

Instead of following the Classical Syllogistic in its search for PCPs which entail an LC, while subjecting both PCPs and LCs to the above and other preconditions, (such as “two negative premises are not allowed” or “the middle term has to be distributed in at least one premise”), one uses the reference system provided by the subsets of the universal set, to tackle the problem of finding LCs out of categorical premises, without any constraints but logical entailment: if the premises are true, the conclusion should be necessarily true, also. One notices that the Barbara and Darapti syllogisms and sorites contain only universal premises *A* and *E*, where, by obversion, the *A* statements can be replaced by *E* statements and vice-versa, since $A(M_j, M_k) = E(M_j, M'_k)$, meaning $M_j M'_k := M_j \cap M'_k = \emptyset$. In a universal set, *U*, containing *n* terms, such a universal premise would empty 2^{n-2} subsets, and would leave $3 * 2^{n-2}$ subsets about which nothing is yet asserted. The only LC one might hope to obtain via emptying subsets is that one or more terms will end up with all their respective subsets emptied, except for one subset about which the premises do not assert anything. The resulting LC is that the whole term reduces to one of its 2^{n-1} subsets. If one wants to impose an ei condition on such a term, it will be clear which one of the term’s subsets will be non-empty: the only one subset which was not emptied, after all the universal premises were applied! In this Section, the numbering of the formulas and figures will contain a (*b*) if they refer to the Barbara type syllogisms or sorites, will contain a (*d*), as in Darapti, if they refer to the

Darapti type syllogisms or sorites, and will contain a (p), as in particular (premise), if they refer to Darii type syllogisms or sorites. The “pure Barbara” type premises amount to a Barbara-chain pattern of inclusions,

$$(b_0) \quad M_1 \subseteq M_2 \subseteq M_3 \subseteq \dots \subseteq M_n.$$

The “pure Darapti” type premises amount to a “Darapti key ring” pattern of inclusions,

$$(d_0) \quad M_1 \subseteq M_2, M_1 \subseteq M_3, \dots, M_1 \subseteq M_n.$$

In (d_0), the M_1 set represents the key ring, all the other terms/sets are “attached” to it via the set inclusion relation. The sets appearing in the Barbara-chain of inclusions can be imagined as interlocking links of a chain. At the end of this section one shows, for any number of terms $k, k \geq 4$, that the Barbara, Darapti and Darii 3-term syllogisms may each be continued via the same general recipe, which may be thought as a mixture of “key rings and chain fragments attached to key rings” such that all the generated sorites still entail a one subset (of U) LC — see below.

The effect of applying the set of $n - 1$ premises of the pure Barbara’s, (resp. Darapti’s), type is that two sets, (resp. one set), will be left with all the subsets emptied — except for one subset about which the emptying premises do not assert anything. Clearly (b_0) implies (d_0) — this premises’ difference results in (b_0) entailing two LCs, while (d_0) entails only one LC — see below.

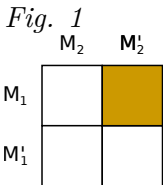
Take, e.g., Celarent’s premises, No M is P , All S is M . (Notice the convention to firstly list the premise containing the P term — since the LC does not depend on the premises’ order, this convention would be pointless — but for the fact that for defining the syllogistic figures, (and for recovering the valid syllogisms’ historical names), the order of the premises has to be uniformly the same; in Classical Syllogistic the real restriction comes from the fact that one accepts as generating valid syllogisms only the PCPs whose LCs have one of the four homologated formats: $A(S, P), E(S, P), I(S, P), O(S, P)$.) Celarent’s premises may be re-written as Barbara’s premises in which P' appears instead of P : All M is P' , All S is M . This also shows that via a relabeling $P \leftrightarrow P'$ Barbara becomes Celarent and vice-versa. Writing Barbara’s premises as $E(M, P')E(M', S)$, and Bramantip premises as $E(M', P)E(M, S')$, one sees that each of these two PCPs may be re-written, (without any content change) as the other one, via a relabeling $P \leftrightarrow P', S \leftrightarrow S', M \leftrightarrow M'$, (or, via a relabeling $P \leftrightarrow S$). The universal LC of Bramantip premises, [(All P is M , All M is S), (therefore)] “All P is S ”, is unacceptable to Classical Syllogistic, even if Bramantip’s premises are clearly of Barbara’s type. Only by adding $P \neq \emptyset$ one can extract Bramantip as an ei valid syllogism! The Darii syllogisms and sorites contain only one particular premise and the rest are the universal premises A and E . As mentioned, Lewis Carroll, expressed very clearly the idea of the 3-term

sylogisms being of only three different patterns. All the premises of the first two types are universal, the Darii type has only one particular premise, which will be always written as $M_1M_2 := M_1 \cap M_2 \neq \emptyset$, (meaning “Some M_1 is M_2 ”). In a Darii sorites with n terms, the LC one can hope for, will say that out of the 2^{n-2} subsets of $M_1M_2 \neq \emptyset$ which are possibly non-empty, the $n-2$ universal premises will empty all but a unique subset of M_1M_2 — which remains non-empty and represents the LC of the Darii sorites.

To graphically represent the syllogisms or biliteral sorites one uses the Karnaugh maps, (or, in short, K-maps), built on only two rows, via mirroring — toward the right — the 2^n partitioning subsets of the universe of discourse with n terms, such that, after mirroring, one obtains the 2^{n+1} partitioning subsets of the universe of discourse with $n+1$ terms (see below). The K-maps were in fact firstly introduced by Alan Marquand [13] in 1881, only one year after John Venn [14] used the 3-circle diagram for representing the three terms of a categorical syllogism; then these maps were rediscovered and improved by Veitch [15] and Karnaugh [16]. For more details about Karnaugh maps on any number of rows, see Abdalla [17].

For a 2-term universe, $n = 2$, the one premise, $M_1M'_2 = \emptyset$, empties one subset. This is equivalent with $M_1 \subseteq M_2$ and $M'_2 \subseteq M'_1$, which, in turn, cf. Stanley Jevons [9] “The Substitution Of Similars, The True Principle Of Reasoning”, can be expressed as: $M_1 = M_1M_2$ and $M'_2 = M'_1M'_2$, which represent two “one partitioning subset” LCs.

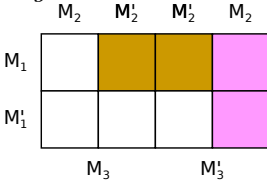
The first premise of the Barbara and Darapti type syllogisms and sorites will be always written as $M_1M'_2 := M_1 \cap M'_2 = \emptyset$, (meaning “All M_1 is M_2 ”). This means that on *Figure 1* below, representing just two terms, the right top subset will be shaded to show its emptiness:



Note that the Karnaugh map for $n = 3$ will be obtained via mirroring the four partitioning subsets of the above universe of discourse for $n = 2$. After mirroring — toward the right edge of the $n = 2$ K-map — the “old subsets” receive M_3 as an extra index, and the “newly minted subsets” receive M'_3 as an added index; the “mirror” was vertically placed at the right edge of the K-map for $n = 2$, (i.e., between the two M'_2 columns from Figure 2b). For a 3-term universe, $n = 3$, Barbara’s premises, $M_1M'_2 = \emptyset$ and $M_2M'_3 = \emptyset$, empty four subsets $M_1M'_2M_3 + M_1M'_2M'_3 = \emptyset$ and

$M_1M_2M'_3 + M'_1M_2M_3 = \emptyset$. One thus gets a Karnaugh map for $n = 3$, having 2^3 subsets:

Fig. 2b



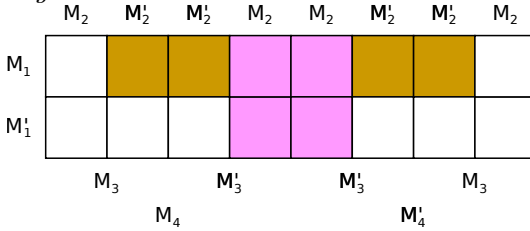
Either graphically, or via Jevons' substitution principle, Barbara's premises provide two LCs: $M_1 = M_1M_2M_3$ and $M'_3 = M'_1M'_2M'_3$. The four subsets about which the premises do not assert anything are: $M_1M_2M_3$, $M'_1M_2M_3$, $M'_1M'_2M_3$ and $M'_1M'_2M'_3$. One can also check that these "remaining" subsets still satisfy $U = M_1 + M'_1 = M_2 + M'_2 = M_3 + M'_3$ and $M_1 \subseteq M_2 \subseteq M_3$. Nothing was explicitly said about the "remaining" four subsets. But graphically one could have gotten two LCs, describing a unique characteristic of two of these subsets, $M_1M_2M_3$ and $M'_1M'_2M'_3$, about which Barbara's premises do not assert anything explicitly, namely,

(b_1) $M_1 = M_1M_2M_3$ and $M'_3 = M'_1M'_2M'_3$ since the sets M_1 and M'_3 have each three empty subsets and one subset which remains, possibly, non-empty. An existential import condition imposed on M_1 or/and M'_3 will assert that $M_1M_2M_3$ or/and $M'_1M'_2M'_3$ are not empty, and thus, e.g., an LC of the type "Some M_2 is M_3 ", or "Some M_1 is M_3 " could be inferred. (One also sees that there is nothing magic about eliminating any of the terms — once the "smallest" set is non-empty any two sets out of three will have Some(thing) in them. But insisting on the elimination of the middle term M_2 , e.g., in Barbara, (resp. Barbari), is, I believe, unfortunate: instead of being told that "All M_1 is $M_1M_2M_3$ ", i.e., $M_1 = M_1M_2M_3$, (resp. $M_1M_2M_3 \neq \emptyset$), one is left with only, in principle, the weaker information that $M_1 = M_1M_3 = M_1M_2M_3 + M_1M'_2M_3$, (resp. $M_1M_3 \neq \emptyset$). Replacing the precise LC, "All M_1 is $M_1M_2M_3$ ", by the standard universal LC, "All M_1 is M_3 ", means that M_1 might be "spread" onto the subsets $M_1M_2M_3$ or $M_1M'_2M_3$ — which is not entirely correct, since one already knows that $M_1M'_2M_3$ is empty. Both LCs from (b_1) satisfy the "one subset LC paradigm": one has an LC if and only if the premises pinpoint a unique subset of U . One can also list the four subsets about which the premises do not explicitly assert anything:

(b_2) $M_1M_2M_3$, $M'_1M_2M_3$, $M'_1M'_2M_3$, $M'_1M'_2M'_3$ — although, as said, graphically, two LCs were obtained: $M_1 = M_1M_2M_3$ and $M'_3 = M'_1M'_2M'_3$.

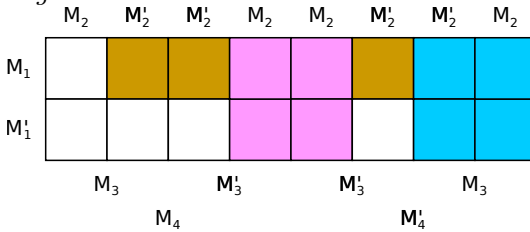
For a Barbara type sorites with $n = 4$, one firstly shades the subsets in Figure 3b corresponding to the emptying action of the first two premises, $M_1M'_2 = \emptyset$ and $M_2M'_3 = \emptyset$:

Fig. 3b



Notice that the empty sets, via mirroring, propagate to the right, since the first two premises, $M_1M'_2 = \emptyset$ and $M_2M'_3 = \emptyset$, empty any set intersection which includes $M_1M'_2$ or $M_2M'_3$, i.e., empty any such subsets which also carry M_4 or M'_4 as an extra index. The third premise, $M_3M'_4 = \emptyset$, empties again the already empty subset $M_1M'_2M_3M'_4$, and, most importantly, empties the three rightmost subsets of the diagram which were “propagated” via “mirroring” from the three leftmost subsets about which the first two premises did not assert anything: now the third premise empties them all at once. Therefore, one obtains the following map of the emptied and not emptied subsets:

Fig. 3b'



It turns out that the general, n -term formulas for the two Barbara’s LCs and for the list of subsets about which the Barbara type sorites with n terms do not assert anything, are:

(b₃) $M_1 = M_1M_2M_3\dots M_{n-1}M_n$ and $M'_n = M'_1M'_2M'_3M'_4\dots M'_{n-1}M'_n$ — these are the two exact LCs: one has found out that each of the M_1 and M'_n sets are made of one (possibly non-empty) subset — all the other $2^{n-1} - 1$ subsets of each of M_1 and M'_n are empty.

(b₄) $M_1M_2M_3\dots M_{n-1}M_n$, $M'_1M_2M_3\dots M_{n-1}M_n$, $M'_1M'_2M_3\dots M_{n-1}M_n$, $M'_1M'_2M'_3M_4\dots M_{n-1}M_n$, \dots , $M'_1M'_2M'_3\dots M'_{n-1}M_n$, $M'_1M'_2M'_3M'_4\dots M'_{n-1}M'_n$, is the complete list of the $(n + 1)$ subsets about which the premises of a Barbara sorites do not assert anything explicitly; the first and the last subsets became the two LCs from the formula (b₃); the total number of emptied subsets is $2^n - (n + 1)$. Note that the union of the first two subsets in (b₄) equals M_2 , the union of the first

three subsets in (b_4) equals M_3 , the union of the first four subsets in (b_4) equals M_4 , ..., the union of the first n subsets in (b_4) equals M_n , while the $(n + 1)^{th}$ subset represents M'_n ; the union of the last two subsets in (b_4) equals M'_{n-1} , ..., and the union of the last n subsets in (b_4) equals M'_1 , which are exactly the relationships necessary to verify that $U = M_1 + M'_1 = M_2 + M'_2 = \dots = M_n + M'_n$. It results that (b_4) spells in a precise, “Karnaugh-map manner”, what Barbara’s chain inclusions initially asserted:

(b_0) $M_1 \subseteq M_2 \subseteq M_3 \subseteq \dots \subseteq M_n$ — and, therefore, $M'_n \subseteq M'_{n-1} \subseteq M'_{n-2} \subseteq \dots \subseteq M'_1$.

Note that (b_4) implies (b_3) , and that (b_3) and (b_4) are valid for $n = 2$: since $M_1M'_2 = \emptyset$, one has two obvious LCs, $M_1 = M_1M_2$ and $M'_2 = M'_1M'_2$; $2^2 - (2 + 1) = 1 =$ one empty subset; a list of $2 + 1 = 3$ subsets about which the universal premise does not assert anything. Also (b_3) and (b_4) are valid for $n = 3$: according to (b_1) there are again two LCs; $2^3 - (3 + 1) = 4 =$ four empty subsets; a list of $3 + 1 = 4$ subsets about which the universal premises do not assert anything. To prove (b_4) by induction, suppose that the list $M_1M_2M_3\dots M_{n-1}M_n$, $M'_1M_2M_3\dots M_{n-1}M_n$, $M'_1M'_2M_3\dots M_{n-1}M_n$, $M'_1M'_2M'_3M_4\dots M_{n-1}M_n$, ... $M'_1M'_2M'_3\dots M'_{n-1}M_n$, $M'_1M'_2M'_3M'_4\dots M'_{n-1}M'_n$ contains the only $n + 1$ subsets which the $n - 1$ premises (b_0) did not empty. One now doubles the number of subsets (or constituents) of U by adding another term, M_{n+1} , and another premise to (b_0) : $M_nM'_{n+1} = \emptyset$. The “old” subsets receive an extra index, M_{n+1} , and the new subsets, mirrored, say, to the right, on a Karnaugh-map with two rows receive the extra index, M'_{n+1} . The newest premise, $M_nM'_{n+1} = \emptyset$, does not act on the subsets which received the extra index M_{n+1} , but it empties all the mirrored subsets which were not emptied by the $n - 1$ premises, and which, at their mirrored creation, received the extra index M'_{n+1} , except for the subset $M'_1M'_2M'_3M'_4\dots M'_{n-1}M'_nM'_{n+1}$ — the mirror image of $M'_1M'_2M'_3M'_4\dots M'_{n-1}M'_nM_{n+1}$, because this latter subset did not contain M_n . QED.

For the 3-term syllogism, applying the two premises of the Darapti’s pattern, $M_1M'_2 = \emptyset$ and $M_1M'_3 = \emptyset$, one sees that only three subsets, (not four), will be emptied — since the subset $M_1M'_2M'_3$ is emptied by each of the two premises. Nothing is explicitly said about the other five subsets. But graphically, or via Jevons’ substitution principle, one has already gotten one LC, describing a unique characteristic of the subset $M_1M_2M_3$, about which Darapti’s premises do not assert anything explicitly, namely:

(d_1) $M_1 = M_1M_2M_3$ since the set M_1 has three empty subsets.

This LC satisfies the “one subset LC paradigm”: one has an LC if and only if the premises pinpoint a unique subset of U . It is simpler, than in Barbara’s syllogism case, to list the five subsets about which the premises do not explicitly assert anything:

(d_2) $M_1M_2M_3$ and M'_1 ,

since Darapti's premises do not assert anything about any of the subsets of M'_1 .

For the Darapti type sorites with n terms, the formulas (d_1) and (d_2) become:

(d_3) $M_1 = M_1M_2M_3\dots M_{n-1}M_n$,

this LC satisfies the "one subset LC paradigm"; the other $2^{n-1} - 1$ subsets of M_1 are empty.

(d_4) $M_1M_2M_3\dots M_{n-1}M_n$ and M'_1 ,

are the sets about which the premises of a Darapti sorites do not assert anything explicitly; the first of the two sets in the (d_4) 2-set list, represents the LC of the Darapti's sorites from the formula (d_3); the total number of emptied subsets is $2^{n-1} - 1$.

Barbara and Darapti sorites have the same first premise. Since Darapti's syllogism and sorites premises do not assert anything about any subset of M'_1 , the second LCs will be lost for any $n > 2$. To prove (d_3) and (d_4) by induction, (even if they are pretty obvious), suppose the $n - 1$ Darapti premises, $M_1M'_2 = \emptyset$, $M_1M'_3 = \emptyset, \dots, M_1M'_n = \emptyset$, empty all the subsets of M_1 except $M_1M_2\dots M_n$, and suppose that nothing is asserted about the set M'_1 . Then, one doubles the number of subsets by adding another term: M_{n+1} . The M_{n+1} index is added to the "old subsets", and the new, mirrored subsets receive the M'_{n+1} additional index. The only two subsets of the set M_1 which were not already emptied by the above $n - 1$ Darapti premises are $M_1M_2\dots M_nM_{n+1}$ and $M_1M_2\dots M_nM'_{n+1}$. The Addition of the n^{th} Darapti premise, $M_1M'_{n+1} = \emptyset$, will empty the subset $M_1M_2\dots M_nM'_{n+1}$. Also, the n^{th} premise does not assert anything about the set M'_1 . QED.

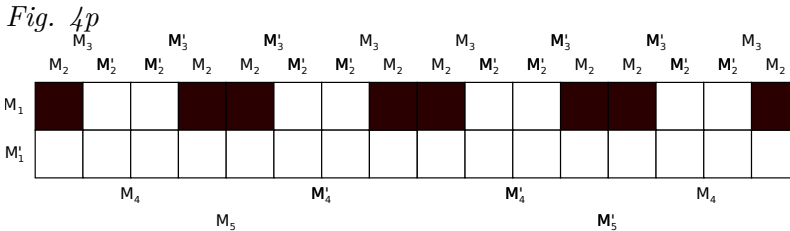
(Note also that any and all of the premises, $M_iM'_k = \emptyset, i = 1, 2, \dots, k - 1$, added together or separately when one adds the k^{th} term to a sorites, would have emptied the subset $M_1M_2\dots M_{k-1}M'_k$, and would have thus created a sorites with a "one subset LC" for each of the k^{th} added term: $M_1 = M_1M_2\dots M_{k-1}M_k$. See below this "uniform general recipe", (R_k), for the continuation of any of the 3-term Darapti, Barbara and Darii type syllogisms, to sorites having a "one subset of U " LC.)

Any Darii sorites has just one particular premise, always chosen to be $M_1M_2 \neq \emptyset$. When a Darii sorites has n terms, the set M_1M_2 contains 2^{n-2} subsets, and $M_1M_2 \neq \emptyset$ means that all these subsets might be non-empty — at least one of them is — and this does not sound like a "one subset paradigm LC". Therefore, to obtain such an LC, one should choose the second, third, etc., universal premises, in such a way that, for any $k > 2$, only, e.g., the subset $M_1M_2M_3\dots M_{k-1}M_k$ will remain non-empty. Thus, the role of the universal premise in a Disamis, Darii, Ferio, Bocardo, Baroco, etc., 3-term syllogism, is to empty the subset $M_1M_2M'_3$. The difference between these syllogisms consists in the terms they use. One can relabel M_1, M_2, M_3 , by using the standard names for the syllogistic variables, S, P,

M, S', P', M' , in various ways such that the Darii's pattern premises, $M_1M_2 \neq \emptyset$ and either $M_1M'_3 = \emptyset$ or $M_2M'_3 = \emptyset$, will represent any and all of the syllogisms Darii, Ferio, Baroco, Disamis and Bocardo. Similarly, all the eight syllogisms of type Barbara, (resp. Darapti), use the same pattern — Barbara, (resp. Darapti).

The difference between a sorites whose all premises are universal and one that has one particular premise, is that the “mission” of all the universal premises in a Barbara or Darapti type sorites is to empty all subsets of a set except one — that one subset is then the subject of the “one subset LC” — for example, one term, such as M_1 , is reduced to just one partitioning subset of U ; the “mission” of all the universal premises in a Darii type sorites is to empty, except one, all the subsets about which the only particular premise asserts that at least one of them is non-empty — this way one obtains, again, a “one subset LC” — which, now, is definitely not empty. In a universe with n terms and 2^n subsets, the Barbara and Darapti sorites empty $2^{n-1} - 1$ subsets of the set which will provide the LC of the sorites. The “mission” of leaving just one non-empty subset of $M_1M_2 \neq \emptyset$ once all the $n - 2$ universal premises of a Darii type sorites are applied, is accomplished, if the $n - 2$ universal premises of the Darii sorites empty $2^{n-2} - 1$ subsets out of the 2^{n-2} subsets of $M_1M_2 \neq \emptyset$.

The Figure 4p shows the 2-row K-map for $n = 5$, with only the effect of the particular premise $M_1M_2 \neq \emptyset$ depicted on it. It was obtained by successively mirroring to the right the K-map for $n = 2$, on which each subset of M_1M_2 was blackened to show that any of its partitioning subsets might be non-empty:



One lists the possible premises of a Darii type sorites for $k = 3, 4, \dots, n$. For each k the pattern is as follows: For $k = 3$ one needs to empty just the subset $M_1M_2M'_3$. This can be done by adding as a 2^{nd} premise either $M_1M'_3 = \emptyset$ or $M_2M'_3 = \emptyset$. In Darii's case, choosing between the two 2^{nd} premises amounts to choosing if either M_1 or M_2 will play the role of the middle term. But when the first premise is the universal premise $M_1M'_2 = \emptyset$, (instead of the particular Darii's first premise $M_1M_2 \neq \emptyset$), then which one of the two 2^{nd} premises is added to the first premise, determines if one obtains either a Darapti or a Barbara 3-term syllogism. One thus obtains a recipe, $[(R_k)]$, see below], for continuing Darii's premises to the premises of

a sorites entailing a one subset LC, and the same recipe can be used for generating sorites out of Barbara's and Darapti's 3-term syllogisms, too. For Darii's extension to $k = 4$, only the subset $M_1M_2M_3M'_4$, the mirror image of the $M_1M_2M_3M_4 \neq \emptyset$ needs to be emptied. This can be accomplished by adding, as a 3^{rd} premise, any one of the following three relations, $M_1M'_4 = \emptyset$, or $M_2M'_4 = \emptyset$, or $M_3M'_4 = \emptyset$. Either one of these three statements, when added as a third premise would assure that $M_1M_2M_3M'_4 = \emptyset$. It results that a Darii type syllogism can be extended to a Darii type sorites via either Darapti key rings or via Barbara interlocked chain-links, "based" (or originating) at any term. A short way to express the multitude of premise choices is as follows.

(R_k) For any term $k \geq 3$ one has $k - 1$ choices for adding the k^{th} term, (and the $(k - 1)^{th}$ premise if one adds only one premise per added term): $M_1M'_k = \emptyset$, $M_2M'_k = \emptyset$, $M_3M'_k = \emptyset$, ..., $M_{k-2}M'_k = \emptyset$, $M_{k-1}M'_k = \emptyset$.

(Taken together, the above premise choices form the Darapti premises whose "subject" is M'_k , the terms are $M_1, M_2, M_3, \dots, M_{k-1}, M_k$ and the LC is $M'_k = M'_1M'_2M'_3 \dots M'_{k-1}M'_k$.) Using the "recipe" (R_k) , one obtains a Darii type sorites, by using any subsets of premise choices, (either one choice or more than one choice per each set of the $k - 1$ premises acceptable for addition when adding a new, k^{th} term M_k , to the sorites, where $3 \leq k \leq n$. Any such "path through the premise choices", or "path of subsets of premise choices", will result in a Darii type sorites with n terms and either $n - 1$ or more premises. As said, the first, particular, premise of a Darii sorites is always $M_1M_2 \neq \emptyset$, and — when one decides to denote the middle term by M_1 , (resp. M_2), — the second premise is $M_1M'_3 = \emptyset$, (resp. $M_2M'_3 = \emptyset$). Each set of two premises are Darii's premises - only the name of the middle term changes from M_1 , to M_2 . Then the three simplest paths for obtaining a Darii sorites would be to use the Darapti key ring type universal premises "originating" either at M_1 , (for each added term, M_k , always choose the first premise from the list of the $k - 1$ possible premises available for addition), or "originating" at M_2 , (for each M_k always choose the second premise from the list of $k - 1$ possible premises), or, one may use the Barbara interlocked links started at M_3 (for each term $M_k, 3 \leq k \leq n$, always choose the last premise from the list of the $k - 1$ premises available for addition):

(p_1) $M_1M'_3 = \emptyset, M_1M'_4 = \emptyset, \dots, M_1M'_{n-1} = \emptyset, M_1M'_n = \emptyset$

(p_2) $M_2M'_3 = \emptyset, M_2M'_4 = \emptyset, \dots, M_2M'_{n-1} = \emptyset, M_2M'_n = \emptyset$

(p_3) $M_2M'_3 = \emptyset, M_3M'_4 = \emptyset, \dots, M_{n-2}M'_{n-1} = \emptyset, M_{n-1}M'_n = \emptyset$

One may declare any of the $(p_1), (p_2), (p_3)$ as being the "pattern Darii". As the (R_k) "recipe" shows, one can also simultaneously use all the premises from all the premise choices: *the one particular premise + two premise choices for the second premise + three premise choices for the third premise + ... + (n-1) premise choices for the (n-1)th premise.* The total number of premises would equal $n(n - 1)/2$. All these

premises together might be declared as the “pattern Darii” — their “one subset LC” is the same as when one chooses just one “path through the premise choices”: $M_1M_2\dots M_{n-1}M_n \neq \emptyset$. The above also shows that, for each added k^{th} term, one may choose to add more than one, but not necessarily all the available $k-1$ premises containing the k^{th} term which are available for addition.

The induction on n : Let suppose that all the subsets of $M_1M_2 \neq \emptyset$ except the subset $M_1M_2M_3\dots M_{n-1}M_n$ are emptied by any “path through the premise choices” for the $n-1$ premises or by any “path of subsets of premise choices” for the given n terms. Once we add a new term, M_{n+1} , the “old subsets” receive the extra index M_{n+1} — thus $M_1M_2\dots M_{n-1}M_nM_{n+1}$ remains definitely non-empty. And only its mirror subset $M_1M_2\dots M_{n-1}M_nM'_{n+1}$ will be definitely non-empty, while the rest of the new subsets of $M_1M_2 \neq \emptyset$, (among the mirrored subsets which received the M'_{n+1} as an extra index), will be definitely empty. Any of the n possible choices for the n^{th} premise of the Darii sorites, $M_1M'_{n+1} = \emptyset, M_2M'_{n+1} = \emptyset, \dots, M_{n-1}M'_{n+1} = \emptyset, M_nM'_{n+1} = \emptyset$, will empty the subset $M_1M_2\dots M_{n-1}M_nM'_{n+1}$. QED. (If the just mentioned premises are added all at once when one adds the $(n+1)^{\text{th}}$ term, then a 2nd LC becomes obviously available: $M'_{n+1} = M'_1M'_2M'_3M'_4\dots M'_nM'_{n+1}$. It is easy to see that the same recipe (R_k) will extend the Darapti 3-term syllogism to a sorites having a “one subset LC”. The above extension recipe (R_k) will also extend the Barbara 3-term syllogism to a sorites having, in general, just one LC — except when the path (p_3) is used for the extension — in that “pure” case, Barbara will also have a second “one subset LC”, $M'_n = M'_1M'_2M'_3M'_4\dots M'_{n-1}M'_n$. Note also that if the path (p_3) is used to extend the Darii 3-term syllogism to an n -term sorites, then, besides the “one subset LC” of the Darii sorites, one also obtains the relationship: $M'_n = M'_2M'_3M'_4\dots M'_{n-1}M'_n = M_1M_2M_3M_4\dots M'_{n-1}M'_n + M'_1M'_2M'_3M'_4\dots M'_{n-1}M'_n$.

5 There exist only three syllogism patterns: one may call them Barbara, Darapti and Darii

Only the indefinite terms used, e.g., in the syllogisms Barbara, Celarent, Camestres, Bramantip, differ, while the syllogisms’ structure is the same: $S^* \subseteq M^* \subseteq P^*$ where $M^* \in \{M, M'\}$, $P^* \in \{P, P'\}$, $S^* \in \{S, S'\}$. The same observation applies to the Darapti, Felapton, and Fesapo, whose terms, (and thus content), differ, but they all follow Darapti’s pattern. Similarly, there is no structural difference either between Darii, Ferio, Baroco, Disamis and Bocardo, which all follow Darii’s pattern. A short argument is as follows: One may relabel the sets M_1, M_2 , and M_3 as S, P, M, S', P', M' in such a way that each syllogism of type Barbara corresponds to a “Barbara pattern of inclusions”, $M_1 \subseteq M_2 \subseteq M_3$, each syllogism of type Darapti

corresponds to a “Darapti pattern of inclusions” $M_1 \subseteq M_2$ and $M_1 \subseteq M_3$, each syllogism of type Darii corresponds to the Darii premises $M_1M_2 \neq \emptyset$ and $M_1M_3' = \emptyset$. Then any difference among syllogisms of the same type is just a difference in the terms employed. One can completely describe all the conclusive syllogisms and valid syllogisms of the same type by a single formula — this also shows that the differences between syllogisms of the same type consist only in which terms they use. Instead of working only with premises formulable via only positive terms, (and thus keeping the terms fixed to only M, P, S), one generalizes to premises containing negative terms, too. The premises and LCs can be collected in one formula per each type, where, as before, $M^* \in \{M, M'\}, P^* \in \{P, P'\}, S^* \in \{S, S'\}$. Note that if $M^* = M'$, then $M^{*'} = M$, etc. Each of the eight PCPs of type Barbara can be written as $E(M^*, P^*)E(M^{*'}, S^*)$, or $S^* \subseteq M^* \subseteq P^{*'}$. Such a PCP entails two, non-independent, but different, LCs — since these PCPs are in fact Barbara type sorites with just one middle term: $S^* \subseteq M^* \subseteq P^{*'}$, (or $P^* \subseteq M^{*'} \subseteq S^{*'}$). Instead of using K-maps or 3-circle Venn diagrams to find the LCs of each type of PCPs, one can directly handle these PCPs using a tree like method, similar to Jevons’ [9] method of decompositions into subsets, or to Lewis Carroll’s [3] method of subscripts, and also similar to Carroll’s [3] own method of trees. Reading the sorites in the Aristotelian way, i.e. starting with the S^* term of the PCP, one gets: $S^* = S^*M^* + S^*M^{*'}$ = $S^*M^* = S^*M^*P^* + S^*M^*P^{*'}$ = $S^*M^*P^{*'}$, where $S^*M^*P^{*'}$:= $M^* \cap S^* \cap P^{*'}$, etc., and the + sign denotes union of (here disjoint) sets. Reading the sorites in the Goclenian way, i.e. starting with the P^* term of the PCP, one gets the 2nd LC: $P^* = P^*M^* + P^*M^{*'}$ = $P^*M^{*'}$ = $P^*M^{*'S^*} + P^*M^{*'S^{*'}}$ = $P^*M^{*'S^{*'}}$. Note that, based on the double inclusions $S^* \subseteq M^* \subseteq P^{*'}$ and $P^* \subseteq M^{*'}$ — inferred from the premises $E(M^*, P^*)E(M^{*'}, S^*)$ — each LC implies the other, i.e., only one LC is independent, but the two LCs pinpoint different subsets of U . The conclusive eight Darapti type PCPs are $E(M^*, P^*)E(M^*, S^*)$. They each entail just one universal LC, which is found in a similar way: $M^* = M^*P^* + M^*P^{*'}$ = $M^*P^{*'}$ = $M^*P^{*'S^*} + M^*P^{*'S^{*'}}$ = $M^*P^{*'S^{*'}}$, or $A(M^*, M^*P^{*'S^{*'}}$ which reflects the inclusions $M^* \subseteq P^{*'}$ and $M^* \subseteq S^{*'}$ asserted by the premises. Thus, for the type Darapti PCPs, the middle term itself is “the subject” of the precise LC; it can be eliminated only via ei on M^* . Also there are eight distinct PCPs of type Darii, $E(M^*, P^*)I(M^*, S^*)$, and eight distinct PCPs of type Disamis, $I(M^*, P^*)E(M^*, S^*)$. The very short trees determining the LCs are: $\emptyset \neq M^*S^* = M^*S^*P^* + M^*S^*P^{*'}$ = $M^*S^*P^{*'}$, and, resp., $\emptyset \neq M^*P^* = M^*S^*P^* + M^*S^{*'P^*}$ = $M^*S^{*'P^*}$.

Listing, after a column sign, and separated by semi-columns, all the possible LCs and LCs’ formats, the formulas describing all the 32 conclusive syllogisms are:

(1) (*Pattern and Type Barbara*) $E(M^*, P^*)E(M^{*'}, S^*) : S^* = S^*M^*P^{*'}$, $P^* = P^*M^{*'S^{*'}}$; $A(S^*, P^{*'}) [= A(P^*, S^{*'}) = E(S^*, P^*)]$ — after M^* is dropped, (classical

Syllogistic style!), the two LCs become identical and the number of all LCs of type Barbara reduces to only four; $I(S^*, P^{*'})$ — after ei on S^* ; $I(P^*, S^{*'})$ — after ei on P^* (for a total of eight ei particular LCs).

(2) (*Pattern and Type Darapti*) $E(M^*, P^*)E(M^*, S^*) : M^* = M^*P^{*'}S^{*'}$, $A(M^*, M^*P^{*'}S^{*'})$; $M^*P^{*'}S^{*'} \neq \emptyset$ if $M^* \neq \emptyset$, or, $I(S^{*'}, P^{*'})$ if $M^* \neq \emptyset$ and after M^* is eliminated.

(3i) (*Pattern and Type Darii*) $E(M^*, P^*)I(M^*, S^*) : M^*S^*P^{*'} \neq \emptyset$; $I(S^*, P^{*'}) [= O(S^*, P^*)]$ - after M^* is dropped from the LC.

(3ii) (*Pattern Darii, Type Disamis*) $I(M^*, P^*)E(M^*, S^*) : M^*P^*S^{*'} \neq \emptyset$; $I(P^*, S^{*'}) [= O(P^*, S^*)]$ — after M^* is dropped from the LC.

By making the substitutions $S^* = M_1, M^* = M_2, P^{*'} = M_3$, where $M^* \in \{M, M'\}, P^* \in \{P, P'\}, S^* \in \{S, S'\}$, all the eight Barbara type syllogisms become the Barbara's pattern 3-term syllogism, showing that all the eight Barbara type syllogisms have the same structure, the differences among them being accounted for by the term changes or replacements $p := P \leftrightarrow P'$, (meaning $p(S^*, P^*, M^*) = (S^*, P^{*'}, M^*)$), $s := S^* \leftrightarrow S^{*'}$, $m := M^* \leftrightarrow M^{*'}$, and their compositions, ps , pm , sm , psm . Add to them the identity transformation, e , and they form a commutative group, G , with eight elements. For example, $spm(S, P, M) = (S', P', M')$, and $s^2 = p^2 = m^2 = e$. The set of eight conclusive syllogisms of type Barbara is left invariant under the action of the group G . The set of eight Darapti type syllogisms is invariant under G , too. By making the substitutions $M^* = M_1, P^{*'} = M_2, S^{*'} = M_3$, where $M^* \in \{M, M'\}, P^* \in \{P, P'\}, S^* \in \{S, S'\}$, all the eight Darapti type syllogisms become the Darapti's pattern 3-term syllogism, showing that all the eight Darapti type syllogisms have the same structure, all the differences among them being accounted for by the transformations of the group G which replaces in all the possible ways the terms M, P, S , by their complementary terms M', P', S' . Similarly, to the Darii's pattern, $M_1M_2 \neq \emptyset, M_1M_3 = \emptyset$, correspond the eight Darii's type syllogisms encompassing Darii, Ferio, Baroco, etc., written as $E(M^*, P^*)I(M^*, S^*) : S^*M^*P^{*'} \neq \emptyset$, where the precise LC is listed after the column sign. By making the substitutions $M^* = M_1, S^* = M_2, P^{*'} = M_3$, where $M^* \in \{M, M'\}, P^* \in \{P, P'\}, S^* \in \{S, S'\}$, all the eight Darii type syllogisms become the Darii 3-term pattern, showing that all the eight Darii type syllogisms have the same structure, the differences among them being accounted for by the differences in the terms employed. To the same Darii's pattern, $M_1M_2 \neq \emptyset, M_1M_3 = \emptyset$, also correspond the eight Disamis type syllogisms, encompassing Disamis, Bocardo, etc., written as $I(M^*, P^*)E(M^*, S^*) : S^{*'}M^*P^* \neq \emptyset$, where the precise LC is listed after the column sign. By making the substitutions $M^* = M_1, P^* = M_2, S^{*'} = M_3$, (which, compared to the Darii's substitutions, also include the $S \leftrightarrow P$ "role switch" between the S and P terms), and where $M^* \in \{M, M'\}, P^* \in \{P, P'\}, S^* \in \{S, S'\}$, all the

eight Disamis type syllogisms become the Darii's pattern 3-term syllogism, showing that all the eight Disamis type syllogisms have the same structure, the differences among them being accounted for by the differences in the terms employed or in the role these terms have: P appears in the universal premise of Darii, but appears in the particular premise of Disamis, etc. Equivalently, one may have observed that any Disamis syllogism becomes a Darii syllogism and vice-versa, via the transformation $S \leftrightarrow P$, which describes a switch in the roles played by the P and S terms in the Darii and Disamis type syllogisms. Note also that instead of using only the formula $(b_0) M_1 \subseteq M_2 \subseteq M_3 \subseteq \dots \subseteq M_n$, to generate a unique Barbara sorites, one could have used the formula $(b'_0) M_1^* \subseteq M_2^* \subseteq M_3^* \subseteq \dots \subseteq M_n^*$, where $M_i^* \in \{M_i, M'_i\}$, $i = 1, 2, \dots, n$, to generate 2^n sorites. One can even apply $n!$ permutations to the n terms $M_1, M_2, M_3, \dots, M_n$ and try to obtain $n!2^n$ Barbara type sorites generated by n terms. All the extra sorites beyond the initial one generated by (b_0) may be considered content variations of the initial sorites, due to either generalizing (b_0) to indefinite terms, or to a changing of the roles played by $M_1, M_2, M_3, \dots, M_n$ in the initial sorites (b_0) .

6 The structure of the 64-element PCP matrix which can replace the moods and figures of Classical Syllogistic

The 64 distinct PCPs obtained by pairing each one of the eight distinct P -premises with each one of the eight distinct S -premises, split into eight groups of eight PCPs per group: the four groups of (1)-Barbara, (2)-Darapti, (3i)-Darii and (3ii)-Disamis type PCPs, plus there are two other PCP patterns — each generating two PCP types per pattern — which do not entail any LC.

The pattern of two particular premises:

(4i) $I(M^*, P^*)I(M^*, S^*)$; **(4ii)** $I(M^*, P^*)I(M^{*'}, S^*)$ — each type contains eight PCPs made of two particular premises, and,

The pattern of one particular and one universal premises, acting one on M and another one on M':

(5i) $E(M^*, P^*)I(M^{*'}, S^*)$; **(5ii)** $I(M^*, P^*)E(M^{*'}, S^*)$ — again, each type contains eight PCPs made of one universal and one particular premises, one “acting on” M and the other on M' , where $M^* \in \{M, M'\}$, $P^* \in \{P, P'\}$, $S^* \in \{S, S'\}$. Since each of the PCP sets (1), (2), (3i), (3ii), (4i), (4ii), (5i), (5ii), is expressed by a formula depending only on S^*, M^* and P^* , it results that each such set is left invariant by the action of the group G. For example, inside the Barbara type conclusive syllogisms, $m(Celarent) = Camestres$, $p(Celarent) = Barbara$, $pm(Camestres) =$

$mp(\text{Camestres}) = \text{Barbara}$. The 64-element PCP matrix may be written as 16 blocks of 2×2 matrices, where inside each block, the four PCPs are all of the same syllogistic type. See *Table 1*.

	$E(M, S^*)$	$E(M', S^*)$	$I(M, S^*)$	$I(M', S^*)$
$E(M, P^*)$	$E(M, P^*)E(M, S^*) :$ $M = MS^*P^*$ <i>Darapti</i>	$E(M, P^*)E(M', S^*) :$ $S^* = MS^*P^*$ $P^* = M'S^*P^*$ <i>Barbara</i>	$E(M, P^*)I(M, S^*) :$ $MS^*P^* \neq \emptyset$ <i>Darii</i>	$E(M, P^*)I(M', S^*)$ (5i)
$E(M', P^*)$	$E(M', P^*)E(M, S^*) :$ $S^* = M'S^*P^*$ $P^* = MS^*P^*$ <i>Barbara</i>	$E(M', P^*)E(M', S^*) :$ $M' = M'S^*P^*$ <i>Darapti'</i>	$E(M', P^*)I(M, S^*)$ (5i)	$E(M', P^*)I(M', S^*) :$ $M'S^*P^* \neq \emptyset$ <i>Darii'</i>
$I(M, P^*)$	$I(M, P^*)E(M, S^*) :$ $MS^*P^* \neq \emptyset$ <i>Disamis</i>	$I(M, P^*)E(M', S^*)$ (5ii)	$I(M, P^*)I(M, S^*)$ (4i)	$I(M, P^*)I(M', S^*)$ (4ii)
$I(M', P^*)$	$I(M', P^*)E(M, S^*)$ (5ii)	$I(M', P^*)E(M', S^*) :$ $M'S^*P^* \neq \emptyset$ <i>Disamis'</i>	$I(M', P^*)I(M, S^*)$ (4ii)	$I(M', P^*)I(M', S^*)$ (4i)

Table 1: The 64-element PCP matrix may be written as 16 blocks of 2×2 matrices, where inside each block, the four PCPs are all of the same syllogistic type.

7 Conclusions

One saw that, based on set inclusions and set intersections, only three distinct patterns, (and four different types), of valid syllogisms or conclusive syllogisms do exist — they represent the $n = 3$ case of the n terms biliteral sorites of the Barbara, Darapti and Darii patterns.

Without ever mentioning the patterns for Barbara, Darapti and Darii n -term sorites, one could still have argued that all the 32 conclusive syllogisms can be reduced to only the Barbara's, Darapti's and Darii's patterns, by using only the group G with 2^3 elements, and the role switch $S \leftrightarrow P$. Quine [18] uses $S \leftrightarrow P$ as a relabeling, which does not change the syllogism's content — one may re-write Darii's

syllogism as a Disamis syllogism via changing the premises' order and by relettering — as Quine calls it — P as S and S as P . But by agreeing that treating the 2^n , (or even $n!2^n$), sorites as being distinct sorites would be a waste of time, one gets the idea that, (after more than 2300 years since Aristotle invented the syllogisms), treating with unchanging awe the 2^3 content variations of each set of the conclusive syllogisms of types Barbara, Darapti, Darii and Disamis, is not such a fruitful idea, either. One may say that in Classical Syllogistic the value of the terms is kept constant — they are always the positive terms S, M, P . Instead, the moods and figures of syllogisms are counted as generating different formats — and names — for the valid syllogisms. This paper's classification of conclusive syllogisms into three patterns and four types, keeps the number of patterns to a minimum and assigns the content variations of the syllogisms to the indefinite terms the syllogisms may employ. Boole [2], Carroll [3], Venn [19] pp.405-406, and especially Reichenbach [12], reached conclusions similar to the ones of the present paper. All these authors thought that, the introductory logic textbooks, should take away the center stage from the valid syllogisms' moods and figures, and hoped that the Classical Syllogistic will be soon presented in a simpler and less unwieldy way — but up to now — nothing changed.

References

- [1] Boole, George: An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities. Macmillan, London (1854). Or: Dover Publications, New York (1958).
- [2] Boole, George: The mathematical analysis of logic, being an essay towards a calculus of deductive reasoning, pp. 34-38. Macmillan, Barclay, & Macmillan, Cambridge (1847).
- [3] Carroll, Lewis: Symbolic Logic, pp. 250, 240. Reprinted. Part I Elementary, 1896. Fifth Edition. Part II, Advanced, never previously published. Edited, with annotations and an introduction, by William Warren Bartley, III. Reprint edition. Clarkson Potter, New York (1986).
- [4] Burris, Stanley: George Boole. The online Stanford Encyclopedia of Philosophy, <https://plato.stanford.edu/entries/boole>.
- [5] Alvarez, Enrique, Correia, Manuel: Syllogistic with Indefinite Terms. History and Philosophy of Logic 33 (4):297-306. (2012). DOI: 10.1080/01445340.2012.680704.
- [6] Read, Stephen, Aristotle's Theory of the Assertoric Syllogism. Available on PhilArchive: <https://philarchive.org/archive/REAAAT0-5>.
- [7] Alex Kruckman, Lawrence S. Moss: Exploring the Landscape of Relational Syllogistic Logics. The Review of Symbolic Logic , Accepted manuscript, pp. 1-35. DOI: <https://doi.org/10.1017/S1755020320000386>.
- [8] van Rooij, Robert. (2012). The propositional and relational syllogistic. Logique et Analyse, 55(217), 85-108. Digital copy: <https://hdl.handle.net/11245/1.353334> or

<http://virthost.vub.ac.be/lnaweb/ojs/index.php/LogiqueEtAnalyse/article/download/1817/1593>.

- [9] Jevons, W. Stanley: The Substitution Of Similars, The True Principle Of Reasoning, Derived From A Modification Of Aristotle 's Dictum. Macmillan And Co., London (1869). Reprint Edition: Wentworth Press (2019).
- [10] Copi, I. M., Cohen, C., McMahon, K: Introduction to Logic. 14th edition. Pearson Education Limited, Harlow, England (2014).
- [11] Hurley, P. J.: A Concise Introduction to Logic. 12th edition. Cengage Learning, Stamford, CT (2015).
- [12] Reichenbach, Hans. The syllogism revised. Philosophy of Science, 19(1):1-16, January 1952. DOI:10.1086/287171.
- [13] Marquand, Allan: On logical diagrams for n terms. Philosophical Magazine (12), 266-270 (1881).DOI: 10.1080/14786448108627104.
- [14] Venn, John: On the Diagrammatic and Mechanical Representation of Propositions and Reasonings. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. Series 5, Vol 9 (No. 59), 1-18 (1880). DOI: 10.1080/14786448008626877.
- [15] Veitch, Edward, W.: A chart method for simplifying truth functions. Proceedings of the Association for Computing Machinery 127-133 (1952). ACM, Pittsburgh. DOI: 10.1145/609784.609801.
- [16] Karnaugh, Maurice: The map method for synthesis of combinational logic circuits. Transactions of the Institute of Electrical Engineers, Part 1. (72), 593-599 (1953). DOI: 10.1109/TCE.1953.6371932.
- [17] Abdalla, Yasser: Introducing the Yasser-map as an improvement of the Karnaugh-map for solving logical problems. In: Proceedings of the First IEEE ICECCT. IEEE, pp. 1-5 (2015). DOI: 10.1109/ICECCT.2015.7226152.
- [18] Quine, Willard Van Orman: Methods of Logic, p. 106. 4th edition. Harvard University Press, Cambridge MA (1982).
- [19] Venn, John: Symbolic Logic. 2nd edition, Macmillan, London (1894). Reprint edition: AMS Chelsea Publishing, Providence RI, pp. 402-406. (2006).

WHAT IS NEGATION IN A SYSTEM 2020?

DOV GABBAY

University of Luxembourg and King's College London

dov.gabbay@kcl.ac.uk

Abstract

The notion of negation is basic to any formal or informal logical system. When any such system is presented to us, it is presented either as a system without negation or as a system with some form of negation. In both cases we are supposed to know intuitively whether there is no negation in the system or whether the form of negation presented in the system is indeed as claimed. To be more specific, suppose Robinson Crusoe writes a logical system with Hilbert type axioms and rules, which includes a unary connective $*A$. He puts the document in a bottle and let it lose at sea. We find it and take a look. We ask: is the connective “ $*$ ” a negation in the system? Yet the notion of what is negation in a formal system is not clear. When we see a unary connective $*A$, (A a wff) together with some other axioms for some additional connectives, how can we tell whether $*A$ is indeed a form of negation of A ? Are there some axioms which the connective “ $*$ ” must satisfy in order to qualify $*$ as a negation?

1 Orientation and overview

The question we are asking now (year 2020) is “what is negation”? The answer is not simple. There are many new negation candidates in the literature such as negation as failure, many paraconsistent negations, negation as inconsistency, as well as the old negations such as classical negation, negation in linear logic, intuitionistic negation, Post negation, etc.

To define what is negation we need to address the following:

1. Give a weakest reasonable definition of a logical system (consequence relation) and address the question of what is negation relative to this notion.

This paper is an expanded version of the 1986 paper [15]. A shorter version of this paper was published as a position paper for new research on the question of What is Negation, for the year 2020, in a Springer Volume in the series, *Outstanding Contributions in Logic*, honouring Arnon Avron. I thank Arnon Avron, Anna Zamansky and Ofer Arieli for reviews and comments.

2. Give an intuitive definition of what is negation and examine for each candidate for negation whether it is indeed a negation.
3. Most importantly, try to show for those candidates which are not negations, what they really are and explain why some communities think that they are negations.
4. Obtain some technical results like what happens to negation in a system when the system is expanded with more axioms and/or more connectives and/or more constants or when it is embedded in another system. Put differently, are there technical aspects influencing the question of what is negation or the answer is purely (a manifestation/implementation of) a conceptual notion.
5. We have already identified that the proof rule of substitution, which is present in many (formulations of) logical systems, causes difficulties. We will allude to this point in due course in this paper.
6. There will be, we hope, a continuation technical paper on this topic.

The answers we offer in this paper are somewhat surprising. Our basic conceptual definition for what is negation is that we need a simple consequence relation between sets or multi-sets of wffs and a single wff of the form

$$A_1, \dots, A_n \models B$$

We need not say how the wff are constructed (the language) or what properties does \models satisfy. We assume $*$ is a unary connective candidate for negation and we assume that intuitively we have a set Θ_* of wffs which we consider as not wanted. We have

- $A_1, \dots, A_n \models *B$ iff for some $\alpha \in \Theta_*$, $A_1, \dots, A_n, B \models \alpha$.

Section 2 discusses properties and examples of this definition. Section 3 outlines the surprising result that negation as failure is indeed a negation. The results are quoted from [24].

Section 4 examines known systems where there is a negation candidate which turns out to be not a negation, especially in the area of paraconsistent negations.

Section 5 discusses future research and the lessons to be learnt from our approach.

2 Negation and consequence

We need to start with a definition of what kind of deductive systems we are going to work with. To choose a definition of a deductive system, we first consider which

candidates for known accepted negations we want to address and how these are presented to us. The main candidates for known negation we consider are classical negation, intuitionistic negation, relevance logic negation, linear Logic negation, Łukasiewicz many valued negation and, last but not least, negation as failure in Logic Programming. Some of these logics are presented as Hilbert systems (such as relevance implication). Some have many representations including Tarski type consequence systems. The best representation from the point of view of answering the question of “What is negation in a deductive system” is for us to look at Tarski systems based on multi-sets.¹

Definition 2.2. Let L be any propositional logical system and let \vdash_L be its provability/consequence relation. We do not specify how L is presented to us, it can be as a Hilbert style system with axioms and rules, or as a natural deduction system or by semantics, etc. The main point is that we have a faithful Tarski style formulation of the provability/consequence relation of L :

$$A_1, \dots, A_k \vdash_L B$$

between a finite multi-set Δ containing the formulas $A_j, j = 1, \dots, k$ and a single B satisfying the following three conditions:

1. $\Delta \vdash A$ for $A \in \Delta$. (reflexivity).
2. If $\Delta \vdash A$ and $\Delta' \supseteq \Delta$ then $\Delta' \vdash A$. (monotonicity).
3. If $\Delta' \vdash A$ and $\Delta \cup \{A\} \vdash B$ then $\Delta \cup \Delta' \vdash B$. (Transitivity, or cut).

¹ The perceptive reader might ask why is it that we are considering "what is negation" in a deductive consequence system, why not present a consequence system semantically?. The answer is not technical but psychological. When the question was considered in 1986 (see [15]), the author had an image of Robinson Crusoe stranded on an island writing a Hilbert System on a sheet of paper, putting it in a bottle, and throwing it into the water. We find it years later and we see a unary connective $*$ in the system and we ask ourselves “Is $*$ a negation in this system?”.

Of course a logic can be presented semantically, but then we can see the intended meaning of the system from the semantics and the challenge is smaller. Consider for example classical logic with the connectives $\{\wedge, \vee, \rightarrow\}$ defined semantically via the traditional truth tables for these connectives. We add a unary symbol “ \neg ” giving it the non-deterministic truth table of Arnon Avron (see [31]), namely:

$$\neg \mathbf{t} = \{\mathbf{f}\} \text{ and } \neg \mathbf{f} = \{\mathbf{t}, \mathbf{f}\}.$$

The consequence relation can be defined semantically.

Question 2.1. *Is this \neg (Avron “negation”) a negation? (We think it is not a negation. See Example 4.8.)*

In fact any relation \vdash on wffs satisfying (1), (2) and (3), can be regarded as a monotonic logical system for sets or multi-sets of data. Note the repetition in rule (3) above for the multi-set case.

Note that in the non-monotonic case [13], condition (2.) above is replaced by condition (2non) (Gabbay called it “Restricted Monotonicity”):

2non. If $\Delta \vdash A$ and $\Delta \vdash B$ then $\Delta \cup \{A\} \vdash B$

Also note that for resource logics, where multi-sets are used (for example, monotonic affine linear logic), we may also wish to investigate the question of what is negation for a consequence relation with condition (3res) instead of condition (3), where we have:

3res. If $\Gamma \vdash A$ and $\Delta \cup \{A\} \vdash B$ then $\Gamma \cup \Delta \vdash B$.

(3res) does not imply (2) for these logics. But note that (3res) implies (3), but the converse is true only if the consequence relation is monotonic, and it is between “sets” and formulas (not multisets).

The fact that we allow Δ or Γ to be a multi-set presents no technical difficulties.²

Our strategy is to give several candidate definitions of what should constitute a negation in a system and test them against our intuitions and against known examples. The examples we look at are as follows:

Example 2.3.

1. Let us consider the following system in a language with \neg and \rightarrow .

² Note that really all we need is to understand, by any precise mathematical- technical means necessary, (proof theoretic, algorithmic, semantic, via translation into another system, via an explicit list/table) the question of when the expression

$$\Delta \vdash A$$

holds.

For Example for the case of relevance implication, in item 2 of Example 2.20, we use a translation into a Hilbert system. I do not know at this stage what axiomatic properties to impose on a Tarski consequence relation in order to make it correspond to the relation obtained from the translation in item 2 of Example 2.20. More future research is required here.

So Definition 2.2 given above is just a very common sample axiomatic definition of a consequence relation. Further note that the author has been claiming for the past 40 years that a logical system should be taken as the declarative set of its theorems as well as an algorithm for demonstrating said theorems. So for example classical logic (perceived as a set of theorems) presented as a Gentzen system is NOT THE SAME LOGIC as classical logic presented via Resolution, which in turn IS NOT THE SAME LOGIC as classical logic presented semantically via Tableaux or truth tables.

- (a) $A \rightarrow (B \rightarrow A)$
- (b) $[A \rightarrow (B \rightarrow C)] \rightarrow [(A \rightarrow B) \rightarrow (A \rightarrow C)]$
- (c) $\neg\neg A \rightarrow A$
- (d) $A \rightarrow \neg\neg A$

Rules

- (e) Modus Ponens

$$\frac{A, A \rightarrow B}{B}$$

Question 2.4. Is \neg to be considered a form of negation in this system (i.e. in item 1 of Example 2.3)?

- 2. Let us make life more difficult by adding more axioms to our system. To get the idea of what to add, first we need disjunctions and conjunctions (the system has only \neg and \rightarrow). So let us see what can be taken as disjunction.

In classical logic (with the connectives $\rightarrow, \neg, \vee, \wedge$, and equivalence \leftrightarrow) we have:

$$(a \rightarrow b) \rightarrow b \leftrightarrow \neg(a \rightarrow b) \vee b \leftrightarrow (a \wedge \neg b) \vee b \leftrightarrow a \vee b.$$

This is in fact a well known definition of \vee in terms of \rightarrow .

Also let $a \wedge b = \text{def.} \neg(\neg a \vee \neg b) = \neg((\neg a \rightarrow \neg b) \rightarrow \neg b)$.

Take the following rule:

- (f)
$$\frac{\vdash A \rightarrow B}{\vdash \neg B \rightarrow \neg A}$$

and the further axioms:

- (g) $\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$
- (h) $\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$
- (i) $((A \rightarrow \neg\neg B) \rightarrow A) \rightarrow A$. (This axiom says $(A \rightarrow B) \vee A$.)

Question 2.5. *Is \neg a negation in this system (i.e. in the system of item 2 of Example 2.3)?³*

3. We can ask further: If we also add the axiom

$$(j) \quad A \rightarrow (\neg A \rightarrow B).$$

Question 2.6. Does the addition of axiom (j) make \neg a negation in the system (of item 3 of Example 2.3)?

(We shall see that answer is no for cases (1) and (2) and yes for case (3).)

It seems from the above Example 2.3 that this question does not have an immediate simple answer. Remember that we cannot just write a set of axioms for negation and say that anything satisfying these axioms is a negation. If we write too many axioms we may get only classical negation, and even that is not guaranteed because maybe we do not know how the negation axioms are supposed to interact with other connectives e.g. with \rightarrow .

Let us look at more examples.

Example 2.7. Consider the system *L3* below of Wajsberg. It axiomatises the 3 valued logic of Łukasiewicz with \rightarrow and \neg .

Axioms:

$$(W1) \quad A \rightarrow (B \rightarrow A)$$

$$(W2) \quad (A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$$

³ Note that axioms (a) and (b), taken together with Modus Ponens define positive intuitionistic implication. To get positive classical implication we need to add Peirce's Rule

$$(P) : ((A \rightarrow B) \rightarrow A) \rightarrow A.$$

Arnon Avron proposed that a better and clearer presentation of the system presented in this item 2 of Example 2.3 would be in a language in which disjunction and conjunction are taken as primitive. The system can then be axiomatised by taking some axiomatisation of positive classical logic which has Modus Ponens as the sole rule of inference, and add to it the axioms (c) and (d), and the rule (f). (Note that if the rule (f) is turned into an axiom, then we get by this a sound and complete system for Classical Logic.) The other items in Example 2.3 can be changed similarly. (That is: in item 1 we take positive intuitionistic logic together with axioms (c) and (d), in item 2 we add (f) and (i), where the latter is taken in a purely positive form, and in item 3 we add (j).

The author prefers the implication based formulation because we need to discuss adding negation as failure to the system.

$$(W3) \quad (\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$$

$$(W4) \quad ((A \rightarrow \neg A) \rightarrow A) \rightarrow A$$

The inference rule is modus ponens.

Question 2.8. Can one determine on the basis of \vdash_{L3} whether $\neg A$ is a negation in $L3$?

Example 2.9. Consider a third system denoted by $LS3$. Its language contains an additional connective Θ besides \neg and \rightarrow . It is obtained from $L3$ by adding the axioms:

$$(\Theta 1) \quad \Theta A \rightarrow \neg \Theta A$$

$$(\Theta 2) \quad \neg \Theta A \rightarrow \Theta A$$

Question 2.10. 1. Is \neg a negation in this system? Is Θ a negation?

2. If \neg is considered a negation in $L3$, does it have to be considered a negation in the extension $LS3$?

Armed with this stock of examples we now move to a formulation and some possible solutions of our problem.

Problem 2.11. *Given a relation \vdash (satisfying (1), (2), (3)) of Definition 2.2 and a connective $*A$ in the language of \vdash , are there any criteria on the relationship between \vdash and $*$ which will agree with our intuitions regarding the question of when $*$ is to be considered a form of negation? Carnap and Church discussed whether a syntactical characterisation of negation was possible. Carnap thought it was possible and Church thought not. A basic intuition regarding the meaning of $\neg A$ is that A does not hold or A is not wanted or A is excluded or even A is not confirmed. Thus if L is a system with a candidate $*A$ for negation, we cannot hope to have A , and its negation $*A$ consistent together (understand “consistent together” intuitively, or maybe “both provable”). This leads us to our first attempt in answering Problem 2.11.*

We must specify a set Θ of unwanted wffs. The wffs of Θ are not allowed to be true (understand “true” intuitively, or maybe “provable”). This is normal and natural for any database. For example we do not want two lecturers to be assigned to the same classroom at the same time. In a formal system L , one can take Θ to be the set containing \perp , i.e. falsity or one can take Θ to be certain conjunctions of atoms, etc.

So to get negation into a system we must have a set of unwanted wffs Θ . This set may be different for different negations. The connective (*1) may be a negation because of Θ_1 and (*2) may be a negation because of Θ_2 , and so on.

We are thus led to the following definition.

Definition 2.12 (Negation as syntactical inconsistency for the monotonic case). Let \vdash be the provability/consequence relation of a system and $*A$ be a connective. We say that $*$ is a form of negation if there is a fixed non-empty set of wffs Θ^* which is not provably equivalent to the set/multi-set of all wffs, such that for any set or multi-set of wffs Δ and any A the following holds:

$$\Delta \vdash *A \text{ iff } \exists y \in \Theta^*(\Delta, A \vdash y).$$

i.e. A is negated by Δ because A leads to some unwanted y in Θ^* .⁴

Lemma 2.13. *Let $*$ be a negation in the logical system \vdash , as defined in Definition 2.12. Then the set $\{x : \emptyset \vdash^* x\}$ is non-empty.*

Proof. Since $*$ is a negation, let $q \in \Theta^*$, then $\emptyset \vdash *q$, since $q \vdash q$.

The above is a purely syntactic (in terms of \vdash) definition. So to check whether $*A$ of an axiom system is a negation, look for a Θ^* and try to prove the above equivalence.

Note that the equivalence must hold for any Δ and A . □

We may ask ourselves, how do we find a Θ^* ? The answer is that if such a Θ^* exists, (i.e. $*$ is a negation according to the above Definition 2.12) then it follows from Lemma 2.13 that Θ^* can be taken as:

$$\Theta^* = \{C \mid \emptyset \vdash *C\}$$

where \emptyset is the empty set.

⁴ If there is disjunction \vee in the language, then note that (for technical reasons) Θ will be closed under disjunctions. We need to check what happens when we are dealing with multi-sets. We expect the differences would be technical, the idea of what is negation would be the same.

Furthermore if we are dealing with multi-sets we may need more copies of Δ , I think we might try

$$\Delta \vdash *A \text{ iff for some } y \in \Theta^* \text{ and some } k, (\Delta \cup \dots (k \text{ times}) \dots \cup \Delta, A \vdash y).$$

This is in the spirit of enhancing the data (via the connective **C**) introduced later in this paper. Note that we can also negate a set/multi-set Γ , namely $\Delta \vdash *\Gamma$ iff for some $y \in \Theta^*$ and some k ; $\Delta \cup, \dots k \text{ times}, \dots \cup \Delta \cup \Gamma \vdash *y$.

Lemma 2.14. *Assume $*$ is a negation with a Θ^* according to Definition 2.12, then for any Δ and any A , (1) is equivalent to (2):*

1. $\Delta, A \vdash C$ for some C such that $\emptyset \vdash *C$.
2. $\Delta, A \vdash B$ for some $B \in \Theta^*$.

Proof. Let $B \in \Theta^*$ then since $B \vdash B$ we get by Definition 2.12 that $\emptyset \vdash *B$. This shows that (2) implies (1).

Assume that for some C such that $\emptyset \vdash *C$, we have $\Delta, A \vdash C$. Since $\emptyset \vdash *C$, we have that for some $B \in \Theta^*$,

$$C \vdash B.$$

then by cut-res (item (3res in Definition 2.2) using $C \vdash B$, we get

$$\Delta, A \vdash B.$$

This proves that (1) implies (2) and we have proved the lemma for the monotonic case. \square

We can thus modify Definition 2.12 as follows:

Definition 2.15 (Definition 2.12 (modified)). Let \vdash be a monotonic logical system and let $*A$ be a connective. We say that $*$ is a form of negation in \vdash iff for any Δ and any A the following holds.

$$\Delta \vdash *A \text{ iff for some } C \text{ such that } \emptyset \vdash *C \text{ we have } \Delta, A \vdash C.$$

The above Definition 2.15 seems theoretically sound and acceptable. All we have to see now is whether it takes care of all the currently known and agreed upon negations.

We will see later that further modifications are necessary. For this reason we continue to use Θ^* itself and not $\{C \mid \emptyset \vdash *C\}$. Note that Θ^* may contain wffs containing $*$ itself (this means that it is built up also by using the negation connective $*$).

We do not need to exclude this possibility. In fact for classical logic we can take Θ^* to be the set $\{q_0 \wedge \neg q_0\}$ for some atom q_0 and we all know that in classical logic $\Delta \vdash \neg A$ iff $\Delta, A \vdash q_0 \wedge \neg q_0$ holds, and so classical negation is a negation. So is intuitionistic negation because the same equivalence holds.

According to Definition 2.12, the \neg defined in Example 2.3 axioms (a) to and including axiom (i), i.e. Question 2.6 is not a negation. One can see this by taking the following interpretation and verifying that all axioms (a) to and including axiom

(i) of Example 2.3 are valid. In this interpretation there are two worlds h and e (heaven for h and earth for e). $\neg A$ is *true* in one if A is *false* in the other. \rightarrow is the usual truth functional implication. All axioms and rules are valid; i.e. we have

$$\vdash A \text{ iff } A \text{ is true in } e \text{ and } h \text{ under any assignment to the atoms.}^5$$

Now we can see that $\neg A$ is not a negation of A , since it just says that A is false in the other world. $A \wedge \neg A$ can be consistent, as A could be true in this world (e.g. e) and false in the other world (e.g. h).

The rule of Definition 2.12 for negation does not apply here. If \neg were a negation, then for some Θ , and for all Δ, A we would have:

$$\Delta \vdash \neg A \Leftrightarrow \Delta, A \vdash y, \text{ for some } y \in \Theta.$$

In particular for any $y \in \Theta$ we get $\vdash \neg y$. Let p be atomic then since

$$\neg p \vdash \neg p$$

we get $\neg p, p \vdash y$ for some $y \in \Theta$ and therefore we get that

$\neg p \wedge p \vdash y$ for some $y \in \Theta$, and hence by definition $\vdash \neg(p \wedge \neg p)$. This means that $p \wedge \neg p$ is false in every model.

⁵ Additional axioms may be needed for this assertion. If we just add the connective \neg to the language of intuitionistic implication we are simply generating repeatedly/recursively new atoms of the form $\neg A$ for any already generated A , using all the wffs of intuitionistic implication as a basis.

- The axiom $\neg\neg A = A$, says the generating is idempotent.
- The axiom $\neg(A \Rightarrow B) = (A \Rightarrow \neg B)$, if added, takes us in the direction of \neg being negation as failure.
- Adding Peirce's rule, i.e. basing the addition of \neg above on classical implications, takes us to the semantics with e and h .

Let me quote Arnon Avron's comment to me, as follows:

I can easily see the "only if" part (i.e. soundness). The converse is not obvious. Have you proved it somewhere? If so, you should add a reference. If not, you should give a proof here.

This point is very important, since your argument for the claim that \neg becomes classical negation depends (so it seems to me) on the completeness part of the above "iff"!

I believe that you can avoid the above problem if you give a direct, syntactic derivation of $\neg A \vee A$ (i.e. excluded middle) in the system given in item 2 of Example 2.3. (This is very easy if you follow my suggestion in Footnote 3) The reason is that it is well known that a complete axiomatization of Classical Logic is obtained by adding to CL+ both (j) and excluded middle. (See P. 27 of our book [3].)

Since p is an atom we cannot have the above since $p \wedge \neg p$ is consistent, meaning that it has a model, for example if p is true at (e) and false at (h) (we can give this assignment since p is atomic) then $p \wedge \neg p$ holds at (e). Thus the \neg above is not a negation according to Definition 2.15.

Turning now to Question 2.6, we add axiom (j), (of Example 2.3) i.e. $A \rightarrow (\neg A \rightarrow B)$ we get $e = h$ and \neg becomes classical negation. We can take $\Theta = \{(\neg q_0 \wedge q_0)\}$ and derive from the axioms that

$$\vdash \neg A \leftrightarrow (A \rightarrow (\neg q_0 \wedge q_0)).$$

In fact the above additional axiom says simply $\vdash \neg(A \wedge \neg A)$.

Let us check now whether Question 2.8, namely whether \neg in the system $L3$ of Example 2.7 is indeed a form of negation. This system axiomatises Łukasiewicz 3 valued logic. There are 3 truth values, 1 (*truth*), $\frac{1}{2}$, and 0 (*falsity*). The truth tables for \neg and \rightarrow are as follows:

$$\neg x = 1 - x \text{ and } x \rightarrow y = \min(1, 1 + y - x).$$

The idea of the definition for $x \rightarrow y$ is that if $x \leq y$ then $x \rightarrow y$ is true. (Like $0 \rightarrow 1$ in classical logic.) If $x > y$ then $x - y$ is the measure of falsity of $x \rightarrow y$ and so the value of $x \rightarrow y$ is $1 - (x - y)$. $\neg x = 1 - x$ is just the mirror image of the truth value.

Conjunction $x \wedge y$ and disjunction $x \vee y$ have the definition below. They are definable from \rightarrow by:

$$\begin{aligned} x \vee y &= \text{def.}(x \rightarrow y) \rightarrow y = \min(x, y). \\ x \wedge y &= \text{def.}\neg(\neg x \vee \neg y) = \max(x, y). \end{aligned}$$

Intuitively, there is no doubt that $\neg x$ is a form of negation in this system, because $\neg x = 1 - x$. The farther x is from the truth the nearer $\neg x$ is to the truth.

Remark 2.16. *The consequence relation for this logic can be defined in two ways, for multi-sets $\Delta = \{A_1, \dots, A_n\} \models B$:*

Option 1. *We can write $A_1, \dots, A_n \models_1 B$ in this system to mean that under any assignment: $\text{Min}(\text{value } A_j) \leq \text{val } B$, and $\models B$ to mean that under any assignment $\text{val } B = 1$.*

Notice that the relation \models_1 , defined semantically above, fulfils the criteria for a logical system. The deduction theorem, however, is not valid for \models .

The Wajsberg axiom system is complete in the sense that the following holds:

$$A_1, \dots, A_n \models B \text{ iff } \vdash \bigwedge A_j \rightarrow B.$$

If we define $A_1, \dots, A_n \vdash B$ to mean that $\vdash \bigwedge A_j \rightarrow B$ it then follows that $\vdash B$ iff $\text{val } B = 1$ under all assignment.

Option 2. *There is another possibility of deriving /attaching a consequence relation to the axioms of L3. We can let databases be multi-sets and let $A_1, \dots, A_n \vDash_2 B$ to mean that*

$$\text{Max}(0, 1 - \Sigma(1 - \text{Value}(A_i))) \leq \text{Value}(B).$$

We can choose an appropriate \mathbf{C} for each case.

For multi-sets we can take the formula $\mathbf{C}(x, y) = \text{def. } \neg(x \rightarrow \neg y)$, which satisfies the equation:

$$\text{Value}(\mathbf{C}(x, y)) = \max(0, \text{Value}(x) + \text{Value}(y) - 1).$$

This is an enhancement over conjunction $x \wedge y$, which has the value $\min(x, y)$.⁶

Our Definition 2.12 of what a negation is should give us that \neg is a negation. Suppose \neg is indeed a negation according to Definition 2.12. Then there exists a fixed Θ such that for any Δ and any A of the logic L3 we have:

$$\Delta \vdash \neg A \text{ iff } \Delta, A \vdash B \text{ (for some } B \in \Theta). \text{ Necessarily } \Theta \neq \emptyset.$$

Take any $B \in \Theta$ and $\Delta = \emptyset$ then $\vdash \neg B$ iff $B \vdash y$ for some $y \in \Theta$; but since $y = B \in \Theta$ and $B \vdash B$ we get $\vdash \neg B$ for all $B \in \Theta$.

One can verify by looking at the axioms the following lemma:

Lemma 2.17. *If $\vdash A$ then value $(A) = 1$ under all assignments.*

Proof. The above is true for the axioms and is preserved under modus ponens and substitution. □

Thus we conclude that for any $B \in \Theta$, value $B = 0$ under all assignments. Now consider an atom q , certainly

$$\neg q \vdash \neg q$$

⁶Arnon Avron commented as follows:

What you call $C(x, y)$ is known as the t -norm that underlies Łukasiewicz logic, and is usually denoted by $\&$. (See Hajek's book on fuzzy logics [27]). Your second Consequence Relation can be characterized as follows:

$$A_1, \dots, A_n \vdash B \text{ iff } v(A_1) \& v(A_2) \& \dots \& v(A_n) \leq v(B)$$

for every valuation v .

A remark: an option you have not mentioned here is the standard one: $A_1, \dots, A_n \vdash B$ iff $v(B) = 1$ for every valuation v s.t. $v(A_i) = 1$ for every i . This option is closely related, of course, to Lemma 2.17.

hence for some $B \in \Theta$,

$$\neg q \wedge q \vdash B$$

hence under all assignments $\text{Min}(\text{value } \neg q, \text{value } q) \leq \text{value } B$. In particular for any assignment h with $h(q) = \frac{1}{2}$. This contradicts the previous conclusion that value $B = 0$ always.

We therefore need to improve our Definition 2.12 of negation.

Our basic idea in defining negation was that $A \vdash \neg B$ holds if A, B together lead to some undesirable result Θ .

$$\text{i.e. } A, B \vdash \Theta.$$

However the way the above is written is that A and B are “combined” together via conjunction, i.e. $A \wedge B$. It is quite possible that A, B can be combined together via a different connective e.g. some connective $\mathbf{C}(A, B)$. Thus $A \vdash \neg B$ holds iff $\mathbf{C}(A, B) \vdash \Theta$. \mathbf{C} is a connective which “brings out” the effect A and B can have together. Of course $\mathbf{C}(x, y)$ is not an arbitrary connective. It must be monotonic and satisfy some obvious properties. $\mathbf{C}(x, y)$ must say more than just $x \wedge y$, and satisfy the conditions listed in Definition 2.18 for it.

Definition 2.18 (Negation as a potential syntactic inconsistency). Let L be a system with a provability relation \vdash and let $*$ be a unary connective of L . We say $*$ is a form of negation in L iff there exist a non empty set of wffs Θ which is not provably equivalent to the set of all wffs, and a binary connective $\mathbf{C}(x, y)$ s.t. the following holds for any D and A :

$$D \vdash^* A \text{ iff } \mathbf{C}(D, A) \vdash y \text{ for some } y \in \Theta.$$

\mathbf{C} must satisfy the following: (*truth* is any provable formula; such formulae exist if $*$ is a negation. See Lemma 2.13).

1. $\mathbf{C}(x, y) \vdash x$
2. $\mathbf{C}(x, y) \vdash y$
3. $\mathbf{C}(\text{truth}, y) = \mathbf{C}(y, \text{truth}) = y$
4. $\frac{x \vdash x'}{\mathbf{C}(x, y) \vdash \mathbf{C}(x', y)} \quad \frac{y \vdash y'}{\mathbf{C}(x, y) \vdash \mathbf{C}(x, y')}$

where $A = B$ abbreviates $A \vdash B$ and $B \vdash A$.

Think of \mathbf{C} as enhanced conjunction.

Remark 2.19.

1. We get from the above that (in case that a falsity can be defined in the logic, with falsity $\vdash A$, for any A):

$$\mathbf{C}(\text{falsity}, y) = \mathbf{C}(x, \text{falsity}) = \text{falsity}.$$

2. Definition 2.18 was given for D a single formula, if L has conjunction then we can take $\Delta \vdash^* A$ as $\bigwedge \Delta \vdash^* A$. See, however, Option 2 of Remark 2.16.

For our negation in the system $L3$, let $\mathbf{C}(x, y) = \neg(x \rightarrow \neg y)$, and let $\Theta = \{\text{falsity}\} = \{\neg(y_0 \rightarrow y_0)\}$. Clearly, by the definition of \vdash , $x \vdash \neg y$ iff $\vdash x \rightarrow \neg y$ iff value $(x \rightarrow \neg y) = 1$ in all assignments, iff value $\neg(x \rightarrow \neg y) = 0$ in all assignments, iff $\neg(x \rightarrow \neg y) \vdash \text{falsity}$.

The truth table for $\mathbf{C}(x, y) = \neg(x \rightarrow \neg y)$ is $\text{Max}(0, \text{value } x + \text{value } y - 1)$.

As can be seen, since the truth function of $\mathbf{C}(x, y)$ is $\text{Max}(0, x + y - 1)$.

We get:

1. $\mathbf{C}(x, y) \leq x$
2. $\mathbf{C}(x, y) \leq y$
3. $\mathbf{C}(1, y) = y = \mathbf{C}(y, 1)$
4. (a) $x \leq x' \Rightarrow \mathbf{C}(x, y) \leq \mathbf{C}(x', y)$
(b) $y \leq y' \Rightarrow \mathbf{C}(x, y) \leq \mathbf{C}(x, y')$

These correspond to the conditions of Definition 2.18, and hence \neg in the 3 valued logic is a negation. In fact the above definitions of \neg , \rightarrow and $\mathbf{C}(x, y)$ as $\neg(x \rightarrow \neg y)$ show that \neg is a negation in all Lukasiewicz many valued logics.

Now that we have changed the definition of negation in a formal system we have to check whether the \neg of Question 2.6, i.e. of item 2 of Example 2.3 is still not considered a negation. So assume that \neg is a negation in the system of Example 2.3, the system with axioms (a) to (i). Then for some Θ and \mathbf{C} the condition of Definition 2.18 holds, namely for all D, A

$$D \vdash \neg A \text{ iff } \mathbf{C}(D, A) \vdash B \text{ for some } B \in \Theta.$$

We shall show that

$$(\dagger) \quad A \rightarrow \neg(x \rightarrow x) \vdash \neg A$$

using \mathbf{C} and Θ , and this is impossible because in our two world model (\dagger) says that if A is false in one world, A is false in the other world also. Thus if we prove (\dagger) then this shows that no \mathbf{C} , Θ can exist and \neg is not a form of negation.

We now proceed to prove (\dagger) :

Since $\mathbf{C}(y, z) \vdash y \wedge z$ we get

$$\mathbf{C}(A \rightarrow \neg(x \rightarrow x), A) \vdash \neg(x \rightarrow x).$$

Hence by definition of \neg

$$\mathbf{C}(\mathbf{C}(A \rightarrow \neg(x \rightarrow x), A), x \rightarrow x) \vdash B, \text{ for some } B \in \Theta.$$

Since $x \rightarrow x$ is *truth* and $\mathbf{C}(y, \text{truth}) = y$ we get:

$$\mathbf{C}(A \rightarrow \neg(x \rightarrow x), A) \vdash B, \text{ for some } B \text{ in } \Theta$$

and hence by definition of \neg we get:

$$(\dagger) \quad A \rightarrow \neg(x \rightarrow x) \vdash \neg A.$$

Example 2.20 (the system of relevant logic R). Consider a language with \rightarrow only and the following set of axioms and rules, defining the system $R \rightarrow$.

Rule. modus ponens

$$\frac{\vdash A, \vdash A \rightarrow B}{\vdash B}$$

Axioms.

R1: $A \rightarrow A$

R2: $(A \rightarrow B) \rightarrow ((C \rightarrow A) \rightarrow (C \rightarrow B))$

R3: $(A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$

R4: $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$

The above system was introduced by Church and Moh [10] and [30] respectively. Church called it “weak positive implicational calculus”. They proved the following deduction theorem for the system.

1. *Deduction theorem for $R \rightarrow$* : If there exists a proof of R from A_1, \dots, A_n in which all A_1, \dots, A_n are used in arriving at B then there exists a proof of $A_n \rightarrow B$ from A_1, \dots, A_{n-1} satisfying the same conditions.⁷

The above calls for the following definition of $\vdash_{R \rightarrow}$

2. *Definition of $\vdash_{R \rightarrow}$*

$$A_1, \dots, A_n \vdash_{R \rightarrow} B \text{ iff} \\ \vdash_{R \rightarrow} A_1 \rightarrow (A_2 \rightarrow \dots \rightarrow (A_n \rightarrow B) \dots)$$

One can see by axiom R3 that the above is independent of the order of $\{A_j\}$. The above system is identical with the implicational relevance logic of Anderson and Belnap. It does not satisfy the conditions of a logical system, but see however, Footnote 2. Negation \neg is introduced into $R \rightarrow$ to obtain $R(\rightarrow, \neg)$, via the Ackermann negation axioms. These axioms are used to introduce negation not only into $R \rightarrow$ but also into all neighbouring systems.

Ackermann axioms for negation

$$\text{AN1: } (A \rightarrow \neg B) \rightarrow (B \rightarrow \neg A)$$

$$\text{AN2: } (A \rightarrow \neg A) \rightarrow \neg A$$

$$\text{AN3: } \neg\neg A \rightarrow A$$

The following can be proved

$$\text{AN4: } A \rightarrow \neg\neg A$$

$$\text{AN5: } (A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$$

See Anderson and Belnap [1, pp. 20–21, 107–109] for details. The above definition of negation is indeed negation according to our Definition 2.18 of negation. Meyer [28] has shown that if we add to $R \rightarrow$ a symbol f (*falsity*) with the additional axiom

$$\text{R5: } ((A \rightarrow f) \rightarrow f) \rightarrow A$$

we get a system equivalent to $R(\rightarrow, \neg)$, with \neg via the interpretation

⁷ Here we need to take the databases Δ as multi-sets. In the Hilbert type formulation this is hidden. When we say that $(A \rightarrow B) \rightarrow (A \rightarrow (A \rightarrow B))$ is not a theorem of $R \rightarrow$, and use the deduction theorem we get that $\{A \rightarrow B, A, A\}$ does not prove B .

1. $\neg A = \text{def. } A \rightarrow f$.

The following must be proved.

2. $D \vdash \neg A$ iff $D, A \vdash f$
i.e. $D \vdash A \rightarrow f$ iff $D, A \vdash f$

or equivalently by definition of $\vdash_{R \rightarrow}$

$$\vdash D \rightarrow (A \rightarrow f) \text{ iff } \vdash_{R \rightarrow} D \rightarrow (A \rightarrow f), \text{ which is correct.}^8$$

Remark 2.21. *Technically, if the system $R(\rightarrow, \neg)$ of Example 2.20 is formulated with \neg and without f , can we find an f such that $\neg A = A \rightarrow f$? In classical logic one can take $f = q_0 \wedge \neg q_0$ or if conjunction is not available, one takes $f = \neg(q_0 \rightarrow q_0)$ for some fixed q . We cannot do the same for $R(\rightarrow, \neg)$, because if we take $f = \neg(q_0 \rightarrow q_0)$ for some fixed atom q_0 , we will not have enough axioms on \rightarrow to be able to use f as needed. We will have to add axiom $R4$ for this new $f = \neg(q_0 \rightarrow q_0)$ and then show that no new theorems can be proved for any wffs not containing q_0 . Thus we see that Definition 2.18 is not quite right in the sense that the system considered may be too weak to show that it has a negation. In other words a connective $*$ may indeed be a negation in the system \vdash , but \vdash may be too weak to prove the Definition 2.18. In fact, a connective $\mathbf{C}(x, y)$ required by Definition 2.18 may not be definable in the language of the system, but only in an extension. Intuitively if $*$ is a negation in a conservative extension, then we can and should regard it a negation in the system itself. We are thus led to the following definition:*

Definition 2.22 (An improved version of Definition 2.18).

1. Let $L1$ and $L2$ be logical systems such that the language of $L2$ extends the language of $L1$.

We say $L2$ is a conservative extension of $L1$ iff the following holds for any Δ, A in the language of $L1$

$$\Delta \vdash_{L1} A \text{ iff } \Delta \vdash_{L2} A.$$

⁸ Arnon Avron commented as follows:

It is not difficult to show that in item 2 above, we can substitute $\neg\neg[(D \rightarrow D) \rightarrow \neg(A \rightarrow A)]$ for f . Since $\neg[(D \rightarrow D) \rightarrow \neg(A \rightarrow A)]$.

This is provable in $R \rightarrow$ for every A, D , you may take Θ as $\{C | \emptyset \vdash \neg C\}$, and there is no need to extend the language in this case.

On the other hand, it is not clear what is $C(x, y)$ in this example, even if you add f ! It should satisfy 1–4 of Definition 2.18. However, since in Definition 2.18 we took “truth” to be *any* provable formula, none of the obvious candidates seems to work.

2. We say that $*$ is a negation in $L1$ iff for some conservative extension $L2$ and some Θ and \mathbf{C} in $L2$ satisfying the conditions of Definition 2.18, we have that for any D, A of $L1$ the following holds:

$$D \vdash^* A \text{ iff } \mathbf{C}(D, A) \vdash_{L1} B, \text{ for some } B \in \Theta.$$

We have now to check whether this new definition of negation turns the connective \neg of Question 2.6, i.e. item 2 of Example 2.3 axioms (a) to (i) inclusive into a negation. (Recall that we found that \neg is not a negation). The answer is no: \neg is still not a negation. The reason is that it can be proved that (think of the consequence semantically) for any conservative extension of the system in Example 2.3 the two world interpretation (with the e world and the h world) is still valid. So the argument for showing that no \mathbf{C} and Θ can make \neg into a negation still goes through.

Example 2.23. We now give another example illustrating the need for Definition 2.22. Consider the language of classical propositional logic and its consequence relation \vdash .

Let \vdash_1 be defined as

$$\begin{aligned} \Delta \vdash_1 A &\text{ iff } \Delta \neq \emptyset \text{ and } \Delta \vdash A. \\ \vdash_1 &\text{ is a consequence relation} \end{aligned}$$

However, \neg is not a negation in \vdash , according to Definition 2.12, since for any non-empty Θ that we choose we would have to have for $B \in \Theta$ that $\emptyset \vdash_1 \neg B$ since certainly $B \vdash_1 B$ contrary to definition of \vdash_1 . But this is counter intuitive since certainly $\Theta = \{q \wedge \neg q\}$ should be acceptable.

The example (which was suggested by the referee) is certainly pathological and Definition 2.22 handles it nicely. However in our view a more satisfactory solution to this particular problem is to require the following additional property to be fulfilled by a consequence relation.

4. $\Delta \vdash A$ iff $\forall x(\Delta, x \vdash A)$. (Coherence).

We now investigate the possibility that there might be negations for which Θ depends on D . This is quite intuitive, since it says that what we do not want, Θ , depends on the data, D , which we have. This is the case for the negation as failure in Logic programming, as shown by Gabbay in [21, Section 4]. Of course logic programming does not satisfy coherence. In fact it turns out that we *cannot* have a notion of Θ dependent on D , for a coherent consequence relation. (See Example 2.23 above).

Proposition 2.24. *Let \vdash be a monotonic logical system with conjunction \wedge and a negation \neg characterised by the following clauses:*

1. *For any D there exists $\Theta(D)$, dependent on D , such that for any A the following hold:*
2. *$D \vdash \neg A$ iff $\exists y \in \Theta(D)(D, A \vdash y)$*
3. *$\Delta \vdash A$ iff $\forall x(\Delta, x \vdash A)$.*

Then there exists an N (independent of D) such that (1) holds, (i.e. $N = \Theta(D)$.)

Proof. We prove Proposition 2.24 by means of two Lemmas.

Proof of Proposition 2.24 Part 1: Two Lemmas

Lemma 2.25. *Let \vdash, \neg and $\Theta(D)$ be as in Proposition 2.24. Let $N(D)$ be the set*

$$N(D) = \{y \mid D \vdash \neg y\}$$

then \neg is a negation satisfying equation (2) of Lemma 2.24 with $N(D)$ as a set of unwanted sentences.

Proof. Very much as in Lemma 2.14, we show that, for any D and A :

- $\exists y \in \Theta(D)(D, A \vdash y)$ iff $\exists z \in N(D)(D, A \vdash z)$
1. Assume $D, A \vdash y$, for some $y \in \Theta(D)$.

By (2) of Proposition 2.24 we get that $D \vdash \neg A$ and hence $A \in N(D)$ and therefore there exists a $z \in N(D)$, namely $z = A$ such that $D, A \vdash z$.

2. Assume $D, A \vdash z$, for some $z \in N(D)$.

Since $z \in N(D)$ we therefore have that $D \vdash \neg z$. Hence by (2) of Lemma 2.24 again, there exists a $y \in \Theta(D)$ such that $D, z \vdash y$. We now have:

$$D, A \vdash z \text{ and } D, z \vdash y$$

and by the cut rule (3res) we get

$$D, D, A \vdash y.$$

This completes the proof of Lemma 2.25.

Note that the proof in part (2) above can be modified to show that $D, A \vdash B$ and $D, A \vdash \neg B$ implies $D, D \vdash \neg A$. \square

Remark 2.26. *We draw several conclusions from Lemma 2.25:*

1. *First that if \neg is indeed a negation dependent on D (via $\Theta(D)$) then equation (2) of Lemma 2.24 is really an uninformative tautology. By Lemma 2.25, $\Theta(D)$ can be taken as $N(D) = \{y \mid D \vdash \neg y\}$ and equation (2) of Lemma 2.24 becomes:*

$$D \vdash \neg A \text{ iff } \exists y (D \vdash \neg y \text{ and } D, A \vdash y)$$

which is trivially true for $y = A$.

Note that for the case where Θ was fixed (independent of D) we got that $D \vdash \neg A$ iff $\exists y (\vdash \neg y \text{ and } D, A \vdash y)$ which is more informative.

2. *The second conclusion is that Θ is dependent on D in a special way.*

As D gets stronger, Θ increases. This is not intuitive! Why should (a priori) what we do not want increase with the database?

This property follows since we have:

$$\frac{D' \vdash D, D \vdash \neg A}{D' \vdash \neg A}$$

3. *The third conclusion follows from the proof of Lemma 2.25 and the assumption (3) of Lemma 2.24 .*

We get the following for \neg :

$$(c1) \frac{D, A \vdash B; D, A \vdash \neg B}{D, D \vdash \neg A \{A, A\}}$$

Furthermore, since we saw in (2) that $D' \vdash D \Rightarrow N(D') \supseteq N(D)$ we can get that (see Footnote 4, and read $A \wedge B$ as $\{A, B\}$). Thus adding \wedge is always conservative):

$$(c2) \frac{D \vdash \neg A}{D \vdash \neg(A \wedge B)}$$

The reason is that if $D, A \vdash y, y \in N(D)$, then certainly $D, B, A \vdash y$ and since $D, B \vdash D$, we have $y \in N(D, B)$ and hence $D \vdash \neg(A \wedge B)$.

We now proceed to use Lemma 2.25 to prove Proposition 2.24 namely that \neg can be taken to be a negation with a fixed Θ (independent of D). We assumed that the language contains conjunction \wedge . \wedge satisfies the three axioms:

$$\begin{aligned} A \wedge B &\vdash A \\ A \wedge B &\vdash B \\ A, B &\vdash A \wedge B. \end{aligned}$$

We proceed now to the second Lemma:

Lemma 2.27. *Let \vdash be a system with negation \neg . satisfying the rule:*

1.
$$\frac{D, A \vdash B; D, A \vdash \neg B}{D \vdash \neg A}$$

Then for $N = \{B \wedge C \mid B \vdash \neg C\}$ we have for any D, A
2. $D \vdash \neg A$ *iff* $\exists y \in N(D, A \vdash y)$.

Proof.

1. Assume $D \vdash \neg A$. We are looking for a y such that $y \in N$ and $D, A \vdash y$. Let $y = D \wedge A$. Certainly $D, A \vdash D \wedge A$ and $D \wedge A \in N$ since $D \vdash \neg A$.
2. Assume that for some $y \in N$, we have $D, A \vdash y$. y is then equal to some $B \wedge C$ with $B \vdash \neg C$. Since $D, A \vdash B \wedge C$ we get $D, A \vdash C$. Since $B \vdash \neg C$ we get $D, A \vdash \neg C$ and hence by rule 1, $D \vdash \neg A$.

□

Part 2 of the proof of Proposition 2.24: Having proved our two Lemmas (2.25 and 2.27) we can proceed. Assume the conditions of Proposition 2.24 for \vdash and \neg hold. By conclusion (c1) of item 3 of Remark 2.26 the conditions of Lemma 2.27 hold and hence \neg is a negation with a fixed $\Theta = N$. ■

The above considerations show that there is no hope for a formulation of a negation \neg with a Θ dependent on the database, within the framework of monotonic logics. The assumption that \wedge is available does not restrict generality since \wedge can always be added to the language and Definition 2.22 for negation be used. ■

3 Negation and failure

We give examples from other papers to show that negation as failure is negation in our sense. We give no proofs. It is too complicated for our current paper which is essentially a position paper, see [14, 24].

3.1 Intuitionistic implication

We need to start with a convenient formulation of a proof system and semantics for intuitionistic implication.

Intuitionistic implication can be presented as a Hilbert system using axioms (a) and (b) and rule MP in Example 2.3. However, we want to add negation as failure to intuitionistic implication and therefore we need to define the notion of "failure".

To achieve this we need to present a proof system for intuitionistic implication algorithmically, so the algorithm can fail.

We make use of the observation [12] that for the case of intuitionistic implication, the proof theory and the Kripke semantics are the same. Definition 3.1 explains this point.

Definition 3.1.

1. A Kripke model for intuitionistic implication has the form $(S, R, 0, h)$ where S is a multiset of possible worlds, $o \in S$ is the actual world and R is a reflexive and transitive relation on S . h is an assignment giving for each atomic q a subset $h(q) \subseteq S$ such that the following holds:

(*) $t \in h(q)$ and tRs implies $s \in h(q)$.

2. Define semantic satisfaction \models of a wff A at $t \in S$ as follows

(a) $t \models q$ iff $t \in h(q)$, for q atomic.

(b) $t \models A \rightarrow B$ iff for all s such that tRs and $t \models A$, we have that $s \models B$.

(c) A holds in the model if $0 \models A$.

3. Consider the following specific generic Kripke model.

(a) Let $S =$ the family of all finite multisets of formulas of the logic. Provability \vdash is the same as for sets.

(b) Let $0 = \emptyset$

(c) Define for atomic q and $\Delta \in S$

$$h(q) = \{\Delta \mid \Delta \vdash q\}.$$

(d) Let R be defined on S as the sub-multiset relation.

The following theorem holds.

Theorem 3.2.

1. Intuitionistic logic is complete for the Kripke semantics in Definition 3.1.

2. In the specific generic model of item (3) of Definition 3.1 we have for any Δ and any wff A :

(**) $\Delta \models A$ iff $\Delta \vdash A$.

The above correspondence (**) in Theorem 3.2 shows that the semantics “ \models ” is essentially proof theoretical tableaux “ \vdash ”. Therefore whatever proof theoretical algorithm we would give to \vdash failure of the algorithm of the query/goal A from any Δ means $\Delta \not\models A$.

The Deduction Theorem for \vdash implies that

(***) $\neg(A \Rightarrow q)$ is equivalent to $A \Rightarrow \neg q$,

where \Rightarrow is intuitionistic implication and \neg is negation as failure.

This is because

$$\Delta \not\models A \Rightarrow q \text{ iff } \Delta, A \not\models q.$$

We are now facing two problems in giving semantics to “ \neg ”.

First when we have embedded negation as failure and in the semantics, the database changes, we need to say from which database the failure is.

So “ $\neg A$ ” becomes “ $\neg_{\Gamma} A$ ” reading “ A fails from Γ ”.

Second, we also know that in intuitionistic logic an item of data may be used more than once. So if it contains negation “ \neg ”, then it may be used relative to different databases. So we cannot give “ $\neg A$ ” a unique meaning.

To solve this problem we need to assume every item of data containing a negation $\neg A$ is used at most once, and if this item is used again, we duplicate it and use the duplicate (this is the reason why we use multi-sets for databases).

So each item of data needs two annotations. One annotation says whether it is in the database or is going to be added later and the second annotation is how many duplicate copies of it are needed.

The next Definition 3.3 and subsequent definitions give databases for intuitionistic implication a new presentation which allows for a Prolog like algorithm. We use the notation “ \Rightarrow ” for intuitionistic implication (and not “ \rightarrow ”) to conform to our Logic Programming notation “ \Rightarrow ”.

Definition 3.3. *Consider a language with a set of propositional atoms Q and the binary connective \Rightarrow . We define recursively the notions of a goal, body, clause and database.*

We use multisets in our definitions to allow us a more refined control of resource considerations. The need for multisets will become apparent later when we give semantics to negation as failure.

1. *An atom $q \in Q$ is a goal.*
2. *The empty multi-set \emptyset is a database.*
3. *If Δ is a database and q is a goal then $\Delta \Rightarrow q$ is a clause.*

4. Any multi-set of clauses is a database.
5. If Δ_1 and Δ_2 are databases. So is $\Delta_1 \cup \Delta_2$.

Definition 3.4. Let Δ be a database. We define $Sub(\Delta)$ the database of all clauses which appear in any of the clauses of Δ as follows:

1. $Sub\{(\emptyset \Rightarrow q)\} = \{(\emptyset \Rightarrow q)\}$
2. $Sub(\Delta_1 \cup \Delta_2) = Sub(\Delta_1) \cup Sub(\Delta_2)$
3. $Sub((\Delta \Rightarrow q)) = \{(\Delta \Rightarrow q)\} \cup \bigcup_{C \in \Delta} Sub\{C\}$.

Example 3.5. Let us compute $Sub\{c \Rightarrow a, (c \Rightarrow a) \Rightarrow c\}$. It is $= \{(c \Rightarrow a)\} \cup \{(c \Rightarrow a) \Rightarrow c\} \cup \{c\} = \{c \Rightarrow a, (c \Rightarrow a) \Rightarrow c, c\}$.

Definition 3.6.

1. A signed clause has the form $\pm(\Delta \Rightarrow q)$. The “+” sign means the clause is active. The “-” sign means the clause is not active.
2. A signed linear logic database is a multi-set of signed clauses.
3. Let Δ be a multi-set database. Let the signed $Sub(\Delta)$ be the (multi-set) database as follows:
 - If $C \in \Delta$ then $+C \in Signed\ Sub(\Delta)$.
 - If $C \in Sub(\Delta) - \Delta$, then $-C \in Signed\ Sub(\Delta)$.

Example 3.7. In Example 3.5 the signed sub-database is

$$\{-c, +(c \Rightarrow a), +((c \Rightarrow a) \Rightarrow c)\}.$$

Definition 3.8.

1. Let Δ be a database and let $\Delta_i \Rightarrow q_i, i = 1, \dots, m$ be all the clauses in Δ . Let α_Δ be an annotation function, giving each clause $\Delta_i \Rightarrow q_i$ in Δ a natural number ≥ 0 , denoted by $\alpha_\Delta(\Delta_i \Rightarrow q_i)$.

We can also present the annotated database $(\alpha_\Delta : \Delta)$ in the form:

$$\Delta = \{k_1 : (\Delta_1 \Rightarrow q_1), \dots, k_m : (\Delta_m \Rightarrow q_m)\},$$

where $k_i = \alpha_\Delta(\Delta_i \Rightarrow q_i)$.

We might (in case there is no misunderstanding) also write “ Δ ” with the annotation implicit.

2. Let $(\alpha_\Delta : \Delta)$ and $(\alpha_\Gamma : \Gamma)$ be two annotated databases. We define the annotated database $(\alpha : \Delta \cup \Gamma)$ as follows (notation: $\Delta + \Gamma$)

- For any clause $C \in \Delta - \Gamma$ let $\alpha(C) = \alpha_\Delta(C)$.
- For any clause C in $\Gamma - \Delta$ let $\alpha(C) = \alpha_\Gamma(C)$.
- For any clause $C \in \Delta \cap \Gamma$ let $\alpha(C) = \alpha_\Delta(C) + \alpha_\Gamma(C)$.

Remark 3.9. For a multi-set database Δ the annotation function is deigned to show in a proof how many times the clause is used. The \pm signed function shows whether a clause is active (can be used) or not active (cannot be used). So when we write $0 : q$ or $n : q$ we mean the clause q is active and has been used n times. This is the same as $(+, n) : q$. When we write $(-, n) : q$, we mean the clause has been used n times and is now not active. In the sequel our algorithm shall use clauses at most once, and so immediately after use they become inactive or are deleted.

Definition 3.10. Let Δ be an annotated database and let q be a goal. Define the notion of $\Delta \vdash_m q$, meaning q can be proved from Δ in m steps. Assume Δ is an annotated multi-set database with annotation α .

Case $m = 1$. q can be proved from Δ in one step if the clause $k : \emptyset \Rightarrow q$ is in Δ . We say that this clause was used in the proof. We change its annotation in Δ from k to $k + 1$ and say that success was achieved from the annotated Δ which yields also an associated annotated database $(\alpha' : \Delta)$, where where α' is like α except that k becomes $k + 1$.

Case $m = n + 1$. Let the clauses of Δ be $\{k_1 : (\Delta_1 \Rightarrow q_1), \dots, k_e : (\Delta_e \Rightarrow q_e)\}$.

We say that q can be proved from Δ in $n + 1$ steps if for some $k_i : (\Delta_i \Rightarrow q_i)$ we have

1. $q_i = q$.
2. $\Delta_i = \{0 : \Delta_{i,1} \Rightarrow p_{i,1}, \dots, 0 : \Delta_{i,s} \Rightarrow p_{i,s}\}$ and for each $1 \leq r \leq s$ we have that $p_{i,r}$ can be proved in n_r steps from $\Delta'_{i,r} = \Delta + \Delta_{i,r}$.

With the new annotation $\alpha'_{i,r}$ for $\Delta'_{i,r}$ provided by the induction hypothesis.

3. $n = \max\{n_e\}$.
4. We say that q is proved in $n + 1$ steps with the new associated database and annotation being $\Delta'_{i,1} + \dots + \Delta'_{i,s}$.

Lemma 3.11.

1. *Implicational propositional intuitionistic logic defined in Example 2.3 (axioms (a), (b) and MP) can also be defined as the smallest consequence relation (in the sense of Definition 2.2) satisfying the deduction theorem for \Rightarrow , namely:*

$$(DT) : \quad \Delta \vdash (A \Rightarrow B) \text{ iff } \Delta \cup \{A\} \vdash B$$

2. *Every intuitionistic wff B can be equivalently rewritten in the form*

$$B = (A_1 \Rightarrow (A_2 \Rightarrow \dots \Rightarrow (A_n \Rightarrow q) \dots)),$$

where q is atomic. Furthermore, B is equivalent to a clause with head q .

3. *Every set of wffs of intuitionistic logic is equivalent to a database.*
4. *Let Δ be a set of intuitionistic wffs and let q be an atom, then: $\Delta \vdash$ in intuitionistic consequence relation (as defined in item 1. above). iff for some m we have that $\alpha' : \Delta' \vdash_m q$, where Δ' is the set of clauses equivalent to the elements of Δ and α' annotates these clauses by the annotation σ .*

Proof. See [25, 17]. □

Example 3.12.

1. Consider in intuitionistic logic

$$\{c \Rightarrow a, (c \Rightarrow a) \Rightarrow c\} \vdash a.$$

If we go forward using modus ponens, we get

Step 1. $c \Rightarrow a, (c \Rightarrow a) \Rightarrow c \vdash c$. We use both clauses.

Step 2. $c, c \Rightarrow a \vdash a$. We use $c \Rightarrow a$ again, as well as c .

Let us now prove a using our computation.

Step 1. We want to show that $\{0 : c \Rightarrow a, 0 : (c \Rightarrow a) \Rightarrow c\} \vdash a$.

Step 2. Use the clause $0 : (c \Rightarrow a)$.

- change its annotation to $1 : (c \Rightarrow a)$
- move to step 3.

Step 3. $\{1 : c \Rightarrow a, 0 : (c \Rightarrow a) \Rightarrow c\} \vdash c$

- use $0 : (c \Rightarrow a) \Rightarrow c$
- change its annotation to $1 : (c \Rightarrow a) \Rightarrow c$.
- Add $0 : c$ to the database
- move to step 4.

Step 4. $\{0 : c, 1 : c \Rightarrow a, 1 : (c \Rightarrow a) \Rightarrow c\} \vdash a$

- Use clause $1 : c \Rightarrow a$.
- Change its annotation to $2 : c \Rightarrow a$.
- Move to Step 5.

Step 5. $\{0 : c, 2 : c \Rightarrow a, 1 : (c \Rightarrow a) \Rightarrow c\} \vdash c$.

- Use clause $0 : c$
- Change its annotation to $1 : c$
- Move to success.

We succeed in 5 steps with associated annotated final database

$$\{1 : c, 2 : c \Rightarrow a, 1 : (c \Rightarrow a) \Rightarrow c\}.$$

Example 3.13. Consider $\{0 : a, 0 : a \Rightarrow b, 0 : a \Rightarrow c, 0 : \{b, c\} \Rightarrow x\} \vdash x$.

Step 1. Use $0 : \{b, c\} \Rightarrow x$

- Change annotation to $1 : \{b, c\} \Rightarrow x$
- Prove $\{0 : a, 0 : a \Rightarrow b, 0 : a \Rightarrow c, 1 : \{b, c\} \Rightarrow x\} \vdash b$
- Prove $\{0 : a, 0 : a \Rightarrow b, 0 : a \Rightarrow c, 1 : \{b, c\} \Rightarrow x\} \vdash c$.
- Go to steps 2.1 and step 2.2.

Step 2.1. Prove $\{0 : a, 0 : a \Rightarrow b, 0 : a \Rightarrow c, 1 : \{b, c\} \Rightarrow x\} \vdash b$.

- Use $0 : a \Rightarrow b$
- change annotation to $1 : a \Rightarrow b$.
- Go to step 3.1.

Step 3.1. Prove $\{0 : a, 1 : a \Rightarrow b, 0 : a \Rightarrow c, 1 : \{b, c\} \Rightarrow x\} \vdash a$

- Use $0 : a$

- Change to $1 : a$

Success of a with the database

$$\Delta_a = \{1 : a, 1 : a \Rightarrow b, 0 : a \Rightarrow c, 1 : \{a, b\} \Rightarrow x\}.$$

Step 2.2. Prove $\{0 : a, 0 : a \Rightarrow, 0 : a \Rightarrow c, 1 : \{a, b\} \Rightarrow x\} \vdash c$.

- Use $0 : a \Rightarrow c$
- Change annotation to $1 : a \Rightarrow c$.
- Go to step 3.2.

Step 3.2. Prove $\{0 : a, 0 : a \Rightarrow b, 1 : a \Rightarrow c, 1 : \{a, b\} \Rightarrow x\} \vdash a$.

- Use $0 : a$
- Change annotation to $1 : a$
Success of c from the database

$$\Delta_c : \{1 : a, 0 : a \Rightarrow b, 1 : a \Rightarrow c, 1 : \{a, b\} \Rightarrow x\}.$$

Therefore success of x with the database

$$\Delta_x = \Delta_a + \Delta_b = \{2 : a, 1 : a \Rightarrow b, 1 : a \Rightarrow c, 2 : \{a, b\} \Rightarrow x\}.$$

Remark 3.14. *We compute with annotations so that we get an idea of how many times we use a clause in the computation. Later on in order to understand negation as failure as added to intuitionistic logic, we need to throw away any clause after its use. The reason is that we want to give a unique semantical meaning to each occurrence of negation as failure and if a clause is used more than once the “failure” will be from possibly two different databases and two different contexts and we will have a problem. So we count and annotate how many times we use each clause and be generous in our counting, and then use duplicates in sufficient number so that each duplicate clause is used at most once.*

So the database

$$\{c \Rightarrow a, (c \Rightarrow a) \Rightarrow c\} \vdash a$$

will become a multiset

$$\{c \Rightarrow a, c \Rightarrow a, (c \Rightarrow a) \Rightarrow c\} \vdash a.$$

Here we use each $c \Rightarrow a$ at most once.

We need the semantics for negation as failure in order to show that it is a negation in our sense.

So the above computation becomes

$$\{c \Rightarrow a, c \Rightarrow a, (c \Rightarrow a) \Rightarrow c\} \vdash a$$

if (use one copy of $(c \Rightarrow a)$ and delete it after use)

$$\{c \Rightarrow a, (c \Rightarrow a) \Rightarrow c\} \vdash c$$

if (use $(c \Rightarrow a) \Rightarrow c$ and delete it and add c).

$$\{c, c \Rightarrow a\} \vdash a$$

if (use $c \Rightarrow a$ and delete it)

$$\{c\} \vdash c$$

Success.

3.2 Negation as failure

We now study the behaviour of negation as failure.

Example 3.15. This example illustrates the idea of why we think negation as failure is a proper negation.

Consider a logic program without loops (where every atom either succeeds or fails):

$$\Delta = \{-b \Rightarrow a\}.$$

From this program, a succeeds and b fails. Let $\Theta(\Delta) = \{y \mid y \text{ fails}\}$.

Add the axioms for a new negation symbol \mathbf{n}_Δ to be

$$\Delta_{\mathbf{n}} = \{y \Rightarrow \mathbf{n}_\Delta \mid y \in \Theta_\Delta\}.$$

In Δ translate any $\neg x$ as $x \Rightarrow \mathbf{n}_\Delta$. This translation gives a new theory Δ' .

We get $\Delta_1 = \Delta' \cup \Delta_{\mathbf{n}}$ to be $\{(b \Rightarrow \mathbf{n}_\Delta) \Rightarrow a, b \Rightarrow \mathbf{n}_\Delta\}$.

This is an intuitionistic theory for intuitionistic \Rightarrow . We have

$$\begin{aligned} \Delta_1 \vdash x &\text{ iff } x \text{ succeeds from } \Delta \\ \Delta_1 \vdash (x \Rightarrow \mathbf{n}_\Delta) &\text{ iff } x \text{ fails from } \Delta. \end{aligned}$$

This was done in my paper [21].⁹

Example 3.16. When we have loops, we can use answer set programming [26].

Consider Γ :

$$\Gamma = \{-a \Rightarrow b, \neg b \Rightarrow a\}.$$

⁹ This is Theorem B on page 29 of [21]. It says and we quote:

Answer set 1. $a = \text{in}, b = \text{out}$. We get Γ_1 .

$$\Gamma_1 = \{b \Rightarrow \mathbf{n}_1, (b \Rightarrow \mathbf{n}_1) \Rightarrow a, (a \Rightarrow \mathbf{n}_1) \Rightarrow b\}.$$

$\Gamma_1 \vdash a$ but $\Gamma_1 \not\vdash b$.

Similarly

Answer set 2. $a = \text{out}, b = \text{in}$.

We get

$$\Gamma_2 = \{a \Rightarrow \mathbf{n}_2, (b \Rightarrow \mathbf{n}_2) \Rightarrow a, (a \Rightarrow \mathbf{n}_2) \Rightarrow b\}.$$

We get $\Gamma_2 \vdash b, \Gamma_2 \not\vdash a$.

Example 3.17. This covers a general loop. Consider the loop

$$\{\neg a \Rightarrow a\}.$$

Answer set programming does not help here, but we can add a historical loop checker. Let us try. We use the notation “ $?a = 1$ ” means $?a = \text{success}$. “ $?a = 0$ ”; means $?a = \text{failure}$.

Part 1: We start the computation with the query $?a = 1$.

$$\neg a \Rightarrow a \quad ?a = 1$$

iff

$$\neg a \Rightarrow a?a = 0$$

iff

$$\neg a \Rightarrow a \quad ?\neg a = 0$$

iff

$$\neg a \Rightarrow a \quad ?a = 1.$$

We loop. So the query $?a = 1$ fails, so $?a = 0$ succeeds.

Theorem B. Let P be any database. Let L be $L = \{yP(?F)y = 0\}$. Assume that P is such that every goal either succeeds or fails. Then for any G $P(?F)G = 1$ iff $(P, L)(?I)G = 1$.

Note the assumption that every goal either succeeds or fails (i.e. no loops). This is noted on the same page of the paper, we Quote further:

Theorem B is important. It says that if our mechanical theorem proving is complete (i.e. $P?A = 0$ or $P?A = 1$), then negation as failure is the truly sound classical negation. This holds because it is equal to negation as inconsistency, which is complete. However, in the case that the theorem prover $P?G$ is not complete, e.g. when we have loops, negation as failure may not behave logically.

Part 2: We start the computation with the query $?a = 0$. We implement this by continuing the computation beyond the loop point of Part 1:

$$?\neg a = 1$$

iff

$$?a = 0$$

we get another looping point where we loop again.

This means that if we start with $?a = 0$ then $?a = 0$ also loops, so the query fails and we get answer $a = 1$.

So we get two possibilities.

We thus get no agreement using the loop checker. If we ask $?a = 1$ we loop and therefore we get that a fails and if we start with the query $?a = 0$ we also loop and get that a succeeds.

In case we consider that a fails, we get we get $\{a\}$ for the fail set and we have:

$$\begin{aligned} \Gamma_1 &= \{a \Rightarrow \mathbf{n}_1, (a \Rightarrow \mathbf{n}_1) \Rightarrow a\} \\ \Gamma_1 &\vdash a \end{aligned}$$

In case a fails and in case a succeeds we get \emptyset_\perp for the fail set. We use \perp

$$\Gamma_2 = \{\perp \Rightarrow \mathbf{n}_2, (a \Rightarrow \mathbf{n}_2) \Rightarrow a\}$$

we have $\Gamma_2 \not\vdash a$.

For the sake of comparison, let us re-do Example 3.16 using a loop checker and see whether we get the same two possibilities or not. In other words independently of whether we ask $?a = 1$ or if we ask $?a = 0$, we get the same result for a , namely either a succeeds or a fails.¹⁰

Example 3.18. Example 3.16 using a loop checker.

Let

$$\Gamma = \{-a \Rightarrow b, -b \Rightarrow a\}.$$

1. Start with $?a = 1$. We get:

$$?a = 1$$

if

$$?\neg b = 1$$

¹⁰ In view of the restriction of Theorem B, (the restriction of no loops, see Footnote 9), we want to eliminate loops by a loop checker, and ask will the theorem go through?

if $?b = 0$
if $\neg a = 0$
if $?a = 1.$

We loop.

Therefore $?a = 1$ fails, so $?a = 0$ succeeds.

If we continue after the loop we get

if $? \neg b = 1$
if $?b = 0$

We see that we get $?b = 0$ looping if we were starting with $b = 0$.

So in this case $b = 0$ fails, so $b = 1$ succeeds. So the answer success set is $b = 1$ and $a = 0$.

- Note that if we start with “ $?a = 1$ ” or with “ $?b = 0$ ” we never get the query “ $?a = 0$ ” or the query “ $?b = 1$ ”.
- Let us start and ask $?b = 1$.

$?b = 1$
if $? \neg a = 1$
if $?a = 0$
if $? \neg b = 0$
if $?b = 1$
if $? \neg a = 1$
if $?a = 0$

The looping elements are “ $?b = 1$ ” and “ $?a = 0$ ”. If the looping elements are failures then the answers are consistent $a = 1$ and $b = 0$ are the successes.

4. We see that (1) and (3) completely agree with the answer set programming answer.

We now want to add negation as failure to full implicational intuitionistic logic.¹¹ This negation, when added to intuitionistic implication, is very difficult to handle mainly because the meaning of “ \neg ” keeps on changing depending on where \neg appears in the formulas. This is different from answer sets where there may be several options for the meaning of “ \neg ”, but once we choose an option, the meaning of \neg gets fixed for all occurrences of “ \neg ”.

The next example 3.19 will illustrate the problem.

Example 3.19. Let Δ_1 be $\{(1) - (4)\}$

1. $(d \Rightarrow (c \Rightarrow \neg a)) \Rightarrow c$
2. $c \Rightarrow a$
3. $\neg d \Rightarrow x$
4. $\neg x \Rightarrow a$

and let Δ_2 be $\{(1) - (6)\}$, where

5. d
6. c

Note that in the logic of intuitionistic implication, we have, for any X, Y, Z and Δ ,

$$\Delta \vdash X \Rightarrow (Y \Rightarrow Z)$$

iff (by definition, or by the deduction theorem)

$$\Delta \cup \{X\} \cup \{Y\} \vdash Z$$

So if we add \neg to \Rightarrow , then for $Z = \neg a$ we get that “ \neg ” says that in our computation, $?a$ is a failure from the database $\Delta \cup \{X, Y\}$.

¹¹ This means that we take implication with Axioms (a) and (b) of Item 1 of Example 2.3, and add the negation as failure symbol and define the computation as in N-Prolog, see [14]. The reader need not reference [14], but follow the computation in the examples which follow. It is very intuitive.

So looking at clause (1), we can see that “ $\neg a$ ” would need to fail from the database Δ_1 with clauses (5) and (6) added i.e. from Δ_2).

While “ \neg ” in clauses (2) and (3) do not add anything to the database, so \neg needs to fail from Δ_1 above.

Let us now do some specific computations to illustrate the problems involved.

Computation, part 1. We ask “ $?a = 1$ ” from database $\Delta_1 = \{(1), (2), (3), (4)\}$ and use clause (4) first. So

$$\Delta_1 \vdash ?a = 1$$

using (4), if

$$\Delta_1 ?\neg x = 1$$

if

$$\Delta_1 ?x = 0$$

using (3), if

$$\Delta_1 ?\neg d = 0$$

if

$$\Delta_1 ?d = 1.$$

We get that the query $?d = 1$ fails from the database Δ_1 , because this database has no clause with head d . Therefore the original query, namely $?a = 1$, fails for the choice of the above initial clause (4) with head a . However, we also have clause (2) with head a and so let us backtrack and ask $?a = 1$ again and this time choose clause (2).

Computation, part 2. Let us backtrack, and ask $\Delta_1 \vdash ?a = 1$ using clause (2).

$$\Delta_1 \vdash ?a = 1$$

using (2), if

$$\Delta_1 ?c = 1$$

using (1), if

$$?(d \Rightarrow (c \Rightarrow \neg a)) = 1$$

and we ask, if

$$\Delta_2 ?\neg a = 1$$

Since

$$\Delta_2 = \Delta_1 \cup \{d, c\}$$

if

$$\Delta_2 ?a = 0$$

from (4), if

$$\Delta_2 ?\neg x = 0$$

Computation, part 3. Note that “ $\neg x$ ” is now asked from Δ_2 and not from the original Δ_1 ! The meaning of “ \neg ” has changed for its same occurrence in “ $\neg x$ ”.

We have two options:

Option 1. Continue and ask $\neg x$ from Δ_1 .

Option 2. Continue and ask $\neg x$ from Δ_2 .

Let us do them in parallel

Option 1	Option 2
from (4) $? \neg x = 0$	from (4), if $? \neg x = 0$
if $?x = 1$	if $?x = 1$
from (3), if $? \neg d = 1$	from (3) if $? \neg d = 1$
if $?d = 0$	if $?d = 0$
Success.	Fail.
Clause (5) not available	Because clause (5) is available.
because we are using Δ_1	

Question. Which option do we adopt?

Answer. Option 1 is better from the point of view of “what is negation” because we want the meaning of each occurrence of negation to become fixed. We write clause 1 as

$$1^*. \quad (d \Rightarrow (c \Rightarrow \neg_{\Delta_2} a)) \Rightarrow c .$$

and clauses (3) and (4) as

$$3^*. \quad \neg_{\Delta_1} d \Rightarrow x$$

$$4^*. \quad \neg_{\Delta_1} x \Rightarrow a.$$

Option 2 is known in the literature as \mathcal{N} =Prolog [14] and its negation as failure was extensively investigated and has complex semantics [24, 26].

Example 3.20. Let us revisit Example 3.19 and be very simple minded about it. We saw in Example 3.15 and Example 3.16. The very simple approach that for a logic program Δ without loops or with semantics where every atom (or literal) x appearing in the program x either succeeds or fails, we can take as Θ_Δ for negation the set of all y which fail.

So let us apply the same procedures to the program Δ_1 of Example 3.19.

The clauses are:

1. $(d \Rightarrow (c \Rightarrow \neg a)) \Rightarrow c$
2. $c \Rightarrow a$
3. $\neg d \Rightarrow x$
4. $\neg x \Rightarrow a.$

The atoms appearing in this program are $\{a, c, d, x\}$. So let us check first whether every atom either succeeds or fails. We get

a succeeds
 c succeeds
 x succeeds
 d fails.

Here we use the \mathcal{N} -Prolog computation, namely Example 3.19 computation Parts 1, 2 and Part 3, option 2.

Second, let Θ_1 be $\{d\}$ (as our recipe dictates) and rewrite the program as (with \mathbf{n}_1 as negation)

- 1*. $(d \Rightarrow (c \Rightarrow (a \Rightarrow \mathbf{n}_1))) \Rightarrow c$
- 2*. $c \Rightarrow a$
- 3*. $(d \Rightarrow \mathbf{n}_1) \Rightarrow x$
- 4*. $(x \Rightarrow \mathbf{n}_1) \Rightarrow a$

and the additional clause for d

- 7*. $d \Rightarrow \mathbf{n}_1$

We now expect the same results of success or failure for $\{a, c, x, d\}$ and of course failure for \mathbf{n} .

Let us check.

Case n_1 :

$$?n_1 = 1$$

using (7), if

$$?d = 1$$

fail.

Case a. Using 4*

$$?a = 1$$

$$?x \Rightarrow n_1$$

add x

$$?n_1$$

using (7), if

$$?d = 1$$

fail.

We backtrack and use clause 2*.

using 2*

$$?a = 1$$

if

$$?c = 1$$

using 1*,

add d

add c

add a , if

$$?n_1 = 1$$

using (7)

$$?d$$

Success for $?a = 1$.

Case c. Success. Follows from case a that $?c = 1$.

Case d.

$$?d = 1$$

fails.

The big question we ask is:

Big Question BQ Is this very simple minded approach an indication of a possible general truth (big theorem) or does it only work sometimes? See [24].

Answer to BQ: It is an accident, as the next Example 3.21 shows. However, there might be a general theorem which is inductive on the structure of nested negations and its proof would certainly be quite complicated.

Example 3.21.

Part 1: The problem. Consider the following program $\Delta_1 = \{(1)\}$:

1. $((a \wedge (\neg a \wedge \neg b \Rightarrow x)) \Rightarrow x) \Rightarrow z$.

Let us query

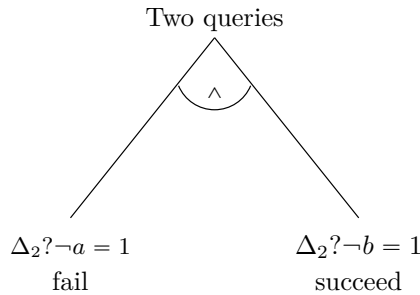
$$\Delta_1?z = 1$$

We use clause (1) and ask if $\Delta_2?x = 1$ where $\Delta_2 = \Delta_1 \cup \{(2), (3)\}$ where

2. a
3. $\neg a \wedge \neg b \Rightarrow x$.

We continue the computation using clause (3) and ask a conjunction if $\Delta_2?\neg a \wedge b = 1$ which splits to two queries.

if



Thus overall we get that $\Delta_1?z$ fails.

Therefore we have that $\Delta_1?z$ fails as well as $\Delta_2?a$, $\Delta_2?b$ and $\Delta_1?x$ all fail because they are not heads of any clauses. So all atoms fail from Δ_1 . So

$$\Theta_{\Delta_1} = \{a, b, x, z\}.$$

We therefore translate Δ_1 to Δ'_1 , namely Δ'_1 includes the following clauses

$$1^*. ((a \wedge ((a \Rightarrow \mathbf{n}) \wedge (b \Rightarrow \mathbf{n}) \Rightarrow x)) \Rightarrow x) \Rightarrow z$$

$$4^*. a \Rightarrow \mathbf{n}$$

$$5^*. b \Rightarrow \mathbf{n}$$

$$6^*. x \Rightarrow \mathbf{n}$$

$$7^*. z \Rightarrow \mathbf{n}.$$

Let us ask

$$\Delta'_1?z = 1$$

if

$$\Delta'_2?x = 1,$$

where $\Delta'_2 = \Delta'_1 \cup \{(2^*), (3^*)\}$, where

$$2^*. a$$

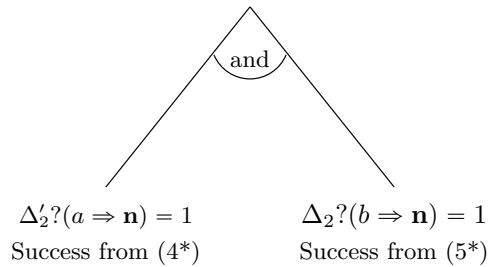
$$3^*. ((a \Rightarrow \mathbf{n}) \wedge (b \Rightarrow \mathbf{n})) \Rightarrow x$$

We continue:

if

$$\Delta'_2?((a \Rightarrow \mathbf{n}) \wedge (b \Rightarrow \mathbf{n})) = 1$$

if



This result does not match.

z should fail from Δ_1 .

Part 2: The remedy. The remedy is that we need to eliminate the negations inductively on their nestings. Suppose we ask as in Part 1, the query $?\neg z$. The ‘neg’ in “ $\neg z$ ” is from Δ_1 , but the “ \neg ” in “ $\neg a$ ” and in “ $\neg b$ ” is from Δ_2 . So from Δ_2 a succeeds. What fails from Δ_2 are b and x and z . So the proper translation of Δ'_2 is

$$\begin{aligned}\Delta''_2 &= \{(1*), (2*), (3*), (5*), *6*), (7*)\} \\ &= \Delta'_2 - \{(4*)\}\end{aligned}$$

Let us now follow the computation of Part 1, up to the point where we have $\Delta''_2?a \Rightarrow \mathbf{n}$. This will fail as we want.

Example 3.22. Consider the intuitionistic implication database $\Delta = \{1a, 2\}$:

(1a) $c \Rightarrow a$

(2) $(c \Rightarrow a) \Rightarrow c$.

This database proves “ a ” but clause 1a needs to be used twice in the proof, as follows:

Step 1. $c \Rightarrow a, (c \Rightarrow a) \Rightarrow c \vdash c$.

Add clause (3): c .

Step 2. $c, c \Rightarrow a \vdash a$.

If we impose affine linear logic resource restrictions, requiring that all clauses can be used at most once, then we cannot prove a .

However, we can use multisets and add duplicates of the copies we need.

In our case we add

(1b) $c \Rightarrow a$.

Let $\Delta = \{c \Rightarrow a, c \Rightarrow a, (c \Rightarrow a) \Rightarrow c\}$.

We have

Step 1. (1a): $c \Rightarrow a$, (2): $(c \Rightarrow a) \Rightarrow c \vdash c$.

We add (3): c and delete the used clauses.

Step 2. (1b): $c \Rightarrow a$ (unused). (3): $c \vdash a$.

The above leads us to the following:

Lemma 3.23 (Simple.). For any Δ in intuitionistic logic there exists a natural number N such that for any atom a appearing in Δ we have

$\Delta \vdash a$ in intuitionistic logic

iff

$\Delta' = \Delta \cup \Delta \cup \dots_{N\text{-times}} \cup \Delta \vdash a$ in affine linear logic.

Remark 3.24. We can be more fine tuned in defining Δ' in Lemma 3.23 and use linear logic.

Example 3.25. Let us add to the database of Example 3.22 the clause

(0) $\neg c \Rightarrow c$

where \neg is negation as well as the clause (4)

(4) $(\alpha_1 \Rightarrow c) \Rightarrow c$.

We need to look at

$\Delta^* = \{(0), (1a), (2), (4)\}$

as a logic program and proceed with a as a goal.

We offer 3 computations for the success of $?a$ which will show that there is no fixed meaning to the negation \neg in $\neg c \Rightarrow c$.

Computation 1

Step 1. $\Delta^*?a = 1$.

Step 2. If, using clause 1a, $\Delta^*?c = 1$.

Step 3. If, using clause 2, $\Delta^* \cup \{(3)\}?a = 1$.

Step 4. Iff, using clause 1a, $\Delta^* \cup \{c\}?c = 1$.

Step 5. Using clause (3), success.

In Computation 1, the meaning of $\neg c$ could only be failure from Δ^* .

However clause (0) is not used! So its meaning does not matter.

Computation 2

Proceed up to Step 3, but continue differently, using clause (0) instead:

Step 4*. Using clause (0), if $\Delta^* \cup \{c\}? \neg c = 1$,
and we continue, if $\Delta^* \cup \{c\}?c = 0$.

Step 5*. Using clause (0), if $\Delta^* \cup \{c\}? \neg c = 0$
and we continue, if $\Delta^* \cup \{c\}? c = 1$.

Step 6*. Using clause (3), if success.

According to Computation 2, $\neg c$ means failure from the database $\Delta^* \cup \{c\}$.
Let us now use clause (4) in Computation 3.

Computation 3

Proceed as in Computation 2, up to Step 5*. Continue with Step 6**:

Step 6**. Using clause (4), if $\Delta^* \cup \{c, \alpha_1\}? c = 1$.

Step 7**. Using clause (0), if $\Delta^* \cup \{c, \alpha_1\}? \neg c = 1$
and we continue, if $\Delta^* \cup \{c, \alpha_1\}? c = 0$.

Step 8**. Using clause (0) if $\Delta^* \cup \{c, \alpha_1\}? \neg c = 0$
and we continue, if $\Delta^* \cup \{c, \alpha_1\}? c = 1$.

Step 9**. Using clause (3), if success.

The problem with the above is that the meaning of “ $\neg c$ ” in clause (0) depends on the computation and is not fixed by the program itself and so we cannot say from which database $\neg c$ is supposed to fail.

The computation uses clause (0) again twice. In fact, we could have had more clauses like

$$(4k) (\alpha_k \Rightarrow c) \Rightarrow c$$

and we could have added more steps which added $\alpha_k, k = 1, 2, 3, \dots$ to the database and got that $\neg c$ is computed from

$$\Delta^* \cup \{c, \alpha_1, \dots, \alpha_k\}.$$

Example 3.26. We have no such problem (of Example 3.25) in the case of affine linear logic, because clauses are used at most once. So to succeed with $\Delta^?a$ we need to duplicate clauses (1a) and (0) because they are used twice. Once we duplicate the negations are also duplicated and thus $\neg c$ appears twice in

$$(0a) (\neg_a c \Rightarrow c)$$

$$(0b) (\neg_b c \Rightarrow c)$$

and each occurrence has a clear meaning.

The above explains the idea behind the following theorem (see paper [24]).

Theorem 3.27. *Let Δ be a database in intuitionistic logic with negation as failure with the N-Prolog goal directed computation as defined in [14], and illustrated in the examples above.*

Then there exists a multiset database Δ' with duplicates from Δ and auxiliary new “administrative” clauses in the language of Δ such that:

For every atom appearing in Δ we have

$$\begin{aligned} \Delta' \vdash \pm & \text{ in linear logic} \\ \text{iff } \Delta \vdash \pm a & \text{ in N-Prolog} \end{aligned}$$

Where $+a$ is a and $-a = \neg a$, and each occurrence “ $O(\neg x, C)$ ” of “ $\neg x$ in a clause C in Δ' is assigned a unique database $\Delta(O(\neg x, C))$ and any such occurrence of $\neg x$ in C is rewritten in Δ' as $\neg_{\Delta}(O(\neg x, C))$, such that in any computation.

$\Delta' \vdash \neg x$ iff x fails from $\Delta(O(\neg x, C))$, where $O(\neg x, C)$ is the occurrence of \neg in $\neg x$ in C .

Proof. Complex construction and induction in [24]. □

We need to define mathematically the induction. If we manage that, then we will get that nested negation as failure is a negation in our sense.

It seems that we need to follow the idea of defining/introducing several negations at once and characterize them together in terms of each other, and they will be negations in our sense. This is a new ball game and is the subject of active research.

Remark 3.28. *Note that in intuitionistic logic with negation as failure the cut theorem does not hold. Let $\Delta_1 = \{\neg b \Rightarrow a\}$*

$$\Delta_2 = \{\neg b \Rightarrow a, a, b\}.$$

We have

$$\begin{aligned} \Delta_1 &= \{\neg b \Rightarrow a\} \vdash a \\ \Delta_2 &= \{\neg b \Rightarrow a, a, b\} \vdash a \end{aligned}$$

but using Δ_1 to cut gives us

$$\Delta_3 = \{\neg b \Rightarrow a, \neg b \Rightarrow a, b\} \not\vdash a.$$

The reason is that in Δ_1 , $\neg b$ means that b fails from Δ_1 and in Δ_2 , $\neg b$ means that b fails from Δ_2 .

So using our ideas we can save the cut theorem. We let

$$\begin{aligned}\Delta_1 &= \{\neg_{\Delta_1} b \Rightarrow a\} \vdash a \\ \Delta_2 &= \{\neg_{\Delta_2} b \Rightarrow a, a, b\} \vdash a\end{aligned}$$

and so we cut

$$\Delta_3 = \{\neg_{\Delta_2} b \Rightarrow a, \neg_{\Delta_2} b \Rightarrow a, b\} \vdash a$$

holds.

3.3 Comparison with semantics of Olivetti Terracini [31]

3.3.1 Background

The problem of the early 1980s was to give semantics to logic programs (whether Prolog or N-Prolog) which contain negation as failure “ \neg ”.

What was needed was a logical system \vdash and a completion process providing a completion $\mathbf{Com}(\Delta)$, for any program Δ such that the following holds:

- If $?q$ succeeds from Δ then $\mathbf{Com}(\Delta) \vdash q$.
- If $?q$ fails from Δ (i.e. $\Delta? \neg q$ succeeds from Δ), then $\mathbf{Com}(\Delta) \vdash \neg q$.

Where \vdash is some logic where the completion is expressed.

There were two answers to this at the time using modal logic.

1. Gabbay for ordinary Prolog [16], using modal logic and [21] using intuitionistic logic.
2. Olivetti-Terracini for general N-prolog, using modal logic [31].

Note that the modal completion does not give proof theory and semantics for “ \neg ”. It may be possible to extract such a semantics from the modal completion. This needs to be investigated.

Example 3.29 (Olivetti–Terracini [31], Example 9.1).

Part 1. The Program

Consider the database (Program)

1. $(a \Rightarrow \neg b) \Rightarrow c$.
2. $\neg a \Rightarrow b$

The Olivetti–Terracini completion is the following conjunction

$$\neg a \wedge b \Leftrightarrow \neg a) \wedge \Box[(\alpha \Rightarrow \neg b) \Rightarrow c] \wedge \Box[\alpha \Rightarrow b) \Rightarrow \neg c],$$

where

$$\alpha = [a \wedge (c \Leftrightarrow \neg b) \wedge (b \Leftrightarrow \neg a)].$$

Part 2: The N-Prolog computation

Let us now compute our own semantics.

First we ask what succeeds and what fails from Δ . The language is $\{a, b, c\}$. $?x = 1$ means success and $?x = 0$ means failure.

Case $?a = 1$. $?a$ fails because there is no clause with head a .

Case $?b = 1$.

Step 1. $\Delta ?b = 1$

Step 2. if, using clause 2, $\Delta ?\neg a = 1$

Step 3. if, $\Delta ?a = 0$

Step 4. if, success since a fails.

Note that the “ \neg ” in clause (2) is failure from Δ .

Case $?c = 1$.

Step 1. $\Delta ?c = 1$

Step 2. if, from clause 1, $\Delta ?(a \Rightarrow \neg b) = 1$.

Step 3. If $\Delta \cup \{a\} ?\neg b = 1$.

Step 4. if $\Delta \cup \{a\} ?b = 0$.

Step 5. if, from clause 2, $\Delta \cup \{a\} ?\neg a = 0$.

Note that in this step clause 2 is used again!

Step 6. if, $\Delta \cup \{a\} ?a = 1$.

Step 7. if, success, because $a \in \Delta \cup \{a\}$.

Note however, that in this computation, the “ \neg ” in clause (2) is failure from $\Delta \cup \{a\}$.

Part 3: The program $\Delta_1 = \Delta \cup \{a\}$.

The clauses of this program are

1. $(a \Rightarrow \neg b) \Rightarrow c$
2. $\neg a \Rightarrow b$
3. a

Let us ask $\Delta ?a = 1, \Delta ?b = 1$ and $\Delta ?c = 1$.

- We have that $\Delta_1?a = 1$ succeeds because $a \in \Delta_1$.
- We have $\Delta_1?b = 1$ fails because a succeeds.
- We have $\Delta_1?c = 1$ succeeds, because of the following
 - $\Delta_1?c = 1$.
 - if, from clause 1, $\Delta_1 \cup \{a\} = \Delta_1?\neg b = 1$
 - if, $\Delta_1?b = 0$
 - if, from clause 2, $\Delta_1?\neg a = 0$
 - if $\Delta_1?a = 1$
 - if, from clause 3, success.

Part 4: Summary.

We summarise as follows:

- (S1) From the program Δ , $?a = 0, ?b = 1, ?c = 1$.
- (S2) From $\Delta_1 = \Delta \cup \{a\}$, we have $?a = 1, ?b = 0, ?c = 1$.
- (S3) Clause (2) is used twice, in different derivations, $\Delta?c$ and $\Delta?b$, with different meaning for negation.
- (S4) In the computation of $\Delta?c = 1$, “ \neg ” is used in each of its occurrences to mean failure from the same database $\Delta \cup \{a\}$.
- (S5) In the query $\Delta?b = 1$, “ \neg ” is used as failure from Δ .

We therefore cannot assign a unique interpretation to “ \neg ” in Δ . Note we have already seen examples (e.g. Example 3.19) where \neg is used twice or more with different meanings in the same computation.

Part 5: The Gabbay completion for Δ .

- Since clause (2) is used twice in $\Delta?c$ we duplicate it, call the duplicate (2.1).
- Add two constants \mathbf{n} and \mathbf{n}_1 , one for failure from Δ (\mathbf{n} represents the disjunction of all atoms failing from Δ), and respectively one for failure from Δ_1 . Replace in Δ each occurrence of $\neg X$ by $X \rightarrow \mathbf{n}$ and in clause (2.1) replace “ $\neg a$ ” by $a \Rightarrow \mathbf{n}_2$.
- View Δ as a multiset. Add to Δ the duplicate clause (2.1) and also add the clause $\Delta_{\mathbf{n}}$ and $\Delta_{\mathbf{n}_1}$, where

$$\begin{aligned}\Delta_{\mathbf{n}} &= \{y \Rightarrow \mathbf{n} \mid \Delta?y = 1 \text{ fails}\} \\ \Delta_{\mathbf{n}_1} &= \{y \Rightarrow \mathbf{n}_1 \mid \Delta_1?y = 1 \text{ fails}\}.\end{aligned}$$

- Therefore the Gabbay completion is the multiset $\Delta^* = \{(1), (2^*), (2.1^*), (4^*), (5^*)\}$.

$$(1^*) \quad (a \Rightarrow (b \Rightarrow \mathbf{n}_1)) \Rightarrow c$$

$$(2^*) \quad (a \Rightarrow \mathbf{n}_1) \Rightarrow b$$

$$(2.1^*) \quad (a \Rightarrow \mathbf{n}) \Rightarrow b$$

$$(4^*) \quad a \Rightarrow \mathbf{n}$$

$$(5^*) \quad b \Rightarrow \mathbf{n}_1.$$

Part 6: Checking the Gabbay completion.

We must get that

$$\Delta^*?a = 1, \text{ fails}$$

$$\Delta^*?b = 1, \text{ succeeds}$$

$$\Delta^*?c = 1, \text{ succeeds}$$

- $\Delta^*?a = 1$ fails because a is not the head of any clause.
- $\Delta^*?b = 1$ succeeds because we can use clauses (2.1^*) and (4^*) .
- $\Delta^*?c = 1$ succeeds because we can use clauses (1^*) , (2^*) and (5^*) .

But note that Δ^* contains no negations. It is a database of intuitionistic implication only!

Part 7: Concluding statement.

We can offer a completion in intuitionistic logic itself for any program with nested negation as failure using the ideas of [24] as illustrated by our above examples and discussion. Note that we require that every literal either succeeds or fails. The case of loops can be problematic. This case also needs to be investigated. In fact what we are doing here is part of a general challenge, as addressed in [24], namely:

The general challenge

1. Given a proof system “ \vdash ” the challenge is to develop a calculus of the form \vdash_N based on \vdash with rules of the form

$$\frac{\Delta_i \vdash_N \Gamma_i, \Delta'_j \not\vdash_N \Gamma'_i}{\pm \Delta \vdash_N \Gamma}$$

where “ \vdash_n ” agrees with “ \vdash ” for rules without “ $\not\vdash_N$ ”.

2. Give semantics for “ \vdash_n ”
3. Bring the metalevel “ $\not\vdash_N$ ” into the object level in the form of a connective “ \neg ”, $\neg A$ (unprovable A) meaning

$$\Delta, \neg A \vdash_N \Gamma \text{ iff } \Delta \not\vdash_N \Gamma, A.$$

See how this is done in [24] for nonmonotonic consequence relation and the conditional.

4. The perceptive reader might observe that the elimination of the negation “ \neg ” in 3. above in terms of constants of the form \mathbf{n} , depends on the algorithmic proof system which we use to realise/compute the provability “ \vdash ” and is not a direct property of the data and “ \vdash ” viewed as a declarative relation. This is not a problem because we can duplicate the data enough times to cater for all possible different proofs. The databases are finite and the provability is decidable and so this can be done. See [24].

4 Sample of other known negations

We examine other known “negations” and see whether they satisfy our definition of what is negation.

4.1 Paraconsistent negations

Definition 4.1 (Finite many valued logics with negation¹²).

1. A finite propositional many valued logic for the set of atomic propositions Q has the form $(V, D, \mathbb{C}, \mathbb{M})$, where V is a finite set of truth values, $D \subsetneq V$ is a non-empty subset of designated values, \mathbb{C} is a set of propositional connectives and \mathbb{M} is a truth table for each connective $\mathbf{c} \in \mathbb{C}$, namely, if \mathbf{c} is an n -place connective then $\mathbb{M}(\mathbf{c})$ is a function from

$$\mathbb{M}(\mathbf{c}) : V^n \mapsto V.$$

2. An assignment h is a function giving for each atomic proposition q a truth value $h(q) \in V$.
3. Given a assignment h to the atoms and a wff A , h can be extended to $h(A)$, for any wff A by induction:

¹²See [2] for a comprehensive study of the connection with paraconsistency.

- (a) $\bar{h}(A) = h(A)$, for A atomic
 (b) $\bar{h}(\mathbf{c}(B_1, \dots, B_m)) = \mathbb{M}(\mathbf{c})(\bar{h}(B_1), \dots, \bar{h}(B_m))$.

4. We define a (minimal) semantic consequence relation \models between sets Δ of wffs and a single wff A as follows:

(*) $\Delta \models A$ iff for every assignment h , if $\bar{h}(X) \in D$ for all $X \in \Delta$, then $\bar{h}(A) \in D$.

5. Let \neg be a unary connective in \mathcal{C} .

- (a) We say that \neg is a possible candidate for negation iff for some $x \in V$ we have $\neg x \notin D$.
 (b) We say that a possible candidate for negation \neg is indeed a negation (in our sense) relative to \models iff there exists a set of wffs Θ_{\neg} , such that for any Δ and any A we have

$$\Delta \models \neg A \text{ iff } \Delta \cup \{A\} \models \alpha$$

for some $\alpha \in \Theta_{\neg}$.

6. We say that a negation \neg , as defined in (4), is minimally paraconsistent relative to \models iff for some atom q and some h we have that both $h(q)$ and $h(\neg q)$ are in D .

Theorem 4.2. *Let \neg be as in item 4 of Definition 4.1. Then \neg cannot be minimally paraconsistent.*

Proof. Let $\alpha \in \Theta_{\neg}$.

1. Since $\emptyset, \alpha \vdash \alpha$ we have that $\emptyset \models \neg \alpha$.
2. Let q be an atom not appearing in α . Then $q, \alpha \models \alpha$.
Therefore $\alpha \models \neg q$.
3. Since we have $\neg q \models \neg q$, we get $q, \neg q \models \alpha$.
4. If there exists a p and h such that $\bar{h}(p)$ and $\bar{h}(\neg p)$ are in D then we get from (3) for $q = p$ that $\bar{h}(\alpha) \in D$.
5. From (2) we have $\alpha \models \neg q$, for atomic q not appearing in α , and from (4) we have that $\bar{h}(\alpha) \in D$. Therefore $\bar{h}(\neg q) \in D$. But $h(q)$ can be any value on q , therefore

$$\mathbb{M}(\neg) : V \mapsto D.$$

This means that for any wff A we have $\models \neg A$. □

Remark 4.3. *Theorem 4.2 relied on its proof on the set of atoms being infinite. That is for any wff $A(x_1, \dots, x_n)$ built up from the atoms x_1, \dots, x_n , we can find a new atom q not appearing in A . We now ask: what if the language contains only one atom p and only the negation connective \neg . Can we prove the theorem for this case?*

The answer is yes, we can.

Step 1. *Since the wffs of the language are formed using only \neg and single atom p , any $\alpha \in \Theta_{\neg}$ has the form $\neg^m p$.*

Step 2. *Since*

$$\neg^m p \models \neg^m p$$

we get

$$\models \neg^{m+1} p.$$

Step 3. *Since*

$$\neg^{m+2} p \models \neg^{m+2} p$$

we get

$$\neg^{m+2} p, \neg^{m+1} p \models \neg^m p.$$

Step 4. *But $\models \neg^{m+1} p$ and $\models \neg^{m+2} p$. Therefore*

$$\models \neg^m p.$$

This means for any h

$$h(\neg^m p) \in D.$$

Step 5. *Therefore, from Step 4*

$$p \models \neg^m p$$

and so

$$\models \neg p.$$

Remark 4.4. *We compare with da-Costa's notion of paraconsistency [11], (as formalised in [9]), which requires the existence of a theory Δ and formulas A and B such that $\Delta \models A$ and $\Delta \models \neg A$ but $\Delta \not\models B$.*

Let us take a logic with $V = \{0, 1\}$ and only one atom p and negation \neg , with:

$$\begin{aligned}\neg 1 &= 1 \\ \neg 0 &= 1.\end{aligned}$$

Let $\Theta_{\neg} = \{\neg p\}$.

Then we have

- $\neg p, \neg\neg p \models \neg p$
- $\neg p, \neg\neg p \models \neg\neg p$
- $\neg p, \neg\neg p \not\models p$

because of the assignment $h(p) = 0$.

The above shows that our definition of negation is compatible with paraconsistency according to da Costa but not necessarily with others (see [3, Chapter 2]). For examples, \neg is not a negation in those logics which allow for $\neg d \notin D$ for some designated $d (d \in D)$ in V .

Example 4.5 (The system J_3). The paraconsistent system J_3 of [29, 32] has 3 values, $0, \frac{1}{2}, 1$ with $\neg 0 = 1, \neg \frac{1}{2} = \frac{1}{2}, \neg 1 = 0$. According to Theorem 4.2, this system does not have a negation in our sense.

Example 4.6. Consider the negation examples discussed in Section 4.5 of [2]. We check them one by one. They all cannot be negations because of Theorem 4.2.

System P_1 . $V = \{t, f, \mathbf{T}\}$. $D = \{t, \mathbf{T}\}$. Table for \neg :

$$\begin{aligned}\neg t &= f \\ \neg f &= \mathbf{T} \\ \neg \mathbf{T} &= t.\end{aligned}$$

System SRM_{\simeq} . $V = \{t, f, \mathbf{T}\}$. $D = \{t, \mathbf{T}\}$. Table for negation:

$$\begin{aligned}\neg t &= f \\ \neg f &= t \\ \neg \mathbf{T} &= \mathbf{T}\end{aligned}$$

(This is also the same as J_3 , at least as far as \neg is concerned.)

System LP. $V = \{t, f, \mathbf{T}\}$. $D = \{t, \mathbf{T}\}$. Table for \neg : Same as SRM_{\simeq} .

They both have conjunction but with different truth tables.

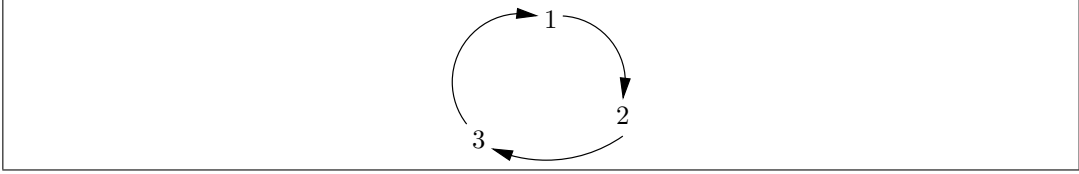


Figure 1

System PAC(RM₃). The table for \neg is the same as **LP** but different table for implication.

System PAC₃. Another extension of **PAC**. Same table for \neg .

Example 4.7. Consider a language with a single propositional atom q and a unary symbol \neg . Assume we have the two axioms

$$\begin{aligned} q &\models \neg\neg\neg q \\ \neg\neg\neg q &\models q. \end{aligned}$$

The semantics for this can be a temporal cycle chain as in Figure 1

We have $t \models \neg A$ iff $t \oplus 1 \models A$ where \oplus is addition modulou 3.

So we have

$$\begin{aligned} 1 \oplus 1 &= 2 \\ 2 \oplus 1 &= 3 \\ 3 \oplus 1 &= 1 \end{aligned}$$

This interpretation of \neg is simply “next in the cycle”. One would not think of it as negation. Let us try to introduce negation.

1. Say we do not want q to be true at node 2. So we do not want $\neg q$ to be true at node 1 and we do not want $\neg\neg q$ to be true at node 3.

We cannot express this negation because we do not have the ability to name worlds.

2. What we can say is something like “we do not want all three worlds to have the same value”. We express this by saying that:

$$\Delta = \{q, \neg q, \neg\neg q\}$$

is inconsistent.

But how do we express “inconsistent”? We cannot unless we add conjunction to the language with the usual axioms

$$\begin{aligned} \{A, B\} &\models A \wedge B \\ A \wedge B &\models B \\ A \wedge B &\models A \\ \neg(A \wedge B) &\models \neg A \wedge \neg B \end{aligned}$$

and we let

$$\Theta_{\neg} = \{p \wedge \neg p \wedge \neg\neg p\}.$$

But we also have

$$p \wedge \neg p \wedge \neg\neg p \models p \wedge \neg p \wedge \neg\neg p.$$

Therefore

$$\models \neg(p \wedge \neg p \wedge \neg\neg p)$$

and therefore

$$\models (p \wedge \neg p \wedge \neg\neg p).$$

A contradiction.

The pure Paraconsistent Enthusiast would say that the contradiction means only that the author’s definition of negation does not apply here, and that they are entitled to view \neg as paraconsistent negation, because we have:

$$\{p, \neg p\} \models p, \{p, \neg p\} \models \neg p,$$

but we DO NOT HAVE

$$\{p, \neg p\} \models \neg\neg p,$$

and this satisfies the da Costa definition of negation. However, we do know what \neg is; it is modal tomorrow, not a negation! This makes our case stronger that \neg is not a negation.

4.2 Other negations

Example 4.8 (Arnon Avron non-deterministic negation). Consider a language with atoms Q and one candidate for negation symbol \neg . Consider the following truth values for atoms

(*1) Truth values: $\{\mathbf{t}, \mathbf{f}\}$.

Truth table of \neg

$$\neg\mathbf{t} = \{\mathbf{f}\}, \neg\mathbf{f} = \{\mathbf{t}, \mathbf{f}\}.$$

(*1) is a nondeterministic system. This means that if we give an atom q the values \mathbf{t} , then $\neg q$ has value \mathbf{f} . If q has value \mathbf{f} then $\neg q$ can be either \mathbf{t} or \mathbf{f} and it is our choice to decide which one!

Observation 1. *The elements of the set $\Delta = \{\neg A, A\}$ cannot be all true. The reason is that if A has value \mathbf{t} then $\neg A$ must have value \mathbf{f} .*

Define a consequence relation between sets Δ of wffs and a single wff A .

(*2) $\Delta \models A$ iff for any assignment of truth values to the atoms, if all elements of Δ get value \mathbf{t} , then A gets value \mathbf{t} .

Observation 2. *For any wffs A and B we have*

$$\{A, \neg A\} \models B.$$

We now show that \neg cannot be a negation in our sense. If \neg were a negation, then for some unwanted wff x we have:

(*3) $\Delta, A \models x$ iff $\Delta \models \neg A$.

*In view of Observation 2, we get, using (*1) and (*2) that*

(*4) $q \models \neg\neg q$.

This, however, is not possible because we can let value q to be \mathbf{t} . Then value $\neg q$ is \mathbf{f} and value $\neg\neg q$ can also be chosen to be \mathbf{f} .

Example 4.9 (Post 3-valued negation). *Consider a language with atoms Q and three truth values $\{0, \frac{1}{2}, 1\}$. Let \neg be post negation with the table*

$$\neg 0 = \frac{1}{2}, \neg \frac{1}{2} = 1, \neg 1 = 0.$$

We show that \neg is not a negation in our sense.

Observation 1. *For any assignment of values and any A*

(*1) *value $A = \text{value } \neg\neg A$.*

We need to define \models .

Let $\Delta \models A$ mean

$$\min\{Val(x) \mid x \in \Delta\} \leq Val(A).$$

*Here we regard 1 as truth and 0 as false. Assume \neg is a negation in our sense, and let $y = \neg^m q$ be the unwanted element. Note that we can assume $0 \leq m \leq 2$ because of (*1).*

Observation 2. $\Delta = \{q, \neg q, \neg\neg q\} \vDash y$.

The reason is that $\min\{Val(x)|x \in \Delta\} = 0$.

Assume without loss of generality that $m = 1$.

Then focussing on $\neg q \in \Delta$ we get $\{q, \neg\neg q\} \vDash \neg q$.

This is not possible because let $Val(q) = 1$. Then $Val(\neg q) = 0$, $Val(\neg\neg q) = \frac{1}{2}$.

We get

$$\min\{Val(x)|x \in \Delta\} = \frac{1}{2}$$

and $Val(\neg q) = 0$.

5 Conclusion and future research

Let us summarise what we have learnt in this position paper about the question of what is negation in a system.

1. Assume a logical language with well formed formulas and a relation \vdash between multisets Δ of wffs and a single wff A of the form

$$\Delta \vdash A.$$

We need not assume any properties of “ \vdash ” nor do we need to know how \vdash is defined.

2. We put forward the basic intuition that a unary connective $*$ in the language, is a negation in \vdash if for every Δ there exists a multiset of wffs $\Theta_*(\Delta)$ such that $\Delta \vdash *A$ iff for some $y \in \Theta_*(\Delta)$

$$\Delta \cup \{A\} \vdash y.$$

3. This intuitive definition works in one form or another also for non-monotonic consequence systems, such as negation as failure (that is why we have that Θ_* depends on Δ). For a monotonic consequence relation we would expect that Θ_* would be the same for any Δ .

We did observe, however, that we might introduce a negation $*(\Delta)$ for each Δ and write some axioms connecting all the $*(\Delta)$ negations. We gave some hints in Section 2, on how this can be done, for the case of negation as failure added to intuitionistic implication.

4. There are systems \vdash with $*$ where $*$ is not a negation.

5. Most of the well known systems with $*$, which are considered as a negation, are also negations according to our definition, but not all of them.
6. There are systems such as paraconsistency systems, where the question of whether their negation candidate is indeed a negation is debated in the literature (see [8]). Our approach might be able to offer a verdict.
7. We show that our definition is compatible with da Costa's definition of paraconsistency but not necessarily with other definition of paraconsistency.
8. There is a need for a systematic examination of all candidates for negation in the literature (especially to map the rich field of paraconsistent negations, see [23]) with a view to improve our definition of what is negation and possibly also refute some community misconceptions.
9. J.Y. Beziau wrote to me on 5.11.2020 (in response to the current paper) as follows:

“Dov is dealing with the question of paraconsistent negation from the point of view of many-valued semantics. I have also been working a bit in this framework, especially studying the question of the replacement theorem see [4]: I think Dov is not dealing with that, so I ask him: does he think that a negation without replacement theorem is a negation? From a philosophical point of view and in relation with the square of opposition I have argued that paracomplete and paraconsistent negations can naturally be considered as kinds of negation, cf. my paper in honor of John Woods [7].”

I will address this in [23].

Meanwhile for the benefit of the reader, let me quote from [4]:

“An important positive property at the level of meta-properties is the replacement property (not to be confused with the substitution property). A logic has this property if it allows to replace an occurrence of a formula by a logically equivalent one, where two formulas ϕ and ψ are called equivalent in a logic (notation: $\phi \dashv\vdash \psi$) if each of them can be deduced in that logic from the other. The replacement property is an important meta-property valid in classical logic, intuitionistic logic, all intermediate logics, all normal modal logics, and many more. In contrast, most of the paraconsistent logics investigated in the literature do not have this property”

References

- [1] A. R. Anderson and N. Belnap. *Entailment 1*, Princeton University Press, 1975.
- [2] Ofer Arieli and Arnon Avron. Three-Valued Paraconsistent Propositional Logics, in J.-Y. Beziau et al. (eds.), *New Directions in Paraconsistent Logic*, Springer Proceedings in Mathematics & Statistics 152, pp 91–128 DOI 10.1007/978-81-322-2719-9_4
- [3] Arnon Avron, Ofer Arieli, Anna Zamansky. *Theory of Effective Propositional Paraconsistent Logics*, 576 pages College Publications (21 May 2018)
- [4] A. Avron and J.-Y. Beziau, “Self-extensional three-valued paraconsistent logics have no implication”, *Logic Journal of the IGPL*, Volume 25, Issue 2 (April 2017), pp.183-194.
- [5] Avron A., Zamansky A. (2011) Non-Deterministic Semantics for Logical Systems. In: Gabbay D., Guenther F. (eds) *Handbook of Philosophical Logic*, vol 16. Springer, Dordrecht, pp 227-304
- [6] Batens, D., de Clercq, K.: A rich paraconsistent extension of full positive logic. *Logique et Analyse* 185D188, 220D257 (2004)
- [7] J.-Y. Beziau, “Cats that are not cats”, in Dov Gabbay, Lorenzo Magnani, Woosuk Park, and Ahti Veikko Pietarinen (eds), *Natural Arguments - A Tribute to John Woods*, College Publications, London, 2019, pp.49-71.
- [8] Jean-Yves Beziau. Are paraconsistent negations negations? <https://www.jyb-logic.org/papers12-11/paraconsistent%20negations.pdf>
- [9] W. A. Carnielli and J. Marcos. A taxonomy of C-systems. In W. A. Carnielli, M. E. Coniglio, and I. D’Ottaviano, editors, *Paraconsistency: The Logical Way to the Inconsistent*, number 228 in Lecture Notes in Pure and Applied Mathematics, pages 1-94. Marcel Dekker, 2002.
- [10] A. Church. The weak theory of implications, *Kontrolliertes Dluken*, pp. 22-37. J. S. L. vol. 16, pp. 236, 1951.
- [11] N. C. A. da Costa. On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic*, 15:497–510, 1974.
- [12] D. Gabbay. *Semantical Investigations in Heyting’s Intuitionistic Logic* (Monograph) Synthese volume 148. D. Reidel, 1981. 300 pp,
- [13] D. Gabbay Theoretical Foundations for Non-monotonic Reasoning in Expert Systems. In *Logics and Models of Concurrent Systems*, K. Apt, ed. pp 439–459. Springer-Verlag, 1985.
- [14] D.M. Gabbay. N-prolog: An extension of prolog with hypothetical implication. Part 2. logical foundations, and negation as failure. *The Journal of Logic Programming*, 2(4):251–283, 1985.
- [15] D. Gabbay. What is Negation in a System? In *Logic Colloquium ’86*, F. R. Drake and J. K. Truss, eds. pp 95–112. Elsevier Science Publishers (North Holland), 1986.
- [16] Gabbay, D.M., 1989, “Modal provability foundations for negation as failure”, in Extensions of Logic Programming, P. Schroeder-Heister, ed., Springer-Verlag, vol. 475 of LNC, pp 179–222

- [17] D Gabbay. Algorithmic Proof with Diminishing Resources Part I In *Proceedings of CSL'90*, pp 156–173. Vol. 533 of LNCS, Springer-Verlag, 1990.
- [18] D. Gabbay. Conditional Implications and Non-monotonic Consequence In *Views on Conditional*, a volume in *Studies in Logic and Computation*, L. Farinas del Cerro et al., eds. pp. 347–369. Oxford University Press, 1995.
- [19] D. M. Gabbay (Editor) *What Is a Logical System?* (Studies in Logic and Computation) Hardcover £ 1 Oct. 1996, 468 pages D. M. Gabbay (Editor)
- [20] Dov M. Gabbay (Editor), Heinrich Wansing (Editor): *What is Negation?* (Applied Logic Series) Hardcover £ 31 Mar. 1999, 356 pages
- [21] D. Gabbay and M. Sergot. Negation as inconsistency. *J Logic Programming*, 1986:1:1-35
- [22] D. Gabbay. What is negation as failure. This is a revised version of an unpublished paper written in 1985. Now published in A. Artikis et al. (Eds.): *Teleo-Reactive Abductive Logic Programs*, Marek Sergot Festschrift, LNAI 7360, pp. 52–78. Springer, Heidelberg (2012)
- [23] Dov Gabbay, What is Para-consistent “Negation”, paper in planning, once I finish my current commitments.
- [24] Dov Gabbay and Ross Horne University of Luxembourg Proof theory of positive linear intuitionistic logic with negation as failure and anti-formulas. Draft paper 637A: May 21, 2020.
- [25] D Gabbay and N Olivetti. *Goal Directed Algorithmic Proof Theory* (Monograph) Kluwer Academic Publishers, 2000. 266 pp.
- [26] Gelfond, Michael (2008). “Answer sets”. In van Harmelen, Frank; Lifschitz, Vladimir; Porter, Bruce (eds.). *Handbook of Knowledge Representation*. Elsevier. pp. 285–316
- [27] Petr Hájek. *Metamathematics of Fuzzy Logic* (Trends in Logic) Paperback, 4 Oct. 2013, 312 pages Springer
- [28] A. Meyer. *Topics in Modal and Many-valued Logics*. Thesis, 1966.
- [29] C.A. Middelburg On the Strongest Three-Valued Paraconsistent Logic Contained in Classical Logic and Its Dual, Arxive 2020
- [30] S. K. Moh. The deduction theorems and two new logical systems. *Methods* vol. 2, pp. 56-75, 1950.
- [31] Nicola Olivetti and Lea Terracini. N-prolog and equivalence of logic programs *Journal of Logic, Language and Information* 1 (4):253-340 (1992)
- [32] Itala M.L. D’Ottaviano. The competeness and compactness of a three-valued first-order logic. *Revista Colombiana de Matematicas* Vol. XIX (1985), pp 77–94
- [33] M. Wajsberg. Axiomatization of the three valued propositional calculus (in Polish). *Comptes rendus des seances de la Societe des Sciences et des Lettres de varsovie*, classe III, vol. 24, pp. 126-148, 1931.
Translation in S. McCall editor *Polish Logic, 1920-1939* Oxford University Press 1967, pp. 264-284.

THE CLASS OF ALL 3-VALUED IMPLICATIVE EXPANSIONS OF KLEENE'S STRONG LOGIC CONTAINING ANDERSON AND BELNAP'S *FIRST DEGREE ENTAILMENT LOGIC*

GEMMA ROBLES

Dpto. de Psicología, Sociología y Filosofía, Universidad de León.

Campus de Vegazana, s/n, 24071, León, Spain

`gemma.robles@unileon.es`

Abstract

We define all 3-valued implicative expansions of Kleene's strong 3-valued matrix (with both only one and two designated values) verifying Anderson and Belnap's *First degree entailment logic*, FDE. Then, the logics determined by each one of these implicative expansions are axiomatized by using a Belnap-Dunn 'two-valued' semantics. This semantics is 'overdetermined' in the case of two designated values and 'underdetermined' when there is only one.

Keywords: 3-valued logic; Kleene's strong 3-valued logic; Belnap-Dunn 'two-valued' semantics; *First degree entailment logic* FDE; implicative logic.

1 Introduction

The aim of this paper is to define the class of all 3-valued implicative expansions of Kleene's strong 3-valued logic containing Anderson and Belnap's *First degree entailment logic*, FDE, also known as Belnap and Dunn's 4-valued logic. To the best of our knowledge, most of these logics have not been given Hilbert-style formulations in the literature (but cf. [26], §6), and some of them have interesting properties such as a considerable functional strength, paraconsistency or paracompleteness and the

I sincerely thank two anonymous referees of the JAL for their comments and suggestions on a previous draft of this paper.

- This work is supported by the Spanish Ministry of Economy, Industry and Competitiveness under Grant [FFI2017-82878-P].

‘quasi variable-sharing property’. Also, we think that they can naturally lead to 4-valued implicative expansions of FDE, as pointed out below in this introduction.

As remarked in [21], FDE is a particularly interesting and useful non-classical logic. FDE is the minimal logic in the De Morgan family of Anderson and Belnap-style relevant logics. It can be viewed as a 4-valued logic in which formulas can be both true and false or neither true nor false, in addition to being true or false (cf. [1], [4], [5], [7], [8]).

The question of expanding FDE with a full implication connective poses itself since, as the name of the logic suggests, formulas of the form $A \rightarrow B$ are not considered in FDE if either A or B contains \rightarrow (cf. [1]). Some full implicative expansions of FDE have been defined in the literature (cf. [6], [27], [18], [19], [21] and references in the last item), but there is still a lot of investigation to be done in the topic (cf. [21]).

Regarding the relationship between Kleene’s strong 3-valued logic and FDE, Fitting notes: “It is well-known that Kleene’s strong 3-valued logic finds a natural extension in Belnap’s four-valued logic” (cf. [11], p. 115; cf. also [10]). In this same sense of the term extension (as the enlargement of the set of truth-values from three to four), implicative expansions of Kleene’s strong 3-valued logic containing FDE could naturally lead to 4-valued implicative expansions of FDE. Actually, we note that the truth-tables representing the implication function in some of the aforementioned implicative expansions of FDE (cf. [6], [27], [18], [19] and [21]) can be ‘divided’ in two 3-valued tables corresponding to the ‘both’ and ‘neither’ parts in the 4-valued tables (cf. [27]).

The present paper is a preliminary study on 3-valued logics containing FDE (cf. [26]). We focus on expansions of Kleene’s strong 3-valued matrix MK3 with two designated values (MK3 with only one designated value is treated in section 5 of the paper). The paper is organized as follows.

In §2, we define the class of all implicative expansions of MK3 (with two designated values) verifying Anderson and Belnap’s logic FDE. (In §5, implicative expansions of MK3 (with only one designated value) verifying FDE are investigated.) By M_i ($1 \leq i \leq 64$), we refer to the 64 implicative expansions of MK3 defined in this section. In §3, the logics determined by each one of the aforesaid implicative matrices are defined. By L_i , we refer to the logic determined by the matrix M_i ($1 \leq i \leq 64$). The L_i -logics are defined in a general and unified way from two different basic logics b_1^3 and b_2^3 . Also, we prove some properties of the L_i -logics that will be useful in the proof of the completeness theorems. In §4, a Belnap-Dunn two-valued overdetermined semantics is provided for each L_i -logic. Following the strategy in [6], as displayed in [29], strong soundness of each L_i -logic w.r.t. its corresponding semantics is proved. Concerning completeness, only that of L_i -logics built upon b_1^3 is

proved: Li -logics defined from b_2^3 present special difficulties and their completeness is proved in §5 along with that of the logics determined by the implicative expansions of MK3 with only one designated value. In order to prove completeness for all the logics treated in the paper, we follow the strategy set up in [34], as applied in [6] and particularly displayed in [29] and [33]. Thus, we think that it is possible to be reasonably general about the details, in this way being able to shorten the otherwise necessary long proofs (cf., e.g., [6], [29] or [33]). In §5, the implicative expansions of MK3 (only one designated value) verifying FDE are investigated. As pointed out above, the completeness of the extensions of b_2^3 is also proved. In §6, we discuss the notion of implication generally characterizing the implicative expansions of Kleene's strong logic containing FDE. Also, we remark some of the features of Brady's method as displayed in [29] and [33] and compare it with the 'correspondence analysis' method used in, e.g., [16] and [38]¹. Finally, in §7, we prove some properties of all logics studied in the paper w.r.t. paraconsistency, paracompleteness, 'variable-sharing property', 'quasi variable-sharing property' and functional strength. As remarked above, the properties of the logics investigated w.r.t. these notions make clear that they are interesting in themselves independently of their applicability in the definition of valuable 4-valued implicative expansions of FDE. An appendix is added remarking some questions related to the topic of the paper.

2 Implicative expansions of MK3 containing FDE

In this section, all 3-valued implicative expansions of Kleene's strong 3-valued matrix MK3 containing Anderson and Belnap's logic FDE are defined. We begin by defining MK3 and FDE.

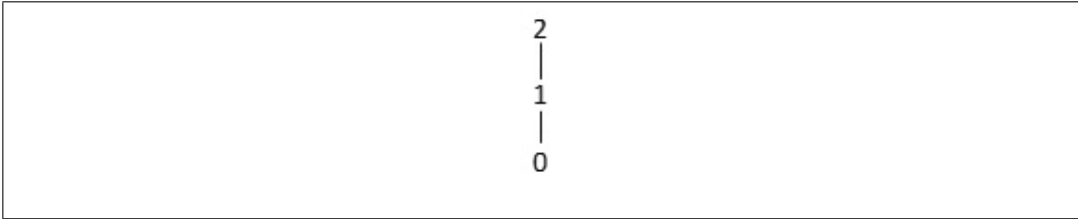
The propositional language consists of a denumerable set of propositional variables $p_0, p_1, \dots, p_n, \dots$, and some or all of the following connectives \rightarrow (conditional), \wedge (conjunction), \vee (disjunction), \neg (negation). The biconditional (\leftrightarrow) and the set of wffs are defined in the customary way. A, B , etc. are metalinguistic variables. Then, logics are formulated as Hilbert-type axiomatic systems, while the following notions are understood in a fairly standard sense (cf., e.g., [29] or [33]): *extension* and *expansion* of a given logic; *logical matrix* M and *M -interpretation*, *M -consequence*, *M -validity*, and finally, *M -determined logic*.

Kleene's strong matrix MK3 can be defined as follows (cf. [15]).

Definition 2.1 (Kleene's strong 3-valued matrix). *The propositional language consists of the connectives \wedge, \vee, \neg . Kleene's strong 3-valued matrix, MK3 (our label),*

¹To address this question was suggested by a referee of the JAL.

is the structure $(\mathcal{V}, D, \mathbf{F})$, where (1) $\mathcal{V} = \{0, 1, 2\}$ and it is ordered as shown in the following lattice



(2) $D = \{1, 2\}$ or $D = \{2\}$; (3) $\mathbf{F} = \{f_\wedge, f_\vee, f_\neg\}$, where f_\wedge and f_\vee are defined as the glb (or lattice meet) and the lub (or lattice joint), respectively, and f_\neg is an involution with $f_\neg(2) = 0, f_\neg(0) = 2$ and $f_\neg(1) = 1$. We display the tables for \wedge, \vee and \neg :

\wedge	0	1	2	\vee	0	1	2	\neg	0
0	0	0	0	0	0	1	2	0	2
1	0	1	1	1	1	1	2	1	1
2	0	1	2	2	2	2	2	2	0

Remark 2.2 (On the ordering of the truth-values in MK3). *Regarding the ordering of the truth-values in MK3, Fitting notes “The informal reading suggests two natural orderings, concerning ‘amount of knowledge’ and ‘degree of truth’ ([10], p. 797; cf. also [11]). Of course, the elements of \mathcal{V} in Definition 2.1 are ordered according to degree of truth.*

On the other hand, Anderson and Belnap’s FDE is defined as follows (cf. [1], p. 158). (Other formulations are provided in [21] and [35]. We note that the rule Con does not appear in the axiomatizations to follow — cf. section 5 on the question of the presence of rules of inference in the *Li*-logics.)

Definition 2.3 (Anderson and Belnap’s FDE). *Anderson and Belnap’s propositional logic FDE can be defined as follows.*

Axioms:

- $\alpha 1.$ $(A \wedge B) \rightarrow A; (A \wedge B) \rightarrow B$
- $\alpha 2.$ $A \rightarrow (A \vee B); B \rightarrow (A \vee B)$
- $\alpha 3.$ $[A \wedge (B \vee C)] \rightarrow [(A \wedge B) \vee (A \wedge C)]$
- $\alpha 4.$ $A \rightarrow \neg\neg A$
- $\alpha 5.$ $\neg\neg A \rightarrow A$

Rules of inference:

Modus Ponens (MP): $A \rightarrow B, A \Rightarrow B$

Conditioned introduction of conjunction ($I \wedge$): $A \rightarrow B, A \rightarrow C \Rightarrow$
 $A \rightarrow (B \wedge C)$

Elimination of disjunction ($E \vee$): $A \rightarrow C, B \rightarrow C \Rightarrow$
 $(A \vee B) \rightarrow C$

Transitivity (Trans): $A \rightarrow B, B \rightarrow C \Rightarrow$
 $A \rightarrow C$

Contraposition (Con): $A \rightarrow B \Rightarrow \neg B \rightarrow \neg A$

We have added the rule MP to Anderson and Belnap's original axiomatization. Of course, MP is added in order to extend FDE (cf. [36]).

Next, we proceed to determine the class of all 3-valued implicative expansions of MK3 verifying FDE. We focus in the case where 1 and 2 are designated values. The case where 2 is the only designated value is treated in §5.

Consider the matrix MK3, 1 and 2 being designated values, and let $a_i (1 \leq i \leq 6) \in \{1, 2\}$. It is clear that in order to verify the self-identity thesis, $A \rightarrow A$, ($\alpha 1$) $(A \wedge B) \rightarrow A$ and the rule MP, \rightarrow -functions have to fulfill the conditions appearing in the table ta.

	\rightarrow	0	1	2
ta.	0	a_1	a_2	a_3
	1	0	a_4	a_5
	2	0		a_6

On the other hand, table ta has to be extended to table tb

	\rightarrow	0	1	2
tb.	0	a_1	a_2	a_3
	1	0	a_4	a_5
	2	0	0	a_6

on pain of the rule Con being falsified. But, given MK3, all \rightarrow -functions in tb verify $I \wedge$, $E \vee$ and Trans. Consequently, there are exactly 64 expansions of MK3 (1 and 2 being designated values) verifying FDE: those built up by adding one of the 64 \rightarrow -functions in table tb to MK3².

²As suggested by a referee of the JAL, the notion of implication tb embodies is discussed in §6.

Next, we display the 64 tables in the general table *tb*. For any i ($1 \leq i \leq 64$), by Mi we refer to the implicative expansion of MK3 built up by adding the \rightarrow -function described by table t_i^3 .

t1	0	1	2	t2	0	1	2	t3	0	1	2	t4	0	1	2
0	2	2	2	0	2	2	1	0	2	1	2	0	2	1	1
1	0	2	2	1	0	2	2	1	0	2	2	1	0	2	2
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2
t5	0	1	2	t6	0	1	2	t7	0	1	2	t8	0	1	2
0	2	2	2	0	2	2	1	0	2	1	2	0	2	1	1
1	0	1	2	1	0	1	2	1	0	1	2	1	0	1	2
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2
t9	0	1	2	t10	0	1	2	t11	0	1	2	t12	0	1	2
0	2	2	2	0	2	2	1	0	2	1	2	0	2	1	1
1	0	2	1	1	0	2	1	1	0	2	1	1	0	2	1
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2
t13	0	1	2	t14	0	1	2	t15	0	1	2	t16	0	1	2
0	2	2	2	0	2	2	1	0	2	1	2	0	2	1	1
1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2
t17	0	1	2	t18	0	1	2	t19	0	1	2	t20	0	1	2
0	2	2	2	0	2	2	1	0	2	1	2	0	2	1	1
1	0	2	2	1	0	2	2	1	0	2	2	1	0	2	2
2	0	0	1	2	0	0	1	2	0	0	1	2	0	0	1
t21	0	1	2	t22	0	1	2	t23	0	1	2	t24	0	1	2
0	2	2	2	0	2	2	1	0	2	1	2	0	2	1	1
1	0	1	2	1	0	1	2	1	0	1	2	1	0	1	2
2	0	0	1	2	0	0	1	2	0	0	1	2	0	0	1
t25	0	1	2	t26	0	1	2	t27	0	1	2	t28	0	1	2
0	2	2	2	0	2	2	1	0	2	1	2	0	2	1	1
1	0	2	1	1	0	2	1	1	0	2	1	1	0	2	1
2	0	0	1	2	0	0	1	2	0	0	1	2	0	0	1

³In the Appendix these tables, as well as those in §5, are classified in ‘subclassical’ and ‘contra-classical’ tables. Also, w.r.t. the presence or absence of Aristotle’s thesis, $\neg(A \rightarrow \neg A)$ or $\neg(\neg A \rightarrow A)$, and Boethius’ thesis, $(A \rightarrow B) \rightarrow \neg(A \rightarrow \neg B)$ or $(A \rightarrow \neg B) \rightarrow \neg(A \rightarrow B)$. (I owe this suggestion to a referee of JAL.)

t29	0	1	2	t30	0	1	2	t31	0	1	2	t32	0	1	2
0	2	2	2	0	2	2	1	0	2	1	2	0	2	1	1
1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
2	0	0	1	2	0	0	1	2	0	0	1	2	0	0	1

Tables t33 through t64 are exactly as t1 through t32, respectively, except that now $f_{\rightarrow}(0,0) = 1$. Thus, these tables are:

t33	0	1	2	t34	0	1	2	t35	0	1	2	t36	0	1	2
0	1	2	2	0	1	2	1	0	1	1	2	0	1	1	1
1	0	2	2	1	0	2	2	1	0	2	2	1	0	2	2
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2

t37	0	1	2	t38	0	1	2	t39	0	1	2	t40	0	1	2
0	1	2	2	0	1	2	1	0	1	1	2	0	1	1	1
1	0	1	2	1	0	1	2	1	0	1	2	1	0	1	2
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2

t41	0	1	2	t42	0	1	2	t43	0	1	2	t44	0	1	2
0	1	2	2	0	1	2	1	0	1	1	2	0	1	1	1
1	0	2	1	1	0	2	1	1	0	2	1	1	0	2	1
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2

t45	0	1	2	t46	0	1	2	t47	0	1	2	t48	0	1	2
0	1	2	2	0	1	2	1	0	1	1	2	0	1	1	1
1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
2	0	0	2	2	0	0	2	2	0	0	2	2	0	0	2

t49	0	1	2	t50	0	1	2	t51	0	1	2	t52	0	1	2
0	1	2	2	0	1	2	1	0	1	1	2	0	1	1	1
1	0	2	2	1	0	2	2	1	0	2	2	1	0	2	2
2	0	0	1	2	0	0	1	2	0	0	1	2	0	0	1

t53	0	1	2	t54	0	1	2	t55	0	1	2	t56	0	1	2
0	1	2	2	0	1	2	1	0	1	1	2	0	1	1	1
1	0	1	2	1	0	1	2	1	0	1	2	1	0	1	2
2	0	0	1	2	0	0	1	2	0	0	1	2	0	0	1

t57	0	1	2	t58	0	1	2	t59	0	1	2	t60	0	1	2
0	1	2	2	0	1	2	1	0	1	1	2	0	1	1	1
1	0	2	1	1	0	2	1	1	0	2	1	1	0	2	1
2	0	0	1	2	0	0	1	2	0	0	1	2	0	0	1

t61	0	1	2	t62	0	1	2	t63	0	1	2	t64	0	1	2
0	1	2	2	0	1	2	1	0	1	1	2	0	1	1	1
1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
2	0	0	1	2	0	0	1	2	0	0	1	2	0	0	1

3 Li -logics and some of their properties

By Li , we refer to the logic determined by the matrix Mi . All Li -logics are extensions of one of the two basic logics b_1^3 and b_2^3 , defined below. In particular, L2, L6, L18, L22, L34, L50 and L54 are extensions of b_2^3 , while the rest of the Li -logics are extensions of b_1^3 . (The reason for distinguishing between two basic logics is explained in section 5.) Both b_1^3 and b_2^3 (the former included in the latter) are built upon the more basic logic b^3 . The label b^3 is intended to abbreviate ‘basic logic contained in all Li -logics’⁴.

Definition 3.1 (The logic b^3). *The logic b^3 is formulated as follows:*

Axioms:

- a1. $(A \wedge B) \rightarrow A; (A \wedge B) \rightarrow B$
- a2. $A \rightarrow (A \vee B); B \rightarrow (A \vee B)$
- a3. $[(A \rightarrow B) \wedge (A \rightarrow C)] \rightarrow [A \rightarrow (B \wedge C)]$
- a4. $[(A \rightarrow C) \wedge (B \rightarrow C)] \rightarrow [(A \vee B) \rightarrow C]$
- a5. $[A \wedge (B \vee C)] \rightarrow [(A \wedge B) \vee (A \wedge C)]$
- a6. $[(A \rightarrow B) \wedge A] \rightarrow B$
- a7. $A \vee (A \rightarrow B)$
- a8. $\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$
- a9. $\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$
- a10. $A \leftrightarrow \neg\neg A$
- a11. $A \vee \neg A$
- a12. $\neg B \vee (A \rightarrow B)$
- a13. $[(A \rightarrow B) \wedge \neg B] \rightarrow \neg A$
- a14. $[(A \wedge \neg A) \wedge B] \rightarrow (A \rightarrow B)$

⁴In the Appendix, this basic logic b^3 is compared to other basic logics previously used in the literature to axiomatize implicative expansions of MK3.

Rules of inference:

$$\begin{aligned} \text{Adjunction (Adj): } & A, B \Rightarrow A \wedge B \\ \text{Modus Ponens (MP): } & A \rightarrow B, A \Rightarrow B \end{aligned}$$

Definition 3.2 (The logic b_1^3). *The logic b_1^3 is axiomatized with (a15) $[(A \rightarrow B) \wedge (B \rightarrow C)] \rightarrow (A \rightarrow C)$ in addition to a1-a14 and Adj and MP of b^3 .*

Definition 3.3 (The logic b_2^3). *The logic b_2^3 is axiomatized with the rule Disjunctive Transitivity (dTrans) $D \vee [(A \rightarrow B) \wedge (B \rightarrow C)] \Rightarrow D \vee (A \rightarrow C)$ in addition to a1-a14 and Adj and MP of b^3 .*

The Li-logics are axiomatized with some subset of the following set of axioms⁵.

- A1. $A \vee \neg(A \rightarrow B)$
- A2. $B \vee \neg(A \rightarrow B)$
- A3. $\neg A \vee \neg(A \rightarrow B)$
- A4. $\neg B \vee \neg(A \rightarrow B)$
- A5. $(A \vee \neg B) \vee \neg(A \rightarrow B)$
- A6. $(\neg A \vee \neg B) \vee \neg(A \rightarrow B)$
- A7. $A \rightarrow [B \vee \neg(A \rightarrow B)]$
- A8. $\neg B \rightarrow [\neg A \vee \neg(A \rightarrow B)]$
- A9. $(A \wedge \neg B) \rightarrow \neg(A \rightarrow B)$
- A10. $(A \wedge \neg A) \rightarrow \neg(A \rightarrow B)$
- A11. $(B \wedge \neg B) \rightarrow \neg(A \rightarrow B)$
- A12. $(A \wedge \neg A) \rightarrow [\neg B \vee \neg(A \rightarrow B)]$
- A13. $(B \wedge \neg B) \rightarrow [A \vee \neg(A \rightarrow B)]$
- A14. $[\neg(A \rightarrow B) \wedge \neg A] \rightarrow A$
- A15. $[\neg(A \rightarrow B) \wedge \neg A] \rightarrow \neg B$
- A16. $[\neg(A \rightarrow B) \wedge \neg A] \rightarrow (A \vee B)$
- A17. $[\neg(A \rightarrow B) \wedge \neg A] \rightarrow (A \vee \neg B)$

⁵A1-A6 are 'contraclassical' wffs, the rest of the theses in the list being 'subclassical' wffs. (The topic is pursued in the Appendix. I owe the suggestion of treating this question to a referee of the JAL.)

- A18. $[\neg(A \rightarrow B) \wedge B] \rightarrow A$
A19. $[\neg(A \rightarrow B) \wedge B] \rightarrow \neg B$
A20. $[\neg(A \rightarrow B) \wedge B] \rightarrow (A \vee \neg B)$
A21. $[\neg(A \rightarrow B) \wedge B] \rightarrow (\neg A \vee \neg B)$
A22. $[\neg(A \rightarrow B) \wedge (A \wedge B)] \rightarrow \neg B$
A23. $[\neg(A \rightarrow B) \wedge (\neg A \wedge B)] \rightarrow C$
A24. $[\neg(A \rightarrow B) \wedge (\neg A \wedge \neg B)] \rightarrow A$
A25. $[\neg(A \rightarrow B) \wedge (A \wedge \neg A)] \rightarrow \neg B$
A26. $[\neg(A \rightarrow B) \wedge (B \wedge \neg B)] \rightarrow A$
A27. $[\neg(A \rightarrow B) \wedge [(A \wedge \neg A) \wedge B]] \rightarrow C$
A28. $[\neg(A \rightarrow B) \wedge [(B \wedge \neg B) \wedge \neg A]] \rightarrow C$
A29. $[\neg(A \rightarrow B) \wedge [(A \wedge \neg A) \wedge (B \wedge \neg B)]] \rightarrow C$

In particular, we have:

Definition 3.4 (Extensions of b_1^3). *The following Li-logics are axiomatized by adding to b_1^3 the following axioms:*

- | | |
|---|--|
| <i>L1: A7, A8, A14, A19, A23.</i> | <i>L16: A5, A9, A10, A11, A16, A21.</i> |
| <i>L3: A7, A8, A13, A16, A19, A27.</i> | <i>L17: A3, A7, A14, A23.</i> |
| <i>L4: A5, A7, A8, A13, A16, A22,</i>
<i>A27.</i> | <i>L19: A3, A7, A13, A15, A16, A27.</i> |
| <i>L5: A9, A14, A19.</i> | <i>L20: A3, A5, A7, A13, A16, A27.</i> |
| <i>L7: A9, A11, A16, A19.</i> | <i>L21: A3, A9, A14, A15.</i> |
| <i>L8: A5, A9, A13, A16, A22.</i> | <i>L23: A3, A9, A11, A15, A16.</i> |
| <i>L9: A7, A8, A12, A14, A18, A21,</i>
<i>A28.</i> | <i>L24: A3, A5, A9, A11, A16, A25.</i> |
| <i>L10: A5, A7, A8, A12, A21, A24,</i>
<i>A28.</i> | <i>L25: A3, A7, A12, A14, A28.</i> |
| <i>L11: A7, A8, A12, A13, A16, A17,</i>
<i>A21, A29.</i> | <i>L26: A3, A4, A7, A24, A28.</i> |
| <i>L12: A5, A7, A8, A12, A13, A16,</i>
<i>A21, A29.</i> | <i>L27: A3, A7, A12, A13, A16, A17,</i>
<i>A29.</i> |
| <i>L13: A9, A10, A14, A21.</i> | <i>L28: A3, A4, A7, A13, A16, A29.</i> |
| <i>L14: A5, A9, A10, A21, A24.</i> | <i>L29: A3, A10, A14.</i> |
| <i>L15: A9, A10, A11, A16, A17, A21.</i> | <i>L30: A3, A5, A10, A24.</i> |
| | <i>L31: A3, A10, A13, A16, A17.</i> |
| | <i>L32: A4, A7, A11, A16.</i> |
| | <i>L33: A2, A8, A13, A19, A23.</i> |
| | <i>L35: A2, A8, A13, A19, A23.</i> |

- | | |
|---|--|
| <p>L36: A1, A2, A8, A22, A27.
 L37: A2, A9, A18, A19.
 L38: A2, A5, A8, A12, A21, A28.
 L39: A2, A11, A19.
 L40: A1, A2, A11, A22.
 L41: A2, A8, A12, A18, A21, A28.
 L42: A2, A5, A8, A12, A21, A28.
 L43: A2, A8, A12, A13, A20, A21,
 A29.
 L44: A1, A2, A8, A12, A21, A29.
 L45: A2, A8, A10, A18, A21.
 L46: A2, A5, A9, A10, A21, A26.
 L47: A2, A9, A10, A20, A21.
 L48: A1, A8, A10, A21.</p> | <p>L49: A2, A3, A23.
 L51: A2, A3, A13, A19, A27.
 L52: A1, A2, A3, A27.
 L53: A2, A3, A9, A15, A18.
 L55: A2, A6, A11, A15.
 L56: A1, A2, A6, A11, A25.
 L57: A2, A8, A12, A18, A21, A28.
 L58: A2, A3, A4, A28.
 L59: A2, A3, A12, A13, A20, A29.
 L60: A1, A2, A4, A8, A28.
 L61: A2, A3, A10, A23.
 L62: A2, A3, A4, A10, A14.
 L63: A2, A3, A10, A11, A20.
 L64: A1, A3, A10.</p> |
|---|--|

Definition 3.5 (Extensions of b_2^3). *The following Li-logics are axiomatized by adding to b_2^3 the following axioms:*

- L2: A5, A7, A8, A22, A24, A27, A28.
L6: A5, A9, A22, A24.
L18: A3, A5, A7, A24, A27, A28.
L22: A3, A5, A9, A24, A25.
L34: A2, A5, A8, A22, A27, A28.
L50: A2, A3, A5, A27, A28.
L54: A2, A3, A5, A9, A25, A26.

In what follows, we prove some properties of the Li-logics. We generally distinguish between extensions of b_1^3 , Eb_1^3 -logics, and extensions of b_2^3 , Eb_2^3 -logics. (By Eb^3 , we generally refer to extensions of b^3 ; cf. Definition 3.1.) Now, let L be an Eb_1^3 -logic. An L-theory is a set of wffs closed under Adj and MP, while if L is an Eb_2^3 -logic, an L-theory is a set of wffs closed under Adj, MP and dTrans (cf. Definitions 3.2 and 3.3). In addition, let L be an Eb^3 -logic. An L-theory is *regular* iff it contains all L-theorems; it is *trivial* iff it contains all wffs; it is *prime* if it has the *disjunction property*, and, finally, it is *complete* if for every wff A, it has either A or $\neg A$. Then, we have:

Proposition 3.6 (Some properties of prime, regular Eb^3 -theories). *Let L be an Eb^3 -logic and t be a prime, regular L-theory. Then, for any wffs A, B, (1) $A \wedge B \in t$ iff $A \in t$ and $B \in t$; (2) $\neg(A \wedge B) \in t$ iff $\neg A \in t$ or $\neg B \in t$; (3) $A \vee B \in t$ iff $A \in t$ or $B \in t$; (4) $\neg(A \vee B) \in t$ iff $\neg A \in t$ and $\neg B \in t$; (5) $A \in t$ iff $\neg\neg A \in t$; (6) $A \in t$ or $\neg A \in t$.*

Proof. Given the properties of t , it is immediate: (1)-(4), by the De Morgan laws (a8-a9); (5) by the double negation axiom (a10), and (6) by the principle of excluded middle (a11). (Notice that regularity is needed only in the last case, case (6).) \square

Concerning the conditional, we prove Propositions 3.7 and 3.8.

Proposition 3.7 (The conditional in prime, regular Eb^3 -theories). *Let L be an Eb^3 -logic and t a prime, regular L -theory. Then, $A \rightarrow B \in t$ iff $A \notin t$ or $\neg B \notin t$ or $(A \in t \ \& \ \neg A \in t \ \& \ B \in t)$.*

Proof. (a) (\Rightarrow) Suppose (1) $A \rightarrow B \in t$ and, for reductio, (2) $A \in t$, $\neg B \in t$ and $\neg A \notin t$ or (3) $A \in t$, $\neg B \in t$ and $B \notin t$. But 2 and 3 are impossible by a13 and a6. (b) (\Leftarrow) If $A \notin t$ or $\neg B \notin t$ or $(A \in t$ and $\neg A \in t$ and $B \in t)$, then $A \rightarrow B \in t$ follows by a7, a12 and a14, respectively. \square

Proposition 3.8 (Negated conditionals in Eb_1^3 -logics). *Let L be an ELi -logic where Li will refer in each case to one of the extensions of b_1^3 displayed in Definition 3.4, and let t be a prime, regular and non-trivial L -theory. We have: $\neg(A \rightarrow B) \in t$ iff*

- *EL1-logics:* $(A \in t \ \& \ B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$.
- *EL3-logics:* $(A \in t \ \& \ B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.
- *EL4-logics:* $(A \in t \ \& \ B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.
- *EL5-logics:* $A \in t \ \& \ \neg B \in t$.
- *EL7:* $(A \in t \ \& \ \neg B \in t)$ or $(B \in t \ \& \ \neg B \in t)$.
- *EL8:* $(A \in t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.
- *EL9:* $(A \in t \ \& \ B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL10:* $(A \in t \ \& \ B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL11:* $(A \in t \ \& \ B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.

- *EL12*: $(A \in t \ \& \ B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL13*: $(A \in t \ \& \ \neg A \in t)$ or $(A \in t \ \& \ \neg B \in t)$.
- *EL14*: $(A \notin t \ \& \ \neg B \notin t)$ or $(A \in t \ \& \ \neg A \in t)$ or $(A \in t \ \& \ \neg B \in t)$.
- *EL15*: $(A \in t \ \& \ \neg A \in t)$ or $(A \in t \ \& \ \neg B \in t)$ or $(B \in t \ \& \ \neg B \in t)$.
- *EL16*: $(A \in t \ \& \ \neg A \in t)$ or $(A \in t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(B \in t \ \& \ \neg B \in t)$.
- *EL17*: $\neg A \notin t$ or $(A \in t \ \& \ B \notin t)$.
- *EL19*: $\neg A \notin t$ or $(A \in t \ \& \ B \notin t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.
- *EL20*: $\neg A \notin t$ or $(A \in t \ \& \ B \notin t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.
- *EL21*: $\neg A \notin t$ or $(A \in t \ \& \ \neg B \in t)$.
- *EL23*: $\neg A \notin t$ or $(B \in t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg B \in t)$.
- *EL24*: $\neg A \notin t$ or $(B \in t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$.
- *EL25*: $\neg A \notin t$ or $(A \notin t \ \& \ B \notin t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL26*: $\neg B \notin t$ or $\neg A \notin t$ or $(A \in t \ \& \ B \notin t)$.
- *EL27*: $\neg A \notin t$ or $(A \in t \ \& \ B \notin t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL28*: $\neg B \notin t$ or $\neg A \notin t$ or $(A \in t \ \& \ B \notin t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.
- *EL29*: $\neg A \notin t$ or $(A \in t \ \& \ \neg A \in t)$.
- *EL30*: $\neg A \notin t$ or $(A \in t \ \& \ \neg A \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$.
- *EL31*: $\neg A \notin t$ or $(A \in t \ \& \ \neg A \in t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.
- *EL32*: $\neg B \notin t$ or $(A \in t \ \& \ B \notin t)$ or $(B \in t \ \& \ \neg B \in t)$.
- *EL33*: $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$.
- *EL35*: $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.

- *EL36*: $A \notin t$ or $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$.
- *EL37*: $B \notin t$ or $(A \in t \ \& \ \neg B \in t)$.
- *EL38*: $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL39*: $B \notin t$ or $(B \in t \ \& \ \neg B \in t)$.
- *EL40*: $A \notin t$ or $B \notin t$ or $(B \in t \ \& \ \neg B \in t)$.
- *EL41*: $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL42*: $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL43*: $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ B \notin t)$.
- *EL44*: $A \notin t$ or $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL45*: $B \notin t$ or $(A \in t \ \& \ \neg A \in t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$.
- *EL46*: $B \notin t$ or $(A \in t \ \& \ \neg A \in t)$ or $(A \in t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$.
- *EL47*: $B \notin t$ or $(B \in t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t)$.
- *EL48*: $A \notin t$ or $(A \in t \ \& \ \neg A \in t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$.
- *EL49*: $\neg A \notin t$ or $B \notin t$.
- *EL51*: $\neg A \notin t$ or $B \notin t$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$.
- *EL52*: $A \notin t$ or $\neg A \notin t$ or $B \notin t$.
- *EL53*: $\neg A \notin t$ or $B \notin t$ or $(A \in t \ \& \ \neg B \in t)$.
- *EL55*: $B \notin t$ or $(B \in t \ \& \ \neg B \in t)$ or $(\neg A \notin t \ \& \ \neg B \notin t)$.
- *EL56*: $A \notin t$ or $B \notin t$ or $(B \in t \ \& \ \neg B \in t)$ or $(\neg A \notin t \ \& \ \neg B \notin t)$.
- *EL57*: $B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$.
- *EL58*: $\neg A \notin t$ or $B \notin t$ or $\neg B \notin t$.

- *EL59*: $\neg A \notin t$ or $B \notin t$ or $(A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t)$ or $(A \notin t \ \& \ B \in t \ \& \ \neg B \in t)$
- *EL60*: $A \notin t$ or $B \notin t$ or $\neg B \notin t$ or $(\neg A \notin t \ \& \ \neg B \in t)$.
- *EL61*: $\neg A \notin t$ or $B \notin t$ or $(A \in t \ \& \ \neg A \in t)$.
- *EL62*: $\neg A \notin t$ or $B \notin t$ or $\neg B \notin t$ or $(A \in t \ \& \ \neg A \in t)$.
- *EL63*: $\neg A \notin t$ or $B \notin t$ or $(A \in t \ \& \ \neg A \in t)$ or $(B \in t \ \& \ \neg B \in t)$.
- *EL64*: $A \notin t$ or $\neg A \notin t$ or $(A \in t \ \& \ \neg A \in t)$.

Proof. It is similar to the proof of Proposition 3.7 (cf. also the proof of Proposition 6.3 in [29]). So, it will suffice to prove one case, say that of EL12-logics

(I) (\Rightarrow) Suppose (1) $\neg(A \rightarrow B) \in t$ and, for reductio, (2) $(A \notin t$ or $B \in t) \ \& \ (\neg A \in t$ or $\neg B \notin t) \ \& \ (A \in t$ or $\neg B \in t) \ \& \ (A \in t$ or $B \notin t$ or $\neg B \notin t) \ \& \ (A \notin t$ or $\neg A \notin t$ or $\neg B \in t)$. There are 72 subcases to consider but each one of them either contravenes the completeness of t or/and contains a contradiction except the following three ones: (a) $A \notin t \ \& \ \neg A \in t \ \& \ B \notin t \ \& \ \neg B \in t$; (b) $A \in t \ \& \ \neg A \in t \ \& \ B \in t \ \& \ \neg B \in t$; (c) $A \in t \ \& \ B \in t \ \& \ \neg A \notin t \ \& \ \neg B \notin t$. But these three subcases are also impossible. Subcase (a): given (1) and A16, $A \in t$ or $B \in t$ follows contradicting (a); subcase (b): by A29, $C \in t$ (for any wff C) follows, contradicting the non-triviality of t ; finally, the impossibility of subcase (c) is proved similarly to that of subcase (a) by using now A21.

(II) (\Leftarrow) Suppose (1) $A \in t \ \& \ B \notin t$ or (2) $\neg A \notin t \ \& \ \neg B \in t$ or (3) $A \notin t \ \& \ \neg B \notin t$ or (4) $A \notin t \ \& \ B \in t \ \& \ \neg B \in t$ or (5) $A \in t \ \& \ \neg A \in t \ \& \ \neg B \notin t$. Then, $\neg(A \rightarrow B) \in t$ follows by A7, A8, A5, A13 and A12, respectively. For example, consider (4): by 4 and A13 we have $A \vee \neg(A \rightarrow B)$, whence by $A \notin t$ and primeness of t $\neg(A \rightarrow B) \in t$ follows. \square

Proposition 3.9 (Negated conditionals in Eb_2^3 -logics). *Let L be an ELi -logic where Li will refer in each case to one of the extensions of b_2^3 displayed in Definition 3.5, and let t be a prime, regular and non-trivial L -theory. We have: $\neg(A \rightarrow B) \in t$ iff*

- *EL2*: $(A \in t \ \& \ B \notin t)$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$.
- *EL6*: $(A \notin t \ \& \ \neg B \notin t)$ or $(A \in t \ \& \ \neg B \in t)$.
- *EL18*: $\neg A \notin t$ or $(A \in t \ \& \ B \notin t)$ or $(A \notin t \ \& \ \neg B \notin t)$.
- *EL22*: $\neg A \notin t$ or $(A \in t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$.

- *EL34*: $B \notin t$ or $(A \notin t \ \& \ \neg B \notin t)$ or $(\neg A \notin t \ \& \ \neg B \in t)$.
- *EL50*: $B \notin t$ or $\neg A \notin t$ or $(A \notin t \ \& \ \neg B \notin t)$.
- *EL54*: $\neg A \notin t$ or $B \notin t$ or $(A \in t \ \& \ \neg B \in t)$ or $(A \notin t \ \& \ \neg B \notin t)$.

Proof. It is similar to that of Proposition 3.8 and it is left to the reader. \square

The section is ended with the proof of the primeness lemma for Eb_1^3 -logics (the case of Eb_2^3 -logics presents some problems —cf. §5).

Lemma 3.10 (Primeness). *Let L be one of the Eb_1^3 -logics in Definition 3.4, t an L -theory and A a wff such that $A \notin t$. Then, there is a prime L -theory \mathcal{T} such that $t \subseteq \mathcal{T}$ and $A \notin \mathcal{T}$.*

Proof. By using Kuratowski-Zorn's Lemma, for example, t is extended to a maximal theory \mathcal{T} such that $A \notin \mathcal{T}$. Then, it is easy to show that \mathcal{T} is prime (cf., for instance, the proof of Lemma 5.13 in [29] and notice that closure under MP is provably guaranteed by the modus ponens axiom a6). \square

4 Belnap-Dunn semantics for the Li -logics

Let T represent truth and F represent falsity. Belnap-Dunn semantics (BD-semantics) is characterized by the possibility of assigning T , F , both T and F or neither T nor F to the formulas of a given logical language (cf. [4], [5], [7], [8]). Concerning 3-valued logics, two variants of BD-semantics: overdetermined BD-semantics (o-semantics) and underdetermined BD-semantics (u-semantics) can be considered. Formulas can be assigned T , F or both T and F in the former; T , F or neither T nor F in the latter (cf. [29], [33]). U-semantics is especially adequate to 3-valued logics determined by matrices with only one designated value; o-semantics, for those determined by matrices where only one value is not designated.

Given an implicative expansion of MK3, M , with 2 as the only designated value, the idea for defining an equivalent u-semantics, M_u , to the matrix semantics based upon M is simple: a wff A is assigned neither T nor F in M_u iff it is assigned 1 in M . Then A is assigned T (resp., F) in M_u iff it is assigned 2 (resp., 0) in M . On the other hand, if M has both 1 and 2 as designated values, then an equivalent o-semantics, M_o , to the matrix semantics based upon M is defined as follows. A is assigned both T and F in M_o iff A is assigned 1 in M . Next, A is assigned T (resp., F) in M_o iff it is not assigned 0 (resp., 2) in M . (Notice that in u-semantics formulas

can be assigned neither T nor F but interpretation of formulas cannot be empty in o-semantics.)⁶

The o-semantics equivalent to the matrix semantics based upon each one of the 64 matrices introduced below, as well as the u-semantics equivalent to the matrix semantics based upon each one of the 8 matrices in §5, have been built up by translating the matrix semantics based upon the matrices in question into an o-semantics (or a u-semantics, as the case may be), according to the simple intuitive ideas just exposed. (Recall that in the present section only implicative expansions of MK3 with two designated values are considered; those with 2 as the only designated value are treated in §5.)

In the sequel, the notion of an Li -model and the accompanying notions of Li -consequence and Li -validity are defined. For each i ($1 \leq i \leq 64$), the Li -model M is an o-semantics equivalent to the matrix semantics based upon the matrix M_i in the sense explained above.

Definition 4.1 (*Li-models*). *For all i ($1 \leq i \leq 64$), an Li -model is the structure (K, I) where (i) $K = \{\{T\}, \{F\}, \{T, F\}\}$, and (ii) I is an Li -interpretation from the set of all wffs to K , according to the following conditions (clauses) for each propositional variable p and wffs A, B : (1) $I(p) \in K$; (2a) $T \in I(\neg A)$ iff $F \in I(A)$; (2b) $F \in I(\neg A)$ iff $T \in I(A)$; (3a) $T \in I(A \wedge B)$ iff $T \in I(A) \ \& \ T \in I(B)$; (3b) $F \in I(A \wedge B)$ iff $F \in I(A)$ or $F \in I(B)$; (4a) $T \in I(A \vee B)$ iff $T \in I(A)$ or $T \in I(B)$; (4b) $F \in I(A \vee B)$ iff $F \in I(A) \ \& \ F \in I(B)$; (5a) $T \in I(A \rightarrow B)$ iff $T \notin I(A)$ or $F \notin I(B)$ or $(T \in I(A) \ \& \ F \in I(A) \ \& \ T \in I(B))$ ⁷. The clause for assigning $\{F\}$ to conditionals is different for each Li -model. Thus, we have the following 64 conditions: $F \in I(A \rightarrow B)$ iff*

- (5b1) $[T \in I(A) \ \& \ T \notin I(B)]$ or $[F \notin I(A) \ \& \ F \in I(B)]$.
- (5b2) $[T \in I(A) \ \& \ T \notin I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$ or $[F \notin I(A) \ \& \ F \in I(B)]$.
- (5b3) $[T \in I(A) \ \& \ T \notin I(B)]$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.
- (5b4) $[T \in I(A) \ \& \ T \notin I(B)]$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.

⁶This topic is pursued in section 6.

⁷Condition (5a) is equivalent to $[(T \in I(A) \text{ or } T \in I(B)) \ \& \ (F \notin I(A) \text{ or } F \in I(B))]$, as well as to $I(A) \leq I(B)$, where \leq is the ordering of truth-values in Remark 2.2 (I owe this remark to a referee of the JAL).

- (5b5) $T \in I(A) \ \& \ F \in I(B)$.
- (5b6) $[T \notin I(A) \ \& \ F \notin I(B)] \text{ or } [T \in I(A) \ \& \ F \in I(B)]$.
- (5b7) $[T \in I(A) \ \& \ F \in I(B)] \text{ or } [T \in I(B) \ \& \ F \in I(B)]$.
- (5b8) $[T \in I(A) \ \& \ F \in I(B)] \text{ or } [T \notin I(A) \ \& \ F \notin I(B)] \text{ or } [T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.
- (5b9) $[T \in I(A) \ \& \ T \notin I(B)] \text{ or } [F \notin I(A) \ \& \ F \in I(B)] \text{ or } [T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b10) $[T \in I(A) \ \& \ T \notin I(B)] \text{ or } [F \notin I(A) \ \& \ F \in I(B)] \text{ or } [T \notin I(A) \ \& \ F \notin I(B)] \text{ or } [T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b11) $[T \in I(A) \ \& \ T \notin I(B)] \text{ or } [F \notin I(A) \ \& \ F \in I(B)] \text{ or } [T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)] \text{ or } [T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b12) $[T \in I(A) \ \& \ T \notin I(B)] \text{ or } [F \notin I(A) \ \& \ F \in I(B)] \text{ or } [T \notin I(A) \ \& \ F \notin I(B)] \text{ or } [T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)] \text{ or } [T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b13) $[T \in I(A) \ \& \ F \in I(A)] \text{ or } [T \in I(A) \ \& \ F \in I(B)]$.
- (5b14) $[T \notin I(A) \ \& \ F \notin I(B)] \text{ or } [T \in I(A) \ \& \ F \in I(A)] \text{ or } [T \in I(A) \ \& \ F \in I(B)]$.
- (5b15) $[T \in I(A) \ \& \ F \in I(A)] \text{ or } [T \in I(A) \ \& \ F \in I(B)] \text{ or } [T \in I(B) \ \& \ F \in I(B)]$.
- (5b16) $[T \in I(A) \ \& \ F \in I(A)] \text{ or } [T \in I(A) \ \& \ F \in I(B)] \text{ or } [T \notin I(A) \ \& \ F \notin I(B)] \text{ or } [T \in I(B) \ \& \ F \in I(B)]$.
- (5b17) $F \notin I(A) \text{ or } [T \in I(A) \ \& \ T \notin I(B)]$.
- (5b18) $F \notin I(A) \text{ or } [T \in I(A) \ \& \ T \notin I(B)] \text{ or } [T \notin I(A) \ \& \ F \notin I(B)]$.
- (5b19) $F \notin I(A) \text{ or } [T \in I(A) \ \& \ T \notin I(B)] \text{ or } [T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.
- (5b20) $F \notin I(A) \text{ or } [T \in I(A) \ \& \ T \notin I(B)] \text{ or } [T \notin I(A) \ \& \ F \notin I(B)] \text{ or } [T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.
- (5b21) $F \notin I(A) \text{ or } [T \in I(A) \ \& \ F \in I(B)]$.

- (5b22) $F \notin I(A)$ or $[T \in I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$.
- (5b23) $F \notin I(A)$ or $[T \in I(B) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ F \in I(B)]$.
- (5b24) $F \notin I(A)$ or $[T \in I(B) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$.
- (5b25) $F \notin I(A)$ or $[T \notin I(A) \ \& \ T \notin I(B)]$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b26) $F \notin I(B)$ or $F \notin I(A)$ or $[T \in I(A) \ \& \ T \notin I(B)]$.
- (5b27) $F \notin I(A)$ or $[T \in I(A) \ \& \ T \notin I(B)]$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b28) $F \notin I(B)$ or $F \notin I(A)$ or $[T \in I(A) \ \& \ T \notin I(B)]$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.
- (5b29) $F \notin I(A)$ or $[T \in I(A) \ \& \ F \in I(A)]$.
- (5b30) $F \notin I(A)$ or $[T \in I(A) \ \& \ F \in I(A)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$.
- (5b31) $F \notin I(A)$ or $[T \in I(A) \ \& \ F \in I(A)]$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.
- (5b32) $F \notin I(B)$ or $[T \in I(A) \ \& \ T \notin I(B)]$ or $[T \in I(B) \ \& \ F \in I(B)]$.
- (5b33) $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$.
- (5b34) $T \notin I(B)$ or $[T \notin I(A) \ \& \ F \notin I(B)]$ or $[F \notin I(A) \ \& \ F \in I(B)]$.
- (5b35) $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.
- (5b36) $T \notin I(A)$ or $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$.
- (5b37) $T \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(B)]$.
- (5b38) $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b39) $T \notin I(B)$ or $[T \in I(B) \ \& \ F \in I(B)]$.
- (5b40) $T \notin I(A)$ or $T \notin I(B)$ or $[T \in I(B) \ \& \ F \in I(B)]$.

- (5b41) $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b42) $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b43) $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b44) $T \notin I(A)$ or $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b45) $T \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(A)]$ or $[F \notin I(A) \ \& \ F \in I(B)]$.
- (5b46) $T \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(A)]$ or $[T \in I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$.
- (5b47) $T \notin I(B)$ or $[T \in I(B) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ F \in I(A)]$.
- (5b48) $T \notin I(A)$ or $[T \in I(A) \ \& \ F \in I(A)]$ or $[F \notin I(A) \ \& \ F \in I(B)]$.
- (5b49) $F \notin I(A)$ or $T \notin I(B)$.
- (5b50) $T \notin I(B)$ or $F \notin I(A)$ or $[T \notin I(A) \ \& \ F \notin I(B)]$.
- (5b51) $F \notin I(A)$ or $T \notin I(B)$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.
- (5b52) $T \notin I(A)$ or $F \notin I(A)$ or $T \notin I(B)$.
- (5b53) $F \notin I(A)$ or $T \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(B)]$.
- (5b54) $F \notin I(A)$ or $T \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(B)]$.
- (5b55) $T \notin I(B)$ or $[T \in I(B) \ \& \ F \in I(B)]$ or $[F \notin I(A) \ \& \ F \notin I(B)]$.
- (5b56) $T \notin I(A)$ or $T \notin I(B)$ or $[T \in I(B) \ \& \ F \in I(B)]$ or $[F \notin I(A) \ \& \ F \notin I(B)]$.
- (5b57) $T \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$.
- (5b58) $F \notin I(A)$ or $T \notin I(B)$ or $F \notin I(B)$.
- (5b59) $F \notin I(A)$ or $T \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(A) \ \& \ F \notin I(B)]$ or $[T \notin I(A) \ \& \ T \in I(B) \ \& \ F \in I(B)]$.

- (5b60) $T \notin I(A)$ or $T \notin I(B)$ or $F \notin I(B)$ or $[F \notin I(A) \ \& \ F \in I(B)]$.
- (5b61) $F \notin I(A)$ or $T \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(A)]$.
- (5b62) $F \notin I(A)$ or $T \notin I(B)$ or $F \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(A)]$.
- (5b63) $F \notin I(A)$ or $T \notin I(B)$ or $[T \in I(A) \ \& \ F \in I(A)]$ or $[T \in I(B) \ \& \ F \in I(B)]$.
- (5b64) $T \notin I(A)$ or $F \notin I(A)$ or $[T \in I(A) \ \& \ F \in I(A)]$.

Definition 4.2 (Li-consequence, Li-validity). *Let M be an Li-model ($1 \leq i \leq 64$). For any set of wffs Γ and wff A , $\Gamma \vDash_M A$ (A is a consequence of Γ in the Li-model M) iff $T \in I(A)$ whenever $T \in I(\Gamma)$ [$T \in I(\Gamma)$ iff $\forall A \in \Gamma (T \in I(A))$]; $F \in I(\Gamma)$ iff $\exists A \in \Gamma (F \in I(A))$]. Then, $\Gamma \vDash_{Li} A$ (A is a consequence of Γ in Li-semantics) iff $\Gamma \vDash_M A$ for each Li-model M . In particular, $\vDash_{Li} A$ (A is valid in Li-semantics) iff $\vDash_M A$ for each Li-model M (i.e., iff $T \in I(A)$ for each Li-model M). (By \vDash_{Li} we shall refer to the relation just defined.)*

Now, given the 64 matrices defined in §2 and Definitions 4.1 and 4.2, we easily have:

Proposition 4.3 (Coextensiveness of \vDash_{M_i} and \vDash_{Li}). *For any i ($1 \leq i \leq 64$), set of wffs Γ and wff A , $\Gamma \vDash_{M_i} A$ iff $\Gamma \vDash_{Li} A$. In particular, $\vDash_{M_i} A$ iff $\vDash_{Li} A$.*

Proof. Cf., e.g., the proof of Proposition 7.4 in [29]. □

The proof of Proposition 4.3 is a mere formalization of the intuitive translation (commented upon above) of a given matrix semantics into its corresponding \circ -semantics. But it greatly simplifies the soundness and completeness proofs since we can focus on the relation \vDash_{M_i} in the former case, while restricting our attention to the relation \vDash_{Li} in the latter one. Thus, we have:

Theorem 4.4 (Soundness of the Li-logics). *For any i ($1 \leq i \leq 64$), set of wffs Γ and wff A , if $\Gamma \vdash_{Li} A$ then (1) $\Gamma \vDash_{M_i} A$ and (2) $\Gamma \vDash_{Li} A$.*

Proof. (1) Given a particular logic Li , it is easy to check that the rules Adj and MP (and dTrans, when present) preserve M_i -validity, whereas the axioms of Li are assigned either 1 or 2 by any M_i -interpretation I . Consequently, if $\Gamma \vdash_{Li} A$, then $\Gamma \vDash_{M_i} A$. Then (2) is immediate by (1) and Proposition 4.3. Finally, if Γ is the empty set, the proof is similar (in case a tester is needed the reader can use that in [12]). □

Turning to the completeness theorem, completeness of Li w.r.t. \models_{Li} is proved by means of a canonical model construction. Then, completeness w.r.t. \models_{Mi} is immediate by coextensiveness of the two consequence relations (Proposition 4.3). As it has been remarked above, we note that in the rest of this section by Li we refer to the extensions of b_1^3 in Definition 3.4 (the extensions of b_2^3 —cf. Definition 3.5— are treated in the following section).

A canonical Li -model is a structure $(K, I_{\mathcal{T}})$ where K is defined as in Definition 4.1, $I_{\mathcal{T}}$ is a \mathcal{T} -interpretation built upon a prime, regular and non-trivial Li -theory \mathcal{T} (cf. the preceding section on the notion of an Li -theory and the classes of Li -theories of interest in the present paper). A \mathcal{T} -interpretation is a function such that for each propositional variable p , we have $T \in I_{\mathcal{T}}(p)$ iff $p \in \mathcal{T}$, and $F \in I_{\mathcal{T}}(p)$ iff $\neg p \in \mathcal{T}$, while complex wffs are assigned a member of K according to conditions 2, 3, 4 and 5 in Definition 4.1.

It is clear that any canonical Li -model is an Li -model. Therefore, completeness actually depends on the possibility of extending the canonical interpretation of propositional variables to all wffs. That is, given the facts proven so far, completeness depends on the following proposition.

Proposition 4.5 (\mathcal{T} -interpreting the set of all wffs). *Let L be an Li -logic and I be a \mathcal{T} -interpretation defined on the L -theory \mathcal{T} . For each wff A , we have: (1) $T \in I(A)$ iff $A \in \mathcal{T}$; (2) $F \in I(A)$ iff $\neg A \in \mathcal{T}$.*

Proof. By induction on the length of A . It is easy by using Propositions 3.6, 3.7, 3.8 and 3.9 (cf., e.g., the proof of Proposition 8.5 in [29]). □

Once Proposition 4.5 is at our disposal, completeness is proved as follows. Let L be an Li -logic. Suppose that Γ is a set of wffs and A is a wff such that $\Gamma \not\vdash_L A$. Then A is not included in the set of consequences derivable in L from Γ (in symbols, $A \notin Cn\Gamma[L]$). Now, given that for any set of wffs Γ , $Cn\Gamma[L]$ is clearly a regular L -theory, by using the primeness lemma (Lemma 3.10), it can be extended to a prime, regular and non-trivial L -theory \mathcal{T} such that $A \notin \mathcal{T}$. Then \mathcal{T} generates a canonical model M with a \mathcal{T} -interpretation $I_{\mathcal{T}}$ such that $T \in I_{\mathcal{T}}(\Gamma)$ (since $T \in I_{\mathcal{T}}(Cn\Gamma[L])$) but $T \notin I_{\mathcal{T}}(A)$, whence $\Gamma \not\vdash_M A$ and finally, $\Gamma \not\vdash_{Li} A$.

Based upon the argumentation just developed, we state the ensuing theorem.

Theorem 4.6 (Completeness of the extensions of b_1^3). *Let Li be any of the extensions of b_1^3 in Definition 3.4. For any set of wffs Γ and wff A , (1) if $\Gamma \models_{Li} A$, then $\Gamma \vdash_{Li} A$; (2) if $\Gamma \models_{Mi} A$, then $\Gamma \vdash_{Li} A$.*

5 Implicative expansions of MK3 with only one designated value. Extensions of \mathbf{b}_2^3

In this section, we investigate all implicative expansions of MK3 (with only 2 as designated value) verifying FDE. Also, we prove completeness of the expansions of \mathbf{b}_2^3 (in Definition 3.5) w.r.t. the BD-semantics defined in section 4. Given that the proof of completeness poses a similar problem in both groups of logics, we begin by showing that L2, L6, L18, L22, L34, L50 and L54 (cf. Definition 3.5) are complete w.r.t. their corresponding BD-semantics (cf. section 4).

As pointed out in the introduction to the paper, we prove completeness by following the method developed in [34] (Chapter 4), as applied in [6] and displayed in particular in [29] and [33]. The key notion in the method is that of ‘canonical interpretation’. As shown in the preceding section, canonical interpretations are functions built upon prime, regular and non-trivial L-theories. But we face a problem for applying the method in the case of some logics. Suppose that L is a logic closed under a rule r but lacking the corresponding axiom. For instance, suppose that L is closed under the rule *Modus Tollens* (MT) $A \rightarrow B, \neg B \Rightarrow \neg A$, but lacks the axiom *Modus Tollens* $[(A \rightarrow B) \wedge \neg B] \rightarrow \neg A$. Then, following the aforementioned method, it is not possible to build prime L-theories closed under MT, in general. Nevertheless, in the items quoted above, it has been shown that, despite the absence of the axiom corresponding to a rule r , prime L-theories are available if in addition to being closed under r , L is also closed under the disjunctive version of r . For instance, if in addition to being closed under MT, L is also closed under disjunctive MT (i.e., $C \vee (A \rightarrow B), C \vee \neg B \Rightarrow C \vee \neg A$).

In this sense, it has to be noted that the Transitivity axiom $[(A \rightarrow B) \wedge (B \rightarrow C)] \rightarrow (A \rightarrow C)$ fails in implicative expansions of MK3 built upon any of the \rightarrow -functions described by the following tables: t2, t6, t18, t22, t34, t50 and t54 (in all cases, it suffices to take an assignment v on the set $V = \{0, 1, 2\}$ such that, for distinct propositional variables p, q and r , $v(p) = 0$, $v(q) = 1$ and $v(r) = 2$. Then, $v([(p \rightarrow q) \wedge (q \rightarrow r)] \rightarrow (p \rightarrow r)) = 0$). Fortunately, the rule Disjunctive Transitivity (dTrans) $D \vee [(A \rightarrow B) \wedge (B \rightarrow C)] \Rightarrow D \vee (A \rightarrow C)$ preserves validity in each one of the implicative expansions of MK3 just referred to. Consequently, the required prime L-theories can be built as shown in [34] (Chapter 4; cf. also [6], [29], [33]) and then completeness of extensions of \mathbf{b}_2^3 is proved similarly as that of extensions of \mathbf{b}_1^3 in section 4.

Let us now turn to implicative expansions of MK3 with only 2 as designated value⁸.

⁸Cf. notes 2, 3.

Reasoning similarly as in the case of 1 and 2 as designated values (cf. section 2), we are left with the following possibilities (a_i ($1 \leq i \leq 3$) $\in \{0, 1\}$).

	\rightarrow	0	1	2
tc.	0	2	2	2
	1	a_1	2	2
	2	a_2	a_3	2

That is,

t65	0	1	2	2	t66	0	1	2	2	t67	0	1	2	2	t68	0	1	2	2
	0	2	2	2		0	2	2	2		0	2	2	2		0	2	2	2
	1	1	2	2		1	1	2	2		1	0	2	2		1	0	2	2
	2	1	0	2		2	1	1	2		2	1	0	2		2	1	1	2

t69	0	1	2	2	t70	0	1	2	2	t71	0	1	2	2	t72	0	1	2	2
	0	2	2	2		0	2	2	2		0	2	2	2		0	2	2	2
	1	0	2	2		1	1	2	2		1	0	2	2		1	1	2	2
	2	0	0	2		2	0	1	2		2	0	1	2		2	0	0	2

As in section 3, let Li refer to the logic determined by the matrix Mi ($65 \leq i \leq 72$). Now, L69, L70, L71 and L72 are investigated in [33]. (We remark that L69, L70 and L71 are $S5^3_E$ —the 3-valued extension of the positive fragment of Lewis’ S5 as axiomatized by Hacking [13]—, Łukasiewicz’s 3-valued logic $\mathbb{L}3$, and the paraconsistent logic $G3^{\leq}_E$, respectively; cf. [33] and references therein.) Then, L65, L66, L67 and L68 can be treated similarly. We begin by defining the required underdetermined BD-semantics.

Definition 5.1 (BD-models for L65, L66, L67 and L68). *For all i ($65 \leq i \leq 68$), an Li -model is the structure (K, I) where (i) $K = \{\{T\}, \{F\}, \emptyset\}$ and (ii) I is an Li -interpretation from the set of all wffs to K , according to the following conditions: (1), (2a), (2b), (3a), (3b), (4a) and (4b) are defined as in Definition 4.1. (5a) $T \in I(A \rightarrow B)$ iff $F \in I(A)$ or $T \in I(B)$ or $[T \notin I(A) \ \& \ F \notin I(B)]$. The clause for assigning $\{F\}$ to conditionals is different for each Li -model. Thus, we have the following 4 conditions: $F \in I(A \rightarrow B)$ iff*

- (5b65) $T \in I(A) \ \& \ T \notin I(B) \ \& \ F \notin I(B)$.
- (5b66) $T \in I(A) \ \& \ T \notin I(A)$.
- (5b67) $[T \notin I(A) \ \& \ F \notin I(A) \ \& \ F \in I(B)]$ or $[T \in I(A) \ \& \ T \notin I(B) \ \& \ F \notin I(B)]$.

- (5b68) $T \notin I(A) \ \& \ F \notin I(A) \ \& \ F \in I(B)$.

The notions of *Li*-consequence and *Li*-validity are defined similarly as in Definition 4.2 (by the way, notice that conditionals are never assigned $\{F\}$ according to L2-models).

On the other hand, the *Li*-logics can be axiomatized as follows. The basic logic b^3 is formulated with a1-a5, a8-a10 of b^3 (cf. Definition 3.1) and the following axioms:

$$\text{A30. } \neg A \rightarrow [A \vee (A \rightarrow B)]$$

$$\text{A31. } B \rightarrow [\neg B \vee (A \rightarrow B)]$$

$$\text{A32. } (A \vee \neg B) \vee (A \rightarrow B)$$

The rules are Adj, MP, Disjunctive Modus Ponens (dMP), Disjunctive Transitivity (dTrans), Disjunctive Modus Tollens (dMT) and Disjunctive 'E contradictione quodlibet' (dECQ)

$$\text{(dMP): } C \vee (A \rightarrow B), C \vee A \Rightarrow C \vee B$$

$$\text{(dTrans): } D \vee [(A \rightarrow B) \wedge (B \rightarrow C)] \Rightarrow D \vee (A \rightarrow C)$$

$$\text{(dMT): } C \vee (A \rightarrow B), C \vee \neg B \Rightarrow C \vee \neg A$$

$$\text{(dECQ): } C \vee (A \wedge \neg A) \rightarrow C \vee B$$

The particular axioms of each *Li*-logic are chosen from the following list⁹:

$$\text{A14. } [\neg(A \rightarrow B) \wedge \neg A] \rightarrow A$$

$$\text{A19. } [\neg(A \rightarrow B) \wedge B] \rightarrow \neg B$$

$$\text{A33. } \neg(A \rightarrow B) \rightarrow A$$

$$\text{A34. } A \rightarrow [(B \vee \neg B) \vee \neg(A \rightarrow B)]$$

$$\text{A35. } [\neg(A \rightarrow B) \wedge \neg B] \rightarrow [(\neg A \vee B) \vee (A \rightarrow B)]$$

$$\text{A36. } \neg(A \rightarrow B) \rightarrow (A \rightarrow B)$$

$$\text{A37. } \neg B \rightarrow [(A \vee \neg A) \vee \neg(A \rightarrow B)]$$

$$\text{A38. } [\neg(A \rightarrow B) \wedge (A \wedge \neg B)] \rightarrow [(\neg A \vee B) \vee (A \rightarrow B)]$$

$$\text{A39. } \neg(A \rightarrow B) \rightarrow \neg B$$

$$\text{A40. } [\neg(A \rightarrow B) \wedge A] \rightarrow [(\neg A \vee B) \vee (A \rightarrow B)]$$

We have:

⁹Cf. note 5.

L1: A19, A33, A34, A35.

L2: A36.

L3: A14, A19, A34, A37, A38.

L4: A14, A37, A39, A40.

Completeness for these four logics can be proved similarly as for the extensions of b_2^3 , since each one of the former logics lacks either the Modus Ponens axiom $[(A \rightarrow B) \wedge A] \rightarrow B$ or the Transitivity axiom $[(A \rightarrow B) \wedge (B \rightarrow C)] \rightarrow (A \rightarrow C)$ or both, but has the corresponding disjunctive rules. The main difference with the proof of completeness for the extensions of b_2^3 lies in the fact that theories are now not complete in general, as the eight expansions of MK3 (2 is the only designated value) lack the principle of excluded middle axiom, $A \vee \neg A$, but are in turn consistent due to the fact that the rule ECQ, $A \wedge \neg A \Rightarrow B$, holds in each one of the eight logics. In this sense, given L_i -models (Definition 5.1) and the formulation of L65-68 provided above, soundness and completeness for these last four logics can proceed similarly as the corresponding proofs for L5-L8 presented in [33].

6 On the notion of implication tables tb and tc embody. On Brady's method

In algebraic and many-valued logic, the conditional (or 'implication', as it is named in [41]; cf. pp. 227, ff.) is traditionally required to meet the following strong condition: $c0$: $a \rightarrow b = t$ iff $a \leq b$, where t , the greatest element in the set of logical values, is the only designated value. In this context, Tomova's notion of 'natural conditional' (cf. [39]) can be viewed as an attempt to liberalize the strong condition $c0$. Given a matrix semantics, a conditional is *natural* if the following three conditions are fulfilled: (1) it is C-extending (i.e., it coincides with the classical conditional when restricted to the classical values T and F); (2) it satisfies the *Modus Ponens*; (3) it is assigned a designated value whenever the value assigned to its antecedent is less than or equal to the value assigned to its consequent. In [39], Tomova studies the lattice of the 30 natural implicative expansions of Kleene's strong 3-valued matrix, MK3, with both only one and two designated values (MK3I and MK3II, respectively). It develops that the 'variable-sharing property' (VSP) is predicable of none of the said 30 natural implicative expansions of MK3 (cf. [28]). The VSP being a necessary property of any relevance logic, as it is known, in order to have 3-valued logics with the VSP, in [28], [32], it is proposed to further liberalize the requisite $c0$ by modifying Tomova's notion of a natural condition in the ensuing sense: conditions (c1) and (c2) are maintained but (c3) is replaced by (c3'): a conditional is assigned a designated value whenever its antecedent and its consequence are assigned the same

value. There are exactly 108 natural (in the new sense of the term) implicative expansions of MK3, 11 of which determine logics with the VSP (cf. [28], [32]). Well then, the 72 implicative expansions of MK3I and MK3II studied in the present paper are generally, and precisely, characterized by conditions (c2) and (c3). (Notice, by the way, that tables tb and tc could not represent any intuitively acceptable notions of conjunction or disjunction: $f_{\rightarrow}(0, 0)$ and $f_{\rightarrow}(0, 2)$ give a designated value, whereas $f_{\rightarrow}(2, 0)$ is not designated.)

More alternative notions of a natural conditional (or 'implication' in Wojcicki's term) can be contemplated in 3-valued logic. Consider the following general tables (b_i ($1 \leq i \leq 3$) $\in \{0, 1, 2\}$; designated values are starred):

td.	\rightarrow 0 1 2 <hr style="border: none; border-top: 1px solid black; margin: 0;"/> 0 2 b_1 2 *1 0 0 b_2 *2 0 b_3 2	te.	\rightarrow 0 1 2 <hr style="border: none; border-top: 1px solid black; margin: 0;"/> 0 2 0 2 *1 0 b_1 b_2 *2 0 b_3 2	tf.	\rightarrow 0 1 2 <hr style="border: none; border-top: 1px solid black; margin: 0;"/> 0 2 b_1 2 *1 0 b_2 0 *2 0 b_3 2
tg.	\rightarrow 0 1 2 <hr style="border: none; border-top: 1px solid black; margin: 0;"/> 0 2 b_1 2 1 b_2 0 b_3 *2 0 0 2	th.	\rightarrow 0 1 2 <hr style="border: none; border-top: 1px solid black; margin: 0;"/> 0 2 0 2 1 b_1 b_2 b_3 *2 0 0 2	ti.	\rightarrow 0 1 2 <hr style="border: none; border-top: 1px solid black; margin: 0;"/> 0 2 b_1 2 1 b_2 b_3 0 *2 0 0 2

The 162 implicative expansions of MK3I and MK3II contained in tables td through ti above are generally, and precisely, characterized by their obeying conditions (c1) and (c2) (each of them falsifies (c3) or (c3')). Moreover, consider the following general tables (b_i ($1 \leq i \leq 8$) $\in \{0, 1, 2\}$; designated values are starred):

tj.	\rightarrow 0 1 2 <hr style="border: none; border-top: 1px solid black; margin: 0;"/> 0 b_1 b_2 b_3 *1 0 b_4 b_5 *2 0 b_6 b_7	tk.	\rightarrow 0 1 2 <hr style="border: none; border-top: 1px solid black; margin: 0;"/> 0 b_1 b_2 b_3 *1 b_4 b_5 b_6 *2 0 b_7 b_8
-----	--	-----	--

The 8748 implicative expansions of MK3I and MK3II contained in tables tj (2187) and tk (6561) are generally, and precisely, characterized by their fulfilling condition (c2) (not all of them satisfy conditions (c1), (c3) or (c3')).

The natural (in Tomova's sense) implicative expansions of MK3I and MK3II are given Hilbert-style formulations from a common basic axiomatic system (see systems b_a^3 and b_b^3 in the Appendix) in [29] and [33]. The natural (in the sense of [28]) implicative expansions of MK3II with the VSP are also given Hilbert-style formulations in [32] from a much weaker common basic axiomatic system than b_a^3 and b_b^3 (see system b_c^3 in the Appendix). Well then, the logics determined by tables td through tk described above can also be formulated in a Hilbert-style way, from

a still weaker system than b_c^3 (see system b_d^3 in the Appendix). Let us propose a couple of examples. Consider the tables (designated values are starred):

t73.	\rightarrow	0	1	2
	0	2	2	2
	*1	0	0	2
	*2	0	0	2

t74.	\rightarrow	0	1	2
	0	1	0	2
	1	2	0	1
	*2	0	0	0

t73 (resp., t74) is one of the tables included in the general table td (resp., tk). Now, b_d^3 takes care of the interpretation of \wedge, \vee and \neg , as well as of the extension and primeness lemmas. So, we only need to show that the canonical interpretation of the conditional holds, since once this is done, completeness can be proved similarly as in section 4.

Table 73: Following the instructions in §4, the interpretation of the conditional is as follows: $T \in I(A \rightarrow B)$ iff $T \notin I(A)$ or $F \notin I(B)$; $F \in I(A \rightarrow B)$ iff $T \notin I(A \rightarrow B)$. Next, given a prime, regular and non-trivial L73-theory \mathcal{T} , we have to prove (a) $A \rightarrow B \in \mathcal{T}$ iff $A \notin \mathcal{T}$ or $\neg B \notin \mathcal{T}$ and (b) $\neg(A \rightarrow B) \in \mathcal{T}$ iff $A \rightarrow B \notin \mathcal{T}$. Now, (a) follows by the ensuing theses: $[(A \rightarrow B) \wedge A] \rightarrow B$; $[(A \rightarrow B) \wedge \neg B] \rightarrow \neg A$; $A \vee (A \rightarrow B)$ and $\neg B \vee (A \rightarrow B)$; and (b) is proved with the following theses: $[(A \rightarrow B) \wedge \neg(A \rightarrow B)] \rightarrow C$ and $A \vee \neg A$ (said theses, verified by Mt73, are added as axioms to b_d^3 in order to axiomatize the logic L73).

Lt73 is an interesting strong (if highly non-standard) logic. Take, for instance, the axiomatization of Anderson and Belnap’s logic of the relevant implication R displayed in [1], p. 341. Mt73 verifies all axioms and rules of this formulation except the self-identity axiom R1, $A \rightarrow A$ (notice that Mt74 also falsifies this axioms).

Table 74: Proceeding similarly as in the case of t73, we have: $T \in I(A \rightarrow B)$ iff $[F \in I(A) \ \& \ T \in I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(A) \ \& \ F \in I(B)]$; $F \in I(A \rightarrow B)$ iff $[T \notin I(B) \ \& \ F \notin I(B)]$ or $T \in I(A)$. Next, given a prime, regular and (now) consistent L74-theory \mathcal{T} , we have to prove (a) $A \rightarrow B \in \mathcal{T}$ iff $[\neg A \in \mathcal{T} \ \& \ B \in \mathcal{T}]$ or $[A \notin \mathcal{T} \ \& \ \neg A \notin \mathcal{T} \ \& \ \neg B \in \mathcal{T}]$, and (b) $\neg(A \rightarrow B) \in \mathcal{T}$ iff $[B \notin \mathcal{T} \ \& \ \neg B \notin \mathcal{T}]$ or $A \in \mathcal{T}$. Now, (a) follows by the rules: $\neg A \wedge B \Rightarrow (A \rightarrow B)$; $\neg B \Rightarrow (A \vee \neg A) \vee (A \rightarrow B)$; $A \rightarrow B \wedge A \Rightarrow \neg A$; $A \rightarrow B \Rightarrow \neg A \vee \neg B$; $A \rightarrow B \wedge \neg A \Rightarrow B$ and $A \rightarrow B \Rightarrow B \vee \neg B$. And (b) follows by the ensuing thesis and rule: $(B \vee \neg B) \vee \neg(A \rightarrow B)$; $A \Rightarrow \neg(A \rightarrow B)$; $\neg(A \rightarrow B) \wedge \neg B \Rightarrow A$; $\neg(A \rightarrow B) \wedge B \Rightarrow A$. The rules and thesis (verified by Mt74) are added to b_d^3 as an axiom and primitive rules of inference in order to axiomatize the logic L74.

The method can also be used to treat non-standard unary and binary connectives. Consider the following truth-table (2 is the only designated value):

t75.	◦	0	1	2
	0	1	0	2
	1	0	0	2
	*2	2	0	2

It is clear that t75 cannot be the representation of any reasonable conjunction, disjunction or implication. Nevertheless, according to the pattern followed in the cases of t74 and t75, we get: (I) $T \in I(A \circ B)$ iff $T \in I(B)$ or $[T \in I(A) \ \& \ F \in I(B)]$; $F \in I(A \circ B)$ iff $[T \notin I(B) \ \& \ F \notin I(B)]$ or $[T \notin I(A) \ \& \ F \notin I(A) \ \& \ F \in I(B)]$. (II) (a) $A \circ B \in \mathcal{T}$ iff $B \in \mathcal{T}$ or $[A \in \mathcal{T} \ \& \ \neg B \in \mathcal{T}]$; (b) $\neg(A \circ B) \in \mathcal{T}$ iff $[B \notin \mathcal{T} \ \& \ \neg B \notin \mathcal{T}]$ or $[A \notin \mathcal{T} \ \& \ \neg A \notin \mathcal{T} \ \& \ \neg B \in \mathcal{T}]$ (\mathcal{T} is a prime, regular and consistent L75-theory). (III) The following rules and thesis are added to b³d in order to axiomatize L75: (a) $B \Rightarrow A \circ B$; $A \wedge \neg B \Rightarrow A \circ B$; $A \circ B \Rightarrow B \vee \neg B$; (b) $B \vee \neg B \vee \neg(A \circ B)$; $\neg B \Rightarrow (A \vee \neg A) \vee \neg(A \circ B)$; $\neg(A \circ B) \wedge B \Rightarrow C$; $\neg(A \circ B) \wedge (A \wedge \neg B) \Rightarrow C$; $\neg(A \circ B) \wedge (\neg A \wedge \neg B) \Rightarrow C$.

Finally, consider the unary table:

t76.		◦
	0	0
	*1	0
	*2	2

We have: (I) $T \in I(\circ A)$ iff $F \notin I(A)$; $F \in I(\circ A)$ iff $F \in I(A)$. (II) $\circ A \in \mathcal{T}$ iff $\neg A \notin \mathcal{T}$; $\neg \circ A \in \mathcal{T}$ iff $\neg A \in \mathcal{T}$ (\mathcal{T} is a prime, regular and non-trivial L76-theory). (III) The following rules and thesis are added to b³d in order to axiomatize L76: $\circ A \vee \neg A$; $\circ A \wedge \neg A \Rightarrow B$; $\neg \circ A \Leftrightarrow \neg A$. (\mathcal{T} is a prime, regular and non-trivial L76-theory.)

In the sequel, we comment a question that has to be addressed. By using the method of ‘correspondence analysis’, all 3-valued unary and binary expansions of the logic of Paradox (LP) in [16] and those of Kleene’s strong logic (K3) in [38] are formulated with natural deduction systems (cf. also [22], [23]). *Grosso modo*, the idea is to characterize each entry in a 3-valued table for a unary or a binary connective by a corresponding rule. However, in the quoted papers, it is not explained how the rule corresponding to the entry in question can be found. Nonetheless, the desired explanation can be provided by using Brady’s method in [6], as displayed in the present paper. (In what follows, we generally employ the symbols and terminology used so far in the paper.)

Consider, e.g., the case of LP. The designated values are 1 and 2. Then, as explained in §4, we have $v(A) = 2$ iff $F \notin I(A)$; $v(A) = 1$ iff $T \in I(A)$ and $F \in I(A)$ and $v(A) = 0$ iff $T \notin I(A)$, where v is any LP-interpretation built upon the set $\{0, 1, 2\}$, I being its corresponding interpretation in the overdetermined BD-

semantics for LP. Let now ELP be an expansion of LP with a binary connective \circ . The proof of completeness for ELP proceeds as shown above in the case of Li -logics and L73, L74, L75 or L76. In particular, we can assume that a prime, regular, non-trivial LP-theory \mathcal{T} has been built, but we need to know what rules (resp., axioms) governing $\circ \mathcal{T}$ has to be closed under (resp., contain) in order to make of \mathcal{T} a prime, regular, non-trivial ELP-theory. Consider then the first entry in the \circ -table and let v be any ELP-interpretation. There are three possibilities: (a) $v(A \circ B) = 0$; (b) $v(A \circ B) = 1$ and (c) $v(A \circ B) = 2$, which, in terms of the corresponding BD-interpretation I are: Suppose $T \notin I(A)$ and $T \notin I(B)$. Then, (a) $T \notin I(A \circ B)$; (b) $T \in I(A \circ B)$ and $F \in I(A \circ B)$, and (c) $F \notin I(A \circ B)$. Conditions (a), (b) and (c) are canonically interpreted as follows: (a) if $A \circ B \in \mathcal{T}$, then $A \vee B \in \mathcal{T}$; (b) if $A \circ B \notin \mathcal{T}$ or $\neg(A \circ B) \notin \mathcal{T}$, then $A \vee B \in \mathcal{T}$; (c) if $\neg(A \circ B) \in \mathcal{T}$, then $A \vee B \in \mathcal{T}$. Finally, the canonical interpretations just provided give us the following rules: (a) $A \circ B \Rightarrow A \vee B$; (b) $\neg[(A \circ B) \wedge \neg(A \circ B)] \Rightarrow A \vee B$ (or the thesis $[(A \circ B) \wedge (\neg A \circ B)] \vee (A \vee B)$); (c) $\neg(A \circ B) \Rightarrow A \vee B$. These rules are precisely the rules corresponding to the v -cases (a), (b) and (c) above provided in [16], p. 722. The rest of the entries in the \circ -table can be treated similarly. Also, expansions of K3 can be handled analogously as those of LP have been addressed.

Thus, we see, ‘correspondence analysis’ and Brady’s procedure, as displayed in [29] and [33] are kindred methods. Nevertheless, it seems that the latter has at least two advantages over the former. Firstly, it is developed upon the basis of a clear and interesting semantics, BD-semantics. Secondly, it supplies an easy heuristics to find the axioms and/or rules corresponding to the truth-table of a unary or binary 3-valued connective. On the other hand, we note that Brady’s method has successfully been applied to 4-valued logics (cf., e.g., [18], [19], [27]) as well as to implicative expansions of basic logics other than Kleene’s strong logic with either one or two designated values (cf., e.g., [25]). In this sense, it would be interesting to investigate to what extent Brady’s method can be expanded.

Let us end this section by remarking that all the systems referred to in the present paper and in its predecessors can possibly be formulated as Gentzen systems following the method in [2] and [3]. (In [24], §8, the reader can find a comparison of the ‘corresponding analysis’ method and the one developed in [2] and [3].)

7 Concluding remarks

In what follows, we use the term Li -logics for generally referring to the 72 logics investigated in the present paper. The terms Li_1 -logics and Li_2 -logics will particularly refer to the 64 logics and the 8 logics in sections 2 and 5, respectively. We briefly

comment upon some of the properties of the Li -logics.

1. All Li_1 -logics are paraconsistent in the sense that rule ECQ, $A \wedge \neg A \Rightarrow B$, does not hold in any of them. The Li_2 -logics are not paraconsistent (ECQ is present in the eight logics), but in turn they are paracomplete in the sense that it is not the case that for any prime, regular and consistent Li_2 -theory t we have $A \in t$ or $\neg A \in t$ for any wff A .
2. None of the Li -logics has the variable-sharing property (vsp) (cf. [1]): $(A \wedge \neg A) \rightarrow (B \vee \neg B)$ is provable in the 72 logics. Nevertheless, the 'quasi variable-sharing property' (qsvp) is a property of the 32 Li_1 -logics the determining matrix of which has $f_{\rightarrow}(1, 1) = 1$ (the proof of this fact is similar to the ones provided in the Appendix II in [28]). The qsvp reads: if $A \rightarrow B$ is provable, then either (i) A and B share a propositional variable or (ii) both $\neg A$ and B are provable (cf. [1], p. 417). Consequently, these 32 Li_1 -logics are quasi-relevant logics in the same sense as the logic R-Mingle (cf [1], §29).
3. As pointed out in §6, the only Li -logics exhibiting a natural conditional (either in the sense of [28] or in that of [39]) are L1, L3, L5, L7, L9, L11, L13 and L15 (Li_1 -logics) and L69-L72 (Li_2 -logics). To the best of our knowledge, the rest of the logics have not been individualized in the literature, though they could be given natural deduction systems or even cut-free sequent calculi (cf. §6; cf. also [26]).
4. We have tried to axiomatize the 72 logics treated in the present paper in the most possible general way. But most of the Li -logics can be given more conspicuous and economic axiomatizations (cf., for instance, the suggestions in the concluding remarks of [29] and [33]).
5. Consider the following general tables T1-T5 (a_i ($1 \leq i \leq 7 \in \{0, 1, 2\}$); b_i ($1 \leq i \leq 4 \in \{0, 2\}$):

<table style="border-collapse: collapse;"> <tr><td style="padding-right: 10px;">T1.</td><td style="border-right: 1px solid black; padding-right: 5px;">\rightarrow</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">2</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">a_1</td><td style="padding: 2px 5px;">a_2</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">1</td><td style="padding: 2px 5px;">a_3</td><td style="padding: 2px 5px;">2</td><td style="padding: 2px 5px;">a_4</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">2</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">a_5</td><td style="padding: 2px 5px;">a_6</td></tr> </table>	T1.	\rightarrow	0	1	2		0	1	a_1	a_2		1	a_3	2	a_4		2	0	a_5	a_6	<table style="border-collapse: collapse;"> <tr><td style="padding-right: 10px;">T2.</td><td style="border-right: 1px solid black; padding-right: 5px;">\rightarrow</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">2</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">0</td><td style="padding: 2px 5px;">b_1</td><td style="padding: 2px 5px;">a_1</td><td style="padding: 2px 5px;">1</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">1</td><td style="padding: 2px 5px;">a_2</td><td style="padding: 2px 5px;">2</td><td style="padding: 2px 5px;">a_3</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">2</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">a_4</td><td style="padding: 2px 5px;">a_5</td></tr> </table>	T2.	\rightarrow	0	1	2		0	b_1	a_1	1		1	a_2	2	a_3		2	0	a_4	a_5	<table style="border-collapse: collapse;"> <tr><td style="padding-right: 10px;">T3.</td><td style="border-right: 1px solid black; padding-right: 5px;">\rightarrow</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">2</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">0</td><td style="padding: 2px 5px;">b_1</td><td style="padding: 2px 5px;">a_1</td><td style="padding: 2px 5px;">b_2</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">1</td><td style="padding: 2px 5px;">a_2</td><td style="padding: 2px 5px;">2</td><td style="padding: 2px 5px;">a_3</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">2</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">a_4</td><td style="padding: 2px 5px;">1</td></tr> </table>	T3.	\rightarrow	0	1	2		0	b_1	a_1	b_2		1	a_2	2	a_3		2	0	a_4	1
T1.	\rightarrow	0	1	2																																																										
	0	1	a_1	a_2																																																										
	1	a_3	2	a_4																																																										
	2	0	a_5	a_6																																																										
T2.	\rightarrow	0	1	2																																																										
	0	b_1	a_1	1																																																										
	1	a_2	2	a_3																																																										
	2	0	a_4	a_5																																																										
T3.	\rightarrow	0	1	2																																																										
	0	b_1	a_1	b_2																																																										
	1	a_2	2	a_3																																																										
	2	0	a_4	1																																																										
<table style="border-collapse: collapse;"> <tr><td style="padding-right: 10px;">T4.</td><td style="border-right: 1px solid black; padding-right: 5px;">\rightarrow</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">2</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">0</td><td style="padding: 2px 5px;">a_1</td><td style="padding: 2px 5px;">a_2</td><td style="padding: 2px 5px;">a_3</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">1</td><td style="padding: 2px 5px;">a_4</td><td style="padding: 2px 5px;">2</td><td style="padding: 2px 5px;">a_5</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">2</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">a_6</td><td style="padding: 2px 5px;">a_7</td></tr> </table>	T4.	\rightarrow	0	1	2		0	a_1	a_2	a_3		1	a_4	2	a_5		2	0	a_6	a_7	<table style="border-collapse: collapse;"> <tr><td style="padding-right: 10px;">T5.</td><td style="border-right: 1px solid black; padding-right: 5px;">\rightarrow</td><td style="padding: 2px 5px;">0</td><td style="padding: 2px 5px;">1</td><td style="padding: 2px 5px;">2</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">0</td><td style="padding: 2px 5px;">b_1</td><td style="padding: 2px 5px;">a_1</td><td style="padding: 2px 5px;">b_2</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">1</td><td style="padding: 2px 5px;">a_2</td><td style="padding: 2px 5px;">a_3</td><td style="padding: 2px 5px;">a_4</td></tr> <tr><td></td><td style="border-right: 1px solid black; padding-right: 5px;">2</td><td style="padding: 2px 5px;">b_3</td><td style="padding: 2px 5px;">a_5</td><td style="padding: 2px 5px;">b_4</td></tr> </table>	T5.	\rightarrow	0	1	2		0	b_1	a_1	b_2		1	a_2	a_3	a_4		2	b_3	a_5	b_4																					
T4.	\rightarrow	0	1	2																																																										
	0	a_1	a_2	a_3																																																										
	1	a_4	2	a_5																																																										
	2	0	a_6	a_7																																																										
T5.	\rightarrow	0	1	2																																																										
	0	b_1	a_1	b_2																																																										
	1	a_2	a_3	a_4																																																										
	2	b_3	a_5	b_4																																																										

In [30] and [31], the following facts are proven:

- (a) Any expansion of MK3 built up by adding any of the 1539 functions in tables T1-T3 is functionally complete for the set of 3 truth-values THREE (cf. [30]). Consequently, according to T1-T3, there are 28 Li_1 -logics functionally complete for THREE: L2, L4, L10, L12, L17-L20, L25-L28, L33-L36, L41-L44, L49-L52, L57-L60.
- (b) Any expansion of MK3 defined by adding any of the 2187 functions in T4 functionally contains Łukasiewicz's 3-valued logic Ł3 (cf. [31]). Therefore, in addition to the 28 Li -logics just quoted in (5) above, 4 Li_1 -logics contain Ł3 according to T4: L1, L3, L9 and L11.
- (c) Any expansion of MK3 defined by adding any of the 3888 functions in T5 is functionally included in Ł3 (cf. [9], [31]). Thus, there are 8 Li_1 -logics functionally included in Ł3 according to T5: L1, L3, L5, L7, L9, L11, L13 and L15.

In addition, we have:

- (d) Let L be an Li_1 -logic whose determining matrix has $f_{\rightarrow}(1,1) = 1$. It is obvious that L does not functionally include Ł3.

Consequently, we have:

- (e) Given (a), (c) and (d), there are exactly 28 Li_1 -logics functionally complete for THREE: those remarked in (a) above.
 - (f) Given (a), (b) and (d), there are exactly 32 Li_1 -logics functionally including Ł3: those remarked in (a) and (b) above.
 - (g) Given (b) and (c), there are 4 Li_1 -logics functionally equivalent to Ł3: L1, L3, L9 and L11.
6. Concerning Li_2 -logics, we have: (a) L69-L72 are functionally equivalent to Ł3 (cf. [28]); (b) following the suggestions in the concluding remarks of [30], it can be shown that L65, L67 and L68 are functionally complete for THREE. Regarding L66, it can be shown that this logic defines Šlupecki's T-operator (cf. [37], [30]), but its relationship to Ł3 is left open.
7. As remarked in the introduction to this paper, we hope that its results can be used to define interesting 4-valued implicative expansions of FDE.

A Appendix

A.1 Basic systems for axiomatizing 3-valued logics

- b_a^3 . The basic system b^3 in [29]: a1-a11 of b^3 and a15 of b_1^3 (cf. Definitions 3.1 and 3.2) together with Adj and MP.
- b_b^3 . The basic system b^3 in [33]: a1-a5, a8-a10 of b_a^3 together with Adj, MP, dMP, dTrans, dEq.
- b_c^3 . The basic system b^3 in [32]:

Axioms: (a1) $A \rightarrow A$; (a2) $(A \wedge B) \leftrightarrow (B \wedge A)$; (a3) $[A \wedge (B \wedge C)] \leftrightarrow [(A \wedge B) \wedge C]$; (a4) $[(A \vee B) \wedge (A \vee C)] \leftrightarrow [A \vee (B \wedge C)]$; (a5) $\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$; (a6) $A \leftrightarrow \neg\neg A$; (a7) $A \vee \neg A$.

Rules: Adj; MP; $I\wedge$ and $E\vee$ of FDE (cf. Definition 2.3) and $(E\wedge) A \wedge B \Rightarrow A, B$; $(IV) A \Rightarrow A \vee B, B \vee A$; $(Trans\leftrightarrow) A \leftrightarrow B, B \leftrightarrow C \Rightarrow A \leftrightarrow C$; $(Pref\leftrightarrow) A \leftrightarrow B \Rightarrow (C \rightarrow A) \leftrightarrow (C \rightarrow B)$; $(Suf\leftrightarrow) A \leftrightarrow B \Rightarrow (A \rightarrow C) \rightarrow (B \rightarrow C)$; $(Fac$ w.r.t. $\leftrightarrow) A \leftrightarrow B \Rightarrow (C \wedge A) \leftrightarrow (C \wedge B)$; $(Con$ w.r.t. $\leftrightarrow) A \leftrightarrow B \Rightarrow \neg A \leftrightarrow \neg B$; $(Metarule)$ If $A, B \Rightarrow C$, then $D \vee A, D \vee B \Rightarrow D \vee C$.

- b_d^3 . The basic system b^3 for implicative expansions of MK3 built upon tables td through tl: Adj; MP; $E\wedge$; IV ; $(DM_1) \neg(A \vee B) \Rightarrow \neg A \wedge \neg B, \neg A \wedge \neg B \Rightarrow \neg(A \vee B)$; $(DM_2) \neg(A \wedge B) \Rightarrow \neg A \vee \neg B, \neg A \vee \neg B \Rightarrow \neg(A \wedge B)$; $(DN) A \Rightarrow \neg\neg A, \neg\neg A \Rightarrow A$; $(Metarule 1 —Mr1)$ If $A_1, \dots, A_n \Rightarrow B$, then $C \vee A_1, \dots, C \vee A_n \Rightarrow C \vee B$; $(Metarule 2 —Mr2)$ If $D_1, \dots, D_n, A \Rightarrow C$ and $E_1, \dots, E_m, B \Rightarrow C$, then $D_1, \dots, D_n, E_1, E_m, A \vee B \Rightarrow C$.

A.2 Subclassical and contraclassical matrices (cf. [14])

Let M be an implicative expansion of MK3. M is a *subclassical matrix* if all the wffs it verifies are classical tautologies when \wedge, \vee, \neg and \rightarrow are understood as the classical connectives. M is a *contraclassical matrix* if it is not a subclassical matrix. (Notice that natural implicative expansions of MK3 (in both senses of ‘natural’ discussed in §6) are subclassical matrices, but that there are non-natural subclassical matrices, as Logic of Paradox, LP, e.g.) Matrices M1, M3, M5, M7, M9, M11, M13, M15, M69, M70, M71 and M72 are natural implicative expansions of MK3 and so, subclassical matrices. The rest of the implicative expansions of MK3 treated in the paper are contraclassical matrices.

A.3 Subclassical and contraclassical wffs (cf. [14])

A wff A is *subclassical* if it is a classical tautology when its connectives are understood as the classical connectives. A is *contraclassical* if it is not a subclassical wff. The contraclassical axioms the logics in this paper have been axiomatized with are: A1, A2, A3, A4, A5, A6, A35, A36, A38 and A40. In the following list, it is noted the contraclassical theses verified by each contraclassical matrix.

A.4 Contraclassical theses verified by each contraclassical matrix

M2: A5.	M39: A2.
M4: A5.	M40: A1, A2, A5.
M6: A5.	M41: A2.
M8: A5.	M42: A2, A5.
M10: A5.	M43: A2.
M12: A5.	M44: A1, A2, A5.
M14: A5.	M45: A2.
M16: A5.	M46: A2.
M17: A6.	M47: A2.
M18: A3, A5, A6.	M48: A1, A2, A5.
M19: A3, A6.	M49: A2, A3, A6.
M20: A3, A5, A6.	M50: A2, A3, A5, A6.
M21: A3, A6.	M51: A2, A3, A6.
M22: A3, A5, A6.	M52: A1, A2, A3, A5, A6.
M23: A3, A6.	M53: A2, A3, A6.
M24: A3, A5, A6.	M54: A2, A3, A5, A6.
M25: A3, A6.	M55: A2, A3, A6.
M26: A3, A4, A5, A6.	M56: A1, A2, A3, A5, A6.
M27: A3, A6.	M57: A2, A3, A6.
M28: A3, A6.	M58: A2, A3, A4, A5, A6.
M29: A3, A6.	M59: A2, A3, A6.
M30: A3, A6.	M60: A1, A2, A3, A4, A5, A6.
M31: A3, A6.	M61: A2, A3, A6.
M32: A3, A6.	M62: A2, A3, A4, A5, A6.
M33: A2.	M63: A2, A3, A6.
M34: A2, A5.	M64: A1, A2, A3, A4, A5, A6.
M35: A2.	M65: A35, A38.
M36: A1, A2, A5.	M66: A35, A36, A39, A40.
M37: A2.	M67: A36.
M38: A2, A5.	M68: A38, A40.

A.5 Aristotle's and Boethius' theses

Aristotle's thesis is $\neg(A \rightarrow \neg A)$ (or $\neg(\neg A \rightarrow A)$). Boethius' thesis is $(A \rightarrow B) \rightarrow \neg(A \rightarrow \neg B)$ (or $(A \rightarrow \neg B) \rightarrow \neg(A \rightarrow B)$). Of course, both theses are contraclassical wffs, but they are important in connexive logic (cf., e.g., [17] or [40]). We remark the implicative expansions of MK3 verifying each one of these theses.

Aristotle's thesis is verified by M6, M8, M14, M16, M22, M24, M30, M32, M38, M40, M46, M48, M54, M56, M62 and M64. (We note that tables with $f_{\rightarrow}(1, 1) = 2$ or $f_{\rightarrow}(0, 2) = 2$ falsify this thesis; also, that M64 is introduced in [20], where Aristotle's thesis is discussed in the context of relevant logics.)

Boethius' thesis is verified by M40, M48, M56 and M64. (We remark that Boethius' thesis is falsified if Aristotle's thesis is falsified; also, if $f_{\rightarrow}(0, 0) = 2$ or $f_{\rightarrow}(0, 1) = 2$, it is also falsified.)

References

- [1] Anderson, A. R., Belnap, N. D. J. (1975). *Entailment. The logic of relevance and necessity*, vol I. Princeton, NJ: Princeton University Press.
- [2] Avron, A., Ben-Naim, J., Konikowska, B. (2007). Cut-free ordinary sequent calculi for logics having generalized finite-valued semantics. *Logica Universalis*, 1(1), 41-70. <https://doi.org/10.1007/s11787-006-0003-6>
- [3] Avron, A., Konikowska, B., Zamansky, A. (2013). Cut-free sequent calculi for C-systems with generalized finite-valued semantics. *Journal of Logic and Computation*, 23(3), 517-540. <https://doi.org/10.1093/logcom/exs039>
- [4] Belnap, N. D. Jr. (1977a). A useful four-valued logic. In G. Epstein, J. M. Dunn (Eds.), *Modern Uses of Multiple-Valued Logic* (pp. 8-37). D. Reidel Publishing Co., Dordrecht.
- [5] Belnap, N. D. Jr. (1977b). How a computer should think. In G. Ryle (Ed.), *Contemporary Aspects of Philosophy* (pp. 30-55). Oriel Press Ltd., Stocksfield.
- [6] Brady, R. T. (1982). Completeness proofs for the systems RM3 and BN4. *Logique et Analyse*, 25, 9-32.
- [7] Dunn, J. M. (1976). Intuitive semantics for first-degree entailments and "coupled trees." *Philosophical Studies*, 29, 149-168.
- [8] Dunn, J. M. (2000). Partiality and its dual. *Studia Logica*, 66(1), 5-40. <https://doi.org/10.1023/A:1026740726955>.
- [9] Finn, V. K. (1969). O predpolnote klasa funktsii, sootvetstvuyushchego trekhznachnoi logike J. Łukasiewicza (The precompleteness of the class of functions that corresponds to the three-valued Logic of J. Łukasiewicz). *Nauchno-tekhnicheskaya informatsiya*. Ser. 2, 10, 35-38.
- [10] Fitting, M. (1991). Kleene's logic, generalized. *Journal of Logic and Computation*, 1(6), 797-810. <https://doi.org/10.1093/logcom/1.6.797>

- [11] Fitting, M. (1994). Kleene's three valued logics and their children. *Fundamenta Informaticae*, 20(1,2,3), 113-131. <https://doi.org/10.3233/FI-1994-201234>
- [12] González, C. (2011). *MaTest*. <https://sites.google.com/site/sefusmendez/matest> (last accessed 20/02/2021).
- [13] Hacking, I. (1963). What is strict implication? *Journal of Symbolic Logic*, 28(1), 51-71.
- [14] Humberstone, L. (2000). Contra-classical logics. *Australasian Journal of Philosophy*, 78(4), 438-474. <https://doi.org/10.1080/00048400012349741>
- [15] Kleene, S. C. (1962). *Introduction to metamathematics*. North Holland. Reprinted Ishi Press, 2009.
- [16] Kooi, B., & Tamminga, A. (2012). Completeness via correspondence for extensions of the logic of paradox. *The Review of Symbolic Logic*, 5(4), 720-730. <https://doi.org/10.1017/S1755020312000196>
- [17] McCall, S. (2012). A history of connexivity. In D. M. Gabbay, F. J. Pelletier, J. Woods (Eds.), *Handbook of the History of Logic* (vol. 11, pp. 415–449). North-Holland. <https://doi.org/10.1016/B978-0-444-52937-4.50008-3>
- [18] Méndez, J. M., Robles, G. (2016). The logic determined by Smiley's matrix for Anderson and Belnap's first degree entailment logic. *Journal of Applied Non-Classical Logics*, 26(1), 47-68. <http://doi.org/10.1080/11663081.2016.1153930>.
- [19] Méndez, J. M., Robles, G. (2015). A strong and rich 4-valued modal logic without Łukasiewicz-type paradoxes. *Logica Universalis*, 9(4), 501-522, <https://doi.org/10.1007/s11787-015-0130-z>.
- [20] Mortensen, C. (1984). Aristotle's thesis in consistent and inconsistent logics. *Studia Logica*, 43(1), 107-116. <https://doi.org/10.1007/BF00935744>
- [21] Omori, H., & Wansing, H. (2017). 40 years of FDE: An introductory overview. *Studia Logica*, 105(6), 1021-1049. <https://doi.org/10.1007/s11225-017-9748-6>
- [22] Petrukhin, Y. (2017). Natural deduction for three-valued regular logics. *Logic and Logical Philosophy*, 26(2), 197-206. <https://doi.org/10.12775/LLP.2016.025>
- [23] Petrukhin, Y., Shangin, V. (2018). Natural three-valued logics characterized by natural deduction. *Logique et Analyse*, 61(244), 407-427.
- [24] Petrukhin, Y., Shangin, V. (2020). Correspondence analysis and automated proof-searching for first degree entailment. *European Journal of Mathematics*, 6, 1452–1495. <https://doi.org/10.1007/s40879-019-00344-5>
- [25] Robles, G. (2014). A simple Henkin-style completeness proof for Gödel 3-valued logic G3. *Logic and Logical Philosophy*, 23(4), 371-390, <https://doi.org/10.12775/LLP.2014.001>
- [26] Robles, G., López, S. M. (2020). Selecting the class of all 3-valued implicative expansions of Kleene's strong logic containing Routley and Meyer's logic B. *Logique et Analyse*, 252, 443-464, <https://doi.org/10.2143/LEA.252.0.3289034>
- [27] Robles, G., Méndez, J. M. (2016). A companion to Brady's 4-valued relevant logic BN4: The 4-valued logic of entailment E4. *Logic Journal of the IGPL*, 24(5), 838-858, <https://doi.org/10.1093/jigpal/jzw011>

- [28] Robles, G., Méndez, J. M. (2019). The class of all natural implicative expansions of Kleene's strong logic functionally equivalent to Łukasiewicz's 3-valued logic Ł3. *Journal of Logic, Language and Information*, Online First, <https://doi.org/10.1007/s10849-019-09306-2>
- [29] Robles, G., Méndez, J. M. (2019). Belnap-Dunn semantics for natural implicative expansions of Kleene's strong three-valued matrix with two designated values. *Journal of Applied Non-Classical Logics*, 29(1), 37-63, <https://doi.org/10.1080/11663081.2018.1534487>
- [30] Robles, G., Méndez, J. M. (2020). A remark on functional completeness of binary expansions of Kleene's strong 3-valued logic. *Logic Journal of the IGPL*, jzaa028 (Advance article), <https://doi.org/10.1093/jigpal/jzaa028>
- [31] Robles, G., Méndez, J. M. (Accepted). A class of implicative expansions of Kleene's strong logic, a subclass of which is shown functionally complete via the precompleteness of Łukasiewicz's 3-valued logic Ł3. *Journal of Logic, Language and Information*.
- [32] Robles, G., Méndez, J. M. (Submitted). Three-valued logics free from paradoxes of relevance.
- [33] Robles, G., Salto, F., Méndez, J. M. (2019). Belnap-Dunn semantics for natural implicative expansions of Kleene's strong three-valued matrix II. Only one designated value. *Journal of Applied Non-Classical Logics* 29(3), 307-325, <https://doi.org/10.1080/11663081.2019.1644079>
- [34] Routley, R. Meyer, R. K., Plumwood, V., Brady, R. T. (1982). *Relevant logics and their rivals*, vol. 1. Atascadero, CA: Ridgeview Publishing Co.
- [35] Shramko, Y. (2019). First-degree entailment and structural reasoning. In H. Omori, H. Wansing (Eds.), *New Essays on Belnap-Dunn Logic* (pp. 311-324). Springer International Publishing. https://doi.org/10.1007/978-3-030-31136-0_18
- [36] Slaney, J. K. (1995). *MaGIC, Matrix Generator for Implication Connectives: version 2.1, notes and guide*. Canberra: Australian National University. <http://users.cecs.anu.edu.au/jks/magic.html> (last accessed 17/03/2020).
- [37] Shupecki, J. (1936). Der volle dreiwertige Aussagenkalkül. *Comptes Rendus Des Séances de La Société des Sciences et des Lettres de Varsovie*, Classe III, 29, 9-11.
- [38] Tamminga, A. (2014). Correspondence analysis for strong three-valued logic. *Logical Investigations*, 20(1), 255-268.
- [39] Tomova, N. (2012). A Lattice of implicative extensions of regular Kleene's logics. *Reports on Mathematical Logic*, 47, 173-182. <http://doi.org/10.4467/20842589RM.12.008.0689>
- [40] Wansing, H. (2021). Connexive logic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2021/entries/logic-connexive/>
- [41] Wójcicki, R. (1988). *Theory of logical calculi: basic theory of consequence operations*. Springer Netherlands. <https://doi.org/10.1007/978-94-015-6942-2>

ENAYAT THEORIES

ALBERT VISSER

*Philosophy, Faculty of Humanities, Utrecht University, Janskerkhof 13,
3512BL Utrecht, The Netherlands*
a.visser@uu.nl

Abstract

In this paper we study solution attempts for a problem posed by Ali Enayat: can there be a finitely axiomatised consistent sequential theory that interprets itself plus the (sentential or non-uniform) Tarski biconditionals? We provide a basic framework for the study of this question and discuss some solution attempts. We connect the question with some interesting conjectures. We briefly touch upon what happens if we consider uniform biconditionals.

1 Introduction: Ali Enayat asks a Question

On January 24, 2014, Ali Enayat sent me an e-mail with subject ‘a simple (!) question’. The content of the mail was as follows.

Suppose \mathbf{B} is a finitely axiomatisable base theory, and \mathbf{B}^\top is \mathbf{B} plus the T-scheme, i.e., biconditionals of the form: S iff $\top(\#S)$, where $\#S$ is the code for S . Question: Is \mathbf{B}^\top interpretable in \mathbf{B} ?

I thought that I would be able to solve the problem, for the most salient case where the base theory is sequential, within a day’s time, but I was sadly mistaken. The problem is still unsolved for the sequential case and even for the more inclusive Vaught case. I give a positive example of a non-Vaught theory that interprets itself plus Tarski biconditionals in Section 4. Admittedly, this example involves a not quite standard Gödel numbering.

Let us say that a theory U that interprets U plus the Tarski biconditionals for the language of U is an *Enayat theory*.¹ Ali’s question suggests the following conjecture.

I thank Ali Enayat for his inspiring question. I am grateful to Ali Enayat, Volker Halbach, Mateusz Łełyk, Bartosz Weisło and Fedor Pakhomov for inspiring discussions.

¹As explained in Section 3, it is somewhat more subtle to get the question right. The notion of Enayat theory *tout court* only makes sense in the case one considers Vaught theories.

Conjecture 1.1. There are no consistent, finitely axiomatised, sequential Enayat theories. ○

The argument in favor of the conjecture is simply the rhetorical question: *what could such an interpretation possibly look like?*

Why publish my failed attempts? I think there are some good reasons to do that.

- It is good to make people aware of the problem. It is a basic and intricate problem from the logico-technical standpoint. I think the related problems concerning the meta-mathematics of first-order theories formulated in this paper illustrate that our problem leads to interesting further questions.

Also, the problem fits, at the lower end, in the broader logico-philosophical program of research into truth theories. It is, in a sense, about the informativeness of the minimal typed truth theory.

Along a different line, I think it is time logicians would look a bit more at sentential schemes (in some broad sense). These often behave differently from their uniform brethren. The primary example of a sentential scheme is parameter-free induction. See, e.g., [10], [1] and [4]. For a slightly different, but related, study, see [21].

- The paper provides the basic framework for the study of Ali's question. It is good to have these things out of the way.
- If a reader would want to try her hand on the problem, the attempts contained in the paper would at least spare her the time to rediscover those.
- In Section 7, we develop Saccheri style what an 'Enayat world' would look like. This provides some further basics that can play a role in a solution.
- In the study of this problem, errors are everywhere dense. Especially, one has to keep the dependencies of the complexities of the various items involved in an argument straight. (I often had the illusion of having solved the problem for days, but then a subtle circularity of dependencies turned up.) The paper provides, I hope, an example of good practice in keeping track of dependencies.
- I feel some of the arguments in the paper are definitely entertaining. A good example is the proof of Theorem 8.8.

In Section 5, we have a brief look at what happens when we consider uniform biconditionals. In a subsequent paper, I hope to give a fuller picture.

2 Basics

In this section, we provide the basic framework for the study of Enayat theories.

2.1 Theories, Translations and Interpretations

Theories are, in our paper, theories of first order predicate logic of finite signature that are given by a sufficiently simple set of axioms, say Δ_1^b . In the few cases where we diverge from this format it will be explicitly mentioned. The axiom set is part of the data of a theory.

We refer the reader for a discussion of translations and interpretations to one of our papers [20] or [23] or [24] or [25]. Here we just fix some notations.

- We write $U \triangleright V$ for U interprets V .
- We write $\Gamma \triangleright_U \Delta$ for Γ interprets Δ over U , i.o.w., $(U + \Gamma) \triangleright (U + \Delta)$. Here Γ and Δ will be, in the typical case, sets of sentences, each with a signature that is an extension of the signature of U .
- We write $\Gamma \blacktriangleright_U \Delta$ for Γ Fujimoto interprets Δ over U . This means that we have an identity-preserving, unrelativised interpretation of $U + \Delta$ in $U + \Gamma$ that preserves the vocabulary of U . Here Γ and Δ have signatures that extend the signature of U .
- $\triangleright_{\text{loc}}$ stands for local interpretability and $\triangleright_{\text{dir}}$ stands for direct interpretability. Direct interpretability is unrelativised and identity preserving interpretability. We allow more-dimensional direct interpretability.
- We use \bowtie for mutual interpretability and \blacktriangleright for mutual Fujimoto interpretability.²

Some knowledge of the book [9] is definitely useful.

2.2 Vaughtness and Sequentiality

In this subsection we define Vaught theories and sequential theories. We refer the reader to [18] and [20] for more information.

²In my earlier papers, I use \equiv for mutual interpretability. However, Lev Beklemishev uses this symbol for other notions of sameness of theories. I think the present notation will eliminate all ambiguity.

2.2.1 Vaught Set Theory

We define *Vaught set theory*, **VS** as follows.

VS1. $\exists x \forall y y \notin x$,

VS2. $\forall u_0 \dots \forall u_{n-1} \exists x \forall y (y \in x \leftrightarrow \bigvee_{i < n} y = u_i)$.

We note that, under the right conventions, **VS1** is the special case for $n = 0$ of **VS2**.

We also note that we do not have extensionality.

We can define Kuratowski pairing in the usual way. Of course, our pairs will not be extensional and the same pair could be implemented by many entities. We define a function as follows:

- f is a function iff $\forall u \in f \forall v \in f (\text{pair}(u) \wedge \text{pair}(v) \wedge ((u)_0 = (v)_0 \rightarrow u = v))$.

We do not just demand the uniqueness of the output but also the uniqueness of the pair that implements a transition. We define $f : x \sim y$ iff f is a bijection between x and y , and $x \sim y$ iff $\exists f f : x \sim y$.

We define **VS⁺** as **VS** plus the axioms saying that \sim is an equivalence relation and that if $f : x \sim y$, then $x \sim f$.

We have the following theorem.

Theorem 2.1. **VS** $\triangleright_{\text{dir}}$ **VS⁺**. *Moreover, the relevant interpretation is one-dimensional.*

We give the proof in Appendix A. Let **R** be the very weak arithmetical theory given formulated by Tarski, Mostowski and Robinson in their classic [14]. We have:

Theorem 2.2. **VS** \triangleright **R**. *Moreover, the relevant interpretation is one-dimensional.*

We give the proof in Appendix A.

A theory U is a *Vaught theory* if $U \triangleright_{\text{dir}}$ **VS**. A theory U is a *Vaught⁺ theory* if $U \triangleright_{\text{dir}}$ **VS⁺**. In these definitions, we allow the direct interpretation to be more-dimensional.

Theorem 2.1 tells us that a Vaught theory is *ipso facto* a Vaught⁺ theory.

2.2.2 Adjunctive Set Theory

We define *adjunctive set theory*, **AS**, as follows:

AS1. $\exists x \forall y y \notin x$,

AS2. $\forall u \forall v \exists x \forall y (y \in x \leftrightarrow (y \in u \vee y = v))$.

A theory U is *sequential* if $U \triangleright_{\text{dir}} \mathbf{AS}$. Here we allow the direct interpretation to be more dimensional.³ Sequentiality is well studied: we refer the reader to [20] for an extensive discussion of the notion.

We present some basic facts concerning sequential theories. We define (for any theory U):

- $\mathfrak{U}(U) := \mathbf{S}_2^1 + \{\diamond_{U,n} \top \mid n \in \omega\}$.
- A theory U is *reflexive* if $U \triangleright \mathfrak{U}(U)$.

Here the $\diamond_{U,n} \top$ are restricted consistency statements for U , where of course U is given by a fixed representation of the axioms and the restriction is both a restriction of the size of the codes of the axioms and of the complexity of the formulas allowed in the proofs. The measure of complexity here is *depth of quantifier alternations*. See, e.g., [19] or [24].

Here are some basic facts.

1. $\mathfrak{U}(U) \triangleright U$ (this holds for any U).
2. If U is sequential, then $U \triangleright_{\text{loc}} \mathfrak{U}(U)$. I.o.w., sequential theories are *locally* reflexive.
3. If U is reflexive, then $U \triangleright_{\text{loc}} V$ implies $U \triangleright V$.
4. Suppose U is sequential and, for all V , we have $U \triangleright_{\text{loc}} V$ implies $U \triangleright V$. Then, U is reflexive.
5. Finitely axiomatised sequential theories are not reflexive. (See [13].)

3 Basic Definitions and Insights

In this section we give a precise definition of Ali's question. My main interest is in Ali's question for sequential theories. However, it is good to have a wider definition. Setting things up with a bit more generality will enable us to apply some informal rigour to the choice of notions. We will see that already for Vaught theories, Ali's question makes good sense.

We consider broader and narrower versions of the question. We will show that the question takes its most natural form in the sequential case.

³Thus, our notion of sequentiality is an extension of the usual one. In [20], I called this notion *polysequentiality*. However, in the light of the facts that (i) we are just looking at a minor extension of the notion and (ii) the modified notion is clearly the right one, it is high time to redefine the traditional notion.

3.1 Basic Formulation of the Question

We first address the treatment of numerals. What is needed here is that the theories we consider interpret at least some minimal theory of successor. For this we choose the theory Succ_0 . The theory Succ_0 has one constant 0 and one binary predicate S. I will use infix notation $\widetilde{}$ for S. The axioms of Succ_0 are as follows. We define $\widetilde{0}(x) := (x = 0)$ and $\widetilde{(n + 1)}(x) := \exists y (\widetilde{n}(y) \wedge y S x)$.

Succ_0 1. $\exists x \widetilde{n}(x)$

Succ_0 2. $\neg (\widetilde{m}(x) \wedge \widetilde{n}(x))$, where $m < n$

Succ_0 3. $(\widetilde{n}(x) \wedge \widetilde{n}(y)) \rightarrow x = y$

So, the theory just says that successor behaves normally as long as we are finitely far removed from 0. We have the following small insight.

Theorem 3.1. *Succ_0 is a sub-theory of R, where we identify $x S y$ with $Sx = y$.*

Consider any theory U and suppose $N : \text{Succ}_0 \rightarrow U$. Suppose N is n -ary. We expand the signature of U with a new n -ary predicate T . We write $\text{T}(\underline{n})$ for:

$$\exists \vec{x} \in \delta_N (\widetilde{n}^N(\vec{x}) \wedge \text{T}(\vec{x})).$$

(In case we want to emphasise the dependence of our numerals on N we write $\text{T}(\underline{n}^N)$.)

The class of sentential Tarski biconditionals TB_N^- consists of the sentences of the form $A \leftrightarrow \text{T}(\ulcorner A \urcorner)$. Here A is a U -sentence and $\ulcorner A \urcorner$ is the Gödel number of A . We will usually omit the underlining and simply write $\text{T}(\ulcorner A \urcorner)$.

We say that N has the *Enayat Property* or that U is an *N -Enayat theory* iff $U \triangleright (U + \text{TB}_N^-)$, or, i.o.w., $\text{T} \triangleright_U \text{TB}_N^-$.

We note that N is part of the data for \underline{n} and U is part of the data for N and the signature of U is part of the data for U . So, indeed, the notation ‘ TB_N^- ’ exhibits all the necessary data, with the exception of the Gödel numbering, to construct the intended set of sentences. In most of the paper, we will treat the Gödel numbering as fixed in the back-ground, where the Gödel numbering is supposed to be a standard efficient Gödel numbering. Exceptions are Subsection 3.3 and Section 4. In Subsection 4, we need a non-standard numbering. In Subsection 3.3, we show that, in the Vaught case, the Enayat property does not depend on the Gödel numbering under the appropriate assumptions of Gödel numberings.

Remark 3.2. We note that, even if our framework is fairly general, the theory Succ_0 may be still too restrictive. The point is that nothing really seems to depend on the uniqueness of the numerals as stipulated in $\text{Succ}_0\mathfrak{3}$. The only advantage of the present approach is that we can use the numeral notation in a meaningful way. One important advantage of the more general approach, where we drop $\text{Succ}_0\mathfrak{3}$, is that also pair theories are covered by the framework. \circ

Here is a first small observation. This observation is well-known. I do not know who first made it.

Theorem 3.3. *Suppose $N : U \triangleright \text{Succ}_0$. Then, $\top \blacktriangleright_{\text{loc},U} \text{TB}_N^-$.*

Proof. We interpret the axioms $A_0 \leftrightarrow \top(\ulcorner A_0 \urcorner)$, \dots , $A_{n-1} \leftrightarrow \top(\ulcorner A_{n-1} \urcorner)$, by defining: $\top(\vec{x}) :\leftrightarrow \bigvee_{k < n} (\widetilde{\ulcorner A_k \urcorner}(\vec{x}) \wedge A_k)$. \square

3.2 Intensionality

Enayatness, as defined here, is an *intensional* property, since, in the general case, it critically depends both on N and on the choice of the Gödel numbering. In Section 4, we will see an example that illustrates these dependencies.

We will see that, in the case that U is a Vaught theory, Enayatness is independent of the Gödel numbering, assuming that all Gödel numberings that we allow are recursively related to some standard Gödel numbering. Secondly, we will see that, for Vaught theories, Enayatness can be considered as a property of theories rather than of interpretations.

In the case of sequential theories, we can even do better: we can give a characterisation of Enayatness in which Gödel numerings nor truth are mentioned!

3.3 Enayatness for Vaught Theories

We show that, for Vaught theories, the property of Enayatness is independent of the choice of numerals. Moreover, we show that, for Vaught theories, Enayatness is independent of the choice of the Gödel numbering, as long as the Gödel numberings considered are recursively related to a standard one.

We have the following theorem:

Theorem 3.4. *Suppose U is a Vaught theory and $N, M : \text{Succ}_0 \triangleleft U$. Then, we have $\text{TB}_N^- \blacktriangleright_U \text{TB}_M^-$. It follows that M has the Enayat Property iff N does.*

Proof. We use a minor adaptation of a well-known argument due to Dedekind and Pudlák. See [13].

We define $\mathfrak{F} := \mathfrak{F}_{N,M}$ between δ_N and δ_M as follows. $\vec{x} \mathfrak{F} \vec{y}$ iff there a partial bijection f between δ_N and δ_M such that (i) $f0_N = f0_M$, (ii) if $\vec{x} \mathfrak{S}_N \vec{y}$ and $f\vec{y}$ is defined, then $f\vec{x}$ is defined and $f\vec{x} \mathfrak{S}_M f\vec{y}$. The definition of partial bijection is provided via the direct interpretation of VS in U .

One now easily shows that

$$\mathfrak{F}(\underline{n}^N) =_M \underline{n}^M \text{ and } \forall \vec{y} ((\delta_M(\vec{y}) \wedge \mathfrak{F}(\underline{n}^N) =_M \vec{y}) \rightarrow \vec{y} =_M \underline{n}^M).$$

We interpret $U + \text{TB}_M^-$ via the translation, say τ , in $U + \text{TB}_N^-$ by taking the identical translation for the U -vocabulary and setting

$$\text{T}_\tau(\vec{y}) := \exists \vec{x} (\delta_N(\vec{x}) \wedge \vec{x} \mathfrak{F} \vec{y} \wedge \text{T}(\vec{y})).$$

The interpretation of $U + \text{TB}_N^-$ in $U + \text{TB}_M^-$ is similar. □

We have found that, if a Vaught theory has the Enayat property for some N , it has the Enayat property for all N . If a Vaught theory has the Enayat property for some N , we will call it simply an *Enayat theory*.

Here is a convenient observation.

Theorem 3.5. *Suppose U is a Vaught theory. Then U is Enayat iff U interprets $U + \text{TB}_N^-$, for some $N : \mathbb{R} \triangleleft U$. Here the dimension of N can be taken to be the dimension of the direct interpretation that establishes Vaughtness.*

Proof. Consider a Vaught theory U . By Theorem 2.2, U interprets \mathbb{R} , say, via N , where N is the composition of the one-dimensional interpretation of \mathbb{R} in VS and the direct interpretation of VS in U . It follows that the dimension of N is the dimension of this direct interpretation.

Since \mathbb{R} extends Succ_0 , by Theorem 3.1, U interprets Succ_0 via the interpretation N' based on τ_N . Clearly, $U + \text{TB}_N^-$ is extensionally the same as $U + \text{TB}_{N'}^-$. So, we are immediately done. □

When considering Vaught theories, we will from now on consider interpretations of \mathbb{R} .

We address the worry that Enayatness for Vaught theories may be crucially dependent on details of the chosen Gödel numbering.

Theorem 3.6. *Suppose we have Gödel numberings ν_0 and ν_1 . We only need to assume that the ν_i assign numbers to sentences. Suppose for some recursive function η we have $\nu_0 = \eta \circ \nu_1$. Suppose $N : R \rightarrow U$. Then, $TB_N^{-\nu_0} \triangleright_U TB_N^{-\nu_1}$.*

It follows that if U is a Vaught theory that is Enayat for ν_0 , then U is Enayat for ν_1 .

Proof. We assume the conditions of the theorem. In R , we can represent the function η by a formula H . We now define a Fujimoto translation τ as follows.

$$T^\tau(\vec{x}) := \delta_N(\vec{x}) \wedge \exists \vec{y} \in \delta_N(H^N(\vec{x}, \vec{y}) \wedge T_0(\vec{y})).$$

It is easy to see that τ delivers the goods. □

We note that we need not impose any *a priori* restriction on the complexity of ν_1 for the theorem to work. Of course, our default assumption is that we are working with a reasonable Gödel numbering.

Here is the general form of our conjecture for Vaught theories.

Conjecture 3.7. No finitely axiomatised consistent Vaught theory is Enayat. ○

We also can ask a more modest question.

Open Question 3.8. Suppose there is a finitely axiomatised, consistent Vaught theory that is Enayat. Can we show, under that assumption, that *all* finitely axiomatised, consistent, Vaught theories are Enayat theories? ○

We end this subsection with one further question.

Open Question 3.9. In Subsection 3.4, we will show that in the recursively enumerable sequential case, we can characterise Enayat theories in a coordinate-free way. Not only is the question of Enayatness independent of the Gödel numbering, but Gödel numberings are not mentioned in the characterisation. Can we do something similar in the Vaught case? ○

3.4 Enayatness for Sequential Theories

In case our theories are sequential and recursively enumerable, we can do better than the previous section by eliminating any reference to numerals and coding from the definition of Enayat theory.

We start with a convenient observation.

Theorem 3.10. *Suppose U is a sequential. Then U is Enayat iff $U \triangleright (U + \text{TB}_N^-)$ for some $N : \mathbf{S}_2^1 \triangleleft U$. We can take the dimension of N to be the dimension of the direct interpretation of AS in U that establishes sequentiality.*

Proof. The proof is entirely analogous to the proof of Theorem 3.5, noting that sequential theories are Vaught and that \mathbf{S}_2^1 extends R and that the interpretation of \mathbf{S}_2^1 in AS is one-dimensional. \square

We say that an interpretation $K : U \triangleleft V$ is *sententially restricted* if, there is an n , such that, for all U -sentences B , there is a V -sentence C , such that $\rho(C) \leq n$ and $V \vdash B^K \leftrightarrow C$. Here ρ is the complexity measure *depth-of-quantifier-alternations*. See [25] for a careful treatment of the measure.

Suppose $N : \mathbf{S}_2^1 \triangleleft V$. We say that an interpretation $K : U \triangleleft V$ is *strongly sententially restricted* w.r.t. N iff, for some V -formula $A(\vec{x})$, where the length of \vec{x} is the dimension of N , we have that, for all U -sentences B , $V \vdash B^K \leftrightarrow A(\ulcorner B \urcorner^N)$.

Theorem 3.11. *Suppose $K : U \triangleleft V$, where V is sequential and recursively enumerable. Let $N : \mathbf{S}_2^1 \triangleleft V$. Then, K is sententially restricted iff K is strongly sententially restricted (w.r.t. N).*

Proof. The right-to-left direction is immediate, noting that numerals only contribute a constant to the complexity independent of the size of the numeral.

We treat left-to-right. Suppose $K : U \triangleleft V$ is sententially restricted. Let the witnessing number be n . Let γ be the the function that takes as input a U -sentence B , searches for the smallest (coded) V -proof with conclusion of the form $B^K \leftrightarrow C$, where $\rho(C) \leq n$, and gives as output C . Clearly γ is a total recursive function.

Let $\text{True}_n(\vec{x})$ be a truth-predicate for V -sentences of complexity $\leq n$ based on a satisfaction predicate $\text{Sat}_n(\vec{s}, \vec{x})$ where we can prove the commutation clauses for formulas of complexity $\leq n$ that are in a suitable V -provable cut J of N . The length of the sequence \vec{x} is the dimension of N . We note that, since standard numbers are in J , we have the Tarski biconditionals for True_n for sentences of the right complexity. See [25] for a detailed treatment of partial truth predicates in sequential theories.

We define $A(\vec{x}) := \text{True}_n(G^N(\vec{x}))$, where G stands for the representation in the arithmetical language of the recursive function γ . Note that we really should have written $\exists \vec{y} \in \delta_N (G^N(\vec{x}, \vec{y}) \wedge \text{True}_n(\vec{y}))$.

Consider any U -sentence B . Suppose $\gamma(B) = C$. We have:

$$\begin{aligned}
 V \vdash A(\ulcorner B \urcorner) &\leftrightarrow \text{True}_n(G^N(\ulcorner B \urcorner)) \\
 &\leftrightarrow \text{True}_n(\ulcorner C \urcorner) \\
 &\leftrightarrow C \\
 &\leftrightarrow B^K
 \end{aligned}$$

□

We now have immediately the following consequence.

Theorem 3.12. *Let V be sequential. Then, V is an Enayat theory iff V has a sententially restricted self-interpretation.*

We note that, for decidable theories, the identity interpretation is restricted. So, having a restricted self-interpretation generally is much broader than being Enayat. We may now formulate the following conjecture:

Conjecture 3.13. No finitely axiomatised consistent sequential theory is Enayat. Equivalently, no finitely axiomatised consistent sequential theory has a sententially restricted self-interpretation. ○

3.5 Preservation over Mutual Interpretability

Surprisingly, Enayatness is preserved for over mutual interpretability if we take \mathbf{R} rather than Succ_0 as the basic arithmetical theory that provides the numerals. More precisely we have the following.

Theorem 3.14. *Suppose $K : U \triangleright V$, $M : V \triangleright U$, $N : V \triangleright \mathbf{R}$, $P : V \triangleright (V + \text{TB}_N^{-V})$. Here the superscript V is there to remind us that we consider TB for the signature of V . Let $\mathcal{E} : V \rightarrow (V + \text{TB}_N^{-V})$ be the identical embedding. Let $N^* := K \circ P \circ \mathcal{E} \circ N$. Then, U is an N^* -Enayat theory.*

Proof. Let $Q_0 := K \circ P \circ \mathcal{E} \circ M$. So, $Q_0 : U \triangleright U$. More graphically, the situation looks like this (using the category theoretical notation for interpretations):

$$U \xrightarrow{M} V \xrightarrow{\mathcal{E}} (V + \text{TB}_N^{-V}) \xrightarrow{P} V \xrightarrow{K} U$$

We extend Q_0 to $Q : U \triangleright (U + \text{TB}_{N^*}^{-U})$ as follows: τ_Q is τ_{Q_0} on the vocabulary of U .

- $\mathsf{T}_{\tau_Q}(\vec{x}) : \leftrightarrow \mathsf{T}_{K \circ P}(\tilde{\tau}_M(\vec{x}))$.

Here $\tilde{\tau}_M$ is the arithmetisation of τ_M in N^* . We have:

$$\begin{aligned}
 U \vdash \mathbb{T}^Q(\ulcorner A \urcorner^{N^*}) &\leftrightarrow \mathbb{T}^{K \circ P}(\tilde{\tau}_M(\ulcorner A \urcorner^{N^*})) \\
 &\leftrightarrow \mathbb{T}^{K \circ P}(\ulcorner A^M \urcorner^{N^*}) \\
 &\leftrightarrow (\mathbb{T}(\ulcorner A^M \urcorner^N))^{K \circ P} \\
 &\leftrightarrow (A^M)^{K \circ P} \\
 &\leftrightarrow A^{K \circ P \circ \mathcal{E} \circ M} \\
 &\leftrightarrow A^{Q_0} \\
 &\leftrightarrow A^Q
 \end{aligned}$$

□

We have the following corollary:

Corollary 3.15. *Suppose U and V are mutually interpretable Vaught theories. Suppose further that V is an Enayat theory. Then, U is an Enayat theory.*

Open Question 3.16. Are there any other (interesting) relations between theories that preserve Enayatness? ○

In case we consider sequential theories, Theorem 3.15 has an important consequence. We remind the reader that every finitely axiomatised sequential A is mutually interpretable with $S_2^1 + \text{con}_{\rho(A)}(A)$. So, our question about examples of sequential Enayat theories reduces to the question whether S_2^1 plus a true Π_1^0 -sentence can be Enayat.

We can strengthen our question as follows.

Conjecture 3.17. Suppose A is finitely axiomatised and consistent and sequential. Let $N : S_2^1 \triangleleft A$. Then, there is no extension of S_2^1 that is mutually interpretable with $A + \text{TB}_N^-$.

It is a well known open question whether every sequential theory is mutually interpretable with an extension-in-the-same-language of S_2^1 . Our conjecture provides a possible example to illustrate a negative answer to this question. ○

We can view the preservation over mutual interpretability a bit more abstractly in the case of Vaught theories. Suppose we work with the degrees of interpretability of Vaught theories. In this case the Tarski functor (based on) $\mathfrak{T}(U) := U + \text{TB}_N^-(U)$, where N is some $N : R \triangleleft U$, makes sense, since we have already shown the independence of N .

We check that \mathfrak{T} is indeed a functor. Suppose $K : U \triangleleft V$. Then, we can extend τ_K to, say, $\tilde{\tau}_K$ as follows. Suppose $N : R \triangleleft U$. We choose \mathbb{T} over V w.r.t. the

NK -numerals. We extend τ_K to τ_K^* by setting $\mathbb{T}_{\tau_K^*}(\vec{x}) := \leftrightarrow \mathbb{T}(\tilde{\tau}_K(\vec{x}))$, where $\tilde{\tau}_K$ is the formalisation of τ_K .

We note that an Enayat Vaught theory is precisely a \mathfrak{T} -algebra. So, it is immediate that being an Enayat theory is preserved under mutual interpretability which is after all the isomorphism of our category.

Perhaps it is possible to make \mathfrak{T} work for a better category, but I did not explore this.

4 A consistent, finitely axiomatised Enayat Theory

Is there a finitely axiomatised theory with the N -Enayat property for appropriate N ? If we ask the question in this generality without further constraints on the admissible theories, there is actually a positive example. The example does depend on what we accept as a Gödel numbering. We discuss the issues here below.

We give an example of a finitely axiomatised theory that is not Enayat for one interpretation (and for any Gödel numbering) and that is Enayat for another interpretation for a special choice of the Gödel numbering.

We consider the theory $W := \text{Th}_{0,S,<}(\mathbb{N})$ of $0, <$ and S in the natural numbers. See [5, Section 3.2] for a careful exposition of this theory. The theory W is a finitely axiomatisable complete theory, to wit, the theory of a discrete linear ordering with initial and without final point. Every definable set of numbers in the language of W over \mathbb{N} is either finite or cofinite. Moreover, inspection of the quantifier elimination shows that the theory has a multi-exponential decision algorithm.⁴

Suppose ι is the direct one-dimensional translation of Succ_0 in W that sends 0 to 0 and xSy to $Sx = y$. Then, clearly, the interpretation K_ι of Succ_0 in W based on ι cannot be Enayat for any Gödel numbering since the set of truths is infinite and co-infinite. We cannot get around this example by tweaking the Gödel numbering.

Let $W' := \text{Th}_{0,S,<,E}(\mathbb{N})$ be the theory of $0, <, S$ and E , for *even*, in the natural numbers. Let j be the following two-dimensional translation of the language of W' in the language of W .

- $\delta_j(x, y) := (y = x \vee y = x + 1)$,
- $Z_j(x, y) := (x = 0 \wedge y = 0)$,
- $(x, y)S_j(x', y') := ((y = x \wedge x' = x \wedge y' = x + 1) \vee (y = x + 1 \wedge x' = x + 1 \wedge y' = x + 1))$,

⁴One further amazing property of $\text{Th}_{0,S,<}(\mathbb{N})$ is the fact that it is a finitely axiomatisable theory that proves full induction.

- $(x, y) <_j (x', y') := ((x' = x \wedge y < y') \vee x < x')$.
- $E_j(x, y) := (x = y)$.

Clearly, this yields an interpretation K_j of W' in W based on j . We note that it follows that W' is multi-exponentially decidable.

Let ν be a standard Gödel numbering for the language of W . We define $\nu^*(A) := 2\nu(A)$ if A is true in \mathbb{N} and $\nu^*(A) = 2\nu(A) + 1$ if A is false. Evidently, ν^* is a multi-exponential Gödel numbering. Let κ translate \mathbb{T} to \mathbb{E} where κ is the identical translation on the vocabulary of W . Clearly, $K_\kappa : (W + \text{TB}_{\text{ID}}^{-\nu^*}) \triangleleft W'$. Hence, $(K_\kappa \circ K_j) : (W + \text{TB}_{\text{ID}}^{-\nu^*}) \triangleleft W$.

Let N be the interpretation of W in W based on j restricted to the language without \mathbb{E} . We have $\mathbb{T} \blacktriangleright_W \text{TB}_N^{-\nu^*}$, showing that Tarski's Theorem on the undefinability of truth fails in our example for a specific choice of Gödel numbering and a specific choice of the numbers. Of course, there is nothing remarkable about this failure, since we do not have the Fixed Point Lemma in this context.

We note that we can do the same trick for, e.g., Presburger Arithmetic. However, Presburger is not finitely axiomatisable.

The reader may object that our Gödel numbering ν^* is contrived, unnatural and an ignoble hack. However, it seems very difficult to exclude it on principled reasons. One may want to demand that Gödel numberings are p-time. However, many of the classical Gödel numberings were exponential or even multi-exponential. This is witnessed by, e.g., the Gödel numbering in Feferman's celebrated arithmetisation paper [6].

Open Question 4.1. Is there an example of a finitely axiomatised theory A with the N -Enayat property for some $N : \text{Succ}_0 \triangleleft A$, when we demand that the Gödel numbering is p-time computable? ○

Open Question 4.2. Is there an example of a finitely axiomatised theory A such that we have the Enayat property for all $N : \text{Succ}_0 \triangleleft A$? ○

5 Neighbours

In this section, we discuss uniform variants of TB^- . We will see that for the uniform variants we have a clear negative answer —quite unlike the stubborn purely sentential case of TB^- .

Remark 5.1. When writing this paper I discovered that much more can be said about uniform biconditionals and Vaught theories. I postpone this to a subsequent paper. ○

We fix a theory U with an interpretation $N : U \triangleright \text{Succ}_0$. In order to avoid heavy and sometimes misleading notations, we assume N to be one-dimensional. Nothing depends on this however.

5.1 Satisfaction

We strengthen TB_N^- to a uniform principle $\text{USB}_{1,N}^-$ in the following way.

$$\text{USB}_{1,N}^- \forall x (\text{sat}(x, \ulcorner A(v) \urcorner) \leftrightarrow A(x)).$$

Here A is a U -formula with at most one free variable v and sat is a new binary predicate. Note that this definition is meaningful also in case our theory is not Vaught.

In case U is Vaught, we also have the following seemingly stronger principle.

$$\text{USB}_N^- \forall a (\text{sat}(a, \ulcorner A(v_0, \dots, v_{n-1}) \urcorner) \leftrightarrow A(a(\ulcorner v_0 \urcorner), \dots, a(\ulcorner v_{n-1} \urcorner))).$$

Here a ranges over assignments, i.e., partial functions from a finite set of variables to domain objects. If v_i is not in the domain, we set value of the variable to some default value x^* . Regrettably, in the general case, x^* must be a parameter, since there need not be definable elements in the ambient theory U .

Theorem 5.2. *Suppose U is a Vaught theory and $N : U \triangleright \text{R}$. Then, we have $\text{USB}_{1,N}^- \blacktriangleright_U \text{USB}_N^-$.*

Proof. If we start with $U + \text{USB}_N^-$, we can use Fujimoto translation τ with:

- $\text{sat}^\tau(x, y) := \text{sat}(\{\ulcorner v \urcorner, x\}, y)$.

We note that we pretended that we have functionality. In reality, we should have said that there is a representative u of $\ulcorner v \urcorner$, a representative w of the pair $\langle u, x \rangle$, a representative z of the set $\{w\}$, such that $\text{sat}(z, y)$. Moreover, what counts as a representative of $\langle u, x \rangle$, should also be further spelled out.

Suppose we start with $U + \text{USB}_{1,N}^-$. We define a recursive function F that sends a code of a formula $A(v_0, \dots, v_{n-1})$ to a code of $A(v(\ulcorner v_0 \urcorner), \dots, v(\ulcorner v_{n-1} \urcorner))$. Here of course the functions should be unraveled to their relational representations. *Par abus de langage*, we use F also for the arithmetisation of F in R . We now use the following Fujimoto translation ν :

- $\text{sat}^\nu(a, y) := \text{sat}(a, F^N(y))$.

Again, unraveling is needed to give the formula its correct form:

$$\exists z \in \delta_N (F^N(y, z) \wedge \text{sat}(a, z)). \quad \square$$

Theorem 5.3. *If U is Vaught and N, N' are interpretations in U of Succ_0 (or, if you wish, \mathbf{R}), then $\text{USB}_{1,N}^-$ and $\text{USB}_{1,N'}^-$ and USB_N^- and $\text{USB}_{N'}^-$ are mutually Fujimoto interpretable over U .*

The proof is entirely analogous to the proof of Theorem 3.4. We also have:

Theorem 5.4. $\top \blacktriangleright_{\text{loc},U} \text{USB}_{1,N}^-$ and $\top \blacktriangleright_{\text{loc},U} \text{USB}_N^-$.

Proof. We treat the case of $\text{USB}_{1,N}^-$. We interpret the axioms

$$A_0(y) \leftrightarrow \text{sat}(y, \ulcorner A_0 \urcorner), \dots, A_{n-1}(y) \leftrightarrow \text{sat}(y, \ulcorner A_{n-1} \urcorner),$$

by defining: $\text{sat}(y, x) := \bigvee_{k < n} (\ulcorner \widetilde{A}_k \urcorner(x) \wedge A_k(y))$. □

Here is a basic insight.

Theorem 5.5. *Suppose $N : U \triangleright \text{Succ}_0$. Then, $\mathcal{U}(U) \triangleright (U + \text{USB}_N^-)$, and, similarly for $\text{USB}_{1,N}^-$.*

Proof sketch. In $\mathcal{U}(U)$ we can build a Henkin interpretation H of U . (See [24].) This Henkin interpretation comes with a satisfaction predicate \mathbf{H} that works on a $\mathcal{U}(U)$ -cut I . Since $\mathcal{U}(U)$ is sequential, there is a definable isomorphism \mathfrak{F} between a cut of I and a cut of NH . We take $\text{sat}(x, y) := \exists x' \exists y' (x' \mathfrak{F} x \wedge y' \mathfrak{F} y \wedge \mathbf{H}(x', y'))$. □

We discuss two alternative forms of USB_N^- . Let us write $\text{comm}(\text{sat}, x)$ for: for all formulas $\leq x$, the predicate sat satisfies the commutation conditions (w.r.t. the signature of U). Suppose U is a Vaught theory and that $N : U \triangleright \mathbf{R}$.

The principle $\text{Comm}_{0,N}^-$ is given by $\text{comm}(\text{sat}, \underline{n}^N)$, for $n \in \omega$.

As is well known the theory \mathbf{R} interprets an extension, say \mathbf{R}^+ , which verifies that \leq is a linear ordering. See [22]. Suppose U is a Vaught theory and that $N : U \triangleright \mathbf{R}^+$.

The principle $\text{Comm}_{1,N}^-$ is given by $\forall x \in \mathfrak{J} \text{comm}(\text{sat}, x)$, $\forall x \in \mathfrak{J} \forall y \leq x \ y \in \mathfrak{J}$, $n \in \mathfrak{J}$, for any $n \in \omega$. Here \mathfrak{J} is a new unary predicate.

We assume that our Gödel coding is monotonic in the sense the the code of a subformula of A is less than the code of A itself.

Theorem 5.6. *Suppose U is a Vaught theory. We have:*

a. *If $N : U \triangleright \mathbf{R}$, then USB_N^- and $\text{Comm}_{0,N}^-$ are interderivable over U .*

b. *If $N : U \triangleright \mathbf{R}^+$, then $\text{Comm}_{1,N}^- \vdash_U \text{Comm}_{0,N}^-$ and $\text{Comm}_{0,N}^- \blacktriangleright_U \text{Comm}_{1,N}^-$.*

Proof. *Ad (a):* The inference from $\text{Comm}_{0,N}^-$ to USB_N is obvious. We treat the case of existential quantification for the other direction. We reason in $U + \text{USB}_N$. We write \check{a} for $a[\ulcorner v_i \urcorner : v_i]$, the result of resetting a on $\ulcorner v_i \urcorner$ to v_i .

$$\begin{aligned} \text{sat}(a, \exists v_i A(v_0, \dots, v_i, \dots, v_{n-1})) &\leftrightarrow \exists v_i A(a(\ulcorner v_0 \urcorner), \dots, v_i, \dots, a(\ulcorner v_{n-1} \urcorner)) \\ &\leftrightarrow \exists v_i A(\check{a}(\ulcorner v_0 \urcorner), \dots, \check{a}(\ulcorner v_i \urcorner), \dots, \check{a}(\ulcorner v_{n-1} \urcorner)) \\ &\leftrightarrow \exists v_i \text{sat}(\check{a}, \ulcorner A(v_0, \dots, v_i, \dots, v_{n-1}) \urcorner) \end{aligned}$$

Ad (b): This is mostly trivial. We interpret $\mathfrak{J}(x)$ as $\text{comm}(\text{sat}, x)$. □

The difference between $\text{Comm}_{0,N}^-$ and $\text{Comm}_{1,N}^-$ may seem somewhat trifling, but the usefulness of $\text{Comm}_{1,N}^-$ lies in the fact that there may be other more interesting interpretations of \mathfrak{J} .

We note that $\text{Comm}_{0,N}^-$ is a restricted axiomatisation of USB_N^- over U . This means that all axioms of $\text{Comm}_{0,N}^-$ have depth-of-quantifier-alternations complexity below a fixed n . This suggests the following question.

Open Question 5.7. Does TB_N^- have a restricted axiomatisation over U ? ○

5.2 Truth

There is also the alternative option of defining a truth principle.

$$\text{UTB}_N^- \quad \forall \vec{x} \in \delta_N (\ulcorner \text{T} \ulcorner A \vec{x} \urcorner \urcorner \leftrightarrow A \vec{x}).$$

We note that to make sense of this we must stipulate that (i) N is an interpretation of S_2^1 and that (ii) we use efficient numerals, since, for ordinary numerals, the mapping from x to the numeral of x is exponential.

It is not clear to me that, in this case, we have an analogue of Theorem 5.3, i.e., that $\text{UTB}_N^- \blacktriangleright_U \text{UTB}_{N'}^-$. However, as we will see, it is immediate from Theorem 5.12, that these theories are mutually interpretable in the sequential case.

Here is a first small insight.

Theorem 5.8. *Let U be sequential and $N : U \triangleright \text{S}_2^1$. Then, $\text{USB}_N^- \blacktriangleright_U \text{UTB}_N^-$.*

Proof sketch. We have to define the predicate T from sat . Consider a number a . In case a is not an N -code of a U -sentence we make $\text{T}(a)$ false. Suppose U is an N -code of a U -sentence. Now we have to analyze a as being a substitution instance of a U -formula b with numerals. There are two obstacles:

- There are not really numerals in b , since we work with a relational signature. So, we have to reverse the term-unwinding translation to obtain the relevant numerals. To do this we need a precise analysis of term-unwinding. Also, we should take care that the reverse algorithm is p-time.
- In the $A(\dot{x})$ of UTB_N^- there could be already numerals in the standard context $A(\cdot)$. However U having just a as input cannot know which numerals are the numerals to replace by variables. Fortunately, it is sufficient to remove numerals maximally. The case where there are some numerals in $A(\cdot)$ can be recovered by substituting some numerals in the result of maximal analysis.

Given that we analyzed a as substitution instance of b where we replace numeral c by variable v , we can compute a corresponding assignment f that sends v to the value of c . Now we define $\text{T}(a)$ by $\text{sat}(f, b)$. □

We do not generally have that $\text{TB}_N^- \triangleright_U \text{UTB}_N^-$.

Theorem 5.9. $\text{TB}_{\text{ID}}^- \not\triangleright_{\text{EA}} \text{UTB}_{\text{ID}}^-$.

Proof. Suppose $\text{TB}_{\text{ID}}^- \triangleright_{\text{EA}} \text{UTB}_{\text{ID}}^-$. Then, it follows that $\text{TB}_{\text{ID}} \triangleright_{\text{PA}} \text{UTB}_{\text{ID}}$. Here the lack of the superscript minus means that we extend induction to the full language. However, there is a model of $\text{PA} + \text{TB}_{\text{ID}}^-$ that is not recursively saturated, where all models of $\text{PA} + \text{UTB}_{\text{ID}}^-$ are recursively saturated. See [3]. □

Open Question 5.10. The argument above works for all subtheories of PA that extend R and more, but still it is rather special. Can we improve it to show that the result holds for all sequential theories? ○

5.3 \mathcal{U}

We now connect uniform biconditionals with the \mathcal{U} -functor.

Theorem 5.11. *Suppose U is sequential and $N : U \triangleright \mathbf{S}_2^1$. Then $(U + \text{UTB}_N^-) \triangleright \mathcal{U}(U)$.*

Proof sketch. In $U + \text{UTB}_N^-$, we can define the intersection \mathcal{J} of all virtual classes $\{x \in \delta_N \mid \text{T}(a(\dot{x}))\}$ that are N -cuts. One can show that \mathcal{J} is an N -cut contained in all U -definable cuts. Thus, in \mathcal{J} , we have all restricted consistency statements of U . □

We partially summarise the above in the following theorem.

Theorem 5.12. *Suppose U is sequential and $N : U \triangleright S_2^1$. Then, following theories are mutually interpretable: $\mathcal{U}(U)$, $U + \text{USB}_N^-$, $U + \text{USB}_{0,N}^-$, $U + \text{Comm}_{0,N}^-$, $U + \text{Comm}_{1,N}^-$, $U + \text{UTB}_N^-$.*

As a consequence, a sequential theory is uniformly Enayat, in any of the possible senses, iff it is reflexive.

Proof. We have:

- $\mathcal{U}(U) \triangleright (U + \text{USB}_N^-)$, by Theorem 5.5;
- $\text{USB}_N^- \bowtie_U \text{USB}_{0,N}^-$, by Theorem 5.2;
- $U + \text{Comm}_{0,N}^-$, $U + \text{Comm}_{1,N}^-$ and $U + \text{USB}_N^-$ are mutually interpretable by Theorem 5.6;
- $(U + \text{USB}_N^-) \triangleright (U + \text{UTB}_N^-)$, by Theorem 5.8;
- $(U + \text{UTB}_N^-) \triangleright \mathcal{U}(U)$, by Theorem 5.11.

The last step completes the circle. □

We note that theories like PRA and PA and ZF are reflexive and, hence, sequential uniform Enayat theories.

We also note that the characterisation of $\mathcal{U}(U)$ as $U + \text{USB}_N^-$ (modulo mutual interpretability) has the advantage of having the Gödel numbering as conventional element, but not the proof system, the arithmetisation of the proof system and the like.

The following corollary is immediate.

Corollary 5.13. *No finitely axiomatised, consistent uniform sequential theory is uniformly Enayat.*

So, for example, none of S_2^1 , EA, ACA_0 , GB is uniformly Enayat.

Open Question 5.14. Can we prove the non-existence of a finitely axiomatised consistent sequential uniform Enayat theory without a detour over the second incompleteness theorem? ○

More can be said about USB in the context of Vaught theories. We hope to do that in a subsequent paper. A salient open question is as follows.

Open Question 5.15. Is there a finitely axiomatised Vaught theory that is uniformly Enayat? Here uniformity is explicated using USB^- . ○

6 Finite Extensions of TB^-

In this section, we formulate two conjectures in the environment of Conjecture 3.13.

Consider a theory U of signature Θ_0 . Let Θ_1 be Θ_0 extended with a unary predicate T and let Θ_2 be binary predicate symbol sat . The variables α, β, \dots range over sentences of Θ_2 . We take as the default that a theory has as signature the minimal signature demanded by its axioms.

In this vocabulary, we can state Tarski's theorem on the undefinability of truth as follows.

Theorem 6.1 (Tarski). *Suppose U is consistent and $N : U \triangleright R$. Then, we have $T \not\blacktriangleright_U TB_N^-$.*

From Tarski's work on truth we also know the following.

Theorem 6.2 (Tarski). *Suppose U is a Vaught theory and $N : U \triangleright R$. Then, $(\forall x \in \delta_N \text{comm}(\text{sat}, x)) \blacktriangleright_U TB_N^-$.*

Of course, this is just a watered down version of Theorem 5.6(a).

In the next theorem, we show that there is no 'best', in the sense of 'weakest', finite extension of a Vaught theory U in an extended signature that Fujimoto interprets TB_N^- over U . So, certainly the commutation conditions, as articulated by $\forall x \in \delta_N \text{comm}(\text{sat}, x)$, are not 'best'.

Theorem 6.3. *Suppose U is a consistent Vaught theory and $N : U \triangleright R$. Suppose $\alpha \blacktriangleright_U TB_N^-$. Then, there is a β with $\beta \blacktriangleright_U \text{USB}_N^-$, but $\beta \not\blacktriangleright_U \alpha$.*

It follows that for $\gamma := (\alpha \vee \beta)$, we have $TB_N^- \triangleleft_U \gamma \triangleleft_U \alpha$. Moreover, in case $\alpha \blacktriangleright_U \text{USB}_N^-$, we find $\text{USB}_N^- \triangleleft_U \gamma \triangleleft_U \alpha$.

The proof is a variation of proof of Theorem 4.1, case (A), of [11].

Proof. Suppose A is of the form $\exists x \in \delta_N A_0(x)$. We write $C(A)$, for 'if there is no witness x of A such that $x \leq y$, then $\text{comm}(\text{sat}, y)$ '. In other words, $C(A)$ is $\neg((\exists y \in \delta_N \neg \text{comm}(\text{sat}, y)) \leq A)$.

We write $\delta \blacktriangleright_U \eta$ as $\exists p \exists \tau \text{proof}_{U+\delta}(p, \eta^\tau)$. By the Fixed Point Lemma, we find B such that $R \vdash B \leftrightarrow C(B^N) \blacktriangleright_U \alpha$. We take $\beta := C(B^N)$.

Suppose $\beta \blacktriangleright_U \alpha$. Then, we find that B is true and, hence, B has a standard witness inside U, N . However, finitely many commutation conditions are Fujimoto-interpretable in U , by the combination of Theorems 5.4 and 5.6(a). Hence, we have $T \blacktriangleright_U \beta \blacktriangleright_U \alpha \blacktriangleright_U TB_N^-$, contradicting Tarski's Theorem on the undefinability of truth. So, $\beta \not\blacktriangleright_U \alpha$.

It follows that B^N has no standard witness inside $U + \beta$. So, $U + \beta$ provides the commutation conditions at all standardly finite levels, i.e., $\beta \blacktriangleright_U \text{Comm}_{0,N}$. Hence, as desired, $\beta \blacktriangleright_U \text{USB}_N^-$. □

Remark 6.4. We note that, in the proof of Theorem 6.3, the sentence B is a fixed point of a formula of essentially the form ‘provable ...’. This might convey the impression that we have Gödelean self-reference here. However, the internal $\mathbb{C}(B)$ has Rosser-form. Thus, it seems very improbable that B is uniquely determined by the equation even if our numbers satisfy \mathbb{S}_2^1 . For the same reason, an ‘explicit’ solution for B seems improbable. ○

Conjecture 6.5. Let U be Vaught. and let $N : \mathbb{R} \triangleleft U$. Suppose $\alpha \blacktriangleright_U \text{TB}_N^-$. Then, $\alpha \blacktriangleright_U \text{USB}_N^-$.

We can put further demands on U and N : that U be sequential, finitely axiomatised, etcetera; that N is an interpretation of \mathbb{S}_2^1 , etcetera. Also, when N interprets \mathbb{S}_2^1 , we may replace USB_N^- by UTB_N^- . ○

We note that the Tarski commutation conditions, are an example of such an α . We note that the construction in the proof of Theorem 6.3 does not immediately help to refute the conjecture. However, it cannot be excluded that some variant of the argument does refute the conjecture.

Remark 6.6. Conjecture 6.5 can be connected to the Davidsonian idea that we need compositionality to obtain a finite axiomatisation of the Tarski Biconditionals. ○

Remark 6.7. The only paper I could find asking a question in the neighbourhood of Conjecture 6.5 is [7]. However, Fine and McCarthy work with what they call *segregated languages*. Their format does not seem to fit ours. Moreover, they do not work with Fujimoto interpretability. It would be well worth exploring what of their work can be adapted to our context. ○

Remark 6.8. Conjecture 6.5 suggests the concept of Fujimoto preservativity \triangleright . We define:

- $\Gamma \triangleright_U \Delta$ iff, for all α such that $\alpha \blacktriangleright_U \Gamma$, we have $\alpha \blacktriangleright_U \Delta$.

Here, the most elegant approach is to take α , Γ and Δ to be in the language of U expanded with a binary predicate. We also want to apply the notion if one of the expansions is to unary, but we can choose some standard way to let a binary predicate pose as a unary one, e.g., we might take Rxx for Px . Alternatively, in a context where we have pairing, we might only consider expansions

with a unary predicate. A final alternative is to assume that intended signatures for α , Γ and Δ are given in the context.

Now Conjecture 6.5 becomes: we have $\text{TB}_N^- \supset_U \text{UTB}_N^-$.

We note that $\Gamma \blacktriangleright_U \Delta$ implies $\Gamma \supset_U \Delta$. Moreover, \supset_U is reflexive and transitive. ○

Here is our second conjecture.

Conjecture 6.9. Suppose A is a finitely axiomatised Vaught in signature Θ_0 . Let $N : A \triangleright \mathbf{R}$. Suppose further that $\top \triangleright_A \text{TB}_N^-$. Then, there is a β such that we have $\top \triangleright_A \beta \blacktriangleright_A \text{TB}_N^-$.

More generally, we may conjecture the following. Suppose A is a finitely axiomatised Vaught theory and $\top \triangleright_A V$. Then, there is a B such that $\top \triangleright_A B \blacktriangleright_A V$. ○

We have:

Theorem 6.10. *The truth of Conjectures 6.5 and 6.9 combined implies the truth of Conjecture 3.13.*

Proof. Suppose $\top \triangleright_A \text{TB}_N^-$. Let β be as promised in Conjecture 6.9, so we have $\top \triangleright_A \beta \blacktriangleright_A \text{TB}_N^-$. It follows, by Conjecture 6.5, that $\top \triangleright_A \beta \blacktriangleright_A \text{USB}_N^-$. However, we have already seen that, in the sequential case, $\top \not\blacktriangleright_A \text{USB}_N^-$. □

7 A Provability Predicate

In the present section we follow the Saccheri strategy. We assume that we have a finitely axiomatised, sequential A that is Enayat. We pretend that it is consistent and explore it as an interesting new world.

7.1 Preliminary Considerations

The first thing that one thinks of is the Liar Paradox for the TB^- -truth predicate. More precisely: suppose we have an interpretation $K : A \triangleright (A + \text{TB}_{N_0}^-)$, where $N_0 : A \triangleright \mathbf{S}_2^1$. Let $\mathfrak{F} := \mathfrak{F}_{N, NK}$ be the usual isomorphism between initial cuts of N_0 and N_0K . We define $\mathbf{K}(\vec{x}) := \exists \vec{y} (\vec{x} \mathfrak{F} \vec{y} \wedge \top^K(\vec{y}))$. We write $\mathbf{K}B$ for $\mathbf{K}(\ulcorner B \urcorner^{N_0})$. By the Gödel Fixed Point Lemma, we find L such that $A \vdash L \leftrightarrow \neg \mathbf{K}L$. We can see that the truth value of L has to alternate if we travel inside K , K^2 , K^3 , \dots . Also, we

have $(\text{ID}_A \langle L \rangle K) : A \triangleright (A + L)$ and $(K \langle L \rangle \text{ID}_A) : A \triangleright (A + \neg L)$. However, nothing paradoxical follows.⁵

We can see that without some further idea nothing paradoxical *can* follow, since if U is e.g. $\text{EA} + \{\text{con}^n(\text{EA}) \mid n \in \omega\}$ we do have that (i) U is a restricted theory in the sense that the complexity of all its axioms is bounded by a fixed n and (ii) U is reflexive for the identical interpretation and, so, $U \triangleright (U + \text{TB}_{\text{ID}_U}^-)$.

Thus, we need to add an ingredient that essentially uses the fact the A is finitely axiomatised rather than just restricted. In this section, this ingredient is the use of a new provability predicate for A , the good properties of which are based on A 's finite axiomatisability. In Section 8, the ingredient is the use of a conjecture that is supposed to hold only for finitely axiomatisable sequential theories.

7.2 What we fix

In this section, we consider a number of things as fixed:

- i. the theory A itself;
- ii. the interpretation \mathcal{S} that witnesses the sequentiality of A ; by cut-elimination, the proof of $\text{AS}^{\mathcal{S}}$ can be taken to have complexity $\max(\rho(A), \rho(\mathcal{S})) + \mathfrak{a}_0$; here \mathfrak{a}_0 is a constant for overhead.
- iii. the interpretation N_0 of \mathbf{S}_2^1 ; we note that, by cut-elimination, the proof of $(\mathbf{S}_2^1)^{N_0}$ can be taken to have complexity $\max(\rho(A), \rho(N_0)) + \mathfrak{a}_1$; we note that, since there is a known interpretation of \mathbf{S}_2^1 in AS , there is an N_0 of complexity $\rho(\mathcal{S}) + \mathfrak{a}_2$;
- iv. the interpretation K of $A + \text{TB}_{N_0}^-$ in A .

We note that the complexity of $\text{K}(\underline{n})$ is $\rho(K) + \rho(N_0) + \mathfrak{a}_3$, where \mathfrak{a}_3 is a constant for overhead.

7.3 The Provability Predicate

Let $n \geq \mathfrak{a}^* := \max(\rho(A), \rho(K) + \rho(N_0)) + \mathfrak{a}$, where \mathfrak{a} is suitable constant number that is needed for the overhead in our reasoning. We note that in the context of our reasoning \mathfrak{a}^* functions as a constant since we treat A , K and N_0 as fixed.

Our new provability predicate is $\Delta_{A,(n)}B := \Box_{A,n}KB$. We note that $n \geq \mathfrak{a}^*$ is precisely what is needed to make $\Delta_{A,(n)}$ a meaningful notion. We use subscript (n) rather than n to remind the reader that the subscript is *not* a constraint on the sentences for which our predicate is meaningful.

⁵We pick up the idea of using an analogue of a semantical paradox in Section 8.

We will omit the subscript A since A is fixed in this section and we will have to work with a whole circus of sub- and superscripts anyway.

We remind the reader of the Löb conditions. For a brief moment, we view Δ as an abstract operator.

$$\text{L1. } \vdash B \Rightarrow \vdash \Delta B,$$

$$\text{L2. } \vdash (\Delta B \wedge \Delta(B \rightarrow C)) \rightarrow \Delta C,$$

$$\text{L3. } \vdash \Delta B \rightarrow \Delta\Delta B,$$

$$\text{L4. } \vdash \Delta(\Delta B \rightarrow B) \rightarrow \Delta B.$$

We will also consider the condition C, to wit:

$$\text{C. } \vdash \Delta B \rightarrow B.$$

In the next subsection, we discuss variants of L1 and L2 for the predicate $\Delta_{(n)}^N$, where N is a cut of N_0 .

7.4 The first two Löb Conditions

Let N be an A -definable, A -verifiable cut of N_0 . We note that N has the same numerals as N_0 . Let $n \geq \mathfrak{a}^*$.

Theorem 7.1. *We have:*

A. *if $A \vdash B$, then $S_2^1 \vdash \Delta_{(n)} B$, and, hence $A \vdash \Delta_{(n)}^N B$. Thus, this gives us L1 in the form:*

$$A \vdash B \Rightarrow A \vdash \Delta_{(n)}^N B.$$

B. *We also have a second form that is a better ‘externalisation’ of L3.*

$$A \vdash_n KB \Rightarrow A \vdash_n K\Delta_{(n)}^N B.$$

Proof. Ad (A): Suppose $A \vdash B$. Then, $A \vdash B^K$ and, hence, $A \vdash KB$. By cut-elimination, we have $A \vdash_n KB$. By Σ_1 -completeness, in the meta-theory, we find $S_2^1 \vdash \square_n KB$, i.o.w., $S_2^1 \vdash \Delta_{(n)} B$.

Ad (B): Suppose $A \vdash_n KB$. Then, $S_2^1 \vdash \Delta_{(n)} B$. It follows that $A \vdash \Delta_{(n)}^{NK} B$ and, hence, $A \vdash K\Delta_{(n)}^N B$. By cut-elimination, we find $A \vdash_n K\Delta_{(n)}^N B$. \square

As the reader can see, n is, in this theorem, not constrained by N . The reason is that, since N is initial in N_0 , the N -numerals simply are the N_0 -numerals.

We note that in both proofs we used cut-elimination. Thus, the proofs can be executed in meta-theory \mathbf{EA}^+ , i.e., $\mathbf{I}\Delta_0 + \text{supexp}$. So, we have:

- $\mathbf{EA}^+ \vdash \Box B \rightarrow \Box \Delta_{(n)}^N B$,
- $\mathbf{EA}^+ \vdash \Delta_{(n)} B \rightarrow \Delta_{(n)} \Delta_{(n)}^N B$.

However, on closer inspection, we have a much better result. We suppose that A , n , N and K are externally given. Then, the first form only requires \mathbf{S}_2^1 and the second form only requires $\mathbf{EA} := \mathbf{I}\Delta_0 + \text{exp}$. We will discuss this in detail when we consider proofs of **L3** in the next subsection.

Theorem 7.2. *We have $\mathbf{S}_2^1 \vdash (\Delta_{(n)} B \wedge \Delta_{(n)}(B \rightarrow C)) \rightarrow \Delta_{(n)} C$. So, a fortiori, the theory A satisfies **L2** for $\Delta_{(n)}^N$.*

Proof. We have:

$$\begin{aligned} A \vdash \mathbf{K}(B \rightarrow C) &\rightarrow (B \rightarrow C)^K \\ &\rightarrow (B^K \rightarrow C^K) \\ &\rightarrow (\mathbf{K}B \rightarrow \mathbf{K}C). \end{aligned}$$

By cut-elimination, $A \vdash_n \mathbf{K}(B \rightarrow C) \rightarrow (\mathbf{K}B \rightarrow \mathbf{K}C)$. It follows by Σ_1 -completeness that $\mathbf{S}_2^1 \vdash \Box_n(\mathbf{K}(B \rightarrow C) \rightarrow (\mathbf{K}B \rightarrow \mathbf{K}C))$. By **L2** for \Box_n , we find:

$$\mathbf{S}_2^1 \vdash \Box_n \mathbf{K}(B \rightarrow C) \rightarrow (\Box_n \mathbf{K}B \rightarrow \Box_n \mathbf{K}C).$$

In other words, $\mathbf{S}_2^1 \vdash (\Delta_{(n)} B \wedge \Delta_{(n)}(B \rightarrow C)) \rightarrow \Delta_{(n)} C$. □

We note that there does not seem to be a way to prove the global version

$$\not\leq \not\leq \quad \mathbf{S}_2^1 \vdash \forall B, C ((\Delta_{(n)} B \wedge \Delta_{(n)}(B \rightarrow C)) \rightarrow \Delta_{(n)} C). \quad \not\leq \not\leq$$

The quantifier over sentences seems essentially external.

Remark 7.3. With the first Löb condition in hand, we can immediately prove the well-known properties of the Gödel sentences. Let N and n be as before. By the Gödel Fixed Point Lemma, we find $G_{N,n} := G$ such that $\mathbf{S}_2^1 \vdash G \leftrightarrow \neg \Delta_{(n)} G^N$. Suppose $A \vdash G^N$. Then, by **L1**, $A \vdash \Delta_{(n)}^N G^N$. On the other hand, by the Fixed Point Equation, $A \vdash \neg \Delta_{(n)}^N G^N$. It follows that $A \vdash \perp$.

We cannot similarly derive Löb's Rule as will be illustrated in Subsection 7.5. The derivation of the Rule does need some form of **L3**. ○

7.5 Guarded Reflection

We have the following theorem:

Theorem 7.4. *For any n , there is an A -definable, A -verifiable N_0 -cut \mathfrak{J}_n with ρ -complexity of order $\mathfrak{b}n + \max(\rho(\mathcal{S}), \rho(N_0)) + \mathfrak{c}$, such that, for all B with $\rho(B) \leq n$, we have $A \vdash \square_n^{\mathfrak{J}_n} B \rightarrow B$.*

See [25], for a careful treatment of this result (or, rather, a result of which this result is an immediate consequence). We now have a form of guarded reflection for Δ .

Theorem 7.5. *Let $n \geq \mathfrak{a}^*$. Then, we have $A \vdash \Delta_{(n)}^{\mathfrak{J}_n} B \rightarrow B^K$.*

As preparation for Remark 7.7, we formulate a well-know lemma that is due to Pudlák. See [13].

Lemma 7.6. *Suppose $N, N' : A \triangleright \mathcal{S}_2^1$. There is an N -cut $\mathfrak{C}_{N,N'}$ and a definable isomorphic embedding $\mathfrak{F}_{N,N'} : \mathfrak{C}_{N,N'} \rightarrow N'$. We have:*

$$\rho(\mathfrak{C}_{N,N'}) = \max(\rho(\mathcal{S}), \rho(N), \rho(N')) + \mathfrak{d} \text{ and } \rho(\mathfrak{F}_{N,N'}) = \max(\rho(\mathcal{S}), \rho(N), \rho(N')) + \mathfrak{c}.$$

We write ∇ for $\neg\Delta\neg$.

Remark 7.7. We have $A \vdash \nabla_{(n)}^{\mathfrak{J}_n} \top$. So, we immediately see that the Second Completeness Theorem fails. Hence, *a fortiori*, Löb's Rule fails.

If we did not have the guard K , we would have A 's inconsistency. Let N be a cut of \mathfrak{J}_n . Then, we have $A \vdash \Delta_{(n)}^N G_{N,n}^N \rightarrow G_{N,n}^N$. Hence, $A \vdash G_{N,n}^N$ and, so $A \vdash \Delta_{(n)}^N G_{N,n}^N$, which gives $A \vdash \neg G_{N,n}^N$. So $A \vdash \perp$.

One thing one could try, in order to get the effect of the above reasoning, is to get under the guard using Lemma 7.6. We have a brief look, to see why this idea fails. Suppose N is a cut both of \mathfrak{J}_n and $\mathfrak{C}_{N_0, N_0 K}$.

We first try $G_{N,n}$. We reason in A . Suppose $\Delta_{(n)}^N G_{N,n}^N$. Then, by guarded reflection, $G_{N,n}^{NK}$. But also $\Delta_{(n)}^{N_0 K} G_{N,n}^N$. Hence, $\neg G_{N,n}^{N_0 K}$. However, since NK is smaller than $N_0 K$, no contradiction materialises.

Next we try $G_{N_0, n}$. We reason in A . Suppose $\Delta_{(n)}^N G_{N_0, n}^{N_0}$. Then, by guarded reflection, $G_{N_0, n}^{N_0 K}$. But also $\Delta_{(n)}^{N_0 K} G_{N_0, n}^{N_0}$. Hence, $\neg G_{N_0, n}^{N_0 K}$. A contradiction. So, canceling the assumption, we find $\neg \Delta_{(n)}^N G_{N_0, n}^{N_0}$, i.e. $G_{N_0, n}^N$. Returning to the meta-language, we see that $A \vdash G_{N_0, n}^N$. Of course, this is still no contradiction.

The result of these two attempts is somewhat disappointing. However, we will see in Subsection 7.7 that a Gödel-style argument does give us some information about K . ○

7.6 The third and fourth Löb Condition

We write \blacksquare for provability in S_2^1 . We define $\text{itexp}(0, x) := x$ and $\text{itexp}(y + 1, x) := 2^{\text{itexp}(y, x)}$.

We start with a lemma.

Lemma 7.8. *Suppose $n \geq \max(\rho(A), \rho(B))$. We have $S_2^1 \vdash \Box B \rightarrow \blacksquare \Box_n B$.*

Proof. We use the version of cut-elimination from Buss' paper [2]. By formalising Buss' result, we have, S_2^1 -verifiably, that whenever p is a proof and whenever $\text{itexp}(\rho(p) + 2, p)$ exists, then we have a cut-free proof q with the same conclusion.

Secondly, we use an insight from Pudlák's paper [13], that,

$$S_2^1 \vdash \exists w 2^y = w \rightarrow \blacksquare \forall x \exists z \text{itexp}(y, x) = z.$$

We reason as follows inside S_2^1 . Suppose p is an A -proof of B . Then, $\rho(p)$ is a logarithmic number. So $\blacksquare \exists z \text{itexp}(\rho(p) + 2, p) = z$. We also find $\blacksquare \text{proof}_A(p, B)$. Hence, inside \blacksquare we have a cut-free proof q that witnesses $\Box_n B$. \square

Let N and N' be A -definable, A -verifiable cuts of N_0 and let $n \geq \mathfrak{a}^*$. We first internalise Theorem 7.1(A).

Theorem 7.9. *We have $S_2^1 \vdash \Box B \rightarrow \blacksquare \Delta_{(n)} B$, and, hence,*

$$S_2^1 \vdash \Box B \rightarrow \Box \Delta_{(n)}^N B \text{ and } A \vdash \Box^{N'} B \rightarrow \Box^{N'} \Delta_{(n)}^N B.$$

Proof. The proof is an internalisation and refinement of the proof of Theorem 7.1(A). We note that externally we have $A \vdash B^K \leftrightarrow KB$, and, thus, $S_2^1 \vdash \Box(B^K \leftrightarrow KB)$. It follows that $(\dagger) S_2^1 \vdash \Box B^K \leftrightarrow \Box KB$.

We reason in S_2^1 . Suppose $\Box B$. Then, since A is finitely axiomatised and interpretations give p-time transformations of proofs, we have $\Box B^K$. Hence, by (\dagger) , $\Box KB$. By Lemma 7.8, it follows that $\blacksquare \Box_n KB$, i.o.w., $\blacksquare \Delta_{(n)} B$. \square

We proceed with an internalisation of Theorem 7.1(B). Let N and N' be as before. Let $n \geq \mathfrak{a}^*$. We use a slightly more general formulation with an extra m for later use.

Theorem 7.10. *Suppose S is in $\exists \Sigma_1^b$ and $m \geq \rho(S) + \rho(K) + \rho(N) + \mathfrak{f}_0$, where \mathfrak{f}_0 is a constant for overhead. We have $S_2^1 \vdash S \rightarrow \Delta_{(m)} S^N$. Hence, it follows that $A \vdash S^{N'} \rightarrow \Delta_{(m)}^{N'} S^N$.*

Proof. We have $A \vdash S^{NK} \rightarrow KS^N$. So, by cut-elimination, $A \vdash_m S^{NK} \rightarrow KS^N$. Note that this makes sense only under our assumption on m . By Σ_1 -completeness, we find: $S_2^1 \vdash \Box_m(S^{NK} \rightarrow KS^N)$. Hence, (\dagger) $S_2^1 \vdash \Box_m S_n^{NK} \rightarrow \Delta_{(m)} S^N$.

We reason in S_2^1 . Suppose S . By $\exists\Sigma_1^b$ -completeness, we have $\Box_m S^{NK}$. By applying (\dagger) , we find $\Delta_{(m)} S^N$. \square

By specialising Theorem 7.10, we find:

Theorem 7.11. *Suppose $m \geq \rho(K) + \rho(N) + \mathfrak{f}$, where \mathfrak{f} is a constant for overhead. We have $S_2^1 \vdash \Delta_{(n)} B \rightarrow \Delta_{(m)} \Delta_{(n)}^N B$. Hence, we have $A \vdash \Delta_{(n)}^{N'} B \rightarrow \Delta_{(m)}^{N'} \Delta_{(n)}^N B$.*

Now, putting $m = n$, L4 follows in the usual way from L1,2,3 in combination with the Fixed Point Lemma.

Theorem 7.12. *Under the assumption that $n \geq \rho(K) + \rho(N) + \mathfrak{f}$, we have:*

$$S_2^1 \vdash \Delta_{(n)}(\Delta_{(n)}^N B \rightarrow B) \rightarrow \Delta_{(n)} B.$$

Hence, we have:

$$A \vdash \Delta_{(n)}^{N'}(\Delta_{(n)}^N B \rightarrow B) \rightarrow \Delta_{(n)}^{N'} B.$$

Assuming A to be consistent, we note that, for any n and N satisfying the assumption of Theorem 7.12, we cannot have guarded reflection for $\Delta_{(n)}^N$. Otherwise, we would have both Löb's Theorem and guarded reflection at the same time. But this is impossible since it would follow that $A \vdash \nabla_{N,(n)} \top$.

We can give an alternative form of L3 where we eliminate the lower bound on n at the cost of relativising the antecedent to a cut.

Theorem 7.13. *1. Consider any m . There is an S_2^1 -cut \mathfrak{J}_m , such that $\rho(\mathfrak{J}_m)$ is of order $\mathfrak{g}n + \mathfrak{h}$ and $S_2^1 \vdash \forall x \in \mathfrak{J}_m \exists w \text{itexp}(m, x) = w$.*

2. Suppose $m, n \geq \mathfrak{a}^$. Then, $S_2^1 \vdash \Delta_{(m)}^{\mathfrak{J}_{m+2}} B \rightarrow \Delta_{(n)} B$.*

Proof. Ad (1): The proof is essentially contained in [12] or [13]. Given any cut I , we consider the virtual class $\{x \mid 2^x \in I\}$. This class is downward closed and closed under successor. We shorten it to a cut J . Inspecting the construction, clearly, $\rho(J) = \rho(I) + \mathfrak{g}$. We obtain \mathfrak{J}_m by iterating the construction starting from the identical cut.

Ad (2): We reason in S_2^1 . Suppose $\Delta_{(m)}^{\mathfrak{J}_{m+2}} B$. This means $\Box_m^{\mathfrak{J}_{m+2}} KB$. Let the witnessing proof be p . Since $p \in \mathfrak{J}_{m+2}$, we find, by (1), that $\text{itexp}(m+2, p)$ exists. So, by the cut-elimination theorem from [2], we find $\Box_{\mathfrak{a}^*} KB$, so, *a fortiori*, $\Box_n KB$, i.e., $\Delta_{(n)} B$. \square

It is always good to have an alternative proof of a result. In the proof of Theorem 7.13, we used S_2^1 -formalisation of Buss' result of [2], a delicate result that involves many details. So, it improves our confidence to have a variant of Theorem 7.13(2), with a different proof.

We write \tilde{K} for $\neg K \neg$. Let $\mathfrak{H}_C : (S_2^1 + \diamond_{\mathfrak{a}^*} \tilde{K}C) \triangleright (A + \tilde{K}C)$ be the Henkin interpretation based on $\diamond_{\mathfrak{a}^*} \tilde{K}C$. See [24]. We note that \mathfrak{H}_C is uniform in C . We can view ' C ' as a variable. Let \mathfrak{D}_m be the common cut in S_2^1 of the identical interpretation and all the $\mathfrak{I}_m \mathfrak{H}_C$, for C such that $\diamond_{\mathfrak{a}^*} \tilde{K}C$. We note that the complexity of \mathfrak{H}_C is a small standard number independent of C . So, $\rho(\mathfrak{D}_m)$ is $\rho(\mathfrak{I}_m)$ plus some standard constant.

Theorem 7.14. *Suppose $m, n \geq \mathfrak{a}^*$. Then, $S_2^1 \vdash \Delta_{(m)}^{\mathfrak{D}_m} B \rightarrow \Delta_{(n)} B$.*

Proof. We reason in S_2^1 . Suppose $\diamond_n \tilde{K}C$. Then, *a fortiori*, we have $\diamond_{\mathfrak{a}^*} \tilde{K}C$ and, hence, $(A + \tilde{K}C)^{\mathfrak{H}_C}$. So, we have, by reasoning inside \mathfrak{H}_C , that $\diamond_{(m)}^{\mathfrak{I}_m \mathfrak{H}_C} \tilde{K}C$. Hence, by the definition of \mathfrak{D}_m , we find $\diamond_{(m)}^{\mathfrak{D}_m} \tilde{K}C$.

We return to the meta-language. We note that $A \vdash K \neg B \leftrightarrow \neg KB$, and, hence $A \vdash_n K \neg B \leftrightarrow \neg KB$ and $A \vdash_m K \neg B \leftrightarrow \neg KB$. So, $S_2^1 \vdash \Box_n K \neg B \leftrightarrow \Box_n \neg KB$ and, similarly, for m . Hence, putting $C := \neg B$, we obtain our desired result. \square

Let S be again $\exists \Sigma_1^b$, let N again be a cut of N_0 and let again $n \geq \mathfrak{a}^*$. Let $m^* := \rho(S) + \rho(K) + \rho(N) + \mathfrak{f}_0$.

Theorem 7.15. *We have $S_2^1 \vdash S^{\mathfrak{J}_{m^*+2}} \rightarrow \Delta_{(n)} S^N$.*

Proof. We reason in S_2^1 . Suppose $S^{\mathfrak{J}_{m^*+2}}$. By Theorem 7.10, we have $\Delta_{(m^*)}^{\mathfrak{J}_{m^*+2}} S^N$. By Theorem 7.13, it follows that $\Delta_n S^N$. \square

By specialising we find the following. Let $m^* := \rho(K) + \rho(N) + \mathfrak{f}$.

Theorem 7.16. *We have $S_2^1 \vdash \Delta_{(n)}^{\mathfrak{J}_{m^*+2}} B \rightarrow \Delta_{(n)} \Delta_{(n)}^N B$.*

We have the following extension of Theorem 7.15. This version could play a role in the development of Rosser arguments. Let S be Σ_1^0 , let N be a cut of N_0 and let $n \geq \mathfrak{a}^*$. Let $m^* := \rho(S) + \rho(K) + \rho(N) + \mathfrak{f}_1$.

Theorem 7.17. *We have $S_2^1 \vdash S^{\mathfrak{J}_{m^*+4}} \rightarrow \Delta_{(n)} S^N$.*

Proof. This uses the well-known fact that $S_2^1 \vdash S^{\mathfrak{J}^2} \rightarrow \blacksquare_{\rho(S)+i} S$. See [9]. Here i is a small constant for overhead. \square

Remark 7.18. Suppose N is shorter than \mathfrak{J}_n . Let $G := G_{N,n}$. We reason in A . Suppose $\Delta_{(n)}^{\mathfrak{J}_n N} G^N$. Then, since $\mathfrak{J}_n N$ is initial in N , we have $\Delta_{(n)}^N G^N$. On the other hand, we have $\Delta_{(n)}^N \Delta_{(n)}^N G^N$, and, hence, $\Delta_{(n)}^N \neg G^N$. Thus, $\Delta_{(n)}^N \perp$. *Quod non.* So, canceling our assumption, we find $\neg \Delta_{(n)}^{\mathfrak{J}_n N} G^N$ and, thus, $G^{\mathfrak{J}_n N}$.

Note the analogy of the present result with the result of Remark 7.7. ○

7.7 Under Guard

In this subsection, we extract some information from the Gödel sentences for $\Delta_{(a^*)}$. Let $\mathfrak{S} := \mathfrak{J}_{a^*} \cap \mathfrak{C}_{N_0, N_0 K}$.

Theorem 7.19. *Suppose A proves that NK is a cut of $N_0 K$. Suppose A is consistent. Then, $A \not\vdash NK \subseteq \mathfrak{F}_{N_0, N_0 K}(\mathfrak{S})$.*

Proof. Suppose $A \vdash NK \subseteq \mathfrak{F}_{N_0, N_0 K}(\mathfrak{S})$. Let $\tilde{N} := \mathfrak{F}_{N_0, N_0 K}^{-1}(NK)$. We note that $\mathfrak{F}_{N_0, N_0 K}$ restricted to \tilde{N} is an isomorphism between \tilde{N} and NK . Moreover, $\tilde{N} \subseteq \mathfrak{S}$. Let $G := G_{N, a^*}$.⁶

We reason in A . Suppose $\Delta_{(a^*)}^{\tilde{N}} G^N$. It follows that $\Delta_{(a^*)}^{NK} G^N$ and, hence $\neg G^{NK}$. On the other hand, by guarded reflection, we have G^{NK} . A contradiction. Hence, by canceling our assumption, $\neg \Delta_{(a^*)}^{\tilde{N}} G^N$, and so G^{NK} .

We return to the meta-language. We have shown that $A \vdash G^{NK}$. It follows that $A \vdash_{a^*} KG^N$ and, hence, $A \vdash \Delta_{(a^*)}^{NK} G^N$. Thus, $A \vdash \neg G^{NK}$. But this contradicts the consistency of A . □

Let $\mathfrak{S}^* := F_{N_0, N_0 K}(\mathfrak{S})$ and let $\mathcal{X} := \{\mathfrak{S}^* \subset NK \mid NK \text{ is } A\text{-provably a cut of } N_0 K\}$.

Theorem 7.20. *Suppose A is consistent. Then, $A + \mathcal{X}$ is also consistent.*

Proof. Suppose $A + \mathcal{X}$ is inconsistent. Then, for some finite subset \mathcal{X}_0 of \mathcal{X} , we have $A \vdash \neg \bigwedge \mathcal{X}_0$.

⁶Strictly speaking, we exceed our earlier framework here, since A need not prove that N is a cut of N_0 . However, it is easy to see that the marginal extensions does no harm. The skeptical reader can always replace N by $N(\text{cut}_{N_0}(N))N_0$, the interpretation that is N if N is indeed an N_0 -cut and that is N_0 otherwise.

Let \mathcal{Y}_0 be the set of N such that $\mathfrak{S}^* \subset NK$ is in \mathcal{X}_0 . We have:

$$\begin{aligned} A \vdash \neg \bigwedge \mathcal{X}_0 &\leftrightarrow \bigvee_{N \in \mathcal{Y}_0} \mathfrak{S}^* \not\subset NK \\ &\leftrightarrow \bigvee_{N \in \mathcal{Y}_0} NK \subseteq \mathfrak{S}^* \\ &\rightarrow \left(\bigcap_{N \in \mathcal{Y}_0} N \right) K \subseteq \mathfrak{S}^* \end{aligned}$$

Writing M for $\bigcap_{N \in \mathcal{Y}_0} N$, we find $A \vdash MK \subseteq \mathfrak{S}^*$. But this contradicts Theorem 7.19. □

Seeing that any NK such that $NK : A \triangleright \mathbf{S}_2^1$ has an initial segment that is K -internally definably isomorphic to an N_0K -cut, Theorem 7.20 tells us that there is an A -model \mathcal{M} in which \mathfrak{S} is below all K -internal interpretations of \mathbf{S}_2^1 . This implies, for example, that for any B and any $m \geq \max(\rho(A), \rho(B))$, we have $\mathcal{M} \models \Box_m^{\mathfrak{S}} B \rightarrow B^K$.

7.8 Rosser?

In this section we discussed the predicate Δ and have shown that it has many good properties. In Remarks 7.3, 7.7, 7.18 and in Subsection 7.7, we explored what information the Gödel sentences of Δ could provide. However, as we seen our information until now makes Δ too well behaved to obtain a contradiction.

There is clearly an infinity of variations on the Rosser sentences and there is some hope that these might lead to the desired contradiction.

8 In Search of Paradox

In this Section we study an attempt to prove Conjecture 3.13 that has some analogies to a paradox, variants of which were independently found by Stephen Yablo and the author.

8.1 Motivating Remarks

We consider a finitely axiomatised, sequential A with the Enayat property. Let $N_0 : \mathbf{S}_2^1 \rightarrow A$ and let K be defined with respect to N_0 as before. Let us briefly dwell on the Liar for K . By the Gödel Fixed Point Lemma, we find L such that $A \vdash L \leftrightarrow KL$. We note that inside A we have that either $L, \neg L^K, L^{KK}, \dots$, or

$\neg L, L^K, \neg L^{KK}, \dots$. We also have $A \vdash L^{\text{ID}\langle L \rangle K}$ and $A \vdash (\neg L)^{K\langle L \rangle \text{ID}}$, and, thus, $A \triangleright (A + L)$ and $A \triangleright (A + \neg L)$. So, L is an Orey sentence for A .

Nothing dangerous seems to follow from the existence of L since we only get the alternations. Can we eliminate the alternations by stipulating that L is false in all iterations of K ? Consider L^* such that $A \vdash L^* \leftrightarrow \forall x \in \delta_{N_0} \neg K^{x+1} L^*$.⁷ We reason in A . Suppose L^* . It follows that $\neg K L^*$ and, thus, $\neg (L^*)^K$. We may conclude $(\exists x \in \delta_K K^{x+1} L^*)^K$. Now if x were in the common cut $\mathfrak{C}_{N_0, N_0 K}$, we would have our contradiction. So, it follows that $N_0 K$ contains an element above the common cut. Suppose we start with $\neg L^*$, we get $\exists x \in \delta_K K^{x+1} L^*$. But what then? We do have $A \vdash (\exists x \in \delta_K K^{x+1} L^*)^{K\langle L^* \rangle \text{ID}}$. Can we tweak the argument in such a way that the existential claim gets a numerical witness? If we could, we would have $A \vdash (K^{n+1} L^*)^{K\langle L^* \rangle \text{ID}}$, and, hence, $A \vdash (L^* \wedge (K^{n+1} L^*)^K) \vee (L^*)^{K^{n+1}}$. It follows that $A \vdash (L^* \wedge K^{n+2} L^*) \vee (L^*)^{K^{n+1}}$. The first disjunct leads to a contradiction, so $A \vdash (L^*)^{K^{n+1}}$. By the fixed point equation, we find $A \vdash (\neg K L^*)^{K^{n+1}}$, and, hence, $A \vdash (\neg L^*)^{K^{n+2}}$. On the other hand, since K interprets A in A , we find $A \vdash (L^*)^{K^{n+2}}$. So, A is inconsistent.

The above program does not quite work. In the first place, we need to add some subtle details. In the second place, and more importantly, we need a substantial conjecture to make it work. However, this conjecture has some interest of its own.

Remark 8.1. The paradoxical reasoning sketched above is reminiscent of the reasoning in [16] concerning a descending hierarchy of truth predicates. (The article was first published in 1989 and the preprint appeared in 1985.) If you forget about the indices, this argument transforms into the well-known Yablo paradox, first published in [26]. ○

8.2 The Small-is-very-small Principle and its Variants

The main idea of our strategy is the use of some kind of numerical existence principle that allows us to replace a provable existence claim by a claim about an external number. This subsection provides the needed existence principle.

A theory is *restricted* if all of its axioms have depth-of-quantifier-alternations complexity below a given standard number k .

The *Small-is-very-small Principle*, or *SIVS*, tells us that, if a restricted theory proves that a number with a certain property exists in a sufficiently small cut (‘is small’) then the theory also believes that the number is standard (‘is very small’). Here the relevant small cut will depend on the property, or, more precisely the complexity of the formula defining the property.

⁷We will worry about details of defining the iteration of K later.

Theorem 8.2 (The Small-is-very-small Principle). *Consider a restricted sequential theory U with bound k and let $N_0 : S_2^1 \triangleleft U$. Let B be of the form $\exists x \in N_0 B_0(x)$. Let ℓ be $\max(k, \rho(B))$ plus some constant j for overhead.*

$$U \vdash \exists x \in \mathfrak{J}_\ell B_0(x) \quad \Rightarrow \quad \text{for some } m \text{ we have } U \vdash \exists x \leq \underline{m} B_0(x).$$

Here \mathfrak{J}_ℓ is the cut that was introduced in Subsection 7.5. The numeral \underline{m} is an N_0 -numeral.

The proof of Theorem 8.2 is given in full detail in [25]. Here we provide a quick sketch. Suppose C is $\exists x \in \delta_{N_0} C_0$ and D is $\exists y \in \delta_{N_0} D_0$. We write $C \leq D$ for $\exists x \in \delta_{N_0} (C_0 \wedge \forall y <^{N_0} x \neg D_0)$.

Proof-sketch. We work under the conditions specified in the theorem.

Suppose $U \vdash \exists x \in \mathfrak{J}_\ell B_0(x)$. It follows that there is a finite subsystem U_0 of U such that $U_0 \vdash \exists x \in \mathfrak{J}_\ell B_0(x)$. We may assume that U_0 is sequential and verifies $(S_2^1)^{N_0}$. Now consider R such that $U_0 \vdash R \leftrightarrow B \leq \square_{U_0, \ell}^{N_0} R$.

Reason in U_0 . In case $\square_{U_0, \ell}^{N_0} R$ is not witnessed in \mathfrak{J}_ℓ , we have R by the fixed point equation. If $\square_{U_0, \ell}^{N_0} R$ is witnessed in \mathfrak{J}_ℓ , we have R by reflection. We return to the meta-language.

We have shown $U_0 \vdash R$. By cut-elimination, we find $U_0 \vdash_\ell R$. Let m witness $U_0 \vdash_\ell R$. By Σ_1 -completeness, we find, in U_0 , that B is witnessed below m . \square

Example 8.3. Let U be a variant of PRA in the arithmetical language. Then U is a consistent restricted sequential theory. Let N_0 be the identical interpretation. Since, U is also reflexive, we can find an interpretation M such that $M : U \triangleright U$ and $U \vdash \square_U^I \perp$, for all definable N_0 -cuts I . It is obvious that we cannot have

$$U \vdash (\exists p \leq \underline{m} \text{ proof}_U(p, \perp))^M.$$

So, we do not have an analogue for Theorem 8.2, if we embed our existential sentence in a self-interpretation M . Intuitively, viewed from the standpoint of the world of U , the cuts definable inside M , seen from the outside, cannot be as small as the cuts we have at the general level of the theory. \circ

We can escape the above example if we restrict ourselves to finitely axiomatised theories A and restrict B to $(\Sigma_1^0)^{N_0}$ -formulas. We then get the following internalised form of Theorem 8.2:

Theorem 8.4. *Consider a consistent finitely axiomatised sequential theory A and let $N_0 : S_2^1 \triangleleft A$. Consider any number n . Then, there is an N_0 -cut \mathcal{I} , such that, for any $M : A \triangleright A$ with $\rho(M) \leq n$, and for any $S \in \Sigma_1^0$, we have: if $A \vdash S^{\mathcal{I}M}$, then S is true.*

We lost the restriction on the complexity of the existentially quantified formula from Theorem 8.2 by a small trick. The proof of this result can be found in [15] or [17]. The essence of the trick is also given in the proof of Theorem 8.7 below.

We note that Theorem 8.4 contains both a restriction of the theory (it has to be finitely axiomatised) and on the formula (it has to be $(\Sigma_1^0)^{N_0}$). One may wonder if the restriction to finitely axiomatised theories suffices. (We have seen, in Example 8.3, that the restriction to $(\Sigma_1^0)^{N_0}$ -sentences does not suffice.) Thus, we are led to the following conjecture.

Conjecture 8.5. Consider a finitely axiomatised sequential theory A and let $N_0 : S_2^1 \triangleleft A$. Consider any number n . There is an N_0 -cut \mathcal{I}_n such that, for any sentence $B := \exists x \in N B_0(x)$ with $\rho(B) \leq n$ and any $M : A \triangleright A$ with $\rho(M) \leq n$, we have:

$$(\dagger) \quad A \vdash (\exists x \in \mathcal{I}_n B_0(x))^M \quad \Rightarrow \quad \text{for some } m \text{ we have } A \vdash (\exists x \leq \underline{m} B_0(x))^M.$$

Here the \underline{m} is an N_0 -numeral.

We note that (\dagger) is equivalent to:

$$(\ddagger) \quad A \vdash (\exists x \in \mathcal{I}_n B_0(x))^M \quad \Rightarrow \quad \text{for some } m \text{ we have } A \vdash \bigvee_{k \leq m} (B_0(\underline{k}))^M.$$

○

The conjecture could just turn out to be provable by a slightly more clever Rosser argument than the ones I employed until now. My first attempts ran into the same kind of problems as my attempts to prove the truth of Conjecture 3.13 directly: somewhere a K on an undesired place. How to get rid of it?

There is an interesting equivalent of Conjecture 8.5.

Conjecture 8.6. Consider a finitely axiomatised sequential theory A and let $N_0 : S_2^1 \triangleleft A$. Consider any number n . There is an N_0 -cut \mathcal{J}_n such that, for any Σ_1^0 -sentence S and for any sentence C with $\rho(C) \leq n$ and any $M : A \triangleright A$ with $\rho(M) \leq n$, we have: $A \vdash (S^{\mathcal{J}_n} \vee C)^M \Rightarrow S$ is true, or $A \vdash C^M$. ○

We note that, if we put \perp for C , then we get something that is known, to wit Theorem 8.4.

Theorem 8.7. *Conjectures 8.5 and 8.6 are equivalent.*

Proof. Conjecture 8.5 implies Conjecture 8.6. Suppose we have Conjecture 8.5. We remind the reader that $S_2^1 \vdash x \in \mathfrak{J}_1 \rightarrow \exists y 2^x = y$. From this, it follows that $S_2^1 \vdash S^{\mathfrak{J}_1} \rightarrow \mathbf{true}(\ulcorner S \urcorner)$, where \mathbf{true} is the usual Σ_1^0 -truth predicate. See [9, Part C,

Chapter V, 5(b)] for a detailed treatment. We will write $\text{true}(S)$ for $\text{true}(\ulcorner S \urcorner)$. We note that $\rho(\text{true}(S))$ is a fixed standard number \mathfrak{z} independent of S . Let $\text{true}(x)$ be $\exists y \text{true}_0(y, x)$, where $\text{true}_0 \in \Delta_0$.

Let $\mathcal{J}_n := \mathfrak{J}_1 \mathcal{I}_{\max(n, \mathfrak{z}+1)}$, where \mathcal{I}_n is provided by Conjecture 8.5. Consider any C and K with complexities below n . Let $n' := \max(n, \mathfrak{z} + 1)$. We have:

$$\begin{aligned}
 A \vdash (S^{\mathcal{J}_n} \vee C)^M &\Rightarrow A \vdash (S^{\mathfrak{J}_1 \mathcal{I}_{n'}} \vee C)^M \\
 &\Rightarrow A \vdash (\text{true}^{\mathcal{I}_{n'}}(S) \vee C)^M \\
 &\Rightarrow A \vdash (\exists y \in \mathcal{I}_{n'} (\text{true}_0^{N_0}(y, S) \vee C))^M \\
 &\Rightarrow \text{for some } m, A \vdash \bigvee_{k \leq m} (\text{true}_0^{N_0}(\underline{k}, S) \vee C)^M \\
 &\Rightarrow S \text{ is true or } A \vdash C^M.
 \end{aligned}$$

The last step uses that if $\neg \text{true}_0(\underline{k}, S)$ is true, then $A \vdash \neg \text{true}_0^{N_0 M}(\underline{k}, S)$.

Conjecture 8.6 implies Conjecture 8.5. Suppose we have Conjecture 8.6. The proof uses an idea from [8]. Let n be given. Let n' be n plus a constant for overhead. We will be more specific about the choice of the constant later. Let $\mathcal{I}_n := \mathcal{J}_{n'}$. Suppose we have B and M , where $\rho(B)$ and $\rho(M)$ are $\leq n$.

Suppose $A \vdash (\exists x \in \mathcal{I}_n B_0(x))^M$. By the Gödel Fixed Point Lemma, we find R such that $A \vdash R \leftrightarrow B \leq \square^{N_0} R^M$. We note that the complexity of R is the complexity of B plus a constant \mathfrak{v} that just depends on N_0 and the arithmetisation of provability. So, traveling back in time, we take $n' := n + \mathfrak{v}$. It is easy to see that $A \vdash (\square^{\mathcal{I}_n} R^M \vee R)^M$. It follows that $\square R^M$ is true or $A \vdash R^M$. So, $A \vdash R^M$. Let m be the Gödel number of a witness of $A \vdash R^M$. We find that $A \vdash \text{proof}^{N_0 M}(\underline{m}, R^M)$. Combining this with $A \vdash R^M$, we find that, in A and inside M , B is witnessed below m . \square

In Appendix B, we formulate weaker versions of Conjecture 8.5 and Conjecture 8.6.

8.3 Conjecture 8.5 implies Conjecture 3.13

We address the matter of defining K^y . Let A be finitely axiomatised, sequential and Enayat. Suppose $N_0 : A \triangleright \mathbb{S}_2^1$ and $K : A \triangleright (A + \text{TB}_{N_0}^-)$. Let $\mathfrak{F} := \mathfrak{F}_{N_0, N_0 K}$ be a definable isomorphism between the cut $\mathfrak{C} := \mathfrak{C}_{N_0, N_0 K}$ of the N_0 -numbers and its image in the $N_0 K$ -numbers.

We define the function $\gamma(y, x)$ as follows:

- $\gamma(0, x) = x$,

- $\gamma(y + 1, x) := \text{subst}(\ulcorner \mathbf{K}(v_0) \urcorner, \text{num}(\gamma(y, x)))$.

Here subst is the substitution function and num assigns to a number the Gödel number of its numeral.

Par abus de langage, we also write γ for the arithmetisation of γ . The function γ is defined on the logarithmic numbers of N_0 . Let N_1 be a logarithmic cut, e.g. $N_1 = \mathfrak{J}_1 N_0$. We define, for $y \in \delta_{N_1}$,

- $\mathbf{K}^y D := (y = 0 \wedge D) \vee \exists z < y (y = z + 1 \wedge \exists u \in N_0 (\gamma(z, \ulcorner D \urcorner) = u \wedge \mathbf{K}(u)))$.

Here is the main result of this section.

Theorem 8.8. *The truth of Conjecture 8.5 implies the truth of Conjecture 3.13.*

Proof. We assume the truth of Conjecture 8.5. Consider a finitely axiomatised, sequential theory A . Suppose $N_0 : A \triangleright \mathbf{S}_2^1$ and $K : A \triangleright (A + \text{TB}_{N_0})$. We derive a contradiction.

Let $n^* := \max(2\rho(K), \rho(\mathbf{K}(x))) + 2$. We clearly may assume that \mathcal{I}_{n^*} is a logarithmic cut in N_0 , by shortening it when needed. We use the fixed point lemma to obtain: $A \vdash L \leftrightarrow \mathbf{K}(\forall w \in \mathcal{I}_{n^*} \neg \mathbf{K}^w L)$. We note that $\rho(L) \leq n^*$, since, generally, $\rho(\mathbf{K}(\underline{s})) = \rho(\mathbf{K}(x)) + 1$.

We have: $A+L \vdash \mathbf{K} \neg L$ and hence $A+L \vdash \neg L^K$. So, $A+L \vdash (\exists w \in \mathcal{I}_{n^*} \mathbf{K}^w L)^{KK}$. Similarly, we have $A + \neg L \vdash (\exists w \in \mathcal{I}_{n^*} \mathbf{K}^w L)^K$. Let $\widetilde{K} := KK\langle L \rangle K$. It follows that $\widetilde{K} : A \triangleright A$ and $A \vdash (\exists w \in \mathcal{I}_{n^*} \mathbf{K}^w L)^{\widetilde{K}}$. We note that $\rho(\widetilde{K}) = \max(\rho(L), 2\rho(K)) + 1$, so $\rho(\widetilde{K}) \leq n^*$.

We apply Conjecture 8.5 to obtain, for some m : $A \vdash (\bigvee_{k \leq m} \mathbf{K}^k L)^{\widetilde{K}}$. Hence,

$$K + L \vdash \bigvee_{k \leq m} L^{K^{k+2}} \text{ and } K + \neg L \vdash \bigvee_{k \leq m} L^{K^{k+1}}.$$

It follows that $(\ddagger) A \vdash \bigvee_{k \leq m+1} L^{K^{k+1}}$.

Since $K : A \triangleright A$, it follows that $A \vdash (\bigvee_{k \leq m+1} L^{K^{k+1}})^{K^{m+2}}$, and, hence, that $A \vdash \bigvee_{k \leq m+1} L^{K^{m+k+3}}$. On the other hand, by the definition of L , and (\ddagger) , we find: $A \vdash \bigwedge_{k \leq m+1} \neg L^{K^{m+k+3}}$. So A is inconsistent. \square

Remark 8.9. We note that the construction $\widetilde{K} := KK\langle L \rangle K$ preserves sentential restrictedness. So, we need Conjecture 8.5 only for a very special kind of interpretation—that does not even need to exist, given the fact that we are looking for a *reductio*. \circ

References

- [1] L.D. Beklemishev. Parameter free induction and reflection. Technical Report LGPS 171, Department of Philosophy, Utrecht University, 1996.
- [2] S. Buss. Cut elimination in situ. In R. Kahle and M. Rathjen, editors, *Gentzen's Centenary*, pages 245–277. Springer International Publishing, 2015.
- [3] C. Cieśliński. *The epistemic lightness of truth: deflationism and its logic*. Cambridge University Press, 2017.
- [4] A. Cordón-Franco, A. Fernández-Margarit, and F.F. Lara-Martín. A note on parameter free Π_1 -induction and restricted exponentiation. *Mathematical Logic Quarterly*, 57(5):444–455, 2011.
- [5] H.B. Enderton. *A mathematical introduction to logic*. Elsevier, 2001.
- [6] S. Feferman. Arithmetically definable models of formalized arithmetic. *Notices of the American Mathematical Society*, 5:679–680, 1958.
- [7] K. Fine and T. McCarthy. Truth without satisfaction. *Journal of Philosophical Logic*, 13(4):397–421, 1984.
- [8] H. Friedman. The disjunction property implies the numerical existence property. *Proceedings of the National Academy of Sciences*, 72(8):2877–2878, 1975.
- [9] P. Hájek and P. Pudlák. *Metamathematics of First-Order Arithmetic*. Perspectives in Mathematical Logic. Springer, Berlin, 1993.
- [10] R. Kaye, J. Paris, and C. Dimitracopoulos. On parameter-free induction schemas. *Journal of Symbolic Logic*, 53(4):1082–1097, 1988.
- [11] F. Pakhomov and A. Visser. On a question of Krajewski's. *The Journal of Symbolic Logic*, 84(1):343–358, 2019.
- [12] P. Pudlák. Some prime elements in the lattice of interpretability types. *Transactions of the American Mathematical Society*, 280:255–275, 1983.
- [13] P. Pudlák. Cuts, consistency statements and interpretations. *The Journal of Symbolic Logic*, 50(2):423–441, 1985.
- [14] A. Tarski, A. Mostowski, and R.M. Robinson. *Undecidable theories*. North-Holland, Amsterdam, 1953.
- [15] A. Visser. The unprovability of small inconsistency. *Archive for Mathematical Logic*, 32(4):275–298, 1993.
- [16] A. Visser. Semantics and the liar paradox. In D. Gabbay and F. Guentner, editors, *Handbook of Philosophical Logic, second edition*, volume 11, pages 149–240. Springer, Heidelberg, 2004.
- [17] A. Visser. Faith & Falsity: a study of faithful interpretations and false Σ_1^0 -sentences. *Annals of Pure and Applied Logic*, 131(1–3):103–131, 2005.
- [18] A. Visser. Pairs, sets and sequences in first order theories. *Archive for Mathematical Logic*, 47(4):299–326, 2008.
- [19] A. Visser. Can we make the Second Incompleteness Theorem coordinate free? *Journal of Logic and Computation*, 21(4):543–560, 2011.

- [20] A. Visser. What is sequentiality? In P. Cégielski, Ch. Cornaros, and C. Dimitracopoulos, editors, *New Studies in Weak Arithmetics*, volume 211 of *CSLI Lecture Notes*, pages 229–269. CSLI Publications and Presses Universitaires du Pôle de Recherche et d’Enseignement Supérieur Paris-est, Stanford, 2013.
- [21] A. Visser. Peano Corto and Peano Basso: A study of local induction in the context of weak theories. *Mathematical Logic Quarterly*, 60(1–2):92–117, 2014.
- [22] A. Visser. Why the theory R is special. In Neil Tennant, editor, *Foundational Adventures. Essays in honour of Harvey Friedman*, pages 7–23. College Publications, UK, 2014. Originally published online by Templeton Press in 2012. See <http://foundationaladventures.com/>.
- [23] A. Visser. On Q. *Soft Computing*, 21(1):39–56, 2017.
- [24] A. Visser. The interpretation existence lemma. In *Feferman on Foundations*, number 13 in Outstanding Contributions to Logic, pages 101–144, New York, 2018. Springer.
- [25] A. Visser. The small-is-very-small principle. *Mathematical Logic Quarterly*, 65(4):453–478, 2019.
- [26] S. Yablo. Truth and reflection. *Journal of Philosophical Logic*, 14(3):297–349, 1985.

A Vaught Set Theory continued

We give the proof of Theorem 2.1, i.e., we show that $\mathbf{VS} \triangleright_{\text{dir}} \mathbf{VS}^+$ via a one-dimensional interpretation.

Proof. We define \mathbf{PC}_0 as the virtual class of all x such that⁸

1. $x \sim x$
2. $\forall y \forall z ((x \sim y \wedge x \sim z) \rightarrow y \sim z)$
3. $\forall y \forall z (x \sim y \sim z \rightarrow x \sim y)$

Suppose x is in \mathbf{PC}_0 and $x \sim y$. We show $y \in \mathbf{PC}_0$.

Ad (1) We have $x \sim y \wedge x \sim y$. Hence, by (2) for x , $y \sim y$.

Ad (2) Suppose $y \sim z$ and $y \sim u$. it follows that $x \sim y \sim z$ and $x \sim y \sim u$. Hence, by (3) for x , we find $x \sim z$ and $x \sim u$. Ergo, by (2), $z \sim u$.

Ad (3) Suppose $y \sim z \sim u$. it follows that $x \sim y \sim z$, hence, by (3) for x , we find $x \sim z$. It follows that $x \sim z \sim u$, hence, by (3) for x , we find $x \sim u$. So we have $x \sim y$ and $x \sim u$, so, by (2) for x , we have $y \sim u$.

⁸‘PC’ stands for pre-cardinals.

The relation \sim is an equivalence relation on PC_0 . The only thing to check is symmetry. Consider $x \in \text{PC}_0$ and suppose $x \sim y$. We have $x \sim y \wedge x \sim x$ and, so, by (2), $y \sim x$.

We define PC_1 as the class of x in PC_0 such that

$$\forall u \forall v \forall f ((x \sim u \sim v \wedge f : u \sim v) \rightarrow u \sim f).$$

We note that \sim is an equivalence relation on PC_1 .

We show that PC_1 is closed under \sim . Suppose $\text{PC}_1(x)$ and $x \sim y$ and $y \sim u \sim v$ and $f : u \sim v$. Then, since $\text{PC}_0(x)$, we have $x \sim u \sim v$ and $f : u \sim v$. So, we may conclude $u \sim f$.

Finally, if $\text{PC}_1(x)$ and $f : x \sim y$, then $x \sim x \sim y$ and $f : x \sim y$, and, so $x \sim f$, and, hence $\text{PC}_1(f)$.

We now define $u \in^+ v$ iff $u \in v \wedge \text{PC}_1(v)$. We claim that we have VS^+ for \in^+ . We note that, whenever x is in PC_1 , it defines the same set for \in and for \in^+ . Also, any non-empty \in^+ -set x will be in PC_1 .

Clearly, any \in -finite set will be in PC_1 . Thus, it will have the same \in^+ -elements. In other words, finite sets are absolute with respect to our interpretation. It follows that we have the VS -axioms for \in^+ .

Since, Kuratowski pairs are constructed using finite sets, we easily see that Kuratowski pairs are absolute too. It follows that, whenever, f is an \in -function that is in PC_1 , then f is also \in^+ -function with the same input-output behavior.

Let $u \sim^+ v$ be defined like \sim with \in replaced by \in^+ . We claim (\dagger) for x in PC_1 , we have $x \sim y$ iff $x \sim^+ y$. First suppose $f : x \sim y$. Clearly, f and y will be in PC_1 , hence $f : x \sim^+ y$. Conversely, suppose $g : x \sim^+ y$. In case x is \in^+ -empty, x will be \in -empty, since $x \in \text{PC}_1$. Hence, $x : x \sim y$. In case, x is \in^+ -non-empty, also y and g will be \in^+ -non-empty. Hence, they are in PC_1 . It follows that $g : x \sim y$.

Consider any x . In case x is \in^+ -empty, we have $x : x \sim^+ x$. In case x is \in^+ -non-empty, it is in PC_1 , hence $x \sim x$, and, so, by (\dagger) , $x \sim^+ x$.

Consider any x and y with $x \sim^+ y$. In case x is \in^+ -empty, we have $x : y \sim^+ x$. Suppose x is \in^+ -non-empty. Then, x is in PC_1 . It follows, by (\dagger) , that $x \sim y$, and hence, that y is in PC_1 and $y \sim x$. Hence, by (\dagger) , $y \sim^+ x$.

Suppose $x \sim^+ y \sim^+ z$. If x is \in^+ -empty, y will be \in^+ -empty and so will be z . Hence, $x : x \sim^+ z$. Suppose x is \in^+ -non-empty. It follows that x is in PC_1 . Hence, by (\dagger) , $x \sim y$, y is in PC_1 and $y \sim z$. So $x \sim z$. Again by (\dagger) , $x \sim^+ z$.

Finally, suppose $f : x \sim^+ y$. In case x is \in^+ -empty, we find that f is \in^+ -empty, and, hence, $x : x \sim^+ f$. Suppose x is \in^+ -non-empty, then so are y and f . Then x , y and f are in PC_1 , and, hence, $f : x \sim y$. It follows that $x \sim f$ and, hence, by (\dagger) , that $x \sim^+ f$. \square

Our next order of business is to prove Theorem 2.2, to wit that $\mathbf{VS} \triangleright \mathbf{R}$ via a one-dimensional interpretation. There are two possible proofs. I will give (a sketch of) both.

First proof: By Theorem 2.1 it suffices to prove that $\mathbf{VS}^+ \triangleright \mathbf{R}$. The basic idea of our interpretation is to give the usual cardinal definitions of the arithmetical operations whenever they work. When they do not work we set them to a default value.

We write $\mathbf{pair}(x, y, z)$ for: z represents a Kuratowski pair with first component x and second component y .

We call our translation ρ . We define:

- $\delta_\rho(x) := (x = x)$,
- $x =_\rho y := x \sim y$,
- $Z_\rho(x) := \forall y y \notin x$,
- $\mathbf{adj}(x, y, z) := \forall u (u \in z \leftrightarrow (u \in x \vee u = y))$,
- $\mathbf{S}_0(x, y) := \exists z (z \notin x \wedge \mathbf{adj}(x, z, y))$,
- $\mathbf{S}_1(x, y) := \exists u \exists v (\mathbf{S}_0(u, v) \wedge x \sim u \wedge v \sim y)$,
- $\mathbf{S}_2(x, y) := \mathbf{S}_1(x, y) \wedge \forall z (\mathbf{S}_1(x, z) \rightarrow y \sim z)$,
- $\mathbf{S}_\rho(x, y) := \mathbf{S}_2(x, y) \vee (\forall z \neg \mathbf{S}_2(x, z) \wedge x \sim y)$,
- $\mathbf{U}(x, y, z) := \forall u (u \in z \leftrightarrow (u \in x \vee u \in y))$,
- $\mathbf{A}_0(x, y, z) := \forall u \neg (u \in x \wedge u \in y) \wedge \mathbf{U}(x, y, z)$,
- $\mathbf{A}_1(x, y, z) := \exists u \exists v \exists w (\mathbf{A}_0(u, v, w) \wedge x \sim u \wedge y \sim v \wedge z \sim w)$,
- $\mathbf{A}_2(x, y, z) := \mathbf{A}_1(x, y, z) \wedge \forall u (\mathbf{A}_1(x, y, u) \rightarrow z \sim u)$,
- $\mathbf{A}_\rho(x, y, z) := \mathbf{A}_2(x, y, z) \vee (\forall u \neg \mathbf{A}_2(x, y, u) \wedge z \sim y)$,
- $\mathbf{M}_0(x, y, z) :=$
 $\quad \forall u \in x \forall v \in y \exists w \in z \mathbf{pair}(u, v, w) \wedge$
 $\quad \forall w \in z \exists u \in x \exists v \in y \mathbf{pair}(u, v, w) \wedge$
 $\quad \forall p \in z \forall q \in z \forall u \in x \forall v \in y ((\mathbf{pair}(u, v, p) \wedge \mathbf{pair}(u, v, q)) \rightarrow p = q)$,
- $\mathbf{M}_1(x, y, z) := \exists u \exists v \exists w (\mathbf{M}_0(u, v, w) \wedge x \sim u \wedge y \sim v \wedge z \sim w)$,
- $\mathbf{M}_2(x, y, z) := \mathbf{M}_1(x, y, z) \wedge \forall u (\mathbf{M}_1(x, y, u) \rightarrow z \sim u)$,

- $M_\rho(x, y, z) := M_2(x, y, z) \vee (\forall u \neg M_2(x, y, u) \wedge z \sim y)$,

It is clear that on the standardly finite sets our operations behave as the ordinary successor, sum and product. Moreover, \leq defined as $x \leq y := \exists z (z + x = y)$ behaves as usual. Thus, ρ carries an interpretation of \mathbf{R} . \square

Remark A.1. We note that we could manipulate the interpretation of \in further in order to interpret principles like:

- $\forall x \forall y \forall z \forall u \forall v ((x \sim y \wedge u \notin x \wedge v \notin y \wedge \text{adj}(x, u, z)) \rightarrow \exists w \text{adj}(y, v, w))$.

Clearly, in the presence of such principles, we can build an interpretation following the above strategy that is simpler. \circ

Second proof. We interpret the theory of a category in \mathbf{VS} . We define \mathbf{Ob} as the class of x such that $\exists i : x \rightarrow x \forall y \in x i(y) = y$ and

$$\forall y \forall z \forall f : x \rightarrow y \forall g : y \rightarrow z \exists h : x \rightarrow z \forall u \in x f(g(u)) = h(u).$$

We define \mathbf{Morph} as the functions between the elements of \mathbf{Ob} and we take as identity on \mathbf{Morph} extensional sameness. We define identity arrows and composition in the obvious way. It is easy to see \mathbf{Ob} and \mathbf{Morph} with the chosen operations define a category in \mathbf{VS} and that the standardly finite sets are in \mathbf{Ob} .

We now define $\text{sum}(x, y, z)$ and $\text{prod}(x, y, z)$ as the category-theoretical sum and product. We note that these are partial operations but have verifiably good properties like commutativity and associativity.

Finally we define our interpretation, say ν , of \mathbf{R} by taking:

- $\delta_\nu := \mathbf{Ob}$,
- $=_\nu$ is isomorphism in our category,
- $Z_\nu(x) := \forall y y \notin x$,
- $A_\nu(x, y, z) := \text{sum}(x, y, z) \vee (\forall w \neg \text{sum}(x, y, w) \wedge z =_\nu y)$,
- $\text{sing}(x) := \exists y \forall z (z \in x \leftrightarrow z = y)$,
- $S_\nu(x, y) := \exists z (\text{sing}(z) \wedge A_\nu(x, z, y))$,
- $M_\nu(x, y, z) := \text{prod}(x, y, z) \vee (\forall w \neg \text{prod}(x, y, w) \wedge z =_\nu y)$.

The rest of the proof is as expected. \square

B Internal SIVS revisited

We have a weaker version of the conjectured internal Small-is-very-small Principle Conjecture 8.5. This version is suggested by attempts to prove Conjecture 8.5. It looks like this. Let A be finitely axiomatised and sequential and let $N_0 : A \triangleright S_2^1$.

Conjecture B.1. Let n be given. Then, there is an N_0 -cut \mathcal{I}_n , such that, for every sentence $B := \exists x \in N_0 B_0(x)$ with $\rho(B) \leq n$, and, for every $M : A \triangleright A$ with $\rho(M) \leq n$, we have:

$$A \vdash (\exists x \in \mathcal{I}_n B_0(x))^M \Rightarrow \text{there are } m \text{ and } k \text{ such that } A \vdash \bigvee_{i \leq m} \bigvee_{0 < j \leq k} (B_0(\underline{i}))^{M^j}.$$

Here M^j means the j -fold iteration of M . ○

Our conjecture also has an equivalent variant (analogous to Conjecture 8.6):

Conjecture B.2. Consider any n . Then, there is an N_0 -cut \mathcal{J}_n , such that for all $S \in \Sigma_1^0$ and for all C with $\rho(C) \leq n$ and for all $M : A \triangleright A$ with $\rho(M) \leq n$ we have: if $A \vdash (S^{\mathcal{J}_n} \vee C)^M$, then, for some k , we have S is true or $A \vdash \bigvee_{0 < j \leq k} C^{M^j}$. ○

Theorem B.3. *Conjectures B.1 and Conjecture B.2 are equivalent.*

The proof is analogous to the proof of the equivalence of Conjectures 8.5 and 8.6. The aim of the proof sketch below is mainly to highlight the differences.

Proof. Conjecture B.1 implies Conjecture B.2. Suppose we have Conjecture B.1.

Let $\mathcal{J}_n := \mathfrak{J}_1 \mathcal{I}_{\max(n, \mathfrak{J}+1)}$, where \mathcal{I}_n is provided by Conjecture B.1. Consider any C and K with complexities below n . Let $n' := \max(n, \mathfrak{J} + 1)$. We have:

$$\begin{aligned} A \vdash (S^{\mathcal{J}_n} \vee C)^M &\Rightarrow A \vdash (S^{\mathfrak{J}_1 \mathcal{I}_{n'}} \vee C)^M \\ &\Rightarrow A \vdash (\text{true}^{\mathcal{I}_{n'}}(S) \vee C)^M \\ &\Rightarrow A \vdash (\exists y \in \mathcal{I}_{n'} (\text{true}_0^{N_0}(y, S) \vee C))^M \\ &\Rightarrow \text{for some } m \text{ and } k, A \vdash \bigvee_{i \leq m} \bigvee_{0 < j \leq k} (\text{true}_0^{N_0}(\underline{i}, S) \vee C)^{M^j} \\ &\Rightarrow S \text{ is true or } A \vdash \bigvee_{0 < j \leq k} C^{M^j}. \end{aligned}$$

Conjecture B.2 implies Conjecture B.1. Suppose we have Conjecture B.2. Let n be given. Let n' be $n + \mathfrak{J}$ and let $\mathcal{I}_n := \mathcal{J}_{n'}$. Suppose we have B and M , where $\rho(B)$ and $\rho(M)$ are $\leq n$.

Suppose $A \vdash (\exists x \in \mathcal{I}_n B_0(x))^M$. We write:

$$\Box \bigvee_j D_{0 < j} := \exists p \exists y \leq |p| \text{proof}(p, \bigvee_{0 < i \leq y} D_i).$$

Here $|p|$ is the entier of the 2-logarithm of p . By the Gödel Fixed Point Lemma, we find R with $A \vdash R \leftrightarrow B \leq \Box^{N_0} \bigvee_{0 < j} R^{M^j}$.

It is easy to see that $A \vdash (\Box^{\mathcal{I}_n} \bigvee_{0 < j} R^{M^j} \vee R)^M$. It follows that $\Box \bigvee_{0 < j} R^{M^j}$ is true or, for some k , we have $A \vdash \bigvee_{0 < j \leq k} R^{M^j}$. So, for some k , we have $A \vdash \bigvee_{0 < j \leq k} R^{M^j}$. Let m be the Gödel number of a witness of $A \vdash \bigvee_{0 < j \leq k} R^{M^j}$. Then, $A \vdash \text{proof}^{N_0 M}(\underline{m}, \bigvee_{0 < j \leq k} R^{M^j})$. Combining this with $A \vdash \bigvee_{0 < j \leq k} R^{M^j}$, we find the desired result. \square

Here is the analogue of Theorem 8.8.

Theorem B.4. *The truth of Conjecture B.1 implies the truth of Conjecture 3.13.*

The proof is just a trivial variant of the proof of Theorem 8.8.

Proof. We assume the truth of Conjecture B.1. Consider a finitely axiomatised, sequential theory A . Suppose $N_0 : A \triangleright \mathbf{S}_2^1$ and $K : A \triangleright (A + \text{TB}_{N_0})$. We derive a contradiction.

Let $n^* := \max(2\rho(K), \rho(\mathbf{K}(x))) + 2$. We clearly may assume that \mathcal{I}_{n^*} is a logarithmic cut in N_0 , by shortening it when needed. We use the fixed point lemma to obtain: $A \vdash L \leftrightarrow \mathbf{K}(\forall w \in \mathcal{I}_{n^*} \neg \mathbf{K}^w L)$. We note that $\rho(L) \leq n^*$, since, generally, $\rho(\mathbf{K}(\underline{s})) = \rho(\mathbf{K}(x)) + 1$.

We have: $A+L \vdash \mathbf{K} \neg L$ and hence $A+L \vdash \neg L^K$. So, $A+L \vdash (\exists w \in \mathcal{I}_{n^*} \mathbf{K}^w L)^{KK}$. Similarly, we have $A + \neg L \vdash (\exists w \in \mathcal{I}_{n^*} \mathbf{K}^w L)^K$. Let $\widetilde{K} := KK \langle L \rangle K$. It follows that $\widetilde{K} : A \triangleright A$ and $A \vdash (\exists w \in \mathcal{I}_{n^*} \mathbf{K}^w L)^{\widetilde{K}}$. We note that $\rho(\widetilde{K}) = \max(\rho(L), 2\rho(K)) + 1$, so $\rho(\widetilde{K}) \leq n^*$.

We apply Conjecture B.1 to obtain,

$$\text{for some } m \text{ and } k: A \vdash \left(\bigvee_{i \leq m} \bigvee_{0 < j \leq k} \mathbf{K}^i L \right)^{\widetilde{K}^j}.$$

Hence,

$$K + L \vdash \bigvee_{j \leq m+k} L^{K^{j+2}} \text{ and } K + \neg L \vdash \bigvee_{j \leq m+k} L^{K^{j+1}}.$$

It follows that $(\ddagger) A \vdash \bigvee_{j \leq m+k+1} L^{K^{j+1}}$.

Since $K : A \triangleright A$, it follows that $A \vdash (\bigvee_{j \leq m+1} L^{K^{j+1}})^{K^{m+k+2}}$, and, hence, that $A \vdash \bigvee_{j \leq m+1} L^{K^{m+k+j+3}}$. On the other hand, by the definition of L , and (\dagger) , we find: $A \vdash \bigwedge_{j \leq m+k+1} \neg L^{K^{m+k+j+3}}$. So A is inconsistent. \square

C Conjectures and Questions

C.1 Conjectures

c1. No finitely axiomatised consistent Vaught theory is Enayat. (Conjecture 3.7.) Equivalently, we have the following conjecture. Suppose U is a consistent Vaught theory. Then $\mathfrak{T}(U)$ is not quasi-finite. (Conjecture ??.)

If this conjecture fails, we conjecture that no finitely axiomatised consistent sequential theory is Enayat. (Conjecture 3.13.)

c2. Suppose A is finitely axiomatised and consistent and sequential. Let $N : \mathbb{S}_2^1 \triangleleft A$. Then, there is no extension of \mathbb{S}_2^1 that is mutually interpretable with $A + \text{TB}_N^-$. (Conjecture 3.17.)

It is a well known open question whether every sequential theory is mutually interpretable with an extension-in-the-same-language of \mathbb{S}_2^1 . Our conjecture provides a possible example to illustrate a negative answer to this question.

c3. Let U be Vaught. and let $N : \mathbb{R} \triangleleft U$. Suppose $\alpha \triangleright_U \text{TB}_N^-$. Then, $\alpha \triangleright_U \text{USB}_N^-$. (Conjecture 6.5.)

c4. Suppose A is a finitely axiomatised Vaught in signature Θ_0 . Let $N : A \triangleright \mathbb{R}$. Suppose further that $\top \triangleright_A \text{TB}_N^-$. Then, there is a β such that $\top \triangleright_A \beta \triangleright_A \text{TB}_N^-$.

More generally, we may conjecture the following. Suppose A is a finitely axiomatised Vaught theory and $\top \triangleright_A V$. Then, there is a B such that $\top \triangleright_A B \triangleright_A V$. (Conjecture 6.9.)

c5. Consider a finitely axiomatised sequential theory A and let $N_0 : \mathbb{S}_2^1 \triangleleft A$. Consider any number n . There is an N_0 -cut \mathcal{I}_n such that, for any sentence $B := \exists x \in N B_0(x)$ with $\rho(B) \leq n$ and any $M : A \triangleright A$ with $\rho(M) \leq n$, we have:

$$(\dagger) \quad A \vdash (\exists x \in \mathcal{I}_n B_0(x))^M \quad \Rightarrow \quad \text{for some } m \text{ we have } A \vdash (\exists x \leq \underline{m} B_0(x))^M.$$

Here the \underline{m} is an N_0 -numeral.

We note that (\dagger) is equivalent to:

$$(\ddagger) \quad A \vdash (\exists x \in \mathcal{I}_n B_0(x))^M \quad \Rightarrow \quad \text{for some } m \text{ we have } A \vdash \bigvee_{k \leq m} (B_0(\underline{k}))^M.$$

(Conjecture 8.5.)

There is an interesting equivalent of Conjecture 8.5. Consider a finitely axiomatised sequential theory A and let $N_0 : \mathbb{S}_2^1 \triangleleft A$. Consider any number n . There is an N_0 -cut

\mathcal{J}_n such that, for any Σ_1^0 -sentence S and for any sentence C with $\rho(C) \leq n$ and any $M : A \triangleright A$ with $\rho(M) \leq n$, we have: $A \vdash (S^{\mathcal{J}_n} \vee C)^M \Rightarrow S$ is true, or $A \vdash C^M$. (Conjecture 8.6.)

- c6. Let n be given. Then, there is an N_0 -cut \mathcal{I}_n , such that, for every sentence $B := \exists x \in N_0 B_0(x)$ with $\rho(B) \leq n$, and, for every $M : A \triangleright A$ with $\rho(A) \leq n$, we have:

$$A \vdash (\exists x \in \mathcal{I}_n B_0(x))^M \Rightarrow \text{there are } m \text{ and } k \text{ such that } A \vdash \bigvee_{i \leq m} \bigvee_{0 < j \leq k} (B_0(\dot{i}))^{M^j}.$$

Here M^j means the j -fold iteration of M . (Conjecture B.1.)

An equivalent conjecture runs as follows. Consider any n . Then, there is an N_0 -cut \mathcal{J}_n , such that for all $S \in \Sigma_1^0$ and for all C with $\rho(C) \leq n$ and for all $M : A \triangleright A$ with $\rho(M) \leq n$ we have: if $A \vdash (S^{\mathcal{J}_n} \vee C)^M$, then, for some k , we have S is true or $A \vdash \bigvee_{0 < j \leq k} C^{M^j}$. (Conjecture B.2)

C.2 Questions

- q1. Suppose there is a finitely axiomatised, consistent Vaught theory that is Enayat. Can we show, under that assumption, that *all* finitely axiomatised, consistent, Vaught theories are Enayat theories? (Question 3.8.)
- q2. In Subsection 3.4, we have shown that in the recursively enumerable sequential case, we can characterise Enayat theories in a coordinate-free way. Not only is the question of Enayatness independent of the Gödel numbering, but Gödel numberings are not mentioned in the characterisation. Can we do something similar in the Vaught case? (Question 3.9.)
- q3. Are there any interesting relations between theories, besides mutual interpretability, that preserve Enayatness? (Question 3.16.)
- q4. Is there an example of a finitely axiomatised theory A with the N -Enayat property for some $N : \text{Succ}_0 \triangleleft A$, when we demand that the Gödel numbering is p-time computable? (Question 4.1.)
- q5. Is there an example of a finitely axiomatised theory A such that we have the Enayat property for all $N : \text{Succ}_0 \triangleleft A$? (Question 4.2.)
- q6. Does TB_N^- have a restricted axiomatisation over U ? (Question 5.7.)
- q7. Can we show that, for no consistent sequential U , we have $\text{TB}_N^- \triangleright_U \text{USB}_N^-$? (Question 5.10.)
- q8. Can we prove the non-existence of a finitely axiomatised consistent sequential uniform Enayat theory without a detour over the second incompleteness theorem? (Question 5.14.)
- q9. Is there a finitely axiomatised Vaught theory that is uniformly Enayat? Here uniformity is explicated using USB^- . (Question 5.15.)

HOW PARADOX FARES IN INCONSISTENCY ROBUST LOGIC AND BEYOND: COMPUTATIONAL AND NATURALIZED APPROACHES

JOHN WOODS*

*The Abductive Systems Group, Department of Philosophy, 1866 Main Mall,
University of British Columbia, Vancouver V6T 1Z1, Canada*

`john.woods@ubc.ca`

“The true mystery of the world is the visible, not the invisible”.
Oscar Wilde¹

“Are we grown up enough to live with uncertainties or will we repeat the mistakes of the twentieth century and pledge blind allegiance to another certainty?”
Davide Malone²

Abstract

A paradox is an antinomy. It is a valid argument from apparently faultless premisses to a logical contradiction. Paradoxes are derivable in natural languages and also in some mathematical languages. But, for all the differences between these language types, they are strikingly alike when it comes to paradox-avoidance. The languages of mathematical logic are artificial. They are purpose-built to further the logicians’ theoretical ends, enabled by how the theorist has designed his language. If a mathematical language blocks a paradox, it may be a pleasant surprise or a fully intended consequence, but either way the blockage is a technician’s provision. On the other hand, natural languages resist paradoxes *naturally*, by way of their largely unspoken logico-semantic

*Author’s note: Arising from some work with Carl Hewitt on the rational management of inconsistent information systems, the present paper was written in 2017 in response to an invitation to place it in a forthcoming edited volume on inconsistency. Following the editors’ abandonment of the project, I have thought it advisable to update my findings and place them elsewhere.

¹Oscar Wilde, *The Picture of Dorian Gray*, published with deletions in the July 1890 issue of *Lipicott’s Monthly Magazine*. Complete version by the Belknap Press in 2011.

²David Malone, *Dangerous Knowledge*, BBC 4 documentary, 2007. http://www.dailymotion.com/playlist/x1cbyd_xSilverPhinx_bbc-dangerous-knowledge/1

conventions. It lies in the linguistic DNA of human language to carry a bias against paradox, save for its uses as jokes or irony. With a natural language, paradox-resistance just come naturally. With an artificial one, it has to be installed, albeit by technicians whose very own mother tongues embed a natural resistance to paradox.

My further purpose is to show that even under dire assumptions — e.g. that a negation-inconsistent system (one in which a derived sentence has an entailed negation) is absolutely inconsistent (one in which every sentence is entailed) — inconsistencies are no impediment to our cognitive advancement in the general case. (This is a generalization of Carl Hewitt’s position on the practical utility of inconsistency robust systems.) In each case, my more general one and Hewitt’s IR (Inconsistency-Robust) one, the cognitive utilities of a system’s inconsistencies, whether absolutely so or only pervasively so, is an empirically discernible fact about human cognitive behaviour.

Since Hewitt’s Inconsistency Robust Direct Logic (IRDL) emphasizes this empirical discernibility point, and my approach also does, my final objective is a twofold one.³ We should equip logics of human reasoning with the means to engage with empirical reality. This might best be done by a working partnership between computational logicians and logicians of a naturalized perspective. So let’s make it official:

- Logicians of inconsistency-robustness and logicians of naturalized and empirical leanings have enough in common to warrant a working alliance or anyhow a limited liability partnership.

Part I

1 Some motivating considerations

A mathematician friend once observed that while computer science could only have been invented by mathematicians, that wasn’t in all respects an entirely good thing.⁴ Computer science was brought about in the lengthy aftermath of what mathematicians had done to logic sixty years earlier.⁵ Since time immemorial, there have been

³Carl Hewitt, “Formalizing common sense reasoning for scalable inconsistency-robust information coordination using Direct Logic TM Reasoning and the Actor Model”, *HAL Archives-Ouvertes*, France, online 2015; also in Hewitt and Woods (2015), 3-103. See also “Strong types for Direct Logic”, *PhiSci Archives*, online 2017.

⁴Ray Reiter in conversation with the author in London in the late 1990s.

⁵Gottlob, *Begriffsschrift: A Formula Language, Modeled Upon that of Arithmetic, for Pure Thought*, in Jean van Heijenoort, editor, *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, pages 5-82, Cambridge, MA: Harvard University Press, 1967; and Alfred Tarski, “The concept of truth in formalized languages”, in his *Logic, Semantics and Metamathematics*:

mathematicians who have hated inconsistency like the plague.⁶ More recently, in the breakthrough that launched arithmetic into the transfinite, diagonalization has been an indispensable vehicle of modern mathematics.⁷ Diagonalization is at the heart of what I will call Church's Paradox, which because it calls into question the very nature of what it means to be a mathematical theory, is the most fundamental of all the mathematical paradoxes, namely, the unenumerability of first-order theorems.

“This, of course, is a deplorable state of affairs . . . Indeed, if there is no formalization of logic as a whole, then there is no exact description of what logic is, for it is in the very nature of an exact description that it implies a formalization. And if there is no exact description of logic, then there is no sound basis for supposing that there is such a thing as logic.”⁸

Church's conclusion here is not much mentioned in the Church literature, and certainly not as a named paradox. Not an antinomy, of course, but his conclusion has the ring of a destructive nihilism.

Church's paradox is important. It presses the question of why it is that the theory that best characterizes the natural numbers has proofs that cannot be written down if the system is consistent.

There may in this result be enough of a resemblance to take note of between the oddity of Peano arithmetic and a characteristic feature of the common law which is the legal system which presides in England (not Scotland), the United States and Canada, and other countries of the Commonwealth. In these systems there is as recognized distinction between statute law and judge-made law. Judges make new laws when, in deciding a case presently before them, they write reasons for judgement (*rationes decidendi*) that create *precedents* or rules of law.⁹ Although reasons for judgement are usually detailed and finely wrought, the precedent created thereby

Papers from 1923-1938, edited by John Corcoran and translation by J. H. Woodger, pages 152-278, Indianapolis: Hackett, 1983. First published in 1933.

⁶For example, Frege: “. . . we must really face the possibility that we may still in the end encounter a contradiction which brings the whole edifice down in ruins.” See Gottlob Frege, *Die Grundlagen der Arithmetik: Eine logisch mathematische Untersuchung über den Begriff der Zahl*, Breslau: Wilhelm Koebner, 1884. Translated into English by J. L. Austin as *The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number*, Oxford: Basil Blackwell, 1950, p. IX.

⁷Georg Cantor, “Über eine elementare Frage der Mannigfaltigkeitslehre”, in William B. Ewald, editor, *From Immanuel Kant to David Hilbert: A Source Book in the Foundations of Mathematics*, volume 2, pages 920-922, Oxford: Oxford University Press, 1996. First published in 1891.

⁸Alonzo Church, “The Richard paradox”, *American Mathematical Monthly*, 41 (1934-361).

⁹Not to overlook regulations imposed by the executive branch under the delegated authority of Congress or Parliament.

is never expressly articulated. A judge-made law is *lex non scripta*. Judge-made laws are never written down, notwithstanding that they are binding on all lower courts and can have “persuasive authority” in like courts in sister jurisdictions. The doctrine of *stare decisis* is not a wholly settled matter in common law jurisprudence, but there is a very large body of tradition and present thought to which I subscribe, according to which

- It is not possible, even in principle, to give full expression to judge-made laws without misdescribing them.

Even so, it is emphasized that

- Judge-made laws are perfectly intelligible to lawyers and legal scholars. A capable judge is able without unnatural effort the general legal principle or rule of law created by a precedent, and can without unnatural effort to apply it in future cases.

The main thing to notice here is the common law’s insistence that sometimes things known implicitly cannot be made explicit without falling into error. This is epistemologically interesting. It reminds us that the common law harbours an epistemology which gives implicit knowledge, belief and understanding its due. I, for one, regard this as cognitive virtue.¹⁰

Of course, it hardly needs saying that Peano arithmetic is at least as different from common law as chalk is from cheese. Still, there are some similarities which might repay some further future attention. The “unwriteability” of judge-made laws causes no injury to the administration of justice. Similarly, the “unwriteability” of the Peano proofs does no damage to work-a-day arithmetic. Arithmetic is certainly as safe as the common law is. That, I say again, might be something to look into. However, I don’t want to leave the comparison of the law’s unwriteability with the unwriteability of the Peano proofs. I have already said that I endorse the view that it is not possible, even in principle, to give to a judge-made law full expression without getting it wrong. Hewitt does not endorse the counterpart view as regards the Peano proofs. Hewitt proposes a solution to Church’s paradox in which the Peano proofs attain full expressibility in a second-order adaptation of Classical Direct Logic. I will come back to this.

Diagonal methods sometimes generate paradoxes, which are inconsistencies in the last circle of mathematical hell. The people who invented computer science

¹⁰A more detailed discussion can be found in my *Is Legal Reasoning Irrational? An Introduction to the Epistemology of Law*, 2nd edition, revised and expanded, volume two of the Law and Society series, London: College Publications, 2017; chapters 3 and 4. First published in 2015. See also my “The role of the common in cognitive prosperity”, forthcoming in *Informal Logic*.

brought to this new frontier the habits of thought of modern mathematics, not least its bias against inconsistency. On the other hand, although invented by people with just this same bias, the remit of computer science far exceeds the digital modelling of the foundations of mathematics. It also aspires to the digitalizations of human reasoning in general, in modalities that are swifter, at least as reliable, and in other ways perhaps of better quality. “Ah, but there’s the rub!” (as the old saying has it). Reasoning even at its humanly possible best doesn’t always dance to mathematical biases of the people who engineer the requisite software. Of particular note are big information systems. Carl Hewitt has characterized a big information system as one that requires multiples of millions of lines of code to computerize. It is pervasively and perpetually inconsistent, and not subject to inconsistency-cleansing. Such systems, says Hewitt, are *inconsistency-robust*. Although indispensably valuable, any device (I speak counterfactually) that removed a system’s every inconsistency would destroy its practicality. “There is no practical way to test for inconsistency”, yet “[e]ven though a system is inconsistent, it is not meaningless.”¹¹ It is important to understand what Hewitt means by the robustness of an inconsistent information-circulation system. By ‘inconsistent’, he means the presence of pervasively and in-eradicably negation-inconsistencies and by ‘robust’, he means not only the absence of absolute consistency, but the presence of inconsistency-management measures that keep the system in good cognitive health. If, as I believe with Hewitt’s concurrence that even at its humanly healthiest, the cognitive economics of humanity are riddled with tolerable levels of pervasive inconsistency, the consequences for logic and computer science alike speak for themselves. For what is wanted is a solid theoretical accounting of these facts in a way that preserves our evident cognitive prospects

- *The question is whether logic or computer science in any form in which we have now are up to the heavy lifting.*

In the majority of logics, e.g. first-order logic (the one most commonly in play) even a smidgeon of inconsistency causes a system to break down. By a theorem known as *ex falso quodlibet*, a negation-inconsistent first order system logically entails the system’s every sentence. These are not the logics that Hewitt or I want for inconsistency robustness, or for paradox either.¹²

Where Hewitt and I agree is on the point of the pervasiveness of negation-inconsistency in robust information systems and, indeed, in any thriving cognitive economy. We agree that classical logics (and intuitionist ones too) are not capable of running the inconsistency-management measures of systems such as these. Many

¹¹Carl Hewitt, IRDL (2017); see early in the conclusion.

¹²John Woods, “How robust can inconsistency get?”, *IfCoLoG Journal of Logics and Their Applications*, 1 (2014), 177-216.

nonclassical logicians are of the view, Hewitt among them, that the best way of keeping systems in good check is by blocking or otherwise impeding certain of the proof-measures of any logic that sanctions *ex falso*. Hewitt and I have different views about whether this is the optimal way to proceed,¹³ but since paradox is our subject here, the debate shifts from inconsistency-inhibition to paradox-avoidance. In a computation approach, only the software engineers can bring this to pass and they can do this only after the concept-engineers have given them something to proceed with. In natural language approaches, or so I will try to show, paradox-avoidance is provided for implicitly by the semantic conventions of human speech. So I think that the value of the computational and naturalistic approaches have for one another lies mainly in the empirically discernible behavioural regularities which each side takes as *data for* its respective theoretical approach.

The expression ‘data for theory’ is a technical term of art. Let \mathcal{T} be a theory with a subject-matter \mathcal{S} . and \mathcal{D} be its data for theory. Then the data for theory with respect to \mathcal{T} are those propositions about \mathcal{S} which satisfy the following two conditions: (i) The community of \mathcal{T} -theorists take the propositions in \mathcal{D} to be true; (ii) they think that the data in question have such importance that should \mathcal{T} ’s *theoretical* provisions conflict with these pre-theoretical ones, the burden is on \mathcal{T} to justify the discompliance. In a manner of speaking data for theory are presumptions that obtain theoretical footfall as a free pass. One of the larger risks of scientific enquiry is the tendency of theorists to cut too much slack to their confirmation-biases. Unless properly restrained, the integrity of can be compromised by the inclusion of propositions that are “friendly” to the theorist’s interest in a positive outcome. Data-corruption is ably discussed under the heading of ‘data-bending’ by Gerd Gigerenzer and others.¹⁴ We should bear in mind, however, that the avoidance of data-bending is much more easily proposed than achieved. But we need not tarry further with it here for what concerns us in this paper.

IRDL is an extension of classical direct logic (CDL), augmented by measures for the management of inconsistency robustness. A key provision of any Direct Logic is its syntactic intolerance of formulaic self-reference. There are no Gödel-like formulas in this logic. One of the central aims of IRDL is to arrive at an account of how such systems work to our (not perfect) advantage. Hewitt’s “Strong types

¹³Carl Hewitt, “Speed bumps in argumentation: Discursive inference for massive pervasively-inconsistent information”, in Dov Gabbay, Lorenzo Magnani, Woosuk Park, and Ahti-Veikko Pietarinen, editors, *Natural Arguments: A Tribute to John Woods*, volume 40 in the Tributes series, pages 35-44, London: College Publications, 2019. See also my “Response to Carl Hewitt”, in Gabbay et al. (2019), pages 45-48.

¹⁴Gerd Gigerenzer, “From tools to theories” in Carl Graumann and Kenneth J. Gergen, editors, *Historical Dimensions of Psychological Discourse*, pages 336-359, Cambridge: Cambridge University Press, 1996. See also Woods (2013/14), section 2. “Data-bending”, pp. 73-78; see also p. 483.

for direct logic” [30] substantially repairs that omission, and will be my focus in the first two parts of this essay. Later on, I’ll consider some alternative possibilities for paradox-abatement.

Classical Direct Logic is Hewitt’s logic of choice for the mathematical parts of inconsistency robust theories. CDL is also proposed as the replacement of first order logic as the standard formulation for computer science. In Hewitt and Woods [34], Hewitt advanced a proof that mathematics proves its own formal consistency.¹⁵ Straightaway this might prompt readers to raise an eyebrow. If Hewitt can learn to live with the pervasive inconsistencies of big information systems such as those that guide the reflections of the Federal Reserve Board or the Bank of England, or indeed humanly own belief-revision systems, why would he care whether *mathematics* is consistent? The question is misguided. Whether he cares or not, Hewitt thinks that mathematics just *is* provably consistent formally. Accordingly, he thinks it desirable to have a logic for mathematics that respects that presumed fact, in the face of Gödel’s results to the contrary.¹⁶ CDL is Hewitt’s logic of choice for that purpose.

This is where paradoxes become important. A system that gives rise to paradox is inconsistent. This means that CDL should be paradox-free. Since many of the logical paradoxes arise in mathematical systems, and all of them either require or admit of “self-referential” diagonal methods to ground their proofs, anyone with a stake in the consistency of mathematics would have good reason to find ways to disarm mischievous diagonalizations, that is, diagonalizations that generate mathematical paradox. Since diagonal methods are themselves mathematical instruments, that would be reason enough to try to sort out those that advance the course of mathematics from those that mistakenly “show” mathematics to be paradoxical.

By far the standard way to disarm a mischief-making diagonalization proof about a subject matter *S* is by imposing *type constraints* on its operating manual. In the close aftermath of the paradox that was widely taken to have blown set theory apart

¹⁵“Inconsistency robustness in foundations: Mathematics proves its own consistency and other matters”, in Hewitt and Woods, (2015) pages 104-157. The proof is re-advanced in “Strong types”. Direct Logic develops axiomatizations for the natural numbers, real numbers, ordinal numbers, set theory and the lambda calculus. The axioms of these systems are categorical. That is to say, up to a unique isomorphism there is just one model that satisfies the respective axioms. By application of “speed-bump” constraints on proof-rules, Classical Direct Logic blocks the known paradoxes of classical mathematics. When Actors are added to the mix, a strongly typed CDL is inferentially decidable. That is to say, every true proposition is provable and every proposition is either provable or disprovable. Furthermore, theorems of these systems are not enumerable by a provably total procedure.

¹⁶Of course, it will be clear to those who read him, that Hewitt’s reason for caring whether mathematics is consistent is that if it isn’t, the unrestrictedly free use of Classical Direct Logic in all intelligent applications would be precluded.

in 1902, recovery measures were proposed in Russell’s theory of ramified types, and in short order the iterative conception of set would take hold. Later, when Tarski turned his mind to the widely-accepted destruction of theories of truth brought about by the Liar paradox, he would set out to do for truth what had already been done for sets. New rules for the truth-predicate would site its attributions in fixed positions in a transfinite array of attribution sites, just as like arrangements had been made for the attribution of the membership predicate. The trouble was, and still is, that building such hierarchical architectures is a lot more difficult to do than aspire to do.¹⁷ A principal objective of “Strong types” is to find a way to keep diagonalization from harm by imposing type constraints in ways that are both efficacious and credible.¹⁸

Part II

1 Paradox and diagonalization

In “Paradoxes, self-reference and truth in the twentieth century”, Andrea Cantini ably summarizes the impact of paradox on mathematical logic. Paraphrasing her closely, the paradoxes motivated important new theoretical developments such as those in type theory, axiomatic set theory, and combinatory logic.¹⁹

Diagonalization was originally introduced by Cantor. It launched transfinite

¹⁷See Fairouz Kamareddine, Twan Laan and Rob Nederpelt, “A history of types”, in Dov M. Gabbay, Francis Jeffrey Pelletier and John Woods, editors, *Logic: A History of its Central Concepts*, volume 11 of Gabbay and Woods, editors, *Handbook of the History of Logic*, pages 451-511, Amsterdam: North-Holland, 2012.

¹⁸Jaakko Hintikka, in conversation, once likened the interplay of diagonalization and typing to an opera about rival geniuses at war. The dramatic tension would be centred in this question: “Are the typists smart enough to restrain what diagonalizers were smart enough to have thought up in the first place?” All the more so, he added, when the geniuses are one and the same.

¹⁹They also prompted new methods for proving metamathematically fundamental results, as with, for example, fixed point theorems and theorems for incompleteness, undecidability and undefinability. They prompted the application of inductive definability and generalized recursion. They stimulated the introduction of new semantic methods, such as revision theory, and semi-inductive definitions, which rest on nontrivial set theoretic developments. They helped promote new axioms for set theory, such as the anti-foundation axiom and those for the mathematics of circular phenomena. They suggested ways to investigate non-classical logics, such as contract-free and many valued logics, and logics with generalized quantifiers. They suggested frameworks of flexible typing for the foundations of mathematics and computer science. They made possible the application of forms of self-referential truth and type-free comprehension in AI and theoretical linguistics. (Dov M. Gabbay and John Woods, editors, *Logic from Russell to Church*, volume 5 of their *Handbook of the History of Logic*, pages 875-1013, Amsterdam: North-Holland, 2009; p. 875.)

mathematics, generalized the power set axiom to the infinite and transfinite case, and yet also provoked as yet unsolved difficulties in transfinite set theory, including the problems induced by paradox. Diagonalizations are also widely taken to provide the logical grounding of the Liar, the Richard and Grelling paradoxes, as well as the Curry and the Russell, and are indispensable to the Gödel theorems and Church's and Rosser's incompleteness results, to name just these. Diagonalizations form a family of like but different proof methods, exhibiting the underlying structure of proof by *reductio absurdum*, some taking a matrix-typological form and some not; some being constructive and others not.

It is important not to confuse self-application with self-reference. "Every non-negative integer has a successor" applies to the number 57 but doesn't refer to it. Even so, sometimes self-application and self-reference are said to come to the same thing. I quote here from Smullyan. To aid the eye I enclose the quoted passage in square brackets. I have slightly reformatted the passage to preserve its quoted-status.

[“We use the symbol “ x ” as a variable ranging over expressions of the English language. By the diagonalization of an expression, we mean the result of substituting the quotation of the expression for every occurrence of the variable in question. For example, consider the following expression

1. John is reading x .

The expression (1) is not a sentence, true or false, but becomes a sentence (true or false) upon substituting the quotation of any expression for “ x ”. If we substitute the quotation of (1) for ‘ x ’, we obtain the diagonalization of (1), which is

2. John is reading ‘John is reading x ’.

Now, (2) is a sentence, and it asserts that John is reading (1). However, (2) is not self-referential; it does not assert that John is reading (1). Let us consider the following expression.

3. John is reading the diagonalization of x .

The diagonalization of (3) is the following

4. John is reading the diagonalization of “John is reading the diagonalization of x .”

Sentence (4) asserts that John is reading the diagonalization of (3), but the diagonalization of (3) is (4) itself. Thus (4) asserts that John is reading the very sentence (4)! Thus (4) is self-referential It might be easier to understand this if we use the following abbreviations. Let us use ‘ \mathcal{J} ’ to abbreviate ‘John is reading’ and ‘ \mathcal{D} ’ to abbreviate ‘the diagonalization of’. Then (3) and (4) assume the following abbreviated forms:

3'. $\mathcal{J}\mathcal{D}x$

4'. $\mathcal{J}\mathcal{D}'\mathcal{J}\mathcal{D}x'$

The sentence (4)' asserts that John is reading the diagonalization of (3)', but the diagonalization of (3)' is (4)'[?] itself.”²⁰

It is open to question whether what Smullyan characterizes as self-reference qualifies as a genuine species of reference and might better thought of as a kind of *syntactic* self-application involving self-intra-substitutions for free variables. Even so, whatever we call it, Smullyan’s characterization doesn’t appear to deliver what the Gödel diagonalization requires for the incompleteness proof. The missing element is negation (or some near thing), reflecting that the basic structure of diagonalization takes the form of a *reductio* proof. Gödel intended his diagonalization to produce sentences that say of themselves that they are not provable in formalized arithmetic. What diagonalization is wanted for here is self-*non*-application. We see both these elements at play in the Cantor diagonalizations that opened up arithmetic to the calls of the transfinite, in proofs of Richard’s paradox, the Liar and the Grelling and those of Russell and Curry.

Dale Jacquette is right to say that, unlike the Cantor matrix-topological diagonalization and later adaptation of it for the Richard paradox, there is little in the make-up of proofs of the logical and semantic-paradoxes that calls out for diagonalizations in that form.

“The difference in the case of Cantor’s original diagonalization and variations on it like Richard’s paradox is that there is an intuitively justified reason for displaying information relevant to the diagonal construction in the form of a matrix, which arises naturally from the denumerably infinite listing of irrational numbers in infinite decimal or binary digital expansions. This is simply what a list of decimal or binary digital expansions must look like. There is nothing contrived about the matrix in

²⁰Raymond Smullyan, *Diagonalization and Self-Reference*, Oxford: Clarendon Press, 1984; pp. 3-4.

that case, but rather a kind of inevitability. The same is not true where paradoxes like the liar, Grelling's, Russell's, Curry's, or other so-called diagonalizations are concerned, which makes it intuitively appropriate to distinguish them as non-matrix-topological."²¹

This might be a good time to say a brief word about the Berry Paradox, concerning which opinions are mixed. Some commentators think that the only thing of note that it shows is how silly natural language is. In the original formulation, the paradox arises from the expression, 'the smallest positive integers are definable in fewer than twelve words'. Assuming that it refers to that number, it does so in eleven words. Hence paradox. There is a good deal of to-ing and fro-ing about the capacity of English to deliver the paradox. Some say that the proof is flawed by the ambiguity of 'definable'. Others pin the blame on its vagueness. Some hold that even if these criticisms hold true of the English word 'definable', it is possible to transform the English version into a formal notation in which the Berry's structure is more convincingly elucidated. Let W be the set of all English words and Z the set of all tuples of W . Put it for set Y that

$$|Y| = |W|, \forall y \in Y, y \in \mathbb{N}, 0 \leq y|W|$$

And define the function f as the bijection of word tuples and integers. The first stage of the paradox is to establish f . Then the next stage contains the assertion that $\exists x \in Z, f(x) \notin Y$, which clearly contradicts the construction.

Of course, the proof still pivots on the English word 'definable' and inherits whatever trouble it had in the paradox's original informal presentation. Perhaps this would be enough to disarm its formal proof.²² George Boolos used similar formulation to give a simplified proof of the Gödel incompleteness result.²³ Related work shows that Kolmogorov complexity is not computable. If it were otherwise, paradoxes like the Berry would come flooding back. This leads some to think that the definition of the Berry number is paradoxical only in the sense that it's not possible to compute it. When we come back to the Berry in Part C, I'll show how a like result can be found in the semantics of ordinary English. Even so, the Berry structure has serious further uses in mathematical logic in which 'definable' is

²¹Dale Jacquette, "Diagonalization in logic and mathematics", in Dov Gabbay and Franz Guentner, editors, *Handbook of Philosophical Logic*, 2nd edition, pages 55-147, Dordrecht: Kluwer, 2002; p. 91.

²²Gregory Chaitin, "The Berry Paradox", *Complexity*, 1 (1995), 26-30.

²³George Boolos, "A new proof of the Gödel incompleteness theorem", *Notices of the American Mathematical Society*, 36 (1989), 338-390. Reprinted in his *Logic, Logic, and Logic*, pages 383-388, Cambridge, MA: Harvard University Press, 1998.

clearly formulated in formal terms, and important limitation results are established. Even so, the Berry structure has serious further uses in mathematical logic in which ‘definable’ is clearly formulated in formal terms, and important limitation results are established.

2 The diagonal lemma and fixed points

It is generally accepted that the mathematical foundations of computer science, while rigorous, realistic and comprehensive, shouldn’t be needlessly complex. For example, although a commonplace of formal semantics, the object-language/metalanguage distinction is difficult for software engineers to handle, and those of them who work on paradoxes prefer to avoid it, as is done, for example in Hewitt’s CDL for intelligent applications. The example of Gödel is especially interesting. The means by which Gödel arrived at g , the coded sentence that truly asserts its own unprovability, is provided by what Carnap would later call the Diagonal Lemma, which gives “the Y untyped fixed point operator” on propositions.²⁴

Let \mathbf{n} be the numeral for the natural number n . The Diagonal Lemma asserts that for any formula $\ulcorner \Phi(x) \urcorner$ of elementary arithmetic with x its sole variable, there exists a number \mathbf{n} named by for n for which the following condition holds:

$$g(\Phi[\mathbf{n}]) = n.$$

The leftmost expression is the Gödel number of the expression arising from $\ulcorner \Phi(x) \urcorner$ by replacement of x by \mathbf{n} in the coding provided by Gödel. In any formalized language for which the Diagonal Lemma holds, the Liar is easily derived if it has its own truth predicate T . When applied to “ $\sim T[\mathbf{n}]$ ”, the Diagonal Lemma produce an n that is the Gödel number of “ $\ulcorner \sim T[\mathbf{n}] \urcorner$ ” says that the sentence with this Gödel number is false. But the sentences are one and the same.

We come now to an interesting reservation of Hintikka, which I myself find persuasive. It is not at all clear that the lemma holds for natural languages, prompting Hintikka to say that “the status of the diagonal lemma strictly speaking needs a review.”²⁵ He allows that the lemma reflects a combinatorial fact about basic arithmetic, registered by “ $\ulcorner \sim T[\mathbf{n}] \urcorner$ ”. But that alone doesn’t show that it has semantic

²⁴Rudolf Carnap, *The Logical Syntax of Language*, London: Routledge & Kegan Paul, 1937. First published, in German, in 1934. Kurt Gödel, “On formally undecidable propositions of *Principia Mathematica* and related system, I” in van Heijenoort, *From Frege to Gödel*, pages 595-617. First published, in German, in 1931.

²⁵Jaakko Hintikka, *The Principles of Mathematics Revised*, Cambridge: Cambridge University Press, 1996; p. 142.

status. As Hintikka notes, ‘true’ seems not to apply to open sentences.²⁶ These seem to me points worth stressing:

- *It is open to question whether the Diagonal Lemma holds for natural languages.*²⁷

Although very useful to mathematicians, fixed points generate inconsistencies when applied to untyped propositions. Unless they are typed using orders these propositions are open to paradoxes such as the Liar, the Curry and the Berry. Assigning orders to each proposition block the mathematical paradoxes by eliminating the fixed points by which apparently self-referential propositions are able to lead to contradiction. Russell tried to reduce all types to propositions with orders, thereby making mathematical practice impossible. Russell sought repairs by proposing that for every higher-order proposition there is a logically equivalent first-order proposition. The ramification proof posed difficulties that comprised the effort. Hewitt’s CDL overcomes this difficulty that only *propositions* that have types with order, and not *procedures* such as lambda calculus, real *numbers*, and *points* in Euclidean space.

Fixed points are specified by the Fixed Point Method. Let FL be a formal language and I an interpretation of it with domain D . Predicates of FL are interpreted as subsets of the Cartesian product $D \times D$; that is, predicates denote n -ary relations on D . Let the interpretation of the expression “ F^n ” be the ordered pair $\langle R_1, R_2 \rangle$ of n -ary relations on D with R_1 the extension of “ F^n ” and R_2 its counterextension.

Axioms are now provided to regulate the behaviour of “ F^n ” in enlargements of FL , by taking an existing model and extending it with the new axioms. Axioms are required to be in conditional or biconditional form, admitting but not requiring quantifiers. The model for the new axioms is specified by transfinite induction. Consider the case in which axioms are biconditionals $\lceil \Phi \text{ iff } \psi \rceil$, where Φ is a sentence ‘ $F_n(a_1, a_2, \dots, a_n)$ ’ and ψ is a formula in which some of the a_i occur. The model expands incrementally by so adjusting the interpretation of “ F ” that instances of Φ take the same truth value as ψ took at the prior stage. At each succeeding stage, new sentences emerge as true instantiations of ψ , and this requires a corresponding expansion of ψ ’s instances to preserve the truth of the axiom.

In due course, “through the magic of infinity”,²⁸ a stage will be reached at which, for each instance of ψ , true or false on the interpretation, Φ will already have been ‘fixed’ in the same truth value. These junctures are called *fixed points*.

²⁶What, we may ask, of “Everything Nixon says is true”? Doesn’t it ascribe the predicate “is true” to “Everything x Nixon says”? It doesn’t. It ascribes “is true” to all the statements in the range of “everything x Nixon says”, not to the expression that refers to them.

²⁷I would say it more strongly. There is no “perhaps” about it. The Diagonal Lemma does not hold for natural languages.

²⁸Chris Mortensen, *Inconsistent Mathematics*, Dordrecht and Boston: Kluwer, 1995; p. 136.

Let $I_0 = I$ be the interpretation with which the stage process begins. The fixed point method spawns a series of interpretations $I_0, I_1, I_2, \dots, I_n$. In successive interpretations of I_0 , the ordered pair $\langle R_1, R_2 \rangle$, the interpretation of “ F^n ”, is altered by extensions of R_1 and R_2 . This succession of alterations can be thought of as governed by the rule:

$$X, I_n + X(I_n).$$

Then a fixed point of X is an interpretation

$$I_\lambda : x(I_\lambda) = I_\lambda$$

which gives a model of the enriched language.

Suppose now that ‘ T_r ’ is Kripke’s truth predicate for his implicitly hierarchical approach to truth in natural language. At the beginning of the construction, the extension and counter extension of “ T_r ” are empty, so all attributions and denials of truth are without truth values.²⁹ Sentences not containing the truth predicate can nevertheless be evaluated “intuitively”, some as true, some as false. These constitute respectively the extension and counterextension of the truth predicate at the next higher level, at which more sentences are evaluated as true or as false at this level. The process continues into the transfinite, and the resulting extensions and counterextensions are swept together into a fixed point \mathfrak{D} the point at which all sentences that would sooner or later belong to the extension (counterextension) of ‘ T_r ’ already do. The consistency of Kripke’s semantics is purchased at the cost of completeness. At its fixed point, application of X tells us no more about the extension and counterextension of ‘ T_r ’ that can be told. Kripke clearly wants to deliver a concept of truth which better resembles the intuitive notion abandoned by Tarski,³⁰ Even so, this leaves plenty of room for a question:

- *Why would we opt for all the ad hocness and counterintuitiveness of Kripke’s molestations of ‘true’ in English instead of such learning to live with paradox and leave the meaning of ‘true’ intact.*³¹

²⁹Saul Kripke, “Outline of a theory of truth”, *Journal of Philosophy*, 72 (1975). Reprinted in his *Philosophical Troubles: Collected Papers Volume 1*, pages 75-98, Oxford: Oxford University Press, 2011; paperback 2013.

³⁰Similar such attempts are made by Hans G. Herzberger, “Paradoxes of grounding in semantics”, *Journal of Philosophy*, 67 (1970), 145-167, and Anil Gupta, “Truth and Paradox”, in Robert L. Martin, editor, *Recent Essays on Truth and the Liar Paradox*, pages 175-236, Oxford: Clarendon Press, 1984.

³¹I have drawn here on my discussion of the Diagonal Lemma and the Fixed Point Method from my *Paradox and Paraconsistency: Conflict Resolution in the Abstract Sciences*, in chapter 7 (268-276) and chapter 5, pp. 175-180, respectively, Cambridge: Cambridge University Press, 2003.

This is a question to which I'll return in later sections. But now for a more technically relevant point.

In “Strong types”, Hewitt points out that parameterized types are instrumental in formalizing classical mathematics. This turns out to matter.

“Using strong parameterized types, it is impossible to construct *I'mUnprovable* [= *g*] because the *Y* untyped fixed point operator does not exist for strongly typed propositions”.³²

Strong typing provides that, although every mathematical object has a type, there is no universal type under which all others are subsumed. A parametrized type is one that distinguishes or ‘specializes’ it. For example, ‘proposition $\langle 2 \rangle$ ’ is a proposition with parameter 2, which makes it a proposition of order 2. ‘Strong \langle proposition $\langle 1 \rangle$ \rangle ’ is a well-formed string of a proposition of order 1. To express that *T* is a type we write ‘*T* ::’. To say that proposition $\langle 2 \rangle$ is a type we write ‘proposition $\langle 2 \rangle$::’. Thanks to Girard’s paradox, if the type *U* of all types exists, it is possible to define the type of all well-foundations over elements of *U* in a way that yields a type-theoretic version of the paradoxes of Burali-Forti and Mirimanoff.³³ Hewitt goes on to say,

“In this way, formal consistency of mathematics is preserved without giving up power because there does not seem to be any *practical uses* for *I'mUnprovable* in Computer Science”.³⁴ (Emphasis in the first instance is added.)

This, too, strikes me as another point worth emphasizing:

- *If computer science has no practical interest in such sentences, why would it need, or tolerate, a base sublogic that mandates the diagonal methods by which they are produced?*

The significance of the diagonal theorem for the incompleteness theorem speaks for itself, and is well worth further examination. Of more immediate interest to our present concerns, is the fact that all paradoxes in diagonal formulation (e.g. Curry’s

³²Hewitt to Woods, personal correspondence, 7 July, 2017

³³See here A. S. Troelstra and Dirk van Dalen, *Constructivism in Mathematics*, volumes 1 and 2, Amsterdam: North-Holland, 1988. The Burali-Forti is the paradox of the greatest ordinal and Mirimanoff’s is a version of it. See Burali-Forti, “Sulle classi ben ordinate”, *Rendicanti del Circolo Matematico di Palermo*, 11 (1897), 260, and “Un question sui numeri transfinito”, *Recondidi del Circolo Matematica di Palermo*, 11 (1897), 154-164; and D. Mirimanoff, “Les antimonies de Russell et de problème fondamentale de la théorie des ensembles”, *L’Enseignement Mathématique*, 19 (1917), 37-52.

³⁴Carl Hewitt, “Strong types for direct logic”.

and Löb's) are blocked in Hewitt's logic. The Curry paradox fails because Curry's methods of generation lack an untyped fixed point. Löb's paradox fails for the same reason.³⁵ So do the others, with the possible exception of Church's paradox,

Hewitt writes:

“Unfortunately we still do not have a good resolution of Church's paradox: Strong types have blocked all other known paradoxes! Why did Church and others drop the ball? Part of the reason is that they may have become distracted by the linguistic form of Richard's paradox. No one continued to focus on the provability of computational enumeration theorems, which is at the heart of the difficulty”.³⁶

As Hewitt points out, these blockages arise from the way in which the CDL approach differs from that of diagonal logics. As the name implies, CDL deals with propositions *directly*, rather than by way of semantically inert strings coded as integers in the manner that engenders the Gödel sentence.

Soon after, Hewitt takes a more upbeat position. He now thinks that the strong types approach “may have a solution to Church's Paradox. However, the solution raises profound issues in the philosophy of mathematics.”³⁷ Hewitt's raised spirits are programmatically structured, offering a pathway to hoped-for relief: I summarize the findings:

- First, mathematics is strongly typed.
- Secondly, mathematics self-proves that it is not open, i.e., proofs are not computationally enumerable. As Hewitt writes, “In other words, the paradox that concerned [Church] (because it could mean the demise of formal mathematical logic) has been transferred into a fundamental theorem of foundations!”
- Thirdly, it used to be thought that mathematics could be based solely on strings of characters. Then diagonalization was discovered and things haven't been the same since. Such strings exist for the second order type axioms of Peano. Even so, the induction axiom quantifies over the uncountably many propositions of first order to categorically characterize the natural numbers.

³⁵Löb's paradox provides that in any formal system FA with Peano arithmetic (PA), for any wff P , if it is provable in FA that ‘if P is provable in FA then P is true’, then P is provable in FA. But from this it follows that from this if $2 + 2 = 5$ is provable in PA, then $2 + 2 = 5$ is not provable in PA. M. H. Löb, ‘Solution of a problem of Leon Henkin’, *Journal of Symbolic Logic*, 20 (1955), 115-118.

³⁶Hewitt to Woods, personal correspondence, 7 July, 2017.

³⁷Hewitt to Woods, personal correspondence, July 19, 2017.

Not only does Peano arithmetic have uncountably many propositions as axioms, they cannot all be obtained by abstraction from character strings. Although proofs abstracted from character strings can be computationally enumerated, but this is not enough to enumerate all of the proofs of Peano arithmetic.³⁸

Hewitt’s solution of the Church won’t sit well in all the high courts of establishment logic.³⁹ But it could be revealing to see how kindly the computer science community would take to it. It shouldn’t be surprising. Mathematical logicians and computer scientists have different interests, and correspondingly different mandates. Computer scientists have a large stake in how minds work, chiefly human ones since they are the minds with which computer scientists are most familiar. Mathematical logicians are interested in mathematical foundations and have no deep stake in disclosing how minds work, unless they are also computer scientists. This division of interests is reflected in their characteristically different approaches to natural language, which nearly everyone agrees is the most publically discernible manifestation of human mentality. The first impulse of the mathematical logician is to slight any natural language of his acquaintance and then to write rules for its replacement. Since they have a stake in knowing how human languages actually work if computer scientists were to do the same, they’d be asking for trouble.

3 Wittgenstein’s Paradox

Writing in 1937, Wittgenstein lodges *a reductio* against the incompleteness theorem.

“Let us suppose I prove the improvability (In Russell’s system) of [Gödel’s] P , then by this proof I have proved P . Now if this proof were one in Russell’s system, I should in this case have proved that it belonged and did not belong to Russell’s system. But there is a contradiction here! [*This*] is what comes from making up such sentences.”⁴⁰
(Emphasis mine)

It is rather striking that Wittgenstein’s paradox has had so little play in establishment circles. The received view is that, in this case, Wittgenstein missed the

³⁸There have been further strengthenings of Hewitt’s solution and I daresay more to come.

³⁹Hewitt’s proof of mathematics’ self-proved consistency is structured as a proof by contradiction in which the interpretation of embedded nonlogical terms have weight. Proofs of this sort have a better name than they deserve. See here my “What did Frege take Russell to have proved?” *Synthese*, 198 (2021), 3949-3977; see especially section 7, “The first principles paradox”.

⁴⁰Ludwig Wittgenstein; *Remarks on the Foundations of Mathematics*, Oxford: Blackwell, 1956; pp. 50-51.

boat because he simply failed to understand Gödel's proof. Of course, there are exceptions. In his book on the Gödel theorem, Franz Berto gives Wittgenstein a chapter-long hearing. He accords Wittgenstein a mixed but sympathetic review, concluding the book with a discussion of "the costs and benefits of making Wittgenstein plausible."⁴¹ A stronger, and I think better, defence of Wittgenstein can be found in Timm Lampert "Wittgenstein and Gödel: An attempt to make 'Wittgenstein's objection' reasonable".⁴²

Wittgenstein's position is that the Gödel sentence g renders mathematics inconsistent. Hewitt's position is that mathematics proves its own consistency. If Hewitt is right, Wittgenstein isn't wrong. Hewitt is convinced that he's right. He writes,

"Let us suppose [Gödel writings are correct and therefore] I prove the improbability (in Russell's system) of [Gödel's *I'munporovable*] P : [i.e., $\vdash \not P$ where $P \Leftrightarrow \not P$] then by this proof I have proved P [i.e., $\vdash P$]. Now if this were one in Russell's system [i.e., $\vdash P$] — I should in this case have proved at once that it belonged [i.e., $\vdash P$] and did not belong [i.e., $\vdash \neg P$ because $\neg P \Leftrightarrow \vdash P$] to Russell's system.

But there is a contradiction here! [i.e., $\vdash P$ and $\vdash \neg P$]. . ."⁴³

But Berto is certainly right to name the first section of his chapter 12 "When geniuses meet . . .".

If Hewitt and I are right to emphasize the ease with which negation inconsistency in big-information computerized systems is tolerated, others are to be credited for having discerned the good that often comes of it. This is not to say that *spotted* inconsistencies can't sometimes be a bit of a challenge, but even there, there is no general cause for alarm. We should pause with this awhile.

⁴¹Francesco Berto, *There's Something About Gödel: The Complete Guide to the Incompleteness Theorem*, Madden, MA: Wiley Blackwell, 2009. First published, in Italian in 2008. See chapter 12, "Gödel versus Wittgenstein and the Paraconsistent Interpretation". See also Victor Rodych, "Misunderstanding Gödel: New arguments about Wittgenstein and new remarks by Wittgenstein", *Dialectica*, 57 (2003), 279-313. In a companion paper, I read Berto as leaning somewhat more favourably to Wittgenstein. See his "The Gödel paradox and Wittgenstein's reasons", *Philosophia*, 17(2009), 208-219.

⁴² *Mathematica Philosophia Mathematica*, 26 (2018), 324-345. Suffice it to say that, as Lampert sees it, there is a difference between the arithmetic formalization of a PM formula Φ , $\lceil \mathcal{F}_A(\Phi) \rceil$, and a metamathematical formalization of it, $\lceil \mathcal{F}_M(\Phi) \rceil$. Wittgenstein's point is that, for the Gödel formula g , $\lceil \mathcal{F}_M(g) \rceil$ is not an admissible item in the formalization of the arithmetical parts of PM. In other words, "Gödel's proof is not a mathematical proof" (p. 325).

⁴³Hewitt to Woods, personal communication, 28 July 2017. Note that Hewitt's parsings of Wittgenstein are placed within square brackets.

4 The sunny side of inconsistency

Let us pick up the thread of a spotted inconsistency in a theory, and attend to the options outlined in Russo *et al.*

“The choice of an inconsistency handling strategy depends on the context and the impact it has on other aspects of the development process. Resolving the inconsistency may be as simple as adding or deleting information from a software description. However, it often relies on resolving fundamental conflicts, or taking important design decisions. In such cases, immediate resolution is not the best option, and a number of choices are available.”

They list the following:

“*Ignore* — it is sometimes the case that the effort of fixing an inconsistency is too great relative to the (low) risk that the inconsistency will have any adverse consequences. In such cases, developers may choose to ignore the existence of the inconsistency in their descriptions. Good practice dictates that such decisions be revisited as a project progresses or as a system evolves.”⁴⁴

Other options include “Defer”, and hope for improved future measures; “Circumvent” as an allowable exception to a software rule; and “Ameliorate” by limiting the inconsistency’s involvement in the system’s inferential provisions.

In a further development, joint work by Dov Gabbay and Anthony Hunter requires notice. For example, Gabbay and Hunter [21, 22] discuss a commonplace example about real-estate listings.

- In knowledge-bases that model everyday human phenomena, inconsistency indicates the desirability of upgrading the model. It is a non-injurious invitation to upgrade.
- Some data-bases contain planned inconsistencies, circulating information of use to different subscribers. For example, a real estate company might value a client’s house at one sum for mortgage qualification, ten percent lower for property tax purposes, and in-between for market-listing purposes. It is interesting that, for those valuations to be sound, there need be no fact of the

⁴⁴Alexandra Russo, Bashar Nusibeh and Steve Easterbrook, “Making inconsistency respectable in software development”, *Journal of Systems and Software*, 56 (2000).

matter about the true value of the house.⁴⁵

- In keeping track of these inconsistent evaluations, realtors not only must keep an idea on which evaluation is assigned for which purpose, but also to adjust the system's inference-engine to protect it from consequence-overload.
- Of particular importance is the Gabbay & Hunter constraint against changing the meaning of the system's classical connectives or crimping the truth-conditions on classical consequence. In shorter words, Gabbay & Hunter inconsistency-management is avowedly anti-paraconsistentist.⁴⁶

From these reflections, we can derive four valuable lessons:

One is that inconsistency really is a routine and unavoidable feature of the human cognitive economy.

The second is that whatever the inconsistency-management devices that may recommend themselves to us, it is a cheat on empirically discernible cognitive reality to choke-off the entailment relation.

Thirdly, and relatedly, the intelligent management of inconsistency is not answerable in any deep way to the exigencies of what inconsistencies entail, and should instead reserve its options for what is rationally inferable from entailment-providing premisses of whatever kind or logical character.

Fourthly and finally, the restabilization of belief in the aftermath of inconsistent incoming information is achievable entirely independently of the restoration of informational consistency.

5 Natural language and paradox

Coming back now to formalized arithmetic, of course, nobody actually does arithmetic in a Gödel-like set-up $\mathcal{F}_A(PM)$ or $\mathcal{F}_A(PA)$.⁴⁷ Nobody needs diagonalization

⁴⁵Dov M. Gabbay and Anthony Hunter, "Making inconsistency respectable 1: A logical framework for inconsistency in reasoning", in Ph. Jourand and T. Kelemen, editors, *Fundamentals of Artificial Intelligence Research*, volume 535 of Lecture Notes in Computer Science, pages 19-32, Berlin: Springer, 1991, and "Making inconsistency respectable, part 2: Meta-level handling of inconsistency", in M. Clarke, R. Kruse and S. Seraffin, editors, *Lecture Notes in Computer Science*, volume 742, pages 129-136, Berlin: Springer, 1992.

⁴⁶However, Hunter appears to have changed his mind about this in "Reasoning with contradictory information using quasi-classical logic", *Journal of Logic and Computation*, 10 (2000), 677-703.

⁴⁷Gödel's proof is based on a formalization of Peano arithmetic (PA) which is taken to stand in for the relevant sections of the second edition of *Principia Mathematica* (1925-1927). It doesn't;

to see the mischief bruted by the sentence spoken by Epimenides (7th or 6th century B.C.) when, speaking as a Cretan, he announced that all Cretans are liars. It is worth nothing that in none of the mother tongues of humanity does ‘All X are liars’ mean that everyone who is a member of X is lying whenever he says anything at all. It takes quite a bit of fiddling before we can get that sentence to provide that any Cretan who says it on that occasion, is saying that what he is saying on that occasion is false. There are two points to take special note of.

- *The semantic paradoxes (so called) have been making the rounds since classical antiquity unabatted by anything remotely like the mathematical mechanisms that launched transfinite mathematics.*
- *There exists within the norms of natural speech an in-built resistance towards paradoxizing its utterances, except as amusements or devices of irony and the like.*

It is perfectly possible in well-regulated English to utter ‘What I am now in process of uttering is false’, without ruffling any feathers, save those with a taste for trouble-making. For the large bulk of neurotypical humanity, including all the Nobel Prize winners, sentences such as this are entertainments, a kind of logic-puzzle. They are the distractions that offer relief to smart youngsters with no interest in or aptitude for the rough and tumble of hockey or football. Of course, when transfinite arithmetic is one’s passion, there will be a perfectly sensible reluctance to employ tools for its advancement that carry a nontrivial risk of damaging it. Even so, what neurotypical human being *really* gives a fig whether or not mathematics is consistent? For the two hundred years after the publication of Newton’s *Principia* in 1687, every young physics student knew that the mathematics of mechanics was inconsistent, as did Newton himself right at the beginning.

The example of Newton also is instructive in another way. Newton was able to sell the idea that the basic structure of the natural world is mathematical. Much later on, economists would make the same pitch for the structure of the social world, at least that part of it that deals with the creation and circulation of wealth. In 1879, Frege did the same for the cognitive world of exact science by structuring a formula language of concept-writing on the model of arithmetic. I have come to think that the best way forward for the logic of human cognition is to put an end to this structural and methodological addiction to mathematics. In this there may lie some inducement for computer science to maintain a watching brief. For if natural-language paradox-avoidance is unneedful of the heavy-equipment mea-

but let that pass.

tures of mathematical methodology, this might occasion some reassessment of their usefulness to computer science.

I have two objectives in mind for the next two parts of this paper. In the part to follow I'll try to expose more of this reluctance of human language to paradoxize its semantically peculiar utterances. In the part after that, I'll attempt to show why it shouldn't matter to neurotypical humanity whether mathematics is inconsistent. In so doing, I'll offer a suggestion concerning why inconsistency robust information systems needn't bother us either. In the last part, I'll offer a thought or two about naturalizing the logic of human inference.

Part III

1 Medicalizing semantic nuisance

In some circles, the semantic paradoxes have come to be called 'semantic pathologies'. We see in this traces of an idea once scorned by Wittgenstein, the idea that the paradoxes are illnesses of our language.⁴⁸ If the diagnosis implied by the name is accurate, we would naturally hope that there might be some cure. Some people are of the view that the search for a cure is made more complex than one might have wished by the perceived inequivalences between and among the various paradoxes, e.g. between the Liar and the Curry. That in turn stirs the hope that there might be a comprehensive unified treatment of the semantic pathologies which bestrides these pairwise similarities and differences.⁴⁹

Some philosophers have resisted the medicalization of semantic nuisance. They lean to the view that, like all the instruments of our cognitive advancement, the instruments of human communication and linguistic construction are imperfect. Even when they work swimmingly in the general case, there are situations in which they seem to let us down. On this non-pathological view, our linguistic conventions aren't sick. What they are is *fallible*. The question that arises for fallibilists is one of a triage that assesses the damage the paradoxes do to the overall efficacy of our instruments of communication and linguistic expression. One of the triagic answers, albeit one which some philosophers see as over-hopeful, is that severally and collectively

⁴⁸Cora Diamond, editor, *Wittgenstein's Lectures on the Foundations of Mathematics*, Cambridge, 1939, Hassocks, UK: Harvester, 1976.

⁴⁹For an attempt to achieve a unified treatment, see for example, Bradley Armour-Garb and James A. Woodbridge, *Pretense and Pathology: Philosophical Fictionalism and its Applications*, Cambridge: Cambridge University Press, 2015. See also these authors and John Woods, Fred Kroon and Jody Azzouni, in a book symposium in *Analysis*, 78 (2018), 713-736.

the semantic paradoxes do no harm that matters, certainly nothing that calls for medical treatment and prognosis. In that case, the *diagnosis* is positive, but the *triage* is negative.⁵⁰

An even stiffer means of resisting the semantic pathologist forgoes the fallibilist alternative. It accepts that there is the appearance of paradox in the situations in question, but it does not accept that its diagnosis is positive for paradox. It could, however, be positive for reference-failure or equivocation, or other kinds of semantic slip-up. In what follows, I'll focus on the negative-diagnosis possibility and will turn to the negative-triage alternative later on. When a negative-for-paradox diagnosis is warranted, I shall say that there is a *dissolving predicate* for the paradox at hand. For ease of reference, what I will mean by a negative diagnosis henceforth is a diagnosis that is negative for paradox, leaving it open that it might be positive for other kinds of semantic bother. But, as we now see, there is an even more basic difference of opinion among students of semantic paradox.

2 Exposing liars without fear or favour

In many precincts of case-making enquiry, it is customary to number the premisses and conclusions of arguments. Numbering is a welcome way of simplifying backwards-reference. In what follows, I shall liberally avail myself of this labour-saving device. In the usual contexts, it functions in the following way:

- Since (*i*) and (*j*) are true, so is (*k*)⁵¹

where the parenthesized numerals refer to the statements to which they were originally applied. In those cases, the first occurrences of parenthesized numerals serve as the number-names of the statements within, the “scope-statements” let's say. In their occurrence in the bulleted passage, the three parenthesized numerals refer to the scope-statements of their original number-namings. One of the virtues of this mode of reference is a generous flexibility and a wide general applicability.

Even so, it is fair and natural to ask whether there might be limits to this desirable generality. Consider now the best-known of the semantic paradoxes, in the classical rendering of the Liar:

(1) (1) is false.

⁵⁰Wittgenstein observes in *Remarks on the Foundations of Mathematics*, that a contradiction can be ‘sealed off’ (p. 104) and so allowed to stand (p. 168). He was also attracted to the view that, strictly speaking, the appearance of a contradiction in a theory does not give rise to inconsistency, for a contradiction “is of no use; it is just a useless language game.”

⁵¹With *i, j, k* varying over the positive integers.

Straightaway we have symptoms of semantic difficulty, beginning with an abiding uncertainty about what the scope statement is about. On the face of it, the second occurrence of the parenthesized numeral-name in the scope's subject-place refers to the statement originally named by its original bestowal. But what statement is that? It is the very statement named by the original application of '(1)' in leftmost position. This leaves us in an uncomfortable position. The scope sentence refers to the statement named in the leftmost parenthesized numeral that names it. So what the scope of '(1)' says is what the leftmost '(1)' names, namely itself. There are quite ordinary situations, such as the one in which Sally asks her Daddy what 'The weather is doubtful' means. Rarely is a Daddy so clueless as to assure his small girl that what it means is what it says. "Please, Daddy", comes the impatient reply, "what does it say?" In our own case, the moral is simple. The structure

(1) (1) is false

is stricken by reference failure, albeit one of a peculiar kind. It cannot tell us what its scope says, and the reason it can't is that its scope says nothing. If it can't say anything, it can't say anything that is true or false of it. The troublesome structure is simply a bit of collateral damage ensuing from the attractive and useful referential liberties of the naming uses of parenthesized numerals in subject position.⁵² Let's call the expression 'doesn't say anything' the dissolving predicate for (1). Philosophers with a thirst for vengeance try to submit the *dissolving predicate* for (1) to the same undoing as the damage purportedly done by (1) to itself, as follows:

(2) (2) doesn't say anything.

Given that everything whatever is true or not (even the uninvited racoon in the rear garden),

(3) (2) is true or not.

If (3) is true, then

(4) (2) doesn't say anything

and so

(5) (2) can't say anything true.

⁵²Wittgenstein also thinks that the self-referential paradoxes don't actually refer, but for reasons having to do with the distinction between algorithmic and metamathematical proofs in calculi of the PM kind. See *Remarks on the Foundations of Mathematics*, VI, §§28. My reference-failure objection arises from the referential conventions of communities of natural speech.

Very well, then,

(6) (2) is not true.

If (6) is true, then

(7) What (6) says is that (2) is not true.

But it does not follow from (6) that what (2) says *is* true. What (2) says is nothing. So

(8) Our dissolving predicate withstands the molestations of Revenge.⁵³
Q. E. D.

The dissolving predicate for the Liar is rooted in the claim that the structure

(1) (1) is false

is guilty of a strange kind of reference-failure, that it fails to tell us what ‘is false’ is predicated of. If reference failure there be, it is a wise thing indeed to add the qualification that it is a strange kind of one. For on the face of it, the parenthesized numeral ‘(1)’ *does* refer. It refers to its own scope ‘(1) is false’, with that same numeral appearing in subject place.

(9) What, then, does (1) refer to?

There would appear to be just two answers to this question:

(10) It refers to the leftmost occurrence of ‘(1)’

or

(11) It refers to what the leftmost occurrence of ‘(1)’ number-names.

If (10) holds, then

(12) Since the leftmost occurrence of ‘(1)’ is a name, it cannot be that it is false (or true either).

If (11) holds, then

(13) The leftmost occurrence of ‘(1)’ refers to its scope.

But

⁵³This retracts an earlier claim to the contrary on p. 244 of *Paradox and Paraconsistency*.

- (14) This is the very answer that didn't answer question (9) in the first place.

What we have here is a failure of backwards reference, owing to the failure of an original one.

Some philosophers find themselves drawn to the view that if a grammatically well-formed statement-making construction says nothing, then it can't be understood and can't have meaning. If this were so, it could threaten the dissolving strategy with unforeseen consequences too costly to bear. For clearly the scope of (1) can be understood and equally does have meaning. Anyone at home with English and French will have no difficulty in making a faultless translation of (1) in flawless French. Accordingly, we would appear to have it that (1) says something after all, and what it says is very bad news for it. So much, then, for 'says nothing' as a dissolving predicate.

Some time ago David Kaplan introduced a distinction that bears nicely on this point, and bequeathed it the name "content v character".⁵⁴ In the present context, Kaplan's distinction offers paradox-deniers the promise of having their cake and eating it too. In Kaplan's own terms it allows the deniers to ascribe character to the likes of (1) and (2) but deny them *content*. Thanks to their character, (1) and (2) are meaningful and perfectly understandable. They are *lexically and syntactically* intelligible. But they lack semantic or propositional content and therefore make no statements and express nothing that bears a truth-value. They are *propositionally* unintelligible and say nothing true or false.

Here, too, there could be some push-back. Suppose a teacher of French sees one of his students at a ball game at which the teacher has arrived a bit late. "What's the score, Charlie?", he asks, and Charlie replies "There's no score yet." "Can you say that in French, Charlie?" and, good student that Charlie is, he says it perfectly in French. The point of this parable is that Charlie knows how to say "There's no score yet" in French, but that doesn't mean that Charlie knows what it says in French. Of course, if he knows what it says in English, he does know what it says in French, never mind whether he himself knows how to say it in French. If not, not. There's little doubt that he does know what it says in both places. But (1) and (2) are conspicuous exceptions, even for English-French bilinguals.

We may now safely conclude that the dissolution strategy is securely reinstated if the Kaplan distinction is allowed to stand. Needless to say, there are critics who

⁵⁴David Kaplan, "Demonstratives", in Joseph Almog, John Perry and Howard Wettstein, editors, *Themes from Kaplan*, pages 481-563, Oxford: Oxford University Press, 1989. Although appearing only slightly less than thirty years ago, pre-publication versions were making the rounds well before then.

accept the conditional but deny its antecedent. These are the philosophers who are phobic about abstract entities. Think here of Quine who will have no truck with them until, as with sets, he needed them for the mathematics that's needed for science. Quine's scepticism has a second source in his purported reservations about translation.⁵⁵ All I will say of these scepticisms is that, first, Quine never had the least worry about the soundness of the map from "The dog is hungry" to "Le chien a faim", and secondly that, for present purposes, I'm content to drop propositions in favour of truth-apt sentences.

Problems, even so, do have a way of piling on. Often enough, proposition deniers also are sceptical about truth and falsehood, as with disquotational, deflationist, prosentential, redundantist, and ideal-limit philosophers of language. There is no time to litigate these matters here, beyond saying that if these scepticisms were bundled together and thrown into the rumble between paradox loyalists and paradox deniers, it would disrespect the plain historical fact that all the semantic paradoxes arose under the implicit and tacit assumption that none of these scepticisms had a seat at that particular table. Of course, it can be said that they are seat-seekers now. Yes, but if admitted, the whole paradox literature could be well and goodly shaken up, perhaps even to the point at which it is up-ended by its own semantically troublesome armchair scepticisms.

In any event, all this is helpful in shaping is not my task here, which is to proceed with the semantic paradoxes under the historical assumptions about propositions, semantic content, translation and truth. So positioned, the dissolution strategy remains a live option.⁵⁶

Where are we now? I should try to be clear about my present position. It is not necessary that I prove that there is no possible way of pinning the Liar on English. It suffices for my purposes to establish that the regulatory regime for predication in English and all the other languages of humanity embeds an in-built resistance to semantic self-paradoxizing. That, if true, would be telling. It would tell us that human languages aren't structured for silly self-dalliance, and it would remind us of Hewitt's insight that sentences such as 'What I am now in process of uttering is false' are no practical interest for computer scientists.

⁵⁵W. V. Quine, *Word and Object*, New York: John Wiley, 1960; chapter 2 on the "Indeterminacy of radical translation".

⁵⁶It is notable that the Armour-Garb and Woodbridge treatment of semantic pathology attempts to make profitable use of these very scepticisms, by way of a philosophical fictionalism about the attendant semantic afflictions, in which the concept of fiction carries a pretendist interpretation. I resist the pretendist twist in my *Truth in Fiction: Rethinking its Logic*, volume 391 in the Synthese Library, Cham: Springer, 2018. See also my remarks in the previously mentioned *Analysis* book symposium on *Pretense and Pathology*, "Pretendism in name only", 713-718.

3 Properties stood for but uninstantiated

Real or apparent, the Liar's reference-failure is similar to what Ryle in a somewhat different but related context calls a "namely-rider" mishap, the error of ceaseless circling, provoking the endless irritations of the semantic run-around.⁵⁷ It would repay us to pause awhile with the case that was on Ryle's mind, to see whether my employment of the namely-rider mishap for the Liar coheres with Ryle's application of it there, concerning a problem advanced by Grelling and Nelson.⁵⁸ Before getting on with it, it is also important to record that neither Grelling nor Nelson had the slightest interest in the length of English expressions. Their whole focus was on the formalized arithmetic of ordinals.

An expression is heterological when it does not instantiate the property it stands for. For example, 'is a mile long' stands for the property of being a mile long, but isn't itself remotely close to being a mile long. On the other hand, 'is an English expression' stands for a property it clearly instantiates. The former expression is heterological. The latter is autological. Consider now

- (1) (1) is heterological.

Before moving on, a reader alert: The naming-by-numeral conventions of this paper reference by numeral-name is *intra-sectional*. If its scope is true, then

- (2) (1) *stands for* the property of being heterological but does not *instantiate* it.

To see what that property is, consider again our two examples.

- (3) The property of being an English expression is uninstantiated by any expression standing for this property that is not an expression of English.
(4) The property of being a mile long is uninstantiated by any expression less than a mile long standing for the property of being a mile long.

It becomes clear that the property shared by (3) and (4) is not the one displayed in either place. What, then, is it? As I read him, Ryle might have been inclined to the view that if there were an answer to this question, it could only be that

- (5) The property common to (3) and (4) is the property H of any property stood-for by an uninstantiating expression of it.

⁵⁷Gilbert Ryle, "Heterologicality", *Analysis*, 11 (1951), 61-69.

⁵⁸Kurt Grelling and Leonard Nelson, "Remarks on the paradoxes of Russell and Burali-Forti", *Abhandlungen der Fries'schen Schule*, (Neue Serie), 2 (1907/1908), 300-334.

H, of course, is not a property of expressions. It is a property of properties of expressions meeting the salient conditions. Accordingly,

- (6) *H* is not heterological.
- (7) It isn't autological either.

That being so, we're back to expressions. What does the property 'heterological' stand for? The answer is

- (8) 'Heterological' stands for no *given* property of each and every heterological expression.

If this lands us in the namely-rider problem envisaged by Ryle for heterologicality, it is not the endlessly circulating problem I ascribed to the Liar earlier on. Even so, we are not quite finished with this. For consider again

- (1) (1) is heterological.

Like anything else, (1) is true or not true. If true

- (9) There is a property that (1) stands for but does not instantiate.

If not true,

- (10) There isn't a property that (1) stands for but doesn't instantiate.

But (10) is plainly false, as witness

- (11) (1) stands for and instantiates the property of being an expression of English.

Even so, given the unspecificity of the indefinite article in (9) and (10), (1) itself doesn't say anything. So

- (12) There is no contradiction in (1). Q. E. D.

The predicate 'doesn't say anything' dissolves the heterological paradox, and does so without exposing itself to avenging self-destruction.

- (13) (13) says nothing

may have Kaplanian character but it embodies no truth-valued semantic content in virtue of which it says anything at all. There is nothing it says that is true or false. Attributions of paradox are spurious, as reflected in a negative diagnosis of indications otherwise.

The namely-riderliness of 'heterological' in Ryle's *Analysis* article is well set out in a formulation of Geach:⁵⁹

⁵⁹P. T. Geach, "Ryle on namely-riders", *Analysis*, 21 (1961), 64-67.

“heterological’ lacks the property for which it stands, namely, that of lacking the property for which it stands, namely, that of lacking”

Geach demurs. He thinks that peculiarities of ‘it’ deny it here the status of a “pronoun of laziness”, in virtue of which

(14) ‘heterological’ lacks the property for which it stands

would be equivalent to

(15) ‘heterological’ lacks the property for which ‘heterological’ stands.

It is not my purpose here to dig into the Geach-Ryle position regarding the equivalence-claim. It suffices for present purposes to note that

(16) (1) implies no contradiction, regardless of its real or apparent namely-riderliness.

The reason why is that although (1) is lexically and syntactically intelligible, it is propositionally empty. I would like to turn now to a further version of semantic paradox, purportedly genuinely contradictory sort, but not self-referential.

4 The meaning of ‘the following’

Consider an infinitely long sequence of sentences:

- (1) All the following sentences are untrue.
- (2) All the following sentences are untrue.
- ⋮
- (*n*) All the following sentences are untrue.
- ⋮

Some people follow Stephen Yablo in thinking that since there is no consistent way to assign truth-values to these sentences, the sequence generates a genuine paradox.⁶⁰ Yablo makes a second observation. The paradox owes nothing to self-reference. I demur from both these claims. The Yablo sequence begets no contradictions. From which we have it that for inconsistent truth-value assignment is none at all. What, then, does (1) actually say? It says that

(3) All the sentences on the list below are untrue.

⁶⁰Stephen Yablo, “Paradox without self-reference”, *Analysis* (53), 1993, 251-252.

Below *what*?

- (4) All the sentences on the list below this very sentence are untrue.

So, contrary to Yablo,

- (5) The self-referential character of (1) is fully discernible there.

Suppose we had an infinitely long sequence of number names prefixing in each case an empty line. Suppose that we were told that the spaces are place-holders for statement-making declarative sentences. It would be clear that no occurrence in the form ‘(n): _____’ has semantic content. When these number-names prefix the *filled-in* lines, those sentences will identify the places where further occurrences of that same sentence can be found. Come back for a moment to

- (1) (1) All the following sentences are untrue.

Granted that we know by inspection what (1) is, there is no way of knowing what all those other sentences are until we examine the places where they themselves crop up. What (1) does is specify a place below. It does not specify the sentence it occupies. (That was Yablo’s own separate doing.) It matters here that ‘untrue’ is satisfied by anything whatever that’s not true, including our unloved racoon in the back garden. But (1) can’t even tell us what *kind* of sentences we’ll find when we check out the specified places. Only upon examination of what we find there do things start to click. This helps us to see that

- (7) (1) itself is referentially opaque.

I mean by this that there is nothing in (1) that specifies the statements that *Yablo* has in mind. In other respects, it is transparent that what (1) specifies is *inter alia* the name of the place where a sentence is to be found, and at that place its *number-name* does indeed vacuously satisfy the predicate of (1) much in the way that the racoon also does. Again, this is not what Yablo is after.

The same confusions crop up at every filled-in line of the sequence. Nothing that occurs there manages to specify the statements that Yablo wants to lay his hands on thenceforth. To do that, he’ll have to do it on his own. Accordingly

- (8) Each line in the sequence is *Yablo-opaque*. Indeed the sequence itself is *transfinitely* Yablo-opaque, opaque without finite end.

How bad a problem is Yablo-opacity? How is to be *triaged*? Some will say that since it is not difficult to remedy, it really isn’t all that troubling. After all, hasn’t Yablo himself provided the very sentences that he’s interested in? To this I reply that

(9) Even with Yablo's help, it remains the case that each occurrence of the expression 'all sentences that follows this one in this sequence' is in its own right Yablo-opaque.

At this point, it would only be natural to seek some relief from paradox-deniers in the well chartered harbours of context. In a long-ago example from Strawson

(8) I am here now

is as it stands, lexically and syntactically intelligible but propositionally empty otherwise. Truth-value bearers require the inputs of context, in this case, of occasions of utterance. In one of them, content expressibility is achieved by knowing the referent of 'I'. In other, more fully realized ones, the referent of 'here' would also be known. Other examples abound, for example,

(9) That fat man in the doorway is another logician from there.

Here, too, referential awareness of 'that' supplies semantic expressibility, and referential acquaintance with 'doorway' and 'there' enlarges it. Other examples hinge upon private unspoken interpersonal conventions. When Sally says to her husband Hal,

(10) "Wow, this is getting to be a good party!",

Hal will reply,

(11) "Right, hon, you get the coats and I'll bring the car around."

Anyone within attentive earshot of that exchange would twig to the convention that animate it, and would be able with ease to discern its propositional content.

Why wouldn't we claim the same for the Yablo-sentences, each expressively realized by Yablo's own specifications of them? Why aren't those specifications expressibility-contextualizations of them? Isn't each line filled in by Yablo an proposition-yielding context for those that follow? I readily concede that Yablo's contributions identify the sentences at each number-named line. But given the semantically unrealized indexicality of each of them, together with the infinite length of the sequence, referential fulfillment is encumbered by transfinite *postponement*, in consequence of which the whole sequence is rendered semantically contentless. It really doesn't matter whether Yablo's specifications are contexts. What matters is that they don't achieve propositional lift-off.

Each line of the Yablo-sequence has Kaplanian character. Each can be faultlessly rendered into faultless French. Any bright six-year old can read the scope-sentences

without difficulty. What they lack is Kaplanian content. There is no truth-value bearer to which any line gives expression. Nothing is said at any line. All intimation of paradox is spurious. There is no contradiction generated by the Yablo-sequence. And, as we've already seen twice before, the predicate that dissolves the purported paradox is itself immune from vengeful demolition. Even if there were a credible answer to this dissolution, it will have taken some artful dodging to bring the paradox off. Sometimes artful dodging can be righteous, as with — as some think — the generation of the Gödel sentence for formalized Peano arithmetic. But in the present case, it's all wind-up and no pitch, all technique and no show.

5 Curry's Paradox⁶¹

Consider

- (1) If (1) is true, then S , for arbitrary S .
- (2) If (1) is true, then if (1) is true, then S .

Eliminating the repetition of the antecedent,

- (3) If (1) is true then S .

So

- (4) (1) is true.

Therefore, by modus ponens,

- (5) S .

Since S is any statement whatever, reapplication of this reasoning gives its negation.

- (6) $\neg S$
- (7) Contradiction.

The Curry is notable. Although it is generated by what Poincaré and Russell call vicious self-reference, in so doing its generation doesn't rely on negation.⁶² Even so, it lies fully exposed to the dissolving predicate "says nothing". It is expressively impotent. Accordingly,

- (8) There is no paradox here. Q. E. D.

⁶¹Haskell B. Curry, "The paradox of Kleene and Rosser", *Transactions of the American Mathematical Society*, 41 (454-516), (1942). See also S. C. Kleene and J. B. Rosser, "The inconsistency of certain formal systems", *Annals of Mathematics*, 36 (1935), 630-637.

⁶²True, $\neg\neg S$ at (6) is the negation of S at (5). But it's a negation that's been generated. No negation played a role in bringing that generation about.

6 Berry's paradox (again)⁶³

Consider the following sentence

- (1) 'The least integer not describable in fewer than twenty syllables' is itself a description of nineteen syllables.

So

- (2) The least integer not describable in fewer than twenty syllables is describable in fewer than twenty syllables. The quoted expression is such a description, and it has only nineteen syllables.

As we saw, the mathematical significance of this paradox requires further elaboration, but for present purposes we needn't set it out here. It suffices to note that (1) is true if the quoted passage is indeed a description, and that (2) is untrue, owing to the reference-failure of its subject term. There is no integer that is not describable in fewer than twenty syllables, or even nineteen. The description 'is odd or even' is satisfied by every integer whatever in five syllables. There is no contradiction between (1) and (2) The problem is reference-failure.

- (3) There is no paradox in Berry. Q. E. D.

I said in Part B that it is possible to show in the semantics of everyday English an analogue of the uncomputability of the Berry number by a Turing machine. Its analogue is reference failure in English.

My overview of how the paradoxes fare in natural language has been neither exhaustive nor highly detailed.⁶⁴ Even so, the coverage is a fairly representative sample, and the discussion has been detailed enough to achieve my objective. Again, that objective was not to prove that natural language is immune from paradox, but rather to show how naturally resistant it is to it.

What Carl Hewitt achieves for the language of CDL by technical provision, human languages try to achieve naturally. There remains the difference that Hewitt's type constraints achieve outright immunity, whereas natural language's mechanisms certainly achieve resistance even if not strict immunity. It is a real difference, and certainly not one to be slighted. But the similarity is also telling. It tells us that

⁶³Attributed by Russell to H. G. Berry of the Bodleian Library. It is a simplified version of Richard's paradox. Bertrand Russell, "Les paradoxes de la logique", *Revue de Metaphysique et de Morale*, 14 (1906), 627-650

⁶⁴For example, none of the paradoxes discussed in Terence Horgan's *Essays on Paradoxes*, New York: Oxford University Press, 2016 is among the ones surveyed here.

there is no *natural* place for paradox in computer science, and stylistics apart, none in natural language either.

Part IV

1 Reference-failure

As we see, the dominant flaw in the cases inspected so far is not paradox, but *reference-failure*. Perhaps we should look at this more closely. Reference-failure is one of the most discussed of semantic infelicities, and what remains the best-known example of it is Russell's golden oldie,

- (1) The present king of France is bald.

It matters that (1) crops up in the quite particular context, in which Russell is re-considering his slightly earlier thesis on denoting concepts.⁶⁵ Contrary to received opinion, the motivating factor in "On denoting" was an epistemological one. Russell distrusted the reliability of knowledge by description. Russell thought that knowledge by acquaintance was much the more solidly secure of the pair. Better, he thought, if our apparent knowings by description could be reconfigured as disguised forms of knowings by quantified acquaintance. Only at that point do the lexical and syntactic characteristics of definite descriptions come into play, by which sentences in the form

- (2) The x that is F is G

is truth-preservingly rewritten as

- (2') One and only one thing is F and that thing is G .

As any competent speaker of English will know, (2') is a sufficient condition of (2)'s truth, but hardly a necessary one. So Russell's putative equivalence fails for English. Mind you, Russell hadn't advanced his contextual eliminability thesis for English, but rather for the formal language of the new pure logic,⁶⁶ inspired by Frege and further developed by Whitehead and Russell in *Principia Mathematica*,

⁶⁵Bertrand Russell, "On denoting", *Mind* N. S. 14 (1905), 479-493.

⁶⁶Strange to say, given all the fuss by logicistically-minded philosophers on the importance of re-proving the truths of arithmetic in "pure logic", very little has been written about what it takes to make a system of logic pure. Russell himself is an exception of sorts. In his 1911 lecture at l'École des Haute Études Sociale, Russell advanced some views on what he called "propositions of pure logic". The paper appeared in *Revue de Métaphysique et de Morale*, 19 (1911), and under the

purpose-built for the foundations of arithmetic.⁶⁷ This is an important qualification. Russell's theory of definite descriptions is one of analytic philosophy's most widely misunderstood doctrines. See here Richard Cartwright, "On the origins of Russell's theory of descriptions", in his *Philosophical Essays*, pages 95-133, Cambridge, MA: MIT Press, 1987. Regrettably the essay was not published earlier. "It emerges, writes Cartwright, "from repeated effort on my part to cover in a [teaching] term the history of analytic philosophy from 1879 to 1929. I always had trouble getting beyond 1905." (p. XX).

2 The "paradox" of fiction⁶⁸

My purpose in discussing 'the present king of France' here is to register the observation that if anyone in a village pub one pleasant early evening in 1905 had said that the present king of France is bald, he'd have been met with open-mouthed incomprehension. France had no king in 1905, and no bald one either. Suppose some nearby wag calls out,

(3) "Don't worry lads, at least Hamlet's uncle was king of Denmark!"

to amused chuckles of approval. People with a reasonable exposure to Shakespeare's *Hamlet* would have no difficulty in knowing to whom (3) refers, even when also knowing that

(4) Uncle Claudius, being fictional, doesn't and never did exist.

If the pub's patrons were left to call the semantic shots, there would have been a comfortably recognized distinction between referring expressions that don't refer to anything at all and referring expressions that refer to things that don't exist. According to our pub-semantics, reference-failure is not occasioned by nonexistent

title "The philosophical implications of mathematical logic", revised by and translated by P. E. B. Jourdain, appeared in *The Monist*, 22 (1913), 481-493. The term 'propositions of pure logic' appears in the second of Russell's Lowell Lectures of 1914, and published in Russell's *Our Knowledge of the External World*. London: Allen and Unwin, 1914. It appears in the singular in the fifth of his articles in "The philosophy of logical atomism, *The Monist*, 28 (1918), 495-527; 29 (1919), 33-63, 190-222, 344-380. An able discussion of the subject can be found in Richard Cartwright, "Propositions of pure logic", in his *Philosophical Essays*, pages 217-236, Cambridge, MA: MIT Press, 1987. The paper originated in 1982 as a talk to the Cambridge Moral Science Club and, save for an abstract in the *Journal of Philosophy* that year, was previously unpublished. Not surprisingly, Russell's idea of purity (for logic) emphasizes the wholly general and thoroughly formal.

⁶⁷A. N. Whitehead and Bertrand Russell, *Principia Mathematica*, Cambridge: University of Cambridge Press, 1910-1913.

⁶⁸Noël Carroll, *The Philosophy of Horror*, New York and London: Routledge, 1990.

referents. It is occasioned by an expression's failure to achieve referential lift-off in the first place, brought about in turn by the absence of a referent. Pub-semanticists are agreed on three further matters.

(5) There are lots of things that don't exist, and Claudius is one of them, and so is Zeus, and Sherlock Holmes is yet another.

(6) Claudius and Zeus are as different from one another as Donald Trump is from the Pope Francis, as indeed Sherlock and Odysseus also are on a different scale of comparison.

(7) There are lots of true statements about Claudius, including that he was Hamlet's mother's second husband and Hamlet's uncle; and lots of false ones, too, including that he was a Cistercian friar. There are lots of truths about Sherlock including the true one that he shared rooms with John H. Watson in 1880s London and the false one that their rooms were in Cheyne Walk.

With regard to these matters, pub-semantics speaks for all of neurotypical humanity. Anyone who speaks a human language will favour in his own home language counterparts of those in view here.

By and large, philosophical semantics is another story altogether. A large majority of philosophical semanticists open their business day with a standing and sometimes unspoken denunciation of the likes of (5), (6) and (7). They think that these sentences fail to refer, and in so thinking, overlook the actual problem they create for them, to wit:

- *Those sentences don't fail to refer; they fail to fail to refer.*

It seems passing strange that philosophical semanticists should stand so easily ready to indict all of humanity for its wall-to-wall semantic misfeasance. They rest their theoretical confidence in so doing on the following assumptions, which some philosophers of language regard as axioms. For ease of reference, I'll call these BLERTs, the Basic Laws of Reference, Existence and Truth:

- I. *The Everything is Something Law*: Everything whatever is something or other.
- II. *The Existential Load Law*: Reference and quantification are existentially loaded.
- III. *The Truth Law*: No sentence discomplying with either I and II can be true.

The first BLERT is true, but if the other two also were, we could easily see the real trouble that the Russell sentence fell into:

(8) The Russell sentence violates the Everything is Something Law.

It is noteworthy that this law alone lays no glove on Hamlet, Claudius, Sherlock or the others. Meinongian semanticists are happy to comply with this law, thinking, as they do, that the golden mountain isn't nothing, or Sherlock Holmes either.⁶⁹ Establishment semanticists can't have their way with our (5), (6) and (7) without also pledging themselves to the Existential Load and Truth Laws.

There is plenty of collateral evidence in linguistic usage to confirm the judgement that the Russell sentence violates the Everything is Something Law. It is impossible to say anything true in English to the effect that the present king of France is nothing at all or that there is no one whomever who is one and the same with him. What I mean by this is that these can't be both said and be true when expressed in what Carnap called the material mode, but only in the formal mode along the lines of

(10) Nothing whatever satisfies the predicate 'is the present king of France.'

Sherlock and Hamlet sentences are subject to no such constraint. That Sherlock and Hamlet is each someone who doesn't exist is effortlessly sayable in the material mode and, by the lights of pub-semantics and the beliefs of everyone whomever who's not a philosopher, is true.

We are now in a good position to know why the problem of negative essentials would be a pseudoproblem if the Existential Load presumption were given up.

(11) Sherlock doesn't exist

would pose no problem at all, whereas

(12) The present king of France doesn't exist

would be a problem, but the problem wouldn't be a negative existential one. (12)'s problem is reference-failure. (12) doesn't tell us what or who instantiates the non-existence predicate. If Existential Loading is dropped, then there is plenty of room for other things to be true, for example, our old friend

(13) There are lots of things that don't exist.

⁶⁹I stand with Meinongians in thinking that Hamlet is not nothing, but I don't accept their ontology of objects. I fail to see the ontic appeal of the thing such that nothing whatever is it. See here Terence Parsons *Nonexistent Objects*, New Haven: Yale University Press, 1980; Dale Jacquette, *Meinongian Logic: The Semantics of Existence and Nonexistence*, Berlin: Walter de Gruyter, 1996; and Francesco Berto, "Modal Meinongianism and fiction: The best of three worlds", *Philosophical Studies*, 153 (2011), 313-334. My differences with Meinongians are further explored in the already-mentioned *Truth in Fiction* [73].

Thanks to a peculiarity of the English lexicon, we can rewrite (13) equivalently as

(13') There *exist* lots of things that don't exist.

It is rather striking how readily smart philosophers see a contradiction here. Apparently they think that the quantifier 'there exists' and predicate 'exist' are cognates of one another. They aren't. 'There exists' functions in English as 'Il y a' does in French. 'Il y a' is unrelated to the verb 'exister'; neither is German's 'es gibt' related to 'existieren'. It is the same way in English, too.

At this point it stands out that but the first the BLERTs are in full-fledged discompliance with a wealth of empirically discernible features of human linguistic behaviour and all like indications of commonplace belief. Since early in the Attic revolution of *logos*, philosophers haven't stinted in ignoring empirical evidence. Some logicians are of the view that the time is long due to call this *hauteur* to account.

Part V

1 Empirically sensitive semantics for naturalized logics

The main problem with the Existential Load and Truth presumptions is their high cost. If given free rein (perhaps I should also say reign), the highest cost would be Big-Box Scepticism. Humanity at large would be described as not knowing a great deal of what it thinks it does know, especially when those offensive presumptions are supplemented by other assumptions equally offensive to the facts of lived experience. For example, if anti-realism were also given its head, then a mathematical statement in the form

(1) 'There exists an x such that F '

would be false every time it was uttered. Further casualties would lie in wait. Nothing said about planet Vulcan, Caesar, Oliver Twist, the largest positive integer, the non-planet Venus, Atlantis, not even that none of them exists, has a chance of survival, short of changing the subject in some more or less empirically unendorsed way. Think here of 'the purported planet Vulcan'. What do we think *its* referent is? Is it the purported planet but no real one? Which one is that? Some say that all this trouble is caused by face-value readings of the offending expressions, and that salvation awaits when read in some other inequivalent way. In so saying, Big-Box Scepticism is redoubled. Not only do we know nothing of things on face-value readings of their reporting sentences, but we remain wholly in the dark about what the face-saving inequivalent alternative readings would be.

There are two prominent sanctuaries in which empirically false semantics seeks harbour. One says that since semantics is the *logic* of reference, existence and truth, and logic is a formal discipline, it owes no allegiance to empirical reality. The other, partly in answer to the first, is that the empirical falsity of the Existence and Truth laws is compensated for by their *normative* authority. Descriptively false of beings like us, they are descriptively valid constraints on reference, ascription and truth-attribution by *ideally rational reasoners*, and thus are descriptively true of them and normatively binding on us. Needless to say, neither of these safe spaces from empirical unruliness is itself empirically endorsed. From which we are at liberty to conclude that

- *No philosophical theory endorsing the Existence and Truth empirical falsehoods has any credible claim to be an authoritative theory of human language use, belief and reasoning.*
- *Any credible such theory will, with requisite circumspection, have to be an empirically sensitive one.*

What this means in turn is that

- *At a minimum, an empirically sensitive philosophical theory of empirically discernible human linguistic and inferential behaviour must attend to the observable data and performance regularities of such behaviour, must carefully analyze those data, and heed the best-confirmed and most successfully replicated empirical sciences of such things, in the absence of proper cause not to.*

The expectation that a philosophical theory should respect empirical reality strikes many philosophers as hopelessly naïve. Don't these empirical-sensitivity theorists understand that philosophy is not an empirical discipline? Of course, not all philosophers over the centuries have endorsed this line, but it cannot be overlooked that from the early stirrings of the revolt of *logos* against cosmogony, a foundational principle of Western intellectual thought has been the fragility of the appearances of things in disclosing the attendant realities. True as that assuredly is (and a good thing too), it remains the case that since Frege's revolt against whatever passed for the mainstream logic of his own day, philosophers have been much in his shadow about matters that strike their own, not necessarily mathematically foundational, interests.⁷⁰

⁷⁰A detailed defence of the case for an empirically sensitive naturalized logic of human reasoning can be found in Woods (2013/14).

A case of particular moment to what interests me in this section is how some post-Fregean philosophers have framed their thinking about the ontological commitments of not just mathematics, but of the natural sciences as well. Let Quine be our guide here, whose long and distinguished career was marked by an ontological more feigned than felt. Quine saw physics, the strongest of the natural sciences, as philosophically diminished by the careless deployment of quantifiers over ontologically questionable domains of discourse. On his most-wanted list of offenders, intensional objects ranked high, followed in no particular order by non-well-individuated ones, abstract ones such as properties, propositions, modalities, dispositions, and more. Sets, however, would be given free passage. They are needed for mathematics, and mathematics is needed by physics. Quine's mission for physical theories was to rid them of the blight of ontologically over-wrought formulation by regimenting their languages into a canonical notation in which they'd find truth-preserving re-expression in a first order extensional language, whose *empirical* objects would be specified by applying the devices of quantification and identity to the theory's empirical predicates. In the end, Quine reluctantly yields to the pressures of quantum mechanics occasioned by the refusal of its objects to play first-order ball. The model theory of Quine's logic derives partly from Frege and mainly from Tarski. So it stands to reason that Quine would impose the BLERTs on the philosophically best languages for science, that is, the ontologically circumspect ones, the ones that kept their ontological commitment to a bare minimum.⁷¹

Quine never thought for one minute that his ontic cleansing of physics was a contribution to physics. Quine never thought that once rendered into canonical notations physics would be made to get better than it was at the time. Neither did he think that the demonstrations of physics would be made richer and more numerous by the use of any of his metaontological claims as additional premisses. Quine's implied view is that the ontology of science doesn't call the shots for empirical science, anyhow for the mature ones. Equally, none of those demonstrations would fall under the gravitational pull of Quine's metaontological framework for science. Quine's framework would carry none of the water for physics' revelations of the natural order. What ontology does is to "stage" the house of science to make it more appealing to the ontologically fastidious consumer.

All this lies in stark contrast to the approach taken by logicians and philosophers of language to the empirical phenomenon of language-use in human communities. One might think that the appropriate approach for a philosopher to take to the empirical sciences of language use is some or other variation of Quine's metaonto-

⁷¹See further to this my "A captious nicety of argument", in Lewis Edwin Hahn and Paul Arthur Schilpp, editors, *The Philosophy of W. V. Quine*, volume 18 of the Library of Living Philosophers, second edition expanded, pages 687-725, Chicago and LaSalle: Open Court, 1998; first edition 1986.

logical approach to physics. One might think, that is to say, that the philosophy of language would stand to empirical linguistics as Quine's philosophy of physics stands to physics itself.

No fact of the matter about nature's physical world owes allegiance to any of Quine's "staging" preferences. No fact of the matter about nature's social world owes allegiance to any of a philosopher's like preferences either. Why would we so much as countenance the contrary idea? Why would we accord it establishment status? Why wouldn't we approach in an empirically sensitive way the apparently paradox-generating linguistic behaviour of human life on the ground.

Concerning why such clever people as paradoxers have been so easily led to favour paradoxical diagnoses over failed-reference ones, my conjecture is that

- *Most by far are in the thrall of the BLERTs which so thoroughly misdescribe the referential goings-on in the speech communities of humanity.*

2 Consequence

Perhaps it's time to give a name to the version of naturalized logic that I espouse. "NLog" fits the bill nicely. A motivating consideration for NLog is the trichotomy that distinguishes

- consequence-*having*

from

- consequence-*spotting*

and those two from

- consequence-*drawing*.

Consequence-havings obtain in logical space. A entails B irrespective of anyone's awareness of it. Consequence-spotting occurs in psychological space, somewhere inside someone's head. Consequence-drawing also happens in the head, in a sector of psychological space that regulates inference and belief-revision. In 1970 Gilbert Harman argued, in effect, that the obtainment conditions on consequence-having or entailment misfire when construed as rules of inference or consequence-drawing.⁷² Taking *modus ponens* as our example, consider the case in which I believe that if p then q and have no overt position on either p or q . One day, information arrives

⁷²Harman's 1970 paper is incorporated into his *Change in View: Principles of Reasoning*, Cambridge, MA: MIT Press, 1986; chapter 1.

which causes me to believe p . In its weakest form, *modus ponens* guarantees that detaching q is allowed. What happened in fact that, in being forced to consider q , I came to think that q is false. So rather than inferring it, I denied it, and also gave up on ‘If p then q ’. This, says Harman, is a perfectly rational accommodation to make. In the section to follow I will advance a proof that *ex falso* is true in English, that is, that a contradiction entails every sentence whatever of the language in which the contradiction is formulated. The distinction between consequence-having and consequence-drawing is crucial here. Given that the obtainment conditions on entailment cannot reliably serve as rules of inference, we have room to concede that whereas *ex falso* is a valid condition on entailment, it is not a valid rule for drawing inferences from contradictory inputs. We have it, then, that a full-service logic for the deductive consequence-logic is one that gives full value to the difference amidst the similarity of havings, spotting, and drawing.

This brings me to another point of similarity between NLog and IRDL. IRDL also acknowledges a tripartite distinction for *inference*, *chaining* and *entailment*. Like NLog’s, IRDL’s inference is the weakest of the three and entailment the strongest, with consequence-spotting assuming intermediate strength in NLog and chaining doing the same in IRDL. Hewitt notes that in classical logic his trichotomy collapses, as indeed does mine in most versions of it. Further inspection reveals however, that IRDL’s three relations aren’t one and the same with NLog’s three. But that IRDL has three and NLog does too, and that they agree as to the extent of *output* strength, indicates a structural similarity of the two approaches.⁷³

3 The negative triage option

In bringing this paper to an end, I turn briefly to the negative triage option. Unlike a negative diagnosis, a negative *trriage* acknowledges paradox but makes light of it. It asks, “what, if it held, how would the Yablo paradox really matter, and if so for what? What would be the good of showing it and what, indeed, the good of knowing it?” The semantic regularities of English and all other mother tongues of humanity are products of evolution. They endure and adjust by trial and error, enabling those who speak them to prosper in ways that enrich human cognition. The rules of English aren’t algorithms. They haven’t been stress-tested for all possible situations, and certainly not for all artful-dodging ones.

Human beings are fallible creatures. We lack the means of universal fact-checking, we lack the capacity for general-purpose inconsistency-searches, and we

⁷³Further details can be found in my *Three Grades of Deductive Involvement: A Full-Service Logic for the Consequence-Relation*, forthcoming in Springer’s Synthese Library.

also lack the capacity to error-proof our error-corrections. We do make errors, lots of them, but overall our cognitive devices have made us right enough of the time about enough of the right things to survive and prosper and to build the great cathedrals of mediaeval France. Our fallibility also has a semantic face. Our truth- and other semantic ascriptions sometimes go awry. When they do, we often heed the advice of Groucho Marx's doctor, of whom it is reported that when Groucho complained that his arm hurt like billy-o when he moved it like so and so, the doctor advised "Stop moving your arm like so and so. That will be \$ 500.00 please." When our ascriptions go semantically awry, the lesson for us is to avoid making them.

Needless to say, there will be push-back, spurred by the conviction that contradiction is not something to be taken lightly. Telling people not to keep on falling into them won't change the damage they've already fallen into after some artful rigging of the very semantic devices that make ascription and truth-value bearing possible. Wouldn't the damage have been already done? Or more gently, doesn't the very fact that it *could* be done show the sorry state our semantic devices are in?

In reply to this reply I ask in return, what is the extent of the damage? This is not not place to take the matter up in any detail.⁷⁴ For now, I'll settle for the simple empirical fact that neurotypical human beings are little disturbed by even the recognized contradictions they've fallen into and don't know how to dissolve.⁷⁵ As mentioned earlier, Newton's calculus is inconsistent. It embeds a contradictory description of infinitesimals, without damage to their operational importance for mechanics. Calculus was cleansed of that contradiction with the help of Weierstrass's limits, and later of Robinson's hyperreals. But in the two hundred years between Newton's infinitesimals and Weierstrass's limits, Newton brought about a revolution in physics, which together with Bacon's emphasis on experimental science, was a foundational event for modern science. What is more, Newton's slip-up was no *rara avis*. Inconsistency is the constant companion of human cognitive life.

Big information systems are perpetually, pervasively and ineliminably inconsistent. They are inconsistency robust. On the face of it, they vastly multiply the purported offences rendered by the one-shot paradoxes on offer here. In the standard reckonings of the semantic pathology crowd, they are an unmitigated and paralyzing disaster. They are, in fact, no such thing. All such systems substantially outperform their known disadvantages. This is something that requires greater theoretical

⁷⁴See, for example, my "Does changing the subject from *A* to *B* really provide an enlarged understanding of *A*?", *Logic Journal of IGPL*, (2016) 24, 456-480 and "The logical structure of strategic reasoning: Inconsistency-management as a test case for logic", *IfCoLoG Journal of Logics and their Applications*, 5 (2018), 1205-1257.

⁷⁵Consider here Terence Horgan's transvaluationism, according to which vague languages such as English are governed by unsatisfiable principles. See *Essays on Paradoxes*, chapters 7 to 9.

attention than it has so far received.

In “The logical structure of strategic reasoning” [74], I argue that an analogue of the *ex falso* theorem is true in English. If this is right, any system expressible in English that contains any inconsistency is one in which corresponding to its every truth there is a validly derivable negation. That would mean that in every practically indispensable inconsistency-robust system, inconsistency is more than pervasive; it is *total*. Yet even in that state, its utility is undisturbed. It may be worth our while to rehearse the proof here. Let S and X be statement-expressing declarative sentences of English.

1. S and not- S . (by assumption)
2. If S and not- S , then it is true that S and not- S . (Condition T)⁷⁶
3. If it is true that S and not- S then, then on the principle that if both of two things hold true so does each, S is true.
4. If S is true then, on the principle that for any pair of sentences containing S at least one of them is true, at least one of S, X is true for arbitrary X .
5. If not- S is true then, on the negation principle, S is false and therefore by bivalence⁷⁷ is not true.
6. If at least one of S, X is true and S is not, then, on the principle that if at least one of two particular sentences is true and this one is not, it’s the other one that’s true, X is true.
7. Since each of these steps save the first arises in a truth-preserving way from prior such lines, we have it that ours is a valid conditional proof of the statement that contradictions logically imply the negations of anything they imply.⁷⁸

I needn’t remind logician readers that the proof of *ex falso* is a hotly contested one. I won’t take the time to litigate the issue here, beyond noticing what bothers

⁷⁶Advanced by Tarski in “The concept of truth in formalized languages” as a condition of “material adequacy” for any theory of natural language truth. The full condition asserts that “ S ” is true if and only if S (“Snow is white” is true if and only if snow is white). The condition’s biconditional structure provides that if S then it is true that S , which is the form in which we have line (2).

⁷⁷Why, then, retain bivalence? If we lost it, we’d still have excluded middle, and negation would flip the truth-value of a true statement to one that’s not true. If it doesn’t flip to falsity, non-truth functions as a third truth, and negation would retain its negational force.

⁷⁸Let’s also note that the proof contains no occurrence of the contested word “or”, and makes no use of the transformation rules of the propositional calculus.

these critics and trying to mitigate their concerns. The key question is whether at (6) ‘not- S ’ can exclude S from the choice-space between S and X in light of the fact that we already have it that S itself verifies the assumption that at least one of S, X is true. The nub of this question — the deep centre of it — is this. At what point of the proof does ‘not-’ lose its *negational* potency? If it loses its power at line (1), we’ll be landed in the “cancellationist” camp, in which a contradictory pair of propositions say nothing at all, and will thereby have dealt a nasty blow to mathematics (which would lose proof by contradiction). This alone is a good reason for thinking that ‘not-’s negational authority is untrifled with at line (1).

Very well, then, suppose that the ‘not-’ of ‘not- S ’ has full negational potency with regard to S . If it *lacked* this feature at line (1), we’d lose all interest in it. From which I conclude that (1)’s interest is wholly centred on ‘not-’s negational powers. The question that now presses in why would ‘not-’ lose its negational potency lower down the proof’s chain of truth-preserving reasoning? The fact that at line (6) it verifies *ex falso* strikes me as no reason at all to think that the S of line (1) doesn’t negate the ‘not- S ’ of the same line, or that lower down the proof goes off the negational track.

Why would I think so? In its present form, the proof centres on the powers of the negation-operator, whose role in life is to *flip* truth-values. If S is true then ‘not- S is false’. If S is false, then ‘not- S ’ is true. Giving the proof this focus helps us see that what’s really on the line here is whether ‘not-’ retains its truth-value flipping powers under the assumption of a contradictory conjunction.⁷⁹ Accordingly,

- *In approaching the S, X pair at line (4), it is necessary to bear in mind that we already have it independently that if “ S and not- S ” is true, so is S . In approaching the S, X pair at line (6), it is also necessary to bear in mind that we already have it independently on this same assumption that S is not true.*
- *In the general case in which we have it by assumption that at least one of two statements S^*, X^* is true and that S^* is not true, we default to the conclusion that X^* is the true member of the pair.*
- *However, ours is not the general case. It is the quite particular case in which on the assumption of the proof we have it independently that if at least one of S, X is true, one of them is S , without the necessity of the other one also being true. On the other hand, this is a case in which on that same assumption we*

⁷⁹In earlier versions of the proof, for example, in my paper for the Payette [50] volume, there is no mention of truth-values, hence no occasion to consider whether ‘not-’ always flips them. Regrettably, this omission helps disguise the fact that flipping is the principal issue of the proof.

also have it independently that if at least one of S, X is true and yet S isn't true, then X is the one that is.

What we have here is the appearance of a standoff. At different validly derived stages of the proof S 's truth makes it the case that at least one of the pair S, X is true, and also that S 's non-truth makes it the case that X is. The question is whether under this assumption we can have it both ways. My answer is the safe one. Either we can have it both ways or we can't. If we can, *ex falso* is secured by a wholly safe conditional proof. If we cannot, we have made negation unrecognizable. We have the word 'not' but there is no truth-value flipping operator it signals. And if that were so line (1), like death, would have no sting. Thinking otherwise, I take it that (1) possesses the sting of negation, that its sting is not erased down the proof's truth-preserving line, and that therefore any inconsistent system, big or small, theoretical or everyday, as a validly derivable negation for each of its derivable sentences, indeed for each of the sentences of the system's language.⁸⁰ If so, every proposition in SHAEF's information-system has a validly implied negation.⁸¹ In one of Leonard Cohen's songs, it is proposed that "first we take Manhattan and then we take Berlin." How in the world would an SHAEF's system for the 1944 invasion of Normandy have made it possible to take Berlin?

Here is a further point not to overlook,

- *Natural languages are paradox-resistant, and human cognitive systems are ex falso-tolerant. What is the likelihood that the two traits are unconnected?*

4 Paraconsistency

It is worth noting that paraconsistent and dialethic logics are built for negation-inconsistent systems. Their purpose is to keep them free of absolute inconsistency.⁸²

⁸⁰For the languages of formal systems — e.g. languages of the first-order functional calculus — it will depend on whether the system's formal representations of the sentences of natural language and of its properties of interest are sufficiently tight to reflect properties of the formal system's linguistic items back onto their natural language counterparts. If so, absolute inconsistency passes to the formally represented natural language system. If not, the cognitive formal representations can't have been of much utility for natural language in the first place. In the case of information-flows in the down-below, it matters whether unconceptualized and nonlinguistic items of information can stand to one another in any relation of incompatibility sufficient to trigger of *ex falso*'s proof. Final answers aren't yet in, but for now I'll give the nod to a qualified Yes.

⁸¹Supreme Headquarters Allied Expeditionary Force (SHAEF).

⁸²Graham Priest, "Paraconsistency and dialetheism", in Dov M. Gabbay and John Woods, editors, *The Many Valued and Nonmonotonic Turn in Logic*, volume 8 of Gabbay and Woods, editors, *Handbook of the History of Logic*, pages 129-204, Amsterdam: North-Holland, 2007, and Bryson Brown, "Preservationism: A short history", in Gabbay and Woods, volume 8, 2007, pages 95-128.

Dialethic logics is special case of paraconsistent logic, made so by its decision to distinguish paradoxical sentences such as the Liar and the Russell sentence as concurrently T and F together. They are at one in thinking that *ex falso* deserves no home in a respectable logic. Why they would think that this would show it false for English escapes me. Dialethism seeks to take the sting out of the Liar and the Russell, by showing how a sentence's joint truth and falsity need not spell trouble for it. This turns on a trick of distinguishing T and F from T -and- F (or, in a variation, $\{T, F\}$). So it is not true, after all, that the Liar is both T and F together. Rather it has the third truth value T -and- F which implies that it is neither T nor F . The trick of it all is that the Liar is made to be classically consistent!

There is no such trick up my sleeve. When I say that in Frege's set theory every validly entailed sentence has a validly entailed (classical) negation, I am not saying that anything true there is also false. The remarkable thing is that no one who still teaches Frege's original system has any trouble sorting out the enduring truths of intuitive set theory, never mind that they, too, have validly entailed negations there. That, I say, is an empirically discernible fact about us, and a fundamental part of our inconsistency-management capabilities.

Paraconsistent logics are built for single-case inconsistency. To the best of my knowledge, none of those logics has been adapted for the management of *pervasive* inconsistency. If what I said earlier holds true, the pervasiveness of IR's inconsistencies is not limited to big information the size of Five Eyes, but also characterizes deep memory, non-occurrent belief, background information, and so on, of the human cognitive agent. If humanity's cognitive systems are pervasively inconsistent, they lie well out of the treatment intentions of paraconsistent logic. The fact that we don't, thus encumbered, crash and burn takes much of the sting out of any purported proof whose corollary is that anything pervasively inconsistent is also absolutely so.

Inconsistency robustness is a phenomenon of computer languages, whereas *ex falso* holds for every human language. To the extent that it repays computer scientists to keep an eye on the workings of human languages, there is reason to devise a comprehensive theory of rational inconsistency-management in which the insights of IRDL and the insights of naturalized logic are given unified expression. So, as I said at the beginning, let's make it official:

- *Logicians of inconsistency robustness and logicians of naturalized and empirical stripe have enough in common to warrant a working alliance, or anyhow a limited liability partnership.*

Acknowledgements

Earlier versions of parts of the paper were delivered as keynote lectures to audiences at the Square of Opposition Conference at the Lateran Pontifical University in Rome in 2014, the Western Canadian Philosophical Association in Vancouver 2014 and the Philosophy Colloquium of the University of Konstanz 2016. For welcome comments on site or valuable follow-up correspondence, my thanks to Lorenzo Magnani, Jean-Yves Beziau, John N. Martin, Jonathan Westphal, Rusty Jones, Sylvia Berryman, Carrie Jenkins, Roberta Ballarin, the late Adam Morton, Mark Migotti, Jeanne Peijnenburg, David Atkinson and Wolfgang Spohn. The paper has also benefited enormously from numerous conversations and email contacts with Carl Hewitt. My embrace of naturalized logic had some early stirrings which I was slow to notice, and gathered force in my years-long work with Dov Gabbay since the latter 1990s. That was when we discovered our shared opinion that inconsistency is overall “no big deal” in human cognitive life. Gillman Payette, Paul Bartha, Chris Mole and Jonathan Ichikawa have done the service of resisting my proof of *ex falso* for English. Payette was instrumental in helping me see the necessity of raising the question of whether the proof degrades the negational potency displayed by ‘not-‘ in the proof’s first line. It is customary to say that any mistakes that remain in my pleadings are mine alone. I am fairly sure that this can hardly be so. My thanks to Jane Spurr in London for her kind forbearance in formatting my file to CP standards. Carol Woods runs the show in Vancouver, without whom, in all things that matter to me, not.

References

- [1] Bradley Armour-Garb and James A. Woodbridge, *Pretense and Pathology: Philosophical Fictionalism and its Applications*, Cambridge: Cambridge University Press, 2015.
- [2] Bradley Armour-Garb, James Woodbridge, John Woods, Fred Kroon and Jody Azzouni, in a book symposium in *Analysis*. 78 (2018), 713-736. Bradley Armour-Garb and James A. Woodbridge, “Replies”, in Armour-Garb et al. (2018), 718-730.
- [3] Francesco Berto, *There’s Something About Gödel: The Complete Guide to the Incompleteness Theorem*, Madden, MA: Wiley Blackwell, 2009. First published, in Italian in 2008.
- [4] Francesco Berto, “The Gödel paradox and Wittgenstein’s reasons”, *Philosophia Mathematica*, 17 (2009), 208-219.
- [5] Francesco Berto, “Modal Meinongianism and fiction: The best of three worlds”, *Philosophical Studies*, 153 (2011), 313-334.

- [6] George Boolos, “A new proof of the Gödel incompleteness theorem”, *Notices of the American Mathematical Society*, 36 (1989), 338-390. Reprinted in his *Logic, Logic, and Logic*, pages 383-388, Cambridge, MA: Harvard University Press, 1998.
- [7] Bryson Brown, “Preservationism: A short history”, in Dov M. Gabbay and John Woods, editors, *The Many Valued and Nonmonotonic Turn in Logic*, volume 8 of Gabbay and Woods, editors, *Handbook of the History of Logic*, Amsterdam: North-Holland, 2007. pages 95-128.
- [8] Cesare Burali-Forti, “Sulle classi ben ordinate”, *Rendicanti del Circolo Matematico di Palermo*, 11 (1897), 260.
- [9] Cesare Burali-Forti, “Una question sui numeri transfinita”, *Recondidi del Circolo Matematico di Palermo*, 11 (1897), 154-164.
- [10] Andrea Cantini, “Paradoxes, self-reference and truth in the 20th century”, in Dov M. Gabbay and John Woods, editors, *Logic from Russell to Church*, volume 5 of their *Handbook of the History of Logic*, pages 875-1013, Amsterdam: North-Holland, 2009.
- [11] Georg Cantor, “Über eine elementare Frage der Mannigfaltigkeitslehre”, in William Bl. Ewald, editor, *From Immanuel Kant to David Hilbert: A Source Book in the Foundations of Mathematics*, volume 2, pages 920-922, Oxford: Oxford University Press, 1996. First published in 1891.
- [12] Rudolf Carnap, *The Logical Syntax of Language*, London: Routledge & Kegan Paul, 1937. First published, in German, in 1934.
- [13] Noël Carroll, *The Philosophy of Horror*, New York and London: Routledge, 1990.
- [14] Richard Cartwright, *Philosophical Essays*, Cambridge, Ma: MIT Press, 1987.
- [15] Gregory Chaitin, “The Berry Paradox”, *Complexity* 1 (1995), 26-30.
- [16] Alonzo Church, “The Richard paradox”, *American Mathematical Monthly*, 41 (1934), 356-361.
- [17] Haskell B. Curry, “The paradox of Kleene and Rosser”, *Transactions of the American Mathematical Society*, 41 (454-516), (1942).
- [18] Cora Diamond, editor, *Wittgenstein’s Lectures on the Foundations of Mathematics*, Cambridge, 1939, Hassocks, UK: Harvester, 1976.
- [19] Gottlob Frege, *Die Grundlagen der Arithmetik: Eine logisch mathematische Untersuchung über den Begriff der Zahl*, Breslau: Wilhelm Koebner, 1884. Translated into English by J. L. Austin as *The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number*, Oxford: Basil Blackwell, 1950, p. IX.
- [20] Gottlob Frege, *Begriffsschrift: A Formula Language, Modeled Upon that of Arithmetic, for Pure Thought*, in Jean van Heijenoort, editor, *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, pages 5-82, Cambridge, MA: Harvard University Press, 1967.
- [21] Dov M. Gabbay and Anthony Hunter, “Making inconsistency respectable 1: A logical framework for inconsistency in reasoning”, in Ph. Jourand and T. Kelemen, editors, *Fundamentals of Artificial Intelligence Research*, volume 535 of *Lecture Notes in Computer Science*, pages 19-32, Berlin: Springer, 1991.

- [22] Dov M. Gabbay and Anthony Hunter, “Making inconsistency respectable, part 2: Meta-level handling of inconsistency”, in M. Clarke, R. Kruse and S. Seraffin, editors, *Lecture Notes in Computer Science*, volume 742, pages 129-136, Berlin: Springer, 1992.
- [23] P. T. Geach, “Ryle on namely-riders”, *Analysis*, 21 (1961), 64-67.
- [24] Gerd Gigerenzer, “From tools to theories” in Carl Graumann and Kenneth J. Gergen, editors, *Historical Dimensions of Psychological Discourse*, pages 336-359, Cambridge: Cambridge University Press, 1996.
- [25] Kurt Gödel, “On formally undecidable propositions of Principia Mathematica and related system, in Jean van Heijenoort, *From Frege to Gödel*, pages 595-617, First published, in German in 1931.
- [26] Kurt Grelling and Leonard Nelson, “Remarks on the paradoxes of Russell and Burali-Forti”, *Abhandlungen der Fries’schen Schule*, (Neue Serie), 2 (1907/1908), 300-334.
- [27] Anil Gupta, “Truth and paradox”, in Robert L. Martin, editor, *Recent Essays on Truth and the Liar Paradox*, pages 175-236, Oxford: Clarendon Press, 1984.
- [28] Gilbert Harman, *Change in View: Principles of Reasoning*, Cambridge, MA: MIT Press, 1986.
- [29] Hans G. Herzberger, “Paradoxes of grounding in semantics”, *Journal of Philosophy*, 67 (1970), 145-167.
- [30] Carl Hewitt, “Inconsistency robustness in foundations: Mathematics proves its own consistency and other matters”, in Carl Hewitt and John Woods, *Inconsistency Robustness*, 2015, pages 104-157.
- [31] Carl Hewitt, “Speed bumps in argumentation: Discursive inference for massive pervasively-inconsistent information”, in Dov Gabbay, Lorenzo Magnani, Woosuk Park, and Ahti-Veikko Pietarinen, editors, *Natural Arguments: A Tribute to John Woods*, volume 40 in the Tributes series, pages 35-44, London: College Publications, 2019.
- [32] Carl Hewitt, “Inconsistency robust direct logic”, a preprint of a paper to appear.
- [33] Carl Hewitt, “Strong types for Direct Logic”, a preprint in *PhilSci Archive*, online, 2017, 1-78.
- [34] Carl Hewitt and John Woods, editors, *Inconsistency Robustness*, volume 52 of *Studies in Logic*, London. College Publications, 2015.
- [35] Jaakko Hintikka, *The Principles of Mathematics Revised*, Cambridge: Cambridge University Press, 1996.
- [36] Terence Horgan, *Essays on Paradoxes*, New York: Oxford University Press, 2016.
- [37] Anthony Hunter, “Reasoning with contradictory information using quasi-classical logic”, *Journal of Logic and Computation*, 10 (2000), 677-703.
- [38] Dale Jacquette, *Meinongian Logic: The Semantics of Existence and Nonexistence*, Berlin: Walter de Gruyter, 1996.
- [39] Dale Jacquette, “Diagonalization in logic and mathematics”, in Dov Gabbay and Franz Guenther, editors, *Handbook of Philosophical Logic*, 2nd edition, pages 55-147, Dordrecht: Kluwer, 2002.
- [40] Fairouz Kamareddine, Twan Laan and Rob Nederpelt, “A history of types”, in Dov M.

- Gabbay, Francis Jeffrey Pelletier and John Woods, editors, *Logic: A History of its Central Concepts*, volume 11 of Gabbay and Woods, editors, *Handbook of the History of Logic*, pages 451-511, Amsterdam: North-Holland, 2012.
- [41] David Kaplan, “Demonstratives”, in Joseph Almog, John Perry and Howard Wettstein, editors, *Themes from Kaplan*, pages 481-563, Oxford: Oxford University Press, 1989.
- [42] S C. Kleene and J. B. Rosser, “The inconsistency of certain formal systems”, *Annals of Mathematics*, 36 (1935), 630-637.
- [43] Saul Kripke, “Outline of a theory of truth”, *Journal of Philosophy*, 72 (1975). Reprinted in his *Philosophical Troubles: Collected Papers Volume 1*, pages 75-98, Oxford: Oxford University Press, 2011; paperback 2013.
- [44] Timm Lampert, “The Gödel paradox and ‘Wittgenstein’s reasons’”, *Philosophia Mathematica*, 26 (2018), 324-345.
- [45] M. H. Löb, “Solution of a problem of Leon Henkin”, *Journal of Symbolic Logic*, 20 (1955), 115-118.
- [46] David Malone, *Dangerous Knowledge*, BBC 4 documentary, 2007. http://www.dailymotion.com/playlist/x1cbyd_xSilverPhinx_bbc-dangerous-knowledge/1
- [47] D. Mirimanoff, “Les antinomies de Russell et de problème fondamentale de la théorie des ensembles”, *L'Enseignement Mathématique*, 19 (1917), 37-52.
- [48] Chris Mortensen, *Inconsistent Mathematics*, Dordrecht and Boston: Kluwer, 1995.
- [49] Terence Parsons, *Nonexistent Objects*, New Haven: Yale University Press, 1980.
- [50] Gillman Payette, editor, “Shut up”, he explained: *Essays in Honour of Peter K. Schotch*, volume 31 of *Tributes*, London: College Publications, 2016.
- [51] Graham Priest, “Paraconsistency and dialetheism”, in Dov M. Gabbay and John Woods, editors, *The Many Valued and Nonmonotonic Turn in Logic*, volume 8 of Gabbay and Woods, editors, *Handbook of the History of Logic*, pages 129-204, Amsterdam: North-Holland, 2007.
- [52] W. V. Quine, *Word and Object*, New York: John Wiley, 1960.
- [53] Victor Rodych, “Misunderstanding Gödel: New arguments about Wittgenstein and new remarks by Wittgenstein”, *Dialectica*, 57 (2003), 279-313.
- [54] Alexandra Russo, Bashar Nusibeh and Steve Easterbrook, “Making inconsistency respectable in software development”, *Journal of Systems and Software*, 56 (2000).
- [55] Bertrand Russell, “On denoting”, *Mind* N. S. 14 (1905).
- [56] Bertrand Russell, “Les paradoxes de la logique”, *Revue de Métaphysique et de Morale*, 14 (1906), 627-650.
- [57] Bertrand Russell, *Revue de Métaphysique et de Morale*, 19 (1911), and under the title “The philosophical implications of mathematical logic”, revised by and translated by P. E. B. Jourdain, appeared in *The Monist*, 22 (1913), 481-493.
- [58] Bertrand Russell, *Our Knowledge of the External World*. London: Allen and Unwin, 1914.
- [59] Gilbert Ryle, “Heterologicality”, *Analysis*, 11 (1951), 61-69.

- [60] Raymond Smullyan, *Diagonalization and Self-Reference*, Oxford: Clarendon Press, 1984.
- [61] Alfred Tarski, “On the concept of truth in formalized languages”, in John Corcoran, editor, and J. H. Woodger, translator, *Logic, Semantics and Metamathematics*, pages 152-278, Indianapolis: Hackett, 1983. First appearance in Polish and German in 1935.
- [62] A. S. Troelstra and Dirk van Dalen, *Constructivism in Mathematics*, volumes 1 and 2, Amsterdam: North-Holland, 1988.
- [63] A.N. Whitehead and Bertrand Russell, *Principia Mathematica*, Cambridge: University of Cambridge Press, 1910-1913.
- [64] Oscar Wilde, *The Picture of Dorian Gray*, published with deletions in the July 1890 issue of Lipicott’s Monthly Magazine. Unexpurgated version by the Belknap Press in 2011.
- [65] Ludwig Wittgenstein, *Remarks on the Foundations of Mathematics*, Oxford: Blackwell, 1956.
- [66] John Woods, “A captious nicety of argument”, in Lewis Edwin Hahn and Paul Arthur Schilpp, editors, *The Philosophy of W. V. Quine*, volume 18 of the Library of Living Philosophers, second, expanded edition, pages 687-725, Chicago and LaSalle: Open Court, 1998; first edition 1986.
- [67] John Woods, *Paradox and Paraconsistency: Conflict Resolution in the Abstract Sciences*, Cambridge: Cambridge University Press, 2003.
- [68] John Woods, *Errors of Reasoning: Naturalizing the Logic of Inference*, volume 45 of Studies in Logic, London: College Publications, 2013; reprinted with corrections in 2014.
- [69] John Woods, *Is Legal Reasoning Irrational? An Introduction to the Epistemology of Law*, 2nd edition, revised and expanded, volume two of the Law and Society series, London: College Publications, 2017. First published in 2015.
- [70] John Woods, “Does changing the subject from A to B really provide an enlarged understanding of A?”, *Logic Journal of IGPL*, (2016) 24, 456-480.
- [71] John Woods, “How robust can inconsistency get?”, *IfCoLoG Journal of Logics and Their Applications*, 1 (2014), 177-216.
- [72] John Woods, “Globalization makes inconsistency unrecognizable”, pages 137-165, in Payette (2016).
- [73] John Woods, *Truth in Fiction: Rethinking its Logic*, volume 391 in the Synthese Library, Cham: Springer, 2018.
- [74] John Woods, “The logical structure of strategic reasoning: Inconsistency-management as a test case for logic”, *IfCoLoG Journal of Logics and their Applications*, 5 (2018), 1205-1257.
- [75] John Woods, “The role of the common in cognitive prosperity”, *Informal Logic*, to appear.
- [76] John Woods, *Three Grades of Deductive Involvement: A Full-Service Logic for the Consequence-Relation*, forthcoming in Springer’s Synthese Library.
- [77] Stephen Yablo, “Paradox without self-reference”, *Analysis* (53), 1993, 251-252.

TWO GENERALIZATIONS OF BOUNDED INVOLUTIVE LATTICES AND OF ORTHOLATTICES

AFRODITA IORGULESCU

*Department of Informatics and Economic Cybernetics
Bucharest University of Economic Studies, Bucharest, RO*
afrodita.iorgulescu@ase.ro

MICHAEL KINYON*

*Department of Mathematics
University of Denver, Denver, CO 80208, USA*
michael.kinyon@du.edu

Abstract

We introduce and study two dual independent absorption laws and two generalizations of bounded lattices and, consequently, two corresponding generalizations of bounded involutive lattices and two corresponding generalizations of ortholattices; a proper subclass of MV algebras is obtained. Many examples are provided.

Keywords: m-MEL algebra, m-BE algebra, m-pre-BCK algebra, m-BCK algebra, MV algebra, lattice, bounded involutive lattice, ortholattice, softlattice, widelattice, bounded involutive softlattice, bounded involutive widelattice, orthosoftlattice, orthowidelattice, Boolean algebra

MSC 2020: 06B75, 06C15, 06D35

1 Introduction

The commutative algebraic structures connected directly or indirectly with classical/nonclassical logics belong to two parallel “worlds”:

1. The “world” of *algebras of logic*, where there are essentially one implication \rightarrow (two, in the noncommutative case) and an element 1 (that can be the last element);

The authors thank the referees for their valuable comments and suggestions, which helped to improve the presentation of the paper.

*Partially supported by Simons Foundation Collaboration Grant 359872

the algebras $(A, \rightarrow, 1)$, verifying the basic property (M): $1 \rightarrow x = x$, are called *M algebras* [10]; among the M algebras with additional operations, there are the algebras $(A, \rightarrow, 0, 1)$ (where a negation can be defined by: $x^- = x \rightarrow 0$), or $(A, \rightarrow, ^-, 1)$, with $1^- = 0$, where 1 is the *last element*, verifying (Ex) (Exchange): $x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$; an internal binary relation can be defined by: $x \leq y \iff x \rightarrow y = 1$; algebras belonging to this “world” are the bounded MEL, BE and aBE, pre-BCK algebras, BCK algebras, bounded BCK algebras, BCK(P) algebras, Hilbert algebras, Wajsberg algebras, implicative-Boolean algebras, etc.

2. The “world” of *algebras*, where there are essentially a product \odot and an element 1 (that can be the last element); the algebras $(A, \odot, 1)$, verifying the corresponding basic properties (PU): $1 \odot x = x$ and (Pcomm): $x \odot y = y \odot x$, are called *commutative unital magmas*. Among the commutative unital magmas with additional operations, there are the algebras [12] $(A, \odot, ^-, 1)$, with $1^- = 0$, where 1 is the *last element*, verifying (Pass) (associativity of product): $x \odot (y \odot z) = (x \odot y) \odot z$; an internal binary relation can be defined by: $x \leq_m y \iff x \odot y^- = 0$, where ‘m’ comes from ‘magma’. Algebras belonging to this “world” are: the m-MEL, m-BE and m-aBE, m-pre-BCK algebras, m-BCK algebras [12], pocrim, (bounded) lattices, residuated lattices, BL algebras, MTL algebras, NM algebras, MV algebras, Boolean algebras, etc.

Between the two parallel “worlds” there are connections, as for examples: the equivalence between BCK(P) algebras and pocrim, in the non-involutive case, and the definitional equivalence between Wajsberg algebras and MV algebras, or between implicative-Boolean algebras and Boolean algebras, in the involutive case ($(x^-)^- = x$). In [12], the two general Theorems 9.1 and 9.3 connect the two ‘worlds’ in the involutive case, by the inverse maps $\Phi (x \odot y \stackrel{def.}{=} (x \rightarrow y^-)^-)$ and $\Psi (x \rightarrow y \stackrel{def.}{=} (x \odot y^-)^-)$ (Theorem 9.1 is for algebras *with last element*, while Theorem 9.3 is for algebras *without last element*). These theorems can be used to prove the *definitional equivalence* between the analogous involutive algebras from the two “worlds” simply by choosing appropriate definitions of these algebras.

Besides the classical and non-classical logics, there exist the quantum logics. Examples of algebraic structures connected with quantum logics are the bounded implicative (implication) lattices, the ortholattices etc. The connections between all these logics are much more clear now.

In this paper, we continue the research from [16], based on [12], in the “world” of *algebras* of the form $(A, \odot, ^-, 1)$, with $1^- = 0$, where 1 is the *last element*, verifying (Pass). We introduce two dual independent absorption laws and two generalizations of bounded lattices and, consequently, two corresponding generalizations of bounded involutive lattices [4] and two corresponding generalizations of ortholattices [18], [4],

[22]; a proper subclass of MV algebras [3], [2] is obtained. Many examples are provided. This paper, like [12] and [16], presents the facts in the same *unifying way*, which consists in fixing unique names for the defining properties, making lists of these properties and then using them for defining the different algebras and for obtaining results.

The paper is organized as follows. In Section 2 (Preliminaries), we recall definitions and results concerning bounded lattices (**BL**), bounded involutive (involution) lattices (**BIL**) and ortholattices (**OL**). We recall also definitions and results concerning the involutive m-MEL algebras and involutive m-BE algebras, defined in [12] and studied in [16]. In Section 3 (Two dual independent absorption laws and two generalizations of bounded lattices), we introduce and study two dual independent absorption laws: (m-Wabs-i) and (m-Vabs-i) and two generalizations of bounded lattices (**BL**): bounded softlattices (**BSL**) and bounded widelattices (**BWL**). In Section 4 (Two generalizations of bounded involutive lattices), we introduce and study two corresponding generalizations of **BIL**: bounded involutive softlattices (**BISL**) and bounded involutive widelattices (**BIWL**). Finally, we put **BIL**, **BISL**, **BIWL**, and their subclasses, on the “involutive Little map”. In Section 5 (Two generalizations of ortholattices), we introduce and study two corresponding generalizations of **OL**: orthosoftlattices (**OSL**) and orthowidelattices (**OWL**). The core of the paper is Theorem 5.14 (based on Theorem 5.13, proved by PROVER9), saying that transitive and antisymmetric orthowidelattices are MV algebras; thus, a proper subclass of MV algebras (**taOWL**) is obtained; a generalization of $(\wedge_m\text{-comm})$ property, called (Δ_m) , is introduced on this occasion. We put **OL**, **OSL**, **OWL**, and their subclasses, on the “involutive Little map”. Finally, future work is presented and the study of the connections between (m-Pabs-i) and (WNM_m) is started. In Section 6 (Examples), we present 11 examples of the various algebras discussed herein.

2 Preliminaries

2.1 Bounded lattices

In order to fix the notations, we recall the following equivalent definitions of lattices.

Definition 2.1. A poset (partially ordered set) $\mathcal{A} = (A, \leq^O)$ will be said to be a *lattice*, if for each two elements $x, y \in A$, there exist $\inf(x, y)$ and $\sup(x, y)$. Moreover, if there exist $0, 1 \in A$ such that $0 \leq^O x \leq^O 1$ for all $x \in A$, then \mathcal{A} is said to be a *bounded lattice* (with last (top) element 1 and first (bottom) element 0) and is denoted by $\mathcal{A} = (A, \leq^O, 0, 1)$.

Definition 2.2. An algebra $\mathcal{A} = (A, \wedge, \vee)$ or, dually, $\mathcal{A} = (A, \vee, \wedge)$, of type $(2, 2)$, will be said to be a *lattice*, if the following dual properties hold: for all $x, y, z \in A$,

(m-Wid)	(idempotency of \wedge)	$x \wedge x = x$,
(m-Vid)	(idempotency of \vee)	$x \vee x = x$;
(m-Wcomm)	(commutativity of \wedge)	$x \wedge y = y \wedge x$,
(m-Vcomm)	(commutativity of \vee)	$x \vee y = y \vee x$;
(m-Wass)	(associativity of \wedge)	$x \wedge (y \wedge z) = (x \wedge y) \wedge z$,
(m-Vass)	(associativity of \vee)	$(x \vee y) \vee z = x \vee (y \vee z)$;
(m-Wabs)	(absorption of \wedge over \vee)	$x \wedge (x \vee y) = x$;
(m-Vabs)	(absorption of \vee over \wedge)	$x \vee (x \wedge y) = x$,

where “m” comes from “magma”, “W” comes from “wedge” (the \LaTeX command for the meet symbol) and “V” comes from “vee” (the \LaTeX command for the join symbol). Moreover, if there exist $0, 1 \in A$ such that: for all $x \in A$,

(m-WU)	$1 \wedge x = x$ and, dually,
(m-VU)	$0 \vee x = x$,

then \mathcal{A} is said to be a *bounded lattice* (with last element 1 and first element 0) and is denoted by $\mathcal{A} = (A, \wedge, \vee, 0, 1)$ or, dually, by $\mathcal{A} = (A, \vee, \wedge, 0, 1)$.

Cf. [22], it was Dedekind [5] who noted that axioms (m-Wid), (m-Vid) can be proved from (m-Wabs), (m-Vabs) ($x \vee x = x \vee (x \wedge (x \vee x)) = x$ and $x \wedge x = x \wedge (x \vee (x \wedge x)) = x$); thus, the above “standard” system **L8** of eight axioms of lattices is equivalent to the system **L6** of six axioms: $\mathbf{L6} = \{(\text{m-Wcomm}), (\text{m-Vcomm}), (\text{m-Wass}), (\text{M-Vass}), (\text{m-Wabs}), (\text{m-Vabs})\}$, and this equivalence was proved by Ore [21].

Naming convention for the dual lattices: (A, \wedge, \vee) is the *left-lattice* and (A, \vee, \wedge) is the *right-lattice* (names coming from the *left-continuity* of a t-norm and the *right-continuity* of a t-conorm; see more on *left- and right- algebras* in [8]).

We shall denote by **L** the class of all left-lattices and by **BL** the class of all bounded left-lattices.

Corollary 2.3. *Let $\mathcal{A} = (A, \wedge, \vee, 0, 1)$ be a bounded left-lattice. Then we have the equivalences:*

$$(m-WU) \iff (m-VL) \quad \text{and} \quad (m-VU) \iff (m-WL)$$

where: $(m-WL)$ $0 \wedge x = 0$ and, dually,
 $(m-VL)$ $1 \vee x = 1$.

Definition 2.4. ([16], Definition 2.9) (See [4] for an equivalent definition) (The dual one is omitted) A *bounded involutive (or involution) left-lattice*, or a *left-BIL* for short, is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, - = {}^{-L}, 0, 1)$ such that the reduct $(A^L, \wedge, \vee, 0, 1)$

is a bounded left-lattice and the unary operation $-$ (called *involution* or *generalized complement*) satisfies the following conditions:

- (DN) $(x^-)^- = x$ (Double Negation) and
- (DeM1) $(x \vee y)^- = x^- \wedge y^-$ (De Morgan law 1) or
- (DeM2) $(x \wedge y)^- = x^- \vee y^-$ (De Morgan law 2).

We shall denote by **BIL** the class of all left-BILs.

Ortholattices form an important example of *sharp* structure (which satisfies the noncontradiction principle) from sharp quantum theory [4] (Birkhoff, 1967; Kalmbach, 1983). Recall the following definition:

Definition 2.5. (See [22], [4]) (The dual one is omitted) A *left-ortholattice*, or a *left-OL* for short, is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, - = {}^{-L}, 0, 1)$ such that the reduct $(A^L, \wedge, \vee, 0, 1)$ is a bounded left-lattice and the unary operation $-$ satisfies (DN), (DeM1), (DeM2) and the complementation laws:

- (m-WRe) $x \wedge x^- = 0$ (noncontradiction principle)
- (m-VRe) $x \vee x^- = 1$ (excluded middle principle).

We shall denote by **OL** the class of all left-OLs.

Definition 2.6. (The dual one is omitted) A *left-Boolean algebra* is a bounded left-lattice that is distributive and complemented, i.e. is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, - = {}^{-L}, 0, 1)$ verifying: (m-Wid), (m-Vid), (m-Wcomm), (m-Vcomm), (m-Wass), (m-Vass), (m-Wabs), (m-Vabs), (m-WU), (m-VU), (m-Wdis), (m-Vdis), (m-WRe), (m-VRe), where:

- (m-Wdis) $z \wedge (x \vee y) = (z \wedge x) \vee (z \wedge y)$ (distributivity of \wedge over \vee)
- (m-Vdis) $z \vee (x \wedge y) = (z \vee x) \wedge (z \vee y)$ (distributivity of \vee over \wedge).

Note that, if \mathcal{A}^L is a left-Boolean algebra, then (DN), (DeM1), (DeM2) hold. We shall denote by **Boole** the class of all left-Boolean algebras. By the above definitions and considerations, we obtain the connections:

$$\mathbf{OL} = \mathbf{BIL} + (\text{m-WRe}) + (\text{m-VRe}) \quad \text{and} \quad \mathbf{Boole} = \mathbf{OL} + (\text{m-Wdis}) + (\text{m-Vdis}).$$

2.2 Involutive m-MEL and m-BE algebras

Recall from [12] the following:

Let $\mathcal{A}^L = (A^L, \odot, - = {}^{-L}, 1)$ be an algebra of type $(2, 1, 0)$ and define $0 \stackrel{\text{def.}}{=} 1^-$. Define an *internal* binary relation \leq_m on A^L by:

$$(\text{m-dfrelP}) \quad x \leq_m y \stackrel{\text{def.}}{\iff} x \odot y^- = 0.$$

Consider the following list **m-A** of basic properties that can be satisfied by \mathcal{A}^L [12]:

(PU) $1 \odot x = x = x \odot 1$ (unit element of product, the *identity*);

(Pcomm) $x \odot y = y \odot x$ (commutativity of product),

(Pass) $x \odot (y \odot z) = (x \odot y) \odot z$ (associativity of product);

(Neg1-0) $1^- = 0$;

(Neg0-1) $0^- = 1$;

(m-An) $(x \odot y^- = 0 \text{ and } y \odot x^- = 0) \implies x = y$ (antisymmetry),

(m-B) $[(x \odot y^-)^- \odot (x \odot z)] \odot (y \odot z)^- = 0$,

(m-BB) $[(z \odot x)^- \odot (y \odot x)] \odot (y \odot z^-)^- = 0$,

(m-*) $x \odot y^- = 0 \implies (z \odot y^-) \odot (z \odot x^-)^- = 0$,

(m-**) $x \odot y^- = 0 \implies (x \odot z) \odot (y \odot z)^- = 0$,

(m-L) $x \odot 0 = 0$ (last element),

(m-Re) $x \odot x^- = 0$ (reflexivity),

(m-Tr) $(x \odot y^- = 0 \text{ and } y \odot z^- = 0) \implies x \odot z^- = 0$ (transitivity);

etc.

Dually, let $\mathcal{A}^R = (A^R, \oplus, - = {}^{-R}, 0)$ be an algebra of type $(2, 1, 0)$ and define $1 \stackrel{\text{def.}}{=} 0^- \neq 0$. Define an *internal* binary relation \geq_m on A^R by:

(m-dfrelS) $x \geq_m y \stackrel{\text{def.}}{\iff} x \oplus y^- = 1$.

The list of dual properties is omitted.

Note that if $\odot = \wedge$ and/or $\oplus = \vee$, then the property (m-An) becomes (m-WAn) and/or the dual property (m-An^R) becomes (m-VAn) and so on.

Recall from [12] the definitions of those algebras needed in this paper (the dual ones are omitted).

Definitions 2.7. Let $\mathcal{A}^L = (A^L, \odot, - = {}^{-L}, 1)$ be an algebra of type $(2, 1, 0)$. Define $0 \stackrel{\text{def.}}{=} 1^-$ (hence (Neg1-0) holds) and suppose that $0^- = 1$ (hence (Neg0-1) holds too). \mathcal{A}^L is a:

- *left-m-MEL algebra*, if (PU), (Pcomm), (Pass), (m-L) hold;
- *left-m-BE algebra*, if (PU), (Pcomm), (Pass), (m-L), (m-Re) hold;
- *left-m-pre-BCK algebra*, if (PU), (Pcomm), (Pass), (m-L), (m-Re), (m-BB) hold;
- *left-m-BCK algebra*, if (PU), (Pcomm), (Pass), (m-L), (m-Re), (m-An), (m-BB) hold.

We shall denote by bold letters the classes of algebras. In ([12], Figures 10, 11), the “Big map” connecting all the new algebras was drawn divided into two pieces, in order to be not too complicated.

Definitions 2.8. Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an algebra of type $(2, 1, 0)$ and let \leq_m be the associated internal binary relation defined by (m-dfrelP). We shall say that \mathcal{A}^L is:

- *antisymmetric*, if \leq_m is antisymmetric (i.e. (m-An) holds);
- *transitive*, if \leq_m is transitive (i.e. (m-Tr) holds).

If \mathbf{X} is a class of algebras, we shall denote by \mathbf{aX} (\mathbf{tX} , $\mathbf{atX=taX}$) the subclass of all antisymmetric (transitive, antisymmetric and transitive, respectively) algebras of \mathbf{X} .

We shall say that an algebra is *involutive*, if it verifies (DN). If \mathbf{X} is a class of algebras, we shall denote by $\mathbf{X}_{(DN)}$ the subclass of all involutive algebras of \mathbf{X} . By ([12], Theorem 6.12), in any involutive m-BE algebra we have the equivalences:

$$(\text{m-BB}) \iff (\text{m-B}) \iff (\text{m-**}) \iff (\text{m-*}) \iff (\text{m-Tr}).$$

Any left-m-BCK algebra is involutive, by ([12], Theorem 6.13), and so we write: $\mathbf{m-BCK} = \mathbf{m-BCK}_{(DN)}$.

In ([16], Figure 5), the hierarchies between $\mathbf{m-MEL}_{(DN)}$, $\mathbf{m-tMEL}_{(DN)}$, $\mathbf{m-aMEL}_{(DN)}$, $\mathbf{m-taMEL}_{(DN)}$ and $\mathbf{m-BE}_{(DN)}$, $\mathbf{m-tBE}_{(DN)} = \mathbf{m-pre-BCK}_{(DN)}$, $\mathbf{m-aBE}_{(DN)}$, $\mathbf{m-taBE}_{(DN)} = \mathbf{m-BCK}_{(DN)}$ are presented under the name “involutive Little map”.

Our definition of MV algebra is as follows (the dual one is omitted): a *left-MV algebra* is an algebra $\mathcal{A}^L = (A^L, \odot, - = {}^{-L}, 1)$ verifying (PU), (Pcomm), (Pass), (m-L), (DN) and $(\wedge_m\text{-comm}) ((x^- \odot y)^- \odot y = (y^- \odot x)^- \odot x)$. Denote by \mathbf{MV} the class of all left-MV algebras. We recall the following important remark, which was the motivation of [12]: a left-MV algebra is just an involutive left-m-MEL algebra verifying $(\wedge_m\text{-comm})$. Since $(\wedge_m\text{-comm})$ implies (m-Re), (m-An), (m-BB), by ([12], (mB1), (mB2), (mCBN1)) respectively, it follows that a left-MV algebra is a left-m-BCK algebra [12]. We write: $\mathbf{MV} \subset \mathbf{m-BCK} = \mathbf{m-BCK}_{(DN)}$.

Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-MEL algebra. Because of the axiom (DN), we introduced in [16] the new operation \oplus , the dual of \odot , by: for all $x, y \in A^L$,

$$x \oplus y \stackrel{\text{def.}}{=} (x^- \odot y^-)^-. \tag{1}$$

By ([16], Corollary 3.5), $(A^L, \oplus, -, 0)$ is an involutive right-m-MEL algebra.

Consider also the following properties [12]:

- (G)=(m-Pid) $x \odot x = x$ (“G” comes from Gödel) and, dually,
- (G^R)=(m-Sid) $x \oplus x = x$ (dual laws of idempotency);
- (m-Pabs) $x \odot (x \oplus y) = x$ and, dually,
- (m-Sabs) $x \oplus (x \odot y) = x$ (dual laws of absorption);
- (m-Pimpl) $[(x \odot y^-)^- \odot x^-]^-= x$ and, dually,
- (m-Simpl) $[(x \oplus y^-)^- \oplus x^-]^-= x$ (dual laws of implicativity);
- (m-Pdis) $z \odot (x \oplus y) = (z \odot x) \oplus (z \odot y)$ and, dually,
- (m-Sdis) $z \oplus (x \odot y) = (z \oplus x) \odot (z \oplus y)$ (dual laws of distributivity);
- (m-Pdiv) $x \odot (x \odot y^-)^- = x \odot y$ and, dually,
- (m-Sdiv) $x \oplus (x \oplus y^-)^- = x \oplus y$ (dual laws of divisibility).

We have obtained in [16] new equivalent definitions (Definitions 2) of BILs, OLs, Boolean algebras, namely:

- A *bounded involutive left-lattice* is a (involutive) left-m-MEL algebra verifying (m-Pimpl).
- A *left-ortholattice* is a (involutive) left-m-BE algebra verifying (m-Pimpl).
- A *left-Boolean algebra* is a (involutive) left-m-BE algebra verifying (m-Pdiv).
- A *left-Boolean algebra* is a (involutive) left-m-BE algebra verifying (m-Pdis).

3 Two dual independent absorption laws and two generalizations of bounded lattices

3.1 Two dual independent absorption laws

Since, in a lattice, the absorption laws (m-Wabs) and (m-Vabs) are not independent (they imply the idempotency laws (m-Wid) and (m-Vid)), we shall introduce the following two dual *independent absorption laws*:

- (m-Wabs-i) $x \wedge (x \vee x \vee y) = x$ and, dually,
- (m-Vabs-i) $x \vee (x \wedge x \wedge y) = x$ (dual laws of independent absorption);
- (m-Pabs-i) $x \odot (x \oplus x \oplus y) = x$ and, dually,
- (m-Sabs-i) $x \oplus (x \odot x \odot y) = x$ (dual laws of independent absorption).

We shall prove that the system of eight axioms:

$$\mathbf{L8-i} = \{(m-Wid), (m-Vid), (m-Wcomm), (m-Vcomm), (m-Wass), (M-Vass), (m-Wabs-i), (m-Vabs-i)\}$$

is equivalent with the “standard” system **L8** of axioms for lattices. First, we prove the following lemma.

Lemma 3.1.

- (a) $(m\text{-Vid}) + (m\text{-Vass}) \implies ((m\text{-Wabs-i}) \iff (m\text{-Wabs}))$ and, dually,
 (a') $(m\text{-Wid}) + (m\text{-Wass}) \implies ((m\text{-Vabs-i}) \iff (m\text{-Vabs}))$.

Proof. First, $(m\text{-Vid}) + (m\text{-Vass}) + (m\text{-Wabs-i})$ imply $(m\text{-Wabs})$:

$$x \wedge (x \vee y) = x \wedge ((x \vee x) \vee y) = x \wedge (x \vee x \vee y) = x$$

and $(m\text{-Vid}) + (m\text{-Vass}) + (m\text{-Wabs})$ imply $(m\text{-Wabs-i})$:

$$x \wedge (x \vee x \vee y) = x \wedge ((x \vee x) \vee y) = x \wedge (x \vee y) = x;$$

thus (a) holds.

Then, dually, $(m\text{-Wid}) + (m\text{-Wass}) + (m\text{-Vabs-i})$ imply $(m\text{-Vabs})$:

$$x \vee (x \wedge y) = x \vee ((x \wedge x) \wedge y) = x \vee (x \wedge x \wedge y) = x$$

and $(m\text{-Wid}) + (m\text{-Wass}) + (m\text{-Vabs})$ imply $(m\text{-Vabs-i})$:

$$x \vee (x \wedge x \wedge y) = x \vee ((x \wedge x) \wedge y) = x \vee (x \wedge y) = x;$$

thus (a') holds. □

Now, the announced result follows by the above Lemma.

Theorem 3.2. *We have: $L8\text{-i} \iff L8$.*

3.2 Two generalizations of bounded lattices

3.2.1 Bounded softlattices

Definitions 3.3. (The dual ones are omitted)

(1) A *left-softlattice* is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee)$ of type $(2, 2)$ such that the axioms $(m\text{-Wid})$, $(m\text{-Vid})$, $(m\text{-Wcomm})$, $(m\text{-Vcomm})$, $(m\text{-Wass})$, $(m\text{-Vass})$ are satisfied.

(2) A *bounded left-softlattice* is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ of type $(2, 2, 0, 0)$ such that the reduct (A^L, \wedge, \vee) is a left-softlattice and the elements 0 and 1 verify the axioms: for all $x \in A^L$,

$$\begin{array}{ll} (m\text{-WU}) & 1 \wedge x = x, & (m\text{-VU}) & 0 \vee x = x, \\ (m\text{-WL}) & 0 \wedge x = 0, & (m\text{-VL}) & 1 \vee x = 1 \end{array}$$

(i.e. such that the reduct $(A^L, \wedge, 0, 1)$ is a bounded meet-semilattice with top element 1 and the reduct $(A^L, \vee, 0, 1)$ is a bounded join-semilattice with bottom element 0).

We shall denote by **SL** the class of all left-softlattices and by **BSL** the class of all bounded left-softlattices.

Note that any left-lattice is a left-softlattice: $\mathbf{L} \subset \mathbf{SL}$. The inclusion is strict since, as the next example shows, there are left-softlattices that are not left-lattices.

Example 3.4. A proper softlattice

The algebra $\mathcal{A}^L = (L_2 = \{a, b\}, \wedge, \vee)$, with the following tables for \wedge and \vee , is a proper left-softlattice, because (m-Wabs-i) and (m-Vabs-i) are not verified for $(x, y) = (a, b)$.

\wedge	a	b	and	\vee	a	b	.
a	a	b		a	a	b	
b	b	b		b	b	b	

Hence, we have: $\mathbf{SL} + (\text{m-Wabs-i}) + (\text{m-Vabs-i}) = \mathbf{L} = \mathbf{SL} + (\text{m-Wabs}) + (\text{m-Vabs})$.

The following example is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ such that the reduct (A^L, \wedge, \vee) is a left-softlattice verifying (m-WU) and (m-VU), but not verifying (m-WL) and (m-VL), therefore it is not a bounded left-softlattice.

Example 3.5. The algebra $\mathcal{A}^L = (L_3 = \{0, a, 1\}, \wedge, \vee, 0, 1)$, with the following tables for \wedge and \vee , is a proper left-softlattice, because (m-Wabs-i) and (m-Vabs-i) are not verified for $(x, y) = (0, a)$. It verifies (m-WU) and (m-VU), but does not verify (m-WL) and (m-VL) for $x = a$.

\wedge	0	a	1	and	\vee	0	a	1
0	0	a	0		0	0	a	1
a	a	a	a		a	a	a	a
1	0	a	1		1	1	a	1

Proposition 3.6. Let $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ be a bounded left-softlattice. Define, for all $x, y \in A^L$,

$$(m\text{-df}O(\wedge)) \quad x \leq_m^{O(\wedge)} y \stackrel{\text{def.}}{\iff} x \wedge y = x \text{ and, dually,}$$

$$(m\text{-df}O(\vee)) \quad x \geq_m^{O(\vee)} y \stackrel{\text{def.}}{\iff} x \vee y = x.$$

Then,

- (1) $(A^L, \leq_m^{O(\wedge)}, 1)$ is an inf-semilattice with top element 1,
- (1') $(A^L, \geq_m^{O(\vee)}, 0)$ is a sup-semilattice with bottom element 0.
- (2) $0 \leq_m^{O(\wedge)} x \leq_m^{O(\wedge)} 1$, for all $x \in A^L$, i.e. the inf-semilattice is bounded,
- (2') $1 \geq_m^{O(\vee)} x \geq_m^{O(\vee)} 0$, for all $x \in A^L$, i.e. the sup-semilattice is bounded.

Proof. (1), (1'): Obvious.

(2): $0 \leq_m^{O(\wedge)} x \iff 0 \wedge x = 0$, which is true by (m-WL).

$x \leq_m^{O(\wedge)} 1 \iff x \wedge 1 = x$, which is true by (m-WU) and (m-Wcomm).

(2'): $1 \geq_m^{O(\vee)} x \iff 1 \vee x = 1$, which is true by (m-VL).

$x \geq_m^{O(\vee)} 0 \iff x \vee 0 = 0$, which is true by (m-VU) and (m-Vcomm). □

Proposition 3.7. *Let $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ be a bounded left-soflattice. Then,*

$$(m\text{-Wabs}) + (m\text{-Vabs}) \iff (x \leq_m^{O(\wedge)} y \iff y \geq_m^{O(\vee)} x).$$

Proof. Suppose (m-Wabs) and (m-Vabs) hold. Then, if $x \leq_m^{O(\wedge)} y$, i.e. $x \wedge y = x$, then

$$y \vee x = y \vee (x \wedge y) \stackrel{(m\text{-Wcomm})}{=} y \vee (y \wedge x) \stackrel{(m\text{-Vabs})}{=} y,$$

i.e. $y \geq_m^{O(\vee)} x$. Conversely, if $y \geq_m^{O(\vee)} x$, i.e. $y \vee x = y$, then

$$x \wedge y = x \wedge (y \vee x) \stackrel{(m\text{-Vcomm})}{=} x \wedge (x \vee y) \stackrel{(m\text{-Wabs})}{=} x,$$

i.e. $x \leq_m^{O(\wedge)} y$.

Conversely, suppose now that $x \leq_m^{O(\wedge)} y \iff y \geq_m^{O(\vee)} x$, i.e. (a) $x \wedge y = x \iff y \vee x = y$ or (b) $y \wedge x = y \iff x \vee y = x$. Then, $x \wedge (x \vee y) = x \wedge (y \vee x) = x \wedge y = x$, by (a), and $x \vee (x \wedge y) = x \vee (y \wedge x) = x \vee y = x$, by (b). Thus, (m-Wabs) and (m-Vabs) hold. □

Proposition 3.8. *Let $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ be a bounded left-soflattice. We have: for all $x, y, z \in A^L$,*

(1) $x \wedge y \leq_m^{O(\wedge)} x, y$,

(1') $x \vee y \geq_m^{O(\vee)} x, y$;

(2) if $x \leq_m^{O(\wedge)} y$, then $x \wedge z \leq_m^{O(\wedge)} y \wedge z$,

(2') if $x \geq_m^{O(\vee)} y$, then $x \vee z \geq_m^{O(\vee)} y \vee z$.

Proof. Routine. □

3.2.2 Bounded widelattices

Definitions 3.9. (The dual ones are omitted)

(1) A *left-widelattice* is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee)$ of type (2, 2) such that the axioms (m-Wcomm), (m-Vcomm), (m-Wass), (m-Vass), (m-Wabs-i), (m-Vabs-i) are satisfied.

(2) A *bounded left-widelattice* is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ of type $(2, 2, 0, 0)$ such that the reduct (A^L, \wedge, \vee) is a left-widelattice and the elements 0 and 1 verify the axioms: for all $x \in A^L$,

$$(m\text{-WU}) \quad 1 \wedge x = x, \quad (m\text{-VU}) \quad 0 \vee x = x.$$

We shall denote by **WL** the class of all left-widelattices and by **BWL** the class of all bounded left-widelattices.

Note that any left-lattice is a left-widelattice: $\mathbf{L} \subset \mathbf{WL}$. The inclusion is strict since, as the next example shows, there are left-widelattices that are not left-lattices.

Example 3.10. A proper widelattice

The algebra $\mathcal{A}^L = (L_3 = \{a, b, c\}, \wedge, \vee)$, with the following tables for \wedge and \vee , is a proper left-widelattice, because (m-Wid) and (m-Vid) are not verified for a .

\wedge	a	b	c		\vee	a	b	c
a	c	a	c	and	a	b	b	a
b	a	b	c		b	b	b	b
c	c	c	c		c	a	b	c

Hence, we have: $\mathbf{WL} + (m\text{-Wid}) + (m\text{-Vid}) = \mathbf{L}$,

$$\mathbf{SL} \cap \mathbf{WL} = \mathbf{L} \quad \text{and} \quad \mathbf{BSL} \cap \mathbf{BWL} = \mathbf{BL}.$$

Corollary 3.11. (See Corollary 2.3) Let $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ be a bounded left-widelattice. Then, we have the equivalences:

$$(m\text{-WU}) \iff (m\text{-VL}) \quad \text{and} \quad (m\text{-VU}) \iff (m\text{-WL}).$$

Proof. First, (m-WU) + (m-Wass) + (m-Vabs-i) imply (m-VL):

Indeed, in (m-Vabs-i) $(x \vee (x \wedge x \wedge y) = x)$, take $x := 1$, to obtain:

$$1 = 1 \vee (1 \wedge 1 \wedge y) = 1 \vee (1 \wedge (1 \wedge y)) = 1 \vee (1 \wedge y) = 1 \vee y;$$

thus (m-VL) holds.

Conversely, (m-VL) + (m-Vass) + (m-Vcomm) + (m-Wcomm) + (m-Wabs-i) imply (m-WU):

Indeed, in (m-Wabs-i) $(x \wedge (x \vee x \vee y) = x)$ take $y := 1$, to obtain:

$$x = x \wedge (x \vee x \vee 1) = x \wedge (x \vee (x \vee 1)) = x \wedge (x \vee (1 \vee x)) = x \wedge (x \vee 1) = x \wedge 1 = 1 \wedge x;$$

thus (m-WU) holds.

Dually, first, (m-VU) + (m-Vass) + (m-Wabs-i) imply (m-WL):

Indeed, in (m-Wabs-i), take $x := 0$, to obtain:

$$0 = 0 \wedge (0 \vee 0 \vee y) = 0 \wedge (0 \vee (0 \vee y)) = 0 \wedge (0 \vee y) = 0 \wedge y;$$

thus (m-WL) holds.

Conversely, (m-WL) + (m-Wass) + (m-Wcomm) + (m-Vcomm) + (m-Vabs-i) imply (m-VU):

Indeed, in (m-Vabs-i), take $y := 0$, to obtain:

$$x = x \vee (x \wedge x \wedge 0) = x \vee (x \wedge (x \wedge 0)) = x \vee (x \wedge (0 \wedge x)) = x \vee (x \wedge 0) = x \vee 0 = 0 \vee x;$$

thus (m-VU) holds. □

3.3 More on involutive m-MEL algebras

Proposition 3.12. *Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-MEL algebra. We have:*

$$(m\text{-Pabs-i}) \iff (m\text{-Sabs-i}).$$

Proof. Routine. □

Proposition 3.13. *We have:*

$$(m\text{CIM13}) \quad (m\text{-Pabs}) + (SU) \implies (m\text{-Pabs-i}),$$

$$(m\text{CIM14}) \quad (G) + (m\text{-Pabs-i}) \implies (m\text{-Pabs}),$$

$$(m\text{CIM15}) \quad (SU) \implies ((m\text{-Pabs}) \iff (G) + (m\text{-Pabs-i})).$$

Proof. Routine. □

Proposition 3.14. *Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-MEL algebra. We have:*

$$(m\text{-Pabs}) \iff (G) + (m\text{-Pabs-i}).$$

Proof. By (mCIM15). □

Proposition 3.15. *Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-MEL algebra. We have:*

$$(m\text{-Pimpl}) \iff (G) + (m\text{-Pabs-i}).$$

Proof. By ([16], Proposition 3.9), (m-Pimpl) \iff (m-Pabs), then apply Proposition 3.14 above. □

4 Two generalizations of bounded involutive lattices

We shall introduce, in the next two subsections, two generalizations of bounded involutive (involution) lattices: bounded involutive softlattices and bounded involutive widelattices.

4.1 Bounded involutive softlattices

Definition 4.1. (Definition 1) (The dual one is omitted) A *bounded involutive left-softlattice*, or a *left-BISL* for short, is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, - = {}^{-L}, 0, 1)$ such that the reduct $(A^L, \wedge, \vee, 0, 1)$ is a bounded left-softlattice (Definitions 3.3) and the unary operation $-$ satisfies (DN), (DeM1) and (DeM2).

Recall that in a left-BIL we have: (m-WU) \iff (m-VL) and (m-VU) \iff (m-WL), by Corollary 2.3. We shall denote by **BISL** the class of all left-BISLs. We have:

$$\mathbf{BISL} + (\text{m-Wabs-i}) + (\text{m-Vabs-i}) = \mathbf{BIL} = \mathbf{BISL} + (\text{m-Wabs}) + (\text{m-Vabs}).$$

Proposition 4.2. Let $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ be a left-BISL. The following are equivalent:

- (i) $y \geq_m^{O(\vee)} x \iff y^- \leq_m^{O(\wedge)} x^-$;
- (ii)=(DeM1) $(x \vee y)^- = x^- \wedge y^-$.

Proof. (i) \implies (ii): Suppose (i) holds. Since $x \vee y \geq_m^{O(\vee)} x$ by Proposition 3.8(1'), then $(x \vee y)^- \leq_m^{O(\wedge)} x^-$, by (i), and similarly, $(x \vee y)^- \leq_m^{O(\wedge)} y^-$; thus $(x \vee y)^-$ is a lower bound of $\{x^-, y^-\}$. Suppose z is a lower bound of $\{x^-, y^-\}$, i.e. $z \leq_m^{O(\wedge)} x^-, y^-$; then by (i) and (DN), $z^- \geq_m^{O(\vee)} x, y$, hence $z^- = z^- \vee z^- \geq_m^{O(\vee)} x \vee z^- \geq_m^{O(\vee)} x \vee y$, by Proposition 3.8 (2'), hence $z \leq_m^{O(\wedge)} (x \vee y)^-$. Hence $(x \vee y)^- = x^- \wedge y^-$, i.e. (ii)=(DeM1) holds.

(ii) \implies (i): Suppose $y \leq_M^{O(\vee)} x$, i.e. $y \vee x = y$; then $y^- \wedge x^- \stackrel{(ii)}{=} (y \vee x)^- = y^-$, i.e. $y^- \leq_m^{O(\wedge)} x^-$. Conversely, suppose $y^- \leq_m^{O(\wedge)} x^-$, i.e. $y^- \wedge x^- = y^-$; then, $y \vee x \stackrel{(DN)}{=} ((y \vee x)^-)^- \stackrel{(ii)}{=} (y^- \wedge x^-)^- = (y^-)^- \stackrel{(DN)}{=} y$, i.e. $y \geq_m^{O(\vee)} x$. \square

We have the following definitional equivalence:

Theorem 4.3.

- (1) Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-MEL algebra verifying (G). Define, for all $x, y \in A^L$, $x \wedge y \stackrel{\text{def.}}{=} x \odot y$, $x \vee y \stackrel{\text{def.}}{=} x \oplus y$, and $0 \stackrel{\text{def.}}{=} 1^-$, where $x \oplus y \stackrel{\text{def.}}{=} (x^- \odot y^-)^-$, by (1). Then $f(\mathcal{A}^L) = (A^L, \wedge, \vee, -, 0, 1)$ is a left-BISL.
- (1') Let $\mathcal{A}^L = (A^L, \wedge, \vee, -, 0, 1)$ be a left-BISL. Define, for all $x, y \in A^L$, $x \odot y \stackrel{\text{def.}}{=} x \wedge y$. Then $g(\mathcal{A}^L) = (A^L, \odot, -, 1)$ is an involutive left-m-MEL algebra verifying (G).

(2) The maps f and g are mutually inverse.

Proof. (1): Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-MEL algebra verifying (G). By ([16], Proposition 3.6), (m-Wcomm), (m-Wass), (m-WU), (m-WL) and (m-Vcomm), (m-Vass), (m-VU), (m-VL) and (DN), (DeM1), (DeM2) hold. By ([16], Proposition 3.7), since (G) holds, then (G^R) holds, hence (m-Wid) and (m-Vid) hold too. Thus, $(A^L, \wedge, \vee, -, 0, 1)$ is a left-BISL.

(1'): Let $\mathcal{A}^L = (A^L, \wedge, \vee, -, 0, 1)$ be a left-BISL. Define $x \odot y \stackrel{\text{def.}}{=} x \wedge y$. Then (PU) (= (m-WU)), (Pcomm) (= (m-Wcomm)), (Pass) (= (m-Wass)), (m-L) (= (m-WL)), (DN), (G) (= (m-Wid)) hold. Thus $(A^L, \odot, -, 1)$ is an involutive left-m-MEL algebra verifying (G).

(2): Routine. □

Theorem 4.3 allows us to give a new, equivalent definition of bounded involutive left-softlattices, as follows:

Definition 4.4. (Definition 2) (The dual one is omitted) A *bounded involutive left-softlattice*, or a *left-BISL* for short, is an involutive left-m-MEL algebra verifying (G), hence $\mathbf{BISL} = \mathbf{m-MEL}_{(DN)} + (G)$.

We shall denote by **aBISL** the class of all left-BISLs verifying (m-An). See Examples 6.1, 6.2 of proper BISL, aBISL respectively.

4.2 Bounded involutive widelattices

Definition 4.5. (Definition 1) (The dual one is omitted) A *bounded involutive left-widelattice*, or a *left-BIWL* for short, is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, - = {}^{-L}, 0, 1)$ such that the reduct $(A^L, \wedge, \vee, 0, 1)$ is a bounded left-widelattice (Definitions 3.9) and the unary operation $-$ satisfies (DN), (DeM1) and (DeM2).

Recall that in a left-BIL we have $(\text{m-WU}) \iff (\text{m-VL})$ and $(\text{m-VU}) \iff (\text{m-WL})$, by Corollary 2.3. We shall denote by **BIWL** the class of all left-BIWLs. We have:

$$\mathbf{BIWL} + (\text{m-Wid}) + (\text{m-Vid}) = \mathbf{BIL}.$$

We have the following definitional equivalence:

Theorem 4.6.

(1) Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-MEL algebra verifying (m-Pabs-i). Define, for all $x, y \in A^L$, $x \wedge y \stackrel{\text{def.}}{=} x \odot y$, $x \vee y \stackrel{\text{def.}}{=} x \oplus y$, and $0 \stackrel{\text{def.}}{=} 1^-$, where $x \oplus y \stackrel{\text{def.}}{=} (x^- \odot y^-)^-$ by (1). Then $f(\mathcal{A}^L) = (A^L, \wedge, \vee, -, 0, 1)$ is a left-BIWL.

(1') Let $\mathcal{A}^L = (A^L, \wedge, \vee, -, 0, 1)$ be a left-BIWL. Define, for all $x, y \in A^L$, $x \odot y \stackrel{\text{def.}}{=} x \wedge y$. Then $g(\mathcal{A}^L) = (A^L, \odot, -, 1)$ is an involutive left-m-MEL algebra verifying (m-Pabs-i).

(2) The maps f and g are mutually inverse.

Proof. (1): Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-MEL algebra verifying (m-Pabs-i). By ([16], Proposition 3.6), (m-Wcomm), (m-Wass), (m-WU), (m-WL) and (m-Vcomm), (m-Vass), (m-VU), (m-VL) and (DN), (DeM1), (DeM2) hold. By Proposition 3.12, since (m-Pabs-i) holds, then (m-Sabs-i) holds, hence (m-Wabs-i) and (m-Vabs-i) hold too. Thus, $(A^L, \wedge, \vee, -, 0, 1)$ is a left-BIWL.

(1'): Let $\mathcal{A}^L = (A^L, \wedge, \vee, -, 0, 1)$ be a left-BIWL. Define $x \odot y \stackrel{\text{def.}}{=} x \wedge y$. Then (PU) (= (m-WU)), (Pcomm) (= (m-Wcomm)), (Pass) (= (m-Wass)), (m-L) (= (m-WL)), (DN), (m-Pabs-i) (= (m-Wabs-i)) hold, by Corollary 3.11. Thus $(A^L, \odot, -, 1)$ is an involutive left-m-MEL algebra verifying (m-Pabs-i).

(2): Routine. □

Theorem 4.6 allows us to give a new, equivalent definition of bounded involutive left-widelattices, as follows:

Definition 4.7. (Definition 2) (The dual one is omitted) A *bounded involutive left-widelattice*, or a *left-BIWL* for short, is an involutive left-m-MEL algebra verifying (m-Pabs-i), hence **BIWL** = **m-MEL**_(DN) + (m-Pabs-i).

We shall denote by **aBIWL** the class of left-BIWLs verifying (m-An). See Examples 6.3, 6.4 of proper BIWL, aBIWL respectively.

4.3 Putting BILs and their two generalizations on the “involutive Little map”

The definitions (Definitions 2) and the results from this section allow us to draw the hierarchies from the following Figure 1, thus putting the mentioned algebras on the “involutive Little map”.

5 Two generalizations of ortholattices

In the next two subsections, we shall introduce two generalizations of ortholattices: orthosoftlattices and orthowidelattices.

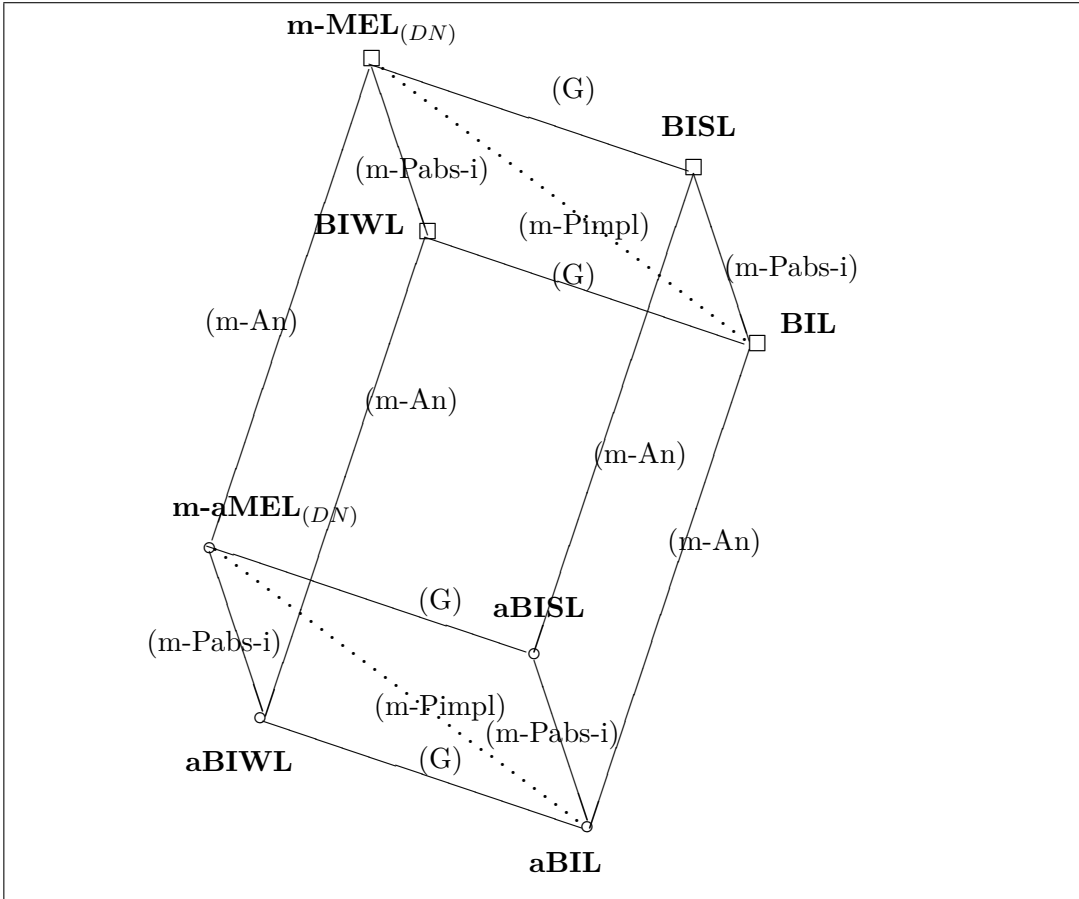


Figure 1: Hierarchies **BISL** vs. **BIWL**

5.1 Orthosoftlattices

Definition 5.1. (Definition 1) (The dual one is omitted) A *left-orthosoftlattice*, or a *left-OSL* for short, is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, - = {}^{-L}, 0, 1)$ such that the reduct $(A^L, \wedge, \vee, 0, 1)$ is a bounded left-softlattice (Definitions 3.3) and the unary operation $-$ satisfies (DN), (DeM1), (DeM2) and (m-WRe), (m-VRe).

We shall denote by **OSL** the class of all left-OSLs. We have:

$$\mathbf{OSL} + (\text{m-Wabs-i}) + (\text{m-Vabs-i}) = \mathbf{OL} = \mathbf{OSL} + (\text{m-Wabs}) + (\text{m-Vabs})$$

and, since $\mathbf{OL} = \mathbf{BIL} + (\text{m-WRe}) + (\text{m-VRe})$, it follows that $\mathbf{OSL} = \mathbf{BISL} + (\text{m-WRe}) + (\text{m-VRe})$.

We have the following definitional equivalence.

Theorem 5.2. (See Theorem 4.3)

- (1) Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-BE algebra verifying (G). Define, for all $x, y \in A^L$, $x \wedge y \stackrel{\text{def.}}{=} x \odot y$, $x \vee y \stackrel{\text{def.}}{=} x \oplus y = (x^- \wedge y^-)^-$, and $0 \stackrel{\text{def.}}{=} 1^-$. Then $f(\mathcal{A}^L) = (A^L, \wedge, \vee, -, 0, 1)$ is a left-OSL.
- (1') Let $\mathcal{A}^L = (A^L, \wedge, \vee, -, 0, 1)$ be a left-OSL. Define, for all $x, y \in A^L$, $x \odot y \stackrel{\text{def.}}{=} x \wedge y$. Then $g(\mathcal{A}^L) = (A^L, \odot, -, 1)$ is an involutive left-m-BE algebra verifying (G).
- (2) The mappings f and g are mutually inverse.

Proof. (1): Let $\mathcal{A}^L = (A^L, \odot, -, 1)$ be an involutive left-m-BE algebra verifying (G). Hence, \mathcal{A}^L is an involutive left-m-MEL algebra verifying (G) and, by Theorem 4.3 (1), $(A^L, \wedge, \vee, -, 0, 1)$ is a bounded involutive left-softlattice (left-BISL). Since (m-Re) and (m-Re^R) hold, then (m-WRe), (m-VRe) hold, hence $(A^L, \wedge, \vee, -, 0, 1)$ is a left-OSL.

(1'): Let $\mathcal{A}^L = (A^L, \wedge, \vee, -, 0, 1)$ be a left-OSL. Hence \mathcal{A}^L is a left-BISL verifying (m-WRe), (m-VRe). Define $x \odot y \stackrel{\text{def.}}{=} x \wedge y$. By Theorem 4.3(1'), $(A^L, \odot, -, 1)$ is an involutive left-m-MEL algebra verifying (G). Since (m-WRe), (m-VRe) hold, it follows that (m-Re), (m-Re^R) hold, thus $(A^L, \odot, -, 1)$ is an involutive left-m-BE algebra verifying (G).

(2): Routine. □

The above Theorem 5.2 allows us to give a new, equivalent definition of left-orthosoftlattices, as follows.

Definition 5.3. (Definition 2) (The dual one is omitted) A *left-orthosoftlattice*, or a *left-OSL* for short, is an involutive left-m-BE algebra verifying (G), hence **OSL** = **m-BE**_(DN) + (G).

Proposition 5.4. Let $\mathcal{A}^L = (A^L, \wedge, \vee, -, 0, 1)$ be a left-OSL. Consider the following dual binary relations:

$$\begin{aligned} (m\text{-dfrel}W) \quad x \leq_m y &\stackrel{\text{def.}}{\iff} x \wedge y^- = 0 \text{ and} \\ (m\text{-dfrel}V) \quad x \geq_m y &\stackrel{\text{def.}}{\iff} x \vee y^- = 1. \end{aligned}$$

Then we have:

- (1) $x \leq_m y \iff y \geq_m x$;
- (2) $x \leq_m^{O(\wedge)} y \implies x \leq_m y$, and
- (2') $x \geq_m^{O(\vee)} y \implies x \geq_m y$.

Proof. (1): By Theorem 5.2 and ([16], Proposition 3.10).

(2): Suppose $x \leq_m^{O(\wedge)} y$, i.e. $x \wedge y = x$ (see Proposition 3.6); then

$$x \wedge y^- = (x \wedge y) \wedge y^- \stackrel{(m\text{-Wass})}{=} x \wedge (y \wedge y^-) \stackrel{(m\text{-WRe})}{=} x \wedge 0 \stackrel{(m\text{-WL})}{=} 0;$$

thus $x \leq_m y$.

(2'): Dually, suppose $x \geq_m^{O(\vee)} y$, i.e. $x \vee y = x$ (see Proposition 3.6); then

$$x \vee y^- = (x \vee y) \vee y^- \stackrel{(m\text{-Vass})}{=} x \vee (y \vee y^-) \stackrel{(m\text{-VRe})}{=} x \vee 1 \stackrel{(m\text{-VL})}{=} 1;$$

thus $x \geq_m y$. □

Proposition 5.5. Let $\mathcal{A}^L = (A^L, \wedge, \vee, 0, 1)$ be a left-OSL.

(1) If (m-Wdiv) $(x \wedge (x \wedge y^-))^- = x \wedge y$ holds, then

$$x \leq_m^{O(\wedge)} y \iff x \leq_m y \quad (\iff y \geq_m x).$$

(1') If (m-Vdiv) $(x \vee (x \vee y^-))^- = x \vee y$ holds, then

$$x \geq_m^{O(\vee)} y \iff x \geq_m y \quad (\iff y \leq_m x).$$

Proof. (1): Suppose $x \leq_m y$, i.e. $x \wedge y^- = 0$; then

$$x \wedge y \stackrel{(m\text{-Wdiv})}{=} x \wedge (x \wedge y^-)^- = x \wedge 0^- \stackrel{(Neg0-1)}{=} x \wedge 1 \stackrel{(m\text{-Wcomm})}{=} 1 \wedge x \stackrel{(m\text{-WU})}{=} x,$$

i.e. $x \leq_m^{O(\wedge)} y$. The converse follows by Proposition 5.4(2).

(1'): Suppose $x \geq_m y$, i.e. $x \vee y^- = 1$; then

$$x \vee y \stackrel{(m\text{-Vdiv})}{=} x \vee (x \vee y^-)^- = x \vee 1^- \stackrel{(Neg1-0)}{=} x \vee 0 \stackrel{(m\text{-Vcomm})}{=} 0 \vee x \stackrel{(m\text{-VU})}{=} x,$$

i.e. $x \geq_m^{O(\vee)} y$. The converse follows by Proposition 5.4(2'). □

We shall denote by **tOSL** the class of all transitive left-OSLs (i.e. verifying (m-Tr)), by **aOSL** the class of all antisymmetric left-OSLs (i.e. verifying (m-An)) and by **taOSL** the class of all transitive and antisymmetric left-OSLs.

Note that by Definition 5.3, ([16], Proposition 4.7) and ([16], Definition 4.19), we have:

$$\mathbf{taOSL} = \mathbf{m\text{-BE}}_{(DN)} + (\mathbf{G}) + (\mathbf{m\text{-An}}) + (\mathbf{m\text{-BB}}) = \mathbf{m\text{-BE}}_{(DN)} + (\mathbf{m\text{-Pdiv}}) = \mathbf{Boole}.$$

See Examples 6.5, 6.6, 6.7 of proper OSL, tOSL, aOSL respectively.

5.2 Orthowidelattices

Definition 5.6. (Definition 1) (The dual one is omitted) A *left-orthowidelattice*, or a *left-OWL* for short, is an algebra $\mathcal{A}^L = (A^L, \wedge, \vee, \bar{} = \bar{}^L, 0, 1)$ such that the reduct $(A^L, \wedge, \vee, 0, 1)$ is a bounded left-widelattice (Definitions 3.9) and the unary operation $\bar{}$ satisfies (DN), (DeM1), (DeM2) and (m-WRe), (m-VRe).

We shall denote by **OWL** the class of all left-OWLs. We have:

$$\mathbf{OWL} + (\text{m-Wid}) + (\text{m-Vid}) = \mathbf{OL}$$

and, since $\mathbf{OL} = \mathbf{BIL} + (\text{m-WRe}) + (\text{m-VRe})$, it follows that $\mathbf{OWL} = \mathbf{BIWL} + (\text{m-WRe}) + (\text{m-VRe})$.

We have the following definitional equivalence:

Theorem 5.7. (See Theorem 4.6)

- (1) Let $\mathcal{A}^L = (A^L, \odot, \bar{}, 1)$ be an involutive left-m-BE algebra verifying (m-Pabs-i). Define, for all $x, y \in A^L$, $x \wedge y \stackrel{\text{def.}}{=} x \odot y$, $x \vee y \stackrel{\text{def.}}{=} x \oplus y$, and $0 \stackrel{\text{def.}}{=} 1^-$, where $x \oplus y \stackrel{\text{def.}}{=} (x^- \odot y^-)^-$ by (1). Then $f(\mathcal{A}^L) = (A^L, \wedge, \vee, \bar{}, 0, 1)$ is a left-OWL.
- (1') Let $\mathcal{A}^L = (A^L, \wedge, \vee, \bar{}, 0, 1)$ be a left-OWL. Define, for all $x, y \in A^L$, $x \odot y \stackrel{\text{def.}}{=} x \wedge y$. Then $g(\mathcal{A}^L) = (A^L, \odot, \bar{}, 1)$ is an involutive left-m-BE algebra verifying (m-Pabs-i).
- (2) The maps f and g are mutually inverse.

Proof. (1): Let $\mathcal{A}^L = (A^L, \odot, \bar{}, 1)$ be an involutive left-m-BE algebra verifying (m-Pabs-i). By ([16], Proposition 4.5), (m-WRe) and (m-VRe) hold. Then, \mathcal{A}^L is an involutive left-m-MEL algebra verifying (m-Pabs-i) and, by Theorem 4.6 (1), $(A^L, \wedge, \vee, \bar{}, 0, 1)$ is a bounded involutive left-widelattice (left-BIWL). Since (m-WRe), (m-VRe) hold, it follows that $(A^L, \wedge, \vee, \bar{}, 0, 1)$ is a left-orthowidelattice.

(1'): Let $\mathcal{A}^L = (A^L, \wedge, \vee, \bar{}, 0, 1)$ be a left-orthowidelattice. Hence, \mathcal{A}^L is a bounded involutive left-widelattice verifying (m-WRe), (m-VRe). Define $x \odot y \stackrel{\text{def.}}{=} x \wedge y$. By Theorem 4.6 (1'), $(A^L, \odot, \bar{}, 1)$ is an involutive left-m-MEL algebra verifying (m-Pabs-i). Since (m-WRe), (m-VRe) hold, it follows that (m-Re), (m-Re^R) hold, thus $(A^L, \odot, \bar{}, 1)$ is an involutive left-m-BE algebra verifying (m-Pabs-i).

(2): Routine. □

The above Theorem 5.7 allows us to give a new, equivalent definition of left-orthowidelattices, as follows:

Definition 5.8. (Definition 2) (The dual one is omitted) A *left-orthowidelattice*, or a *left-OWL* for short, is an involutive left-m-BE algebra verifying (m-Pabs-i), hence $\mathbf{OWL} = \mathbf{m-BE}_{(DN)} + (\text{m-Pabs-i})$.

We shall denote by \mathbf{tOWL} the class of all transitive left-OWls (i.e. verifying (m-Tr)), by \mathbf{aOWL} the class of all antisymmetric left-OWls (i.e. verifying (m-An)) and by \mathbf{taOWL} the class of all antisymmetric and transitive left-OWls.

See Example 6.8 of proper OWL and Examples 6.9 of two proper tOWls.

A problem we have not been able to resolve is the following.

Open problem 5.9. *Find an example of antisymmetric orthowidelattice which does not verify (m-Tr) ($\iff \dots(\text{m-BB})$), i.e. a proper element of \mathbf{aOWL} (using Mace4, we have searched exhaustively for an example up through and including size 24), or prove that an involutive left-m-aBE algebra satisfying (m-Pabs-i) satisfies also (m-Tr) (we have also tried to find a proof using Prover9, but despite letting it run for several days, it was unable to find one).*

5.2.1 Transitive, antisymmetric OWls: taOWL

Here we shall prove Theorem 5.14, saying that taOWls are MV algebras, which is *the core of this paper*. A direct proof of Theorem 5.14 by PROVER9 took 58321 seconds (about 17 hours) and had 205 lines (length of the proof). The shortest proof that we were able to find has two steps:

Step 1: in an involutive m-BE algebra verifying (m-Pabs-i) and (m-BB) ($\iff \dots(\text{m-Tr})$), but not necessarily (m-An), we have the property:

$$(\Delta_m) \quad (x \wedge_m y) \odot (y \wedge_m x)^- = 0,$$

where $x \wedge_m y \stackrel{\text{def.}}{=} (x^- \odot y)^- \odot y$; this is Theorem 5.13, whose proof by PROVER9 had 'only' 178 lines (intermediary results = steps of the proof) which, after making a graph of their dependencies, were grouped into Lemmas 5.10, 5.11, 5.12 and Theorem 5.13 below, with their corresponding "humanized" proofs.

Step 2: In an involutive m-BE algebra verifying (m-Pabs-i), (m-BB) and (m-An), we have $(\wedge_m\text{-comm})$; this is Theorem 5.14, which follows easily by Theorem 5.13, since $(\Delta_m) + (\text{m-An}) \iff (\wedge_m\text{-comm})$ (indeed, $(\wedge_m\text{-comm})$ implies (m-An), by ([12], (mB2))).

Lemma 5.10. *Let $\mathcal{A}^L = (A^L, \odot, ^-, 1)$ be an involutive left-m-BE algebra. We have:*

$$x \odot (y \odot z) = y \odot (x \odot z), \tag{2}$$

$$x \odot (y \oplus x^-) = y \wedge_m x, \tag{3}$$

$$(x \odot (y \odot z^-))^- = (x \odot y)^- \oplus z, \tag{4}$$

$$\begin{aligned}
 x^- \odot (x \wedge_m y) &= 0, & (5) \\
 (x \odot (y^- \odot z))^- &= y \oplus (x \odot z)^-, & (6) \\
 x \odot ((y \oplus x^-) \odot z) &= (y \wedge_m x) \odot z, & (7) \\
 x \odot (y \odot (z \oplus (x \odot y)^-)) &= z \wedge_m (x \odot y), & (8) \\
 x \odot (y \odot (z \oplus x^-)) &= y \odot (z \wedge_m x), & (9) \\
 x \oplus (y \oplus x)^- &= (y \wedge_m x^-)^-, & (10) \\
 x \odot (x \oplus y)^- &= 0, & (11) \\
 x \oplus (x \odot y)^- &= 1, & (12) \\
 x^- \odot ((x \wedge_m y) \odot z) &= 0, & (13) \\
 x \odot (y \wedge_m x^-) &= 0, & (14) \\
 x \oplus (y \odot x)^- &= 1, & (15) \\
 x \odot (y \odot x)^- &= y^- \wedge_m x, & (16) \\
 x \oplus (x \wedge_m y)^- &= 1, & (17) \\
 (x \oplus y^-) \odot (z \wedge_m y) &= (x \wedge_m y) \odot (z \oplus y^-), & (18) \\
 x \odot (x \odot y)^- &= y^- \wedge_m x, & (19) \\
 (x^- \wedge_m y^-)^- &= y \oplus (x \odot y^-), & (20) \\
 x \odot (y \wedge_m x)^- &= (y^- \odot x) \wedge_m x, & (21) \\
 x \oplus ((y \wedge_m x) \odot z)^- &= 1, & (22) \\
 x \wedge_m (y \odot x) &= y \odot x, & (23) \\
 (x \oplus (y \odot x^-))^- &= y^- \wedge_m x^-, & (24) \\
 x \odot (y \oplus (z \odot x)^-) &= (z^- \oplus y) \wedge_m x, & (25) \\
 x^- \wedge_m (x \oplus y)^- &= (x \oplus y)^-, & (26) \\
 x \wedge_m ((x \wedge_m y) \odot z) &= (x \wedge_m y) \odot z, & (27) \\
 x \odot ((x^- \oplus y) \wedge_m z) &= y \wedge_m (x \odot z), & (28) \\
 (x^- \wedge_m y)^- \wedge_m x &= x, & (29) \\
 x \odot (y \oplus (x^- \oplus z)) &= (y \oplus z) \wedge_m x, & (30) \\
 x \odot ((x \odot y)^- \oplus z) &= (y^- \oplus z) \wedge_m x, & (31) \\
 (x \wedge_m y)^- \wedge_m x^- &= x^-, & (32) \\
 (x \oplus y) \wedge_m x &= x, & (33) \\
 (x \wedge_m y^-)^- \wedge_m y &= y, & (34) \\
 ((x^- \odot y) \wedge_m y) \odot z &= y \odot ((x \wedge_m y)^- \odot z), & (35) \\
 x \oplus (y \wedge_m (z \wedge_m x))^- &= 1, & (36)
 \end{aligned}$$

$$x \wedge_m (y \wedge_m (x \wedge_m z)) = y \wedge_m (x \wedge_m z), \quad (37)$$

$$x \wedge_m (y \odot (z \wedge_m (x \wedge_m u))) = y \odot (z \wedge_m (x \wedge_m u)), \quad (38)$$

$$(x \oplus (y \odot x^-)) \wedge_m y = y, \quad (39)$$

$$x \odot (y \oplus (z \wedge_m x)^-) = (y \oplus (z^- \odot x)) \wedge_m x, \quad (40)$$

$$(x \odot y)^- \wedge_m x^- = x^-, \quad (41)$$

$$x \odot (y \odot ((z \oplus x^-) \wedge_m u)) = y \odot (z \wedge_m (x \odot u)), \quad (42)$$

$$x \wedge_m (y \wedge_m (z \wedge_m x)) = y \wedge_m (z \wedge_m x), \quad (43)$$

$$x \wedge_m (y \wedge_m (z \wedge_m (x \wedge_m u))) = y \wedge_m (z \wedge_m (x \wedge_m u)). \quad (44)$$

Proof. We shall prove most of the above properties:

$$(13): x^- \odot ((x \wedge_m y) \odot z) \stackrel{(7)}{=} x^- \odot [y \odot ((x \oplus y^-) \odot z)] = x^- \odot y \odot (x^- \odot y)^- \odot z \stackrel{(m-Re), (m-L)}{=} 0.$$

$$(18): (x \oplus y^-) \odot (z \wedge_m y) \stackrel{(9)}{=} y \odot ((x \oplus y^-) \odot (z \oplus y^-)) \text{ and } (x \wedge_m y) \odot (z \oplus y^-) \stackrel{(7)}{=} y \odot ((x \oplus y^-) \odot (z \oplus y^-)).$$

$$(21): x \odot (y \wedge_m x)^- \stackrel{(10)}{=} x \odot (x^- \oplus (y \oplus x^-)^-) = x \odot (x^- \oplus (y^- \odot x)) \text{ and } (y^- \odot x) \wedge_m x = x \odot (x^- \oplus (y^- \odot x)).$$

$$(22): x \oplus ((y \wedge_m x) \odot z)^- \stackrel{(7)}{=} x \oplus [x \odot ((y \oplus x^-) \odot z)]^- \stackrel{(12)}{=} 1.$$

$$(23): x \wedge_m (y \odot x) \stackrel{(8)}{=} y \odot (x \odot (x \oplus (y \odot x)^-)) \stackrel{(15)}{=} y \odot (x \odot 1) = y \odot x.$$

$$(25): (z^- \oplus y) \wedge_m x = x \odot ((z^- \oplus y)^- \odot x)^- = x \odot (z^- \oplus y \oplus x^-) = x \odot (y \oplus (z^- \oplus x^-)) = x \odot (y \oplus (z \odot x)^-).$$

$$(26): x^- \wedge_m (x \oplus y)^- = x^- \wedge_m (x^- \odot y^-) = x^- \wedge_m (y^- \odot x^-) \stackrel{(23)}{=} y^- \odot x^- = (x \oplus y)^-.$$

$$(27): \text{Take } X := (x \wedge_m y) \odot z \text{ in (16); we obtain: } x \wedge_m ((x \wedge_m y) \odot z)$$

$$\stackrel{(16)}{=} X \odot (x^- \odot X)^- = X \odot (x^- \odot ((x \wedge_m y) \odot z))^- \stackrel{(13)}{=} X \odot 0^- = X \odot 1 = X = (x \wedge_m y) \odot z.$$

$$(28): y \wedge_m (x \odot z) \stackrel{(8)}{=} x \odot (z \odot (y \oplus (x \odot z)^-)) \stackrel{(25)}{=} x \odot ((x^- \oplus y) \wedge_m z).$$

$$(29): (x^- \wedge_m y)^- \wedge_m x = x \odot (x^- \oplus (x^- \wedge_m y)^-) \stackrel{(17)}{=} x \odot 1 = x.$$

$$(30): x \odot (y \oplus (x^- \oplus z)) = x \odot (x^- \oplus (y \oplus x)) = (y \oplus z) \wedge_m x.$$

$$(31): x \odot ((x \odot y)^- \oplus z) \stackrel{(4)}{=} x \odot (x \odot (y \odot z^-))^- (y \odot z^-)^- \wedge_m x = (y^- \oplus z) \wedge_m x.$$

$$(32): (x \wedge_m y)^- \wedge_m x^- = x^- \odot (x^- \odot (x \wedge_m y))^- \stackrel{(5)}{=} x^- \odot 0^- = x^-.$$

$$(33): (x \oplus y) \wedge_m x = x \odot (x \odot (x \oplus y)^-)^- \stackrel{(11)}{=} x \odot 0^- = x \odot 1 = x.$$

$$(34): (x \wedge_m y^-)^- \wedge_m y = y \odot (y \odot (x \wedge_m y^-))^- \stackrel{(14)}{=} y \odot 0^- = y.$$

$$(35): ((x^- \odot y) \wedge_m y) \odot z \stackrel{(21)}{=} (y \odot (x \wedge_m y)^-) \odot z.$$

$$(36): x \oplus (y \wedge_m (z \wedge_m x))^- \stackrel{(3)}{=} x \oplus (X \odot (y \oplus X^-))^- = x \oplus ((z \wedge_m x) \odot (y \oplus (z \wedge_m x)^-))^- \stackrel{(22)}{=} 1.$$

$$(37): x \wedge_m (y \wedge_m (x \wedge_m z)) \stackrel{(3)}{=} x \wedge_m (X \odot (y \oplus X^-)) = x \wedge_m ((x \wedge_m z) \odot (y \oplus (x \wedge_m z)^-)) \stackrel{(27)}{=} (x \wedge_m z) \odot (y \oplus (x \wedge_m z)^-) \stackrel{(3)}{=} y \wedge_m (x \wedge_m z).$$

$$(38): \text{Put } X := x \wedge_m u \text{ in (9); we obtain: } y \odot (z \wedge_m (x \wedge_m u)) \stackrel{(9)}{=} X \odot (y \odot (z \oplus X^-)) = (x \wedge_m u) \odot (y \odot (z \oplus (x \wedge_m u)^-)) \stackrel{(27)}{=} x \wedge_m ((x \wedge_m u) \odot Z) = x \wedge_m ((x \wedge_m u) \odot (y \odot (z \oplus (x \wedge_m u)^-))) \stackrel{(9)}{=} x \wedge_m (y \odot (z \wedge_m X)) = x \wedge_m (y \odot (z \wedge_m (x \wedge_m u))), \text{ for } Z := y \odot (z \oplus (x \wedge_m u)^-) \text{ and } X := x \wedge_m u \text{ again.}$$

$$(39): (x \oplus (y \odot x^-)) \wedge_m y \stackrel{(20)}{=} (y^- \wedge_m x^-)^- \wedge_m y \stackrel{(29)}{=} y.$$

$$(40): \text{Put } Z := z^- \odot x; \text{ then, we have: } (y \oplus (z^- \odot x)) \wedge_m x \stackrel{(30)}{=} x \odot (y \oplus (x^- \oplus Z)) = x \odot (y \oplus (x^- \oplus (z^- \odot x))) = x \odot (y \oplus x^- \oplus (z \oplus x^-)^-) \stackrel{(10)}{=} x \odot (y \oplus (z \wedge_m x)^-).$$

$$(41): (x \odot y)^- \wedge_m x^- = (x^- \oplus y^-) \wedge_m x^- \stackrel{(33)}{=} x^-.$$

$$(42): \text{Put } Z := (z \oplus x^-) \wedge_m u, \text{ to obtain: } x \odot (y \odot ((z \oplus x^-) \wedge_m u)) \stackrel{(2)}{=} y \odot (x \odot Z) = y \odot (x \odot ((z \oplus x^-) \wedge_m u)) = y \odot (x \odot ((x^- \oplus z) \wedge_m u)) \stackrel{(28)}{=} y \odot (z \wedge_m (x \odot u)).$$

$$(43): x \wedge_m (y \wedge_m (z \wedge_m x)) \stackrel{(3)}{=} (y \wedge_m (z \wedge_m x)) \odot (x \oplus (y \wedge_m (z \wedge_m x))^-) \stackrel{(36)}{=} (y \wedge_m (z \wedge_m x)) \odot 1 = y \wedge_m (z \wedge_m x).$$

$$(44): y \wedge_m (z \wedge_m (x \wedge_m u)) \stackrel{(37)}{=} y \wedge_m [x \wedge_m (z \wedge_m (x \wedge_m u))] \stackrel{(37)}{=} x \wedge_m (y \wedge_m [x \wedge_m (z \wedge_m (x \wedge_m u))]) = x \wedge_m (y \wedge_m (z \wedge_m (x \wedge_m u))). \quad \square$$

Lemma 5.11. *Let $\mathcal{A}^L = (A^L, \odot, ^-, 1)$ be an involutive left- m -BE algebra. If (m -BB) holds (i.e. if \mathcal{A}^L is a left- m -pre-BCK $_{(DN)}$ algebra), then we have:*

$$(x \odot y^-)^- \odot (z \odot ((y \odot z)^- \odot x)) = 0, \tag{45}$$

$$(x^- \oplus y) \odot (z \odot ((y \odot z)^- \odot x)) = 0, \tag{46}$$

$$(x^- \oplus y) \odot (z \odot (x \odot (y \odot z)^-)) = 0, \tag{47}$$

$$x \odot ((x^- \oplus y) \odot (y^- \wedge_m z)) = 0, \tag{48}$$

$$(x \wedge_m y) \odot (x^- \wedge_m z) = 0, \tag{49}$$

$$(x^- \wedge_m y)^- \wedge_m (x \wedge_m z) = x \wedge_m z, \tag{50}$$

$$x \odot ((x^- \wedge_m y) \wedge_m z) = 0, \tag{51}$$

$$x^- \odot ((x \wedge_m y) \wedge_m z) = 0, \tag{52}$$

$$(x \oplus (y \odot x^-)) \wedge_m (y \wedge_m z) = y \wedge_m z, \tag{53}$$

$$x \wedge_m ((x \wedge_m y) \wedge_m z) = (x \wedge_m y) \wedge_m z, \tag{54}$$

$$(x \wedge_m y)^- \odot (x \wedge_m (y \odot z)) = 0, \quad (55)$$

$$x \oplus ((x \wedge_m y) \wedge_m z)^- = 1, \quad (56)$$

$$x \wedge_m ((x \odot y) \wedge_m z) = (x \odot y) \wedge_m z, \quad (57)$$

$$x \wedge_m ((y \wedge_m x) \wedge_m z) = (y \wedge_m x) \wedge_m z, \quad (58)$$

$$(x \odot y) \wedge_m (y \odot (x \wedge_m z)) = y \odot (x \wedge_m z), \quad (59)$$

$$x \wedge_m ((y \wedge_m (x \wedge_m z) \wedge_m u) = (y \wedge_m (x \wedge_m z)) \wedge_m u, \quad (60)$$

$$(x \wedge_m y) \oplus (x \wedge_m (y \odot z))^- = 1, \quad (61)$$

$$(x \wedge_m y)^- \odot (x \wedge_m (z \wedge_m y)) = 0, \quad (62)$$

$$x \odot (y \odot (y \wedge_m x)^-) = 0, \quad (63)$$

$$(x \oplus ((x \wedge_m y)^- \odot z)) \wedge_m z = z, \quad (64)$$

$$x \wedge_m ((y \wedge_m (z \odot x)) \wedge_m u) = (y \wedge_m (z \odot x)) \wedge_m u, \quad (65)$$

$$(x \wedge_m y) \wedge_m (x \wedge_m (z \wedge_m y)) = x \wedge_m (z \wedge_m y), \quad (66)$$

$$(x \wedge_m y)^- \odot ((x \wedge_m z) \wedge_m y) = 0, \quad (67)$$

$$(x \wedge_m y)^- \odot ((z \wedge_m x) \wedge_m y) = 0, \quad (68)$$

$$x^- \wedge_m (y \odot (y \wedge_m x)^-) = y \odot (y \wedge_m x)^-, \quad (69)$$

$$(x^- \oplus ((x \oplus y) \odot z)) \wedge_m z = z, \quad (70)$$

$$((x \wedge_m y)^- \oplus (x \odot z)) \wedge_m z = z, \quad (71)$$

$$(x \wedge_m y)^- \wedge_m (z \odot (z \wedge_m x)^-) = z \odot (z \wedge_m x)^-, \quad (72)$$

$$((x \oplus y) \odot z) \wedge_m (x \odot z) = x \odot z, \quad (73)$$

$$((x \oplus y) \odot z) \oplus ((x \odot z) \wedge_m u)^- = 1. \quad (74)$$

Proof. (48): $x \odot ((x^- \oplus y) \odot (y^- \wedge_m z)) \stackrel{(16)}{=} x \odot ((x^- \oplus y) \odot z \odot (y \odot z)^-) \stackrel{(47)}{=} 0.$

(49): $(x \wedge_m y) \odot (x^- \wedge_m z) \stackrel{(7)}{=} y \odot ((y^- \oplus x) \odot (x^- \wedge_m z)) \stackrel{(48)}{=} 0.$

(50): $(x^- \wedge_m y)^- \wedge_m (x \wedge_m z) \stackrel{(19)}{=} (x \wedge_m z) \odot ((x \wedge_m z) \odot (x^- \wedge_m y))^- \stackrel{(49)}{=} (x \wedge_m z) \odot 0^- = (x \wedge_m z) \odot 1 = x \wedge_m z.$

(51): $x \odot ((x^- \wedge_m y) \wedge_m z) \stackrel{(29)}{=} ((x^- \wedge_m y)^- \wedge_m x) \odot ((x^- \wedge_m y) \wedge_m z) \stackrel{(49)}{=} 0.$

(52): $x^- \odot ((x \wedge_m y) \wedge_m z) \stackrel{(32)}{=} ((x \wedge_m y)^- \wedge_m x^-) \odot ((x \wedge_m y) \wedge_m z) \stackrel{(49)}{=} 0.$

(53): $(x \oplus (y \odot x^-)) \wedge_m (y \wedge_m z) \stackrel{(20)}{=} (y^- \wedge_m x^-)^- \wedge_m (y \wedge_m z) \stackrel{(50)}{=} y \wedge_m z.$

(54): $x \wedge_m ((x \wedge_m y) \wedge_m z) \stackrel{(32)}{=} ((x \wedge_m y)^- \wedge_m x^-)^- \wedge_m ((x \wedge_m y) \wedge_m z) \stackrel{(50)}{=} (x \wedge_m y) \wedge_m z.$

(55): $(x \wedge_m y)^- \odot (x \wedge_m (y \odot z)) \stackrel{(42)}{=} y \odot ((x \wedge_m y)^- \odot ((x \oplus y^-) \wedge_m z))$

$= y \odot ((x \wedge_m y)^- \odot ((x^- \odot y)^- \wedge_m z)) \stackrel{(35)}{=} ((x^- \odot y) \wedge_m y) \odot Z$

$= ((x^- \odot y) \wedge_m y) \odot ((x^- \odot y)^- \wedge_m z) \stackrel{(49)}{=} 0, \text{ for } Z := (x^- \odot y)^- \wedge_m z.$

$$\begin{aligned}
 (56): & x \oplus ((x \wedge_m y) \wedge_m z)^- = x^- \odot ((x \wedge_m y) \wedge_m z) \stackrel{(51)}{=} 0. \\
 (57): & x \wedge_m ((x \odot y) \wedge_m z) \stackrel{(41)}{=} ((x \odot y)^- \wedge_m x^-)^- \wedge_m ((x \odot y) \wedge_m z) \stackrel{(50)}{=} (x \odot y) \wedge_m z. \\
 (58): & x \wedge_m ((y \wedge_m x) \wedge_m z) \stackrel{(34)}{=} ((y \wedge_m x)^- \wedge_m x^-)^- \wedge_m ((y \wedge_m x) \wedge_m z) \stackrel{(50)}{=} (y \wedge_m x) \wedge_m z. \\
 (59): & (x \odot y) \wedge_m (y \odot (x \wedge_m z)) \stackrel{(28)}{=} y \odot ((y^- \oplus (x \odot y)) \wedge_m (x \wedge_m z)) \stackrel{(53)}{=} y \odot (x \wedge_m z). \\
 (60): & (y \wedge_m (x \wedge_m z)) \wedge_m u \stackrel{(37)}{=} (x \wedge_m (y \wedge_m (x \wedge_m z))) \wedge_m u \\
 & \stackrel{(54)}{=} x \wedge_m ((x \wedge_m (y \wedge_m (x \wedge_m z))) \wedge_m u) \stackrel{(37)}{=} x \wedge_m ((y \wedge_m (x \wedge_m z)) \wedge_m u). \\
 (62): & (x \wedge_m y)^- \odot (x \wedge_m (z \wedge_m y)) \stackrel{(3)}{=} (x \wedge_m y)^- \odot (x \wedge_m (y \odot (z \oplus y^-))) \stackrel{(55)}{=} 0. \\
 (63): & x \odot (y \odot (y \wedge_m x)^-) = (y \wedge_m x)^- \odot (x \odot y) \stackrel{(23)}{=} (y \wedge_m x)^- \odot (y \wedge_m (x \odot y)) \stackrel{(55)}{=} 0. \\
 (64): & (x \oplus ((x \wedge_m y)^- \odot z)) \wedge_m z \stackrel{(40)}{=} z \odot (x \oplus ((x \wedge_m y) \wedge_m z)^-) \stackrel{(56)}{=} z \odot 1 = z. \\
 (65): & (y \wedge_m (z \odot x)) \wedge_m u \stackrel{(58)}{=} (z \odot x) \wedge_m ((y \wedge_m (z \odot x)) \wedge_m u) \\
 & = (x \odot z) \wedge_m ((y \wedge_m (z \odot x)) \wedge_m u) \stackrel{(57)}{=} x \wedge_m ((z \odot x) \wedge_m ((y \wedge_m (z \odot x)) \wedge_m u)) \\
 & \stackrel{(58)}{=} x \wedge_m ((y \wedge_m (z \odot x)) \wedge_m u). \\
 (66): & (x \wedge_m y) \wedge_m (x \wedge_m (z \wedge_m y)) \stackrel{(16)}{=} (x \wedge_m (z \wedge_m y)) \odot ((x \wedge_m y)^- \odot (x \wedge_m (z \wedge_m y)))^- \\
 & \stackrel{(62)}{=} (x \wedge_m (z \wedge_m y)) \odot 0^- = (x \wedge_m (z \wedge_m y)) \odot 1 = x \wedge_m (z \wedge_m y). \\
 (67): & (x \wedge_m y)^- \odot ((x \wedge_m z) \wedge_m y) \stackrel{(54)}{=} (x \wedge_m y)^- \odot (x \wedge_m ((x \wedge_m z) \wedge_m y)) \stackrel{(62)}{=} 0. \\
 (68): & (x \wedge_m y)^- \odot ((z \wedge_m x) \wedge_m y) \stackrel{(58)}{=} (x \wedge_m y)^- \odot (x \wedge_m ((z \wedge_m x) \wedge_m y)) \stackrel{(62)}{=} 0. \\
 (69): & x^- \wedge_m (y \odot (y \wedge_m x)^-) \stackrel{(16)}{=} (y \odot (y \wedge_m x)^-) \odot (x \odot (y \odot (y \wedge_m x)^-))^- \stackrel{(63)}{=} \\
 & (y \odot (y \wedge_m x)^-) \odot 0^- = y \odot (y \wedge_m x)^-. \\
 (70): & (x^- \oplus ((x \oplus y) \odot z)) \wedge_m z \stackrel{(26)}{=} (x^- \oplus ((x^- \wedge_m (x \oplus y)^-)^- \odot z)) \wedge_m z \stackrel{(64)}{=} z. \\
 (71): & ((x \wedge_m y)^- \oplus (x \odot z)) \wedge_m z \stackrel{(32)}{=} ((x \wedge_m y)^- \oplus (((x \wedge_m y)^- \wedge_m x^-)^- \odot z)) \wedge_m z \stackrel{(64)}{=} z. \\
 (72): & (x \wedge_m y)^- \wedge_m (z \odot (z \wedge_m x)^-) \stackrel{(69)}{=} (x \wedge_m y)^- \wedge_m [x^- \wedge_m (z \odot (z \wedge_m x)^-)] \\
 & \stackrel{(50)}{=} x^- \wedge_m (z \odot (z \wedge_m x)^-) \stackrel{(69)}{=} z \odot (z \wedge_m x)^-. \\
 (73): & \text{Put } X := x, Y := (x \oplus y) \odot z \text{ and } Z := z \text{ in (28), to obtain:} \\
 & ((x \oplus y) \odot z) \wedge_m (x \odot z) \stackrel{(28)}{=} x \odot ((x^- \oplus ((x \oplus y) \odot z)) \wedge_m z) \stackrel{(70)}{=} x \odot z. \\
 (74): & \text{Put } X := (x \oplus y) \odot z \text{ and } Y := x \odot z \text{ in (73); hence, } X \wedge_m Y = x \odot z. \text{ Then} \\
 & ((x \oplus y) \odot z) \oplus ((x \odot z) \wedge_m u)^- \stackrel{(73)}{=} ((x \oplus y) \odot z) \oplus (((x \oplus y) \odot z) \wedge_m (x \odot z)) \wedge_m u)^- \stackrel{(56)}{=} 1. \quad \square
 \end{aligned}$$

Lemma 5.12. *Let $\mathcal{A}^L = (A^L, \odot, ^-, 1)$ be an involutive left- m -BE algebra. If (m -Pabs- i) holds (i.e. if \mathcal{A}^L is a left-OWL), then we have:*

$$x \odot (x \oplus (y \oplus x)) = x, \quad (75)$$

$$x \odot (x \oplus x) = x, \quad (76)$$

$$x \oplus (x \odot x) = x, \quad (77)$$

$$(x \oplus x) \odot (y \wedge_m x) = y \wedge_m x, \quad (78)$$

$$(x \odot y) \oplus (x \odot (y \odot (x \odot y))) = x \odot y, \quad (79)$$

$$x \odot (x \odot x) = x \odot x, \quad (80)$$

$$(x^- \oplus y) \wedge_m (x \odot x) = y \wedge_m (x \odot x), \quad (81)$$

$$x \odot (y \wedge_m (x \odot x)) = y \wedge_m (x \odot x), \quad (82)$$

$$(x \odot y) \wedge_m (y \odot x) = y \odot x. \quad (83)$$

Proof. (75): By (Scomm).

(76): From (m-Pabs-i), taking $y = 0$.

(77): It is the dual of (76).

$$(78): (x \oplus x) \odot (y \wedge_m x) = (x \oplus x) \odot (x \odot (y^- \odot x^-)) \stackrel{(Pass)}{=} ((x \oplus x) \odot x) \odot (y^- \odot x)^- \stackrel{(76)}{=} x \odot (y^- \odot x)^- = y \wedge_m x.$$

(79): By (77), for $X := x \odot y$.

$$(80): x \odot (x \odot x) \stackrel{(77)}{=} (x \oplus (x \odot x)) \odot (x \odot x) \stackrel{(77)}{=} (x \oplus (x \odot x) \oplus (x \odot x)) \odot (x \odot x) \stackrel{(m-Pabs-i)}{=} x \odot x.$$

$$(81): \text{Put } X := x \odot x \text{ in (25); we obtain: } (x^- \oplus y) \wedge_m (x \odot x) \stackrel{(25)}{=} X \odot (y \oplus (x \odot X)^-) = (x \odot x) \odot (y \oplus (x \odot (x \odot x))^-) \stackrel{(80)}{=} (x \odot x) \odot (y \oplus (x \odot x)^-) \stackrel{(3)}{=} y \wedge_m (x \odot x).$$

$$(82): y \wedge_m (x \odot x) \stackrel{(80)}{=} y \wedge_m (x \odot (x \odot x)) \stackrel{(28)}{=} x \odot ((x^- \oplus y) \wedge_m z) = x \odot ((x^- \oplus y) \wedge_m (x \odot x)) \stackrel{(81)}{=} x \odot (y \wedge_m (x \odot x)).$$

$$(83): (x \odot y) \wedge_m (y \odot x) \stackrel{(79)}{=} [(y \odot x) \oplus (x \odot (y \odot (x \odot y)))] \wedge_m (y \odot x) \stackrel{(33)}{=} y \odot x. \quad \square$$

Theorem 5.13. Let $\mathcal{A}^L = (A^L, \odot, ^-, 1)$ be an involutive left- m -BE algebra. If (m-BB) and (m-Pabs-i) hold (i.e. if \mathcal{A}^L is a left-tOWL), then (Δ_m) holds.

Proof. The proof has 47 steps. Here they are:

$$x^- \wedge_m ((x \wedge_m y) \odot (x \wedge_m y)) = 0. \quad (84)$$

(84): Put $X := x \wedge_m y$ and $Y := x^-$ in (82), to obtain:

$$x^- \wedge_m ((x \wedge_m y) \odot (x \wedge_m y)) \stackrel{(82)}{=} (x \wedge_m y) \odot (x^- \wedge_m ((x \wedge_m y) \odot (x \wedge_m y))) \stackrel{(49)}{=} 0.$$

$$((x \oplus x)^- \wedge_m y) \wedge_m x = 0. \quad (85)$$

(85): Put $Y := (x \oplus x)^- \wedge_m y$ in (78), to obtain:

$$((x \oplus x)^- \wedge_m y) \wedge_m x \stackrel{(78)}{=} (x \oplus x) \odot (Y \wedge_m x) = (x \oplus x) \odot (((x \oplus x)^- \wedge_m y) \wedge_m x) \stackrel{(51)}{=} 0.$$

$$((x \odot x) \wedge_m y) \wedge_m x^- = 0. \quad (86)$$

$$(86): ((x \odot x) \wedge_m y) \wedge_m x^- = ((x^- \oplus x^-)^- \wedge_m y) \wedge_m x^- \stackrel{(85)}{=} 0.$$

$$(x \odot y)^- \oplus ((y \odot x) \wedge_m x) = 1. \quad (87)$$

$$(87): (x \odot y)^- \oplus ((y \odot x) \wedge_m x) \stackrel{(83)}{=} ((y \odot x) \wedge_m (x \odot y))^- \oplus ((y \odot x) \wedge_m x) \stackrel{(61)}{=} 1.$$

$$(x \oplus ((x \odot x) \wedge_m y))^- \wedge_m x^- = x^-. \quad (88)$$

$$(88): (x \oplus ((x \odot x) \wedge_m y))^- \wedge_m x^- =$$

$$\begin{aligned} &= (x^- \odot ((x \odot x) \wedge_m y)^-) \wedge_m x^- = (((x \odot x) \wedge_m y)^- \odot x^-) \wedge_m x^- \\ &\stackrel{(21)}{=} x^- \odot [((x \odot x) \wedge_m y) \wedge_m x^-]^- \stackrel{(86)}{=} x^- \odot 0^- = x^- \odot 1 = x^-. \end{aligned}$$

$$(x \odot (x \wedge_m y)) \wedge_m x^- = 0. \quad (89)$$

$$(89): (x \odot (x \wedge_m y)) \wedge_m x^- \stackrel{(59)}{=} [(x \odot x) \wedge_m (x \odot (x \wedge_m y))] \wedge_m x^- \stackrel{(86)}{=} 0.$$

$$(x^- \oplus ((x \odot y) \wedge_m y)) \wedge_m y = y. \quad (90)$$

$$(90): (x^- \oplus ((x \odot y) \wedge_m y)) \wedge_m y \stackrel{(31)}{=} y \odot [(y \odot x)^- \oplus ((x \odot y) \wedge_m y)] \stackrel{(87)}{=} y \odot 1 = y.$$

$$((x \odot (x \wedge_m y)) \wedge_m z) \wedge_m x^- = 0. \quad (91)$$

(91): In (67), put $X := x \odot (x \wedge_m y)$, $Y := x^-$ and $Z := z$, to obtain

$$((x \odot (x \wedge_m y)) \wedge_m x^-)^- \odot [((x \odot (x \wedge_m y)) \wedge_m z) \wedge_m x^-] = 0,$$

hence, by (89), $0^- \odot [((x \odot (x \wedge_m y)) \wedge_m z) \wedge_m x^-] = 0$, i.e.

$$1 \odot [((x \odot (x \wedge_m y)) \wedge_m z) \wedge_m x^-] = 0,$$

thus $((x \odot (x \wedge_m y)) \wedge_m z) \wedge_m x^- = 0$.

$$x \odot (((x \odot y) \wedge_m y)^- \odot (y \wedge_m z)) = 0. \quad (92)$$

$$(92): A := x \odot (((x \odot y) \wedge_m y)^- \odot (y \wedge_m z))$$

$$\stackrel{(Pass)}{=} (x \odot ((x \odot y) \wedge_m y)^-) \odot (y \wedge_m z) = (x^- \oplus ((x \odot y) \wedge_m y))^- \odot (y \wedge_m z);$$

since $y \stackrel{(90)}{=} (x^- \oplus ((x \odot y) \wedge_m y)) \wedge_m y$, it follows that

$$A \stackrel{(90)}{=} (x^- \oplus ((x \odot y) \wedge_m y))^- \odot (((x^- \oplus ((x \odot y) \wedge_m y)) \wedge_m y) \wedge_m z) \stackrel{(52)}{=} 0.$$

$$((x \odot y)^- \odot (x \wedge_m z)) \wedge_m (y \odot y) = 0. \quad (93)$$

(93): Note that $y \stackrel{(71)}{=} ((x \odot y) \oplus (x \wedge_m z)^-) \wedge_m y = ((x \odot y)^- \odot (x \wedge_m z))^- \wedge_m y$; we put $X := ((x \odot y)^- \odot (x \wedge_m z))^-$, hence $y = X \wedge_m y$. Then

$$((x \odot y)^- \odot (x \wedge_m z)) \wedge_m (y \odot y) = X^- \wedge_m (y \odot y) = X^- \wedge_m ((X \wedge_m y) \odot (X \wedge_m y)) \stackrel{(84)}{=} 0.$$

$$(x \odot ((x \oplus y) \wedge_m z)) \wedge_m (x \oplus y)^- = 0. \quad (94)$$

(94): In (91), put $X := x \oplus y$, $Y := z$ and $Z := x \odot ((x \oplus y) \wedge_m z)$, to obtain

$$[((x \oplus y) \odot ((x \oplus y) \wedge_m z)) \wedge_m (x \odot ((x \oplus y) \wedge_m z))] \wedge_m (x \oplus y)^- = 0,$$

hence, by (73), to obtain $[x \odot ((x \oplus y) \wedge_m z)] \wedge_m (x \oplus y)^- = 0$.

$$(x \wedge_m y)^- \odot (((z \wedge_m x) \wedge_m y) \wedge_m u) = 0. \quad (95)$$

(95): In (92), put $X := (x \wedge_m y)^-$, $Y := (z \wedge_m x) \wedge_m y$ and $Z := u$, to obtain

$$(x \wedge_m y)^- \odot [(((x \wedge_m y)^- \odot ((z \wedge_m x) \wedge_m y)) \wedge_m ((z \wedge_m x) \wedge_m y))]^- \odot (((z \wedge_m x) \wedge_m y) \wedge_m u) = 0,$$

hence by (68), $(x \wedge_m y)^- \odot [0 \wedge_m ((z \wedge_m x) \wedge_m y)]^- \odot (((z \wedge_m x) \wedge_m y) \wedge_m u) = 0$, hence $(x \wedge_m y)^- \odot [0]^- \odot (((z \wedge_m x) \wedge_m y) \wedge_m u) = 0$, i.e. $(x \wedge_m y)^- \odot (((z \wedge_m x) \wedge_m y) \wedge_m u) = 0$.

$$(x \odot y) \wedge_m (y^- \wedge_m x^-) = 0. \quad (96)$$

(96): In (94), put $X := x$, $Y := y \odot x^-$ and $Z := y$, to obtain

$$(x \odot ((x \oplus (y \odot x^-)) \wedge_m y) \wedge_m (x \oplus (y \odot x^-)))^- = 0,$$

hence, by (39), to obtain $(x \odot y) \wedge_m (x \oplus (y \odot x^-)) = 0$, i.e. $(x \odot y) \wedge_m (y^- \wedge_m x^-) = 0$, by (24).

$$(x \odot (y \wedge_m z)) \wedge_m (y^- \wedge_m x^-) = 0. \quad (97)$$

(97): In (94), put $X := x$, $Y := y \odot x^-$ and $Z := y \wedge_m z$, to obtain

$$(x \odot [(x \oplus (y \odot x^-)) \wedge_m (y \wedge_m z)]) \wedge_m (x \oplus (y \odot x^-))^- = 0,$$

hence, by (53), to obtain $(x \odot [y \wedge_m z]) \wedge_m (x \oplus (y \odot x^-))^- = 0$, hence, by (24), $(x \odot (y \wedge_m z)) \wedge_m (y^- \wedge_m x^-) = 0$.

$$(x \oplus y)^- \wedge_m (y \wedge_m x) = 0. \quad (98)$$

$$(98): (x \oplus y)^- \wedge_m (y \wedge_m x) = (x^- \odot y^-) \wedge_m (y \wedge_m x) \stackrel{(96)}{=} 0.$$

$$((x \oplus y)^- \wedge_m z) \wedge_m (y \wedge_m x) = 0. \quad (99)$$

(99): In (67), put $X := (x \oplus y)^-$, $Y := y \wedge_m x$ and $Z := z$, to obtain

$$[(x \oplus y)^- \wedge_m (y \wedge_m x)]^- \odot (((x \oplus y)^- \wedge_m z) \wedge_m (y \wedge_m x)) = 0,$$

hence, by (98), to obtain $0^- \odot (((x \oplus y)^- \wedge_m z) \wedge_m (y \wedge_m x)) = 0$, i.e.

$$((x \oplus y)^- \wedge_m z) \wedge_m (y \wedge_m x) = 0.$$

$$(x \oplus y)^- \wedge_m ((y \wedge_m z) \wedge_m x) = 0. \quad (100)$$

$$(100): (x \oplus y)^- \wedge_m ((y \wedge_m z) \wedge_m x) = (x^- \odot y^-) \wedge_m ((y \wedge_m z) \wedge_m x) \\ \stackrel{(32)}{=} (x^- \odot ((y \wedge_m z)^- \wedge_m y^-)) \wedge_m ((y \wedge_m z) \wedge_m x) \stackrel{(97)}{=} 0.$$

$$((x \odot y^-) \wedge_m z) \wedge_m (y \wedge_m x^-) = 0. \quad (101)$$

$$(101): ((x \odot y^-) \wedge_m z) \wedge_m (y \wedge_m x^-) = ((x^- \oplus y)^- \wedge_m z) \wedge_m (y \wedge_m x^-) \stackrel{(99)}{=} 0.$$

$$(x \wedge_m ((y \oplus x)^- \wedge_m z)) \wedge_m y = 0. \quad (102)$$

$$(102): (x \wedge_m ((y \oplus x)^- \wedge_m z)) \wedge_m y \stackrel{(60)}{=} (y \oplus x)^- \wedge_m ((x \wedge_m ((y \oplus x)^- \wedge_m z)) \wedge_m y) \stackrel{(100)}{=} 0.$$

$$x^- \wedge_m ((x \odot x) \wedge_m y) = 0. \quad (103)$$

(103): $A := x^- \wedge_m ((x \odot x) \wedge_m y) \stackrel{(88)}{=} ((x \oplus ((x \odot x) \wedge_m y))^- \wedge_m x^-) \wedge_m ((x \odot x) \wedge_m y)$; put $Y := (x \odot x) \wedge_m y$ and by (57), we obtain: $Y = x \wedge_m Y$; hence, we have: $A \stackrel{(57)}{=} ((x \oplus Y)^- \wedge_m x^-) \wedge_m (x \wedge_m Y) \stackrel{(99)}{=} 0$.

$$(x^- \odot (y \wedge_m z)) \wedge_m (x \wedge_m y^-) = 0. \quad (104)$$

$$(104): (x^- \odot (y \wedge_m z)) \wedge_m (x \wedge_m y^-) \stackrel{(59)}{=} [(y \odot x^-) \wedge_m (x^- \odot (y \wedge_m z))] \wedge_m (x \wedge_m y^-) \stackrel{(101)}{=} 0.$$

$$(x \wedge_m ((y \odot x^-) \wedge_m z)) \wedge_m y^- = 0. \quad (105)$$

$$(105): (x \wedge_m ((y \odot x^-) \wedge_m z)) \wedge_m y^- = (x \wedge_m ((y^- \oplus x)^- \wedge_m z)) \wedge_m y^- \stackrel{(102)}{=} 0.$$

$$(x \odot x) \wedge_m (x^- \wedge_m y) = 0. \quad (106)$$

$$(106): (x \odot x) \wedge_m (x^- \wedge_m y) \stackrel{(37)}{=} x^- \wedge_m ((x \odot x) \wedge_m (x^- \wedge_m y)) \stackrel{(103)}{=} 0.$$

$$(x \oplus y)^- \wedge_m (x \wedge_m (y \wedge_m z)) = 0. \quad (107)$$

(107): In (104), put $X := x$, $Y := (y \wedge_m z)^-$ and $Z := y^-$, to obtain

$$(x^- \odot [(y \wedge_m z)^- \wedge_m y^-]) \wedge_m (x \wedge_m (y \wedge_m z)) = 0,$$

hence, by (32), $(x^- \odot y^-) \wedge_m (x \wedge_m (y \wedge_m z)) = 0$, i.e. $(x \oplus y)^- \wedge_m (x \wedge_m (y \wedge_m z)) = 0$.

$$(x \wedge_m (x^- \odot (y \wedge_m z))) \wedge_m y^- = 0. \quad (108)$$

$$(108): (x \wedge_m (x^- \odot (y \wedge_m z))) \wedge_m y^- \stackrel{(59)}{=} (x \wedge_m [(y \odot x^-) \wedge_m (x^- \odot (y \wedge_m z))]) \wedge_m y^- \stackrel{(105)}{=} 0.$$

$$(x \odot x) \wedge_m (y \odot (y \wedge_m x)^-) = 0. \quad (109)$$

$$(109): (x \odot x) \wedge_m (y \odot (y \wedge_m x)^-) \stackrel{(69)}{=} (x \odot x) \wedge_m (x^- \wedge_m (y \odot (y \wedge_m x)^-)) \stackrel{(106)}{=} 0.$$

$$(x \oplus y)^- \wedge_m (y \wedge_m (z \wedge_m x)) = 0. \quad (110)$$

$$(110): (x \oplus y)^- \wedge_m (y \wedge_m (z \wedge_m x)) \stackrel{(43)}{=} (x \oplus y)^- \wedge_m (x \wedge_m (y \wedge_m (z \wedge_m x))) \stackrel{(107)}{=} 0.$$

$$(x \wedge_m (x \oplus y)^-) \wedge_m (y \wedge_m z) = 0. \quad (111)$$

$$(111): (x \wedge_m (x \oplus y)^-) \wedge_m (y \wedge_m z) = (x \wedge_m (x^- \odot y^-)) \wedge_m (y \wedge_m z)$$

$$\stackrel{(32)}{=} (x \wedge_m (x^- \odot ((y \wedge_m z)^- \wedge_m y^-))) \wedge_m (y \wedge_m z) \stackrel{(108)}{=} 0.$$

$$(x \odot x) \wedge_m ((y \odot x)^- \odot (y \wedge_m z)) = 0. \quad (112)$$

(112): In (109), put $X := y \odot y$ and $Y := (x \odot y)^- \odot (x \wedge_m z)$, to obtain

$$\begin{aligned} ((y \odot y) \odot (y \odot y)) \wedge_m (((x \odot y)^- \odot (x \wedge_m z)) \\ \odot [((x \odot y)^- \odot (x \wedge_m z)) \wedge_m (y \odot y)]^-) = 0, \end{aligned}$$

hence, by (93), $((y \odot y) \odot (y \odot y)) \wedge_m (((x \odot y)^- \odot (x \wedge_m z)) \odot 0^-) = 0$, hence $((y \odot y) \odot (y \odot y)) \wedge_m ((x \odot y)^- \odot (x \wedge_m z)) = 0$, hence by (Pass) and (80),

$$(y \odot y) \wedge_m ((x \odot y)^- \odot (x \wedge_m z)) = 0.$$

$$x \wedge_m ((y \oplus x)^- \wedge_m y) = 0. \quad (113)$$

$$(113): x \wedge_m ((y \oplus x)^- \wedge_m y) \stackrel{(37)}{=} (y \oplus x)^- \wedge_m (x \wedge_m ((y \oplus x)^- \wedge_m y)) \stackrel{(110)}{=} 0.$$

$$(x \odot y)^- \odot (x \wedge_m ((y \odot y) \wedge_m z)) = 0. \quad (114)$$

(114): Put $X := y \odot y$, $Y := (x \odot y)^-$, $Z := x$ and $u := z$ in (38), to obtain

$$\begin{aligned} (x \odot y)^- \odot (x \wedge_m ((y \odot y) \wedge_m z)) &\stackrel{(38)}{=} \\ &(y \odot y) \wedge_m [(x \odot y)^- \odot (x \wedge_m ((y \odot y) \wedge_m z))] \stackrel{(112)}{=} 0. \\ x \wedge_m (y \odot ((z \oplus x)^- \wedge_m z)) &= 0. \end{aligned} \quad (115)$$

(115): In (55), put $X := x$ and $Y := (y \oplus x)^- \wedge_m y$, to obtain

$$(x \wedge_m ((y \oplus x)^- \wedge_m y))^- \odot (x \wedge_m (z \odot ((y \oplus x)^- \wedge_m y))) = 0,$$

hence, by (113), $0^- \odot (x \wedge_m (z \odot ((y \oplus x)^- \wedge_m y))) = 0$, hence $x \wedge_m (z \odot ((y \oplus x)^- \wedge_m y)) = 0$, hence $x \wedge_m (y \odot ((z \oplus x)^- \wedge_m z)) = 0$.

$$x \wedge_m (y \wedge_m ((z \wedge_m (x \oplus z)^-) \wedge_m u)) = 0. \quad (116)$$

$$\begin{aligned} (116): x \wedge_m (y \wedge_m ((z \wedge_m (x \oplus z)^-) \wedge_m u)) \\ \stackrel{(44)}{=} (z \wedge_m (x \oplus z)^-) \wedge_m (x \wedge_m (y \wedge_m ((z \wedge_m (x \oplus z)^-) \wedge_m u))) \stackrel{(111)}{=} 0. \end{aligned}$$

$$x \wedge_m ((y \wedge_m z) \odot (z \odot (z \oplus x))^-) = 0. \quad (117)$$

(117): First, $A := (y \wedge_m z) \odot (z \odot (z \oplus x))^-$
 $= (y \wedge_m z) \odot (z^- \oplus (z \oplus x)^-) = (y \wedge_m z) \odot ((z \oplus x)^- \oplus z^-)$;

put $X := y$, $Y := z$, $Z := (z \oplus x)^-$ in (18), to obtain $A \stackrel{(18)}{=} (y \oplus z^-) \odot ((z \oplus x)^- \wedge_m z)$;
 then

$$x \wedge_m ((y \wedge_m z) \odot (z \odot (z \oplus x))^-) = x \wedge_m A = x \wedge_m ((y \oplus z^-) \odot ((z \oplus x)^- \wedge_m z)) \stackrel{(115)}{=} 0.$$

$$x \wedge_m ((y \wedge_m ((x \wedge_m z) \oplus y)^- \wedge_m z)) = 0. \quad (118)$$

(118): Put $X := x$, $Y := z$ and $Z := y \wedge_m ((x \wedge_m z) \oplus y)^-$ in the right side of (66);
 then $x \wedge_m ((y \wedge_m ((x \wedge_m z) \oplus y)^- \wedge_m z)$

$$\stackrel{(66)}{=} (x \wedge_m z) \wedge_m [x \wedge_m ((y \wedge_m ((x \wedge_m z) \oplus y)^-) \wedge_m z)] \stackrel{(116)}{=} 0.$$

$$(x \wedge_m y) \odot (y \odot (y \oplus x))^- = 0. \quad (119)$$

(119): Put $X := x$, $Y := y$ and $Z := (y \odot (y \oplus x))^-$ in the right side of (27); then
 $(x \wedge_m y) \odot (y \odot (y \oplus x))^- = (X \wedge_m Y) \odot Z \stackrel{(27)}{=} x \wedge_m [(x \wedge_m y) \odot (y \odot (y \oplus x))^-] \stackrel{(117)}{=} 0$.

$$x \wedge_m ((y^- \wedge_m ((x \wedge_m z)^- \odot y)) \wedge_m z) = 0. \quad (120)$$

(120): $x \wedge_m ((y^- \wedge_m ((x \wedge_m z)^- \odot y)) \wedge_m z) = x \wedge_m ((y^- \wedge_m ((x \wedge_m z) \oplus y^-)^-) \wedge_m z) \stackrel{(118)}{=} 0$.

$$(x \odot (x \oplus y)) \wedge_m (y \wedge_m x) = y \wedge_m x. \quad (121)$$

(121): Put $X := y \wedge_m x$ and $Y^- := x \odot (x \oplus y)$ in the right side of (19); then

$$\begin{aligned} (x \odot (x \oplus y)) \wedge_m (y \wedge_m x) &= Y^- \wedge_m X \stackrel{(19)}{=} X \odot (X \odot Y)^- \\ &= (y \wedge_m x) \odot [(y \wedge_m x) \odot (x \odot (x \oplus y))^-]^- \stackrel{(119)}{=} (y \wedge_m x) \odot 0^- = y \wedge_m x. \\ (x^- \wedge_m (x \odot (x \wedge_m y)^-)) \wedge_m y &= 0. \end{aligned} \quad (122)$$

(122): $A := (x^- \wedge_m (x \odot (x \wedge_m y)^-)) \wedge_m y \stackrel{(Pcomm)}{=} (x^- \wedge_m ((x \wedge_m y)^- \odot x)) \wedge_m y$;
 put $X := x$, $Y := x^-$, $Z := (x \wedge_m y)^-$ and $U := y$ in the right side of (65); then

$$\begin{aligned} A &= (Y \wedge_m (Z \odot X)) \wedge_m U \stackrel{(65)}{=} X \wedge_m ((Y \wedge_m (Z \odot X)) \wedge_m U) \\ &= x \wedge_m [(x^- \wedge_m ((x \wedge_m y)^- \odot x)) \wedge_m y] \stackrel{(120)}{=} 0. \\ (x \wedge_m y)^- \oplus ((y \oplus z) \odot (y \oplus x)) &= 1. \end{aligned} \quad (123)$$

(123): Since $x \wedge_m y \stackrel{(121)}{=} (y \odot (y \oplus x)) \wedge_m (x \wedge_m y)$, it follows that

$$\begin{aligned} (x \wedge_m y)^- \oplus ((y \oplus z) \odot (y \oplus x)) &\stackrel{(121)}{=} [(y \odot (y \oplus x)) \wedge_m (x \wedge_m y)]^- \oplus ((y \oplus z) \odot (y \oplus x)) \\ &\stackrel{(Scomm)}{=} ((y \oplus z) \odot (y \oplus x)) \oplus [(y \odot (y \oplus x)) \wedge_m (x \wedge_m y)]^- \stackrel{(74)}{=} 1, \end{aligned}$$

for $X := y$, $Y := z$, $Z := y \oplus x$ and $U := x \wedge_m y$.

$$((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-) \wedge_m x = 0. \quad (124)$$

(124): Put $X := x$, $Y := y$ and $Z := x \wedge_m y$ in (72); then

$$\begin{aligned} A &:= (x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^- = Z \odot (Z \wedge_m X)^- \\ &\stackrel{(72)}{=} (X \wedge_m Y)^- \wedge_m (Z \odot (Z \wedge_m X)^-) \\ &= (x \wedge_m y)^- \wedge_m ((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-); \end{aligned}$$

then, $((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-) \wedge_m x = A \wedge_m x$
 $= [(x \wedge_m y)^- \wedge_m ((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-)] \wedge_m x \stackrel{(122)}{=} 0.$

$$((x \oplus y) \odot (x \oplus z)) \wedge_m (z \wedge_m x) = z \wedge_m x. \quad (125)$$

(125): $((x \oplus y) \odot (x \oplus z)) \wedge_m (z \wedge_m x) \stackrel{\text{def.}}{=} (z \wedge_m x) \odot [(z \wedge_m x)^- \oplus ((x \oplus y) \odot (x \oplus z))] \stackrel{(123)}{=} (z \wedge_m x) \odot 1 = z \wedge_m x.$

$$(x \wedge_m (y \wedge_m z)) \odot (x \odot (z \oplus y))^- = 0. \quad (126)$$

(126): Since $A := y \wedge_m z \stackrel{(125)}{=} ((z \oplus y) \odot (z \oplus y)) \wedge_m (y \wedge_m z)$, then

$$(x \wedge_m (y \wedge_m z)) \odot (x \odot (z \oplus y))^- \stackrel{(\text{Pcomm})}{=} (x \odot (z \oplus y))^- \odot (x \wedge_m A) \\ = (x \odot (z \oplus y))^- \odot (x \wedge_m [(z \oplus y) \odot (z \oplus y)] \wedge_m (y \wedge_m z)) \stackrel{(114)}{=} 0,$$

with $X := x$, $Y := z \oplus y$, $Z := y \wedge_m z$.

$$(x \wedge_m y)^- \odot (y \wedge_m (y^- \wedge_m x)) = 0. \quad (127)$$

(127): $(x \wedge_m y)^- \odot (y \wedge_m (y^- \wedge_m x)) \stackrel{(3)}{=} (y \odot (x \oplus y^-))^- \odot (y \wedge_m (y^- \wedge_m x)) \stackrel{(\text{Pcomm}), (126)}{=} 0,$

for $X := y$, $Y := y^-$, $Z := x$.

$$(x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^- = 0. \quad (128)$$

(128): Put $X := (x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-$ and $Y := x$ in (127), to obtain

$$[((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-) \wedge_m x]^- \odot [x \wedge_m (x^- \wedge_m ((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-))] = 0;$$

then, by (124), $[0]^- \odot [x \wedge_m (x^- \wedge_m ((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-))] = 0$, hence $x \wedge_m (x^- \wedge_m ((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-)) = 0$, hence, by (69), $x \wedge_m ((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-) = 0$, hence, by (27), $(x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^- = 0.$

$$((x \wedge_m y) \wedge_m x) \wedge_m (x \wedge_m y) = x \wedge_m y. \quad (129)$$

(129): $((x \wedge_m y) \wedge_m x) \wedge_m (x \wedge_m y) \stackrel{(19)}{=} (x \wedge_m y) \odot ((x \wedge_m y) \odot ((x \wedge_m y) \wedge_m x)^-)^- \stackrel{(128)}{=} (x \wedge_m y) \odot 0^- = x \wedge_m y.$

$$(\Delta_m) \quad (x \wedge_m y) \odot (y \wedge_m x)^- = 0. \tag{130}$$

$$(130): (x \wedge_m y) \odot (y \wedge_m x)^- \stackrel{(P_{\text{comm}})}{=} (y \wedge_m x)^- \odot (x \wedge_m y) \stackrel{(129)}{=} (y \wedge_m x)^- \odot [(x \wedge_m y) \wedge_m x] \wedge_m (x \wedge_m y) \stackrel{(95)}{=} 0.$$

This completes the proof. □

Now, we are able to prove:

Theorem 5.14. *Let $\mathcal{A}^L = (A^L, \odot, ^-, 1)$ be an involutive left- m -BE algebra. If $(m\text{-}BB)$, $(m\text{-}Pabs\text{-}i)$ and $(m\text{-}An)$ hold (i.e. if \mathcal{A}^L is a left-taOWL), then $(\wedge_m\text{-}comm)$ $(x \wedge_m y = y \wedge_m x)$ holds (i.e. \mathcal{A}^L is a left-MV algebra).*

Proof. By Theorem 5.13, the property (Δ_m) holds, hence, for any $x, y \in A^L$, $(x \wedge_m y) \odot (y \wedge_m x)^- = 0$ and $(y \wedge_m x) \odot (x \wedge_m y)^- = 0$, i.e. $x \wedge_m y \leq_m y \wedge_m x$ and $y \wedge_m x \leq_m x \wedge_m y$; since $(m\text{-}An)$ holds, it follows that $x \wedge_m y = y \wedge_m x$, i.e. $(\wedge_m\text{-}comm)$ holds. □

Thus $(\Delta_m) + (m\text{-}An) \iff (\wedge_m\text{-}comm)$ and, consequently, antisymmetric and transitive left-OWLs (= left-taOWLs) are particular left-MV algebras. See Examples 6.10 and 6.11 of taOWLs and proper MV algebra, respectively. Hence, we have: **taOWL** \subset **MV**.

5.3 Putting OLs and their two generalizations on the “involutive Little map”

The definitions (Definitions 2) and the results from this section allow us to draw the hierarchies from the following three Figures, 2, 3 and 4, thus putting all the mentioned algebras on the “involutive Little map”.

Resuming, the connections in $m\text{-}BE_{(DN)}$ algebras and in $m\text{-}aBE_{(DN)}$ algebras are presented in Figures 5 and 6, respectively.

5.4 Future work

This paper is part of a large research project, started in 2019 with the long paper [12], then continued with paper [16] and now with this paper. In the next paper, [17], we prove, among many other things, that any quantum MV algebra verify the property (Δ_m) introduced in this paper. Further, we investigate the orthomodular algebras introduced in [17] in [13]. Then we use the OSLs, the OWLs and Theorem 5.13 from this paper in [14].

Consider the following property:

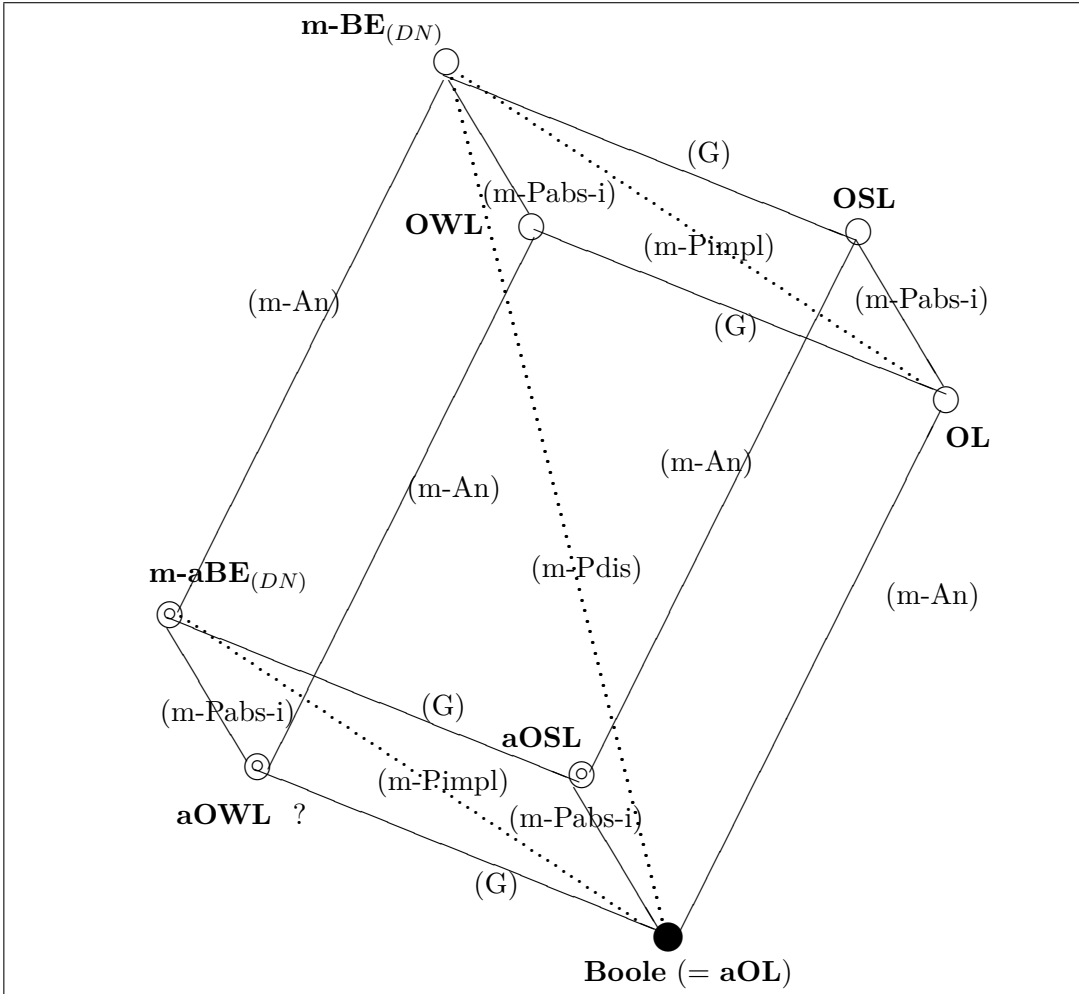


Figure 2: Hierarchies **OSL** vs. **OWL**

(WNM) (Weak Nilpotent Minimum) $(x \odot y)^- \vee [(x \wedge y) \rightarrow (x \odot y)] = 1$, which was introduced in paper [6] in 2001 (see also [7], [20]), where \wedge and \vee are the lattice operations from a bounded residuated lattice $(A, \wedge, \vee, \odot, \rightarrow, 0, 1)$. The property (WNM) was then used in more general cases in [8].

Here, we shall consider, formally, that:

$$x \wedge y \stackrel{\text{def.}}{=} y \wedge_m x = x \odot (x \odot y^-)^- = x \odot (x \rightarrow y) \text{ and}$$

$$x \vee y \stackrel{\text{def.}}{=} y \vee_m x = x \oplus (x \oplus y^-)^- = (y^- \wedge_m x^-)^- = (x^- \wedge y^-)^-,$$

hence $x \wedge y = (x^- \vee y^-)^-$, therefore here the property will be denoted by (WNM_m).

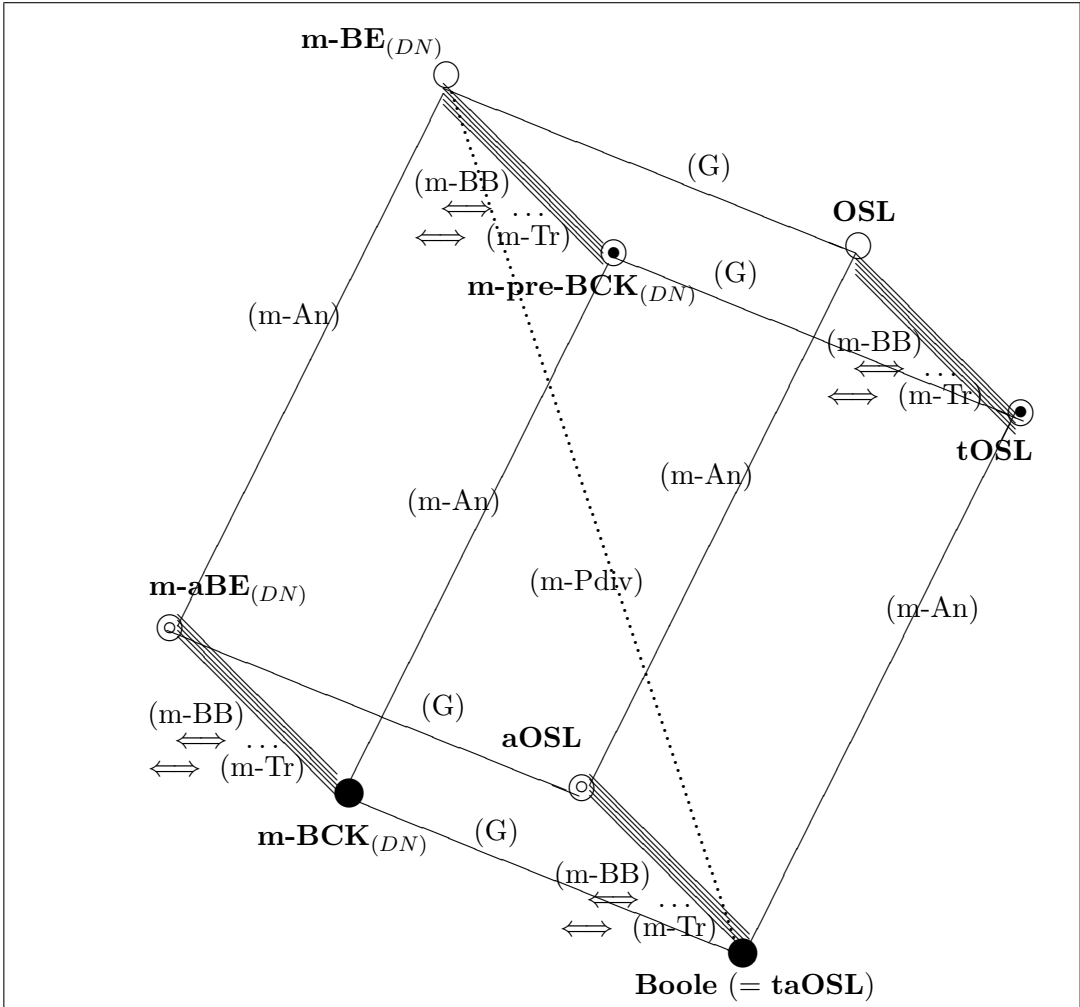


Figure 3: Hierarchies **OSL** vs. **m-pre-BCK_(DN)**

Note that in MV algebras, we have $(WNM_m) = (WNM)$.

Hence, (WNM_m) becomes, equivalently:

$$(x \odot y)^- \vee [(x \odot (x \odot y^-)^-) \odot (x \odot y)^-]^- = 1 \text{ or, equivalently,}$$

$$(x \odot y) \wedge [x \odot (x \odot y^-)^- \odot (x \odot y)^-] = 0 \text{ or, equivalently,}$$

$$(x \odot y) \odot [(x \odot y) \odot [x \odot (x \odot y^-)^- \odot (x \odot y)^-]^-]^- = 0 \text{ or, equivalently,}$$

$$x \odot y \leq_m (x \odot y) \odot [x \odot (x \odot y^-)^- \odot (x \odot y)^-]^-.$$

But, we also have $(x \odot y) \odot [x \odot (x \odot y^-)^- \odot (x \odot y)^-]^- \leq_m x \odot y$.

Hence, if (m-An) holds, then we obtain the property:

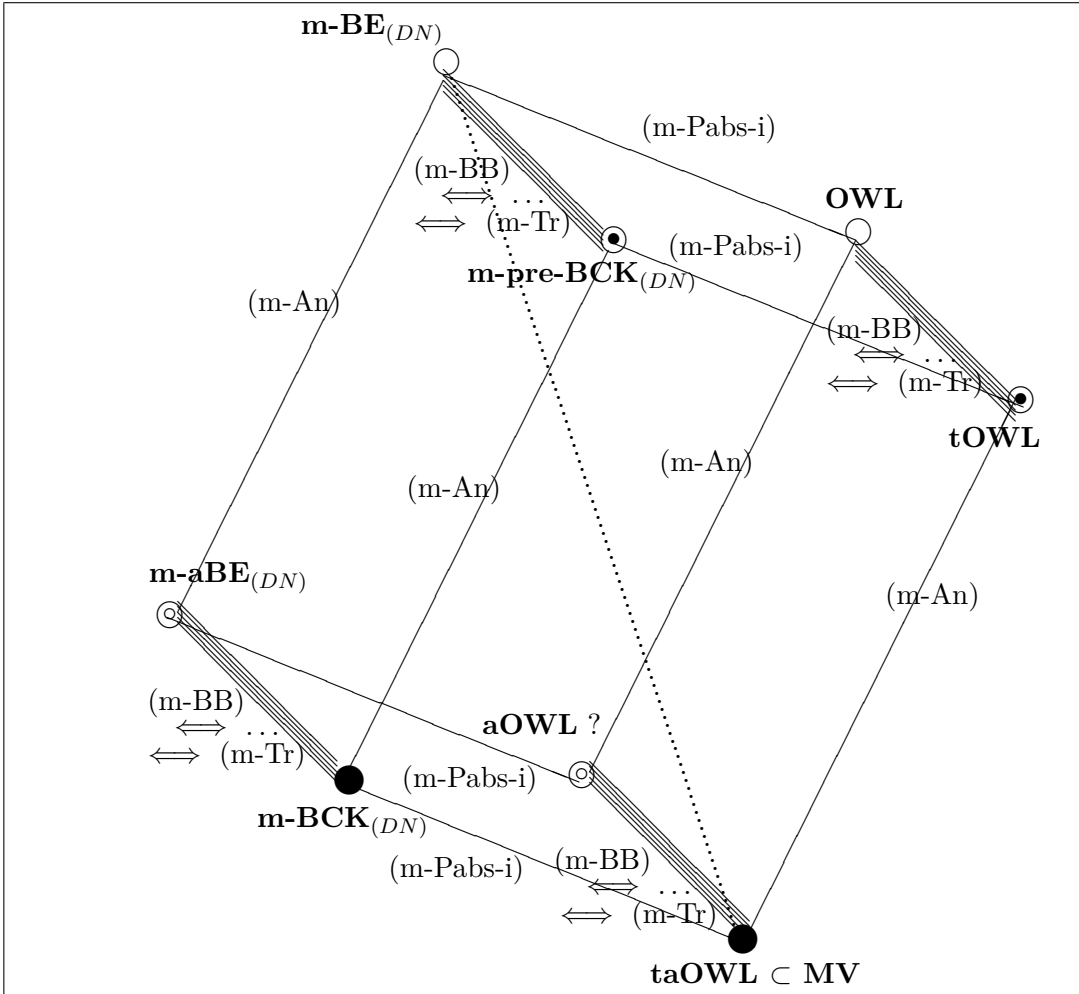


Figure 4: Hierarchies **OWL** vs. **m-pre-BCK**_(DN)

$$(aWNM_m) \quad (x \odot y) \odot [x \odot (x \odot y^-)^- \odot (x \odot y)^-]^- = x \odot y.$$

Thus, we have proved the following result:

Lemma 5.15. *Let $\mathcal{A}^L = (A^L, \odot, ^-, 1)$ be an involutive m -aBE algebra. Then*

$$(WNM_m) \implies (aWNM_m).$$

Lemma 5.16. *Let $\mathcal{A}^L = (A^L, \odot, ^-, 1)$ be an involutive m -BE algebra. Then:*

$$(aWNM_m) \implies (WNM_m).$$

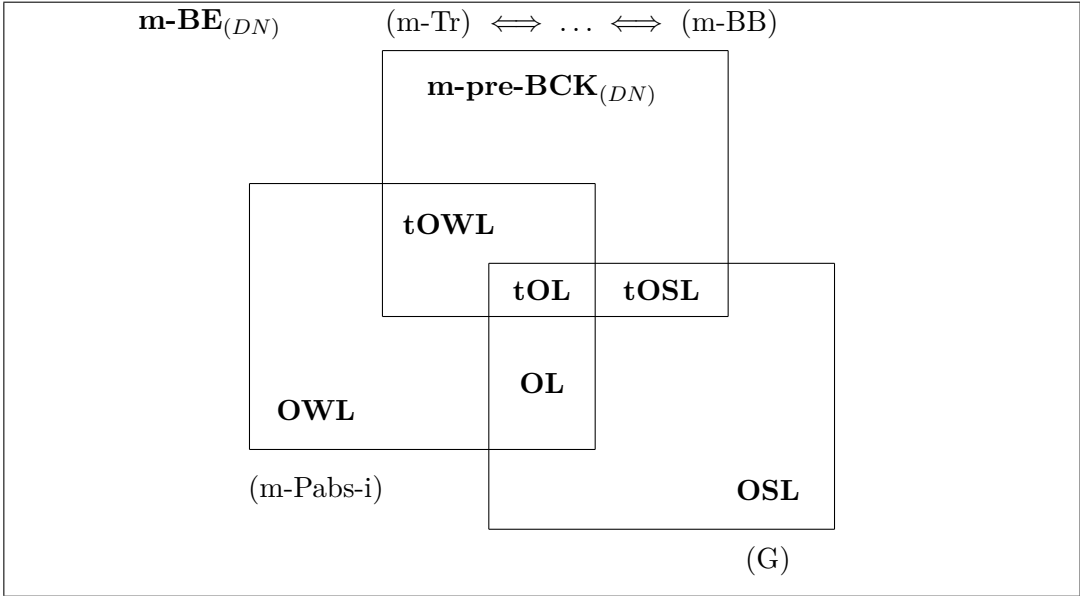


Figure 5: Resuming connections in $\mathbf{m-BE}_{(DN)}$

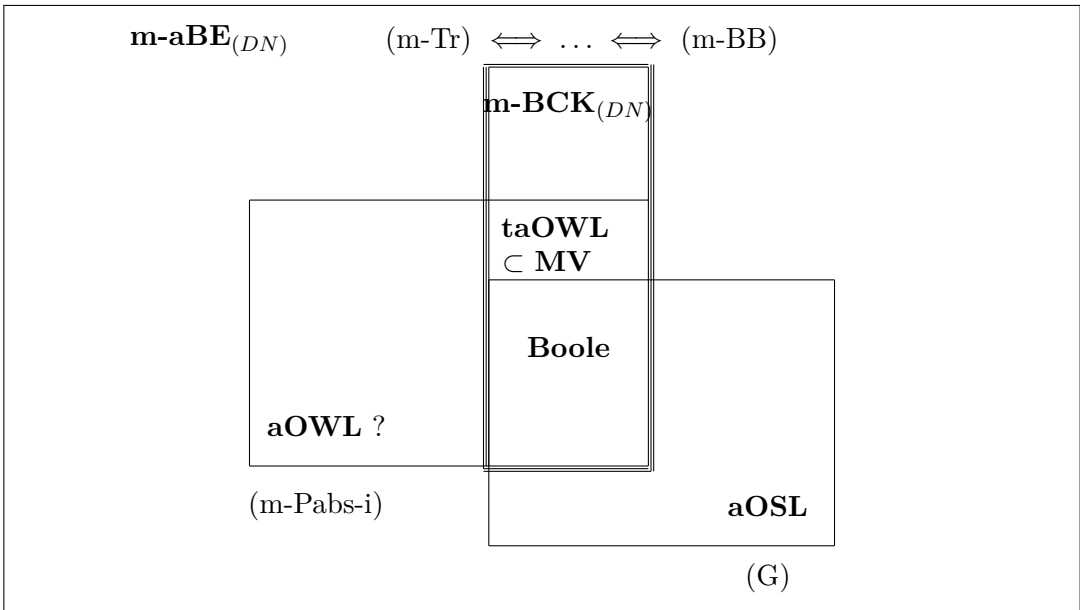


Figure 6: Resuming connections in $\mathbf{m-aBE}_{(DN)}$

Proof. If (aW NM_m) holds, then

$$(x \odot y) \odot [(x \odot y) \odot [x \odot (x \odot y^-)^- \odot (x \odot y)^-]^-]^- = (x \odot y) \odot [x \odot y]^- \stackrel{(m-Re)}{=} 0,$$

hence (W NM_m) holds. □

Proposition 5.17. *Let $\mathcal{A}^L = (A^L, \odot, ^-, 1)$ be an involutive m -aBE algebra. Then:*

$$(aW\mathcal{N}M_m) \iff (W\mathcal{N}M_m).$$

Proof. By Lemmas 5.15 and 5.16. □

Remarks 5.18. Note that:

- the proper OWL from Example 6.8 verifies (aW NM_m) and (hence) (W NM_m),
- the proper tOWL from Examples 6.9(1) verifies (aW NM_m) and (hence) (W NM_m), while the proper tOWL from Examples 6.9(2) verifies (W NM_m), but does not verify (aW NM_m) for (a, e) ;
- both taOWLs (MV algebras) from Examples 6.10 verify (aW NM_m) and (hence) (W NM_m);
- the proper MV algebra from Example 6.11 does not verify (aW NM_m) and (W NM_m) for (b, b) .

We continue this subject (the connections between (m-Pabs-i) and (W NM_m)) in manuscript [15].

6 Examples

Example 6.1. A proper BISL

Using a PASCAL program, we found that the following algebra $\mathcal{A}^L = (A_6 = \{0, a, b, c, d, 1\}, \odot, ^-, 1)$, with the given tables for $\odot, ^-$ and the additional operation \oplus , is an involutive left- m -MEL algebra verifying (G), and not verifying (m-Re) for b ; (m-B), (m-BB), (m-*), (m-**); (m-Tr) for (a, b, d) , (m-An) for (a, b) , (m-Pimpl) for (c, a) , (m-Pabs-i) for (b, a) , (m-Pdis) for (a, b, b) . Hence, \mathcal{A}^L is a proper bounded involutive left-softlattice (Definition 2) (i.e. without (m-An), (m-Tr)).

\odot	0	a	b	c	d	1	x	x^-	\oplus	0	a	b	c	d	1
0	0	0	0	0	0	0	0	1	0	0	a	b	c	d	1
a	0	a	0	0	0	a	a	d	a	a	a	a	1	1	1
b	0	0	b	b	0	b	b	c	b	b	a	b	c	1	1
c	0	0	b	c	d	c	c	b	c	c	1	c	c	1	1
d	0	0	0	d	d	d	d	a	d	d	1	1	1	d	1
1	0	a	b	c	d	1	1	0	1	1	1	1	1	1	1

Example 6.2. A proper aBISL

Using a PASCAL program, we found that the following algebra $\mathcal{A}^L = (A_6 = \{0, a, b, c, d, 1\}, \odot, -, 1)$, with the given tables for $\odot, -$ and the additional operation \oplus , is an involutive left-m-aMEL algebra verifying (G), and not verifying (m-Re) for b ; (m-B), (m-BB), (m-*), (m-**); (m-Tr) for (a, b, d) , (m-Pimpl) for (b, a) , (m-Pabs-i) for (b, a) , (m-Pdis) for (a, b, b) . Hence, it is a proper antisymmetric bounded involutive left-softlattice (Definition 2) (i.e. without (m-Tr)).

\odot	0	a	b	c	d	1		x	x^-		\oplus	0	a	b	c	d	1
0	0	0	0	0	0	0		0	1		0	0	a	b	c	d	1
a	0	a	0	0	0	a	and	a	d	,	a	a	a	a	1	1	1
b	0	0	b	d	d	b		b	c		b	b	a	b	a	1	1
c	0	0	d	c	d	c		c	b		c	c	a	a	c	1	1
d	0	0	d	d	d	d		d	a		d	d	1	1	1	d	1
1	0	a	b	c	d	1		1	0		1	1	1	1	1	1	1

Example 6.3. A proper BIWL

Using the *Mace4* finite model builder, we found that the following algebra $\mathcal{A}^L = (A_4 = \{0, a, b, 1\}, \odot, -, 1)$, with the given tables for $\odot, -$ and the additional operation \oplus , is an involutive left-m-MEL algebra verifying (m-Pabs-i) and not verifying (m-Re) for a , (m-B), (m-BB), (m-*), (m-**); (m-Tr) for (a, b, a) , (m-An) for (a, b) , (m-Pimpl) for $(b, 0)$, (G) for b , (m-Pdis) for (a, b, b) . Hence, it is a proper bounded involutive left-widelattice (Definition 2) (i.e. without (m-An), (m-Tr)).

\odot	0	a	b	1		x	x^-		\oplus	0	a	b	1
0	0	0	0	0		0	1		0	0	a	b	1
a	0	a	0	a	and	a	a	,	a	a	a	1	1
b	0	0	0	b		b	b		b	b	1	1	1
1	0	a	b	1		1	0		1	1	1	1	1

Example 6.4. A proper aBIWL

Using the *Mace4* finite model builder, we found that the following algebra $\mathcal{A}^L = (A_7 = \{0, a, b, c, d, e, 1\}, \odot, -, 1)$, with the given tables for $\odot, -$ and the additional operation \oplus , is an involutive left-m-MEL algebra verifying (m-Pabs-i) and (m-An), and not verifying (m-Re) for a , (m-B) for $(a, 0, a)$, (m-BB) for $(a, a, 0)$, (m-*) for $(0, 0, a)$, (m-**) for (a, e, a) , (m-Tr) for (d, a, e) , (m-Pimpl) for $(a, 0)$, (G) for a , (m-Pdis) for (a, a, a) . Hence, it is a proper antisymmetric bounded involutive left-widelattice (Definition 2) (i.e. without (m-Tr)).

\odot	0	a	b	c	d	e	1	x	x^-	\oplus	0	a	b	c	d	e	1
0	0	0	0	0	0	0	0	0	1	0	0	a	b	c	d	e	1
a	0	c	a	c	0	c	a	a	a	a	a	b	b	a	b	1	1
b	0	a	b	c	d	c	b	b	c	b	b	b	b	b	b	1	1
c	0	c	c	c	0	c	c	c	b	c	c	a	b	c	b	e	1
d	0	0	d	0	d	0	d	d	e	d	d	b	b	b	d	1	1
e	0	c	c	c	0	e	e	e	d	e	e	1	1	e	1	e	1
1	0	a	b	c	d	e	1	1	0	1	1	1	1	1	1	1	1

Example 6.5. A proper OSL

Using a PASCAL program, we found that the following algebra $\mathcal{A}^L = (A_6 = \{0, a, b, c, d, 1\}, \odot, ^-, 1)$, with the given tables for \odot , $^-$ and the additional operation \oplus , is a proper involutive left-m-BE algebra verifying (G), and not verifying (m-B) for (a, b, a) , (m-BB) for (a, a, b) , (m-*) for (a, c, b) , (m-**) for (a, b, a) , (m-Tr) for (a, b, d) , (m-An) for (a, b) , (m-Pimpl) for (d, b) , (m-Pabs-i) for (a, b) . Hence, it is a proper left-orthosoftlattice (Definition 2).

\odot	0	a	b	c	d	1	x	x^-	\oplus	0	a	b	c	d	1
0	0	0	0	0	0	0	0	1	0	0	a	b	c	d	1
a	0	a	0	0	0	a	a	d	a	a	a	b	1	1	1
b	0	0	b	0	0	b	b	c	b	b	b	b	1	1	1
c	0	0	0	c	c	c	c	b	c	c	1	1	c	1	1
d	0	0	0	c	d	d	d	a	d	d	1	1	1	d	1
1	0	a	b	c	d	1	1	0	1	1	1	1	1	1	1

Example 6.6. A proper tOSL

Using a PASCAL program, we found that the following algebra $\mathcal{A}^L = (A_6 = \{0, a, b, c, d, 1\}, \odot, ^-, 1)$, with the following tables for \odot , $^-$ and the additional operation \oplus , is an involutive left-m-BE algebra verifying (G) and (m-Tr), and not verifying (m-An) for (a, c) , (m-Pimpl) for (b, a) , (m-Pabs-i) for (b, d) . Hence, it is a proper transitive left-orthosoftlattice (Definition 2).

\odot	0	a	b	c	d	1	x	x^-	\oplus	0	a	b	c	d	1
0	0	0	0	0	0	0	0	1	0	0	a	b	c	d	1
a	0	a	0	a	0	a	a	d	a	a	a	1	a	1	1
b	0	0	b	0	d	b	b	c	b	b	1	b	1	d	1
c	0	a	0	c	0	c	c	b	c	c	a	1	c	1	1
d	0	0	d	0	d	d	d	a	d	d	1	d	1	d	1
1	0	a	b	c	d	1	1	0	1	1	1	1	1	1	1

Example 6.7. A proper aOSL

Using a PASCAL program, we found that the following algebra $\mathcal{A}^L = (A_6 = \{0, a, b, c, d, 1\}, \odot, ^-, 1)$, with the given tables for \odot , $^-$ and the additional operation

\oplus , is a proper involutive left-m-aBE algebra verifying (G), and not verifying (m-B) for (a, b, a) , (m-BB) for (a, a, b) , (m-*) for (b, d, a) , (m-**) for (a, b, a) , (m-Tr) for (a, b, d) , (m-Pimpl) for (d, b) , (m-Pabs-i) for (a, b) . Hence, it is a proper antisymmetric left-orthosoftlattice (Definition 2).

\odot	0	a	b	c	d	1		x	x^-		\oplus	0	a	b	c	d	1
0	0	0	0	0	0	0		0	1		0	0	a	b	c	d	1
a	0	a	0	0	0	a		a	d		a	a	a	b	c	1	1
b	0	0	b	0	b	b	and	b	c	, with	b	b	b	b	1	1	1
c	0	0	0	c	c	c		c	b		c	c	c	1	c	1	1
d	0	0	b	c	d	d		d	a		d	d	1	1	1	d	1
1	0	a	b	c	d	1		1	0		1	1	1	1	1	1	1

Example 6.8. A proper OWL

Using the *Mace4* finite model builder, we found that the following algebra $\mathcal{A}^L = (A_5 = \{0, a, b, c, 1\}, \odot, -, 1)$, with the given tables for \odot , $-$, and the additional operation \oplus , is an involutive left-m-BE algebra verifying (m-Pabs-i), while (m-B) does not hold for (b, a, c) , (m-BB) for (b, c, b) , (m-*) for (a, c, b) , (m-**) for (b, a, c) , (m-Tr) for (b, a, c) , (m-An) for (a, b) , (m-Pimpl) for $(a, 0)$, (G) for a . Hence, it is a proper left-orthowidelattice (Definition 2).

\odot	0	a	b	c	1		x	x^-		\oplus	0	a	b	c	1
0	0	0	0	0	0		0	1		0	0	a	b	c	1
a	0	0	0	0	a	and	a	b	, with	a	a	1	1	b	1
b	0	0	0	a	b		b	a		b	b	1	1	1	1
c	0	0	a	0	c		c	c		c	c	b	1	1	1
1	0	a	b	c	1		1	0		1	1	1	1	1	1

Examples 6.9. Proper tOWLs

1. Using the *Mace4* finite model builder, we found the following algebra $\mathcal{A}^L = (A_8 = \{0, a, b, c, d, e, f, 1\}, \odot, -, 1)$ with the given tables for \odot , $-$ and the additional operation \oplus , is an involutive left-m-BE algebra verifying (m-Pabs-i) and (m-BB) $\iff \dots$ (m-Tr), and not verifying (m-An) for (a, d) , (m-Pimpl) for $(c, 0)$, (G) for c . Hence, it is a proper left-tOWL (Definition 2).

\odot	0	a	b	c	d	e	f	1		\oplus	0	a	b	c	d	e	f	1
0	0	0	0	0	0	0	0	0		0	0	a	b	c	d	e	f	1
a	0	a	0	0	a	a	0	a		a	a	a	1	1	d	e	1	1
b	0	0	b	c	0	0	f	b	and	b	b	1	b	b	1	1	b	1
c	0	0	c	f	0	0	f	c		c	c	1	b	b	1	1	c	1
d	0	a	0	0	a	d	0	d		d	d	d	1	1	e	e	1	1
e	0	a	0	0	d	e	0	e		e	e	e	1	1	e	e	1	1
f	0	0	f	f	0	0	f	f		f	f	1	b	c	1	1	f	1
1	0	a	b	c	d	e	f	1		1	1	1	1	1	1	1	1	1

and $(0, a, b, c, d, e, f, 1)^- = (1, b, a, d, c, f, e, 0)$.

2. Using the *Mace4* finite model builder, we found the following algebra $\mathcal{A}^L = (A_{14} = \{0, a, b, c, d, e, f, g, h, i, j, k, m, 1\}, \odot, ^-, 1)$, with the given tables for \odot and $^-$, is an involutive left-m-BE algebra verifying (m-Pabs-i) and (m-BB) $\iff \dots$ (m-Tr), and not verifying (m-An) for (a, e) , (m-Pimpl) for $(a, 0)$, (G) for a . Hence, it is a proper left-tOWL (Definition 2).

\odot	0	a	b	c	d	e	f	g	h	i	j	k	m	1		x	x^-
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		0	1
a	0	c	0	c	b	g	0	c	k	g	f	0	c	a		a	b
b	0	0	0	0	b	0	0	0	k	0	f	0	0	b		b	a
c	0	c	0	c	0	c	0	c	0	c	0	0	c	c		c	d
d	0	b	b	0	d	b	f	0	h	0	j	k	b	d		d	c
e	0	g	0	c	b	i	0	g	k	i	f	0	c	e		e	f
f	0	0	0	0	f	0	0	0	f	0	f	0	0	f	and	f	e
g	0	c	0	c	0	g	0	c	0	g	0	0	c	g		g	h
h	0	k	k	0	h	k	f	0	j	0	j	f	k	h		h	g
i	0	g	0	c	0	i	0	g	0	i	0	0	c	i		i	j
j	0	f	f	0	j	f	f	0	j	0	j	f	f	j		j	i
k	0	0	0	0	k	0	0	0	f	0	f	0	0	k		k	m
m	0	c	0	c	b	c	0	c	k	c	f	0	c	m		m	k
1	0	a	b	c	d	e	f	g	h	i	j	k	m	1		1	0

Examples 6.10. taOWls (= MV algebras + (m-Pabs-i))

1. The following algebra $\mathcal{A}^L = (L_3 = \{0, a, 1\}, \odot, ^-, 1)$, with the given tables for \odot , $^-$ and for the additional operation \oplus , is the linearly ordered left-MV algebra with the smallest number of elements verifying (m-Pabs-i), and not verifying (m-Pimpl) for $(a, 0)$, (G) for a , (m-Pdis) for (a, a, a) .

\odot	0	a	1	and	x	x^-	,	with	\oplus	0	a	1
0	0	0	0		0	1			0	0	a	1
a	0	0	a		a	a			a	a	1	1
1	0	a	1		1	0			1	1	1	1

2. The following algebra $\mathcal{A}^L = (A_6 = \{0, a, b, c, d, 1\}, \odot, -, 1)$, with the given tables for $\odot, -$ and the additional operation \oplus , is a left-MV algebra verifying (m-Pabs-i), and not verifying (m-Pimpl) for $(b, 0)$, (G) for b , (m-Pdis) for (b, b, b) . It is isomorphic with $\mathcal{L}_{3 \times 2}$ from ([8], page 165).

\odot	0	a	b	c	d	1	and	x	x^-	,	with	\oplus	0	a	b	c	d	1
0	0	0	0	0	0	0		0	1			0	0	a	b	c	d	1
a	0	a	0	a	0	a		a	d			a	a	a	c	c	1	1
b	0	0	0	0	b	b		b	c			b	b	c	d	1	d	1
c	0	a	0	a	b	c		c	b			c	c	c	1	1	1	1
d	0	0	b	b	d	d		d	a			d	d	1	d	1	d	1
1	0	a	b	c	d	1		1	0			1	1	1	1	1	1	1

Example 6.11. A proper MV algebra

The following algebra $\mathcal{A}^L = (L_4 = \{0, a, b, 1\}, \odot, -, 1)$, with the given tables for $\odot, -$ and the additional operation \oplus , is a proper left-MV algebra, not verifying (m-Pabs-i) for $(a, 0)$, (m-Pimpl) for $(a, 0)$, (G) for b , (m-Pdis) for (a, a, b) , (aWNM_m) for (b, b) . \mathcal{A}^L is the chain $0 \leq_m a \leq_m b \leq_m 1$, i.e. \mathcal{A}^L is \mathcal{L}_4 from [8].

\odot	0	a	b	1	and	x	x^-	,	with	\oplus	0	a	b	1
0	0	0	0	0		0	1			0	0	a	b	1
a	0	0	0	a		a	b			a	a	b	1	1
b	0	0	a	b		b	a			b	b	1	1	1
1	0	a	b	1		1	0			1	1	1	1	1

References

- [1] Garrett Birkhoff, *Lattice Theory*. (3rd edition), American Mathematical Society Colloquium Publications, Vol. XXV, American Mathematical Society, Providence RI 1967.
- [2] Roberto L. O. Cignoli, Itala M. L. D’Ottaviano, and Daniele Mundici, *Algebraic Foundations of Many-valued Reasoning*. Trends in Logic - Studia Logica Library **7**, Kluwer Academic Publishers & Springer Science, Dordrecht 2000.
- [3] Chen Chung Chang, Algebraic analysis of many valued logics. *Trans. Amer. Math. Soc.* **88** (1958), 467–490.
- [4] Maria Luisa Dalla Chiara, Roberto Giuntini, and Richard Greechie, *Reasoning in Quantum Theory. Sharp and Unsharp Quantum Logics*. Trends in Logic – Studia Logica Library **22**, Kluwer Academic Publishers, Dordrecht, 2004

- [5] Richard Dedekind, Über Zerlegung von Zahlen durch ihre grössten gemeinsamen Teiler, *Festschrift Tech. Hochschule Braunschweig*, 1-40. = *Gesammelte Math. Werke*, vol. II, 1897, pp. 103–147.
- [6] Francesc Esteva and Lluís Godó, Monoidal t-norm based logic: towards a logic for left-continuous t-norms, *Fuzzy Sets and Systems* **124** (2001), no. 3, 271–288.
- [7] Francesc Esteva, Lluís Godó, and Carles Noguera, On rational weak nilpotent minimum logics, *J. Mult.-Valued Logic Soft Comput.* **12** (2006), 9–32.
- [8] Afrodita Iorgulescu, *Algebras of logic as BCK algebras*, Bucharest University of Economic Studies Press, Bucharest 2008.
- [9] — New generalizations of BCI, BCK and Hilbert algebras, Parts I, II (Dedicated to Dragoş Vaida), *J. Mult.-Valued Logic Soft Comput.* **27** (2016), no. 4, 353–406 and 407–456. (A previous version available from December 6, 2013 at <http://arxiv.org/abs/1312.2494>.)
- [10] — Implicative-groups vs. groups and generalizations, *Matrix Rom, Bucharest* 2018.
- [11] — Generalizations of MV algebras, ortholattices and Boolean algebras, *ManyVal 2019, Bucharest, Romania*, November 1-3, 2019.
- [12] — Algebras of logic vs. Algebras, in *Landscapes in Logic*, Vol. 1: Contemporary Logic and Computing, Ed. Adrian Reuş, College Publications, 2020, 157–258.
- [13] — On Quantum-MV algebras - Part I: Orthomodular algebras, submitted.
- [14] — On Quantum-MV algebras - Part II: Orthomodular lattices, softlattices and wide-lattices, submitted.
- [15] — On Quantum-MV algebras - Part III: The properties (m-Pabs-i) and (WNM_m), manuscript.
- [16] Afrodita Iorgulescu and Michael Kinyon, Putting bounded involutive lattices, De Morgan algebras, ortholattices and Boolean algebras on the “map”, *Journal of Applied Logics (JALs-ifCoLog), Special Issue on Multiple-Valued Logic and Applications*, Vol. 8, No. 5, June 2021, 1169 - 1213.
- [17] — Putting Quantum MV algebras on the “map”, submitted.
- [18] Gudrun Kalmbach, *Orthomodular Lattices*, London Mathematical Society Monographs **18**, Academic Press, London, New York, etc. 1983.
- [19] William W. McCune, *Prover9 and Mace4*, available at <http://www.cs.unm.edu/~mccune/Prover9>.
- [20] Carles Noguera, Francesc Esteva, J. Gispert, On triangular norm based axiomatic extensions of the Weak Nilpotent Minimum Logic, *MLQ Math. Log. Q.* **54** (2008), 387–409.
- [21] Óystein Ore, On the foundations of abstract algebra, I. *Ann. Math.* **36** (1935), 406–437.
- [22] Ranganathan Padmanabhan and Sergiu Rudeanu, *Axioms for Lattices and Boolean Algebras*, World Scientific, Singapore, etc. 2008.

A SCIENCE LIKE ANY OTHER

ANDREW DAVID IRVINE
University of British Columbia, Canada
andrew.irvine@ubc.ca

Abstract

Scientific realism holds that the world is composed of natural, physical, observer-independent particulars, properties and relations, and that science is a fallible but reliable guide for discovering their existence and nature. This article defends the claim that formal logic is a part of science, that it has as its subject structures and attributes present in the natural, physical world and that logical truths, like other scientific truths, are discovered *a posteriori*. In short, it defends the view that logic is a science like any other.

1 Two Theorems

Here is an interesting theorem in modal logic:

$$\vdash \diamond \sim \Box (p \vee \sim p)$$

The theorem states that it is possible that the law of excluded middle is not necessary.

Here is a second theorem:

$$\vdash \diamond \sim \Box \sim (p \wedge \sim p)$$

This theorem states that it is possible that the law of non-contradiction is not necessary.

Neither theorem is a theorem of the modal systems S4 or S5, but both are theorems of the modal system S7. Both appear to assert something relevant to the debate about whether logical truths are necessary truths.

For helpful comments on early drafts of this paper, I would like to thank Jim Franklin, Jim Robinson, Victor Rodych, John Woods and Barrie McCullough, as well as audiences at the University of Oxford, the University of Athens, the National Autonomous University of Mexico and the annual Dubrovnik Philosophy of Science Conference.

2 A Challenge

In his *Philosophy of Logic*, Quine considers a conversation between a classical logician and a non-classical (or what Quine calls a *deviant*) logician. The classical logician accepts the law of non-contradiction. The non-classical logician rejects it, believing instead that some sentences and their negations are both true. The non-classical logician, Quine notes, does more than just advocate a non-standard method for generating the familiar class of logical truths. Instead, the non-classical logician advocates “a change of that class itself.”¹ As Quine also notes, “It is not just a change of demarcation, either, between what to call logical truth and what to call extra-logical truth. It is a question rather of outright rejection of part of our logic as not true at all.”²

In effect, the non-classical logician is offering the classical logician a challenge. The challenge is this: Is the classical logician able to give an argument for preferring classical logic over an alternative, non-classical logic? Is the classical logician able to give evidence as to why classical logic is true and why a competing, non-classical logic is not? When faced with such a challenge, how should the classical logician reply?³

The first response that comes to mind is what we might call the logician’s reply. The reply goes something like this: Yes, of course we have good reason for preferring classical logic over alternative, non-classical logics. The reason is that for any classical logical truth, there is a proof of that truth.⁴ For someone who might be skeptical of the law of non-contradiction, for example, we can construct the following derivation:

(Pf 1)	1.	$(p \wedge \sim p)$	Assump IP
	2.	p	1 Simp
	3.	$p \vee (q \wedge \sim q)$	2 Add
	4.	$\sim p$	1 Simp
	5.	$(q \wedge \sim q)$	3, 4 DS
	6.	$\sim(p \wedge \sim p)$	1 – 5 IP

¹Quine (1986), 80.

²Quine (1986), 80-1.

³Woods (2003) helpfully gives this type of challenge, the type of challenge that requires “adjudicating in a principled way the conflict between supposing that [a given argument] is a sound demonstration of a counterintuitive truth, as opposed to seeing it as a counterexample of its premises,” the name “Philosophy’s Most Difficult Problem” (14). Additional discussion can be found, not only in Woods (2003) but in Haack (1974), Hjortland (2017), Quine (1951), (1960) and (1986), and Priest (2008) and (2014).

⁴For current purposes, we need not distinguish between the existence of a proof and the ability to generate such a proof mechanically. As is well known, classical propositional logic is complete, sound and decidable, while classical predicate logic is complete, sound and only semi-decidable.

It then turns out that rejecting the law of non-contradiction will be, in Quine's words, "absurd on the face of it."⁵ A proof like this decides the question. As Quine puts it, "If sheer logic is not conclusive, what is? What higher tribunal could abrogate the logic of truth functions or of quantification?"⁶ Of course, we might also offer a shorter proof. For example, for logics that include the rule of repetition, we might offer the following derivation:

$$\begin{array}{ll}
 \text{(Pf 2)} & 1. \mid (p \wedge \sim p) \quad \text{Assump IP} \\
 & 2. \mid (p \wedge \sim p) \quad 1 \text{ Rep} \\
 & 3. \sim(p \wedge \sim p) \quad 1 - 2 \text{ IP}
 \end{array}$$

For logics that allow a more direct version of indirect proof, even (Pf 2) can be shortened:

$$\begin{array}{ll}
 \text{(Pf 3)} & 1. \mid (p \wedge \sim p) \quad \text{Assump IP} \\
 & 2. \sim(p \wedge \sim p) \quad 1 - 1 \text{ IP}
 \end{array}$$

For logics that have the law of non-contradiction as an axiom, the following proof is even shorter:

$$\text{(Pf 4)} \quad 1. \sim(p \wedge \sim p) \quad \text{Axiom}$$

The problem with all such proofs is that, although they are valid, they all beg the question against the non-classical logician. In one way or another, each assumes the very point at issue, namely that contradictions need to be outlawed. This assumption is built into classical logic, either as an axiom (as in Pf 4) or in the rules, for example in the rule of indirect proof (as in Pfs 1, 2 and 3). As a result, the classical logician's reply will be unconvincing to the non-classical logician. Stated more carefully, the non-classical logician's challenge to the classical logician is really this: Is the classical logician able to give a *non-question-begging* argument for preferring classical logic over an alternative, non-classical logic? Is the classical logician able to give *non-question-begging* evidence as to why classical logic is true and why a competing, non-classical logic is not?

To this revised challenge we can offer a second reply, something we might call the *philosopher's reply*. This reply doubles down against the non-classical logician. Instead of merely *assuming* the law of non-contradiction, this reply argues that denying the law of non-contradiction involves a significant error. As Quine puts it,

My view of this dialogue is that neither party knows what he is talking about. They think they are talking about negation, \sim , 'not'; but surely the notation ceased to be recognizable as negation when they took to

⁵Quine (1986), 81.

⁶Quine (1986), 81.

regarding some conjunctions of the form ' $p \cdot \sim p$ ' as true, and stopped regarding such sentences as implying all others. Here, evidently, is the deviant logician's predicament: when he tries to deny the doctrine he only changes the subject.⁷

Once again, though, the non-classical logician will not be satisfied. The non-classical logician will want to insist that he and the classical logician in fact are still arguing over the same sentence. He will argue that believing otherwise results from a confusion on the part of Quine, not on the part of the non-classical logician. A stalemate seems to have been reached.

Or has it? Is there no argument that will carry the discussion forward? Given our revised challenge — the challenge of finding a non-question begging way of adjudicating between two competing interpretations of the same logical sentence — can nothing more be said? Is there no evidence that will help resolve our revised challenge to the satisfaction of both parties?

At this point we will want to consider a third reply, a response we can call the *scientist's reply*. This is a reply that accepts Quine's claim that we now have two different theories discussing two different phenomena. It is a reply that agrees with Quine that the non-classical logician in fact has changed the subject. Even so, the reply proposes that it is still possible to ask which of the two theories more accurately mirrors or describes the world in which we live. It is still possible to ask which of the two theories, if either, correctly integrates logic into our broader theories of the world, broader theories that will include both our physics and our metaphysics, among others.

Put another way, Quine's objection is one that can be made, not only to non-classical theories of logic but to *any* competing theory in *any* discipline. Perhaps because of this very generality, the objection is rarely if ever conclusive. In comparisons between classical mechanics and today's standard model, for example, none of Newton's four key concepts — space, time, matter and force — retain their original meaning. All have undergone significant revision. Space and time have become a single bundle of relations now referred to as space-time. Matter, which Newton defined as having both mass and volume, is no longer a fundamental concept, since the elementary constituents of atoms are no longer understood to have size or volume in any ordinary sense of the word. Force is now described by the Schrödinger equation, rather than by Newton's equations, and it is "quantized" in a way that would have been foreign to Newton. And in the special theory of relativity, mass and energy turn out to be inter-definable. Clearly, by denying classical mechanics, modern physicists seem to have "changed the subject" — but this hardly served

⁷Quine (1986), 81.

to end debate. In chemistry, competing models of the atom, including Dalton's solid-sphere model, Thomson's plum-pudding model, Rutherford's nuclear model, Bohr's planetary model and Schrödinger's three-dimensional model all, in Quine's sense, referred to different subjects. Even so, science progressed. In all such cases, progress was made because there was a real-world subject that could be consulted, a subject that served as a theory-neutral arbitrator for deciding between competing scientific theories.

In many of the physical sciences, finding this theory-neutral arbitrator is relatively straightforward. Perhaps you think the moon is made of rock and dirt. I think it is made of a semi-luminous aether whose vibrations result in various forms of electromagnetic radiation, including light. Perhaps we even have an old-fashioned friend who thinks we are both wrong and that the moon is made of green cheese. Or perhaps I think that quantum theory correctly describes an inherent randomness that exists in the world. You disagree and suggest that the postulated randomness is merely superficial, that it arises from some hidden, yet-to-be-discovered variable.

To decide such issues, we do more than just consult the internal structure of our current theories. We do more than simply conclude that, since our theories have different propositional content, they must be incommensurable. In our first example, we go to the moon and check. In our second, we consult with John Bell and he shows us that there are a series of inequalities that need to be obeyed by every deterministic theory but that will be violated by quantum theory, should it in fact describe an inherently random world. We then take the time and effort to construct the appropriate experiments.

When we do, it turns out that our observations show you to be right in the case of the moon (where not a trace of an aether and not a speck of coagulated dairy products has been found) and me to be right in the case of quantum mechanics (in which Bell's inequalities repeatedly have been observed to be violated). Of course, any such observation *might* turn out to be misleading. None of our intellectual activities, even those in geology and physics, are infallible; but this is no objection to the principle that it is the subject that determines the theory, not the theory that determines the subject.

This third reply presents us with a model for understanding theory choice in logic. In this model, logic will have a subject about which any given logical theory (and any particular sentence within that theory) will be right or wrong, correct or incorrect. This subject may turn out to be rather different from that of other sciences. For example, it may be more formal than that of geology and more abstract than that of physics. But this only shows that although certain aspects of logic may be different in degree from other sciences, they need not be different in kind. It does not show that logical theories are justifiable *a priori* and it does not mean that

logical theories are true only in the (very weak, Quinean) sense that any competing logical theory will turn out to be theoretically incommensurable. Instead, logic will be a science like any other.

3 Necessity

A system of modal logic is normal whenever it contains both $\Box p \supset p$ and $\Box(p \supset q) \supset (\Box p \supset \Box q)$ as either theorems or axioms, and both *modus ponens* and necessitation, or their equivalents, as rules of inference.⁸ Necessitation (Nec) is the rule that

$$\text{If } \vdash p, \text{ then } \vdash \Box p$$

Because it is a rule whose antecedent refers only to theorems, necessitation is distinct from the invalid formula $p \supset \Box p$, a formula that will be incompatible with all but the most deterministic of worlds.

Necessitation leads directly to theorems containing iterated modalities. It generates theorems asserting that the truths of logic are not just necessary but necessarily necessary, for example that $\Box\Box(p \vee \sim p)$ and $\Box\Box\sim(p \wedge \sim p)$. A question then naturally arises about the equivalence (or non-equivalence) of various sentences. Should $\Diamond\Diamond p$ be understood to be equivalent to $\Diamond p$? Should $\Box\Box p$ be thought to be equivalent to $\Box p$? Should $\Box\Diamond p$ be understood to be equivalent to $\Diamond p$? Should $\Diamond\Box p$ be thought to be equivalent to $\Box p$? S5, perhaps the most popular account of alethic modality proposed to date, answers in the affirmative. Because of S5's universal accessibility relation, in which every possible world is accessible from every possible world, iterated modalities collapse. S5 thus fits neatly with a positivist view of logic.

Logical positivism (or logical empiricism) differed from earlier forms of empiricism because of its commitment to the idea that meaningful statements must be either empirically verifiable or tautological, and to the idea that knowledge is based on publicly verifiable experiment, not just individual experience. Contingent sentences, sentences that are true or false depending on how the world in fact is constituted, are then justified *a posteriori*. Statements in logic and mathematics, however, posed a problem: If not empirically verifiable, what justification can they have?

The positivist answer was that such statements did not need to be justified *a posteriori* since they turned out to be true (or false) regardless of how the world might be organized. Mere tautologies tell us nothing about the world. The sentence $p \vee \sim p$ tells us neither that p is true nor that p is false. It tells us nothing about p ,

⁸The definition is due to Kripke, although the terminology used here is from Hughes and Cresswell (1968), 31, 237. Cf. Kripke (1963), 67. Some authors prefer to use a definition involving the formulation, If $\models \alpha$, then $\models \Box\alpha$.

or about q , for that matter. Put in other words, necessary sentences turn out to be empty of content. They stand *unconnected* to the world. They need not be justified *a posteriori* because they cannot be falsified *a posteriori*.

Stated another way, contingently true sentences exclude possibilities. “Socrates is older than Plato” excludes the possibility that Plato is older than Socrates. In contrast, being empty of content, necessarily true sentences seem to exclude nothing. It is for this reason that $\Box\Box(p \vee \sim p)$ is thought to be equivalent to $\Box(p \vee \sim p)$. The sentence $p \vee \sim p$ excludes nothing. Thus, it must be true in all possible worlds. Thus, it follows that $\Box(p \vee \sim p)$ must itself be necessary. It then follows that both $\Box(p \vee \sim p) \supset \Box\Box(p \vee \sim p)$ and $\Box\Box(p \vee \sim p) \supset \Box(p \vee \sim p)$. So it follows that $\Box\Box(p \vee \sim p) \equiv \Box(p \vee \sim p)$. The logic of alethic modality thus appears to require a universal accessibility relation. Every possible world must be accessible from every possible world. If $p \vee \sim p$ is necessary at one possible world, it will be necessary at all possible worlds. All possible worlds not only have the same logic, they *must* have the same logic. Thus, not only does the sentence $\Box\Box p$ imply $\Box p$, the sentence $\Box p$ implies $\Box\Box p$. Hence, $\Box p \equiv \Box\Box p$. Not only does the sentence $\Diamond\Diamond p$ imply $\Diamond p$, the sentence $\Diamond p$ implies $\Diamond\Diamond p$. Hence, $\Diamond p \equiv \Diamond\Diamond p$.

Unfortunately, because it includes these equivalences, S5 turns out to be incompatible with the scientist’s reply to the challenge posed by the non-classical logician. If the classical logician accepts, and the nonclassical logician rejects, the sentence $\Box(p \vee \sim p)$, and if we want to turn to the natural world to adjudicate between the two interpretations offered for this sentence, it will need to be possible both for $\Box(p \vee \sim p)$ to be true and for $\Box(p \vee \sim p)$ to be false. Unless $\Box(p \vee \sim p)$ is both possibly true and possibly false ($\Diamond\Box(p \vee \sim p) \wedge \Diamond\sim\Box(p \vee \sim p)$), we will be forced to return to the philosopher’s reply. We will be forced to return to the idea that, not only is the law of excluded middle a logical truth, it is a necessary logical truth at every possible world. The belief that $\sim(p \vee \sim p)$ will then have to be a belief about something other than what is found in the world. Once again, the non-classical logician merely will have changed the subject and the classical logician will be unable to give a non-question begging answer to the non-classical logician’s challenge.

In contrast, accepting the scientist’s reply means there must be at least one possible world in which $\Diamond\sim\Box(p \vee \sim p)$. The law of excluded middle (and other logical laws) will still be necessary. But they will not be necessarily necessary. Having accepted an equivalence between $\Box(p \vee \sim p)$ and $\Box\Box(p \vee \sim p)$, systems such as S5 disallow this possibility. We therefore turn to the non-normal system, S7.

4 S7

S7 shares several features with the modal system S5.⁹ Both systems are extensions of S3, which like all alethic modal systems supplements classical logic with the operators \Box (for necessity) and \Diamond (for possibility). Both systems allow for the traditional interdefinability of necessity and possibility: $\Box p =_{df} \sim \Diamond \sim p$ and $\Diamond p =_{df} \sim \Box \sim p$. Both allow for the definition of strict implication: $p \Rightarrow q =_{df} \Box(p \supset q)$. And both include the familiar axioms of necessity ($\Box p \supset p$), possibility ($p \supset \Diamond p$) and distribution ($\Box(p \supset q) \supset (\Box p \supset \Box q)$).

While S5 contains as one of its defining features the axiom $\Diamond p \supset \Box \Diamond p$, S7 is the result of supplementing S3 with the axiom $\Diamond \Diamond p$. The axiom is easily seen to be incompatible with S5, since from $\Diamond \Diamond p$ it follows that no propositions are necessarily necessary, while from the rule of necessitation (If $\vdash p$, then $\vdash \Box p$) together with classical logic, it follows that many propositions are necessarily necessary.

The main difference between S5 and S7 thus concerns necessitation. Motivating necessitation is the observation that logical truths are somehow different from ordinary, empirical truths. Logical truths are necessary. Contingent truths are not. Even so, it is not immediately clear that necessitation properly captures this difference.

S7 lacks necessitation, but since it is an extension of S3, it still has in its place the slightly weaker rule, Nec(S7):

$$\text{Nec(S7) } \text{ If } \vdash_{CPL} p, \text{ then } \vdash_{S7} \Box p.$$

This rule states, not that every theorem of S7 is necessary, but only that every theorem of classical (propositional) logic is necessary. In other words, it includes all that is essential for capturing the observation that logical truths are somehow different from ordinary, contingent truths. Anyone who doubts this should ask the following question: Why accept as necessary anything other than logical truths? If this is a difficult question to answer, the acceptance of this restricted form of necessitation becomes not just plausible but mandatory. The acceptance of Nec(S7) also allows us to satisfy our intuition that the scientist's reply requires us to have at least one possible world in which $\Diamond \Diamond \sim (p \vee \sim p)$ and our intuition that laws such as the law of excluded middle and the law of non-contradiction will still remain necessary.

For example, given S7's restricted rule of necessitation, we can still derive $\Box(q \vee \sim q)$:

⁹For a more detailed description of S7, see Halldén (1950), Halldén (1951), Sobocinski (1962) and Irvine (2013).

- (Pf 5) 1. $(q \vee \sim q)$ CPL
 2. $\Box(q \vee \sim q)$ 1 Nec(S7)

Given the axiom $\Diamond\Diamond p$, we can also derive $\Diamond\sim\Box(q \vee \sim q)$:

- (Pf 6) 1. $\Diamond\Diamond p$ Axiom
 2. $\Diamond\Diamond(q \wedge \sim q)$ 1 Sub
 3. $\Diamond\sim\Box\sim(q \wedge \sim q)$ 2 $\Diamond - \Box$ Interdefinability
 4. $\Diamond\sim\Box(\sim q \vee \sim \sim q)$ 3 DeM
 5. $\Diamond\sim\Box(\sim q \vee q)$ 4 DN
 6. $\Diamond\sim\Box(q \vee \sim q)$ 5 Comm

This makes possible a more serious consideration of the scientist’s reply.

5 Logical Content

Although the scientist’s reply gives us a model for understanding theory choice in logic and S7 gives us an understanding of the modality of these choices, one question still remains: What is the theory-neutral arbitrator for deciding between alternative logical theories? Put another way, if Quine is right that when the non-classical logician denies the truth of classical logic “he only changes the subject,” then we are entitled to ask, what is the subject of logic?

The answer is that logic, mathematics and science all attempt to discover and study regularities (or patterns) found in nature.¹⁰ Science aims to discover and study regularities like $F = dv/dt(m) = ma$ (Newton’s second law) or $E = mc^2$ (Einstein’s mass-energy equivalence). Mathematics aims to discover and study regularities like $1 < n < p < 2n$ (Bertrand’s postulate) or $a^2 + b^2 = h^2$ (the Pythagorean theorem). Logic aims to discover and study regularities like $(p \wedge (p \supset q)) \supset q$ (*modus ponens*) and $\sim(p \wedge \sim p)$ (the law of non-contradiction). In the case of logic, the regularities studied are connected directly to the formal features of propositions, hence the commonly noted claim that logic studies propositional (or logical) form.¹¹

Discovering these regularities (or patterns) allows for the articulation of laws, universal regularities of various kinds. In the case of science, nomic possibility requires nothing more than being consistent with the laws of nature. In mathematics, mathematical possibility requires nothing more than being consistent with the laws of mathematics. In logic, logical possibility requires nothing more than being consistent with the laws of logic. In all three cases, the laws in question describe the

¹⁰All three do other things as well, but these investigations are central.

¹¹For an example of a naturalist theory of propositions, see Joaquin and Franklin (2021), in which a connection is postulated between a logical consequence relation and an inclusion relation among states of affairs (e.g., 27ff).

(single, physical) world which we inhabit. In science, this is a widely accepted view. In mathematics and logic, less so.¹²

Some sentences, as we have noted, exclude possibilities. “Socrates is older than Plato” excludes the possibility that Plato is older than Socrates. Because of the influence of positivism, the class of sentences that exclude possibilities has often been identified with the class of contingent sentences, but this is a mistake. Necessary sentences also exclude possibilities. For example, the sentence “ $11 + 3 = 14$ ” excludes the possibility that $11 + 3 = 2$. In other words, the adoption of ordinary Peano Arithmetic (PA) excludes (in the same application) the adoption of a modular arithmetic. We learn through trial and error that we need to adopt PA when describing the coins in my pocket, or the movement of particles through a selectively permeable membrane, or the addition of astronomical units. In contrast, we adopt mod-12 (or a “clock-face” arithmetic) rather than PA when describing meeting times based on a 12-hour time schedule in which a three-hour meeting that begins at 11 o’clock ends at 2 o’clock. Thus, a necessary truth like “ $11 + 3 = 14$ ” excludes possibilities (other potential necessary truths) in the same way as a contingent truth. It is only by investigating the world that we learn what is necessary and what is not, and which necessary truths accurately describe which kinds of phenomena.¹³

In logic, formally valid arguments are valid, not because of the particular terms they contain but because of their form. The actual truth or falsity of their component (atomic) sentences turns out not to be relevant to the study of implication or of other purely logical relations. The easiest way to discover the form of a proposition is to look for properties and relations that remain invariant upon substitution.¹⁴ The sentence “Socrates is older than Plato,” for example, asserts a relation between two

¹² Among the exceptions are both Russell and Prior. As Russell (1919) notes, “logic is concerned with the real world just as truly as zoology, though with its more abstract and general features” (169); and as Copeland (2020) notes, for Prior, “the point of a logical calculus was always that it had a subject matter, be it time, obligation, agency, or even biology, and a concern for philosophical problems never lay far below his theorems. It was the extra-symbolic world that mattered to Prior, not the formal results per se.” Or as Prior (1996) writes, “Philosophy, including Logic, is not primarily about language, but about the real world. . . Formalism, i.e. the theory that Logic is just about symbols and not things, is false. Nevertheless, it is important to ‘formalise’ as much as we can, i.e. to state truths about things in a rigorous language with a known and explicit structure” (45).

¹³ This example intentionally disregards the fatuously postulated distinction between pure and applied mathematics. It does so on the ground that it is only through so-called applications that a mathematical theory is able to receive a semantic interpretation. Merely postulating a model for PA, for example, without a real-world application has, as Russell (1919) notes in a slightly different context, “many advantages; they are the same as the advantages of theft over honest toil” (71).

¹⁴ More economically, we might refer only to relations (rather than to both properties and relations), taking properties to be one-place relations.

individuals. In this case, the constituents of the proposition being asserted turn out to be two terms and a relation: Socrates, Plato and older than. The form of a proposition is that which remains unchanged when every constituent of the proposition has been replaced by another.¹⁵ Thus the two propositions expressed by the sentences “Socrates is older than Plato” and “Alexander is taller than Aristotle” share the same form, Rxy , even though they share no constituents in common.

In this context, the reference to propositional (as opposed to sentential) form indicates that the form being studied is not that of an individual sentence token, such as the English sentence “Socrates lives in Athens” or the French sentence “Socrate vit à Athènes.” Instead, the form being studied is that of the proposition (or, when the sentence is true, the corresponding fact) being expressed by such sentences. And instead of distinguishing logical from non-logical terms on the basis of their presumed content neutrality, logical terms simply turn out to be terms that have the same semantic value under every interpretation of their corresponding language, L .

If this account of logical form is correct, logic not only studies something objective, it also studies something abstract. Logical forms are abstract, not in the Platonic sense that they are something separate from and independent of the (physical) world, but in the Aristotelian sense that they are *abstracted from* the (physical) world, in much the same way that ordinary part-whole relations allow us to focus on, and refer to, just one part of a more complex whole. Just as the physical properties of an object (such as shape, mass, colour and temperature) can be distinguished from the object as a whole, the form of a proposition can be distinguished from its propositional content. Both are *abstracted from* (that is, considered in isolation from) the rest of the physical world.¹⁶ The form of a proposition (or fact) is then something just as objective as the viscosity of a fluid, the charge of an electron or the spin of an elementary particle.

Scientific realism is the view that the world is composed of natural, physical, observer-independent particulars, properties and relations,¹⁷ and that science is a fallible but reliable guide to discovering their existence and nature. In logic, the nat-

¹⁵For example, see Russell (1919), 196-205.

¹⁶For discussion, see Armstrong (1978a), Armstrong (1978b) and Franklin (2014). As Franklin points out, on this view mathematics turns out to be “a science of the real world, just as much as biology or sociology are. . . . Aristotelianism regards mathematics as literally being about some aspect of reality, but about certain kinds of properties and relations rather than about individual objects” (1-2).

¹⁷In other words, it is the view that the world is a totality of *facts* (things having properties and standing in relations to one another), not just of *things* (mere atoms in the void). See Russell (1918/19), (1919) and (1924), Wittgenstein (1922), Armstrong (1978a) and (1978b) and Irvine (1990) and (2010). Quine (1948) gives a classic statement of the alternative, nominalistic view.

ural, physical, observer-independent objects being studied are the formal, abstract structures found in facts in the natural, physical world. It uses these structures to create accurate formation rules for language and transformation rules for logical relations such as implication. As Frege remarks, “It is applicability alone which elevates arithmetic from a game to the rank of a science.”¹⁸ The same is true of logic.

Although this view emphasizes realism with regard to logical structures within the natural, physical world, it need not preclude logical pluralism. It is consistent with the suggestion that there may turn out to be multiple logical structures ready to be discovered and applied in different contexts. What it does preclude is the traditional view that logical knowledge is justified *a priori*.¹⁹

That logical knowledge is justified *a posteriori* follows from the fact that logic, like science and mathematics, is never completely topic neutral. Just as contingent sentences exclude possibilities, so too do logical truths. Just as the contingent sentence “Socrates is older than Plato” is not topic neutral, neither is the logical truth, $\Box \sim (p \wedge \sim p)$. The sentence “Socrates is older than Plato” precludes the possibility that Plato is older than Socrates. The sentence “ $\Box \sim (p \wedge \sim p)$ ” precludes the possibility of dialetheism, the theory that there are some true (declarative) sentences whose negations are also true. To borrow Almog’s helpful phrase, just as ordinary, contingent sentences fail to be topic neutral about facts, logical sentences fail to be topic neutral about *pre-facts*, something Almog defines in terms of “permutation resistance.”²⁰ Thus, the ordinary fact that Quine is a philosopher is not invariant upon substitution. If we permute Quine with Mt Blanc, “the fact is gone.”²¹ In contrast, the pre-fact that Quine exists remains invariant upon substitution. If we permute Quine with Mt Blanc, the pre-fact remains. Here are three kinds of cases in which even the necessary truths of logic fail to be topic-neutral:

Terms, Properties and Relations. Just as Newtonian mechanics can be thought of as a special or limiting case of relativity theory, propositional logic can be thought of as a special or limiting case of first-order logic. The need to use predicate logic in place of propositional logic then tells us something about the world, just as

¹⁸Frege (1893/1903), 187.

¹⁹For example, this view precludes the suggestion expressed in Russell (1919) to the effect that “Logical propositions are such as can be known *a priori*, without study of the actual world. We only know from a study of empirical facts that Socrates is a man, but we know the correctness of the syllogism in its abstract form (i.e. when it is stated in terms of variables) without needing any appeal to experience” (204).

²⁰Almog (1989).

²¹Almog (1989), 201.

the need to use relativity theory in place of classical mechanics tells us something about the world. In one sense, the expressive power of propositional logic is just as great as that of predicate logic: The inference from “Socrates is older than Plato” to “Someone is older than Plato” can be fully *expressed* using the symbolism “ A , therefore B ,” but it cannot be fully *explained* by this symbolism, in the way that it is with the symbolism “ Osp , therefore $(\exists x)Oxp$.” Similarly, Aristotle’s principles of conversion, contraposition, obversion, contradiction, contrariety and subcontrariety successfully exhibit many logical relations, but they do so without the full theoretical generality of predicate logic. The choice between various logics — term logic, propositional logic and predicate logic — thus turns out to be just as informative about the detailed physical structure of facts in the world as the choice between classical mechanics and relativity theory.

Existence, Subsistence and Being. Decisions regarding predication likewise turn out to have significant ontological consequences. Rather than admitting only one form of predication, Zalta’s introduction of a second type of predication (*encoding* versus *exemplifying*) allows for a distinction to be made between intensional and extensional entities, should this be discovered to be necessary.²² The decision about whether to accept existence as a predicate has similarly significant implications. Anselm’s traditional ontological argument – the argument that since perfection implies existence, the idea of a non-existent, perfect God must be self-contradictory²³ – turns out to be valid in logics in which existence is accepted as a predicate. It becomes invalid in modern, first-order logic, a logic in which, as Russell reminds us, adding existence as a predicate turns out to be “metaphysically impossible.”²⁴

Bivalence, Dialetheism and Paraconsistency. Bivalence is the principle that every (declarative) sentence is either true or false, always one and never both. Dialetheism is the principle that some (declarative) sentences are both true and false or, alternatively, that there are some true (declarative) sentences whose negations are also true. A logic is paraconsistent if and only if its logical consequence relation fails to be explosive.²⁵ Accepting one or another such theory thus says something significant about the world that we are intending to describe. At the moment, it appears unlikely that a quantum superposition of (even infinitely many) wave functions will require a radical shift in logic away from classical bivalence to dialetheism,

²²Zalta (1988). Cf. Russell (1919), 169, Berto (2013) and Berto and Jago (2019).

²³Anselm (1078).

²⁴Russell (1900), 185. Cf. Moore (1936).

²⁵For example, see Priest (1987), Priest, Routley and Norman (1989) and Schotch, Brown and Jennings (2009).

but it is still early days. As we learn more, a full theory of quantum computing might require the modification of classical logic, just as quantum mechanics has led to changes in classical mechanics. But regardless, the choice of logic in all such cases can hardly be described as something content-neutral.

Numerous other choice points also help us distinguish between competing logical systems. In developing a theory of logical consequence, do we need to follow Frege and Russell and accept that $p \wedge \sim p$ implies everything? Or is it more accurate to follow Aristotle and Bolzano and accept that $p \wedge \sim p$ implies nothing?²⁶ In developing a theory of time, will a block universe lead us toward an untensed logic? Or will the arrow of time and McTaggart's B series lead us toward a tensed logic?²⁷ If it turns out that talk of properties of properties is unavoidable in science, will this require the adoption of a second-order or higher-order logic?²⁸ And will advances in quantum computing show that although classical second-order logic need not admit of a complete proof theory, these kinds of limitations will eventually be overcome and that changes in logic will be seen as the discovery of real features within the natural world?²⁹

Answering questions like these is a part of science. This type of intellectual work requires the integration of logic and mathematics with other scientific disciplines. It is thus natural to conclude that logic is not, as the positivists (and others, including Russell) thought, an *a priori* discipline. Instead, as Russell pointed out over a century ago, "Philosophy, if what has been said is correct, becomes indistinguishable from logic as that word has now come to be used,"³⁰ for it is logic that embodies our most fundamental physical and metaphysical assumptions about the world.

References

- [1] Almog, Joseph (1989) "Logic and the World," *Journal of Philosophical Logic*, 18, 197-220.
- [2] Anselm (1078) *Proslogium*, Sidney Norton Deane (trans.), New York: Open Court, 1903.
- [3] Armstrong, David (1978a) *Nominalism and Realism* (Universals and Scientific Realism, vol. 1), Cambridge: Cambridge University Press.
- [4] Armstrong, David (1978b) *A Theory of Universals* (Universals and Scientific Realism, vol. 2), Cambridge: Cambridge University Press.
- [5] Berg, J. (1962) *Bolzano's Logic*, Stockholm: Almqvist Wiksell.

²⁶For example, see Russell (1906), 201, Berg (1962) and Berg (1992), 82.

²⁷For example, see McTaggart (1908).

²⁸For example, see Shapiro (1991).

²⁹For example, see Lupacchini (2018).

³⁰Bertrand Russell (1914), 65.

- [6] Berg, J. (1992) *Ontology without Ultrafilters and Possible Worlds*, Sankt Augustin: Akademia.
- [7] Berto, Francesco (2013) *Existence as a Real Property*, Dordrecht: Springer.
- [8] Berto, Francesco and Mark Jago (2019) *Impossible Worlds*, Oxford: Oxford University Press.
- [9] Copeland, B. Jack (2020) “Arthur Prior,” *The Stanford Encyclopedia of Philosophy* (Spring 2020 Edition), Edward N. Zalta (ed.), plato.stanford.edu/archives/spr2020/entries/prior.
- [10] Franklin, James (2014) *An Aristotelian Realist Philosophy of Mathematics*, London: Palgrave MacMillan.
- [11] Frege, Gottlob (1893/1903) “Grundgesetze der Arithmetik, Begriffsschriftlich Abgeleitet,” 2 vols, trans. in part in *Translations from the Philosophical Writings of Gottlob Frege*, Peter Geach and Max Black (eds), New York: Philosophical Library, 1952, 137–58.
- [12] Haack, Susan (1974) *Deviant Logic*, Cambridge: Cambridge University Press.
- [13] Halldén, S. (1950) “Results concerning the Decision Problem of Lewis’ Calculi S3 and S6,” *Journal of Symbolic Logic*, 14, 230–36.
- [14] Halldén, S. (1951) “On the Semantic Non-completeness of Certain of Lewis’ Calculi,” *Journal of Symbolic Logic*, 16, 127–29.
- [15] Hjortland, Ole Thomassen (2017) “Anti-exceptionalism about Logic,” *Philosophical Studies*, 174, 631–58.
- [16] Hughes, G.E. and M.J. Cresswell (1968) *An Introduction to Modal Logic*, London, New York: Methuen.
- [17] Irvine, Andrew David (1990) “Nominalism, Realism and Physicalism in Mathematics,” in *Physicalism in Mathematics*, Dordrecht: Kluwer Academic Publishers, ix–xxvi.
- [18] Irvine, Andrew David (2010) “Frege on Number Properties,” *Studia Logica*, 96, 239–60.
- [19] Irvine, Andrew David (2013) “S7,” *Journal of Applied Logic*, 11, 523–29.
- [20] Joaquin, J.J. B and James Franklin (2021) “A Causal-Mentalist View of Propositions,” *Organon*, 28, 2 (March 23), 1–31.
- [21] Kripke, S. (1963) “Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi,” *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9, 67–96.
- [22] Lupacchini, Rossella (2018) “Church’s Thesis, Turing’s Limits, and Deutsch’s Principle,” in Cuffaro, M.E., and S.C. Fletcher (eds.) *Physical Perspectives on Computation, Computational Perspectives on Physics*, Cambridge: Cambridge University Press, 2018, 60–82.
- [23] McTaggart, J.M.E. (1908) “The Unreality of Time,” *Mind*, 17, 457–74.
- [24] Moore, G.E. (1936) “Is Existence a Predicate?” *Aristotelian Society*: Supplementary Volume 15, 154–88.
- [25] Priest, Graham (1987) *In Contradiction*, Dordrecht: Martinus Nijhoff Publishers.
- [26] Priest, Graham (2008) *An Introduction to Non-Classical Logic*, 2nd edn, Cambridge:

- Cambridge University Press.
- [27] Priest, Graham (2014) “Revising Logic”, in Penelope Rush (ed.), *Metaphysics of Logic*, Cambridge: Cambridge University Press, 2014, 211–23.
- [28] Priest, Graham, R. Routley and J. Norman (eds) (1989) *Paraconsistent Logic*, Munich: Philosophia Verlag.
- [29] Prior, Arthur (1996) “A Statement of Temporal Realism,” in Copeland, B.J. (ed.), *Logic and Reality: Essays on the Legacy of Arthur Prior*, Oxford: Clarendon Press, 1996, chap. 1.
- [30] Quine, W.V. (1948) “On What There Is”, *Review of Metaphysics*, 2, 21-38.
- [31] Quine, W.V. (1951) “Two Dogmas of Empiricism”, *Philosophical Review*, 60, 20-43.
- [32] Quine, W.V. (1960) *Word and Object*, Cambridge: MIT Press.
- [33] Quine, W.V. (1986) *Philosophy of Logic*, 2nd edn, Cambridge: Harvard University Press.
- [34] Russell, Bertrand (1900) *A Critical Exposition of the Philosophy of Leibniz*, Cambridge: Cambridge University Press.
- [35] Russell, Bertrand (1906) “The Theory of Implication,” *American Journal of Mathematics*, 28, 159-202.
- [36] Russell, Bertrand (1914) “On Scientific Method in Philosophy,” in John G. Slater (ed.), *The Philosophy of Logical Atomism and Other Essays, 1914-19*, (The Collected Papers of Bertrand Russell, vol. 8), New York: George Allen & Unwin Ltd., 55-73.
- [37] Russell, Bertrand (1918/19), “The Philosophy of Logical Atomism,” *Monist*, 28, 495-527; 29, 32-63, 190-222, 345-80; reprinted in Bertrand Russell, *Logic and Knowledge*, London: Allen and Unwin, 1956, 177-281.
- [38] Russell, Bertrand (1919) *Introduction to Mathematical Philosophy*, London: George Allen and Unwin Ltd.
- [39] Russell, Bertrand (1924) “Logical Atomism,” in J.H. Muirhead (ed.), *Contemporary British Philosophers*, London: Allen and Unwin, 1924, 356–83; reprinted in Bertrand Russell, *Logic and Knowledge*, London: Allen & Unwin, 1956, 323–43.
- [40] Schotch, P., B. Brown and R. Jennings (eds) (2009) *On Preserving*, Toronto: University of Toronto Press.
- [41] Shapiro, Stewart (1991) *Foundations without Foundationalism: A Case for Second-Order Logic*, New York: Oxford University Press.
- [42] Sobocinski, B. (1962) “A Note on the Regular and Irregular Modal Systems of Lewis,” *Notre Dame Journal of Formal Logic*, 3, 109–13.
- [43] Wittgenstein, Ludwig (1922) *Tractatus Logico-Philosophicus*, London: Kegan Paul.
- [44] Woods, John (2003) *Paradox and Paraconsistency*, Cambridge: Cambridge University Press.
- [45] Zalta, Edward N. (1988) *Intensional Logic and the Metaphysics of Intentionality*, Cambridge, MA: The MIT Press.

GROUPS DEFINABLE IN O-MINIMAL STRUCTURES: VARIOUS PROPERTIES AND A DIAGRAM

ANNALISA CONVERSANO
Massey University Auckland
a.conversano@massey.ac.nz

Abstract

We present a diagram surveying equivalence or strict implication for properties of different nature (algebraic, model theoretic, topological, etc.) about groups definable in o-minimal structures. All results are well-known and an extensive bibliography is provided.

This survey has been mostly prepared during the “Logic and Algorithms in Group Theory” Program (September 3 – December 20, 2018) at the Hausdorff Institute in Bonn. Many thanks to the organizers for the kind invitation and the great scientific program.

1 Introduction

Groups definable in o-minimal structures have been studied by many authors in the last 30 years and include algebraic groups over algebraically closed fields of characteristic 0, semialgebraic groups over real closed fields, important classes of real Lie groups such as abelian groups, compact groups and linear semisimple groups. See [52] for a nice introduction to the topic.

The previous diagram shows connections between several properties of a definably connected group G definable in a (sufficiently saturated) o-minimal expansion \mathcal{M} of a real closed field \mathcal{R} (although most results are known in arbitrary o-minimal structures – or o-minimal expansions of groups).

Notation in the diagram is as follows: we denote by $\mathcal{N}(G)$ the maximal normal definable torsion-free subgroup of G , by \bar{G} the quotient $G/\mathcal{N}(G)$, and by C the maximal normal definably connected definably compact subgroup of \bar{G} . Subgroups $\mathcal{N}(G)$ and C are proved to exist in [25, Proposition 2.6].

$E(G)$ denotes the *o-minimal Euler characteristic* of G (see Section 2).

Assuming G is definable with parameters over the reals, $G(\mathbb{R})$ denotes the corresponding real Lie group (see [67, Prop 2.5 & Rem 2.6]).

By a beautiful conjecture of Pillay from [68] (see [57] for a survey about its solution), the quotient G/G^{00} (of a definably connected group G by the smallest type-definable subgroup of bounded index G^{00}) is a compact Lie group, when equipped with the logic topology [17]. Finally, we denote by G^{000} the smallest subgroup of bounded index in G which is $\text{Aut}(\mathcal{M})$ -invariant.

Properties in each box are equivalent. An arrow indicates that a property in a box implies the properties in the other one. The dotted lines show that there is no implication in either direction. All implications are strict.

The diagram is explained as follows: Section 2 introduces the left-most column of torsion-free and definably compact groups. In Section 3 solvable groups G and their quotient $\bar{G} = G/\mathcal{N}(G)$ are discussed. Section 4 covers the right part of the diagram. Finally, Section 5 provides counterexamples for the implications that do not hold.

Further work on definable groups outside the scope of this diagram can be found in [4, 7, 8, 18, 22, 35, 36, 39, 34, 40, 43, 44, 46, 47, 48, 50, 51, 69, 73]. In recent years, the investigation has been extended by several authors to the wider class of locally definable groups. See, for instance, [5, 14, 31, 37, 38].

2 Torsion-free and definably compact groups

The left side of the diagram shows classes of torsion-free and definably compact groups. Roughly speaking, torsion-free groups resemble triangular groups of upper triangular matrices, and definably compact groups are very closely related to closed subgroups of orthogonal groups. Definable torsion-free groups have been studied, for instance, in [10, 24, 64] and definably compact groups, much more extensively, in [3, 9, 32, 41, 49, 54, 59].

If \mathcal{P} is a cell decomposition of a definable set X , the *o-minimal Euler characteristic* $E(X)$ is defined as the number of even-dimensional cells in \mathcal{P} minus the number of odd-dimensional cells in \mathcal{P} , and it does not depend on \mathcal{P} (see [28], Chapter 4). Strzebonski proved in [71, Prop 2.5] that a definable group G is **torsion-free** if and only if $E(G) = \pm 1$, and in [71, Prop 4.1] he deduced that such groups are uniquely divisible. By [64, 2.4 and 2.11] torsion-free groups G are also definably connected and solvable. In fact G is definably completely solvable (or triangular) [24, Theo 4.4]. Namely, G contains a chain of normal (in G) definably connected subgroups

$$\{e\} = G_0 < G_1 < \cdots < G_{n-1} < G_n = G$$

where $\dim G_i = i$. Note that since 1-dimensional definably connected groups are abelian [70], definable groups with such a chain are solvable. Abelian definably compact groups are not, in general, definably completely solvable. See, for instance, Examples 5.1 and 5.2 in [66]. By [21, Prop 2.5] and [64, Cor 5.7], a definable group is definably contractible if and only if it is torsion-free.

In \aleph_0 -saturated ordered structures, only sequences that are eventually constant converge, so the standard notion of compactness by sequences is not very useful in a model-theoretic context. Definable compactness has been introduced by Peterzil and Steinhorn in [66, Def 1.1] as a better analogue to compactness. A definable set X is **definably compact** if for every definable continuous function $f: (a, b) \rightarrow X$, the limits of $f(x)$, as x tends to a or to b , exist in X . Over any o-minimal structure, if X has the topology induced by the order of the ambient structure, this is equivalent to say that X is closed and bounded [66, Theo 2.1]. Thus over the reals, this coincides with the usual notion of compactness. The equivalent condition using definable open coverings is given in [59, Cor 2.3].

By [41, Theo 8.1 & Rem3 pg.588], G is definably compact if and only if G has *fsg* (finitely satisfiable generics). That is, there is a global type $p(x)$ and a small model \mathcal{R}_0 such that for every $g \in G$ the left translate $gp = \{\varphi(x) : \varphi(g^{-1}) \in p\}$ is finitely satisfiable in \mathcal{R}_0 .

If G is definably compact, then the o-minimal dimension of G as a definable set coincide with the dimension of G/G^{00} as a Lie group [41, Theo 8.1] (see [1] for another proof, and [32] for the case of an arbitrary o-minimal structure). Moreover, when G is defined over the reals, $G(\mathbb{R})$ is Lie isomorphic to G/G^{00} via the standard part map [41].

Conversely, if G is not definably compact, then by [66, Theo 1.2] G contains infinite definable torsion-free subgroups H , for which $H = H^{00} \subset G^{00}$ [26, 2.4], and so $\dim G/G^{00} < \dim G$. Moreover, when G is defined over the reals, the Lie group $G(\mathbb{R})$ contains the torsion-free closed subgroup $H(\mathbb{R})$ (where H is the torsion-free definable subgroup mentioned above). Since closed subgroups of compact groups are compact too, and compact Lie groups have torsion, it follows that in this case $G(\mathbb{R})$ cannot be Lie isomorphic to the compact G/G^{00} .

Both classes of definably compact and torsion-free definable groups are closed by definable subgroups and definable quotients [64, 2.3].

Since definably compact groups have torsion [33, 56], it follows that definably compact and torsion-free definable subgroups of a definable group always have trivial intersection. If a definable group is not definably compact, then it contains a 1-dimensional torsion-free definable subgroup [66, Theo 1.2]. Therefore, for both classes, the condition of being torsion-free or definably compact is equivalent to not having any non-trivial definable subgroup of the other class.

Every definable linear group G can be decomposed into a product $G = KH$ of a (maximal) definable torsion-free subgroup H and a (maximal) definably compact subgroup K [20, Theo 4.1]. If G is not linear, G may not have maximal definably compact subgroups, but a similar decomposition holds where K is abstractly compact. That is, it is isomorphic to a definably compact subgroup of $G/\mathcal{N}(G)$ [20, Theo 1.5]. On the other hand, maximal definable torsion-free subgroups always exist [21, Cor 2.4] and they are conjugate to each other [65, Theo 3.26].

3 Solvable groups

Solvable definable groups have been first studied by Edmundo in [29]. As observed in [25, Prop 2.2], the quotient \bar{G} of a definable solvable group G by its maximal normal definable torsion-free subgroup $\mathcal{N}(G)$ is definably compact. Moreover, if G is not definably compact then $\mathcal{N}(G)$ is infinite (unlike semisimple groups that are not definably compact). We give below a direct proof of both facts (see [19, Theo 2.5.1]):

Theorem 3.1. *Let G be a solvable definable group and let $N = \mathcal{N}(G)$ be its maximal normal definable torsion-free subgroup. If G is not definably compact, then N is infinite and G/N is a definably compact group.*

Proof. Because $N \subseteq G^0$ and G/G^0 is finite (so definably compact), we can suppose G is definably connected. We proceed by induction on $n = \dim G$. The case $n = 1$ is obvious, so let $\dim G = n > 1$.

If G is abelian, the theorem can be extracted from [64, 2.6]. The argument is that if G/N is not definably compact, then by [66, Theo 1.2] there is a definable 1-dimensional torsion-free subgroup H in G/N , and the pull-back of H in G is a definable torsion-free subgroup of G , in contradiction with the maximality of N .

Let G be now non-abelian. Since G is solvable and definably connected, there is a normal solvable definable subgroup $S < G$ such that G/S is abelian and infinite (definability of S follows from [60, 1.17]). We distinguish the cases where S is definably compact and where S is not definably compact.

- If S is definably compact then G/S is not. By the abelian case, the maximal normal definable torsion-free subgroup N_1 of G/S is infinite and $(G/S)/N_1$ is definably compact. If $\pi: G \rightarrow G/S$ is the canonical projection, let $N' = \pi^{-1}(N_1)$. By [21, Lemma 2.3] the definable exact sequence

$$1 \longrightarrow S \xrightarrow{i} N' \xrightarrow{\pi} N_1 \longrightarrow 1$$

splits definably in a direct product, thus G contains a definable subgroup N definably isomorphic to N_1 such that $N' = S \times N$. Since $S \cong N'/N$ which is definably compact, it follows that N is the maximal normal definable torsion-free subgroup of N' , and it is normal in G as well.

To show that G/N is definably compact, it is enough to provide a normal definable subgroup of G/N which is definably compact, such that quotienting by it we obtain a definably compact group. One such subgroup is N'/N which is definably isomorphic to S , and the quotient $(G/N)/(N'/N)$ is definably isomorphic to $(G/S)/N_1$, as the following diagram shows by “the 3×3 lemma”.

$$\begin{array}{ccccccc}
 & & & 1 & & 1 & \\
 & & & \downarrow & & \downarrow & \\
 1 & \longrightarrow & S & \longrightarrow & N' & \longrightarrow & N_1 & \longrightarrow & 1 \\
 & & \uparrow & & \downarrow & & \downarrow & & \\
 1 & \longrightarrow & S & \longrightarrow & G & \longrightarrow & G/S & \longrightarrow & 1 \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & G/N' & \longleftrightarrow & (G/S)/N_1 & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 1 & & 1 & &
 \end{array}$$

Thus N is the maximal normal definable torsion-free subgroup of G , it is infinite and the theorem is proved for the case where S is definably compact.

- If S is not definably compact, then by induction the maximal normal definable torsion-free subgroup N_1 of S is infinite (possibly $N_1 = S$) and S/N_1 is definably compact. Note that N_1 is normal in G as well. If G/N_1 is definably compact then N_1 is the maximal normal definable torsion-free subgroup of G and we are done. Otherwise, again by induction, its infinite maximal normal definable torsion-free subgroup N_2 is such that $(G/N_1)/N_2$ is definably compact.

Let N be the pull-back in G of N_2 . Note that N is torsion-free and G/N is definably isomorphic to $(G/N_1)/N_2$ which is definably compact. Hence N is the maximal normal definable torsion-free subgroup of G and it is infinite, since it contains N_1 .

□

While any connected solvable real Lie group splits into a product of 1-dimensional connected subgroups [45, Lem 3.6], definable solvable groups with torsion are not, in general, definably completely solvable. See Example 5.3.

If G is solvable, then $[G, G]$ is nilpotent by [29, Theo 6.9]. Since $G/[G, G]$ is abelian for any group G , the converse is obvious.

If G is solvable, then $[G, G]$ is definable by [6, Theo 1.3]. Moreover, definably connected definably compact solvable groups are abelian by [63, Cor 5.4]. Therefore, if G is solvable, then \bar{G} is abelian and $[G, G] \subseteq \mathcal{N}(G)$ is torsion-free. Conversely, if $[G, G]$ is definable and torsion-free, then it is solvable by [64, 2.11], and G is solvable as well. The condition of $[G, G]$ being definable is necessary to conclude that G is solvable, since $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ is torsion-free (see Example 5.7 where $[G, G] = \widetilde{\mathrm{SL}}_2(\mathbb{R})$).

4 Definable amenability, exactness, and connected components

Recall that G is said to be *definably amenable* if it has a left invariant Keisler measure and G has *bounded orbit* if there is some complete type $p \in S_G(\mathcal{M})$ whose stabilizer $\mathrm{Stab}(p) = \{g \in G : gp = p\}$ has bounded index in G . In [41] groups with *fsg* in complete NIP theories are shown to be both definably amenable and with bounded orbit, by lifting the Haar measure of the compact Lie group G/G^{00} to a left invariant Keisler measure on G , making use of a global generic type p , whose stabilizer is G^{00} . The two classes of groups are indeed shown to be the same in the o-minimal context [25, Cor 4.12], and to coincide with the class of groups G such that $\bar{G} = G/\mathcal{N}(G)$ is definably compact [25, Prop 4.6]. In short, if \bar{G} is definably compact, then G is definably amenable because torsion-free and definably compact groups are definably amenable. Conversely, if \bar{G} is not definably compact, then there is a definable quotient of G that is a definably simple not definably compact group, and such groups are not definably amenable (such as $\mathrm{PSL}_2(\mathbb{R})$ [41, Rem 5.2]).

Before discussing the remaining equivalent conditions for \bar{G} to be definably compact, let us consider the upper right part of the diagram.

A non-abelian definable group is said to be **definably simple** if it does not have any non-trivial normal definable subgroup. By [62, Theo 5.1], definably simple groups are exactly the definable groups that are elementarily equivalent to a (non-abelian) simple real Lie group.

As the center $Z(G)$ is a normal definable subgroup of any definable group G , it follows that definably simple groups are centerless.

Given any definable group G , the quotient $G/Z(G)$ can be definably embedded in some $\mathrm{GL}_n(\mathcal{R})$ through the adjoint representation [55, Cor 3.3]. Therefore **centerless groups** are linearizable.

An infinite definable group is said to be **semisimple** if it does not have any infinite abelian (or, equivalently, solvable) – definable or not – normal subgroup. Semisimple definable groups have been studied by Peterzil, Pillay and Starchenko in [60, 61]. They prove that given a semisimple group G definable in an arbitrary o-minimal structure \mathcal{M} , the quotient by the center $G/Z(G)$ is a direct product of definably simple groups H_i , and each H_i is definably isomorphic to a definable subgroup of $\mathrm{GL}_n(\mathcal{R}_i)$, for some real closed field \mathcal{R}_i definable in \mathcal{M} .

It is well-known that every definable group G has a maximal normal definably connected solvable subgroup R (called *the solvable radical*) and the quotient G/R is a semisimple definable group. Since a proof does not appear in the literature, as far as we know, we add it below together with the proof of the existence of a maximal normal solvable subgroup:

Remark 4.1. Let G be a definably connected definable group. Then G has a unique normal solvable definably connected subgroup R such that G/R is trivial or semisimple. The subgroup R is the maximal normal solvable definably connected subgroup of G . If G/R is semisimple and $\pi: G \rightarrow G/R$ is the canonical projection, then $\pi^{-1}(Z(G/R))$ is the maximal normal solvable subgroup of G .

Proof. By induction on $n = \dim G$. If G is not semisimple, let $A < G$ be an infinite normal definable subgroup. If G/A is semisimple, take $R = A^0$. Otherwise, by induction there is a normal solvable definably connected subgroup $S < G/A$ such that $(G/A)/S$ is semisimple. Then take R to be the definably connected component of the identity of the pre-image of S in G .

If S is another normal solvable definably connected subgroup of G , then RS/R is a normal solvable definably connected subgroup of the semisimple group G/R . Therefore $RS = R$ and $S \subseteq R$.

If $S \subsetneq R$, then $\dim S < \dim R$ and R/S is an infinite solvable definable subgroup of G/S . Therefore G/S cannot be semisimple and R is unique.

Let $H = \pi^{-1}(Z(G/R))$. As mentioned above, G/H is a direct product of definably simple groups and does not contain any solvable (definable or not) normal subgroup. Therefore H is the maximal normal solvable subgroup of G . □

Semisimple groups G are **perfect** by [42, 3.1]. That is, G is equal to its commutator subgroup $[G, G]$ or, equivalently, G does not have any proper abelian quotient.

By [62, Theo 4.5], **linear** groups G have a definable Levi decomposition. That is, G contains (maximal) semisimple definable subgroups S (all conjugate) such that $G = RS$, where R is the solvable radical. See [27, Theo 1.1] for a Levi decomposition

of an arbitrary definably connected group G , where S is in general a countable union of definable sets.

Very recently Baro proved in [2, Theo 3.1] that the commutator subgroup of a linear group is definable. In general, this is not the case. See Examples 5.7 and 5.8, where the commutator subgroup of G is isomorphic to the universal cover of $SL_2(\mathbb{R})$. These are also examples of groups without a definable Levi decomposition. Another definable group with no definable Levi subgroups is Example 5.9, whereas its commutator subgroup is definable. On the other hand, we expect that definability of Levi subgroups implies definability of the commutator subgroup.

Remark 4.2. If \bar{G} is definably compact, then G has a definable Levi decomposition.

Proof. If \bar{G} is definably compact, then by [42, Cor 6.4], $\bar{G} = Z(\bar{G}) \cdot [\bar{G}, \bar{G}]$ is a definable Levi decomposition of \bar{G} . Let H be the pull-back of $[\bar{G}, \bar{G}]$ in G . So H is a definable extension of a definably compact semisimple group by a definable torsion group $\mathcal{N}(G)$. The extension splits definably by [27, Prop 5.1] and $H = \mathcal{N}(G) \times S$, for some definably compact semisimple definable subgroup S . As $G/H \subset Z(\bar{G})$ is abelian, then clearly S is a definable Levi subgroup of G . \square

If $[G, G]$ is definable, then clearly $[\bar{G}, \bar{G}]$ is definable as well.

Remark 4.3. $[\bar{G}, \bar{G}]$ is definable if and only if \bar{G} has a definable Levi decomposition.

Proof. By Theorem 3.1, $\mathcal{N}(\bar{G}) = \{e\}$ and the solvable radical \bar{R} of \bar{G} is definably compact. By [63, Theo 4.4] \bar{R} is abelian. To see that \bar{R} is central in \bar{G} , let $s: \bar{G}/\bar{R} \rightarrow \bar{G}$ be a definable section of the canonical projection $\pi: \bar{G} \rightarrow \bar{G}/\bar{R}$.

For every $g \in \bar{G}$, the conjugation map $f_g: \bar{R} \rightarrow \bar{R}$ mapping $a \mapsto gag^{-1}$ is a definable automorphisms of \bar{R} , so there is a homomorphism $\Phi: \bar{G} \rightarrow \text{Aut}(\bar{R})$, given by $g \mapsto f_g$ such that $\bar{R} \subseteq \ker \Phi$, being \bar{R} abelian. Thus Φ induces the definable homomorphism

$$\begin{aligned} \varphi: \bar{G}/\bar{R} &\longrightarrow \text{Aut}(\bar{R}) \\ x &\mapsto (a \mapsto s(x)as(x)^{-1}) \end{aligned}$$

which does not depend on the choice of the section s . Since \bar{R} has no definable families of definable automorphisms by [63, Cor 5.3], it follows that $\varphi(\bar{G}/\bar{R}) = \{e\}$, and so $\bar{R} \subseteq Z(\bar{G})$. By [27, Lem 3.2] $\bar{G} = Z(\bar{G})[\bar{G}, \bar{G}]$ and $[\bar{G}, \bar{G}]$ coincides with the (unique) Levi subgroup of \bar{G} . Hence $[\bar{G}, \bar{G}]$ is definable if and only if \bar{G} has a definable Levi decomposition. \square

Finally, by [26, Prop 2.6], whenever \bar{G} has a definable Levi decomposition, then $G^{00} = G^{000}$.

We now go back to the class of groups G such that \bar{G} is definably compact. In Remark 4.2 we observed that G has a definable Levi decomposition $G = RS$, and clearly *Levi subgroups S are definably compact*.

Conversely, if $G = RS$ and S is definably compact, then \bar{G} must be definably compact, as $R/\mathcal{N}(G)$ is definably compact by Theorem 3.1.

When G is defined over the reals, then \bar{G} is definably compact if and only if G/G^{00} is isomorphic to a maximal compact subgroup of $G(\mathbb{R})$ [26, Prop 2.10]. Moreover, \bar{G} definably compact is also equivalent for G^{00} to be torsion-free and for G and G/G^{00} to have same homotopy and cohomology types [19]. A last equivalent condition regards exactness of sequences. Given an exact sequence of definably connected groups

$$1 \rightarrow H \rightarrow G \rightarrow Q \rightarrow 1$$

the induced sequences $1 \rightarrow H^{00} \rightarrow G^{00} \rightarrow Q^{00} \rightarrow 1$ and $1 \rightarrow H^{000} \rightarrow G^{000} \rightarrow Q^{000} \rightarrow 1$ turn out to be exact if and only if $H \cap G^{00} = H^{00}$ and $H \cap G^{000} = H^{000}$ (see [26, Lemma 2.1]). So we have the following definition from [26]:

Definition 4.4. Let G be a definable group. We say that

1. G has *the almost exactness property* if for every normal definable subgroup H of G , H^{00} has finite index in $G^{00} \cap H$.
2. G has *the exactness property* if for every normal definable subgroup H of G , $H^{00} = G^{00} \cap H$.
3. G has *the strong exactness property* if for every definable subgroup H of G , $H^{00} = G^{00} \cap H$.

When G is definably compact, then G **has the strong exactness property** by [12]. By [26, Theo 4.11] G has the strong exactness property if and only if \bar{G} is definably compact.

If G **has the exactness property**, then G/G^{00} is isomorphic to C/C^{00} by [26, Prop 4.10]. If G is defined over the reals, this is equivalent for G/G^{00} to be isomorphic to the maximal normal compact connected subgroup of $G(\mathbb{R})$ [26].

Finally, G **has the almost exactness property** if and only if $G^{00} = G^{000}$ ([26, Theo 4.4]) if and only if G/G^{00} is isogenous to C/C^{00} ([26, Rem 4.7]).

5 Counterexamples

In this last section we describe 9 semialgebraic groups proving that all implications in the diagram are strict. The enumeration coincides with the number on the corresponding arrows – or dotted lines – in the diagram.

Example 5.1. : $\boxed{\mathcal{R}^2 \rtimes \text{SO}_2(\mathcal{R})}$

$\text{SO}_2(\mathcal{R})$ denotes the special group of orthogonal matrices 2×2 with coefficients in \mathcal{R} . That is,

$$\text{SO}_2(\mathcal{R}) = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : a, b \in \mathcal{R}, a^2 + b^2 = 1 \right\}$$

$\text{SO}_2(\mathcal{R})$ is an abelian 1-dimensional definably compact group (abstractly) isomorphic to the unit circle.

Let now G be the following group of matrices 3×3 :

$$G = \left\{ \begin{pmatrix} a & -b & x \\ b & a & y \\ 0 & 0 & 1 \end{pmatrix} : a, b, x, y \in \mathcal{R}, a^2 + b^2 = 1 \right\}$$

The group G is definably isomorphic to the semidirect product $\mathcal{R}^2 \rtimes \text{SO}_2(\mathcal{R})$ where the action is the matrix multiplication of $\text{SO}_2(\mathcal{R})$ on $(\mathcal{R}^2, +)$.

Note that $\mathcal{N}(G) \cong \mathcal{R}^2$ and $G/\mathcal{N}(G) \cong \text{SO}_2(\mathcal{R})$ is definably compact, but G is not definably compact nor torsion-free.

Moreover, G is centerless but it is not perfect nor definably simple.

Example 5.2. : $\boxed{\mathcal{R}^3 \rtimes \text{SO}_3(\mathcal{R})}$

A similar group G of matrices 4×4 can be obtained considering the action given by the matrix multiplication of $\text{SO}_3(\mathcal{R})$, the special orthogonal group of matrices 3×3 with coefficients in \mathcal{R} , on $(\mathcal{R}^3, +)$.

Note that $\mathcal{N}(G) \cong \mathcal{R}^3$ and $G/\mathcal{N}(G) \cong \text{SO}_3(\mathcal{R})$ is a definably compact definably simple group. Thus G is not solvable.

Moreover, G is a perfect group that is not semisimple.

Example 5.3. : $\boxed{n\text{-dimensional torus}}$

This is Example 5.2 in [66]. Given $\mathcal{B} = \{v_1, \dots, v_n\}$ a set of linearly independent vectors in \mathbb{R}^n , let $L = v_1\mathbb{Z} + \dots + v_n\mathbb{Z}$ be the integral lattice generated by \mathcal{B} , and E the equivalence on \mathbb{R}^n induced by L . That is, for each $a, b \in \mathbb{R}^n$

$$a E b \iff a - b \in L$$

For every bounded box D containing the fundamental parallelogram of L , there is a finite sublattice $L' \subset L$ such that for all $a, b \in D$ we have

$$a E b \iff a - b \in L'$$

Thus, even though the quotient group $(\mathbb{R}^n, +)/L$ is not definable, it is abstractly isomorphic to the definable group $G = (S, \oplus)$, where $S \subset D$ is a definable n -dimensional set of representatives of $(\mathbb{R}^n, +)/L$ and the group operation \oplus on S is defined as

$$x \oplus y = z \iff (x + y) E z$$

Peterzil and Steinhorn prove in [66] that \mathcal{B} can be chosen so that the proper definable subgroups of G are all finite. When this is the case and $n > 1$, G is an example of an abelian group that is not definably completely solvable.

Moreover G is definably compact, but it is not definably isomorphic to a definable subgroup of the general linear group, as abelian definably compact linear groups split in a direct product of 1-dimensional definable subgroups by [62, Lem 3.9].

Example 5.4. : $\mathrm{SO}_2(\mathcal{R}) \times \mathrm{SL}_2(\mathcal{R})$

$\mathrm{SL}_2(\mathcal{R})$ denotes the special linear group 2×2 , that is the group of matrices with coefficients in \mathcal{R} with determinant equal to 1. It is a semisimple definably connected group that is not definably compact.

Let G be the direct product of $\mathrm{SO}_2(\mathcal{R})$ by $\mathrm{SL}_2(\mathcal{R})$. The group G is linear and is not centerless. Moreover G has a definable Levi decomposition (the unique Levi subgroup is SL_2) and $\bar{G} = G$ is not definably compact. Its commutator subgroup $[G, G]$ is again $\mathrm{SL}_2(\mathcal{R})$, so it is definable, but G is not perfect. Finally, note that $G^{00} = \mathrm{SO}_2(\mathcal{R})^{00} \times \mathrm{SL}_2(\mathcal{R})$, so G has the exactness property by [26, 4.3].

Example 5.5. : $\mathrm{SO}_2(\mathcal{R}) \times_{\{I, -I\}} \mathrm{SL}_2(\mathcal{R})$

Let now consider the group G obtained from $\mathrm{SO}_2(\mathcal{R}) \times \mathrm{SL}_2(\mathcal{R})$ above by identifying the common central subgroup $H = \{I, -I\}$. A way to define G is the following:

Fixed a definable choice of representatives $s: \mathrm{PSL}_2(\mathcal{R}) \rightarrow \mathrm{SL}_2(\mathcal{R})$ (for instance, $s(A)$ the matrix $[a_{ij}]$ such that $a_{11} > 0$ or $a_{11} = 0$ and $a_{12} > 0$), let $G = \mathrm{SO}_2(\mathcal{R}) \times \mathrm{PSL}_2(\mathcal{R})$ be with the group operation given by

$$(X_1, A_1) * (X_2, A_2) = \begin{cases} (X_1 X_2, A_1 A_2) & \text{if } s(A_1 A_2) = s(A_1) s(A_2), \\ (-X_1 X_2, A_1 A_2) & \text{otherwise.} \end{cases}$$

G contains normal definable subgroups $R = \mathrm{SO}_2(\mathcal{R}) \times \{I\}$ and $S = \{\pm I\} \times \mathrm{PSL}_2(\mathcal{R})$ (S is definably isomorphic to $\mathrm{SL}_2(\mathcal{R})$), such that $RS = G$ and $R \cap S = \{(I, I), (-I, I)\}$. Note that $R^{00} = \mathrm{SO}_2(\mathcal{R})^{00} \times \{I\}$, $G^{00} = \pm \mathrm{SO}_2(\mathcal{R})^{00} \times \mathrm{PSL}_2(\mathcal{R})$, and $G^{00} \cap R = \pm \mathrm{SO}_2(\mathcal{R})^{00} \times \{I\}$. Therefore R^{00} is properly contained in $G^{00} \cap R$ and G does not have the exactness property. Moreover $G/G^{00} \cong \mathrm{SO}_2(\mathcal{R})/\mathrm{SO}_2(\mathcal{R})^{00} \cong C/C^{00}$.

Example 5.6. : $\boxed{\mathrm{SU}_2(\mathcal{K}) \times_{\{\pm I\}} \mathrm{SL}_2(\mathcal{R})}$

Denoted by $\mathcal{K} = \mathcal{R}(\sqrt{-1})$ the algebraic closure of the real closed field \mathcal{R} , every semialgebraic subgroup of the general linear group $\mathrm{GL}_n(\mathcal{K})$ can be viewed as a definable subgroup of $\mathrm{GL}_{2n}(\mathcal{R})$. One such group is the special unitary group:

$$\mathrm{SU}_n(\mathcal{K}) = \{X \in \mathrm{GL}_n(\mathcal{K}) : X \bar{X}^T = I, \det X = 1\}$$

When $n = 2$, $\mathrm{SU}_2(\mathcal{K})$ is the universal cover of $\mathrm{SO}_3(\mathcal{R})$ with kernel $\{\pm I\}$, therefore it is a semisimple definably compact group.

Let now G be the amalgamated direct product of $\mathrm{SU}_2(\mathcal{K})$ and $\mathrm{SL}_2(\mathcal{R})$ over the common central subgroup $\{\pm I\}$, obtained definably as the previous example. Note that $\mathcal{N}(G) = \{e\}$ and the maximal normal definably compact definably connected subgroup of $\bar{G} = G$ is $C = \mathrm{SU}_2(\mathcal{K})$. Therefore $G/G^{00} \cong \mathrm{SO}_3(\mathbb{R})$, while $C/C^{00} \cong \mathrm{SU}_2(\mathbb{C})$. Moreover, G is a perfect group that is not centerless.

Example 5.7. : $\boxed{\mathbb{R} \times_{\mathbb{Z}} \widetilde{\mathrm{SL}}_2(\mathbb{R})}$

We now describe a small modification of [25, Example 2.10] and [19, Example 3.1.7]. The idea is to define a semialgebraic group that is isomorphic to the amalgamated direct product of $(\mathbb{R}, +)$ and the universal cover of $\mathrm{SL}_2(\mathbb{R})$ over the common subgroup $(\mathbb{Z}, +) \cong Z(\widetilde{\mathrm{SL}}_2(\mathbb{R}))$.

Let $\pi: \widetilde{\mathrm{SL}}_2(\mathbb{R}) \rightarrow \mathrm{SL}_2(\mathbb{R})$ be the universal cover of $\mathrm{SL}_2(\mathbb{R})$ and fix $s: \mathrm{SL}_2(\mathbb{R}) \rightarrow \widetilde{\mathrm{SL}}_2(\mathbb{R})$ a definable section. Recall that $\widetilde{\mathrm{SL}}_2(\mathbb{R})$ is a semisimple Lie group with

infinite center $Z(\widetilde{\mathrm{SL}}_2(\mathbb{R})) = \ker \pi$ and π is a homomorphism. Thus the image of the 2-cocycle $h_s: \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}) \rightarrow \widetilde{\mathrm{SL}}_2(\mathbb{R})$ given by $h_s(x_1, x_2) = s(x_1)s(x_2)s(x_1x_2)^{-1}$ is contained in $\ker \pi$. Hrushovski, Peterzil and Pillay prove in [42, 8.5] that it is actually a definable map and takes only finitely many values. Note that $\ker \pi$ is isomorphic (as an abstract group) to $(\mathbb{Z}, +)$. Fixed a generator v of $\ker \pi$, if $h_s(x_1, x_2) = kv$, $k \in \mathbb{Z}$, with abuse of notation we write $h_s(x_1, x_2)$ meaning the corresponding $k \in \mathbb{Z}$.

Now let $G = \mathbb{R} \times \mathrm{SL}_2(\mathbb{R})$. Fixed $a \neq 0$, consider the definable group operation on G given by

$$(t_1, x_1) * (t_2, x_2) = (t_1 + t_2 + h_s(x_1, x_2)a, x_1x_2).$$

Then $(G, *)$ is a semialgebraic group such that $[G, G]$ is not definable and with no definable Levi subgroups. To see this, note that the center of G (whose connected component coincides with the solvable radical) is the subgroup $Z = Z(G) = \mathbb{R} \times \{\pm I\}$. We can identify the subgroup $\langle a \rangle \times \mathrm{SL}_2(\mathbb{R})$ with $\widetilde{\mathrm{SL}}_2(\mathbb{R})$, which is isomorphic to it by construction. Every element g of G is a product $g = zx$, for some $z \in Z$ and $x \in \widetilde{\mathrm{SL}}_2(\mathbb{R})$. Therefore every commutator of G is a commutator of $\widetilde{\mathrm{SL}}_2(\mathbb{R})$, and $[G, G] = \widetilde{\mathrm{SL}}_2(\mathbb{R})$ is the unique Levi subgroup of G . However, $G/\mathcal{N}(G) = \mathrm{SL}_2(\mathbb{R})$, so $\bar{G} = [\bar{G}, \bar{G}]$.

Example 5.8. : $\mathrm{SO}_2(\mathcal{R}) \times_{\mathbb{Z}} \widetilde{\mathrm{SL}}_2(\mathcal{R}), \mathbb{Z} \subset \mathrm{SO}_2^{00}(\mathcal{R})$

Let us now consider the semialgebraic group G in a sufficiently saturated elementary extension \mathcal{R} of the reals, obtained as the previous example by replacing \mathbb{R} with $\mathrm{SO}_2(\mathcal{R})$ and by taking $a \in \mathrm{SO}_2(\mathcal{R})^{00}$, so that $\mathcal{N}(G) = \{e\}$ and $G = \bar{G}$.

As before, $[G, G] = \langle a \rangle \times \mathrm{SL}_2(\mathcal{R})$ is not definable, and $G^{00} = G^{000} = \mathrm{SO}_2(\mathcal{R})^{00} \times \mathrm{SL}_2(\mathcal{R})$. Therefore $G/G^{00} \cong \mathrm{SO}_2(\mathcal{R})/\mathrm{SO}_2^{00}(\mathcal{R}) \cong C/C^{00}$.

Example 5.9. : $\mathrm{UT}_3(\mathbb{R}) \times_{\mathbb{Z}} \widetilde{\mathrm{SL}}_2(\mathbb{R})$

Let $\mathrm{UT}_3(\mathbb{R})$ be the group of real unipotent matrices (that is, upper triangular matrices with 1's on the diagonal) 3×3 . Given $h_s: \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}) \rightarrow \widetilde{\mathrm{SL}}_2(\mathbb{R})$ as in Example 5.7, one can define on $G = \mathrm{UT}_3(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ the following group operation:

$$(A, X) * (B, Y) = \left(AB + \begin{pmatrix} 0 & 0 & h_s(X, Y) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, XY \right)$$

so that $Z(G)^0 = Z(\mathrm{UT}_3(\mathbb{R})) \cong (\mathbb{R}, +)$ and $H = \mathbb{Z} \times \mathrm{SL}_2(\mathbb{R}) \cong \widetilde{\mathrm{SL}}_2(\mathbb{R}) \triangleleft G$ is the unique Levi subgroup of G . Note that $[G, G] = Z(G)^0 \times \mathrm{SL}_2(\mathbb{R})$ is isomorphic to Example 5.7. Therefore the commutator subgroup of G is definable, while Levi subgroups are not.

In [23] it is proved that the nilpotent Lie group G/H interprets the real field expanded with a predicate for the integers, and therefore it interprets every real Lie group.

References

- [1] A. Achille and A. Berarducci, A Vietoris-Smale mapping theorem for the homotopy of hyperdefinable sets, *Selecta Mathematica* 4(2018), 3445–3473.
- [2] E. Baro, The derived subgroup of linear and simply-connected o-minimal groups, *Journal of Algebra*, 527 (2019), 1–19.
- [3] E. Baro and A. Berarducci, Topology of definable abelian groups in o-minimal structures, *Bull. Lond. Math. Soc.* 44 (2012), no. 3, 473–479.
- [4] E. Baro, A. Berarducci, M. Otero, Cartan subgroups and regular points of o-minimal groups, *J. London. Math. Soc.*, to appear.
- [5] E. Baro, P. Eleftheriou, Y. Peterzil, Locally definable and approximate subgroups of semialgebraic groups, *Journal of Mathematical Logic*, 20 (2020).
- [6] E. Baro, E. Jaligot, M. Otero, Commutators in groups definable in o-minimal structures, *Proc. Amer. Math. Soc.* 140 (2012), no. 10, 3629–3643.
- [7] E. Baro, E. Jaligot, M. Otero, Cartan subgroups of groups definable in o-minimal structures, *J. Inst. Math. Jussieu* 13 no. 4 (2014) 849–893.
- [8] E. Baro and M. Otero, On o-minimal homotopy groups, *Quart. J. Math.* 61(2010) 275–289.
- [9] E. Barriga, Definably compact groups definable in real closed fields, *Isr. J. Math.* 238(2020), 121–166.
- [10] E. Barriga, A. Onshuus, Classifying torsion free groups in o-minimal expansions of real closed fields. *Ann. Pure Appl. Logic* 167 (2016), 1267–1297.

- [11] A. Berarducci, Zero-groups and maximal tori, *Logic Colloquium '04* (Andretta, A., Kearnes, K., and Zambella, D., editors), *Lecture Notes in Logic*, vol. 29, ASL, 2007, pp. 33–45.
- [12] A. Berarducci, O-minimal spectra, infinitesimal subgroups and cohomology, *Journal of Symbolic Logic*, 72(4), 1177–1193, 2007.
- [13] A. Berarducci, Cohomology of groups in o-minimal structures: acyclicity of the infinitesimal subgroup, *Journal of Symbolic Logic*, 74(3), 891–900, 2009.
- [14] A. Berarducci, M. Edmundo, M. Mamino. Discrete subgroups of locally definable groups, *Selecta Mathematica*, 19(3), 719–736, 2013.
- [15] A. Berarducci, M. Mamino, On the homotopy type of definable groups in an o-minimal structure, *Journal of the London Mathematical Society*, 83 (2011), 563–586.
- [16] A. Berarducci, M. Mamino, M. Otero. Higher homotopy of groups definable in o-minimal structures, *Israel Journal of Mathematics*, 180(1), 143–161, 2010.
- [17] A. Berarducci, M. Otero, Y. Peterzil, and A. Pillay, A descending chain condition for groups definable in o-minimal structures, *Annals of Pure and Applied Logic*, 134 (2005), 303–313.
- [18] A. Berarducci, Y. Peterzil and A. Pillay, Group covers, o-minimality, and categoricity, *Confluentes Mathematici* 2 (04), 473–496, 2010.
- [19] A. Conversano, On the connections between definable groups in o-minimal structures and real Lie groups: the non-compact case, Ph.D. thesis, University of Siena, 2009.
- [20] A. Conversano, Maximal compact subgroups in the o-minimal setting, *Journal of Mathematical Logic*, vol. 13 (2013), 1–15.
- [21] A. Conversano, A reduction to the compact case for groups definable in o-minimal structures, *Journal of Symbolic Logic*, 79 (2014), 45–53.
- [22] A. Conversano, Nilpotent groups, o-minimal Euler characteristic and linear algebraic groups, preprint arXiv:1904.09738v2, 2020.
- [23] A. Conversano and M. Mamino, One Lie group to define them all, preprint arXiv:2010.10579, 2020.

- [24] A. Conversano, A. Onshuus, and S. Starchenko, Solvable Lie groups definable in o-minimal theories, *Journal of the Institute of Mathematics of Jussieu*, (2018) 1–12.
- [25] A. Conversano and A. Pillay, Connected components of definable groups and o-minimality I, *Advances in Mathematics*, 231 (2012), 605–623.
- [26] A. Conversano and A. Pillay, Connected components of definable groups and o-minimality II, *Annals of Pure and Applied Logic*, 166 (2015), 836–849.
- [27] A. Conversano and A. Pillay, On Levi subgroups and the Levi decomposition for groups definable in o-minimal structures, *Fundamenta Mathematicae*, 222 (2013), 49–62.
- [28] L. van den Dries, *Tame topology and o-minimal structures*, Cambridge University Press, Cambridge, 1998.
- [29] M. Edmundo, Solvable groups definable in o-minimal structures, *J. Pure and Applied Algebra*, 185 (2003), 103–145.
- [30] M. Edmundo, A remark on divisibility of definable groups, *Mathematical Logic Quarterly*, vol. 51 (2005), no. 6, pp. 639–641.
- [31] M. Edmundo, Locally definable groups in o-minimal structures, *J. Algebra* 301(1)(2006) 194–223. Corrigendum (with E. Baro) *J. Algebra* 320(7) (2008) 3079–3080.
- [32] M. Edmundo, M. Mamino, L. Prelli, J. Ramakrishnan and G. Terzo, On Pillay’s conjecture in the general case, *Adv. Math.* 310 (2017) 940–992.
- [33] M. Edmundo and M. Otero, Definably compact abelian groups, *J. Math. Logic*, 4 (2004), 163–180.
- [34] M. Edmundo and G. Terzo, A note on generic subsets of definable groups, *Fundamenta Mathematicae* 215 (2011), 53–65.
- [35] P. Eleftheriou, A semi-linear group which is not affine, *Annals of Pure and Applied Logic* 156 (2008), 287–289.
- [36] P. Eleftheriou, A. Hasson and Y. Peterzil, Strongly minimal groups in o-minimal structures, *Journal of the European Mathematical Society*, to appear.
- [37] P. Eleftheriou and Y. Peterzil, Definable quotients of locally definable groups, *Selecta*

- Mathematica N.S., Volume 18 (2012), 885–903.
- [38] P. Eleftheriou and Y. Peterzil, Lattices in locally definable subgroups of $(R^n, +)$, *Notre Dame Journal of Formal Logic*, 54 (2013), 449–461.
 - [39] P. Eleftheriou, Y. Peterzil and J. Ramakrishnan, Interpretable groups are definable, *Journal of Mathematical Logic*, Volume 14, No. 1 (2014).
 - [40] J. Gismatullin, D. Penazzi and A. Pillay, Some model theory of $SL(2, \mathbb{R})$, *Fund. Math.* 229 (2015), no. 2, 117–128.
 - [41] E. Hrushovski, Y. Peterzil, and A. Pillay, Groups, measures and the NIP, *Journal of the American Mathematical Society*, 21 (2008), 563–596.
 - [42] E. Hrushovski, Y. Peterzil, and A. Pillay, On central extensions and definably compact groups in o-minimal structures, *Journal of Algebra*, 327 (2011), 71–106.
 - [43] E. Hrushovski and A. Pillay, Affine Nash groups over real closed fields, *Confluentes Mathematici* 3 (2011), 577–585.
 - [44] E. Hrushovski and A. Pillay, On NIP and invariant measures, *Journal of the European Math. Society*, 13 (2011), 1005–1061.
 - [45] K. Iwasawa, On some types of topological groups, *Annals of Math.* 50 (1949), 507–558.
 - [46] G. Jagiella, Definable topological dynamics and real Lie groups, *Math. Logic Quarterly* 61(1-2) (2015), 45–55.
 - [47] G. Jagiella, The Ellis group conjecture and variants of definable amenability, *Journal of Symbolic Logic* 83 (2018), 1376–1390.
 - [48] D. Macpherson, A. Mosley and K. Tent, Permutation Groups in o-Minimal Structures, *Journal of the London Mathematical Society*, 62(2000), 650–670.
 - [49] M. Mamino, Splitting definably compact groups in o-minimal structures. *J. Symbolic Logic* 76 (2011), 973–986.
 - [50] A. Nesin, A. Pillay, and V. Razenĵ, Groups of dimensions two and three over o-minimal structures, *Annals of Pure and Applied Logic*, 53(3) (1991), 279–296.
 - [51] A. Onshuus and S. Post, A definability criterion for connected Lie groups,

arXiv:1910.11287

- [52] M. Otero, A survey on groups definable in o-minimal structures, in *Model Theory with Applications to Algebra and Analysis*, vol. 2, Editors: Z. Chatzidakis, D. Macpherson, A. Pillay and A. Wilkie, LMS LNS 350, Cambridge University Press, 2008, pp. 177–206.
- [53] M. Otero, On divisibility in definable groups, *Fundamenta Mathematicae* 202, n.3 (2009) 295–298.
- [54] M. Otero and Y. Peterzil, G-linear sets and torsion points in definably compact groups, *Archive for Mathematical Logic* 48 (2009) 387–402.
- [55] M. Otero, Y. Peterzil and A. Pillay, On groups and rings definable in o-minimal expansions of real closed fields, *Bulletin of the London Math. Soc.* 28 (1996) 7-14.
- [56] Y. Peterzil, Some topological and differentiable invariants in o-minimal structures, a survey of the solution to the torsion point problem, in *Model Theory and Applications*, Caserta 2002.
- [57] Y. Peterzil, Pillay’s conjecture and its solution—a survey, *Logic colloq. 2007*, Ed: F. Delon, U. Kohlenbach, P. Maddy and F. Stephan, pp. 177–203.
- [58] Y. Peterzil, Returning to semi-bounded sets, *J. Symbolic Logic*, 74 No. 2 (2009), 597–617.
- [59] Y. Peterzil and A. Pillay, Generic sets in definably compact groups, *Fundamenta Mathematicae*, 193 (2007), 153–170.
- [60] Y. Peterzil, A. Pillay and S. Starchenko, Definably simple groups in o-minimal structures, *Transactions of the American Mathematical Society*, 352(2000), 4397-4419.
- [61] Y. Peterzil, A. Pillay, and S. Starchenko, Simple algebraic and semialgebraic groups over real closed fields, *Transactions of the American Mathematical Society*, 352 (2000), 4421-4450.
- [62] Y. Peterzil, A. Pillay, and S. Starchenko, Linear groups definable in o-minimal structures, *Journal of Algebra*, 247 (2002), 1–23.
- [63] Y. Peterzil and S. Starchenko, Definable homomorphisms of abelian groups in o-minimal structures, *Annals of Pure and Applied Logic*, 101 (2000), 1–27.

- [64] Y. Peterzil and S. Starchenko, On torsion-free groups in o-minimal structures, *Illinois J. Math.*, 49(4)(2005), 1299–1321.
- [65] Y. Peterzil and S. Starchenko, Topological groups, μ -types and their stabilizers, *Journal of European Math. Society*, 19 (10) (2017), 2965–2995.
- [66] Y. Peterzil and C. Steinhorn, Definable compactness and definable subgroups of o-minimal groups, *J. London Math. Soc.*, 59(1999), 769–786.
- [67] A. Pillay, On groups and fields definable in o-minimal structures, *J. Pure and Applied Algebra*, 53 (1988), 239–255.
- [68] A. Pillay, Type-definability, o-minimality, and compact Lie groups, *J. Math. Logic*, 4 (2004), 147–162.
- [69] A. Pillay and N. Yao, On minimal flows, definably amenable groups, and o-minimality, *Advances in Mathematics* 290 (2016), 483–502.
- [70] V. Razenj, One-dimensional groups over an o-minimal structure, *Annals of Pure and Applied Logic*, 53 (1991), 269–277.
- [71] A. Strzebonski, Euler characteristic in semialgebraic and other o-minimal groups, *Journal of Pure and Applied Algebra*, 86 (1994), 173–201.
- [72] A. Strzebonski, One-dimensional groups definable in o-minimal structures, *Journal of Pure and Applied Algebra*, 1994.
- [73] K. Tent, Sharply n -transitive groups in o-minimal structures, *Forum Math.* 12(2000), 65–75.

ADDENDA: NOTE ON ‘NORMALISATION FOR BILATERAL CLASSICAL LOGIC WITH SOME PHILOSOPHICAL REMARKS’

NILS KÜRBIS
University of Lodz, Poland
nils.kurbis@filhist.uni.lodz.pl

My claim that the reduction step for case (2) of maximal formulas with reductio and non-contradiction ‘does only what it is supposed to do: it removes one maximal formula and introduces no complications’ [1, 548] was rash, as pointed out to me by Pedro del Valle-Inclan.¹ If there are more assumptions in the assumption class $[+ A]^i$ than the one displayed in the original deduction (occurring as top formulas in Π' and Π''), then these remain undischarged in the reduced deduction. In this note I give a solution to the problem. But first, two further corrections:

p.549, line 9f: ‘the right premise is not conclusion of an elimination rule’ should read ‘the right premise is not maximal’ (that is, it is not conclusion of $+\vee E$, $-\wedge E$ or reductio).

p.550, line 11: ‘reductio’ should be ‘non-contradiction’.

Reduction step (1) for maximal formulas with reductio and non-contradiction also works in case some of the formulas discharged by reductio are premises of non-contradiction, if it is applied with strategy. I repeat it here for convenience:

$$\begin{array}{ccc}
 & [+ A]^i & \\
 & \Pi & \\
 \Sigma & \frac{\perp}{- A} & \\
 + A & \hline & \\
 & \perp & \\
 & \Xi & \\
 & & \sim \\
 & & \Sigma \\
 & & [+ A] \\
 & & \Pi \\
 & & \perp \\
 & & \Xi
 \end{array}$$

and also the situation under consideration:

¹To whom also many thanks for discussion.

$$\frac{\frac{\frac{\Sigma}{+ A} \quad \frac{\frac{[+ A]^i \quad \Pi'}{- A}}{\perp} \quad \Pi''}{\perp} \quad - A}{\perp} \quad - A}{\Xi} \quad i$$

with further formulas in assumption class i left implicit. The case to be avoided is that applying the reduction step introduces maximal formulas of the same degree as the one removed: in the example, this happens if Π' and Σ conclude $- A$ and $+ A$ by introduction rules, or if Π' concludes $- A$ by reductio, $+\vee E$ or $-\wedge E$ and Σ concludes $+ A$ by one of these rules, too. If both are concluded by introduction rules, remove them as part of the reduction step by applying the appropriate procedure given under case (d) of the reduction steps for maximal formulas; the latter break up formulas into subformulas and thus any resulting new maximal formulas are of lower degree than the one removed. In the other cases, the conclusion $- A$ of Π' is itself maximal, but of one degree lower than the maximal formula $- A$ to be removed. Recall that if both premises of non-contradiction are maximal, the degree of the right premises is the degree of the formula plus 1. Thus applying the reduction procedure increases the degree of the conclusion $- A$ of Π' by one, as afterwards it stands to the right of another maximal formula. The strategy of the proof of Theorem 1 requires applying the reduction steps to maximal formulas of highest degree such that none others of that degree stand above it. So all maximal formulas above the lower $- A$ have at most its degree *qua* formula (i.e. counting only the number of connectives). Thus one way of dealing with this problem is to remove all maximal formulas of that degree that stand above the lower $- A$ before applying the reduction step that removes it. More economical would be to focus only on the troublesome cases and to remove all and only those maximal formulas that are premises of non-contradiction the other premise of which is in assumption class i before applying the reduction step that removes the lower $- A$. A better solution altogether may, however, be to introduce a special measure taking care of maximal formulas that are conclusions of reductio and premises of non-contradiction. The above strategy effectively requires a subsidiary induction to show that, while the lower maximal formula $- A$ is kept fixed, the other maximal formulas of its degree *qua* formula that stand above it are removable from the deduction. The more economical strategy could use a method similar to that employed by Stålmarmark in his normalisation proof for unilateral classical logic [2] and associate the maximal formulas that are the conclusion of reductio and premise of non-contradiction with those assumptions discharged by reductio that are premises

of non-contradiction and stand next to formulas that are themselves maximal.

References

- [1] Nils Kürbis. Normalisation for bilateral classical logic with some philosophical remarks. *The Journal of Applied Logics*, 8(2):531–556, 2021.
- [2] Gunnar Stålmarek. Normalisation theorems for full first order classical natural deduction. *Journal of Symbolic Logic*, 52(2):129–149, 1991.